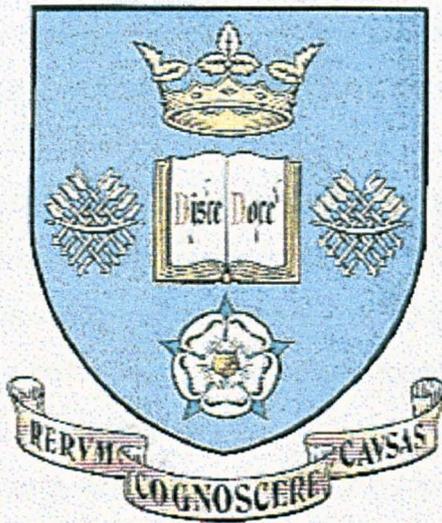


**ERROR ANALYSIS AND SYSTEM
IMPROVEMENTS IN PHASE-STEPPING
METHODS FOR PHOTOELASTICITY**



WENYAN JI

Thesis Submitted for the Degree of
DOCTOR OF PHILOSOPHY

Department of Mechanical Engineering
University of Sheffield

November 1997

Abstract

In the past automated photoelasticity has been demonstrated to be one of the most efficient technique for determining the complete state of stress in a 3-D component. However, the measurement accuracy, which depends on many aspects of both the theoretical foundations and experimental procedures, has not been studied properly. The objective of this thesis is to reveal the intrinsic properties of the errors, provide methods for reducing them and finally improve the system accuracy.

A general formulation for a polariscope with all the optical elements in an arbitrary orientation was deduced using the method of Mueller Matrices. The deduction of this formulation indicates an inherent connectivity among the optical elements and gives a knowledge of the errors. In addition, this formulation also shows a common foundation among the photoelastic techniques, consequently, these techniques share many common error sources.

The phase-stepping system proposed by Patterson and Wang was used as an exemplar to analyse the errors and provide the proposed improvements. This system can be divided into four parts according to their function, namely the optical system, light source, image acquisition equipment and image analysis software. All the possible error sources were investigated separately and the methods for reducing the influence of the errors and improving the system accuracy are presented.

To identify the contribution of each possible error to the final system output, a model was used to simulate the errors and analyse their consequences. Therefore the contribution to the results from different error sources can be estimated quantitatively and finally the accuracy of the systems can be improved.

For a conventional polariscope, the system accuracy can be as high as 99.23% for the fringe order and the error less than 5 degrees for the isoclinic angle. The PSIOS system is limited to the low fringe orders. For a fringe order of less than 1.5, the accuracy is 94.60% for fringe orders and the error about 6 degrees for the isoclinic angle if the undefined zones are ignored.

Acknowledgements

Firstly I am deeply indebted to Prof. Eann A. Patterson, for his supervision, encouragement, motivation and constant support.

I would like to thank the head of department, Prof. R. A. Smith and Prof. G. Tomlinson for allowing me to use the department facilities.

My gratitude extends to the members of the Experimental Stress Analysis Research Laboratory: to Dr. Zhifan Wang for his great help on my research; to Drs. Steve Haake, Richard Burguete, Rachel Tomlinson and Sandro Barone for their technical advice; and to John Driver and Richard Kay for making specimens for me and their kindly help.

Heartfelt thanks are given to those people who have been worked in room RC7, particularly: Mark Pacey, Elizabeth Olden, Mark Bateman, Juan Carazo, Matt Carre, Guan Chew and James Hobb, for their generally help, improving my English and doing proof reading for me. For those I met and I am not mentioned here, thank you all!

Thanks are also due to Rolls Royce Plc for providing experimental data.

Additionally thanks are given to CVCP for ORS awards and the University Bursary Scheme for the financial support.

Finally, special thanks to my grandparents and my parents for their unfailing love, continuous encouragement and constant support since I was born. The thesis would have not been possible without my husband, Shan, for his love, support, and always being there for willing help.

Contents

ABSTRACT	i
ACKNOWLEDGEMENTS	ii
CONTENTS	iii
LIST OF FIGURES	ix
NOMENCLATURE	xvi

CHAPTER 1

INTRODUCTION	1
1.1 Introduction.....	2
1.2 Objectives of the Research	2
1.3 Procedure	3
1.4 Approach.....	4
1.5 Structure of the Thesis.....	5

CHAPTER 2

LITERATURE REVIEW.....	6
2.1 Introduction.....	7
2.2 Point-by-point Analysis.....	8
2.3 Early Full-field Analysis System	9
2.4 Phase-stepping System.....	11
2.5 Fast Fourier Transform	13
2.6 White Light Source and RGB camera.....	15
2.7 Object Step-Loading Method	16
2.8 Reflection Photoelasticity.....	17

2.9 Conclusions	19
 CHAPTER 3	
FUNDAMENTAL THEORY FOR PHOTOELASTICITY	20
3.1 General Formulation in Photoelasticity	21
3.1.1 Mueller Matrix and Formulae for the Light Intensity	21
3.1.2 Automated Phase-stepping System	25
3.2 Theory of Automated Phase-stepping System.....	26
3.2.1 The Conventional Transmission Polariscope	26
3.2.2 The PSIOS System.....	29
3.3 System Arrangement of Automated Photoelasticity	32
3.3.1 Optical System and Light Source	33
3.3.2 Image Acquisition.....	34
3.3.3 Image Analysis	35
3.3.3.1 Estimate.....	36
3.3.3.2 Wrapping.....	37
3.4 Mathematical Modelling.....	40
3.4.1 Introduction	40
3.4.2 Modelling of A Circular Disc under Diametral Compression Load	41
3.5 Conclusions	44
Figures	45
 CHAPTER 4	
OPTICAL SYSTEM.....	52
4.1 Errors in Conventional Transmission Phase-stepping System.....	53
4.1.1 Effect of the Quarter-wave Plate.....	54
4.1.1.1 The Optical Property of Quarter-Wave Plate	54
4.1.1.2 Effect of Quarter-Wave Plate	55
4.1.1.3 Results and Discussion.....	56
4.1.2 Linear Alignment and the Error Adjustment.....	57
4.1.2.1 Introduction.....	57

4.1.2.2	Adjustment of the Camera According to the Polariser Position	58
4.1.2.3	Adjustment of the Other Optical Elements.....	63
4.1.2.4	Conclusion.....	63
4.1.3	Angular Alignment	64
4.1.3.1	Angular Misalignment due to Output Optical Elements	64
4.1.3.2	Angular Misalignment Introduced by the Input Quarter-Wave Plate ..	66
4.1.3.3	Angular Misalignment by Input and Output Elements	66
4.1.3.4	Conclusion.....	66
4.1.4	Variation of the Light Source and Test Environment.....	67
4.1.5	Comparison on Modelling Results and Discussion	68
4.1.6	Conclusions.....	70
4.2	Errors in the PSIOS System	71
4.2.1	Spatial Alignment of Four Beams	72
4.2.1.1	Target Design	73
4.2.1.2	Magnification Calibration.....	74
4.2.1.3	Translational Misalignment.....	75
4.2.1.4	Rotational Misalignment	77
4.2.1.5	Calibration of the Translational and Rotational Misalignment	79
4.2.1.6	Conclusion.....	82
4.2.2	Intensity Calibration of Four Beams.....	83
4.2.2.1	Transmission and Reflection Ratio for Single Beamsplitter Cube	83
4.2.2.2	Combined Effect of the Beamsplitters.....	84
4.2.2.3	Effects by Allowed Tolerance in Four Beams	86
4.2.2.4	Discussion and Conclusion.....	87
4.2.3	Oblique Viewing.....	88
4.2.4	Conclusions.....	88
Figures	90
CHAPTER 5		
LIGHT SOURCE		
5.1	Introduction.....	107
5.1	Introduction.....	108

5.2. Errors in Non-matching Monochromatic Light Source	109
5.2.1 Mismatch of the Quarter-Wave Plate	110
5.2.2 Dispersion of Birefringence	110
5.2.3 Theoretical Prediction for the Non-matching Monochromatic Light Case	112
5.2.4 Comparison of the Experimental Data and Theoretical Predication	113
5.2.5 Conclusion	116
5.3 Errors Associated with the Use of the White Light Source	116
5.3.1 Theoretical Analysis of Errors in White Light Polaroscope.....	116
5.3.2 Results of Theoretical Analysis	119
5.3.3 Experimental Results	120
5.3.4 Discussion	121
5.4 Medium-Band Filters in White Light Polaroscope.....	122
5.5 Conclusions	125
Figures	126

CHAPTER 6

IMAGE ACQUISITION & ANALYSIS.....	142
6.1 Introduction.....	143
6.2 Camera	143
6.2.1 The Linearity of the CCD Camera in the Brightness Level	144
6.2.2 Sensitivity of the Individual Cell to the Light Intensity	146
6.2.3 Errors in the Geometry	148
6.2.4 Simulation the Effect of the Camera Nonlinearity.....	148
6.2.5 Discussion on the Nonlinearity of the Camera	149
6.3 Amplifier.....	150
6.3.1 Introduction	150
6.3.2 Function of Brightness and Contrast Control	150
6.3.2.1 Change of the Brightness Parameter	151
6.3.2.2 Change of the Contrast Parameter	152
6.3.2.3 Function of Brightness and Contrast Control.....	154
6.3.3 Effect of Brightness and Contrast	154

6.3.3.1 Overview on Adjusting Brightness and Contrast Parameters	154
6.3.3.2 Effect in theory.....	155
6.3.3.3 Truncating error	156
6.3.3.4 Conclusion on the Effect of the Brightness and Contrast.....	157
6.3.4. Saturation.....	157
6.3.4.1 Introduction	157
6.3.4.2 Detection of Saturation.....	158
6.3.5 Conclusions on Amplifier Control.....	159
6.4 A/D Converter	159
6.4.1 Spatial Quantisation.....	160
6.4.2 Digitisation of the Brightness Level.....	160
6.4.3 Conclusions	163
6.5 Image Analysis and Image Pre-processing.....	163
6.5.1 Accuracy of Image Analysis	164
6.5.2 Image Noise.....	164
6.5.3 Analytical Analysis on the Error due to the Image Noise	165
6.5.4 Example of the Noise Effect.....	166
6.5.5 Image Pre-processing.....	166
6.5.5.1 Image Average	167
6.5.5.2 Gaussian Smoothing Filter.....	167
6.5.6 The Effect of Gaussian Filter.....	168
6.5.7 Discussion and Conclusions	170
Figures.....	171

CHAPTER 7

SYSTEM DISCUSSION AND EVALUATION.....	190
7.1 Overview of Error Sources in Phase-stepping Systems.....	191
7.1.1 Optical System.....	192
7.1.2 Light Source	194
7.1.3 Image Acquisition.....	195
7.1.3.1 CCD Camera.....	196

7.1.3.2 Amplifier	197
7.1.3.3 A/D Converter.....	197
7.1.4 Image Analysis.....	198
7.1.5 Discussion	198
7.2 System Evaluation.....	200
7.2.1 Total Effect on the Measured Light Intensity.....	200
7.2.2 Error Distribution for the Isochromatic and Isoclinic Data.....	201
7.2.2.1 The Digitisation Error in the Isochromatic and Isoclinic Data.....	203
7.2.3 System Evaluation in Current Conditions	204
7.2.3.1 The Conventional Polariscope	204
7.2.3.2 The PSIOS System.....	205
7.3 Conclusion	206
Figures.....	208

CHAPTER 8

CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH	214
8.1 Conclusions.....	215
8.1.1 Fundamental Theory for Photoelasticity	216
8.1.2 Error Analysis and Improvement in Each Stage.....	216
8.1.3 System Evaluation	218
8.1.4 A Simulation of Errors.....	219
8.2 Suggestions for the Further Research.....	219

REFERENCES	221
-------------------------	------------

APPENDICES.....	233
------------------------	------------

APPENDIX A Published Papers Arising from the Thesis.....	233
--	-----

APPENDIX B Distribution of Multivariate Normal Density Function	234
---	-----

List of Figures

Chapter 3 FUNDAMENTAL THEORY FOR PHOTOELASTICITY

- 3.1 The schematic diagram of a polariscope showing the angular orientations of the optical elements. The fast axis of the polariser is being used as a reference axis. 45
- 3.2 The schematic diagram showing the arrangement for the automated photoelasticity system 46
- 3.3 Optical arrangement for normal transmission and reflection polariscope 46
- 3.4 A schematic diagram showing the output section of the optical head 47
- 3.5 The schematic arrangement of image acquisition section 48
- 3.6 A flow diagram of image analysis procedure (PECAN) 48
- 3.7 A flow diagram showing the simulation procedure 49
- 3.8 a) The experimental circular disc under compression; b) Six raw images of the simulated circular disc and c) its fringe order and isoclinic angle 50

Chapter 4 OPTICAL SYSTEM

- 4.1 A plane-polarised light vector entering a doubly refracting plate 90
- 4.2 Effect of the wavelength mismatch on the isoclinic and isochromatic parameters 91
- 4.3 The distorted unitary circle with 5° misalignment to the incident light 92
- 4.4 Linear misalignment of the quarter-wave plate with an angle χ for the phase-stepping system 92
- 4.5 The effects on the isoclinic and isochromatic parameters of the quarter-wave plates being at 1 (left) and 5 (right) degrees to the perpendicular to the light path. 93
- 4.6 Transformation of the two co-ordinate systems and relative six parameters. X-Y-Z is associated with the polariser and the X_c - Y_c - Z_c 94

	is associated with the camera.	
4.7	Translation of the co-ordinate system X-Y-Z to the origin of the camera co-ordinate system $X_c-Y_c-Z_c$ to create a new co-ordinate system $X_1-Y_1-Z_1$	94
4.8	Three sequential angular rotation of the co-ordinate systems.	95
4.9	Rotation through an angle ϖ_x about the X_1 -axis of the co-ordinate system $X_1-Y_1-Z_1$ to form a new co-ordinate system $X_2-Y_2-Z_2$	95
4.10	The effects on the isoclinic (left) and isochromatic (right) parameters due to the angular misalignment of the optical elements. a) errors due to 0.1° misalignment of the output elements; b) errors due to 0.1° misalignment of the input element; c) errors due to 0.1° misalignment of both the input and the output elements.	96
4.11	Variation of the light intensity using an arbitrary pixel sampling 100 times.	97
4.12	The modelling results of the fringe order and isoclinic angle for a circular disc due to the use of sodium light source.	98
4.13	The modelling results of the fringe order and isoclinic angle for a circular disc due to 5° linear misalignment.	99
4.14	The modelling results of the fringe order and isoclinic angle for a circular disc due to 0.1° angular misalignment a) output elements; b)input quarter-wave plate; c) both output and input elements	100
4.15	The comparison of the modelling results on the fringe order and isoclinic angle along the central line of the circular disc for above three cases.	101
4.16	An alignment target used for spatial calibration of the four beams. size: 86 mm \times 54 mm	102
4.17	A schematic representation of the effect of a rotational misalignment of a camera with respect to the X_1 -axis.	102
4.18	Schematic representation of a translational alignment.	103

4.19	Schematic representation of a rotational alignment.	103
4.20	Assigning the light intensity value to the new co-ordinate system X_2'' - Y_2'' - Z_2'' .	104
4.21	Cube transmittance curve (reproduced from Melles Griot catalogue)	104
4.22	The modelling results of the fringe order and the isoclinic angle for a circular disc due to intensity tolerance allowed in the cube beamsplitter	105
4.23	The comparison of the modelling results on the fringe order and isoclinic angle along the central line of the disc due to intensity tolerance in four beams	106

Chapter 5 LIGHT SOURCE

5.1	The modelling isochromatic fringe order and isoclinic angle for the circular disc at three wavelengths of 489.3, 560 and 689.3 nm.	127
5.2	The comparison of the fringe order and isoclinic angle along the central line of the disc.	128
5.3	The comparisons of the theoretical prediction and experimental results for the isoclinic angle measured using six narrow band sources. The theoretical data and experimental are represented by solid and broken lines, respectively.	129
5.4	The comparisons of the theoretical prediction and experimental results for the isochromatic fringe order measured using six narrow band sources. The theoretical data and experimental are represented by solid and broken lines, respectively.	130
5.5	Spectra for three white light sources. The normalisation was performed using the maximum value in each case.	131
5.6	Spectra for the monochromatic camera provided by manufacturers.	131
5.7	Combined responses of the white light sources and the monochromatic camera.	132
5.8	Predicted errors in the isoclinic angle (left panel) and isochromatic fringe order (right panel) due to the three white light sources, as function of the isochromatic and isoclinic parameters in a specimen.	133

-
- The units of fringe order errors is 0.001 fringe.
- 5.9 Predicted errors in isoclinic angle and isochromatic fringe order along the centre line of a disc subject to diametral compression ($N_{\max} = 5$) at three white light sources, where the top panel is the fringe order from the stress-optic law and elasticity theory, the middle and bottom panels show the predicted errors in the isoclinic angle and fringe order for three cases. 134
- 5.10 Theoretical modelling for a circular disc subject to diametral compression with the white light source S1. The isochromatic fringe order and isoclinic angle are illustrated. 135
- 5.11 The isochromatic and isoclinic parameters across a diameter perpendicular to the loaded diametry in a disc subject to diametral compression. Experimental and theoretical predictions are shown for (a) the isochromatic and (b) isoclinic parameters for the disc in a polariscope using the light source S1 and the sodium light source. 136
- 5.12 Spectra of the medium-band filters and the light source S1 in the band. 137
- 5.13 Combined response curves of the medium-band filters and the source S1. 137
- 5.14 Predicted errors in the isoclinic angle (left panel) and the isochromatic fringe order (right panel) in light source S1 combined with 80 nm (top) and 40 nm (bottom) bandwidth filters, as function of the isochromatic and isoclinic parameters in a specimen. The units of fringe order errors are 0.001 fringe. 138
- 5.15 Predicted errors in the isoclinic angle and isochromatic fringe order along the central line of a disc subject to diametral compression ($N_{\max} = 10$). The worst light source S was combined with the use of medium-band filters. 139
- 5.16 Theoretical modelling for a circular disc subject to diametral compression with two medium-band filters and the light source S1. The isochromatic fringe order and isoclinic angle are illustrated. 140
- 5.17 The comparison of the isochromatic fringe order and isoclinic angle 141

along central line of a compressive disc in the light source S1 and with medium-band filters

Chapter 6 IMAGE ACQUISITION AND ANALYSIS

- | | | |
|------|---|-----|
| 6.1 | The experimental configuration for testing the transmission ratio of the CCD camera. | 171 |
| 6.2 | The linearity of the CCD camera tested by Neutral density filters. | 171 |
| 6.3 | A typical $H & D$ curve (reproduced from Gonzalez, 1993). | 172 |
| 6.4 | The distribution of the light source along an arbitrary line. | 172 |
| 6.5 | The raw image 1 and image 2 and their sum for the disc model obtained from simulation due to nonlinearity of the CCD camera are illustrated in the left panel. The right panel shows their profiles along the line of $Y=127$. | 173 |
| 6.6 | The raw image 1 and image 2 and their sum for the disc from experiments are illustrated in the left panel. The right panel shows their profiles along the line of $Y=127$. | 174 |
| 6.7 | The Modelling results for the isochromatic fringe order and the isoclinic angle due to nonlinearity of the CCD camera. | 175 |
| 6.8 | The comparison of the modelling results on the isochromatic and isoclinic data due to the nonlinearity of the camera. The data are shown along the central line of the circular disc. | 176 |
| 6.9 | Light intensities along an arbitrary line with the same contrast level (127) and different brightness levels. | 177 |
| 6.10 | The difference of the light intensity for an increment of 5 brightness levels | 177 |
| 6.11 | The average value and the variance for each of 5 increments in brightness levels | 178 |
| 6.12 | Light intensities along an arbitrary line with the same brightness level (127) and different contrast levels. | 179 |
| 6.13 | The difference of the light intensity for an increment of 5 contrast levels | 179 |
| 6.14 | The contrast coefficient for each increment of 5 levels | 180 |

-
- 6.15 The average value and the variance for each of 5 increments in contrast levels 180
- 6.16 The modelling of a circular disc with different light intensity range. 181
a Raw images (image 1) are illustrated in the left panel and their profiles along the line of $Y=127$ are shown in the right panel.
- 6.16 The isochromatic fringe order (left) and isoclinic angle (right)for a 182
b modelling disc with light intensity range.
- 6.17 Probability density function for the digitisation 183
- 6.18 Flow diagram shows the effect of the Gaussian filter in the phase- 183
stepping system.
- 6.19 The raw images for a blade captured by the PSIOS system. The 184
a bottom panel shows the profile along the line of $Y=51$.
- 6.19 The fringe order and isoclinic angle for the blade in the left panel and 185
b their profiles along the line of $Y=51$ in the right panel. Wrapping error was observed due to various kinds of noise.
- 6.20 The effect of Gaussian filtering. Perfect data, data with 10% noise 186
and data after processing are illustrated in the left panel. Their light intensities along the line of $Y=127$ are shown in the right panel.
- 6.21 The isochromatic fringe order and isoclinic angle from perfect data, 187
data with 10% noise and Gaussian filtered noisy data.
- 6.22 The raw images showed in Fig. 6.19a were smoothed by the 188
a Gaussian filter. The bottom panel shows the comparison of the light intensity along the line of $Y=51$ before smoothing and after smoothing.
- 6.22 The fringe order and isoclinic angle for the blade after processing by 189
b a Gaussian filter are showed in the left panel. Their profiles along the line of $Y=51$ are shown in blue line in the right panel, while cyan lien are profiles before processing.

Chapter 7 SYSTEM DISCUSSION AND EVALUATION

- 7.1 A histogram showing the distribution of variation in the measured 208

	light intensities as compared to the mean value for 100 samples. The data is aggregated in the graph. The variance of the data shown is 3.51.	
7.2a	Error distribution for the fringe order due to errors on the measured light intensity.	208
7.2b	Error distribution for the isoclinic angle due to errors on the measured light intensity.	209
7.3	The disc model shows the system evaluation for the conventional polariscope.	210
7.4	The system evaluation on the fringe order and isoclinic angle along the central line of the disc for the conventional polariscope.	211
7.5	The disc model shows the system evaluation for the PSIOS system.	212
7.6	The system evaluation on the fringe order and isoclinic angle along the central line of the disc for the PSIOS system.	213

Nomenclature

Chapter 3

- α : relative phase retardation;
- β : angles between the fast axis of the analyser and the fast axis of the polariser;
- C : stress-optic coefficient;
- C_0, C_λ : stress-optic coefficient at matching wavelength, and other wavelengths;
- ε : quarter-wave plate error
- f : fringe constant ($\text{Nm}^{-1}/\text{fringe}$);
- ϕ : angles between the fast axis of the second quarter-wave plate and the fast axis of the polariser;
- i_0 : light intensity when all the fast axes of the polariscope are parallel to each other;
- i_1 to i_6 : light intensity for phase steps No. 1 to 6;
- i_m : stray light intensity
- i_v : light intensity when all the optical elements are parallel to each other;
- K_x, K_y : transmission ratio of the cube beamsplitter
- λ : wavelength (nm);
- λ_m : matching wavelength for the quarter-wave plates (nm);
- N : isochromatic fringe order;
- N_m : fringe order at the matching wavelength of the quarter-wave plate;
- θ : isoclinic angle (degree);
- ρ : angles between the fast axis of the first quarter-wave plate and the fast axis of the polariser;
- σ_1, σ_2 : principal stresses;

t : thickness of the specimen;

ξ_1, ξ_2 : retardation caused by first and second quarter-wave plate, respectively;

Chapter 4

χ : linear misalignment by quarter-wave plates;

χ' : relative retardation error caused by linear misalignment of quarter-wave plates;

C_x, C_y, C_z : translation parameters in two co-ordinates systems

δ : relative linear phase shift;

Δ, Δ' : relative angular phase shift;

ε : angular error in quarter-wave plate;

E_t : light vector

M : lens magnification

$\omega_x, \omega_y, \omega_z$: rotation parameters in two co-ordinates systems

ζ : angular misalignment in optical elements

Chapter 5

$F(\lambda)$: response curve of the monochromatic camera;

$I_v(\lambda)$: spectrum of the white light source;

$(I)_\lambda$: light intensity for non-matching monochromatic light source;

$(I)_w$: light intensity for white light source;

$S(\lambda)$: combined response of light source and monochromatic camera;

$T(\lambda)$: products of transmission ratio of every element in the polariscope;

Chapter 6

b : brightness coefficient;

B : light intensity value when only brightness parameter changes;

ΔB : the difference of light intensity between brightness changes;

$\overline{\Delta B}$: mean value of the brightness coefficient;

σB : standard deviation of the brightness coefficient;

-
- c : contrast coefficient;
- C : light intensity value when only contrast parameter changes;
- ΔC : the difference of light intensity between contrast changes;
- $\overline{\Delta c}$: mean value of the contrast coefficient;
- σc : standard deviation of the contrast coefficient;
- d : the length of each side of square tube;
- D : common logarithm of the ratio of incident to transmitted light;
- I_o : the light intensity of incident;
- I_t : the light intensity of transmitted;
- L : length of tube;

CHAPTER 1

INTRODUCTION

A brief introduction to photoelasticity is given in the beginning of the chapter, followed by the objectives, procedure and approach of this research. The structure of the thesis is outlined.

1.1 Introduction

Photoelasticity was first observed by Sir David Brewster in 1816, as a phenomenon of temporary double diffraction induced in amorphous transparent materials by mechanical stress. The formulation of Maxwell-Neumann Stress-Optic Law [Maxwell 1853] was the next milestone during the history of the photoelasticity development. According to this law, changes in refractive indices are related linearly to the stress or strains developing in a linearly elastic material, and based on such a relationship, models of structural components can be constructed and analysed so as to determine the state of stress in the model. In the early 1930's photoelasticity was developed as a powerful stress analysis tool in experimental mechanics. However, its use was seriously hampered by the tedious and time consuming nature of photoelasticity.

Much effort has been expended in the last 20 years on the automated evaluation of the isochromatic fringe order and isoclinic angle in photoelasticity. This effort is justified by the fact that photoelasticity remains virtually the only available technique for determining by experiment the three-dimensional stress state in a component. Automation offers the prospect of rapid, accurate measurements made over a full-field view.

1.2 Objectives of the Research

The measurement accuracy of these automated photoelastic techniques depends on many aspects of both their theoretical foundations and experimental procedures. A comprehensive study of the error analyses in photoelasticity will reveal the intrinsic

properties of the errors, provide methods for reducing them and finally improve the system accuracy. Previously, this study has not been essentially aware and properly addressed, although a number of authors have reported their views on system accuracy analysis [Allison *et. al.* 1990, Allison 1989, Haake *et. al.* 1993 and Ramesh *et. al.* 1995]. The objective of this research is therefore, using the phase-stepping system proposed by Patterson and Wang [1991, 1997b] as an exemplar, to investigate all the possible error sources and quantitatively analyse their consequences, and then an approach to improve the accuracy of photoelastic stress analysis is to be proposed.

1.3 Procedure

In order to analyse the errors and improve the system accuracy, first the system is separated into consecutive sections according to their function; the nature of every section is then investigated.

The deduction of the formulation for a general system set-up is very important because it can indicate the inherent connection between each system element and system output and may reveal knowledge about errors due to unexpected conditions.

By investigating all the errors in each section and their contributions to the final system output, the methods to improve the accuracy of system output can be proposed. To identify the contribution of each possible error to the final system output, modelling can be used to simulate the errors.

The system accuracy is then evaluated based on the combined influence of all sections of the system on the system output.

1.4 Approach

The analysis of the errors can be approached in two ways,

1. Experimental Method

Each section of the system is isolated and examined under specific conditions to study the possible errors and their properties. Methods are then proposed for the removal or reduction of the errors.

The system, which is formed by combining all sections under normal conditions, is used to examine the accuracy of the system with a known object.

2. Numerical Modelling

Numerical modelling can provide a virtual way to separate properly the possible errors and influences of each section from the others, while the experimental methods may not always do so. The sequence of reducing errors is determined based on their contributions.

The basic model of the system is based on the assumption that all the system components are in perfect conditions and placed exactly according to the system arrangement. This model can be used as a presentation of the ideal system.

All kinds of possible noise and misconduction in the system can be simulated by changing the parameters of the basic model of the system, so that all the error sources could be studied separately and precisely.

The comparison of the experimental and modelling results can verify the error analysis and give a confidence in error properties and finally improve the system accuracy.

1.5 Structure of the Thesis

Chapter 2 will provide a literature review regarding the recent developments in the automated photoelasticity.

The fundamental theory is deduced in **Chapter 3** and an introduction to the system hardware is also given. The numerical modelling procedure used through the whole research is presented.

The system can be divided into four sections according to their functions, namely optical system, light source, image acquisition and image analysis. **Chapter 4** investigates the errors in the optical system; **Chapter 5** concentrates on the light source; image acquisition and image analysis are discussed in **Chapter 6**.

Based on the discussion of the errors and their contributions in each section, the system accuracy is evaluated in **Chapter 7**.

Conclusions are then drawn and suggestions for further research are described in **Chapter 8**.

CHAPTER 2

LITERATURE REVIEW

This chapter is reviewed the development of the automated photoelasticity in the last two decades. The advantages and disadvantages for each method are described and the phase-stepping method proposed by Patterson and Wang is chosen as an exemplar to analyse the errors.

2.1 Introduction

The science of photoelasticity is based on the Maxwell Stress-Optic Law [Maxwell 1853]. The photoelastic phenomenon was first observed by Sir David Brewster in 1816 and then was developed in the early thirties as an experimental method for analysing stress fields in mechanics. A description of the method was written in 1940's by Frocht [Frocht 1941]. Now photoelasticity has been proven to be one of the most efficient methods for determining the complete state of stress in a three-dimensional component. However, its use was severely hampered by the tedium involved and the time required for a well trained operator to obtain and interpret results. The research for a new automated photoelastic technique has been promoted.

With the advent of a small powerful computers and numerical techniques, a number of successful automated methods have been developed. Early transmission systems for automated photoelasticity have been reviewed by Patterson in 1988 [Patterson 1988] and can be divided into two categories: the point-by-point analysis and the full-field systems. The spinning element is a early point-by-point method while the spectral contents analysis is a later point-by-point method. The full-field type includes those that use the image thinning technique, phase-stepping or phase-shifting technique, the fast Fourier transform, and the colour image processing technique which uses white light as the source and a RGB camera as a sensor. The literature review on transmission polariscope will mainly concentrate on the methods developed after 1988 and the development on the reflection photoelasticity will be briefly surveyed.

2.2 Point-by-point Analysis

Two types of technique are normally used in point-by-point analysis. The first one involves rotating the polariser and analyser [Allison *et. al.* 1972, Redner 1974 and Fessler *et. al.* 1983] while the second one analyses the spectral contents to identify isochromatic parameters at a point on a sample. One of the first proponents of the method of spectral contents analysis was Redner [1985] and later Voloshin and Redner [1989]. A similar system was developed by Sanford and Iyengar [1985]. A white light circular polariscope was used, the fringe order being found from spectra of the light emerging from the analyser. This was done either by fitting the theoretical light intensity equation to the spectrum or by comparing a series of theoretical spectra with the experimental spectrum and choosing the one with the least difference. Haake and Patterson [1992] demonstrated the use of Spectral Contents Analysis (SCA) to stress frozen specimens and extended it to a three-dimensional components. Haake and Patterson [1993] also extended SCA to a full-field method and used the spectrum of the light emerging from a plane polariscope to measure the isoclinic angle. Pacey *et. al.* [1997] compared the accuracy and the speed of spectral search method, such as summation, data base, Newton-Raphson and Golden section and proposed three new methods based on maximum entropy method, a genetic algorithm and a memic algorithm and found that by combining of the maximum entropy method with either the genetic algorithm or the memic algorithm, spectra could be analysed up to 30 times faster with comparable accuracy.

The major advantage of spectral analysis is that it does not require the well-trained specialist to acquire and interpret results and it can automatically extract the value of retardation from the light intensity of beam. Although a full-field capability

has been demonstrated, spectral analysis is largely limited to point-by-point method by processing time required for each method.

2.3 Early Full-field Analysis System

Two similar full-field systems were developed by Müller and Saackel [1979], and Seguchi *et. al.* [1979] independently. In these systems a TV camera was used as an image sensor and a fringe thinning technique was used to obtain the integer fringes. Müller and Saackel achieved this by fitting circles into the fringe and expanding the circles until the fringe boundaries were reached. Seguchi *et. al.* extracted the centre lines by progressively thinning fringes through removing the outer layer of the points until only the centre line was left. The identification and assignment of the fringe orders to the isochromatic fringes was carried out by the operators.

Developments based on numerical techniques and fringe thinning techniques have been proposed by Chen and Taylor in 1989 [Chen *et. al.* 1989], and by Ramesh, Ganesan and Mullick in 1991 [Ramesh *et. al.* 1991]. Chen and Taylor have used image processing techniques to analyse the fringe patterns produced in photoelasticity, holographic interferometry and speckle interferometry. The different quality of the fringe images was considered. Ramesh, Ganesan and Mullick used the minimum polarised light intensity criterion to detect the edge and extract the fringe skeleton. The common problem in their methods was these required the operator to identify and assign the fringe orders to isochromatic fringes.

Umezaki, Tamaki and Takahashi [1989] improved their system by taking images in two or more different wavelengths. The zero order fringe was identified as

the only fringe that suffered no displacement with the change in wavelength. However, if there is no zero order fringe presented, then the technique fails. This situation may arise in the analysis of slices from three dimensional models.

All of the systems in the full-field analysis described so far are based on fringe thinning. The major disadvantage is that they only obtain the information about integer fringe order and a lot of useful information is lost, such as the isochromatic information between the fringe skeleton. The method of "Half-Fringe Photoelasticity" proposed by Voloshin and Burger [1983] intended to overcome this problem. This method operated with no more than a half fringe order presented in the model and avoided the necessity of manual identification of the fringes. In order to stay within the half fringe order, it was required the combination of the following: low loads, high material fringe value and thin slices. High modulus glass models were chosen in preference to polymeric models and then the problem of large deformation was overcome. The obvious disadvantage of this method is it can not identify relative retardation exceeding half a fringe order so the effective region of this method was limited.

Brown and Sullivan [1990] developed "The Computer-aided Holophotoelastic method". In this approach, the isopachic data (the sum of the absolute retardations) was obtained from phase-stepping procedure applied to a real time holographic fringe system. The isochromatic and isoclinic fringe measurements were obtained by employing half-fringe photoelasticity and polarisation-stepping photoelasticity. It can provide accurate stress measurements for a transparent structure subjected to a two dimensional state of stress. The system was validated using a diametral loaded disc. The error for measurement of the isochromatic was up to 10% of usable fringe value,

with an absolute error of about 0.06 in measurement of isopachics, and 0.5 to 1.5 degrees obtained for isoclinics. All of these errors could be significant in low-stress region.

2.4 Phase-stepping System

The phase-stepping or phase-shifting concept for photoelasticity was first introduced by Hecker and Morche in 1986. This method was used to determine two data fields of relative phase α and the isoclinic data θ . The relative phase α of the two refracted waves, transmitted through the specimen, are related to the fringe order N and the principal-stress difference $(\sigma_1 - \sigma_2)$ by:

$$\alpha = 2\pi N = (\sigma_1 - \sigma_2) \cdot f \cdot t \quad (2-1)$$

where t is the specimen thickness, and f is the material fringe constant. The isoclinic angle is the direction of the principal stress.

Hecker and Morche [1986] chose a left circular polariscope to determine the relative phase α and a plane polariscope to determine the isoclinic data θ . By taking account of the local mean illumination and local amplitude of the intensity component, the light intensity for a circular polariscope emerging from an analyser can be deduced. Then they chose five phase steps to obtain the relative phase α . Information about the isoclinic parameter θ was collected in a white light polariscope with a plane polariser in three analyser positions. The intensities of the photoelastic pattern were recorded by a TV camera, digitised and stored in a computer. The accuracy of this method was only moderate in the view of the authors, because they did not consider the inhomogeneities of the optical components.

Based on these principles, Patterson and Wang further developed this technique in 1991 [Patterson and Wang 1991]. They used a left circular polariscope and selected six different phase steps to determine both the isochromatic and the isoclinic data in the full-field. With a knowledge of the fringe order at one point, the absolute fringe order and isoclinic angle can be obtained. Wang and Patterson [1995] provided further description of their method including the use of the demodulation and the fuzzy sets method. The major advantage of this method is that both the isoclinic and the fractional isochromatic parameters can be established at all individual points in the field without reference to neighbouring points. The accuracy is around ± 0.007 fringes for the isochromatics and 0.75 degrees for the isoclinics [Haake, Wang and Patterson, 1993].

Sarma *et. al.* [1992] proposed a new computerised method using Stokes representation of polarised light for the computations of the fractional fringe order and the isoclinic data. The light intensity data were recorded in three analyser positions in plane polariscope set-up. The relationship between the intensity values and the isochromatic, isoclinic parameters were derived from Stokes vector representation of polarised light and its transformation. The accuracy of within ± 3 degrees for isoclinics and ± 0.05 fringe for isochromatics was achieved. Because only polarisers were used, the biggest advantage is that no errors were introduced due to dispersion of the quarter-wave plate.

Asundi [1993] used the phase-stepping method for demonstrating photoelasticity to undergraduate students. A standard circular polariscope was used in this method, all elements of the polariscope were rotated so that the polariser axis was aligned parallel to one of the principal stress axis. Then rotated the analyser to

four positions to obtain the fractional fringe order. The drawback of this method is that only the isochromatics may be found and only along lines of constant isoclinic angle.

Ramesh and Ganapathy [1995] compared all of the algorithms from the phase stepping or shifting methods mentioned before and concluded that only the method proposed by Patterson and Wang [Patterson and Wang 1991] has the full-field ability and the best accuracy.

The phase-stepping technique has the following advantages over the other techniques:

1. Both the fractional isochromatic and isoclinic parameters at all points in the field of view are determined independent of their neighbouring points so no accumulation errors will be produced.
2. It contains all of the information about the fractional isochromatic parameters and isoclinic parameters at all points in the full-field unlike the fringe thinning technique.
3. It is an automatic full-field system capable of handling slices from complex three dimensional photoelastic models with the minimum of operator interaction.

The disadvantages are the periodic nature of the relative retardation and the undefined zones in the isoclinic angle.

2.5 Fast Fourier Transform

More recently, the use of Fourier transforms in photoelasticity has been proposed. Morimoto *et. al.* [1994, 1993] in Japan presented this method to separate the

isochromatics and isoclinics. To avoid the errors due to the ellipticity of circular polariscopes, they proposed the method using the plane polariscope and combined the co-ordinates (x, y) and the angle θ of the crossed polaroids into the 3-D (x, y, θ) image data. Ninety sequential 2-D images were recorded at every 1 degree angle of θ from (-45) degree to 44 degree. Then using the Fourier transform in the θ direction, the complex function of Fourier spectrum is obtained, that is the 3-D (x, y, θ) image. The 2-D image extracted using the filter in the frequency at $\omega=0$, shows the isochromatics. The arctangent of the ratio of the imaginary to the $\omega=-\omega_0$ gives the direction of principal stress. The accuracy of the isoclinic data for this method, tested by a disc in diametral compression, is better than 10 degrees for integer fringe order, and less than 5 degrees for other values. The greatest advantage is that this method is insensitive to high frequency noise.

Using the Fourier transform and carrier fringe technique, Quan *et. al.* [1993] developed another method. They used a quartz wedge to generate a linear carrier fringe and superimposed the fringes formed by the stressed model in the circular polariscope. The combined fringe pattern was captured by a CCD camera and analysed by the Fourier transform. The inverse Fourier transform of the intensity distribution leads to the photoelastic data. The isochromatic data can be obtained by the arctangent of the ratio of the imaginary to the real part and the quantitative 3-D principal stress difference can be produced. By properly selecting the spatial cut-off frequencies, the unwanted low frequency background and high frequency speckle noise can be filtered out. The key of this method is that the carrier fringe frequency must be high enough to modulate the phase, otherwise, large errors will be introduced

to the phase maps. Only one frame is needed to form the fringe information using this method. The disadvantage is that high stress concentration areas can not be solved completely even using high density carrier fringes and no isoclinic information can be obtained. A circular disc and a circular ring in diametrical compression were used to test the method and errors of 3% were experienced.

The use of Fourier transforms have been seen to be more resistant to noise but to require a large number of images to be collected or to rely on the availability of high density carrier fringes.

2.6 White Light Source and RGB camera

Carazo-Alvarez *et. al.* [1994] combined the SCA and phase-stepping technique to completely automate the photoelastic analysis. They used the SCA to determine the absolute fringe order at just one point in fringe field which is used as a calibration point to the whole field analysis. The white light source used emits the wavelength in the range of 450 to 750 nm. They also concluded that the use of 8 points spectra can measure the fringe order up to 7.5.

Extending this idea, Haake and Patterson [1995] proposed a new approach wherein the use of a spectroscope is replaced by using 8 high quality optical filters in conjunction with a black and white CCD camera and made the SCA method full field.

Ajovalasit *et. al.* [1995a, b], however, proposed using the RGB camera to collect the colour images in a white light source and uniquely determine retardations in the range of 0 to 3 fringe orders by directly comparing the intensity values of R, G and B with the calibration array stored in the computer.

Ramesh and Deskmukh [1996] proposed the method called Three Fringe Photoelasticity, so called because the fringe order can be measured up to 3. They used the same hardware as Ajovalasit and determined the fringe order by a similar database search method. Instead of directly using RGB values corresponding to a dark field, the difference of RGB values between bright and dark field can reduce the amount of noise and the use of $(R-G)$, G and B can further reduce the noise influence.

RGB camera was also proposed to overcome the undefined zones in the isoclinic angle and calibrate the unwrapped relative retardation by Patterson and Wang [1997a].

2.7 Object Step-Loading Method

To solve the problems of the discontinuity arise from the automated photoelasticity, Ekman and Nurse [1997] and Ng [1997] proposed the method called object step-loading. The former based on the theory of the phase-stepping method, while the latter based on the use of the Fourier transform. Suppose P denoted a load applied for a specimen and ∂P denoted an increment of the load for a circular transmission polariscope, the former compared the light intensity with load P to the $P + \partial P$, and $P - \partial P$ to obtain the absolute fringe order and the increment of the load was directly linked with the maximum fringe order that can be measured. The latter, however, used a successive load to obtain the stress that applied in the specimen. The increment of the loading ∂P is proportional to the phase difference, and both systems

limited the phase difference less than π . A circular disc under compression was used to verify their systems and both showed a good reliability.

2.8 Reflection Photoelasticity

In the previous discussion, all the experimental set-up use the transmission polariscope and these will limit the specimen to be transparent. The method of birefringent coatings allows the evaluation of the strain distribution on the surface of an opaque body and this method may be considered as an application and extension to opaque material of two-dimensional photoelasticity, which is applicable to only transparent models. The method presents the great advantages, when compared with transmission photoelasticity, that no model need be made of the structural member under investigation and that it can be used for strain determination of structures in situ.

The idea of using birefringent coatings to measure the surface strain of an opaque body was first conceived by Mesnager in 1930 [Mesnager 1930]. He also established the basic principles of the method. Oppel [1936], some years later, applied the method to the case of a photoelastic strain gauge for local determination of the state of stress in a body. However, due to the low optical sensitivity of the material of the coating used by Oppel many difficulties were encountered in the interpretation of the optical patterns obtained. Almost two and a half decades after its invention, the method of birefringent coatings received considerable attention by many investigators who further elaborated and developed the method as a potential tool for stress analysis. Many publications aiming to establish and apply the method to

engineering problems were published in the late fifties [D'Agostino *et. al.* 1955, Duffy *et. al.* 1961 and Zandman *et. al.* 1956].

Zandman described the development of the birefringent coating in his monograph “photoelastic coatings” in 1970 [Zandman 1970] and concluded that the thickness of the coating is extremely critical in the interpretation of the fringe patterns. As the coating becomes thinner, the number of the observed fringes decreases, whereas as the thickness of the coating increases the parasitic birefringence, which makes difficult the interpretation of the induced fringe patterns, increase. The effect of the thickness of the coating on the accuracy of the results was studied by many investigators. The following factors influence the interpretation of the fringe pattern: a) the reinforcing effect of the coating; b) the transition of strains from the surface of the body to the surface of the coating; c) the gradients over the surface in question; and d) the difference between Poisson ratios of the materials of the main body and the coating. All of these factors reduce the accuracy of interpretation of the fringe patterns. Their significance increases as the thickness of the coating increases. The influence of all of these factors on the accuracy of the results obtained has been studied by various authors [Zandman *et. al.* 1962, Post *et. al.* 1961, Duffy 1961, Lee *et. al.* 1961, Day 1962, Theocaris *et. al.* 1964 and Yew *et. al.* 1968].

The thickness of the coating became less important since the low fringe order can be observed by the automated methods. Ramesh and Mangal [1997] extended the phase-stepping method proposed by Patterson and Wang to the reflection photoelasticity, while Patterson and Wang [1997b] developed a new instrument (Phase-Stepped Images Observed Simultaneously – PSIOS), which had the capability to observe and capture phase-stepped images simultaneously and therefore offered the

potential for real-time automated photoelasticity. Both papers showed the good reliability.

2.9 Conclusions

The development of the automated transmission photoelasticity and reflection photoelasticity was reviewed. Their advantage and drawback were described. Two phase-stepping methods proposed by Patterson and Wang [1991, 1997b] are then accepted as exemplars for their effectiveness and availability to analyse the errors.

CHAPTER 3

FUNDAMENTAL THEORY FOR PHOTOELASTICITY

In **Chapter 2**, the development of automated photoelasticity has been reviewed. All these techniques were discussed using the specific orientations of the polariscope. The formulation of a general system set-up, within which the parameters of the optical elements can be changed arbitrarily, is very important because it indicates an inherent connection between each optical element and the resulting phase-stepped images. Such a formulation can be used to reveal information about the errors resulting due to any unexpected orientation and/or imperfect elements. This general formulation is deduced using the method of Mueller Matrices [Theocaris and Gdoutos, 1979]. Subsequently, the fundamental theory for the phase-stepping system is described. Throughout this thesis a model of a circular disc under compression is used to simulate error influence. The theory for this model is briefly introduced at the end of this chapter.

3.1 General Formulation in Photoelasticity

Early systems for automated photoelasticity have been reviewed in **Chapter 2**. It appeared that all these systems were based on a special arrangement of the polariscope. The general formulation of an arbitrary set-up is considered very important because the investigation of such a general relationship can indicate any inherent connection between each optical element and the resulting phase-stepped image. In addition, since it is not possible to set the system exactly as specified, the deviation of the arrangement will cause the specific arrangement to lose some geometric constraints which are basic conditions for further analysis. In this case, there is a tendency for random errors due to setting up the experiment. The study of such a general situation can give information about this error. In the following section the general formulation for a system in which all the optical elements in the arbitrary orientation is deduced using the method of Mueller Matrices.

3.1.1 Mueller Matrix and Formulae for the Light Intensity

Consider a two-dimensional plane stress state in a specimen which has a relative phase retardation, α and an isoclinic angle, θ that may vary with spatial co-ordinates (x, y) within the plane of the specimen. Relative retardation is related to the isochromatic fringe order, N by,

$$\alpha = 2\pi \cdot N = 2\pi \cdot \frac{(\sigma_1 - \sigma_2) \cdot C \cdot t}{\lambda} \quad (3-1)$$

where σ_1 and σ_2 are the maximum and minimum principal stresses, respectively, t is the thickness of the specimen, C is the stress-optic coefficient and λ is the wavelength of light used.

In a polariscope, suppose the fast axes of the first and second quarter-wave plates and an analyser make arbitrary angles ρ , ϕ and β to the fast axis of the polariser, as shown in Fig. 3.1. A Cartesian co-ordinate system **X-Y-Z** was selected such that the **X**-axis was coincident with the fast axis of the polariser and the **Z**-axis was coincident with the direction of the light source. The retardations caused by the first and second quarter-wave plates are ξ_1 and ξ_2 , respectively. The Mueller calculus is used to deduce a general equation for a polariscope with a general set-up.

The normalised Stokes vector **S** of the unpolarised light that emerges from the monochromatic light source is expressed as:

$$\mathbf{S} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (3-2)$$

The Mueller matrix of the polariser \mathbf{P}_β , with its axis at an angle β to the **OX** axis, is,

$$\mathbf{P}_\beta = \frac{1}{2} \begin{bmatrix} 1 & \cos 2\beta & \sin 2\beta & 0 \\ \cos 2\beta & \cos^2 2\beta & \sin 2\beta \cos 2\beta & 0 \\ \sin 2\beta & \sin 2\beta \cos 2\beta & \sin^2 2\beta & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (3-3)$$

For a birefringent plate with retardation α , whose fast axis at an angle θ with the **OX** axis, its Mueller matrix $\mathbf{R}_\theta(\alpha)$ can be expressed as:

$$\mathbf{R}_d(\alpha) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos^2 2\theta + \sin^2 2\theta \cos \alpha & (1 - \cos \alpha) \sin 2\theta \cos 2\theta & -\sin 2\theta \sin \alpha \\ 0 & (1 - \cos \alpha) \sin 2\theta \cos 2\theta & \sin^2 2\theta + \cos^2 2\theta \cos \alpha & \cos 2\theta \sin \alpha \\ 0 & \sin 2\theta \sin \alpha & -\cos 2\theta \sin \alpha & \cos \alpha \end{bmatrix} \quad (3-4)$$

Hence, for a wave plate with retardation of ξ , its Mueller matrix $\mathbf{Q}_\rho(\xi)$ is then expressed as:

$$\mathbf{Q}_\rho(\xi) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos^2 2\rho + \sin^2 2\rho \cos \xi & (1 - \cos \xi) \sin 2\rho \cos 2\rho & -\sin 2\rho \sin \xi \\ 0 & (1 - \cos \xi) \sin 2\rho \cos 2\rho & \sin^2 2\rho + \cos^2 2\rho \cos \xi & \cos 2\rho \sin \xi \\ 0 & \sin 2\rho \sin \xi & -\cos 2\rho \sin \xi & \cos \xi \end{bmatrix} \quad (3-5)$$

In the case of wave plate with a quarter-wave retardation, i.e., $\xi = 90^\circ$, then $\mathbf{Q}_\rho(\xi)$ can be simplified as:

$$\mathbf{Q}_\rho(90) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos^2 2\rho & \sin 2\rho \cos 2\rho & -\sin 2\rho \\ 0 & \sin 2\rho \cos 2\rho & \sin^2 2\rho & \cos 2\rho \\ 0 & \sin 2\rho & -\cos 2\rho & 0 \end{bmatrix} \quad (3-6)$$

According to the Mueller calculus, the Stoke vector \mathbf{C} of the light that emerges from the analyser is,

$$\mathbf{C} = \mathbf{P}_\beta \mathbf{Q}_\rho(\xi_2) \mathbf{R}_d(\alpha) \mathbf{Q}_\rho(\xi_1) \mathbf{P}_0 \mathbf{S}$$

or:

$$\begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 1 & \cos 2\beta & \sin 2\beta & 0 \\ \cos 2\beta & \cos^2 2\beta & \sin 2\beta \cos 2\beta & 0 \\ \sin 2\beta & \sin 2\beta \cos 2\beta & \sin^2 2\beta & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos^2 2\phi + \sin^2 2\phi \cos \xi_2 & (1 - \cos \xi_2) \sin 2\phi \cos 2\phi \\ 0 & (1 - \cos \xi_2) \sin 2\phi \cos 2\phi & \sin^2 2\phi + \cos^2 2\phi \cos \xi_2 \\ 0 & \sin 2\phi \sin \xi_2 & -\cos 2\phi \sin \xi_2 & \cos \xi_2 \end{bmatrix} \\
 \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos^2 2\theta + \sin^2 2\theta \cos \alpha & (1 - \cos \alpha) \sin 2\theta \cos 2\theta & -\sin 2\theta \sin \alpha \\ 0 & (1 - \cos \alpha) \sin 2\theta \cos 2\theta & \sin^2 2\theta + \cos^2 2\theta \cos \alpha & \cos 2\theta \sin \alpha \\ 0 & \sin 2\theta \sin \alpha & -\cos 2\theta \sin \alpha & \cos \alpha \end{bmatrix} \\
 \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos^2 2\rho + \sin^2 2\rho \cos \xi_1 & (1 - \cos \xi_1) \sin 2\rho \cos 2\rho & -\sin 2\rho \sin \xi_1 \\ 0 & (1 - \cos \xi_1) \sin 2\rho \cos 2\rho & \sin^2 2\rho + \cos^2 2\rho \cos \xi_1 & \cos 2\rho \sin \xi_1 \\ 0 & \sin 2\rho \sin \xi_1 & -\cos 2\rho \sin \xi_1 & \cos \xi_1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (3-7)$$

Assuming all the optical elements are parallel to each other, the light intensity is expressed as i_0 . Since the light intensity I is proportional to the term c_0 , the emerging light intensity is given by,

$$I = \frac{i_0}{2} - \frac{i_0}{2} \cos 2(\beta - \phi) \left\{ \begin{array}{l} -\cos 2(\phi - \theta) \cos 2(\theta - \rho) \cos 2\rho \\ + \cos 2(\phi - \theta) \sin 2(\theta - \rho) \sin 2\rho \cos \xi_1 \\ + \sin 2(\phi - \theta) \sin 2(\theta - \rho) \cos 2\rho \cos \alpha \\ + \sin 2(\phi - \theta) \cos 2(\theta - \rho) \sin 2\rho \cos \xi_1 \cos \alpha \\ - \sin 2(\phi - \theta) \sin 2\rho \sin \xi_1 \sin \alpha \end{array} \right\} \\
 + \frac{i_0}{2} \sin 2(\beta - \phi) \left\{ \begin{array}{l} -\sin 2(\phi - \theta) \cos 2(\theta - \rho) \cos 2\rho \cos \xi_2 \\ + \sin 2(\phi - \theta) \sin 2(\theta - \rho) \sin 2\rho \cos \xi_2 \cos \xi_1 \\ - \cos 2(\phi - \theta) \sin 2(\theta - \rho) \cos 2\rho \cos \xi_2 \cos \alpha \\ + \sin 2(\theta - \rho) \cos 2\rho \sin \xi_2 \sin \alpha \\ - \cos 2(\phi - \theta) \cos 2(\theta - \rho) \sin 2\rho \cos \xi_2 \cos \alpha \cos \delta_1 \\ + \cos 2(\theta - \rho) \sin 2\rho \sin \xi_2 \sin \alpha \cos \xi_1 \\ - \cos 2(\theta - \rho) \sin 2\rho \cos \xi_2 \sin \alpha \sin \xi_1 \\ + \sin 2\rho \sin \xi_2 \cos \alpha \sin \xi_1 \end{array} \right\} \quad (3-8)$$

For the quarter-wave plate with 90° retardations, the above formula can be simplified as:

$$I = \frac{i_0}{2} [1 + \cos 2(\beta - \phi) \cos 2(\phi - \theta) \cos 2(\theta - \rho) \cos 2\rho \\
 - \cos 2(\beta - \phi) \sin 2(\phi - \theta) \sin 2(\theta - \rho) \cos 2\rho \cos \alpha \\
 + \cos 2(\beta - \phi) \sin 2(\phi - \theta) \sin 2\rho \sin \alpha \\
 + \sin 2(\beta - \phi) \sin 2\rho \cos \alpha \\
 + \sin 2(\beta - \phi) \sin 2(\theta - \rho) \cos 2\rho \sin \alpha] \quad (3-9)$$

In a plane polariscope no quarter-wave plate is used and so the Stoke vector \mathbf{E} of the light emerging from the analyser is,

$$\mathbf{E} = \mathbf{P}_\beta \mathbf{R}_\alpha(\alpha) \mathbf{P}_0 \mathbf{S}$$

or

$$\begin{bmatrix} e_0 \\ e_1 \\ e_2 \\ e_3 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 1 & \cos 2\beta & \sin 2\beta & 0 \\ \cos 2\beta & \cos^2 2\beta & \sin 2\beta \cos 2\beta & 0 \\ \sin 2\beta & \sin 2\beta \cos 2\beta & \sin^2 2\beta & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos^2 2\theta + \sin^2 2\theta \cos \alpha & (1 - \cos \alpha) \sin 2\theta \cos 2\theta & -\sin 2\theta \sin \alpha \\ 0 & (1 - \cos \alpha) \sin 2\theta \cos 2\theta & \sin^2 2\theta + \cos^2 2\theta \cos \alpha & \cos 2\theta \sin \alpha \\ 0 & \sin 2\theta \sin \alpha & -\cos 2\theta \sin \alpha & \cos \alpha \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (3-10)$$

The light intensity I is given by the term e_0 :

$$I = k \cdot e_0 = \frac{i_0}{2} [1 + \cos 2(\beta - \theta) \cos 2\theta - \sin 2(\beta - \theta) \sin 2\theta \cos \alpha] \quad (3-11)$$

The techniques including SCA, phase-stepping, FFT and RGB which were reviewed in **Chapter 2** appear to use different formulae. However, as observed by Patterson *et al.* [Patterson, Ji and Wang 1997], all of these formulae can be deduced from Eq. 3-8. Hence the techniques could be considered as variants on a single technique. The phase-stepping technique has been selected to represent these variants and is used in considering the sources and effects of errors throughout this thesis.

3.1.2 Automated Phase-stepping System

The general formulae derived are for a polariscope with an arbitrary set-up. Several phase-shifting or phase-stepping systems have been developed based on these

formulae [Hecker *et. al.* 1986, Patterson *et. al.* 1991, Sarma *et. al.* 1992, Asundi 1993 and Patterson *et. al.* 1997b]. The phase-stepping concept has been simply described by Hecker *et. al.* as changing the absolute phase of the reference wave in equal steps and measuring the local light intensity after each step. The change in phase is in practice achieved by rotation of the input or output optical elements of the polariscope. The following section will review two phase-stepping systems developed by Patterson and Wang [1991, 1997b]. These were chosen for their availability and effectiveness. Ramesh and Ganapathy [1995] reviewed phase-stepping algorithms for photoelasticity and found those proposed by Patterson and Wang [1991] to be the most effective.

3.2 Theory of Automated Phase-stepping System

Two systems have been developed by Patterson and Wang for the development of the automated phase-stepping system, namely the Conventional Phase-stepping Polariscope, and the PSIOS (Phase-Stepped Image Observed Simultaneously) system. The former is based on the use of six phase steps and can only be used in sequential recording, while the latter is based on the four phase steps and can capture four images simultaneously.

3.2.1 The Conventional Transmission Polariscope

The conventional phase-stepping polariscope can be used both in transmission and reflection photoelasticity [Ramesh and Mangal 1997, Barone and Patterson, 1996]. This section discusses its use in transmission photoelasticity only. In a conventional

transmission polariscope monochromatic light was used as a source and the input-quarter-wave plate was placed so that its fast axis had (-45°) with respect to the fast axis of the polariser, i.e., left circular polarised light was generated. The light intensity observed through the polariscope at any point on the specimen was governed by Eq. 3-9 with corresponding parameters,

$$I = \frac{i_0}{2} [1 - \cos 2(\beta - \phi) \sin 2(\phi - \theta) \sin \alpha - \sin 2(\beta - \phi) \cos \alpha] \quad (3-12)$$

By taking account of the local mean illumination, denoted by I_m , Eq. 3-12 can be rewritten as,

$$I = i_m - i_v [\cos 2(\beta - \phi) \sin 2(\phi - \theta) \sin \alpha + \sin 2(\beta - \phi) \cos \alpha] \quad (3-13)$$

where β and ϕ , the angles between the fast axes of output elements and reference respectively, are two controlled variables and the light intensities i_m , i_v , the isochromatic angle α and the isoclinic data θ are four unknowns. In order to solve these four unknowns, at least four independent equations are required. The four required equations will be determined according to the arrangement of ϕ and β . To minimise the influence of the noise, six phase-steps were selected by Patterson and Wang and are shown on Table 3.1.

Table 3.1 The angular orientations of the output elements of the polariscope proposed by Patterson and Wang [1991]

	step 1	step 2	step 3	step 4	step 5	step 6
angle ϕ	0	0	0	$\pi/4$	$\pi/2$	$3\pi/4$
angle β	$\pi/4$	$-\pi/4$	0	$\pi/4$	$\pi/2$	$3\pi/4$

The correspondent light intensities were expressed as:

$$i_1 = i_m - i_v \cos \alpha \quad (3-14a)$$

$$i_2 = i_m + i_v \cos \alpha \quad (3-14b)$$

$$i_3 = i_m + i_v \sin \alpha \sin 2\theta \quad (3-14c)$$

$$i_4 = i_m - i_v \sin \alpha \cos 2\theta \quad (3-14d)$$

$$i_5 = i_m - i_v \sin \alpha \sin 2\theta \quad (3-14e)$$

$$i_6 = i_m + i_v \sin \alpha \cos 2\theta \quad (3-14f)$$

The six intensity values can be measured at each point in the field of view to form six images. These intensities can be combined as follows:

$$i_d = i_2 - i_1 = 2 i_v \cos \alpha \quad (3-15a)$$

$$i_s = i_3 - i_5 = 2 i_v \sin \alpha \sin 2\theta \quad (3-15b)$$

$$i_c = i_6 - i_4 = 2 i_v \sin \alpha \cos 2\theta \quad (3-15c)$$

Thus, the periodic isoclinic parameter at any point (x, y) over the surface of the slice can be expressed as:

$$\theta_f = \frac{1}{2} \arctan \frac{i_s}{i_c} = \frac{1}{2} \arctan \frac{i_3 - i_5}{i_6 - i_4} \quad (3-16)$$

This expression produces values of θ_f which are periodic with a wavelength of $\pi/2$ and a range from $-\pi/4$ to $\pi/4$ which corresponds to the range 0 to $\pi/2$ obtained from the classical definition of the isoclinic angle given by Frocht [1941]. The relative phase or retardation is also obtained as a periodic function and is given by:

$$\alpha_f = \arctan \frac{i_s}{i_d \cdot \sin 2\theta} = \arctan \frac{i_c}{i_d \cdot \cos 2\theta} \quad (3-17)$$

Consequently, from Eqs. 3-17 and 3-1 the isochromatic parameter, N is obtained as a periodic function with a wavelength of half a fringe order with $\alpha = 0$ when $N = 1/2, 1, 3/2, \dots$ etc. It should be noted that at $\alpha_f = 0$, $i_3 = i_4 = i_5 = i_6 = i_m$ and the isoclinic parameter, θ_f is undefined. Since the isoclinic parameter is only undefined over relatively thin bands it is possible to extrapolate its values by cubic spline function over them [Kihara 1990 and Wang *et. al.* 1995].

The analysis described above will yield the fractional isoclinic parameter and the fractional retardation at individual points in the field of view without reference to neighbouring points.

3.2.2 The PSIOS System

Because six images are collected sequentially in the conventional transmission polariscope, automated photoelasticity is restricted to static analysis or at least readily repeatable cyclic or dynamic events. A new instrument called PSIOS (Phase-Stepped Images Observed Simultaneously) was designed by Patterson and Wang [1997b]. It is capable of collecting four phase-stepped images simultaneously and can be used for transmission and reflection in both static and cyclic applications. Dynamic applications are also possible if high speed cameras are fitted.

The PSIOS system uses a beam-splitting system to simultaneously grab four phase-stepped images. The cube beamsplitters are within the polariscope and affect the state of polarisation. In optical terms they can be considered as partial linear polarisers with transmission coefficients, K_1 and K_2 , along their principal axes. The beamsplitters are arranged such that their principal axes are aligned parallel and

perpendicular to the reference system and for simplicity the transmission coefficients in the **OX** and **OY** directions have been used, i.e., K_x and K_y respectively. Using described notation, the Mueller matrix, $\mathbf{P}_{K_x K_y}$, for a partial linear polariser is:

$$\mathbf{P}_{K_x K_y} = \frac{1}{2} \begin{bmatrix} K_x + K_y & K_x - K_y & 0 & 0 \\ K_x - K_y & K_x + K_y & 0 & 0 \\ 0 & 0 & 2\sqrt{K_x K_y} & 0 \\ 0 & 0 & 0 & 2\sqrt{K_x K_y} \end{bmatrix} \quad (3-18)$$

In this case, the Stokes vector of the light emerging from the PSIOS system can be expressed as:

$$\mathbf{C}' = \mathbf{P}_\beta \mathbf{Q}_\phi(\xi_2) \mathbf{P}_{K_x K_y} \mathbf{R}_d(\alpha) \mathbf{Q}_\phi(\xi_1) \mathbf{P}_0 \mathbf{S} \quad (3-19)$$

Substituting Eq. 3-18 into Eq. 3-19, the similar expression for the light emerging from each beam,

$$I' = k \cdot c_0' = \frac{i_0}{2} [A + B \sin 2\theta \sin \alpha + C \cos 2\theta \sin \alpha + D \cos \alpha] \quad (3-20)$$

$$\text{where } A = (K_x + K_y) + (K_x - K_y) \cos 2\phi \cos 2(\beta - \phi),$$

$$B = (K_x - K_y) + (K_x + K_y) \cos 2\phi \cos 2(\beta - \phi),$$

$$C = -2\sqrt{K_x K_y} \sin 2\phi \cos 2(\beta - \phi),$$

$$\text{and } D = -2\sqrt{K_x K_y} \sin 2(\beta - \phi).$$

Hence for the orientations of the output elements given in Table 3.1, four of the phase-steps 2, 3, 4 and 5, the intensities observed on each arm of the instrument are:

$$i_a = \frac{i_0}{4} [K_{xa} + K_{ya} + (K_{xa} - K_{ya}) \sin 2\theta \sin \alpha + 2\sqrt{K_{xa} K_{ya}} \cos \alpha] \quad (3-21a)$$

$$i_b = \frac{i_0}{2} [K_{xb} + K_{xb} \sin 2\theta \sin \alpha] \quad (3-21b)$$

$$i_c = \frac{i_0}{2} [K_{yc} - K_{yc} \sin 2\theta \sin \alpha] \quad (3-21c)$$

$$i_d = \frac{i_0}{4} [K_{xd} + K_{yd} + (K_{xd} - K_{yd}) \sin 2\theta \sin \alpha - 2\sqrt{K_{xd}K_{yd}} \cos 2\theta \sin \alpha] \quad (3-21d)$$

In the PSIOS system the light in each arm is split twice. The four light intensities and the corresponding relationship with cubes are shown in Table 3.2.

Table 3.2 Cube beamsplitters and their correspondent coefficients in the PSIOS system

Camera No.	Cube 1	Cube 2	Cube 3	K_x	K_y
<i>a</i>	<i>T</i>	<i>R</i>		$T_x R_x$	$T_y R_y$
<i>b</i>	<i>T</i>	<i>T</i>		$T_x T_x$	$T_y T_y$
<i>c</i>	<i>R</i>		<i>R</i>	$R_x R_x$	$R_y R_y$
<i>d</i>	<i>R</i>		<i>T</i>	$T_x R_x$	$T_y R_y$

The light reached cameras *a* and *d* through transmission and reflection where it is reasonable to assume that $K_x = K_y$ in cameras *a* and *d* (the details will be discussed in Section 4.2.1). To reach the other cameras the light passed through either transmission twice or reflection twice. The four light intensities can be calculated from Eq. 3-21 as:

$$i_a = \frac{i_0}{2} [K_{xa} + K_{xa} \cos \alpha] \quad (3-22a)$$

$$i_b = \frac{i_0}{2} [K_{xb} + K_{xb} \sin 2\theta \sin \alpha] \quad (3-22b)$$

$$i_c = \frac{i_0}{2} [K_{yc} - K_{yc} \sin 2\theta \sin \alpha] \quad (3-22c)$$

$$i_d = \frac{i_0}{2} [K_{xd} - K_{xd} \cos 2\theta \sin \alpha] \quad (3-22d)$$

This leads to a solution for α and θ which is slightly different to those described by Eqs. 3-16 and 3-17:

$$\theta_f = \frac{1}{2} \arctan \frac{\frac{2i_b}{K_{xb} \cdot P} - 1}{1 - \frac{2i_d}{K_{xd} \cdot P}} \quad (3-23)$$

and,

$$\alpha_f = \arctan \frac{1 - \frac{2i_d}{K_{xd} \cdot P}}{\left(\frac{2i_a}{K_{xa} \cdot P} - 1\right) \cdot \cos 2\theta} = \arctan \frac{1 - \frac{2i_c}{K_{xc} \cdot P}}{\left(\frac{2i_a}{K_{xa} \cdot P} - 1\right) \cdot \sin 2\theta} \quad (3-24)$$

where, $P = \frac{i_b}{K_{xb}} + \frac{i_c}{K_{yc}}$.

3.3 System Arrangement of Automated Photoelasticity

Due to rapid development of the small powerful computer and improved numerical techniques in the last two decades, many tedious and time consuming photoelastic techniques have been automated. An automated system is illustrated in Fig. 3.2. A signal is generated by an optical system, recorded by a camera and then transferred via a digitiser to a computer. A computer then processes the data to give the result of the analysis. The following section discusses each of these parts separately.

3.3.1 Optical System and Light Source

The optical system for a circular polariscope generally consists of an polariser, an input quarter-wave plate, a specimen, an output quarter-wave plate, and an analyser. Alternatively, for a plane polariscope the system comprises a polariser, a specimen and an analyser. In phase-stepping systems, the light source should be a monochromatic light with the wavelength matched to the wave retarder plates.

The only difference between transmission and reflection polariscopes is that in a transmission polariscope all the optical elements are arranged on the same axis and on both sides of the specimen; whilst in a reflection polariscope all the elements lie on the same side of the specimen. Fig. 3.3 illustrates two typical arrangement for transmission and reflection polariscopes.

In the transmission system proposed by Patterson and Wang a standard diffuse light polariscope was used with a 15 inch field of view and a sodium light source (Sharples Stress Engineers Ltd, Preston, England). The wavelength of the sodium light is 589.3 nm and the matching wavelength for the quarter-wave plate is 560 nm.

The PSIOS system was capable of observing and capturing all the phase-stepped information simultaneously. Fig. 3.4 is a schematic arrangement of the output sections of the instrument. Circularly polarised light was used as an input to the instrument. The prototype was configured for use in reflection photoelasticity, however, the instrument could be modified to fit in a transmission task by replacing its integral light source with the source and input elements of a transmission polariscope. Having passed through the specimen or a birefringent coating the light is elliptically polarised. A set of lenses were arranged so that a collimated light was produced. The collimated light was then split into four beams of identical intensity which were then

passed through four sets of quarter-wave plates and analysers. These four sets were arranged with different orientations according to the phase-stepped arrangement. Consequently, four images carrying the phase-stepped information were recorded simultaneously by four cameras.

A 300W projector lamp (Thorn-EMI, type: ELH) was selected which best satisfies the requirements for power, portability and integration with optical head. A cold mirror was placed in front of the light source to direct the light beam and to remove the near-infrared radiation. Circularly polarised light was produced by a polariser and a quarter-wave plate positioned in front of the cold mirror. Other optical elements are same as for the transmission polariscope.

An hybrid cube beamsplitter (03 BSC 007, Melles Griot) is a particular requirement for this instrument, having a $45\pm 5\%$ transmission and reflection ratio to the light with a wavelength range from 400 to 700 nm.

3.3.2 Image Acquisition

The image acquisition system comprises a CCD camera, a monitor and a digitiser. The schematic arrangement is shown in Fig. 3.5. The fringe pattern generated by the optical system was focused by the camera lens and projected on to a CCD chip. The CCD camera converted an optical signal to an electronic signal. The signal at this stage was discretised in spatial co-ordinates and analogue with respect to brightness. The output from the camera was then supplied to a Micro Eye TM card (Data translation) which was a transputer-based digitiser. The card was housed in the

computer and provided an analogue signal to a monitor for display and a digitised signal to the computer to form a digital image for further analysis.

The digitiser used here performed two roles, one as an amplifier and the other as an A/D converter. It provided functions to adjust brightness and contrast of the input signal.

The camera used in the transmission polariscope was a Panasonic WV-BP100 and had a resolution of 380×380 pixels. The resolution of the image was 256×256 pixels. The brightness value of each pixel occupied an 8-bit word which can express 256 distinct levels.

Due to the requirements for small size and portability, the Pulnix, TM 526 was chosen for the PSIOS instrument. The system also used the same digitiser (Micro Eye TM 425, Data Translation) and a Multivision Monitor (Taxan 725). A multiplexer (Panasonic Quad WJ 410) device could combine four input signals from the four cameras into a single output signal. The image had a 512×512 resolution and 256 grey levels in this case.

3.3.3 Image Analysis

After data acquisition, 6 or 4 images were obtained by the conventional transmission polariscope or the PSIOS system. Software named PECAN¹ was used for processing the data, its flow diagram is illustrated in Fig. 3.6. Four steps were normally needed

¹ PECAN – Software, named from PhotoElasticity Computer ANalysis, was programmed by Dr. Z. F. Wang and Prof. E. A. Patterson. Their work is gratefully acknowledged by the author.

for processing the raw data to obtain the isochromatic fringe order and isoclinic data. These steps are identified below by the name of the corresponding subroutine used to perform the step.

1. Estimate was for calculating the periodic isochromatic and isoclinic data;
2. Markfind was used to identify marks on the optical elements or specimen system and the jig for holding specimen;
3. Background identified the background to the specimen within the field of view and deleted the data which lay outside the boundary of the slice;
4. Wrapping function extended the range of the isoclinic so that the sign and derivative of the retardation can be found correctly. It also wrapped the periodic values of isochromatic parameter to produce a map of continuous relative retardation.

All the calculations for stress analysis, strain analysis and determination for stress intensity factors were based on the wrapping results of the isochromatic fringe order and the isoclinic angle.

Slightly different processing methods were used in the “Estimate” subroutine for the conventional transmission polariscope compared to the PSIOS system but the other procedures were exactly the same. In the following two sections only the Estimate and Wrapping function were discussed.

3.3.3.1 Estimate

The “Estimate” function was used for calculating the periodic isochromatic and isoclinic data. According to the images obtained, Eqs. 3-16, 3-17 or Eqs. 3-23, 3-24

were used to yield the fractional isoclinic parameter and the fractional retardation at individual points in the field of view without reference to neighbouring points. Consequently the transformation of the periodic form into a continuous function will be considered.

3.3.3.2 Wrapping [Wang and Patterson 1995, Wang 1991]

In the preceding analysis, both the relative retardation, α_f and the isoclinic angle, θ_f were found using an arctangent operator. This resulted in data which lied in the ranges $-\pi/4 \leq \theta_f \leq \pi/4$ and $-\pi/2 \leq \alpha_f \leq \pi/2$.

- Extending the range of the isoclinic angle

The periodic nature of θ_f , which was required in Eqs. 3-17 or 3-24 to evaluate the denominators, caused the reversal of the sign of the relative retardation and its first derivative at locations where $\theta_f = \pm 45^\circ$. These reversals did not exist in stress or strain distributions to which the retardation is related. To overcome this difficulty, the range of the periodic isoclinic angle should be extended to $-\pi/2 \leq \theta_f \leq \pi/2$ before it was used to generate the relative retardation. The field of the isoclinic angle can be considered to consist of two families of curves C_x and C_y which were parallel to the x and y axes respectively. Each curve was made up of a period of data values θ_f and had a discontinuity at each end of height $\pi/2$. The period extension can be easily achieved by scanning the isoclinic data continuously from any arbitrary point in the X - Y plane, and on reaching the discontinuity at the end of a curve. The data of the next

curve was given a $\pi/2$ translation in the Z -direction so the θ , $d\theta/dx$ and $d\theta/dy$ were continuous across the boundary. The scan was continued for the second curve until the discontinuity at its boundary with the next curve was located. The scan continued repeating the whole process until all the data had been considered. It should be pointed out that this procedure generated one of two fields of data depending on the location of the arbitrary starting point relative to the periods of θ_f , i.e., if the starting point was in the second period described above the extended isoclinic angle would be $\pi/2$ out of phase with the one generated by the process described. Physically the two fields of extended data were related to the two principal stress directions respectively.

When the extended data were used to generate the relative retardation, the resulting data was proportional to $(\sigma_1 - \sigma_2)$ everywhere or $(\sigma_2 - \sigma_1)$ everywhere, i.e., the relative retardation data was devoid of any influence from the isoclinic parameter but remained at this stage discontinuous.

- Generating continuous isochromatic retardation

The field of relative retardation data can be considered as two families of curves Dx and Dy parallel to the x and y axes respectively. Each curve was made up of data values α_f and has a discontinuity at each end where $\tan\alpha = \pm\infty$. These curves can be identified by scanning the data from an arbitrary point, although in practice the same starting point was used as that employed when extending the isoclinic data. When a discontinuity was identified all the data already scanned was given a $\pm\pi$ translation in the Z -direction. The direction of translation was dependent on achieving constant the α , $d\alpha/dx$ and $d\alpha/dy$ at the location of the discontinuity. The result of this process was

a continuous field of relative retardation data which may be related to $(\sigma_1 - \sigma_2)$ or $(\sigma_2 - \sigma_1)$. These two possible fields were the inverse of one another.

The relative retardation can be converted to isochromatic fringe order by the provision of the fringe order at two calibration points. The use of a single calibration point allowed relative retardation to be converted to absolute retardation and hence isochromatic fringe order, the provision of the second calibration point allowed the gradient of the $(\sigma_1 - \sigma_2)$ field to be identified so that the appropriate one of the two data fields can be selected, provided the isochromatic fringe order was not identical at two calibration points.

- Fuzzy set classification

The families of curves C and D were rarely well-defined with discontinuities of height precisely $\pi/2$ and π respectively, as a result of noise from various sources. In practice the ends of periods in the relative retardation data were defined by:

$$\left| \frac{d\alpha}{dx} \text{ or } \frac{d\alpha}{dy} \right| \geq \frac{7}{8} \pi \quad (3-25)$$

and the isoclinic parameter by:

$$\left| \frac{d\theta}{dx} \text{ or } \frac{d\theta}{dy} \right| \geq \frac{4}{9} \pi \quad (3-26)$$

The shape of each period of data can be described by its first derivative and value at the mid-point and ends of the period, therefore six fuzzy subsets were defined. By using fuzzy set classification, the isoclinic parameter can be extended and a continuous map of relative retardation can be achieved.

The presence of spikes of noise in the data can also be removed at the same time by following condition,

$$\frac{\pi}{24} < \left| \frac{dz}{dx} \text{ or } \frac{dz}{dy} \right| < \frac{\pi}{4} \quad (3-27)$$

where z could be either the relative retardation or the isoclinic angle. So the spike can be removed if satisfying the above condition and replaced by a value which is the average of the adjacent data points.

3.4 Mathematical Modelling

3.4.1 Introduction

In last section, two methods using the phase-stepping technique were discussed in detail. The validation of the methods could be done by comparing the experimental results obtained with other well-established experimental methods or by using the simulation method. The latter method is preferable because the evaluation of the system accuracy using the experimental method requires a detailed knowledge of the accuracy of the other methods. Comparison of the simulation results and the artificial data is more straightforward.

In this particular case, first the photoelastic parameters (the isochromatic fringe order and the isoclinic angle) were calculated from the theory of elasticity solution and Stress-Optic Law for the specimen being considered. Six or four images were then generated based on the known photoelastic parameters with the assumption of the light intensity values i_m and i_v . The processing of the six or four images based on the use of the PECAN can obtain the fringe order and the isoclinic angle.

Comparing the results from the system with the known information from the simulation allowed the difference in the two sets of data to be identified.

Also, the noises are inevitable in practice and there are multiple stages in processing procedures from photoelastic information in the specimen to the final results. Noise may arise at each stage, and hence the results include all the noise influences. It is therefore difficult to distinguish the effect of each source of noise and so their individual influences are unknown at the moment. It is proposed to model all the sources of noise in order to increase the understanding of the interaction and total effect. The comparison with noise introduced and without noise will clarify the noise influences on the isochromatic fringe order and isoclinic angle and this will allow the procedure to be improved so that noise is reduced. Fig. 3.7 shows a flow diagram of the simulation procedure.

3.4.2 Modelling of A Circular Disc under Diametral Compression Load

A mathematical model was established to simulate the analysis procedure, a circular disc subjected to a diametral compressive load was employed as an example. A Cartesian co-ordinate system was selected such that the origin was at the centre of the disc and the Y-axis was coincident with the loading direction. Its fringe order N and isoclinic angle θ in the disc can be then deduced [Frocht, 1948] as,

$$\frac{2PR}{\pi f t} = N \cdot \frac{(R^2 + x^2 + y^2)^2 - 4R^2 y^2}{(R^2 - x^2 - y^2)} \quad (3-28a)$$

$$\tan\theta = \frac{2xy}{(R^2 - y^2 + x^2)} \quad (3-28b)$$

where P is the load on the specimen, R and t are the radius and the thickness of the disc, respectively, and f is defined as the fringe value of model, which is constant for any particular model and light source and is given by:

$$f = \frac{\lambda}{2 C t} \quad (3-29)$$

where C is stress-optic coefficient. The fringe order N_c at the origin can be calculated by putting $x = y = 0$ in Eq. 3-28a:

$$N_c = \frac{2P}{\pi f t R} \quad (3-30)$$

Referring to the fringe order at the origin, Eq. 3-28a can be rewritten as:

$$N = N_c \cdot R^2 \cdot \frac{(R^2 - x^2 - y^2)}{(R^2 + x^2 + y^2)^2 - 4R^2 y^2} \quad (3-31)$$

An image with 256×256 pixels was employed to describe the fringe order in the circular disc. To maximise the disc dimensions, the diameter of the disc was taken as 255 pixels. A default co-ordinate system $X'-Y'$ in the computer image had $(0, 0)$ at top left which resulted in the centre of the disc having co-ordinate $(127.5, 127.5)$. The Cartesian co-ordinate system $X-Y$ was selected such that the centre of the disc was at $(0, 0)$. The fringe order in the $X'-Y'$ system was described as:

$$N = N_c \cdot 127.5^2 \cdot \frac{(127.5^2 - (x'-127.5)^2 - (y'-127.5)^2)}{(127.5^2 + (x'-127.5)^2 + (y'-127.5)^2)^2 - 4 \cdot 127.5^2 (y'-127.5)^2} \quad (3-32)$$

where $x' = 0, 1, 2 \dots 255,$
 $y' = 0, 1, 2 \dots 255$

In order to restrict the data within the area of the circular disc, an additional condition was applied, such that,

$$\text{when and only when } (x'-127.5)^2 + (y'-127.5)^2 \leq 127.5^2,$$

$$N = N_c \cdot 127.5^2 \cdot \frac{(127.5^2 - (x'-127.5)^2 - (y'-127.5)^2)}{(127.5^2 + (x'-127.5)^2 + (y'-127.5)^2)^2 - 4 \cdot 127.5^2 (y'-127.5)^2}; \quad (3-33)$$

N is undefined else where.

A similar model for describing the isoclinic angle can be derived as,

$$\text{when and only when } (x'-127.5)^2 + (y'-127.5)^2 \leq 127.5^2,$$

$$\tan \theta = \frac{2 \cdot (x'-127.5) \cdot (y'-127.5)}{(127.5^2 - (y'-127.5)^2 + (x'-127.5)^2)}; \quad (3-34)$$

θ is undefined else where.

The geometry and dimension of the disc used during the experimental investigation is shown in Fig. 3.8a. Six raw images of a simulated circular disc with central fringe order $N_c = 1.5$ is illustrated in Fig. 3.8b. The fringe order and isoclinic angle of the disc are illustrated in Fig. 3.8c. These raw images, which do not include any error influences, were generated by combining Eqs. 3-33 and 3-34 with Eq. 3-14 or Eq. 3-22 depending whether the simulated case is for the conventional polariscope or the PSIOS system. The simulation was implemented in MathCAD Version Plus 6.0 Professional Edition (Adept Software Inc., USA). The artificial data was then processed using the subroutine described in outline in Section 3.3.3. As shown in Fig. 3.8c, there is no difference between two sets of data N_n, θ_n and N, θ except in the region having high fringe gradients. The difference in this region is caused by the limitation of the image resolution. So the fringe resolution must be less than 0.26 fringes per pixel and this is comparable with Haake's conclusion, 0.1 fringes per pixel, [Haake *et. al.*, 1993] considering that it was obtained from experiments.

With an error introduced, the raw images are generated based on the general formulation of photoelasticity. The difference between data N_e, θ_e and N_n, θ_n is therefore the influence caused by this kind of error. The possible errors are

investigated in the following chapters so that their consequences could be separated and analysed quantitatively by the simulation method.

3.5 Conclusions

The general formulation was derived based on the method of Mueller Matrices. The theory of phase-stepping systems and their arrangement for hardware were then described. A circular disc under compression is used to simulate the error influence throughout this thesis. The method for generating disc model was provided.

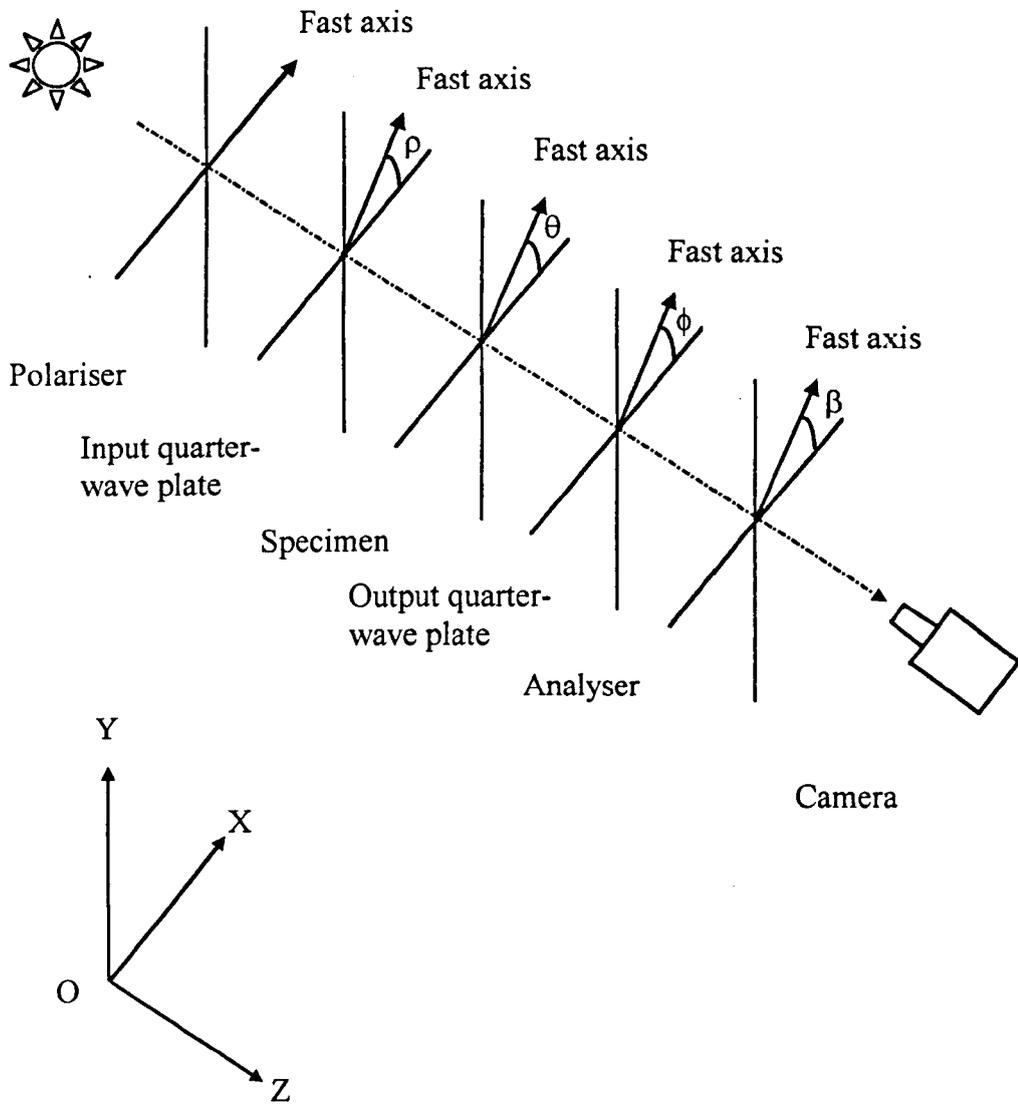


Fig. 3.1 Schematic diagram of a polariscope showing the angular orientations of the elements. The fast axis of the polariser is being used as reference axis.

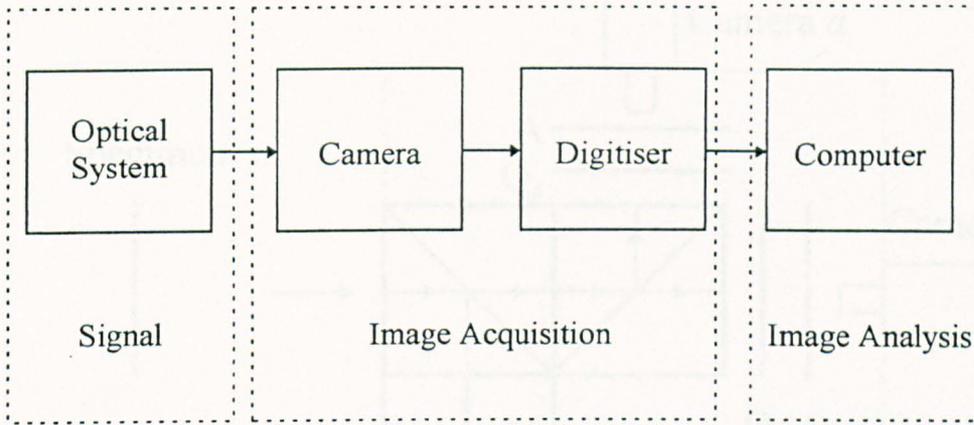
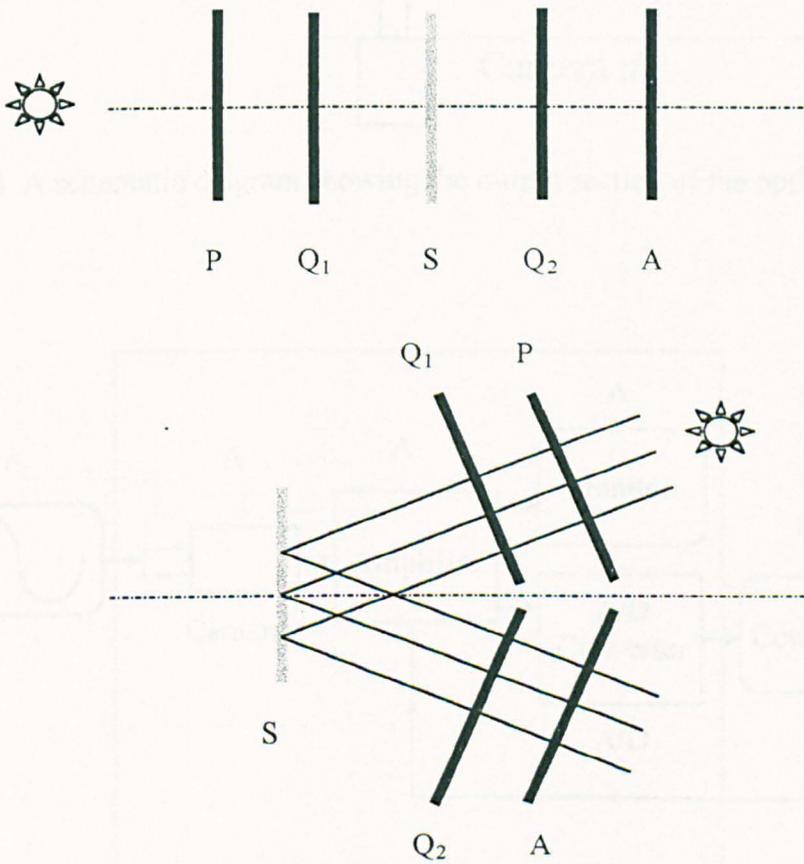


Fig. 3.2 The schematic diagram showing arrangement for the automated photoelasticity system



P: Polariser A: Analyser
 Q₁, Q₂: Input & Output quarter-wave plates
 S: Specimen

Fig. 3.3 Optical arrangement for normal transmission and reflection polariscope

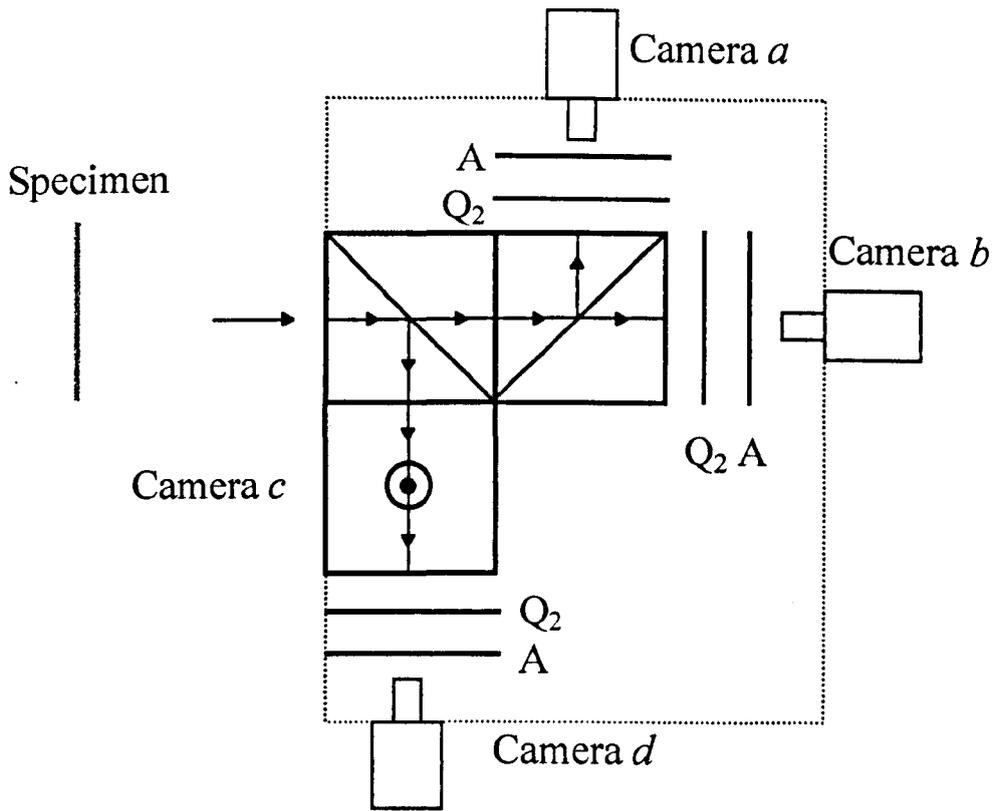
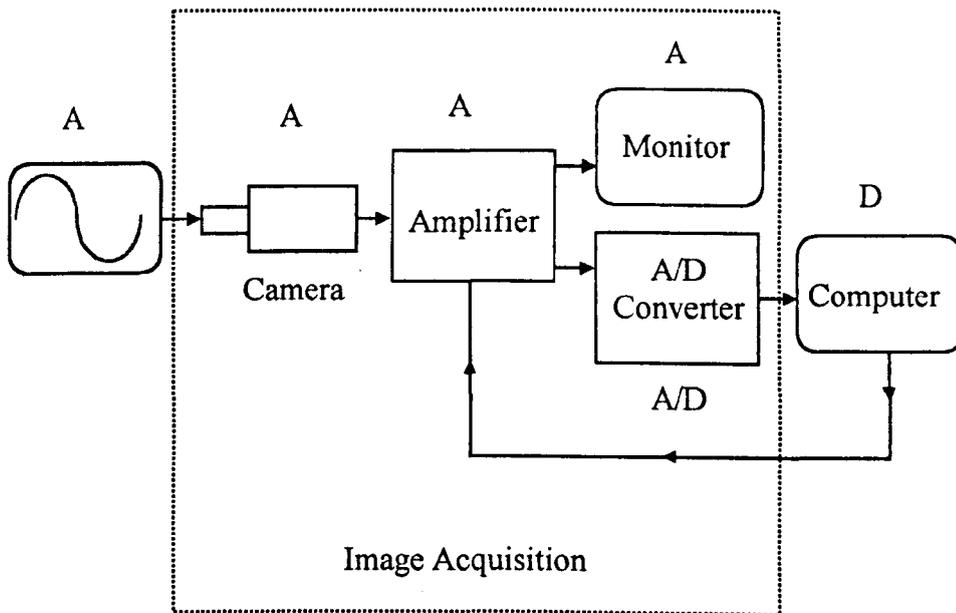


Fig. 3.4 A schematic diagram showing the output section of the optical head



A: Analogue signal

D: Digital signal

Fig. 3.5 A schematic arrangement of the image acquisition

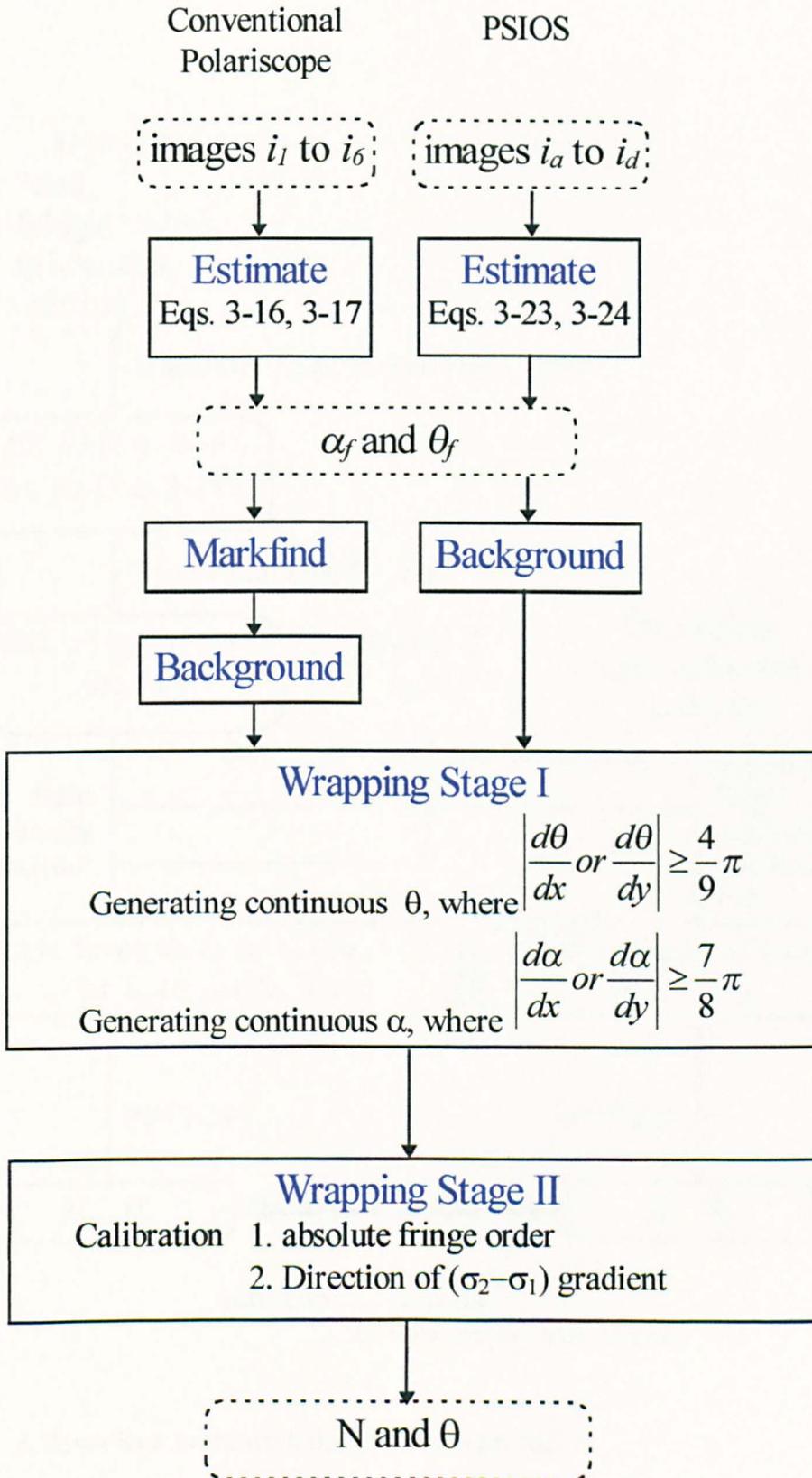


Fig. 3.6 A flow diagram of image analysis procedure (PECAN)

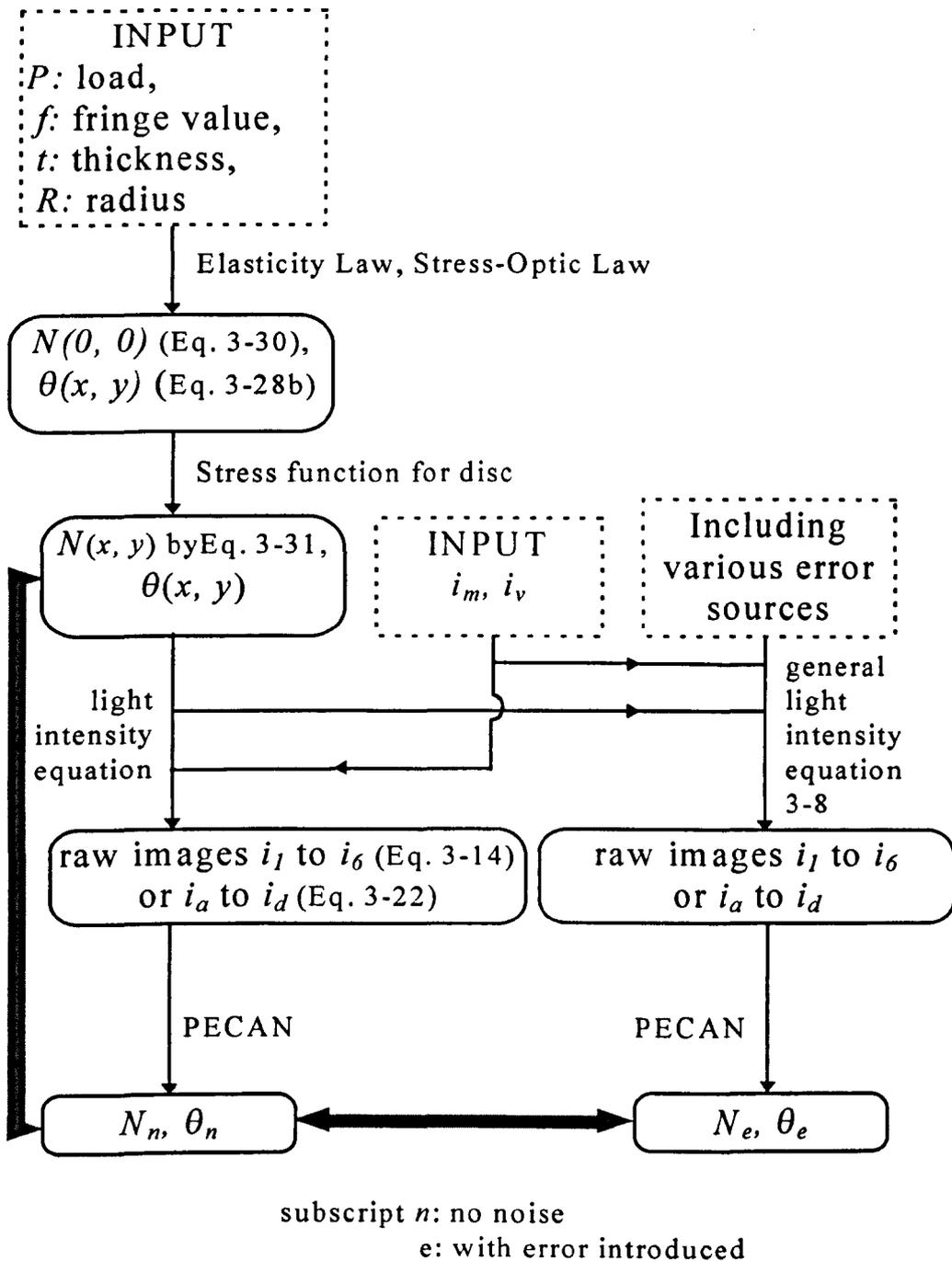


Fig. 3.7 A flow diagram showing the simulation procedure

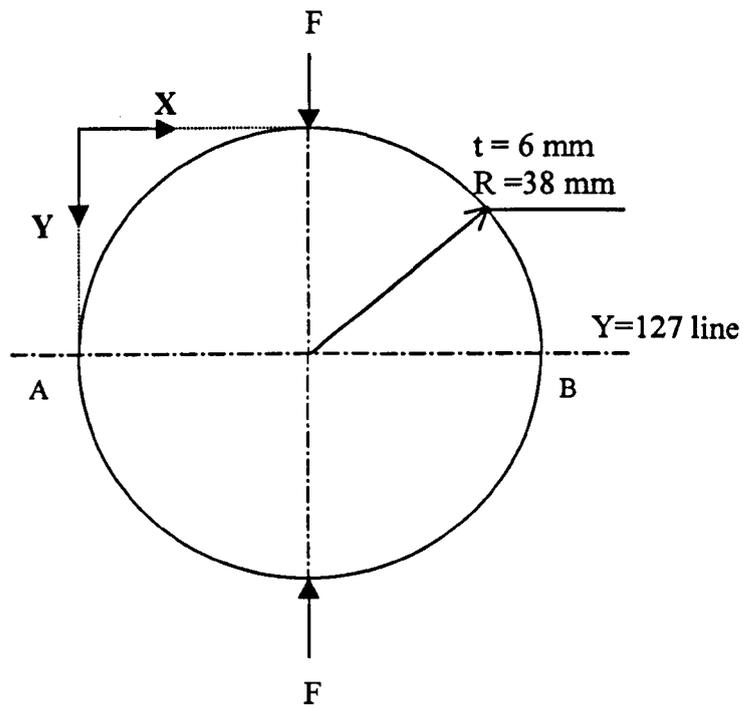


Fig. 3.8a The experimental circular disc under compression, its central line ($Y=127$) AB is perpendicular to the load direction.

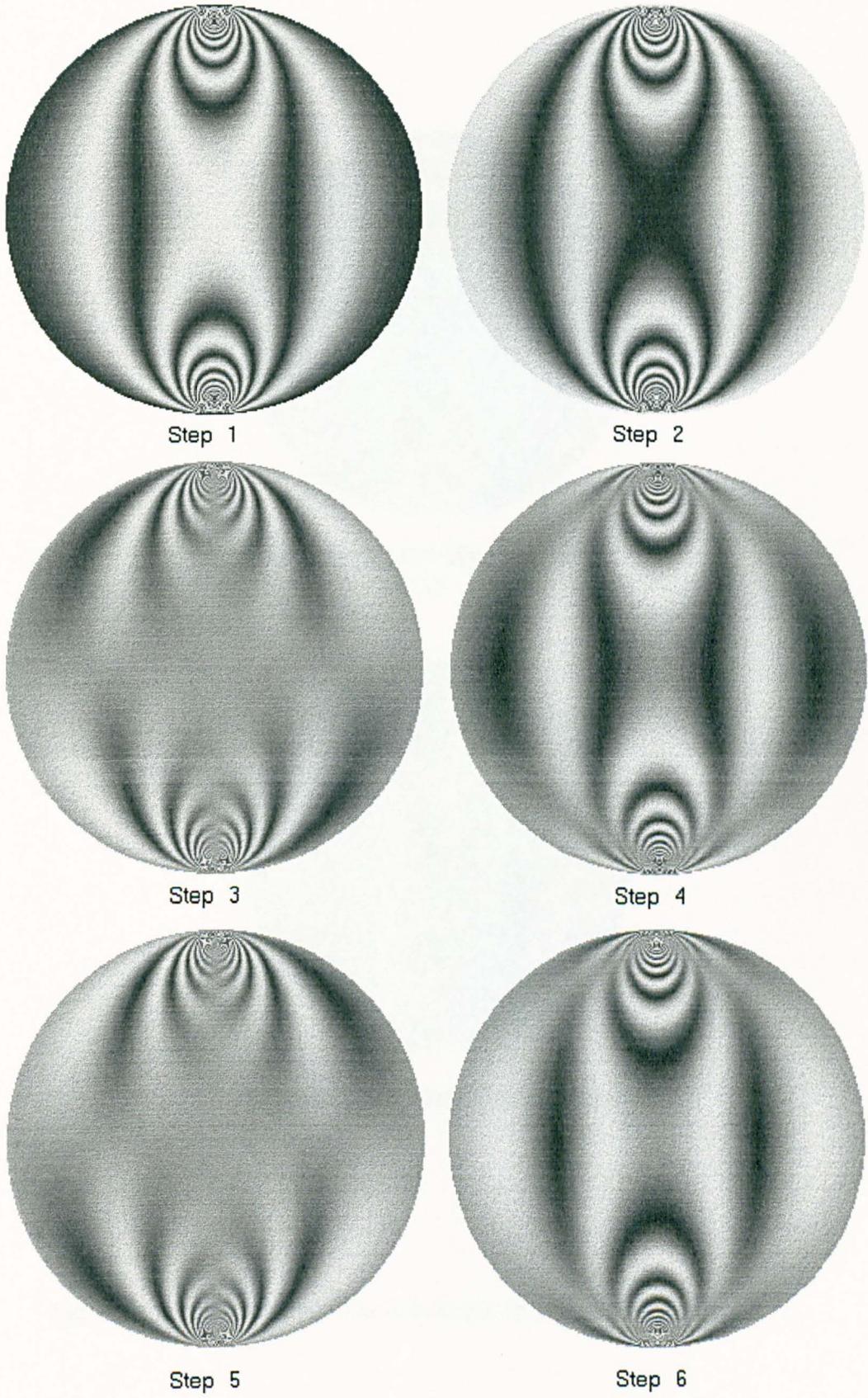


Fig. 3.8b Six raw images of the simulated circular disc. The angles defining each step are detailed in Table 3.1.

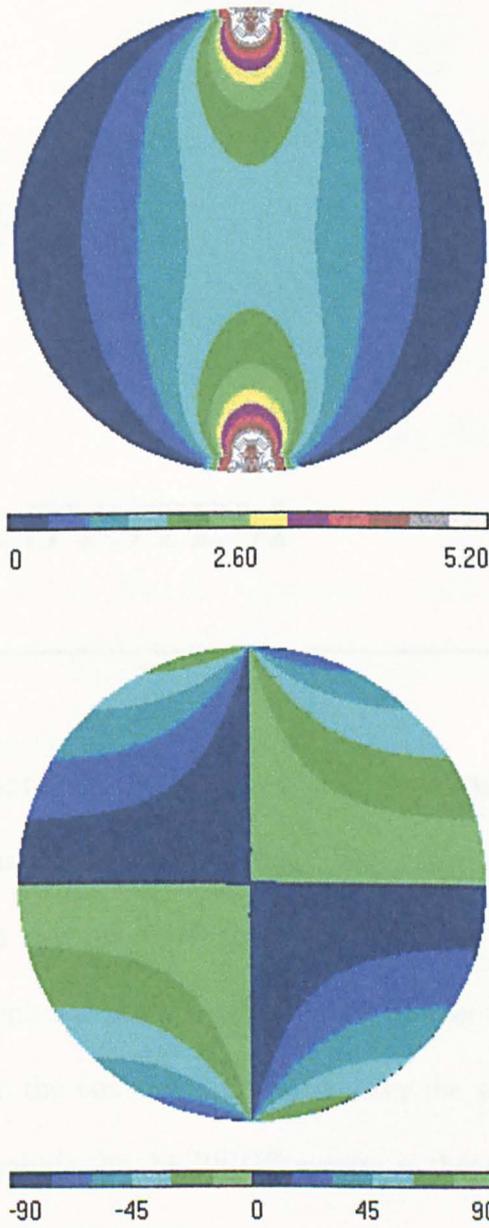


Fig. 3.8b Fringe order and isoclinic angle for the simulated disc

CHAPTER 4

OPTICAL SYSTEM

The subject of this chapter is a discussion of the errors in the optical system, referring to both the conventional polariscope and the PSIOS system. In the latter system, four phase stepped images can be observed simultaneously. All the optical elements employed in the conventional polariscope are also used by the PSIOS system. So all the errors discussed in the conventional system have the same effects in the PSIOS system. The error analysis for the PSIOS system is therefore concentrated on the errors that only existed in this system.

4.1 Errors in Conventional Transmission Phase-stepping System

The optical system normally consists of a polariser, an input-quarter-wave plate, a specimen, an output-quarter-wave plate and an analyser for a circular polariscope or a polariser, a specimen and an analyser for a plane polariscope. For a circular polariscope, errors may be introduced by the quarter-wave plates due to dispersion effects due the fact that the fringe order varies with the wavelength of the light; errors may also arise from a lack of parallelism among the optical elements and it may cause distortion with respect to geometry and the light intensity; from angular errors introduced when the optical elements are rotated to each phase step; from a spatial and temporal variation of the light source and test environment; and finally, from the inhomogeneous properties of the optical elements.

How large are these errors? How do they influence the final results? And how great are their effects? These questions are not essentially answered in the literature. Since material inhomogeneities are difficult to identify unless they are totally nonbirefringent or opaque, they must be avoided when employing automated fringe analysis. The following sections will take the phase-stepping systems proposed by Patterson and Wang [1991] as an exemplar to analyse the first four errors quantitatively and then to develop some approach to modify the experimental system and finally, to improve the accuracy of the stress analysis.

4.1.1 Effect of the Quarter-wave Plate [Dally *et. al.*, 1991]

In the literature review, both a plane polariscope and a circular polariscope were used in the phase-stepping method. The advantage of the plane polariscope is that the errors caused by quarter-wave plates can be removed. The effect of the quarter-wave plate is investigated first.

4.1.1.1 The Optical Property of Quarter-Wave Plate

A wave-plate is an optical element which has the ability to resolve a light vector into two orthogonal components and to transmit the components with different velocities. Such a material is called *doubly refracting* or *birefringent*. The doubly refracting plate illustrated in Fig. 4.1 has two principal axes labelled 1 and 2. The transmission of light along axis 1 proceeds at velocity c_1 and along axis 2 at velocity c_2 . Since c_1 is greater than c_2 , axis 1 is called the *fast axis* and axis 2 the *slow axis*.

If this doubly refracting plate is placed in a field of plane-polarised light so that the light vector E_t makes an angle υ with axis 1 (the fast axis), then on entering the plate the light vector is resolved into two components E_{t1} and E_{t2} along axes 1 and 2 with magnitudes E_{t1} and E_{t2} given by,

$$\begin{aligned} E_{t1} &= E_t \cos \upsilon \\ E_{t2} &= E_t \sin \upsilon \end{aligned} \tag{4-1}$$

The light components E_{t1} and E_{t2} travel through the plate with different velocities c_1 and c_2 ; therefore, the two components emerge from the plate at different times. In other words, one component is retarded in time relative to the other. This retardation produces a relative phase shift between the two components. The linear

phase shifts for components E_{11} and E_{12} with respect to a wave in air can be expressed as:

$$\begin{aligned}\delta_1 &= t(n_1 - n) \\ \delta_2 &= t(n_2 - n)\end{aligned}\quad (4-2)$$

where n is the index of refraction of air, n_1, n_2 are the indices of refraction along axis 1, axis 2 and t is a thickness of the material.

The relative linear phase shift is then simply calculated as:

$$\delta = \delta_2 - \delta_1 = t(n_2 - n_1) \quad (4-3)$$

The relative angular phase shift Δ between the two components as they emerge from the plate is given by,

$$\Delta = \frac{2\pi}{\lambda} \delta = \frac{2\pi t}{\lambda} (n_2 - n_1) \quad (4-4)$$

The relative angular phase shift Δ produced by a doubly refracting plate is dependent upon its thickness t , the wavelength λ , and the properties of the plate as described by $(n_2 - n_1)$. When the doubly refracting plate is designed to give $\Delta = \pi/2$, i.e., $\delta = \lambda/4$, it is called a *quarter-wave plate*.

4.1.1.2 Effect of Quarter-Wave Plate

The quarter-wave plates used to generate circular polarised light are manufactured for a specific wavelength, therefore variation of the light wavelength could cause the relative retardation error ϵ .

In the phase-stepping system investigated here the quarter-wave plates are designed to generate a quarter wave retardation for a light with wavelength 560 nm.

So,

$$\Delta = \frac{\pi}{2} = \frac{2\pi t}{560}(n_2 - n_1) \quad (4-5)$$

The wavelength of the sodium light source can be measured using the Spectral Contents Analysis (SCA) method and it was found to be 592.25 nm. The corresponding phase shift Δ' is given by,

$$\Delta' = \frac{2\pi t}{592.25}(n_2 - n_1) = \frac{\pi}{2} \cdot \frac{560}{592.25} \quad (4-6)$$

Therefore the error ε can be obtained as:

$$\varepsilon = \Delta' - \Delta = \frac{\pi}{2} \left(\frac{560}{592.25} - 1 \right) = -1.485^\circ \quad (4-7)$$

In the formulation for the light intensity Eq. 3-13, the relative phase retardation of the two quarter-wave plates is assumed to be $\pi/2$. When the error ε is included the expression for the intensity is calculated from Eq. 3-8 by,

$$\begin{aligned} I = & i_m + i_v \cdot \sin 2(\beta - \phi) \cdot \left[-\sin 2(\phi - \theta) \cos 2\theta \sin^2 \Delta' - \cos 2(\phi - \theta) \sin 2\theta \sin^2 \Delta' \cos \alpha - \cos^2 \Delta' \cos \alpha \right] \\ & + i_v \cdot \cos 2(\beta - \phi) \cdot \left[-\cos 2(\phi - \theta) \cos 2\theta \sin \Delta' + \sin 2(\phi - \theta) \sin 2\theta \sin \Delta' \cos \alpha - \sin 2(\phi - \theta) \cos \Delta' \sin \alpha \right] \end{aligned} \quad (4-8)$$

The consequence of this error for both photoelastic parameters is plotted as contours in Fig. 4.2 as a function of isoclinic angle and fringe order for one period of each parameter. Owing to the periodic properties of isoclinic angle and relative retardation, only the ranges 0 to 45 degrees for isoclinic angle and 0 to 0.5 fringe order, have been illustrated.

4.1.1.3 Results and Discussion

From Fig. 4.2, the errors in the isochromatic data due to the effect of the quarter-wave plate is quite small, it is always less than 0.015 fringes. But the errors in the isoclinic data are up to 45 degrees at the points of integer and half fringe orders. So the quarter-wave plate makes a big contribution to the error in the isoclinic data.

4.1.2 Linear Alignment and the Error Adjustment

4.1.2.1 Introduction

A circular polariscope was used in the phase-stepping system proposed by Patterson and Wang. The underlying theory for the phase-stepping systems is based on the assumptions that the planes of all the optical elements and of the specimen are parallel to each other and perpendicular to the incident light; and that the optical axis of the camera is parallel to the incident light. In practice, commonly these conditions are achieved by the manual inspection rather than by any specific equipment or procedure and so a considerable lack of parallelism may be introduced. The non-parallelism could cause distortion both in the geometry and in the light intensity.

For example, if a camera with a 5° misalignment to the incident light was used to record an image, the image of a circular specimen would be in deformed shape as shown in Fig. 4.3. Consequently using this arrangement, the measured isochromatic and isoclinic parameters would be distorted.

Meanwhile, the recorded light intensities would also be changed and would also influence the values of isochromatic and isoclinic data. For example, if the quarter-wave plates are not perpendicular to the light beam by an angle χ , as shown in

Fig. 4.4, the distance within the quarter-wave plate that the light components travel is increased. Because the relative retardation is dependent on its thickness, as described in Eq. 4-4, the relative retardation is no longer equal to $\pi/2$. Assuming that the thickness of the wave-plate is denoted by t and the length of the light path is denoted by d , the relation between t and d is given by,

$$d = t / \cos\chi \quad (4-11)$$

Then the corresponding relative retardation is given by χ' ,

$$\chi' = \Delta / \cos\chi \quad (4-12)$$

For a 5° misalignment of the quarter-wave plate, the relative retardation is equal to 90.344° . When this error is added to the retardation for the quarter-wave plates the expression for the light intensity emitted from the polariscope is obtained by letting $\chi' = \Delta'$ in Eq. 4-8. The errors in both the isoclinic and isochromatic parameters resulting from the linear misalignment of 1° and 5° are shown in Fig. 4.5, respectively. The errors in the isochromatic data due to the linear misalignment of the quarter-wave plate are quite small, they are less than 0.00005 fringes for 1° misalignment and 0.001 fringes for 5° misalignment. The errors in the isoclinic data are up to 45 degrees at the points of integer and half fringe orders.

Since the alignment error is a primary error source in the measurement procedure, the resultant error would be amplified in the subsequent stages and cause more serious effects in the final results. It is necessary to make more effort to reduce it.

4.1.2.2 Adjustment of the Camera According to the Polariser Position

Adjustment of the optical system involves the transformation of two co-ordinate systems. A co-ordinate system $X-Y-Z$ is selected such that the $X-Y$ plane is coincident with the plane of the polariser and the Z -axis is perpendicular to the plane of the polariser. Another co-ordinate system $X_c-Y_c-Z_c$ is chosen with respect to the camera, i.e., the origin is at the lens centre and the Z_c -axis is coincident with the optical axis of the camera and its positive direction is towards the light source. As illustrated in Fig. 4.6, adjustment of the optical system requires the transformation of the co-ordinates of points from the $X-Y-Z$ to the $X_c-Y_c-Z_c$ system. The orientation and position of the $X_c-Y_c-Z_c$ system can be expressed in terms of six parameters: three translational parameters (C_x, C_y, C_z), which are the co-ordinates of the lens centre in the $X-Y-Z$ co-ordinate system, and three rotational parameters: ϖ_x, ϖ_y and ϖ_z , which are rotational angles about the X, Y , and Z axes. Before preceding with the development of transformation equations, it is important to define sign conventions. All of the co-ordinate systems should be defined as right-handed. Rotational angles ϖ_x, ϖ_y and ϖ_z are defined as positive if they are counter clockwise when viewed from the positive end of their respective axes. Positive ϖ_x -rotation about the X -axis, for example, is shown in Fig. 4.6.

The transformation of the two co-ordinate systems is developed in the following two basic steps: (1) translation and (2) rotation.

- **Step 1. Translation**

Translation of the origin of the $X-Y-Z$ co-ordinate system to the origin of the $X_c-Y_c-Z_c$ system, representing a new co-ordinate system by $X_1-Y_1-Z_1$, as illustrated in Fig.

4.7. An arbitrary point P with co-ordinates (x, y, z) in the X - Y - Z system can be expressed as $P(x_1, y_1, z_1)$ in the camera co-ordinate system. The relationship between the two co-ordinates should be satisfied by the following equation:

$$\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} -C_x \\ -C_y \\ -C_z \end{pmatrix} \quad (4-13)$$

- **Step 2. Rotation**

The rotational equations are developed in a sequence of three independent two-dimensional rotations. These rotations are illustrated in Fig. 4.8. First the co-ordinate system X_1 - Y_1 - Z_1 is rotated an angle, ϖ_x about the X_1 -axis to form the co-ordinate system X_2 - Y_2 - Z_2 . Second, this new system is rotated ϖ_y about the Y_2 -axis to form the co-ordinate system X_3 - Y_3 - Z_3 . Third, this system is rotated ϖ_z about Z_3 -axis to form the co-ordinate system X_c - Y_c - Z_c . The exact amount and direction will depend upon the orientation relationship between the X - Y - Z and X_c - Y_c - Z_c co-ordinate systems.

The development of the rotational formulae is as follows:

First, rotate through the angle ϖ_x about the X_1 -axis, as illustrated in Fig. 4.9. Since this rotation was about X_1 , the X_1 and X_2 axes are coincident and therefore the x co-ordinate of P is unchanged. The co-ordinates of point $P(x_1, y_1, z_1)$ in first rotated system X_2 - Y_2 - Z_2 , as shown graphically in Fig. 4.9, are:

$$\begin{aligned} x_2 &= x_1 \\ y_2 &= y_1 \cos \varpi_x + z_1 \sin \varpi_x \\ z_2 &= z_1 \cos \varpi_x - y_1 \sin \varpi_x \end{aligned}$$

This equation can be rewritten in matrix form as,

$$\begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \varpi_x & \sin \varpi_x \\ 0 & -\sin \varpi_x & \cos \varpi_x \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} = R_1 \cdot \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} \quad (4-14)$$

In a similar way, rotate through an angle ϖ_y about the Y_2 -axis and third rotate an angle ϖ_z about the Z_3 -axis and then the secondly co-ordinates of P are X_3 - Y_3 - Z_3 and third rotation X_c - Y_c - Z_c system are:

$$\begin{pmatrix} x_3 \\ y_3 \\ z_3 \end{pmatrix} = \begin{pmatrix} \cos \varpi_y & 0 & -\sin \varpi_y \\ 0 & 1 & 0 \\ \sin \varpi_y & 0 & \cos \varpi_y \end{pmatrix} \cdot \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = R_2 \cdot \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = R_2 \cdot R_1 \cdot \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} \quad (4-15)$$

and,

$$\begin{pmatrix} x_c \\ y_c \\ z_c \end{pmatrix} = \begin{pmatrix} \cos \varpi_z & \sin \varpi_z & 0 \\ -\sin \varpi_z & \cos \varpi_z & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_3 \\ y_3 \\ z_3 \end{pmatrix} = R_3 \cdot \begin{pmatrix} x_3 \\ y_3 \\ z_3 \end{pmatrix} = R_3 \cdot R_2 \cdot R_1 \cdot \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} \quad (4-16)$$

respectively. The rotational matrix is then obtained as:

$$R = R_3 \cdot R_2 \cdot R_1 = \begin{pmatrix} \cos \varpi_z & \sin \varpi_z & 0 \\ -\sin \varpi_z & \cos \varpi_z & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \cos \varpi_y & 0 & -\sin \varpi_y \\ 0 & 1 & 0 \\ \sin \varpi_y & 0 & \cos \varpi_y \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \varpi_x & \sin \varpi_x \\ 0 & -\sin \varpi_x & \cos \varpi_x \end{pmatrix} \quad (4-17)$$

The transformation equation from the polariser to camera axis can be obtained by combining the translational and rotational procedures,

$$\begin{pmatrix} x_c \\ y_c \\ z_c \end{pmatrix} = R \cdot \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} = R \cdot \left(\begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} -C_x \\ -C_y \\ -C_z \end{pmatrix} \right) \quad (4-18)$$

Because the rotational equations are multiplicative and the translational equations are additive, it is not convenient for further research. Using Maxwell's method [Maxwell, 1951], Eq. 4-18 can be rewritten as:

$$\begin{pmatrix} x_c \\ y_c \\ z_c \\ 1 \end{pmatrix} = \begin{pmatrix} \cos \varpi_z & \sin \varpi_z & 0 & 0 \\ -\sin \varpi_z & \cos \varpi_z & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \varpi_y & 0 & -\sin \varpi_y & 0 \\ 0 & 1 & 0 & 0 \\ \sin \varpi_y & 0 & \cos \varpi_y & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \varpi_x & \sin \varpi_x & 0 \\ 0 & -\sin \varpi_x & \cos \varpi_x & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & -C_x \\ 0 & 1 & 0 & -C_y \\ 0 & 0 & 1 & -C_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \quad (4-19)$$

So for an arbitrary point on the polariser, the co-ordinates of this point on the camera co-ordinate system can be expressed by the six parameters, C_X , C_Y , C_Z , ϖ_x , ϖ_y and ϖ_z .

In practice the adjustment of the misalignment can be implemented by putting an alignment target with a standard grid on the plane of the polariser so that it is coincident with the plane of the polariser. For an arbitrary point $P(x, y, z)$ on the target, its co-ordinates in the camera system are $P_c(x_c, y_c, z_c)$. Assuming $P_c(x_c, y_c, z_c)$ is observed at $p(x, y)$ in an image plane, the relation between $P_c(x_c, y_c, z_c)$ and $p(x, y)$ is simply formulated as:

$$\begin{aligned} \frac{x}{f} &= \frac{x_c}{z_c - f} \\ \frac{y}{f} &= \frac{y_c}{z_c - f} \end{aligned} \quad (4-20)$$

where f is the focal length of the lens. The above expression can be rewritten in matrix form by,

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{f}{f - z_c} \cdot \begin{pmatrix} x_c \\ y_c \end{pmatrix} \quad (4-21)$$

Because the co-ordinates of the points on the target and their projections on the image are known, it is relatively straightforward to solve Eqs. 4-19 and 4-21 for the six unknown parameters C_X , C_Y , C_Z , ϖ_x , ϖ_y and ϖ_z .

The six parameters represent the distances of translation and angles of rotation of the polariser such that the plane of the polariser is perpendicular to the optical axis of the camera. In practice, the polariser is fixed as a reference, the camera is translated by distance (C_x, C_y, C_z) and rotated by angle $(-\varpi_x)$, $(-\varpi_y)$ and $(-\varpi_z)$ about the **X**, **Y** and **Z** axes, thus the position and orientation of the camera is adjusted for perfect alignment.

4.1.2.3 Adjustment of the Other Optical Elements

In the last procedure, the camera was adjusted so that its optical axis is perpendicular to the plane of the polariser. After fixing the camera, the alignment target can be put on the input quarter-wave plate and the process repeated to align the plate. Because the co-ordinates of points on the target and their projections on the image are known, the orientation of the input quarter-wave plate can be evaluated and adjusted to be perpendicular to the camera axis. The same method is used to adjust the output quarter-wave plate, the analyser and the specimen.

In this manner the distortions caused by the non-parallelism of the optical elements can be reduced to a minimum. The misalignment can be achieved is less than 1 degree. The resulting error are shown in the left panel of Fig. 4.5.

4.1.2.4 Conclusion

Alignment error could cause distortion with respect to the geometry and the light intensity and then influence the shape and accuracy of the isochromatic and isoclinic data, i.e., 5° linear misalignment of the quarter-wave plates could cause 0.001 fringes

error and up to 45 degrees isoclinic angle at the points of integer and half fringe orders.

The method for adjusting the system has been provided. The best alignment that can be achieved in practice is less than 1 degree. The errors resulting from a misalignment of 1° are shown in Fig. 4.5.

4.1.3 Angular Alignment

4.1.3.1 Angular Misalignment due to Output Optical Elements

For transmission photoelasticity in the phase-stepping system, the light intensities are measured at six phase steps, e.g., the light intensity i_1 is measured when $\phi = 0$ and $\beta = 45^\circ$. In practice, it is hard to achieve the precise angles due to the limitation of the instruments and other factors. In all of these cases, the error in the angular position of the optical elements would influence the values of the light intensity hence influence the isochromatic and isoclinic data. When the angles of the output elements have an error $\zeta = 0.1^\circ$ from the standard angles, on the horizontal centre line of a compression disc, loaded vertically, where the isoclinic angle is equal to 0, then the intensities are given in Table 4.2.

Because it is impossible to know exactly the errors in the angular position in practice, otherwise the errors could be directly eliminated, the misalignment error was modelled by using a theoretical value. Assuming that the output quarter-wave plate has $\zeta = 0.1^\circ$ larger than the proposed angle and the analyser has $\zeta = -0.1^\circ$ less than the proposed angle during all the rotating procedures, then the errors in the

isochromatic fringe order and the isoclinic angle for the fringe order of 0 to 0.5 and isoclinic angle of 0 to 45 degree are simulated in Fig. 4.10a.

Table 4.2 The change of light intensity i_1 due to the angular error

i_1	ϕ	β	light intensity
	ζ	$\pi/4 + \zeta$	$i_1 = i_m - i_v \cos\alpha$
	ζ	$\pi/4 - \zeta$	$i_1 = i_m - i_v (\cos 4\zeta \cdot \cos\alpha + \sin 4\zeta \cdot \sin 2\zeta \cdot \sin\alpha)$
	$-\zeta$	$\pi/4 + \zeta$	$i_1 = i_m - i_v (\cos 4\zeta \cdot \cos\alpha + \sin 4\zeta \cdot \sin 2\zeta \cdot \sin\alpha)$
	$-\zeta$	$\pi/4 - \zeta$	$i_1 = i_m - i_v \cos\alpha$
i_2			
	ζ	$-\pi/4 + \zeta$	$i_2 = i_m + i_v \cos\alpha$
	ζ	$-\pi/4 - \zeta$	$i_2 = i_m + i_v (\cos 4\zeta \cdot \cos\alpha + \sin 4\zeta \cdot \sin 2\zeta \cdot \sin\alpha)$
	$-\zeta$	$-\pi/4 + \zeta$	$i_2 = i_m + i_v (\cos 4\zeta \cdot \cos\alpha + \sin 4\zeta \cdot \sin 2\zeta \cdot \sin\alpha)$
	$-\zeta$	$-\pi/4 - \zeta$	$i_2 = i_m + i_v \cos\alpha$
i_3			
	ζ	ζ	$i_3 = i_m - i_v \sin 2\zeta \cdot \sin\alpha$
	ζ	$-\zeta$	$i_3 = i_m + i_v (\sin 4\zeta \cdot \cos\alpha - \cos 4\zeta \cdot \sin 2\zeta \cdot \sin\alpha)$
	$-\zeta$	ζ	$i_3 = i_m - i_v (\sin 4\zeta \cdot \cos\alpha - \cos 4\zeta \cdot \sin 2\zeta \cdot \sin\alpha)$
	$-\zeta$	$-\zeta$	$i_3 = i_m + i_v \sin 2\zeta \cdot \sin\alpha$
i_4			
	$\pi/4 + \zeta$	$\pi/4 + \zeta$	$i_4 = i_m - i_v \cos 2\zeta \cdot \sin\alpha$
	$\pi/4 + \zeta$	$\pi/4 - \zeta$	$i_4 = i_m + i_v (\sin 4\zeta \cdot \cos\alpha - \cos 4\zeta \cdot \cos 2\zeta \cdot \sin\alpha)$
	$\pi/4 - \zeta$	$\pi/4 + \zeta$	$i_4 = i_m - i_v (\sin 4\zeta \cdot \cos\alpha + \cos 4\zeta \cdot \cos 2\zeta \cdot \sin\alpha)$
	$\pi/4 - \zeta$	$\pi/4 - \zeta$	$i_4 = i_m - i_v \cos 2\zeta \cdot \sin\alpha$
i_5			
	$\pi/2 + \zeta$	$\pi/2 + \zeta$	$i_5 = i_m + i_v \sin 2\zeta \cdot \sin\alpha$
	$\pi/2 + \zeta$	$\pi/2 - \zeta$	$i_5 = i_m + i_v (\sin 4\zeta \cdot \cos\alpha + \cos 4\zeta \cdot \sin 2\zeta \cdot \sin\alpha)$
	$\pi/2 - \zeta$	$\pi/2 + \zeta$	$i_5 = i_m - i_v (\sin 4\zeta \cdot \cos\alpha + \cos 4\zeta \cdot \sin 2\zeta \cdot \sin\alpha)$
	$\pi/2 - \zeta$	$\pi/2 - \zeta$	$i_5 = i_m - i_v \sin 2\zeta \cdot \sin\alpha$
i_6			
	$3\pi/4 + \zeta$	$3\pi/4 + \zeta$	$i_6 = i_m + i_v \cos 2\zeta \cdot \sin\alpha$
	$3\pi/4 + \zeta$	$3\pi/4 - \zeta$	$i_6 = i_m + i_v (\sin 4\zeta \cdot \cos\alpha + \cos 4\zeta \cdot \cos 2\zeta \cdot \sin\alpha)$
	$3\pi/4 - \zeta$	$3\pi/4 + \zeta$	$i_6 = i_m - i_v (\sin 4\zeta \cdot \cos\alpha - \cos 4\zeta \cdot \cos 2\zeta \cdot \sin\alpha)$
	$3\pi/4 - \zeta$	$3\pi/4 - \zeta$	$i_6 = i_m + i_v \cos 2\zeta \cdot \sin\alpha$

4.1.3.2 Angular Misalignment Introduced by the Input Quarter-Wave Plate

The above case illustrates the errors in the isochromatic and isoclinic data when only the output optical elements have angular errors. In practice, angular errors may also arise from the position of the input quarter-wave plate. In the phase-stepping system, the axis of the polariser is used as a reference axis. For a left circular polariscope, the angle between the fast axis of the input quarter-wave plate and the reference axis should be (-45^0) . The errors are introduced when this angle is not exactly (-45^0) . Assuming ζ degree errors occurred in this angle, the general form for the light intensity is revised from Eq. 3-8 as:

$$I = i_m + i_v (\sin 2(\beta - \phi) \cdot A_1 - \cos 2(\beta - \phi) \cdot \sin 2(\theta - \phi) \cdot A_2 + \cos 2(\beta - \phi) \cdot \cos 2(\theta - \phi) \cdot A_3)$$

$$\begin{aligned} \text{Where } A_1 &= -\cos 2\zeta \cdot \cos \alpha + \cos 2(\theta - \zeta) \cdot \sin 2\zeta \cdot \sin \alpha \\ A_2 &= -\cos 2\zeta \cdot \sin \alpha - \cos 2(\theta - \zeta) \cdot \sin 2\zeta \cdot \cos \alpha \\ A_3 &= -\sin 2(\theta - \zeta) \cdot \sin 2\zeta \end{aligned} \quad (4-22)$$

When the error ζ is 0.1^0 , the errors of photoelastic data is simulated in

Fig. 4.10b.

4.1.3.3 Angular Misalignment by Input and Output Elements

It is possible that both the input quarter-wave plate and the output elements have angular misalignment in practice. Their effects can be obtained by combining Table 4.2 and Eq. 4-22 and are illustrated in Fig. 4.10c.

4.1.3.4 Conclusion

The three cases of angular misalignment have slightly different influences on the isochromatic fringe order and the isoclinic angle. The misalignment of the output

elements has a larger influence on the fringe order when the isoclinic angle is about 45 degrees and the maximum error could be ± 0.06 fringes; while the misalignment of the input quarter-wave plate has a larger effect on the fringe order when isoclinic angle is about 0 degree and the maximum error could be 0.003 fringes. When both the output and input elements have misalignment errors, the error in the isochromatic fringe order is about 0.006 fringes due to some of the effects cancelling out. For the isoclinic data, the errors are up to 45 degrees in the points near the half and integer fringe orders, while less than 2 degrees in the other areas.

4.1.4 Variation of the Light Source and Test Environment

In Eq. 3-13, which describes the light intensity emitted from the polariscope, the term i_v is the light intensity emerging from the analyser when all the axes of the optical elements and of the specimen are parallel to each other while i_m is to account for the stray light. Based on these assumptions, the light intensities of i_v and i_m should be constant during the sampling procedure. Due to the limitation of the experimental conditions, however, the values of these light intensities may change slightly.

The light source was switched on and its light intensity was measured after half an hour so as to achieve a stable state. The camera was focused on the diffuser of the light source and 100 times samplings were taken. The average values of 100 samples were taken as the true image of the light source. From this “true” image, the light intensities were not uniformly distributed and the central area was brighter than the edge. Meanwhile, the light source changed its intensity value randomly during the sampling time. For an arbitrary pixel, a variation of its intensity value about its mean

is shown in Fig. 4.11. The variations may have come from the camera, the digitiser or from the variations of the light source and test environment. It showed that more than 95% of the variation is in the range of $(-2, +2)$ intensity units over a maximum range of 256 grey levels.

Moving all the apparatus to a dark room may reduce the scatter in the experimental data, but will not eliminate it entirely. The variations of the light intensities i_v and i_m will influence the values of the isochromatic and isoclinic data. Because the variation occurred randomly with time, their effect can not be estimated accurately. The total effect of errors from the measured light intensity are discussed in Section 7.2.1.

4.1.5 Comparison on Modelling Results and Discussion

In the preceding sections, the errors due to the mismatch of the quarter-wave plate, the linear alignment, the angular alignment and the variation of the light source were discussed and their influence on the isochromatic fringe order and isoclinic angle were simulated by MathCad Plus 6.0 Professional Edition. These simulations described the exact errors occurring in the range of 0 to 45 degrees for the isoclinic angle and 0 to 0.5 for the fringe order. It is concluded that the errors in the fringe order are generally small and less than 0.06 fringes and the errors in the isoclinic angle are very serious at the half and integer fringe points but can be neglected in the other areas.

In practice, since the isoclinic angle is undefined at the location of the half and integer fringes, the actual isochromatic and isoclinic data are interpolated over the undefined zones based on their neighbouring points. In order to simulate the actual

processing results, the compressive circular disc was used as a model as discussed in **Section 3.4.2**. The isochromatic fringe order and isoclinic angle contaminated with the errors described in **Sections 4.1.1, 4.1.2, and 4.1.3** are shown in Figs. 4.12-14. The comparison were made along the central line of the disc and is shown in Table 4.3 and Fig. 4.15.

Table 4.3 Optical error sources in the conventional polariscope and their influence on the isochromatic fringe order, N and isoclinic angle, θ along the central line of the disc ($N_c = 1.5$)

Optical Error Sources	Particular Case	Maximum Errors in	
		N (fringe)	θ (degree)
mismatch of $\lambda/4$,	Sodium 592	0.092	31
linear misalignment	5° of $\lambda/4$ misalignment	0.017	4
angular misalignment	0.1° in output elements	0.002	2
	0.1° in input $\lambda/4$	0.013	2
	0.1° in input & output	0.013	2
spatial and temporal variation of light source	95% of the variation is in the range of (-2, +2)	will be discussed in Section 7.2.1	
Material inhomogeneities	Should be avoid in automated fringe analysis since they are difficult to identify unless they are nonbirefringent or opaque		

The simulation of the disc shows that the maximum error results from the use of sodium light and the consequent mismatch with the quarter-wave plates. For the disc with a fringe order of 1.5 at the centre, the maximum error is about 0.092 fringes, which is about 6% in the isochromatic fringe order, and 31 degrees in the isoclinic

angle along the central line of the disc. The 5° linear misalignment would cause errors of 0.017 in the fringe order and 4 degrees in the isoclinic angle. The angular misalignment of the input element has a larger influence than the output elements. The reason is because the comparison is made along the central line of the disc where the isoclinic angle is zero and the input misalignment has a larger effect than the output misalignments in this case (see Fig. 4.10a, b). For instance, for a misalignment 0.1° of the input element the errors are less than 0.013 on the fringe order and 2 degrees on the isoclinic angle.

4.1.6 Conclusions

The influence of the optical errors in a conventional polariscope on the isochromatic fringe order and the isoclinic angle were described by a contour plot for the range of 0 to 45 degrees for the isoclinic angle and 0 to 0.5 for the fringe order and also simulated by a circular disc. The contour plot describes the influence of the errors as a function of the fringe order and the isoclinic angle. The simulation using the disc incorporated real processing through the use of software PECAN. It can be concluded that the use of the sodium light causes the maximum errors in the isochromatic fringe order and the isoclinic angle, namely 0.092 fringes and 31 degrees respectively for the central line of the disc, while the influences of 5° linear misalignment and 0.1° angular misalignment are less than 0.02 on the fringe order and 4 degrees on the isoclinic angle.

4.2 Errors in the PSIOS System

In the conventional polariscope the images were collected sequentially, with the orientation of the elements within the polariscope being changed for each image. This restricts automated photoelasticity to static analysis or at least readily repeatable cyclic or dynamic events. In three-dimensional photoelasticity such a restriction does not represent a constraint due to the nature of the stress freezing process. However the restriction is constraining in two-dimensional photoelasticity and reflection photoelasticity, where fracture and impact studies could be considered.

The PSIOS system [Patterson *et. al.*, 1997b] aims to capture in real-time all the information needed to produce full-field maps of the isochromatic fringe order and the isoclinic angle with a minimum of operator interaction.

Because all the optical elements used in the conventional polariscope were included in the PSIOS system, the errors identified in the conventional polariscope have the same effect in the PSIOS system. Cube beamsplitters were specially required by the PSIOS for generating four beams that are to be phase-stepped. The orientation of these beamsplitters is important for three reasons:

1. The path lengths for all the beams must be the same so that the four images recorded by the CCDs are identical in size.
2. The number of divisions or splits experienced by each beam must be the same so that prior to the output quarter-wave plate, i.e., prior to phase-stepping, the intensities are identical.

3. The beamsplitters have to be arranged so that their effect on the state of polarisation is complementary to that of the output quarter-wave plate and analyser.

Since the last point has been considered during the building of the system, two aspects, namely spatial and intensity are analysed in the following sections.

4.2.1 Spatial Alignment of Four Beams

As illustrated in Fig. 3.4, the circularly polarised light is passed through a birefringent specimen or coating and becomes elliptically polarised then the image is collected by the focal lens. Five identical objective lenses were used within the PSIOS system. The first objective lens is used to generate a collimated light. The collimated light is then split into four beams by three beamsplitters before passing through four sets of quarter-wave plates and analysers respectively. The phase-stepped information is focused by four objective lenses on to four CCD chips. Due to different orientation of the output quarter-wave plates and the analysers, four phase-stepped images are collected simultaneously by the four CCD cameras.

The four images should be identical in spatial position, i.e., any point on the specimen should be always observed at the same co-ordinates in four images. This is very difficult to achieve due to tolerances in the process of manufacturing. In order to minimise the misalignment introduced by operation, the set-up of optical elements should be carefully adjusted with help of a specially designed alignment target which will be described in **Section 4.2.1.1**. The magnification should be first adjusted so that the four images have the identical size for the specimen of investigation. Then on

the basis of best accuracy possible by mechanically aligning, the residual misalignment can be adjusted by the software.

Each camera and its rigidly attached lens were considered as a unit, and their relative alignment was investigated with respect to six degrees of freedom. They were defined as translations along the **X**, **Y** and **Z** axes and rotations about the **X**, **Y** and **Z** axes where the **X**, **Y** and **Z** axes are part of a Cartesian co-ordinate system aligned such that the direction of propagation of light is coincident with the **Z**-direction and the axis of the polariser is coincident with the **X**-direction.

4.2.1.1 Target Design

In the prototype of the PSIOS system, all the optical components, including the camera lenses were contained in a rigid box. The cameras and the single objective lens were screwed onto the outside of the box at their aperture rings. The aim of this design is to achieve a robust structure and keep the alignment firmly fixed.

The alignment target¹ is to provide sufficient spatial sensitivity so that the effective translations and rotations can be measured from its images. To achieve this, the target should have some obvious features. The target size is 86 mm × 54 mm, a standard business card size. The target is split into four small squares by two orthogonal lines with the scale markings. In each square, a special geometrical shape is plotted, as shown in Fig. 4.16.

¹ The author is grateful to Dr. Richard L. Burguete for the useful discussion about the alignment target.

4.2.1.2 Magnification Calibration

All the recorded images should have the same magnification which was defined as the ratio between the real object length and its length represented in the image. For example, there are two points, say A and B, on the specimen, the distance between the projections of A and B on all the recorded images should be the same in terms of number of pixels.

In order to calibrate the magnification ratio, two assumptions were used here. First, plane perspective projection because the object distance is considerable larger than the imaging distance, no spherical distortion is introduced; second, homogeneous magnification over the image plane.

The same magnification ratio for the four images requires a certain level accuracy from the four objective lenses. To check this, two points were chosen in the alignment target, which were at opposite extremes of all four images. The distances between them were calculated for each image. The criteria for choosing the points were, 1) The two points must appear on the four images; 2) They should be near the ends of a diagonal so the magnification ratio on both the X- and the Y-direction are considered.

Points A and B were selected in the target, and their co-ordinates in the four images are shown in Table 4.4. If (X_A, Y_A) and (X_B, Y_B) are used to denote the co-ordinates of points A and B respectively, the distance AB between two points is then equal to ,

$$AB = \sqrt{(X_B - X_A)^2 + (Y_B - Y_A)^2} \quad (4-23)$$

The co-ordinates of these two points and the distance between them are shown in Table 4.4. The maximum difference in the distance AB was between image 1 and image 2, and was 1.01 pixels. Since the distance between AB is almost the longest distance occurring in the field of view and the error is less than 0.5% compared to images 3 and 4 it has been ignored in the phase-stepping analysis.

Table 4.4 The co-ordinates of two points and their distance in four images

	Point A	Point B	Distance AB	Relative error	Percentage error
image 1	(53, 54)	(243, 196)	237.2003	-0.2065	-0.087%
image 2	(61, 72)	(253, 213)	238.2121	0.8053	0.339%
image 3	(16, 61)	(207, 202)	237.4068	0	0
image 4	(43, 62)	(234, 203)	237.4068	0	0

4.2.1.3 Translational Misalignment

Each camera and its rigidly attached lens were considered as a unit, and their relative alignment was investigated with respect to the six degrees of freedom. The degrees of freedom were defined as translation along the **X**, **Y** and **Z** axes and rotations about three axes, where **X**, **Y** and **Z** are a Cartesian system aligned such that the direction of propagation of light is in the **Z**-direction and the axis of the polariser is coincident with the **Y**-direction. The translational tolerances of the camera in the **X**, **Y** and **Z**-axes were all 0.05 mm.

- Translation with respect to the **Z**-axis

The translation in the **Z**-axis would influence the focusing of the images but would not influence the position of the image. For a camera lens with a focal length f , the depth of the field (*DOF*) is obtained as [Grant 1997]:

$$DOF = 4(1 + 1/M)^2 \cdot (f^\#)^2 \cdot \lambda \quad (4-24)$$

or alternatively,

$$DOF = \frac{2F(1 + 1/M)^2 \cdot (f^\#)^2}{10^6 \cdot M} \quad (4-25)$$

where M represents the lens magnification and $f^\#$ is determined by the camera aperture. For example, if the lens aperture $f^\#$ is 8, sodium light is used as source ($\lambda = 589.3$ nm) and magnification is 1, then the depth of field is 0.716 mm. The tolerance on camera position in the **Z**-direction, which is 0.05 mm, is far less than this requirement. So no software is needed to adjusting the translation on the **Z**-axis.

- Translation with respect to the **X**-axis

A translation of any camera in the **X**-axis direction would not affect the focusing of the image from that camera but would make the image shift along the **X**-axis, i.e., the point (x, y) would be $(x-\Delta x, y)$ in the new co-ordinate system. Due to the tolerances during the manufacturing process, ± 0.05 mm is allowed in the **X**-direction, the translation of the camera is inevitable. The ideal is to achieve mechanical alignment of the images to the best accuracy

possible and then correcting any residual misalignment using software. The method for correcting misalignment is discussed in **Section 4.2.1.5**.

- Translation with respect to the **Y**-axis

The translation of any camera along the **Y**-axis has the same tolerance as allowed in the **X**-axis, and would have the similar effect as along the **X**-axis. The same rule is applied to the **Y**-axis. Details are discussed in **Section 4.2.1.5**.

4.2.1.4. Rotational Misalignment

- Rotation with respect to the **X**-axis

The rotation of the camera about the **X**-axis would distort the image. Fig. 4.17 shows a schematic representation of the effect of a rotational misalignment of a camera with respect to the **X**-axis. It can be easily seen that the magnification M is equal to,

$$\begin{aligned} M_1 &= \frac{e_1}{O_1} = \frac{f \cdot \cos \varpi_x}{D_1} = \frac{f \cdot \cos \varpi_x}{D + d \cdot \cos \varpi_x} \\ M_2 &= \frac{e_2}{O_2} = \frac{f \cdot \cos \varpi_x}{D_2} = \frac{f \cdot \cos \varpi_x}{D - d \cdot \cos \varpi_x} \end{aligned} \quad (4-26)$$

where ϖ_x is the angle of the inclination of the optical axis of the camera to the **X**-axis, e and O are the size of the image and object respectively, D is the object distance and f is the focal length of the lens. The subscripts 1 and 2 denote the two corresponding objects. The objects 1 and 2 have the same size in the object plane but different in the image plane, so the ratio between two magnification coefficients is,

$$\frac{M_1}{M_2} = \frac{e_1}{e_2} = \frac{D_2}{D_1} = \frac{D - d \cdot \cos \varpi_x}{D + d \cdot \cos \varpi_x} \quad (4-27)$$

So if a square object is observed by a inclined camera, the image would be a trapezium and the shorter side would correspond to the side with the longer object distance. A simple experiment can be used to check if the camera has inclination with respect to the **X**-axis. A standard graph paper was used as a calibration target, and a square, as large as possible, was chosen in the graph paper. If two sides, which are parallel to the **x**-axis in the image plane, have the same length in terms of the numbers of the pixels, the camera does not have any rotational misalignment; otherwise the camera must be inclined with respect to the **X**-axis. In practice since the camera was fixed on a rigid box, the rotation of the camera about the **X**-axis was restricted by the parallel tolerance of the rigid box. The experiment showed that the error caused by the rotation of the camera about the **X**-axis was negligible.

- Rotation with respect to the **Y**-axis

A similar examination can be made for the two sides of the square which is parallel to the **Y**-axis. Because the rotation of the camera about the **Y**-axis had the same restraints as the **X**-axis, the error caused by the rotation of the camera about the **Y**-axis was also negligible.

- Rotation with respect to the **Z**-axis

The rotational freedom of the camera was fixed by four holes in the **X**- and **Y**-direction and by the screws in the **Z**-direction. In a strict sense, the screw is

not a fixing element, therefore the camera has a much greater freedom about the **Z**-axis compared with the other directions.

The rotation of the camera about the **Z**-axis is directly related to the relative inclination of their corresponding images, as shown in Fig. 4. 19.

4.2.1.5 Calibration of the Translational and Rotational Misalignment

With regard to the misalignment of the cameras, it can be concluded that,

1. The translational misalignment of the cameras along the **X**- and **Y**-axes can not be avoided due to the tolerance in manufacturing process. All these misalignments could cause the translation of the image in the **X**- and **Y**-directions;
2. The translational misalignment of the cameras along the **Z**-axis would affect the focusing of the image. But a comparison of the tolerance allowed in the **Z**-axis with the normal value of *DOF* indicates that this effect can be neglected;
3. The rotational misalignment of the cameras along the **X**- and **Y**-axes was restricted by the mountings of the cameras and their influences are negligible;
4. The rotational misalignment of the cameras about the **Z**-axis, however, could cause the rotation of the corresponding images.

The alignment is first achieved by careful mechanical adjustment of the camera by using the alignment target and the best accuracy possible is obtained. Then the residual misalignment in the image can be corrected using the software. The former

should be a “once-a-lifetime” operation during the building of the instrument, whereas the latter process could be performed each time the instrument is used.

Both image translations and rotations need to be aligned. Two alignment points are chosen and the first point is used for aligning the image translation. After translational alignment, the second point is used to align the image rotations.

- Alignment for image translations

The alignment point A was chosen on the target and was observed on four images at $(a_n, b_n) \mid n=1, 2, 3 \text{ and } 4$, respectively. They should have the same co-ordinates in the four images. Supposing that the image 1 is used as an

alignment datum, $\begin{bmatrix} a_n - a_1 \\ b_n - b_1 \end{bmatrix} \mid n=2, 3 \text{ and } 4$ gives the relative translation of other

images with respect to the image 1. The distances of transform for images 2, 3 and 4 are then given by,

$$\begin{pmatrix} x_n' \\ y_n' \end{pmatrix} = \begin{pmatrix} x_n \\ y_n \end{pmatrix} - \begin{pmatrix} a_n - a_1 \\ b_n - b_1 \end{pmatrix}, \quad n = 2, 3 \text{ and } 4 \quad (4-28)$$

Then the pixels with the same co-ordinates in four images were expected to represent the same point on the specimen. Fig. 4.18 illustrates a schematic diagram of a translational alignment. The shaded area is not common to all the images and hence can not be in the subsequent processing.

- Alignment for image rotations

An arbitrary point P other than A was chosen on the alignment target and its corresponding projection were observed on four images at $(s_n, t_n) |_{n=1, 2, 3 \text{ and } 4}$, respectively. If the co-ordinates are the same, then clearly there is no rotational misalignment in the images, otherwise, the rotational error needs be found in the following way. A schematic diagram of the rotational alignment is shown in Fig. 4.19.

a) Calculating the rotational angle

In image 1, the angle, φ_1 , between the line AP to the horizontal axis is calculated as:

$$\varphi_1 = \arctan \frac{t_1 - b_1}{s_1 - a_1} \quad (4-29)$$

Similarly, the angle φ_2 in image 2 is obtained by,

$$\varphi_2 = \arctan \frac{t_2 - b_1}{s_2 - a_1} \quad (4-30)$$

the difference in the angles is then expressed as:

$$\varphi = \varphi_2 - \varphi_1 = \arctan \frac{t_2 - b_1}{s_2 - a_1} - \arctan \frac{t_1 - b_1}{s_1 - a_1} \quad (4-31)$$

b) Deducing the relationship between two co-ordinates systems

The co-ordinate system $X_2''-Y_2''$, after rotating the angle φ , has the relation with the co-ordinate system $X_2'-Y_2'$ before rotation, as follows:

$$\begin{pmatrix} x_2' - a_1 \\ y_2' - b_1 \end{pmatrix} = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} x_2'' - a_1 \\ y_2'' - b_1 \end{pmatrix} \quad (4-31)$$

c) Assigning the light intensity value

The light intensity value I at the point (x_2'', y_2'') is equal to the values at (x_2', y_2') . The light intensity only has the value at the pixel (integer co-ordinates) and it is possible that the co-ordinates (x_2', y_2') are not integer. Two ways were used for assigning the light intensity value. First, the value x_2' is rounded to m and the value y_2' is rounded to n , then,

$$I(x_2'', y_2'') = I(x_2', y_2') \approx I(m, n) \quad (4-32)$$

Second, the co-ordinates (x_2', y_2') must fall into four pixels, as shown in Fig. 4.20. If m denotes the truncated value of x_2' and n denotes the truncated value of y_2' , then the light intensity at (x_2', y_2') can be expressed by weighted average value as,

$$\begin{aligned} I(x_2'', y_2'') &= I(x_2', y_2') \\ &\approx I(m, n) \cdot (m - x_2' + 1)(n - y_2' + 1) + I(m + 1, n) \cdot (x_2' - m)(n - y_2' + 1) \\ &\quad + I(m, n + 1) \cdot (m - x_2' + 1)(y_2' - n) + I(m + 1, n + 1) \cdot (x_2' - m)(y_2' - n) \end{aligned} \quad (4-33)$$

Obviously, the second method has a better accuracy than the first one but it takes a longer time to calculate.

4.2.1.6 Conclusion

In this section, the alignment target was designed and the methods for aligning the magnification, translational and rotational misalignment were provided. After alignment, a pixel in different images has the one to one correspondent relationship.

4.2.2 Intensity Calibration of Four Beams

Three cube beamsplitters were employed within the optical head of the PSIOS system. All the light beams suffered twice splits prior to the quarter-wave plates and the light intensity must be the same according to the requirement of the theoretical formula. The transmittance and reflectance properties of a single beamsplitter will be discussed first, and then the combined effect of three beamsplitters. The net coefficients on each beam of the instrument will be provided and the allowed tolerance on the four beams will be simulated in the end of this section.

4.2.2.1 Transmission and Reflection Ratio for Single Beamsplitter Cube

The cube beamsplitter is made from BK7 glass and provides equal transmission and reflection. The percent transmittance was provided by the manufacturer (Melles Griot, 03 BSC007) and is described in Fig. 4.21. The transmittance is a function of wavelength and also slightly different along the p-plane and the s-plane². The average transmittance is approximately 48% in the p-plane and 40% in the s-plane. The average reflectance, according to the information provided by the manufacturer, is about 42% in the p-plane and 50% in the s-plane. The allowed tolerance for an individual cube is $\pm 1\%$. If T and R are used to denote the transmission and reflection ratio respectively and a Cartesian co-ordinate system is established such that the Y-axis is coincident with the s-plane, then the above ratio can be expressed as Table 4.5.

² The p-plane is defined where the plane of polarisation is parallel to the plane of incidence while the s-plane is the plane when the polarisation plane is perpendicular to the plane of incidence.

Table 4.5 The transmission and reflection ratios of the cube beamsplitter

T_x	T_y	R_x	R_y	tolerance	absorption
48%	40%	42%	50%	$\pm 1\%$	10%

4.2.2.2 Combined Effect of the Beamsplitters

The cube beamsplitter can be considered as a partial linear polariser and its Stokes vector can be expressed as,

$$P_{K_x, K_y} = \frac{1}{2} \begin{bmatrix} K_x + K_y & K_x - K_y & 0 & 0 \\ K_x - K_y & K_x + K_y & 0 & 0 \\ 0 & 0 & 2\sqrt{K_x K_y} & 0 \\ 0 & 0 & 0 & 2\sqrt{K_x K_y} \end{bmatrix} \quad (3-18)$$

where K_x and K_y are the transmission coefficients along the **X** and **Y** axes, respectively. Because the cube beamsplitters were connected in series and they can be taken as a set, the output from the set is the product of the transmission ratios [Levi, 1968]. Thus, the combined effects are expressed as,

$$T \cdot T = \frac{1}{2} \begin{bmatrix} T_x^2 + T_y^2 & T_x^2 - T_y^2 & 0 & 0 \\ T_x^2 - T_y^2 & T_x^2 + T_y^2 & 0 & 0 \\ 0 & 0 & 2\sqrt{T_x^2 \cdot T_y^2} & 0 \\ 0 & 0 & 0 & 2\sqrt{T_x^2 \cdot T_y^2} \end{bmatrix} \quad (4-34)$$

$$R \cdot R = \frac{1}{2} \begin{bmatrix} R_x^2 + R_y^2 & R_x^2 - R_y^2 & 0 & 0 \\ R_x^2 - R_y^2 & R_x^2 + R_y^2 & 0 & 0 \\ 0 & 0 & 2\sqrt{R_x^2 \cdot R_y^2} & 0 \\ 0 & 0 & 0 & 2\sqrt{R_x^2 \cdot R_y^2} \end{bmatrix} \quad (4-35)$$

$$T \cdot R = R \cdot T = \frac{1}{2} \begin{bmatrix} T_x R_x + T_y R_y & T_x R_x - T_y R_y & 0 & 0 \\ T_x R_x - T_y R_y & T_x R_x + T_y R_y & 0 & 0 \\ 0 & 0 & 2\sqrt{T_x R_x \cdot T_y R_y} & 0 \\ 0 & 0 & 0 & 2\sqrt{T_x R_x \cdot T_y R_y} \end{bmatrix} \quad (4-36)$$

Four CCD cameras were orientated as shown in Fig. 3.4, then the image in camera *a* was obtained by transmission in cube 1 and reflection in cube 2, i.e., the transmission coefficients equals $T_x R_x$ in the X-direction and $T_y R_y$ in the Y-direction. All the coefficients in the four images were deduced in a similar way and are shown in Table 4.6.

Table 4.6 The net coefficients in four images

Camera No.	Cube 1	Cube 2	Cube 3	net coefficients	
				K_x	K_y
<i>a</i>	<i>T</i>	<i>R</i>		$T_x R_x$	$T_y R_y$
<i>b</i>	<i>T</i>	<i>T</i>		$T_x T_x$	$T_y T_y$
<i>c</i>	<i>R</i>		<i>R</i>	$R_x R_x$	$R_y R_y$
<i>d</i>	<i>R</i>		<i>T</i>	$T_x R_x$	$T_y R_y$

The transmission coefficients of the individual cube beamsplitters were specified by the manufacturer. Therefore the net coefficients for each arm of the instrument can be evaluated and are shown in Table 4.7.

Table 4.7 The numerical values of transmission : reflection ratio in four images within the polariscope

Camera	net coefficients		allowed tolerance	
No.	K_x	K_y	K_x	K_y
<i>a</i>	20.16%	20%	(19.27~21.07)%	(19.11~20.91)%
<i>b</i>	23.04%	16%	(22.09~24.01)%	(15.21~16.81)%
<i>c</i>	17.64%	25%	(16.81~18.49)%	(24.01~26.01)%
<i>d</i>	20.16%	20%	(19.27~21.07)%	(19.11~20.91)%

Referring back to Eqs. 3-21, 3-23 and 3-24, it is reasonable to assume that $K_{xz} = K_{yz}$ and $K_{xd} = K_{yd}$ to simplify the calculation.

4.2.2.3 Effects by Allowed Tolerance in Four Beams

The nominal or specified transmission and reflection ratio for a cube are listed in Table 4.5. As a result of manufacturing, individual cubes may have a $\pm 1\%$ tolerance in transmission or reflection ratio, the “true” splitting ratio on each camera axis could be any value lying in the range in Table 4.7. This section is therefore to discuss the influence of this tolerance on phase-stepping.

Again, the circular disc is used in the simulation model. For the case without any error, referring back to Fig. 3.7, the four images coefficients used in Eq. 3-22 were correspondent to the coefficients accepted by PECAN, and they are listed in Table 4.8 in the left column. Hence no errors would be introduced. However, for the case considering the tolerance allowed in the cube, the “true” ratios are within the range of tolerance. The worst case was chosen in the range of the tolerance and is

expressed by bold font, as listed in the right column of Table 4.8. So the value in bold font was used in Eq. 3-22, but the data processing through the use of PECAN still used the coefficients as listed in the left column of Table 4.8. The analysis of results for the isochromatic fringe order and isoclinic angle while considering the allowed tolerance are shown in Fig. 4.22 and a comparison along the central line of disc for two cases is shown in Fig. 4.23.

Table 4.8 The net coefficients of four images within the polariscope

Camera No.	net coefficients		true ratio range	
	K_x	K_y	K_x	K_y
<i>a</i>	20%	20%	(19.27~21.07)%	(19.11~20.91)%
<i>b</i>	23%		(22.09~24.01)%	(15.21~16.81)%
<i>c</i>		25%	(16.81~18.49)%	(24.01~26.01)%
<i>d</i>	20%	20%	(19.27~21.07)%	(19.11~20.91)%

4.2.2.4 Discussion and Conclusion

From Figs. 4.22 and 4.23, it can be seen that the tolerance of the beamsplitters results in almost no errors in the isochromatic fringe order. However, there is obviously an influence on the isoclinic angle. And the error is less than 5 degrees if the undefined zones of the isoclinic angle are ignored.

4.2.3 Oblique Viewing

In the reflection polariscope, as shown in Fig. 3.3b, the light source and the camera were put on the same side of the specimen and this introduced an incidence and observation angle. A number of researchers [Redner 1963, Hung *et. al.* 1990 and Chaufhari *et. al.* 1990] investigated this problem and Patterson and Wang [1997b] reported their view that the angles of viewing or obliquity should be maintained below 10 degrees so that no correction is necessary.

When the PSIOS system is used for the reflection, oblique viewing is introduced. The distance between the light source and the camera is 75 millimetres. For 1 meter distance between the instrument and the specimen, which is normally used in practice, 2.15 degrees oblique viewing angle was introduced. Since this angle is less than 10 degrees, no correction is needed for this case.

4.2.4 Conclusions

The great advantage of the PSIOS system is that it can capture four phase-stepped images simultaneously. The use of the four cameras requires a careful alignment for the spatial position. A target was designed for an initial visual alignment and any residual misalignment was calibrated by the software.

First, the same magnification is required by four recorded images so that the images of the specimen are identical in size. Normally this can be achieved by using the same object lenses. Six degrees freedom were investigated respectively. The translational misalignment along the **Z**-axis and rotational misalignment about the **X**- and **Y**-axes are sufficiently small to be neglected; the translational misalignment along

the **X**- and **Y**-axes could cause the translation of the images in the **X**- and **Y**-direction, and the rotational misalignment about the **Z**-axis could result in the inclination of the corresponding image. Two points are needed for aligning the image for translation and rotation and the method for calibration and assigning the intensity values for the new pixels were provided.

The same light intensity was also required by the four beams and the modified coefficients were calculated for the beamsplitting unit. The errors in the isochromatic and the isoclinic data were estimated using the simulation of a disc in compression. This error has almost no influence on the fringe order and typically a maximum of 5 degrees error in the isoclinic angle if the undefined zones are ignored.

Oblique viewing is a problem for the reflection polariscope but for less than 10 degrees incident angle the effects were thought to be negligible.

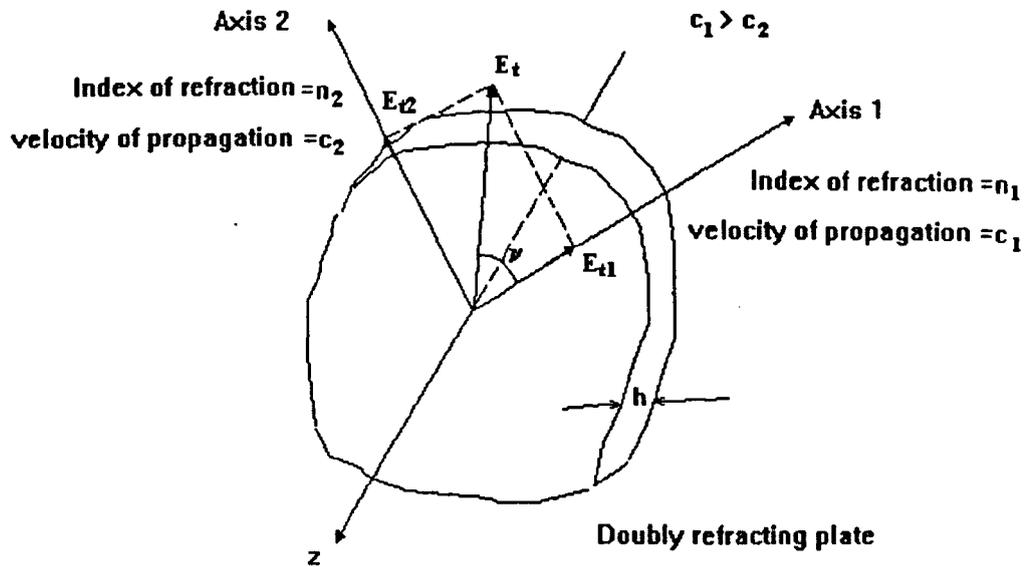


Fig. 4.1 A plane-polarised light vector entering a doubly refracting plate

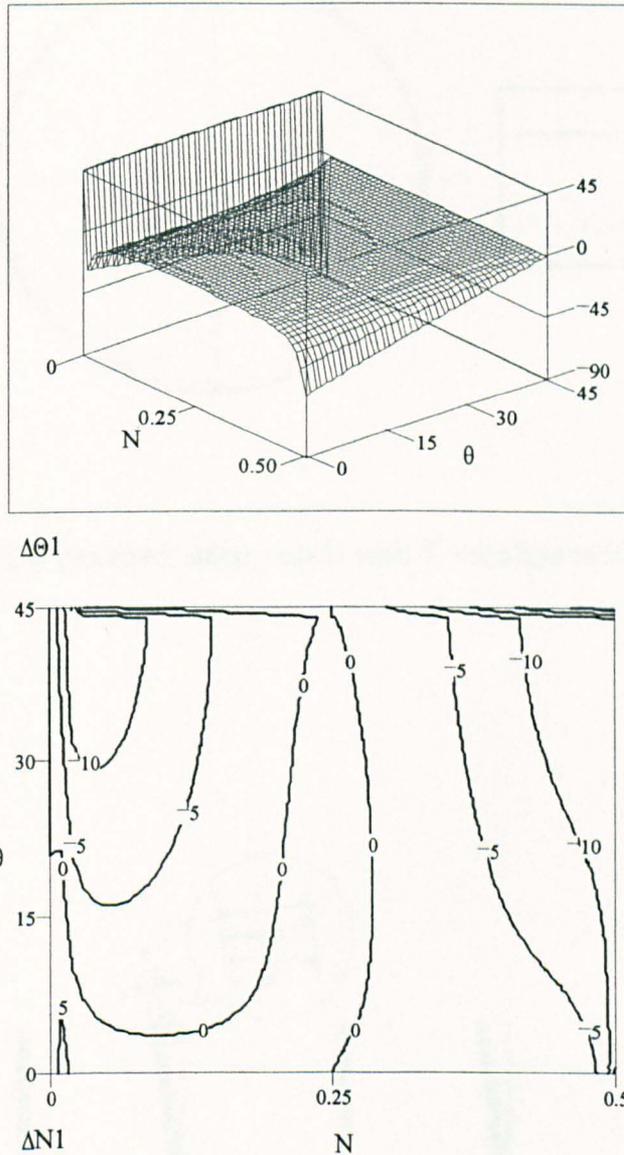


Fig. 4.2 Effect of wavelength mismatch on the isoclinic (top) and isochromatic (bottom) parameters. The wavelength of sodium light is 592.25 nm while the matching wavelength for quarter-wave plates is 560 nm. The errors are shown in absolute terms on the z-axis as functions of both isoclinic and isochromatic parameters on the x- and y- axes respectively. The units of fringe order errors are 0.001 fringes.

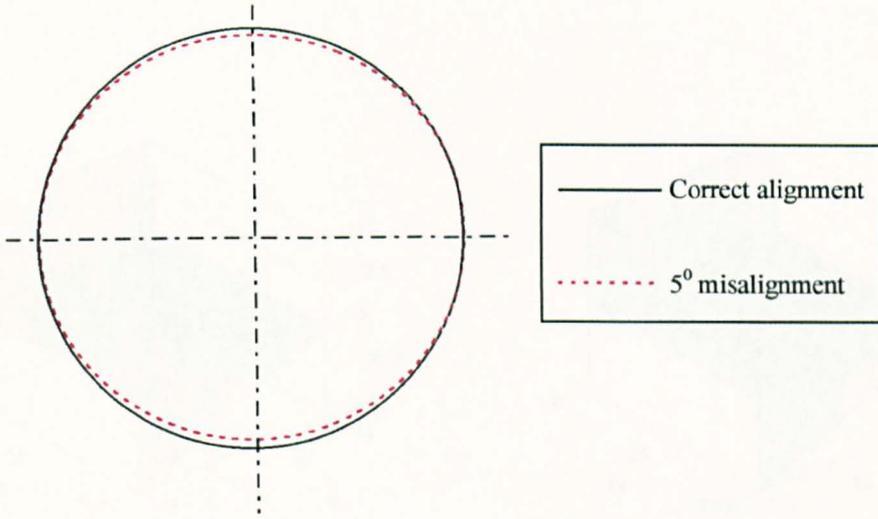


Fig. 4.3 The distorted unitary circle with 5° misalignment to the incident light

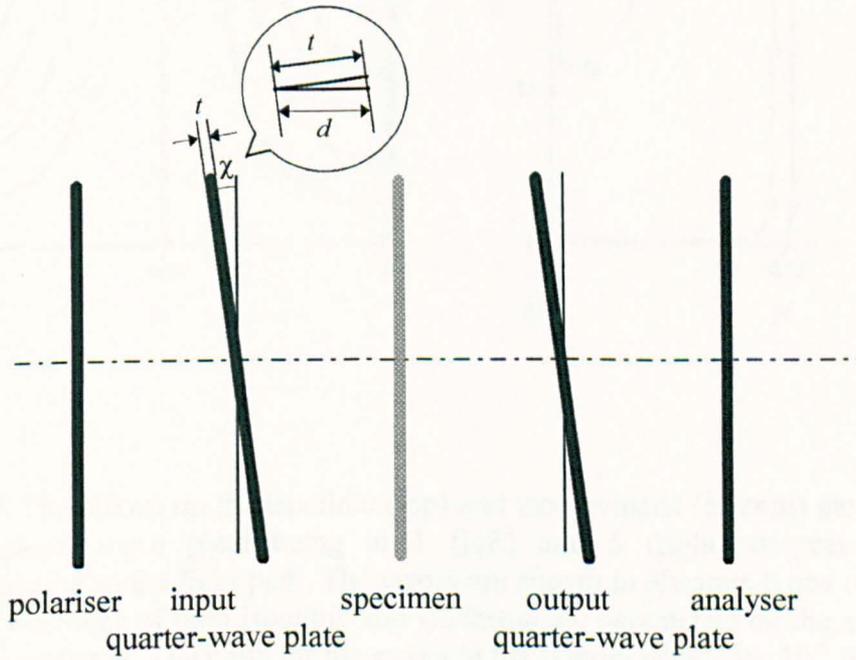


Fig. 4.4 Linear misalignment of the quarter-wave plate with an angle χ for the phase-stepping system

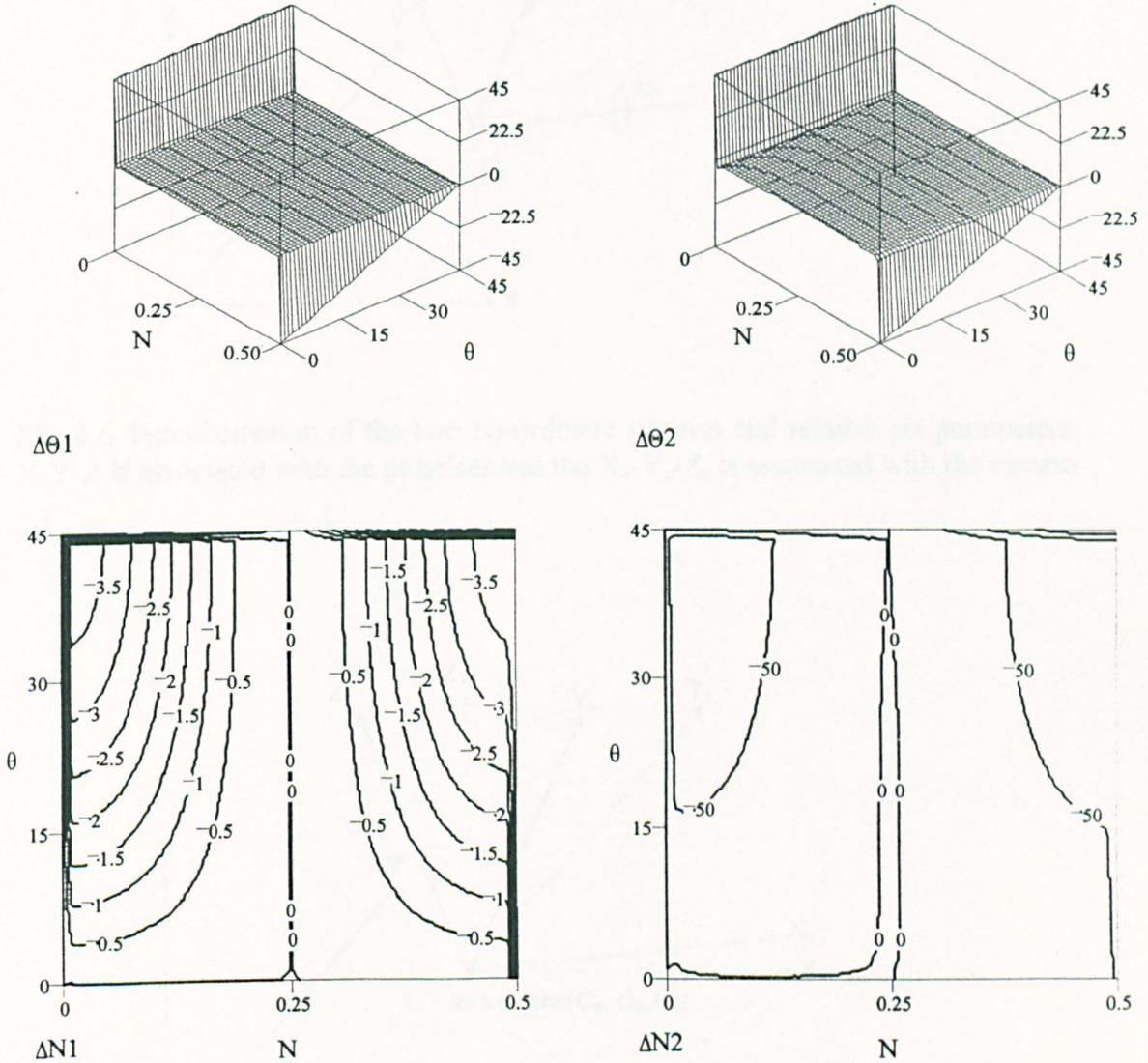


Fig. 4.5 The effects on the isoclinic (top) and isochromatic (bottom) parameters of a quarter-wave plate being at 1 (left) and 5 (right) degrees to the perpendicular to the light path. The errors are shown in absolute terms on the z-axis as functions of both isoclinic and isochromatic parameters on the x- and y-axes respectively. The units for the errors in the bottom panels are 10^{-5} fringes.

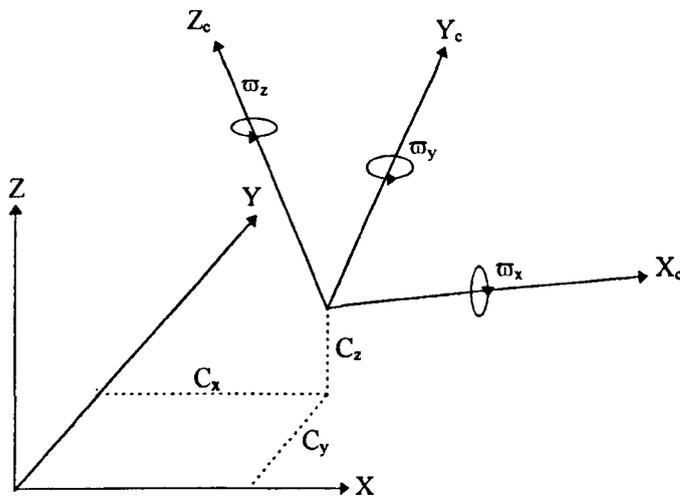


Fig. 4.6 Transformation of the two co-ordinate systems and relative six parameters. X-Y-Z is associated with the polariser and the X_c-Y_c-Z_c is associated with the camera

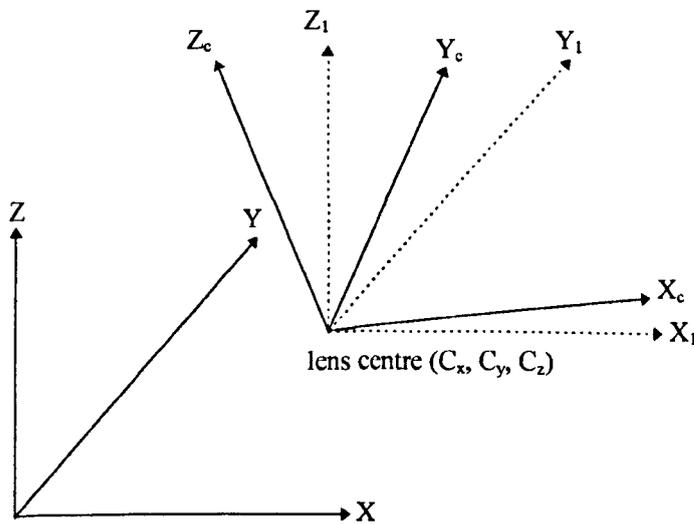


Fig. 4.7 Translation of the co-ordinate system X-Y-Z to the origin of the camera co-ordinate system X_c-Y_c-Z_c to create a new co-ordinate system X₁-Y₁-Z₁

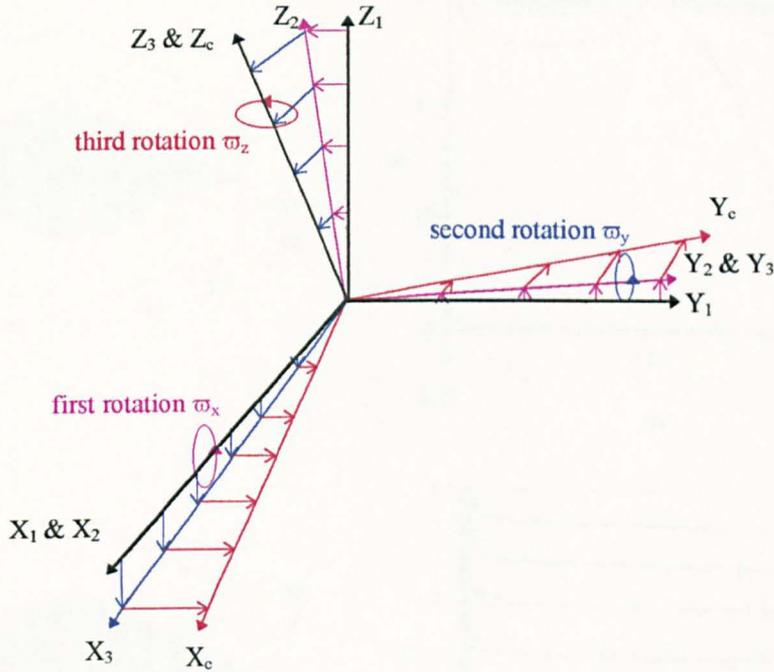
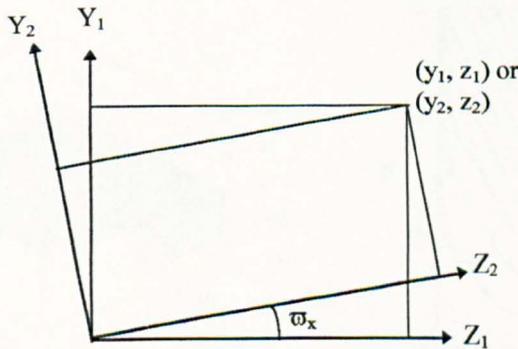


Fig. 4.8 The three sequential angular rotation of the co-ordinate systems. First the co-ordinate system $X_1-Y_1-Z_1$ is rotated ω_x about the X_1 -axis to form the co-ordinate system $X_2-Y_2-Z_2$. Second this new system is rotated ω_y about the Y_2 -axis to form the co-ordinate system $X_3-Y_3-Z_3$. Third, this system is rotated ω_z about Z_3 -axis to form the co-ordinate system $X_c-Y_c-Z_c$.



$$y_2 = y_1 \cos \omega_x + z_1 \sin \omega_x$$

$$z_2 = z_1 \cos \omega_x - y_1 \sin \omega_x$$

Fig. 4.9 Rotation through an angle ω_x about the X_1 axis of the co-ordinate system $X_1-Y_1-Z_1$ to form a new co-ordinate system $X_2-Y_2-Z_2$

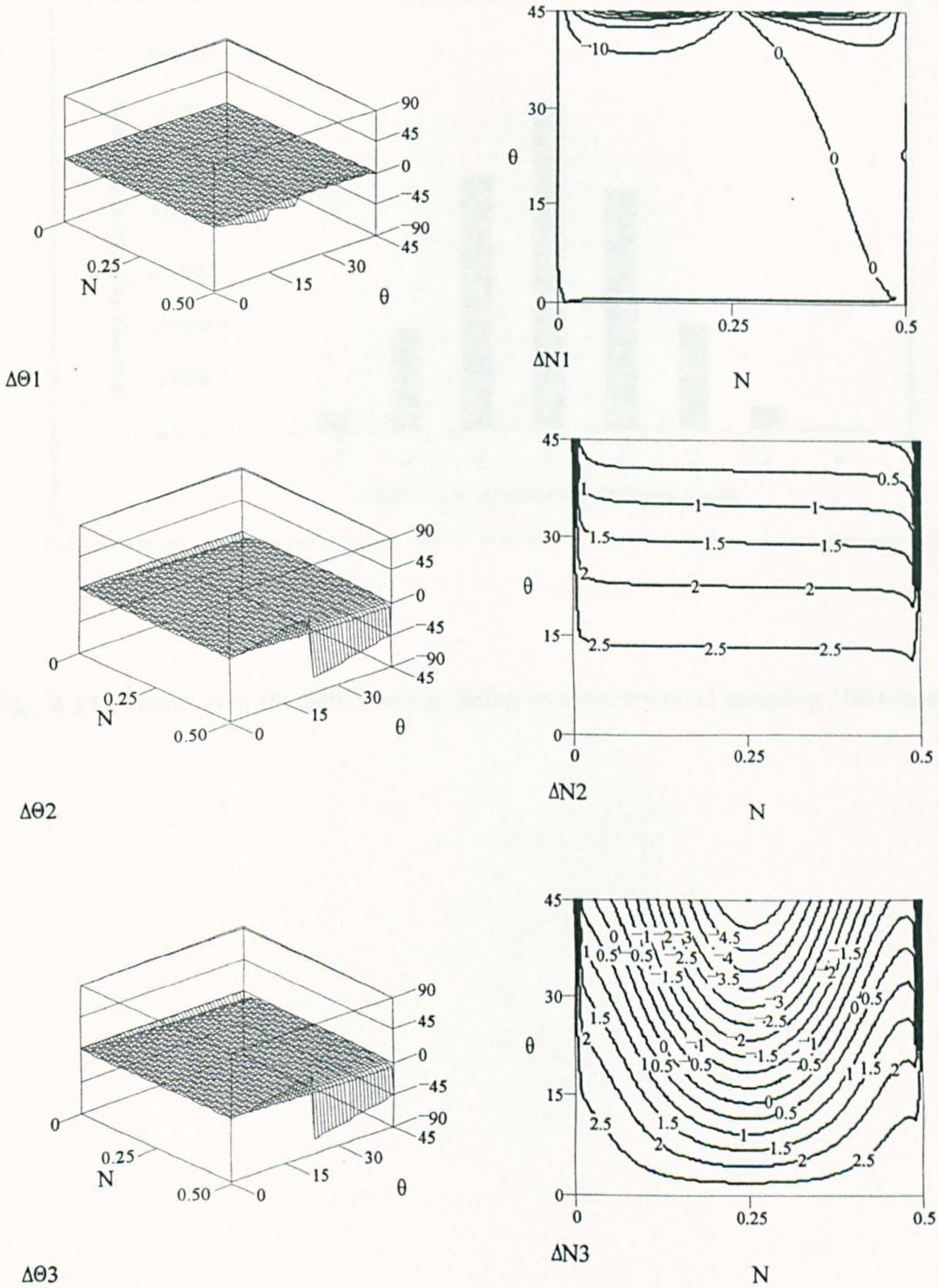


Fig. 4.10 The effects on the isoclinic (left) and isochromatic (right) parameters due to the angular misalignment of optical elements. The units for the isochromatic fringe order are 0.01 fringes.

- a: (top panel) errors due to 0.1° misalignment of the output elements;
- b: (middle panel) errors due to 0.1° misalignment of the input element;
- c: (bottom panel) errors due to 0.1° misalignment of both the input and the output elements;

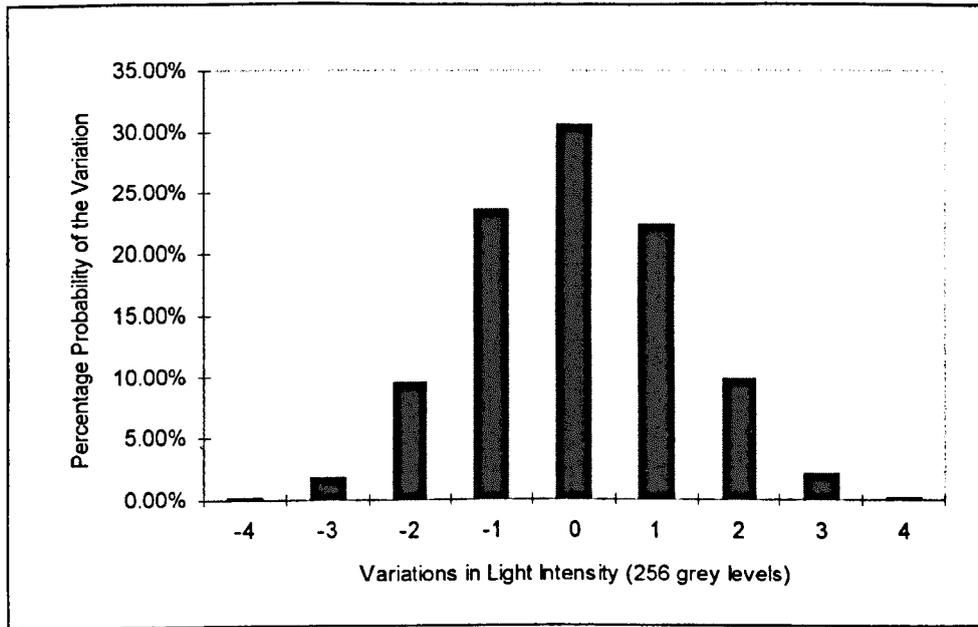


Fig. 4.11 Variation of the light intensity using an arbitrary pixel sampling 100 times.

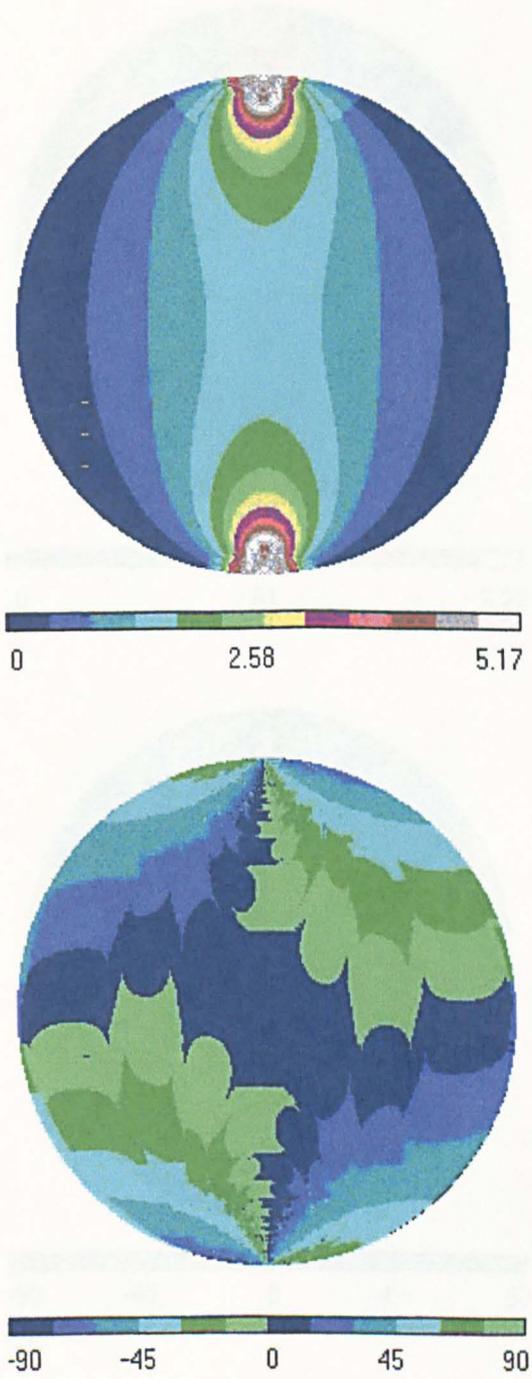


Fig. 4.11 The modelling results of the fringe order and isoclinic angle for a circular disc under a sodium light source

Fig. 4.12 The modelling results of the fringe order and isoclinic angle for a circular disc due to the use of sodium light source

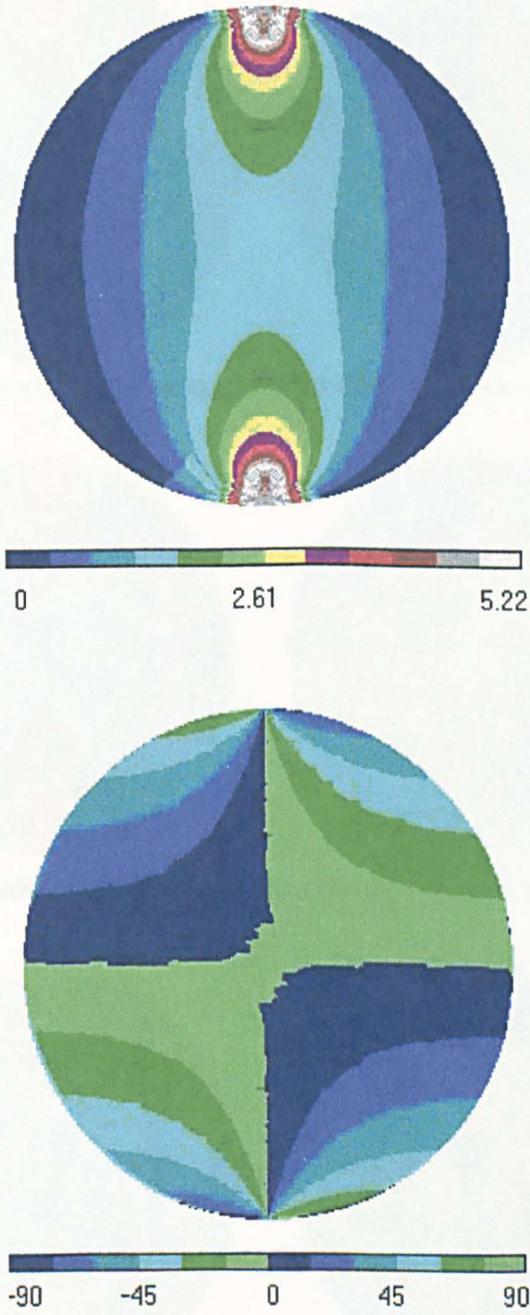


Fig. 4.13 The modelling results of the fringe order and isoclinic angle for a circular disc due to 5° linear misalignment

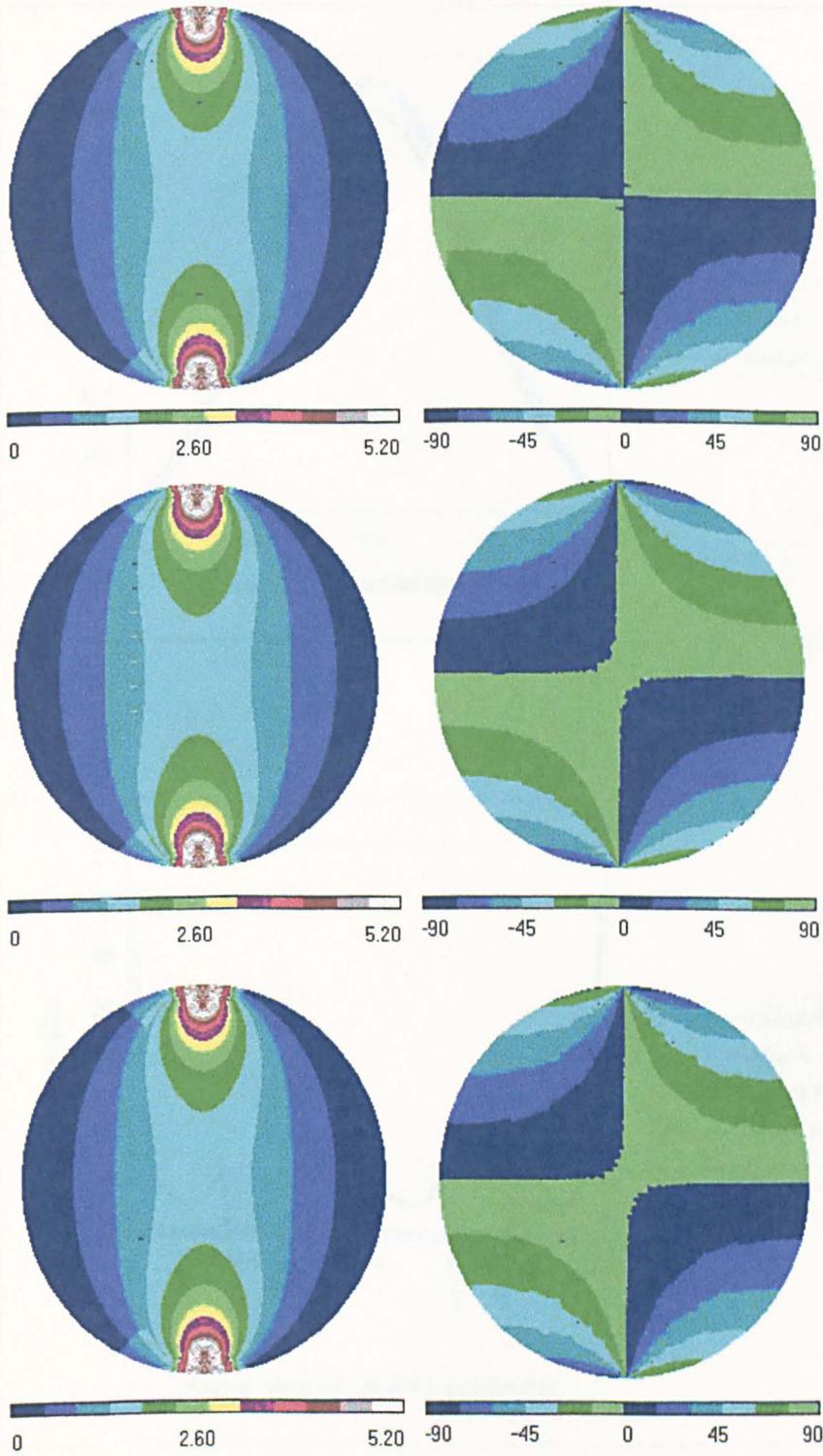


Fig. 4.14 The modelling results of the fringe order and isoclinic angle for a circular disc due to 0.1° angular misalignment a) output elements; b) input quarter-wave plate; c) both output and input elements

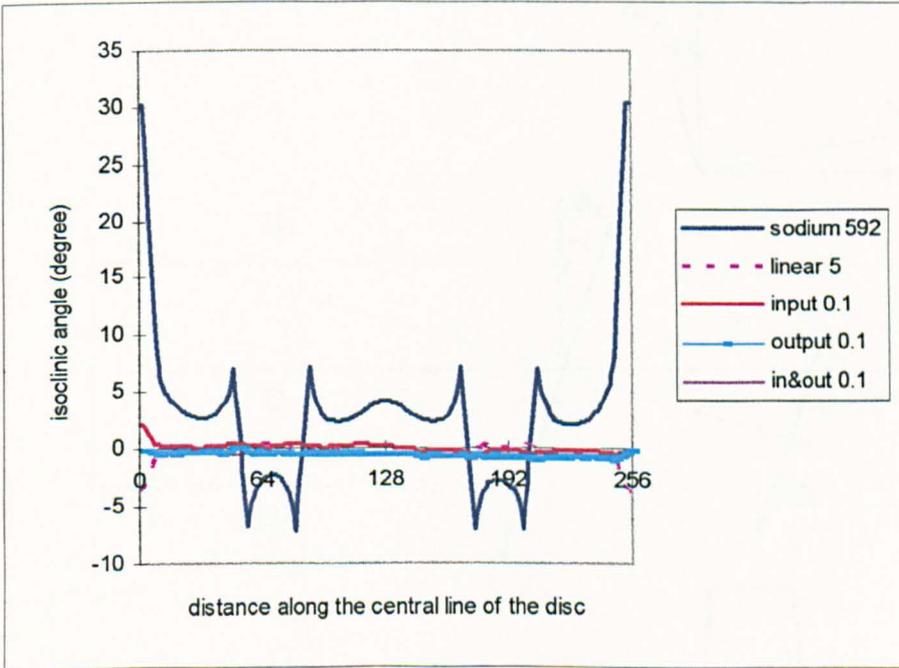
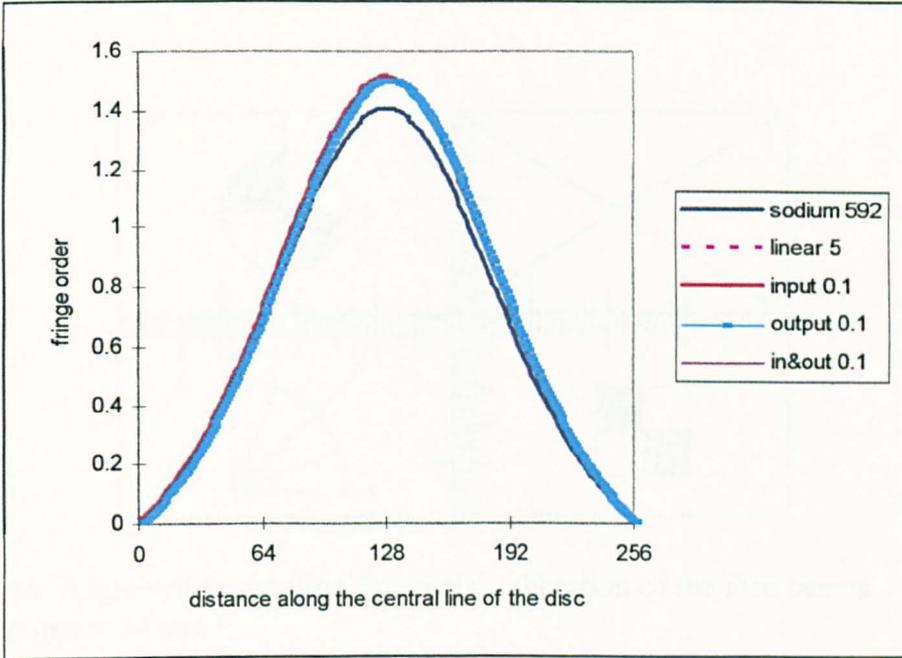


Fig. 4.15 The comparison of the modelling results on the fringe order and isoclinic angle along the central line of the circular disc for the above three cases

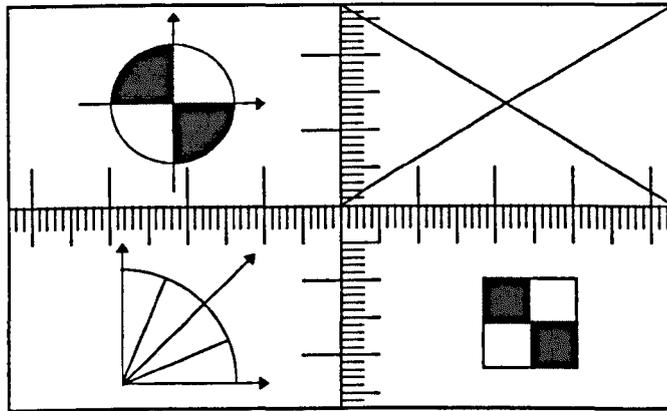


Fig. 4.16 Alignment target used for spatial calibration of the four beams
size: 86 mm × 54 mm

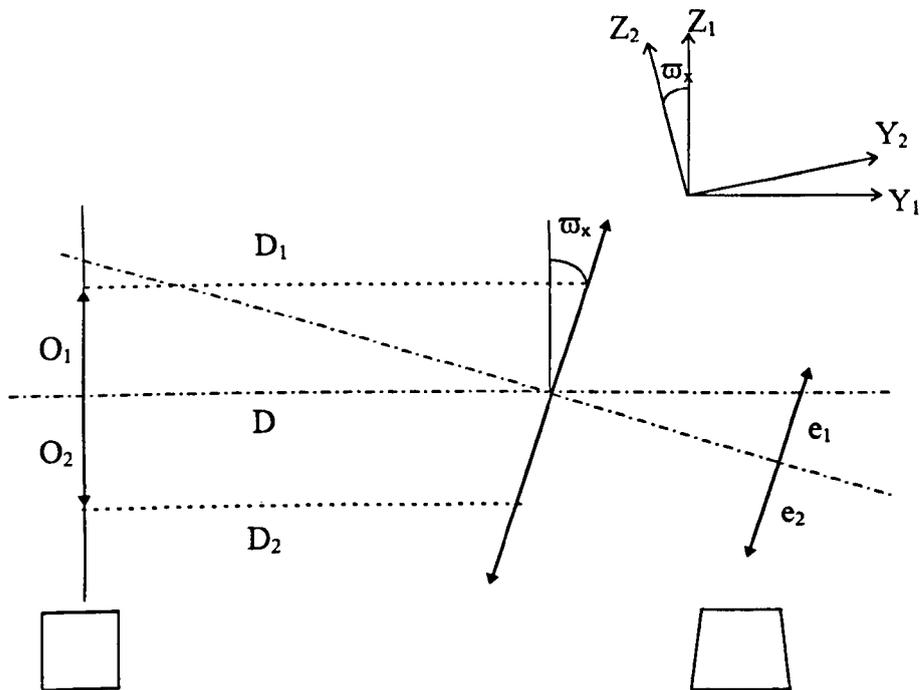


Fig. 4.17 A schematic representation of the effect of a rotation misalignment of a camera with respect to X_1 -axis

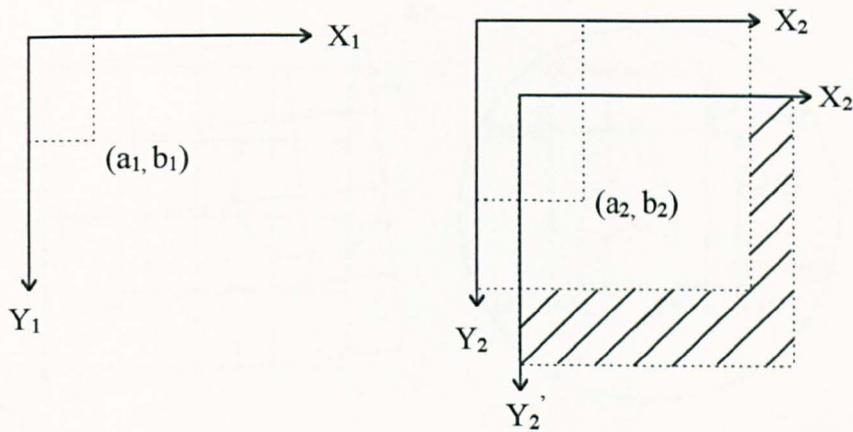


Fig. 4.18 Schematic representation of a translational alignment

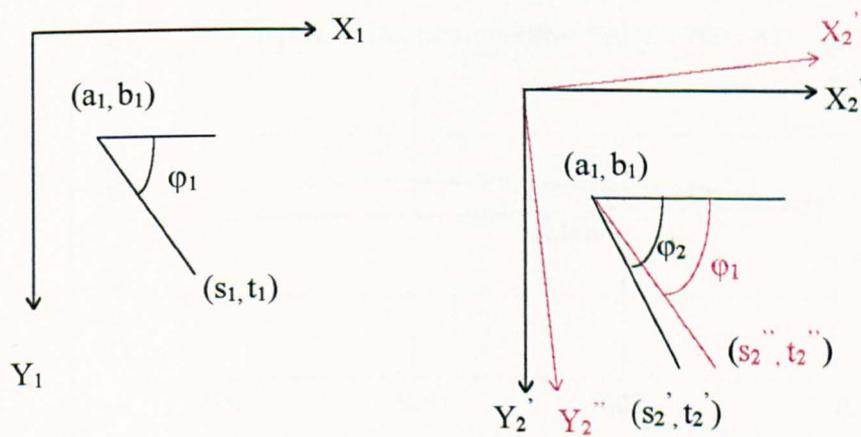


Fig. 4.19 Schematic representation of a rotational alignment

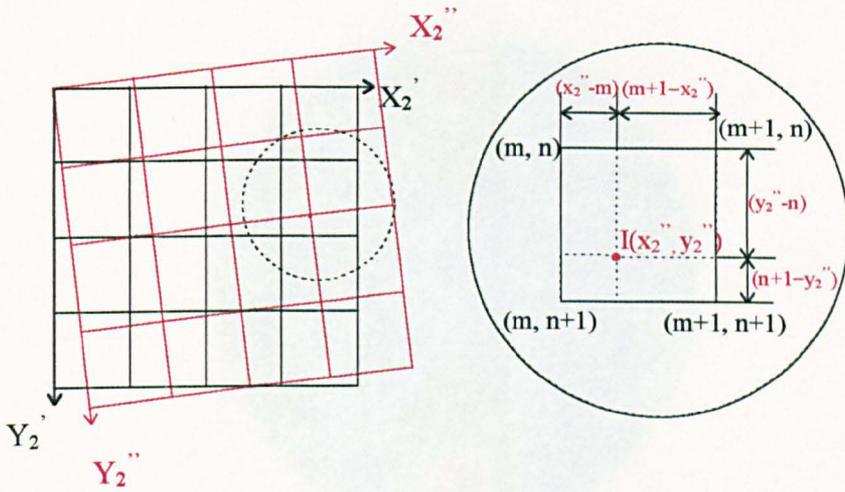


Fig. 4.20 Assigning the light intensity value to the new co-ordinate system $X_2''-Y_2''-Z_2''$

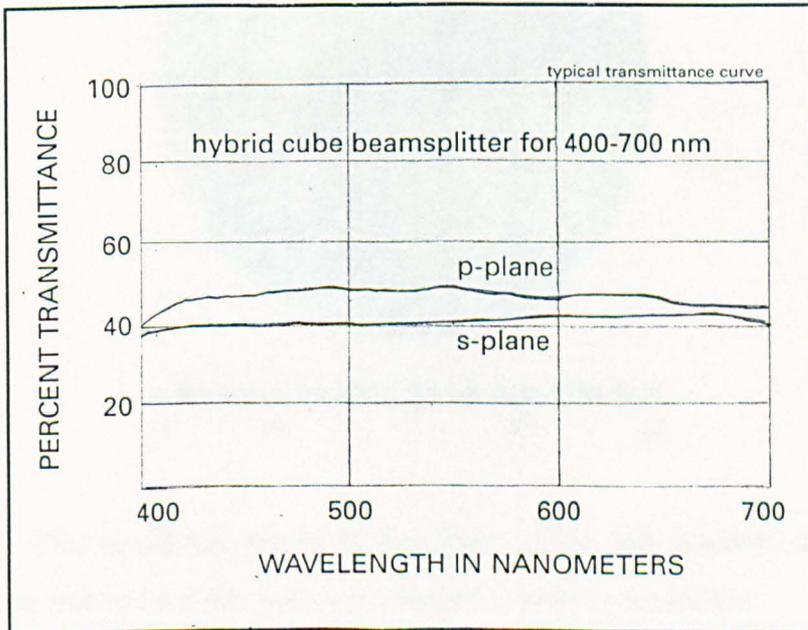


Fig. 4.21 Cube transmittance curve (reproduced from Melles Griot catalogue)

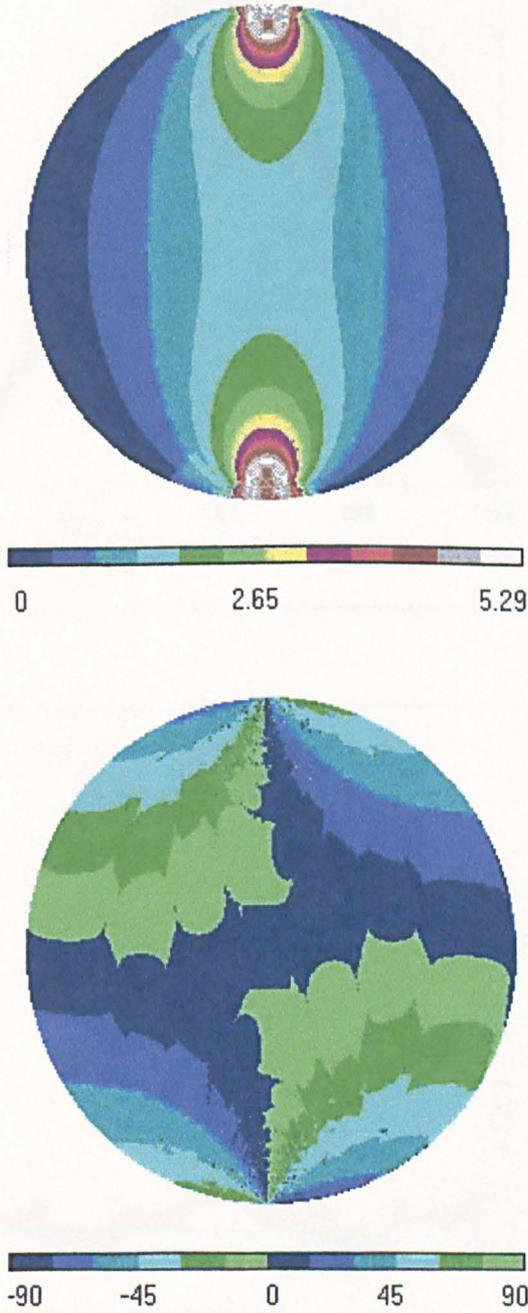


Fig. 4.22 The modelling results of the fringe order and isoclinic angle for a circular disc due to intensity tolerance allowed in cube beamsplitter.

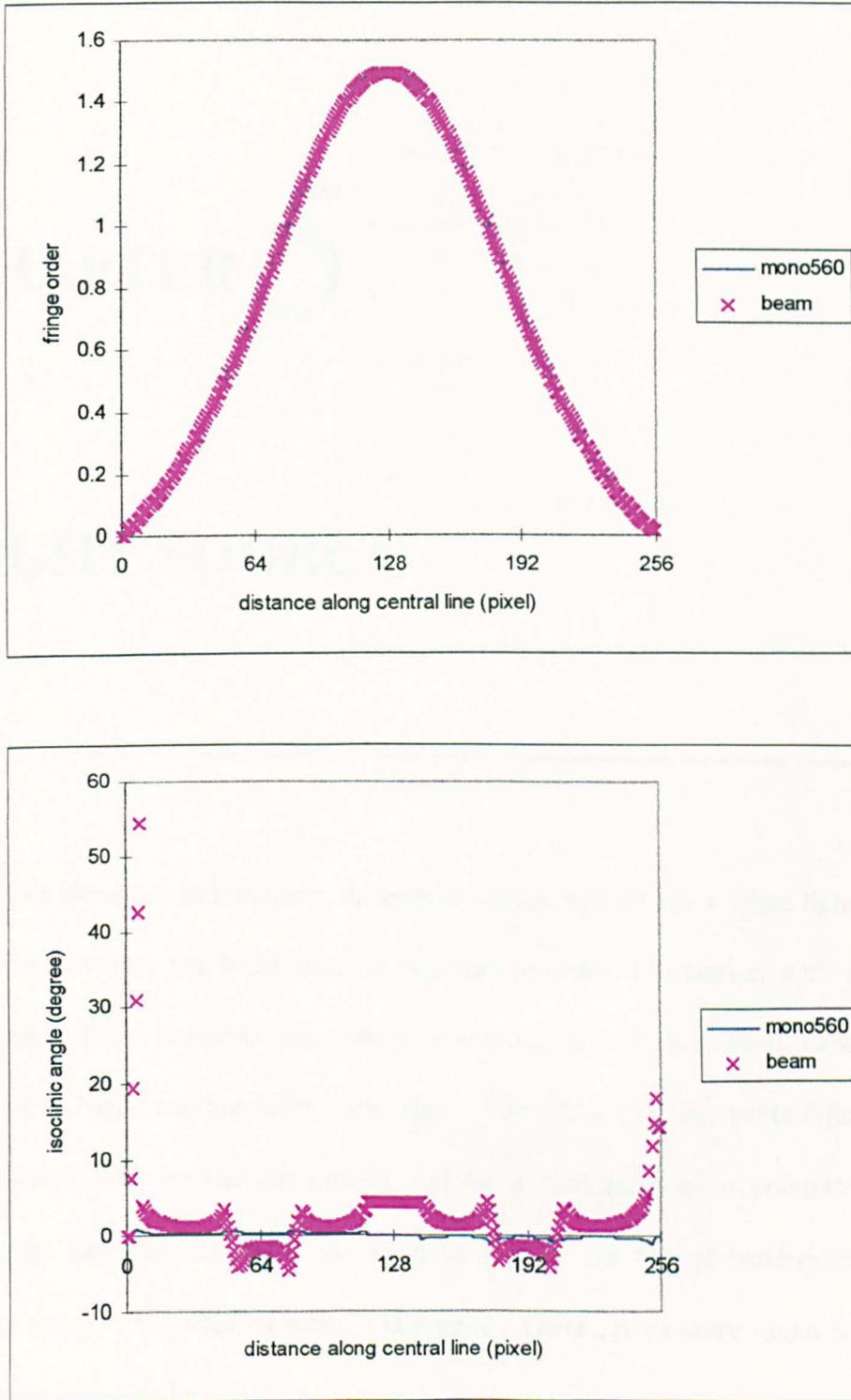


Fig. 4.23 The comparison of the modelling results on the fringe order and isoclinic angle along the central line of the disc due to intensity tolerance in four beams

CHAPTER 5

LIGHT SOURCE

It is often desirable in a modern automated polariscope to use a white light source in order to combine the RGB data or spectral contents information with the phase-stepping system. However, the theory underlying many such systems, namely phase-stepping is based on monochromatic light. The effect of using white light has been investigated both by the experiment and by a simulation of a polariscope. The simulation has been validated and used to explore the use of bandwidth filters to control the errors caused by using white light. These errors were found to be larger for higher fringe orders.

5.1 Introduction

As described before, the PSIOS system, which contains a four-way beamsplitter within the polariscope, is capable of collecting simultaneously the four phase-stepped images required to evaluate relative retardation and the isoclinic angle. The splitting of the beam leads to an inevitable reduction in the intensity of light being measured by the recording device. In addition, the PSIOS instrument was designed for use in an industrial environment and so portability is important and makes high intensity monochromatic sources impractical. The use of a white light source could overcome many difficulties associated with the low intensity at the recording devices. The use of white light would also offer an opportunity to calibrate the phase-stepped maps using spectral analysis [Redner 1984] or RGB method [Patterson *et. al.* 1997a]. Hence, the capability to use white light would be an advantage but the principle of phase stepping in photoelasticity is founded on the use of a monochromatic light source. The effect of using the white light and bandwidth filters have been investigated using a mathematical simulation package and experiments have been performed to verify the results of theoretical analysis.

Since the white light can be considered as the integral of the monochromatic light source over the whole wavelength range of visible light, the investigation of errors associated with the use of the white light were started by a consideration of non-matching monochromatic light. Followed by the white light error, the effect of the medium band filter is studied.

5.2. Errors in Non-matching Monochromatic Light Source

In Chapter 3, the light intensity emitted from the polariscope has been deduced for an arbitrary polariscope. In a polariscope with left circular light input and the first quarter-wave plate with its fast axis at $\pi/4$ to that of the polariser, i.e., $\rho = -\pi/4$, the light intensity is given by,

$$\begin{aligned}
 I = I_m + I_v \cdot [& -\cos 2(\beta - \phi) \cdot \sin 2(\phi - \theta) \quad \cdot \sin \alpha \cdot \sin \xi_1 \\
 & + \cos 2(\beta - \phi) \cdot \cos 2(\phi - \theta) \cdot \cos 2\theta \quad \cdot \cos \xi_1 \\
 & - \cos 2(\beta - \phi) \cdot \sin 2(\phi - \theta) \cdot \sin 2\theta \cdot \cos \alpha \cdot \cos \xi_1 \\
 & - \sin 2(\beta - \phi) \cdot \sin 2(\phi - \theta) \cdot \cos 2\theta \quad \cdot \cos \xi_1 \cdot \cos \xi_2 \\
 & - \sin 2(\beta - \phi) \cdot \cos 2(\phi - \theta) \cdot \sin 2\theta \cdot \cos \alpha \cdot \cos \xi_1 \cdot \cos \xi_2 \\
 & + \sin 2(\beta - \phi) \quad \cdot \sin 2\theta \cdot \sin \alpha \cdot \cos \xi_1 \cdot \sin \xi_2 \\
 & - \sin 2(\beta - \phi) \quad \cdot \sin 2\theta \cdot \sin \alpha \cdot \sin \xi_1 \cdot \cos \xi_2 \\
 & - \sin 2(\beta - \phi) \quad \cdot \cos \alpha \cdot \sin \xi_1 \cdot \sin \xi_2]
 \end{aligned} \tag{5-1}$$

In this equation, the orientation of the polariscope and the retardation of the quarter-wave plates are known and the light intensity can be measured, so there are four unknowns I_m , I_v , θ and α remaining. Therefore, a minimum of four images are required to solve the equations. The phase-stepping system is based on this equation and consists of changing the absolute phase of reference wave in equal steps and measuring the local light intensity after each step. The change in phase is, in practice, achieved by rotating of the output or input optical elements of the polariscope. To facilitate simultaneous recording of the phase-steps it is preferable to fix the input optical elements.

If an arbitrary wavelength of monochromatic light is used, rather than the matched wavelength of monochromatic light required by the quarter-wave plates, some errors will be introduced including an error due to the mismatch of the quarter-wave plates and due to dispersion of the birefringence.

5.2.1 Mismatch of the Quarter-Wave Plate

In Section 4.1.1, the effect of the quarter-wave plate due to the mismatch of wavelength was investigated. When the wavelength of monochromatic light is unmatched for the quarter-wave plates, the relative retardations caused by the quarter-wave plates are changed, i.e., $\xi_1 = \xi_2 \neq \pi/2$. Assuming that the matching wavelength for the quarter-wave plates is λ_0 then for an arbitrary wavelength, λ the error ε caused by the quarter-wave plates is given by,

$$\varepsilon = \frac{\pi}{2} \left(\frac{\lambda_0}{\lambda} - 1 \right) \quad (5-2)$$

The relative retardations caused by the quarter-wave plates are equal to,

$$\xi_1 = \xi_2 = \varepsilon + \frac{\pi}{2} = \frac{\pi}{2} \cdot \frac{\lambda_0}{\lambda} \quad (5-3)$$

5.2.2 Dispersion of Birefringence

Some materials exhibit a significant dispersion of birefringence with the wavelength of incident light, examples could be found in reference [Haake and Patterson, 1992]. If C_m is a material constant at the matching wavelength λ_m and the constant C_m becomes C_λ at an arbitrary wavelength λ , it follows from Eq. 3-1 that,

$$(\sigma_1 - \sigma_2) \cdot t = \frac{N \cdot \lambda}{C_\lambda} = \frac{N_m \cdot \lambda_m}{C_m} \quad (5-4)$$

so the fringe order observed in the wavelength λ is,

$$N = \frac{N_m \cdot \lambda_m}{\lambda} \cdot \frac{C_\lambda}{C_m} \quad (5-5)$$

where N_m is the fringe order at the matching wavelength λ_m , and the fringe order, N , is for the wavelength λ .

By considering both the mismatch of the quarter-wave plates and the dispersion effect, the light intensity emitted from the polariscope for a non-matching wavelength $(I)_\lambda$ can be obtained by substituting Eq. 5-3 and Eq. 5-5 into Eq. 5-1:

$$\begin{aligned}
 (I)_\lambda = I_m + I_v [& -\cos 2(\beta - \phi) \cdot \sin 2(\phi - \theta) \cdot \sin(2\pi \frac{N_m \lambda_m}{\lambda} \cdot \frac{C_\lambda}{C_m}) \cdot \sin(\frac{\pi}{2} \cdot \frac{\lambda_m}{\lambda}) \\
 & + \cos 2(\beta - \phi) \cdot \cos 2(\phi - \theta) \cdot \cos 2\theta \cdot \cos(\frac{\pi}{2} \cdot \frac{\lambda_m}{\lambda}) \\
 & - \cos 2(\beta - \phi) \cdot \sin 2(\phi - \theta) \cdot \sin 2\theta \cdot \cos(2\pi \frac{N_m \lambda_m}{\lambda} \cdot \frac{C_\lambda}{C_m}) \cdot \cos(\frac{\pi}{2} \cdot \frac{\lambda_m}{\lambda}) \\
 & - \sin 2(\beta - \phi) \cdot \sin 2(\phi - \theta) \cdot \cos 2\theta \cdot \cos^2(\frac{\pi}{2} \cdot \frac{\lambda_m}{\lambda}) \\
 & - \sin 2(\beta - \phi) \cdot \cos 2(\phi - \theta) \cdot \sin 2\theta \cdot \cos(2\pi \frac{N_m \lambda_m}{\lambda} \cdot \frac{C_\lambda}{C_m}) \cdot \cos^2(\frac{\pi}{2} \cdot \frac{\lambda_m}{\lambda}) \\
 & - \sin 2(\beta - \phi) \cdot \cos(2\pi \frac{N_m \lambda_m}{\lambda} \cdot \frac{C_\lambda}{C_m}) \cdot \sin^2(\frac{\pi}{2} \cdot \frac{\lambda_m}{\lambda})]
 \end{aligned}
 \tag{5-6}$$

where C_λ is the stress optic coefficient at the central wavelength of the narrow band filter. If the polariscope has the orientations shown in Table 3.1, then the six light intensities can be expressed as $(I_1)_\lambda$, $(I_2)_\lambda$, ... and $(I_6)_\lambda$, and the isoclinic angle θ_λ and isochromatic data α_λ can be expressed in an identical way as Eqs. 3-16 and 3-17 by substituting $(I_1)_\lambda$ to $(I_6)_\lambda$ for I_1 to I_6 ,

$$\theta_f = \frac{1}{2} \arctan \frac{(i_3)_\lambda - (i_5)_\lambda}{(i_6)_\lambda - (i_4)_\lambda} \tag{5-7}$$

$$\alpha_f = \arctan \frac{[(i_3)_\lambda - (i_5)_\lambda]}{[(i_2)_\lambda - (i_1)_\lambda] \cdot \sin 2\theta} = \arctan \frac{[(i_6)_\lambda - (i_4)_\lambda]}{[(i_2)_\lambda - (i_1)_\lambda] \cdot \cos 2\theta} \tag{5-8}$$

Based on the dispersion effect investigated by Haake and Patterson [1992], the fringe constant f is linearly related to the wavelength for stress frozen material. Its relationship for the photoelastic epoxy resin MY750 [Kenny, 1965] is expressed as

$$f = 0.93 \cdot \lambda - 117.81 \quad (5-9)$$

where the unit for the fringe constant is $\text{Nm}^{-1}/\text{fringe}$, and the wavelength is in nm. The relationship for the widely used epoxy resin CT200, CT1200 can also be found [Haake and Patterson 1992].

Since the matching wavelength for the quarter-wave plates used in the phase-stepping system is 560 nm, the dispersion effect for MY750 can be expressed as

$$\frac{C_\lambda}{C_m} = \frac{\lambda / f}{\lambda_m / f_m} = \frac{\lambda}{1.29\lambda - 163.16} \quad (5-10)$$

Substitution of this dispersion expression into Eqs. 5-7 and 5-8 allows the isoclinic and isochromatic parameters to be obtained at different wavelengths.

5.2.3 Theoretical Prediction for the Non-matching Monochromatic Light Case

In the previous discussion, the errors due to the mismatch of the quarter-wave plates and dispersion of birefringence were identified as two factors that contribute to the errors in the isoclinic and isochromatic parameters when an unmatched light source is used. A circular disc subjected to compressive stress was used as a model to simulate the isochromatic and isoclinic data produced at different wavelengths. Fig. 5.1 shows the simulated fringe order and isoclinic angle at three wavelengths, namely: 489.3; 560; and 689.3 nm. The comparison of the fringe order along the central line of the disc is shown in Fig. 5.2. The maximum fringe order along the central line of the disc,

which is 1.5 in the matching wavelength, became 1.79 at 489.3 nm and 1.18 at 689.3 nm. Although the large difference was existed in the fringe order, the effect to the fringe order is not significant, because a correction can be applied to obtain the fringe order at the matching wavelength by substituting Eqs. 5-9 and 5-10 to Eq. 5-5 as,

$$N_m = N \cdot \frac{\lambda}{\lambda_m} \cdot \frac{C_m}{C_\lambda} = N \cdot \frac{f}{f_m} = N \cdot \frac{0.93 \cdot \lambda - 117.81}{0.93 \cdot 560 - 117.81} \quad (5-11)$$

and then the errors in the fringe order are (0.02) and (-0.01) in these two cases. So, as long as the fringe constant f and the matching wavelength of the quarter-wave plates required are known, then the fringe order at the matching wavelength can always be obtained from the monochromatic light being used. However, the effect of the non-matching wavelength on the isoclinic angle was not recoverable and the influence to the area, where the integer and half fringe order was located, was 45 degrees, and around 10 to 15 degrees for any other regions.

5.2.4 Comparison of the Experimental Data and Theoretical Predication

To verify the results of the theoretical model, a circular disc subjected to compressive stress was tested in polariscopes using different light sources. The disc was made from epoxy resin MY750 using the procedures described by Kenny [1965]. The disc was machined from a solid block of resin and had a diameter 76 mm and thickness 6 mm. A point load of 33.56 N was applied at room temperature across a diameter. To achieve a variety of monochromatic sources, a white light source was used with a set of narrow band filters to limit the bandwidth. The wavelengths of the filters are shown in Table 5.1.

Table 5.1 The wavelengths of narrow band filters

Filter	1	2	3	4	5	6
Central wavelength (nm)	449.3	489.5	569.5	649.5	689.5	724.3
Half Peak Bandwidth (nm)	10.1	9.0	10.0	11.0	10.0	10.5

The bandwidth of the filters is less than 12 nm, and the influence of this range to the fringe order can be obtained by using Eq. 5-11. For example, for the filter with a central wavelength of 569.5 nm, the 10 nm range could cause a difference of 0.011 N_m of fringe order. Hence, it is reasonable to assume that the light emitted from the filters is monochromatic with a wavelength equal to that of the central wavelength of the filter. The light intensity was measured for the six cases listed in Table 5.1 and the isoclinic and isochromatic parameters were obtained using Eqs. 5-7 and 5-8. For convenience, the data is shown along the centre line of the disc where the direction of the principal stresses are 0° and 90° and hence the isoclinic data should be zero or ninety degrees. Figs. 5.3 and 5.4 show the experimental and theoretical data for the isoclinic angle and the isochromatic fringe order, respectively, where the experimental and theoretical data are represented by broken and solid lines respectively.

For the matching wavelength, 560 nm in this case, the isoclinic angle is obtained from Eqs. 3-7 and 3-8. Obviously, at the integer and half fringe points where $I_3 = I_4 = I_5 = I_6$, the isoclinic angle is undefined because both the numerator and the denominator are equal to zero. Since the undefined zone is a thin band compared with the whole field, the undefined points can be extrapolated from the adjacent areas [Kihara 1990, Wang and Patterson 1995]. But for the non-matching

wavelengths, due to the errors from the quarter-wave plates and dispersion effects, the undefined zones are very large in the vicinity of the integer and half fringe points. As seen in Fig. 5.3., both the experimental and theoretical isoclinic data have a very thin band undefined zones when $\lambda = 569.5$ nm, which is the nearest to the matching wavelength, while the undefined zones show a clear tendency to increase when moving away from the matching wavelength. When these undefined points are ignored, the amplitude of the plot for the isoclinic data are smaller for filters of approximately 560 nm than for those both with much longer and shorter wavelengths; that is the errors increase with the absolute value of the wavelength difference between the matched wavelength and the wavelength of the monochromatic source. The experimental results and theoretical predications do not correlate perfectly, however there is reasonable agreement and both approaches support the conclusion that errors in the isoclinic angle and the size of the undefined zones increase with deviation from the matching wavelength of the quarter-wave plates.

Comparing the fringe order at different wavelengths in Fig. 5.4, the isochromatic fringe orders are greater for shorter wavelengths and lower at longer wavelengths in both the experimental and theoretical results. Discontinuities are found at integer and half fringe values. It can be seen that these are related to the undefined isoclinic angles since they occur at the same points in data. Examination of Eqs. 5-7, 5-8 reveals that the undefined isoclinic zone will give discontinuities in the isochromatic data. This inter-relationship has been described in detail by Wang and Patterson [1995]. Ignoring the edge effects occurring in Fig. 5.4, the experimental results agree well with the theoretical data although the experimental data is slightly higher than predicated by the theory for the 489.5 nm filter.

5.2.5 Conclusion

In conclusion, two kinds of error, i.e., mismatch of the quarter-wave plates and the dispersion error, make significant contributions to the isoclinic and isochromatic data in the case of unmatched monochromatic light; and can be simulated using the model presented here.

5.3 Errors Associated with the Use of the White Light Source

5.3.1 Theoretical Analysis of Errors in White Light Polariscopes

When white light is employed the error due to the mismatch of the quarter-wave plates and the dispersion, discussed above, are introduced. Besides these, the intensity of the light source I_v varies with wavelength. When the light passes through the polariscopes, the transmission ratios of every element are wavelength dependent. In addition the recording device, here a monochromatic camera, has a response that is wavelength dependent. Since the elements of the polariscopes are connected in series, the recorded intensity is then the product of the response of all the elements [Levi, 1968]. If enough continuous narrow band sources were to be used in the wavelength interval of the white light source, then the total light intensity should be the sum of the individual light intensities from the sources. Extending to a more general case, the light intensity should be the integration of the light intensity from every monochromatic state. Assuming that the wavelength interval of the white light is λ_i

to λ_2 , then the intensity at the output for the polariscope with a white light source, $(I)_{W}$ is,

$$(I)_{W} = \int_{\lambda_1}^{\lambda_2} I_v(\lambda) \cdot T(\lambda) \cdot F(\lambda) \cdot (I)_{\lambda} d\lambda \quad (5-12)$$

where $I_v(\lambda)$ is the intensity of the light source,

$T(\lambda)$ is the product of the transmission ratio of every element in a polariscope,

$F(\lambda)$ is the response curve of the monochromatic CCD camera,

$(I)_{\lambda}$ is the light intensity in every monochromatic case.

If the forms of $I_v(\lambda)$, $T(\lambda)$, $F(\lambda)$ and $(I)_{\lambda}$ are known for all wavelengths, then the calculation method of the isoclinic and isochromatic parameters would be the same as for the monochromatic light.

The spectrum $I_v(\lambda)$ of the four white light sources currently used in author's laboratory have been measured using a spectral contents analysis system consisting of a Monolight 6000 and 6800 (Monolight Instruments Ltd, England) and are shown in Fig. 5.5. The sources have been labelled S1, S2, O1 and L1. Sources S1 and S2 are nominally identical and were manufactured by Sharples Stress Engineers Ltd (Preston, England) for 15' diameter diffuse light polariscope. These sources use fluorescent tubes. Source O1 is an Oriel source, Model No. 66181 (USA) using a 250W QTH (Quartz Tunsten Halogen) bulb. Source L1 was designed and manufactured for the PSIOS system and includes a 300W projector lamp (type: ELH) with a cold mirror to reduce infrared emission.

The transmission ratios $T(\lambda)$ from the elements in the polariscope are wavelength dependent. Since the spectra of white light sources were measured after

they passed the optical elements, the spectra are actual the combined effect of the transmission ratios of the polariscope and the light sources.

The response curve $F(\lambda)$, of the monochromatic CCD camera (Pulnix, TM526) was obtained from the manufacturers and is shown in Fig. 5.6.

The combined spectra of the white light sources and the camera in Fig. 5.7, can be obtained by multiplication and expressed as,

$$S(\lambda) = I_v(\lambda) \cdot T(\lambda) \cdot F(\lambda) \quad (5-13)$$

In practice, the combined response curves were obtained graphically and then described by fitting sixth order polynomials to them using a least square approach. So for the sources S1 and S2 we have,

$$S_s(\lambda) = -0.0185X^6 + 0.6477X^5 - 9.2927X^4 + 67.91X^3 - 288.12X^2 + 622.33X - 549.4 \quad (5-14)$$

For the source O1,

$$S_{o1}(\lambda) = 0.1167X^5 - 3.0334X^4 + 31.008X^3 - 156.03X^2 + 387.49X - 380.93 \quad (5-15)$$

and for the source L1,

$$S_{L1}(\lambda) = 0.001X^6 - 0.0198X^5 + 0.1702X^4 - 1.2709X^3 + 8.8013X^2 - 33.909X + 48.77 \quad (5-16)$$

where in last three equations, $X = \lambda / 100$ nm.

The coefficients of the polynomials are also shown in Table 5.2. The intensity of the white light over the wavelength interval can be obtained by substituting these coefficients into Eq. 5-12. The orientations of the six images were chosen to be the same as before, and the six light intensities can be expressed as $(I_1)_w$, $(I_2)_w$, ... and

$(I_\theta)_w$, and then the isoclinic angle θ_w and the isochromatic retardation α_w can be deduced in a similar way as in Eqs. 5-7 and 5-8.

5.3.2 Results of Theoretical Analysis

The theoretical calculations were performed using a mathematical software package (MathCAD Plus 6.0 Professional Edition, Adept Scientific Plc, USA) by substituting Eqs. 5-13, 5-14 and 5-15 to Eq. 5-12 to obtain the isochromatic and isoclinic data in three white sources. The results are shown in Fig. 5.8 for the case of low fringe orders while results for high fringe orders are shown in Fig. 5.9. The photoelastic data for a wavelength of 560 nm were used as the standard for all the cases. The fringe order and the isoclinic angle in a specimen were taken as the variables on the horizontal and vertical axes, respectively, and the errors for the isoclinic and isochromatic parameters were plotted as contours in Fig. 5.8 for the three different white light sources. Owing to the periodic properties of the isoclinic angle and the relative retardation, only the ranges 0 to 45 degrees for the isoclinic angle and 0 to 0.5 for the fringe order, have been illustrated. Errors at higher fringe orders can be simulated along the centre line of a compressive disc, where the isoclinic angle should be zero. Fig. 5.9 shows the predicted errors for the case of a maximum fringe order of 5 observed in the three different white light sources. In this figure errors in the isochromatic fringe order and the isoclinic angle are much larger at higher fringe orders than at lower ones. When the fringe order is 5, the errors in the fringe order were 1.73, 0.60 and 0.82 for sources S1, O1 and L1, respectively. So even the best source, O1 has an error larger than 10%. Hence white light sources must be limited

to low fringe orders if reliable results are to be acquired. If 5% errors for fringe order are permitted, then the maximum fringe order that could be measured would be 0.5, 3, and 2 using S1, O1 and L1 sources respectively, with a 560 nm quarter-wave plate. Discontinuities are also found at the location of half and integer fringe orders, as observed in the data for the non-matching monochromatic light. Once again, the undefined isoclinic angle is identified to be the reason of this phenomenon. The increase in the size of the undefined zones and the amplitude of error for the isoclinic angle are quite obvious as the fringe order increases. The results from the theoretical modelling for the circular disc with the light source S1, which has the largest errors in above analysis, are shown in Fig. 5.10.

Therefore the light sources can be ranked in terms of increasing error as follows: O1, L1, S1(S2). This ranking is probably related to the relative bandwidth and the offset of the central wavelength from the matched wavelength for the quarter-wave plates. In these respects, it can be seen from Fig. 5.5 that the sources S1 and S2 are the worst.

5.3.3 Experimental Results

Experimental tests have been carried out for the light source which showed the largest errors in the theoretical analysis, i.e., S1. The same circular disc was used again. The disc was viewed in a 15' diameter diffuse light polariscope in which was installed the light source S1. The installation also contained an alternative sodium light source. The alternative source was monochromatic with a central wavelength of 592.25 nm (see Section 4.1.1.2). The fringe order and the isoclinic angle along the diameter of

the disc perpendicular to the load for the sodium and the white light source S1, are shown in Fig. 5.11. From this figure, it can be seen that the fringe order in sodium and white light exhibit a large difference at high fringe orders, e.g., $N_{S1} / N_{Na} = 0.75$ when $N_{Na} = 2$. Discontinuities are found at the location of the half and integer fringe orders, as predicted by the theoretical model.

The corresponding results for the isoclinic angle are shown in Fig. 5.11. The undefined zones in the isoclinic angle in the white light S1 are larger than in the sodium light. The isoclinic angle should be zero along the centre line of the disc and the amplitude of errors in the white light are greater than in the sodium light. In addition more scatter occurred in the data collected using the white light. This supports the conclusions drawn from the theoretical predictions.

5.3.4 Discussion

The results from the theoretical model and the experiments show close correlation for both the isochromatic and isoclinic data. Therefore the most significant error source in the white light polariscope can be identified as the mismatch of the quarter-wave plates and dispersion errors. These two kinds of error can not be removed when the white light source is used. However, there are two methods to reduce the white light errors, and more details are discussed in the next section.

1. Selection of the most appropriate light source. The light source should have a high intensity close to the matching wavelength and low intensity at other wavelengths. In the sources tested here source O1 meets this criterion and the results obtained using it have the smallest errors as seen in Figs. 5.8 and 5.9.

2. Choosing the appropriate filters to limit the wavelength bandwidth of the light source, assuming that the spectrum is approximately flat within the bandwidth of the filter.

The spectra for both should be also symmetrical about the matching wavelength.

5.4 Medium-Band Filters in White Light Polariscopes

Again the worst source, S1 has been used to illustrate the effect of a medium-band filter.

Commercially available medium-band filters, with a central wavelength of 550 nm and FWHM bandwidth (the Full-Width, Half Maximum transmission bandwidth) of 40 nm and 80 nm were used (35-5065, 35-5073, Ealing Electro-optics., Inc.). The spectra of the filters were provided by manufacturers and are shown in Fig. 5.12. The combined response of the light source, monochromatic camera and each of the filters are given in Fig. 5.13. The expressions to describe the spectrum of the filters were obtained by the least square method as before. So for the 80 nm filter,

$$S_{80}(X) = 2.653X^4 - 0.0999X^3 - 1.928X^2 + 0.084X + 0.4191 \quad (5-17)$$

and for the 40 nm filter,

$$S_{40}(X) = 1003X^5 - 117.36X^4 - 86.234X^3 - 0.9793X^2 + 1.4119X + 0.5724 \quad (5-18)$$

where $X = \lambda / 100$ nm in both equations.

For comparison, the coefficients for all the light sources and filters are given in Table 5.2.

Table 5.2. Coefficients for polynomials in $X (= \lambda / 100 \text{ nm})$ describing combined response curve.

Light Source	Coefficients						
	6th	5th	4th	3rd	2nd	1st	constant
S1, S2	-0.0185	0.6477	-9.2927	67.91	-288.12	622.33	-549.4
O1	0	0.1167	-3.0334	31.008	-156.03	387.49	-380.93
L1	0.001	-0.0198	0.1702	-1.2709	8.8013	-33.909	48.77
S(80)	0	0	2.653	-0.0999	-1.928	0.084	0.4191
S(40)	0	1003	-117.36	-86.234	-0.9793	1.4119	0.5724

By substituting these coefficients into Eq. 5-12, the isoclinic angle and the isochromatic fringe order were simulated. Fig. 5.14 show the errors in the isoclinic data and the isochromatic fringe order when the white light source S1 is combined with the 80 and 40 nm filters. Significant improvements can be clearly seen when comparing the results with Fig. 5.8. The errors for high fringe orders are simulated in Fig. 5.15, which corresponds to Fig. 5.9. The effects of medium band filters are obvious. When the fringe order is 5, the error in the fringe order is 1.73 fringes for the light source S1 without filter and it is reduced to 0.063 and 0.032 respectively with 80 and 40 nm filters. So, if an accuracy of $\pm 5\%$ fringe order is required, then the maximum fringe order that could be measured would be 0.5 fringe for source S1, but 6 and 10 fringes respectively with 80 and 40 nm filters. For the isoclinic angle, both the undefined zones in half and integer fringes and the amplitude on the other positions decrease rapidly with the use of the filters. If the same accuracy is required for the isoclinic angle when the undefined zones are ignored, the 40 nm filter can be used to measure up to 10 fringes while up to 5 fringes for 80 nm filter.

It is interesting to note in Fig. 5.15 that the errors in the fringe order give positive values first and then decrease to negative values when the fringe order in a specimen increases. The same tendency is found if much higher fringe orders are investigated. This is almost certainly due to the difference between the central wavelength of the bandwidth and the matched wavelength of the quarter-wave plate being positive for the white light (600-560 nm) and negative for the medium bandwidth filter (560-550 nm). This trend can be observed in Fig. 5.4.

The theoretical modelling of a circular disc for the 80 and 40 nm filters are shown in Fig. 5.16 and the comparison for with two filters and without filter are made in Fig. 5.17 along the central line of the disc. Significant improvement through the use of the medium-band filters can also be observed by comparing the modelling disc shown in Fig. 5.16 with the one shown in Fig. 5.10.

Referring back to **Section 4.1.1**, the disc modelling for the sodium light source with unmatched quarter-wave plates are shown in Figs. 4.12, 4.15. By comparing with Figs. 5.16, 5.17, it can be concluded that the combined use of the medium band filter and the white light source has less errors than the combination with the sodium light source and unmatched quarter-wave plate, assuming that the medium band filter has the matching wavelength for the quarter-wave plates.

5.5 Conclusions

The use of white light could overcome difficulties associated with low intensity at the recording devices in the PSIOS system, but it introduces some errors to the fringe analysis as a consequence. A mathematical model of the phase-stepping technique has

been developed and used to analyse the errors generated by the use of various bandwidth spectra in the polariscope light sources. There is good correlation between results from the model and experiments. The errors due to mismatch of quarter-wave plates and dispersion are thought to make significant contributions to the errors in the isoclinic angle and the isochromatic fringe order when a white light source is employed.

The form of the spectrum of white light directly influences the errors in the isoclinic angle and isochromatic fringe order. The ideal have high density in the area adjacent to the matching wavelength of the quarter-wave plates and low density elsewhere and also the spectrum should be symmetrical about the matching wavelength.

The errors due to the use of white light have a much greater influence on the measurement of high isochromatic fringe orders than low orders. So white light sources are limited to low fringe orders if reliable results are to be acquired. For the three white light sources, discussed, if $\pm 5\%$ accuracy is required for the isochromatic fringe order, then the maximum fringe order that could be measured was 0.5, 3, and 2 respectively for the sources supplied by Sharples Stress Engineers Ltd (source S1), Oriel Inc. (source O1) and a source made in the author's laboratory (source L1), when used with a 560 nm quarter-wave plate.

The use of medium-band filters greatly decreases the errors generated by the white light. If the accuracy of $\pm 5\%$ in the fringe order is required, then the maximum fringe order that can be measured for using the worst light source was increase to 10 and 6 fringe orders for 40 and 80 nm bandwidth filters, respectively.



IMAGING SERVICES NORTH

Boston Spa, Wetherby

West Yorkshire, LS23 7BQ

www.bl.uk

**PAGE MISSING IN
ORIGINAL**

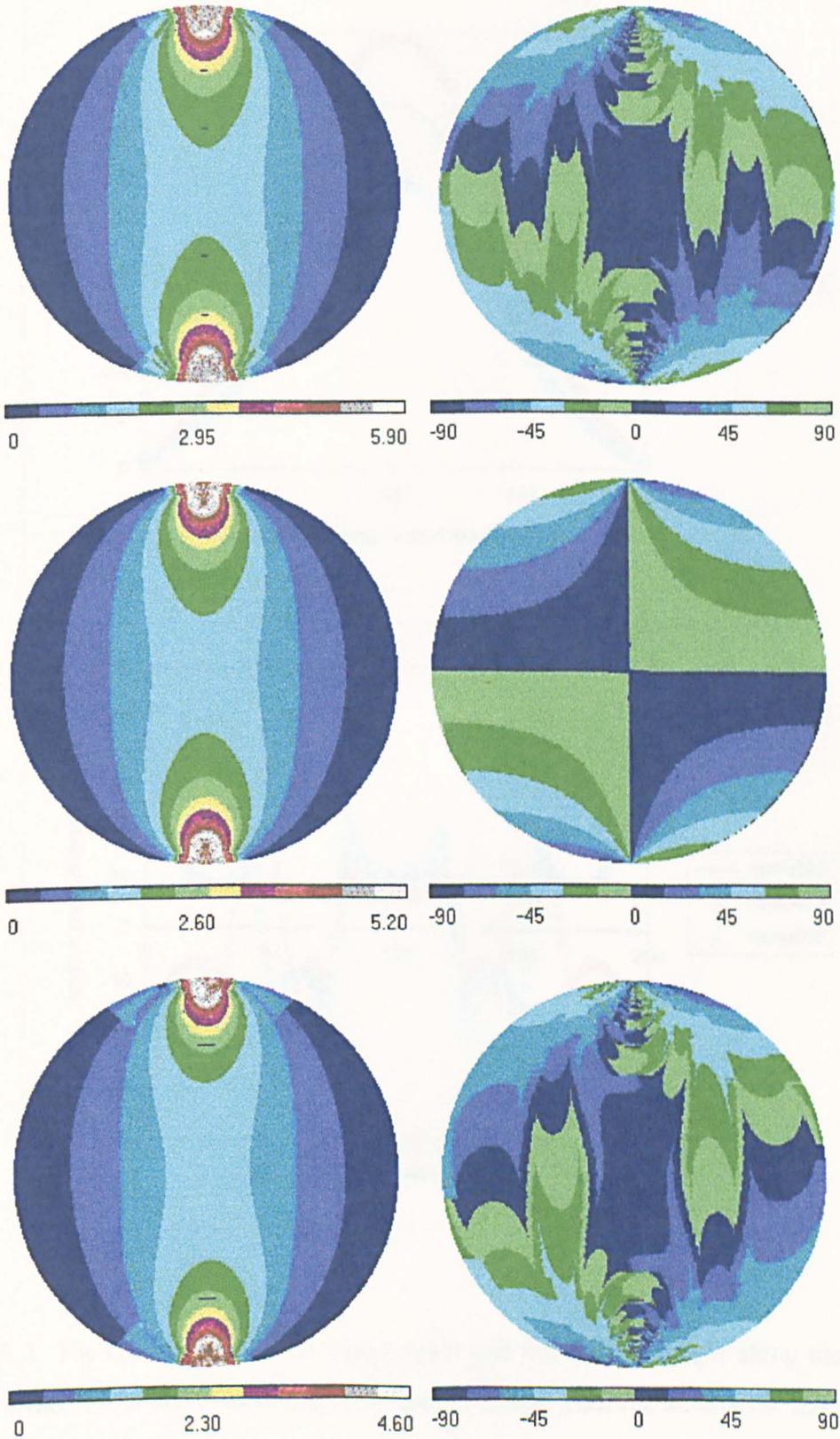


Fig. 5.1 The modelling isochromatic fringe order and isoclinic angle for the circular disc at the three wavelengths of 489.3, 560 and 689.3 nm

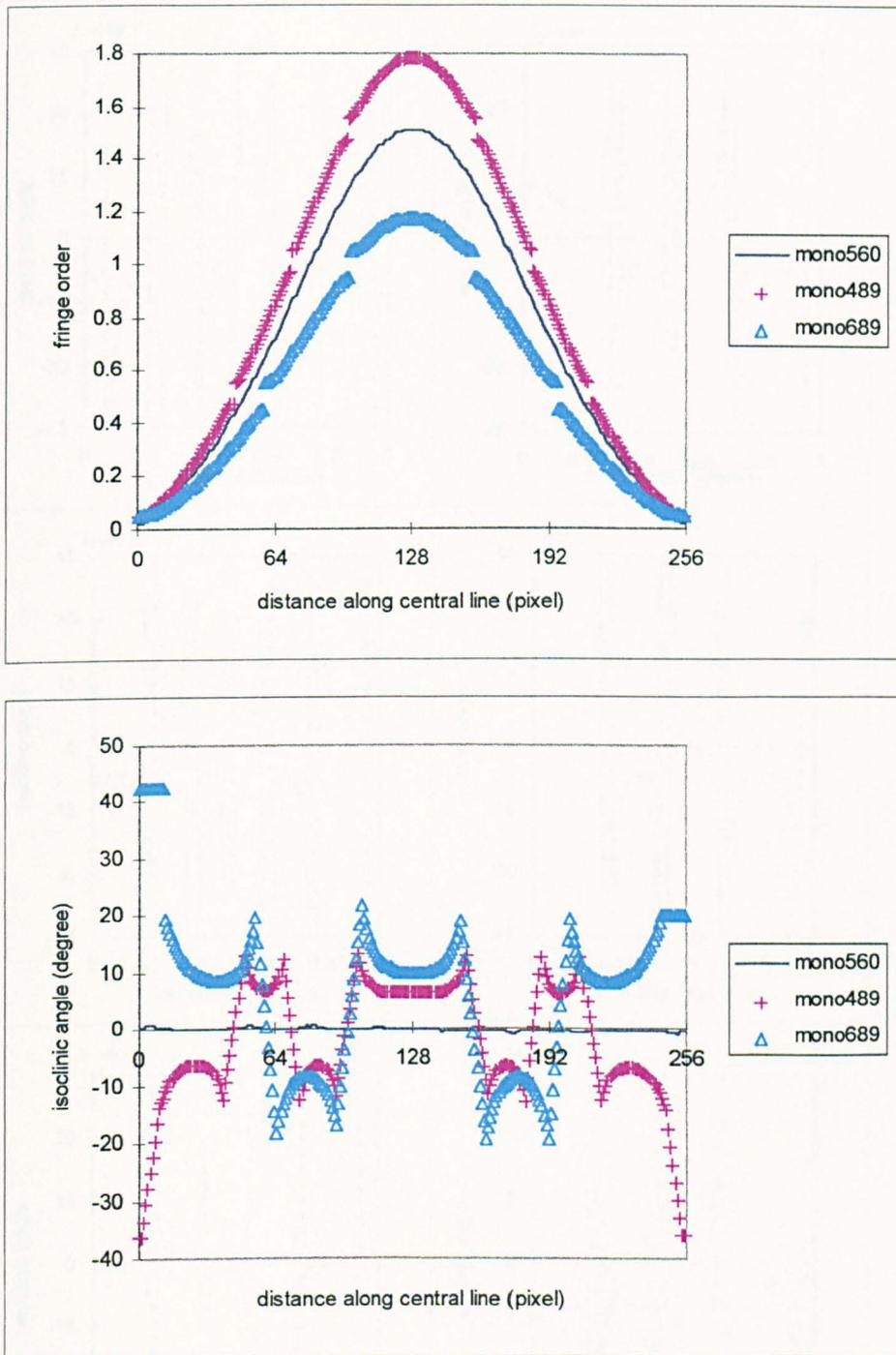


Fig. 5.2 The comparison of the fringe order and the isoclinic angle along the central line of the disc at three wavelengths. Data obtained from the simulation discussed in Section 5.2.3

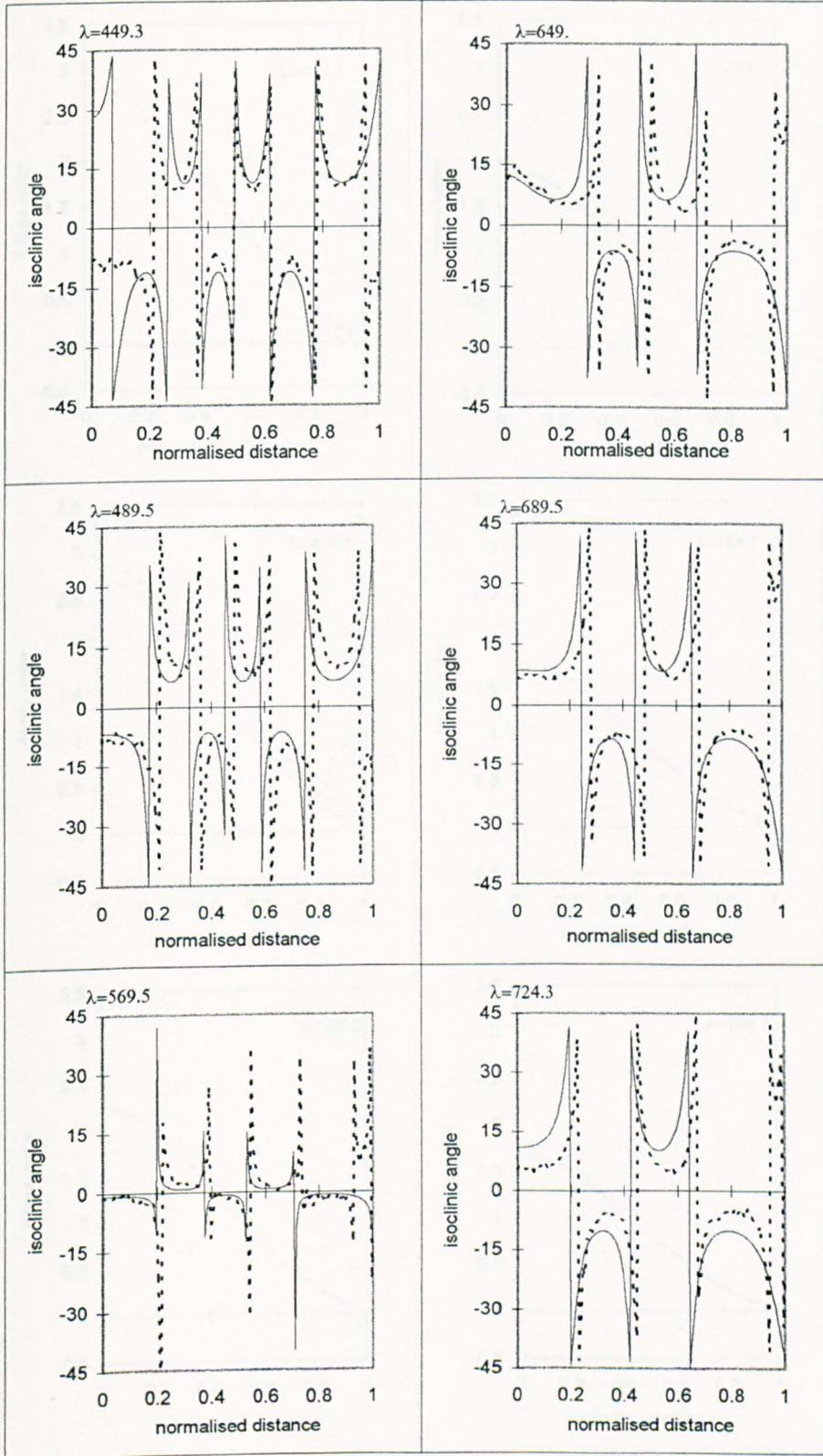


Fig. 5.3 The comparisons of the theoretical prediction and experimental results for isoclinic angle measured using six narrow band sources. The theoretical data and experimental are represented by solid and broken lines, respectively.

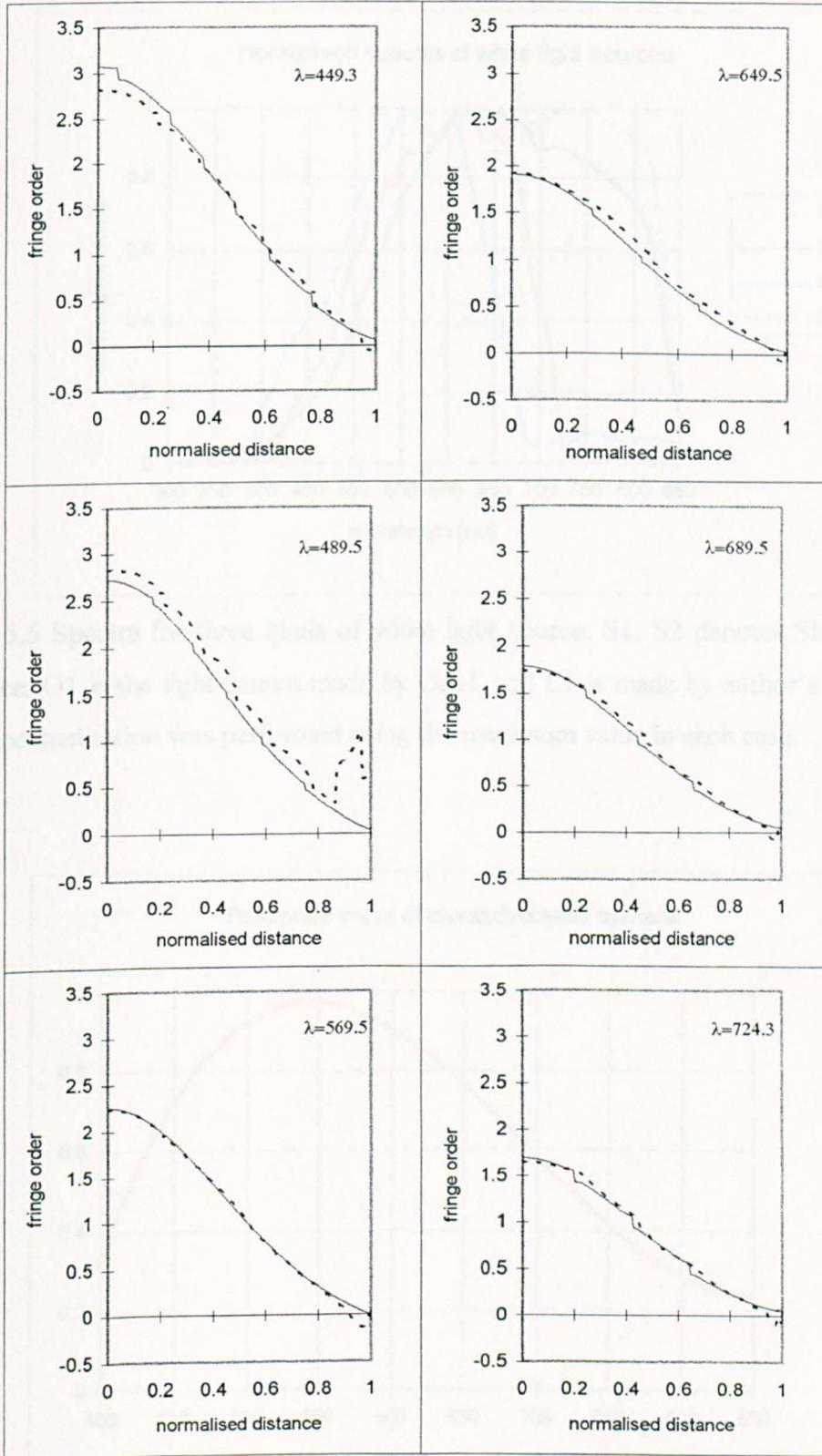


Fig. 5.4 The comparisons of the theoretical prediction and experimental results for isochromatic fringe order measured using six narrow band sources. The theoretical data and experimental are represented by solid and broken lines, respectively.

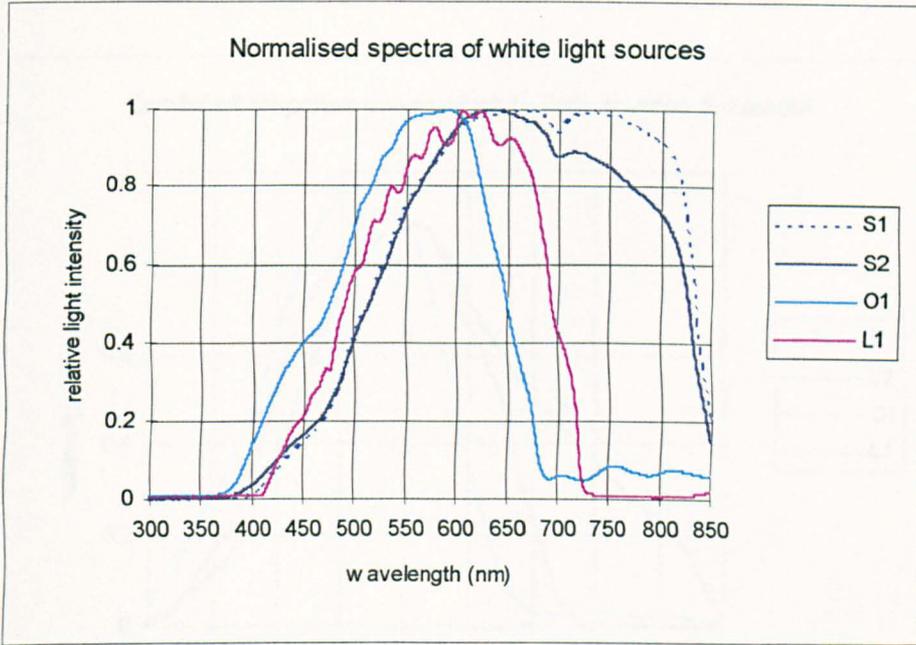


Fig. 5.5 Spectra for three kinds of white light source. S1, S2 denotes Sharples light source, O1 is the light source made by Oriel, and L1 is made by author's laboratory. The normalisation was performed using the maximum value in each case.

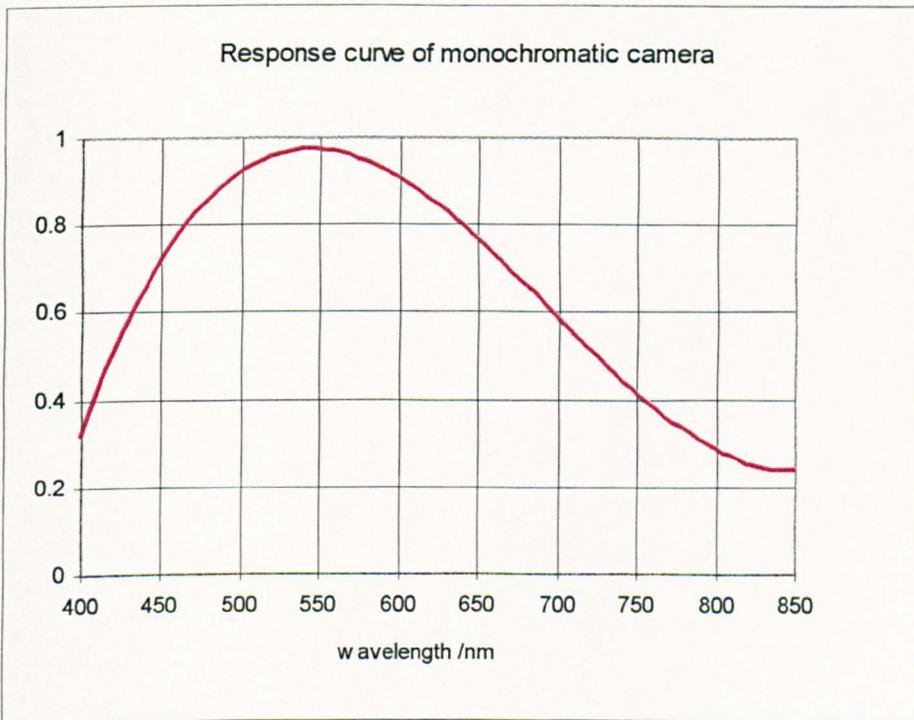


Fig. 5.6 Spectra for the monochromatic camera provided by the manufacturers.

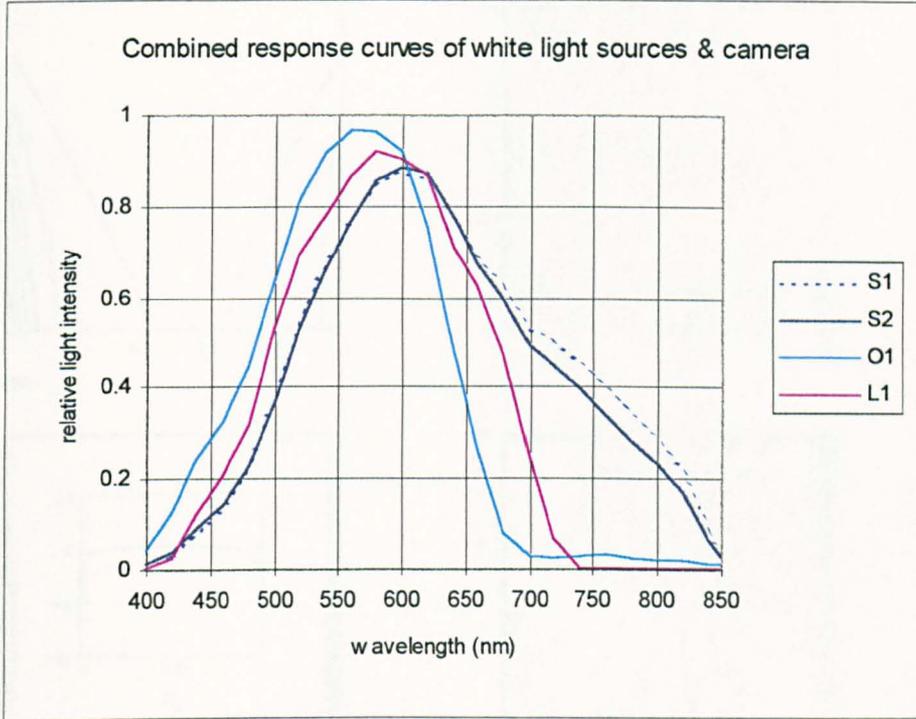


Fig. 5.7 Combined responses of the white light sources and the monochromatic camera

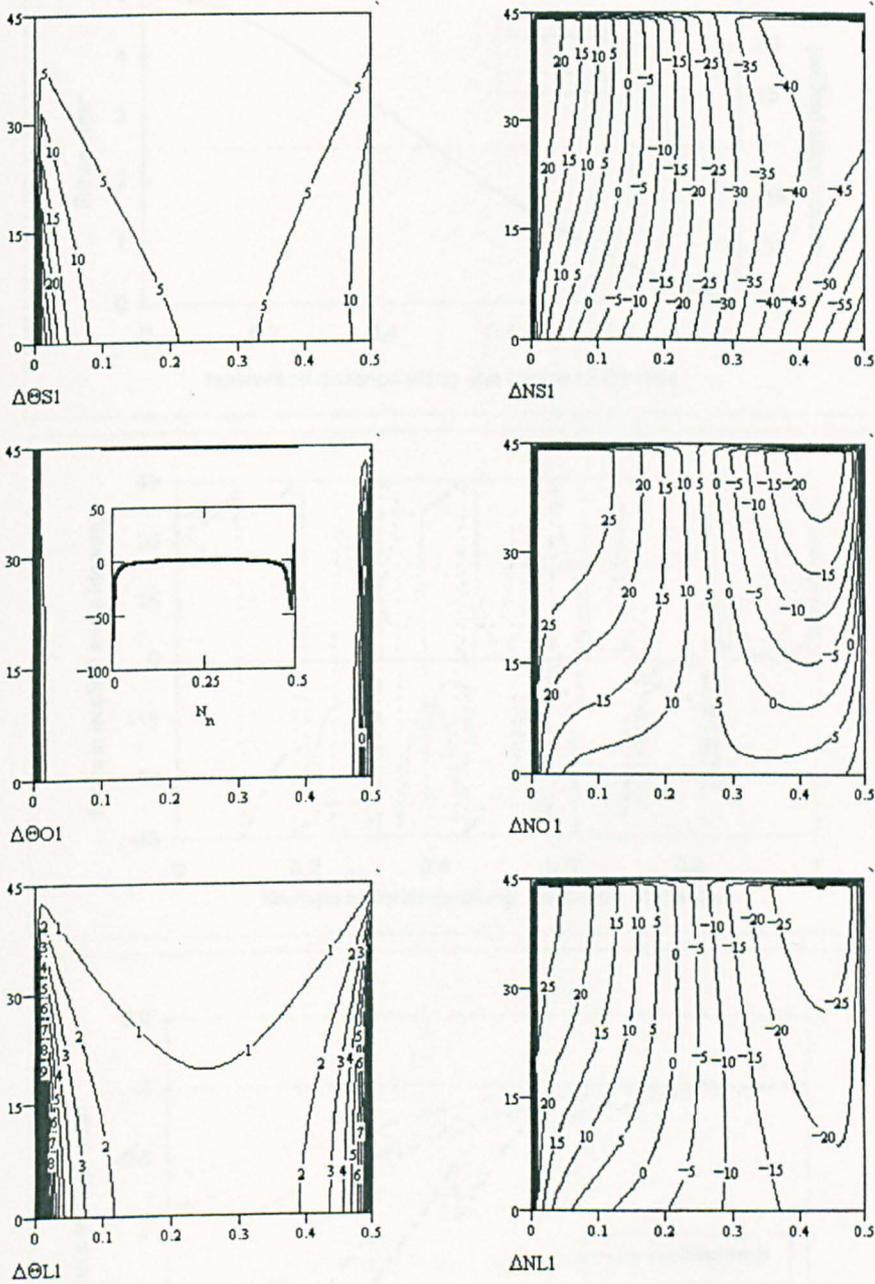


Fig. 5.8 Predicted errors in the isoclinic angle (left panel) and isochromatic fringe order (right panel) due to the three white light sources S1, O1 and L1, as function of the isochromatic and isoclinic parameters in a specimen. The units of fringe order errors is 0.001 fringe. The inset graphs in the left panel show the isoclinic errors as a function of the fringe order for isoclinic angle 15° .

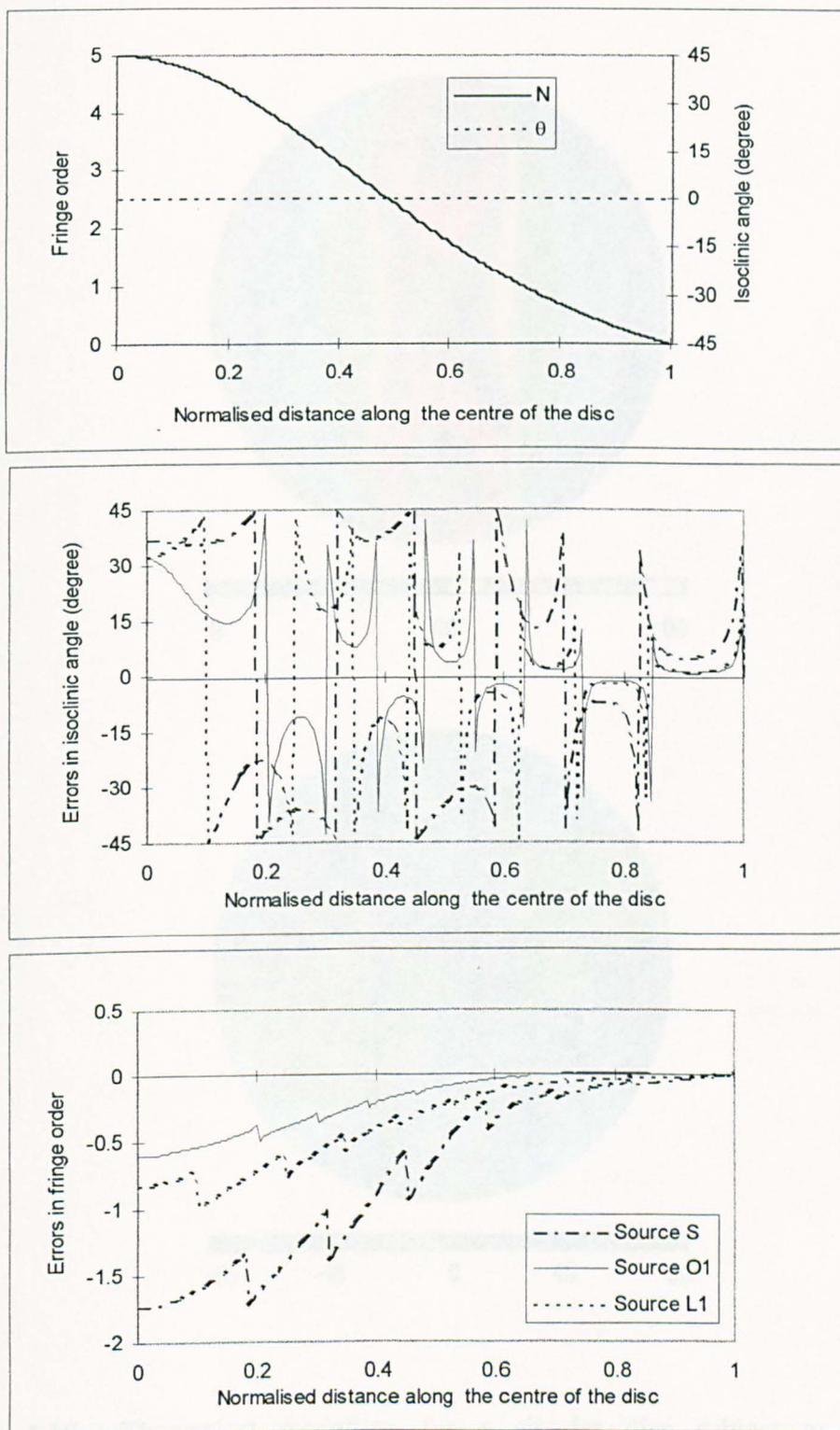


Fig. 5.9 Predicted errors in the isoclinic angle and isochromatic fringe order along the centre line of a disc subject to diametral compression ($N_{\max} = 5$) at three white light sources, where the top panel is the fringe order from the stress-optic law and elasticity theory, the middle and bottom panels show the predicted errors in the isoclinic angle and the fringe order for three cases.

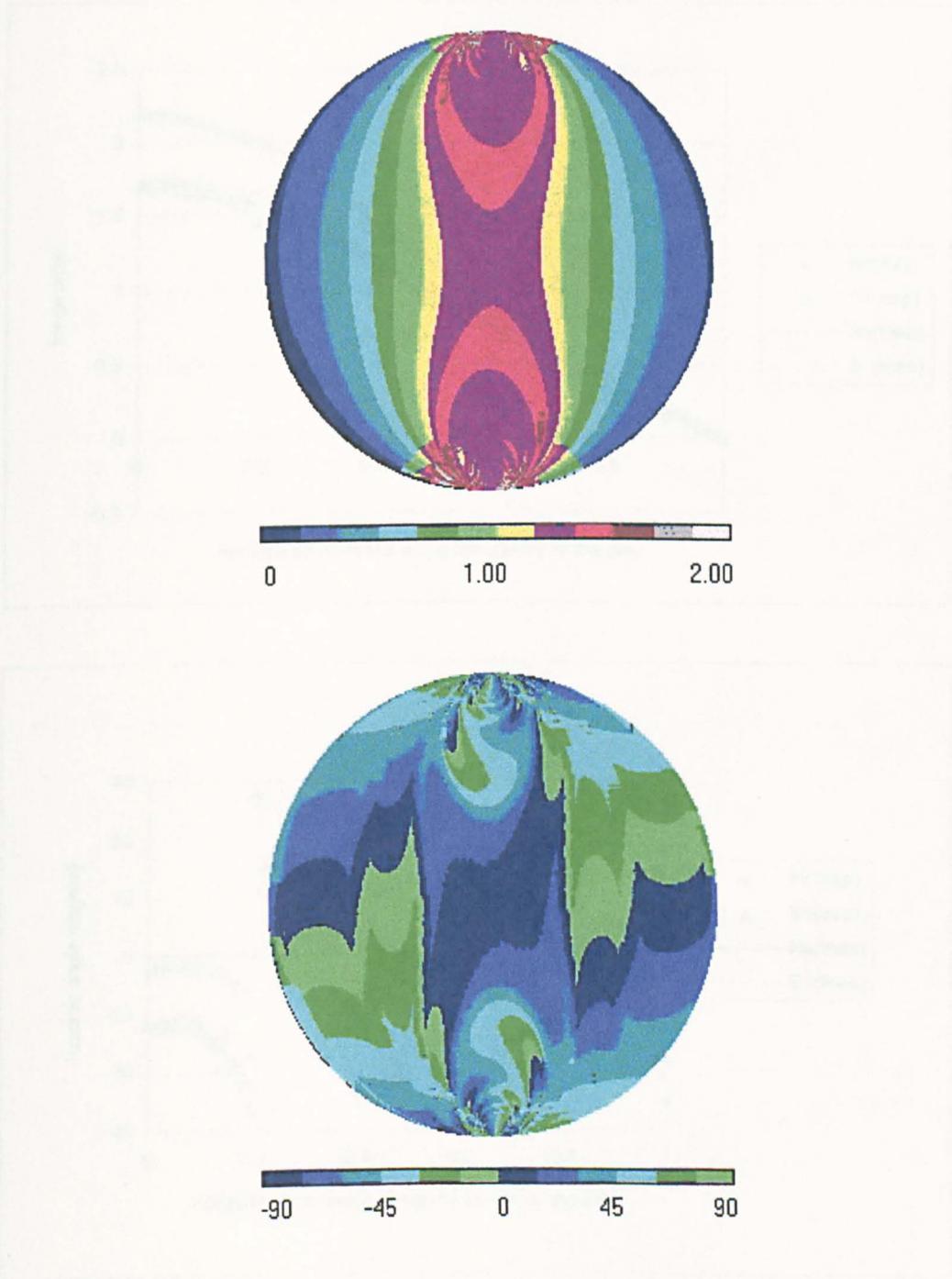


Fig. 5.10 Theoretical modelling for a circular disc subject to diametral compression with white light source S1. The isochromatic fringe order and isoclinic angle are illustrated.

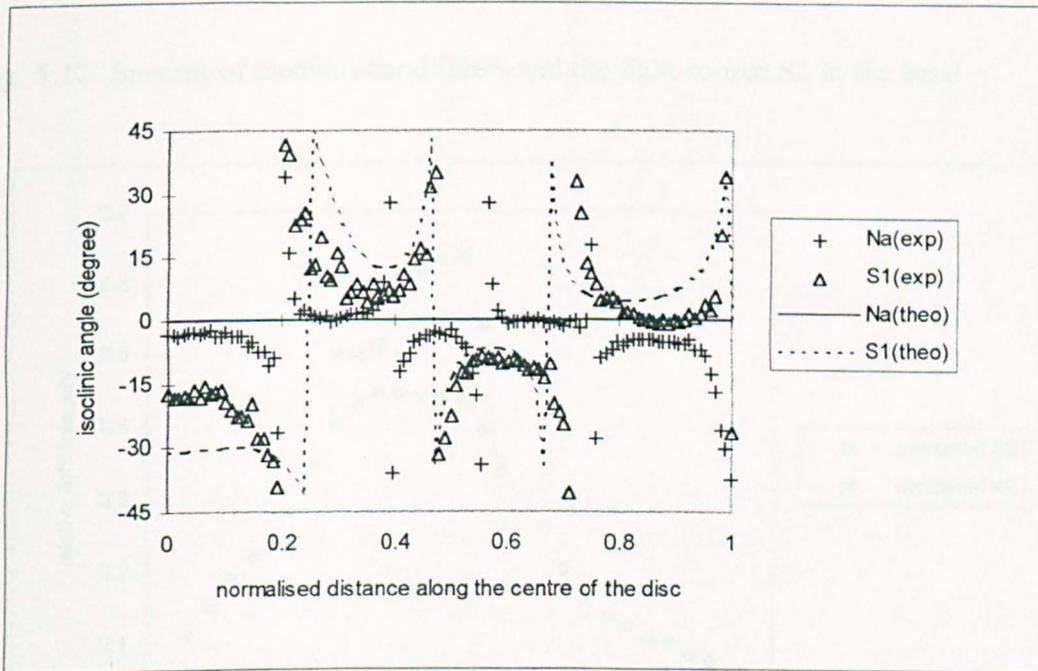
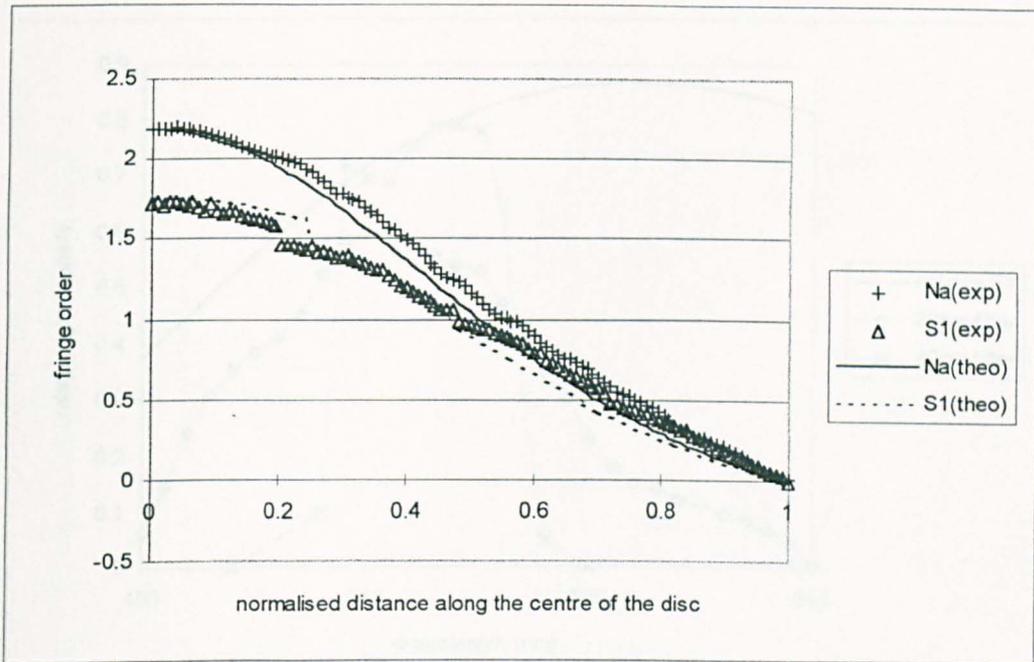


Fig. 5.11 The isochromatic and isoclinic parameters across a diameter perpendicular to the loaded diametry in a disc subject to diametral compression. Experimental and theoretical predictions are shown for (a) isochromatic and (b) isoclinic parameters for the disc in a polariscope using light source S1.

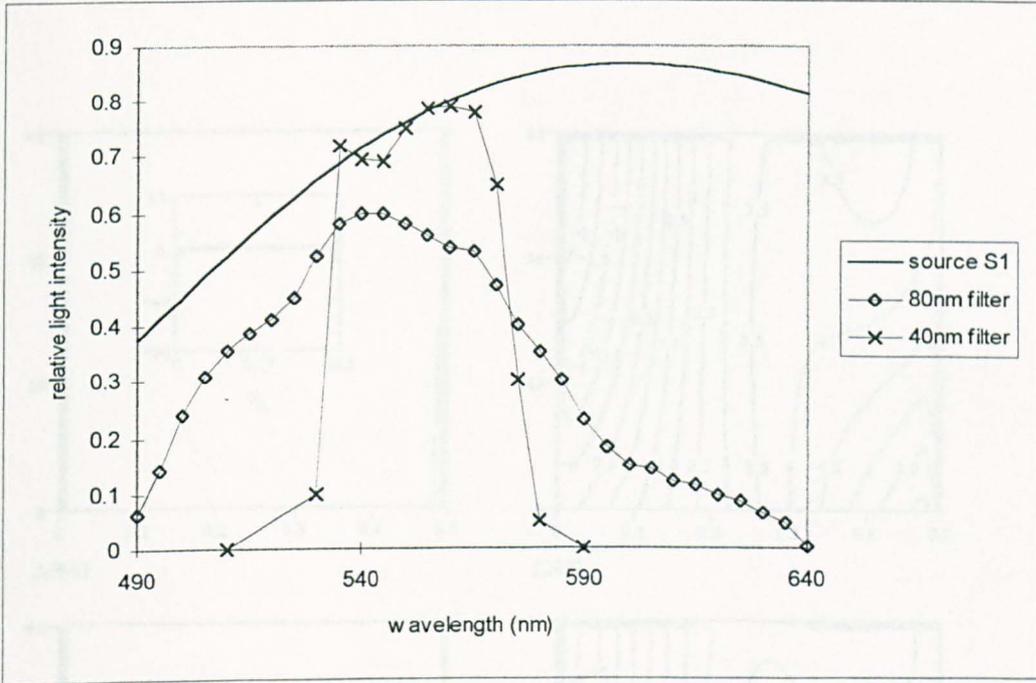


Fig. 5.12 Spectra of medium-band filters and the light source S1 in the band

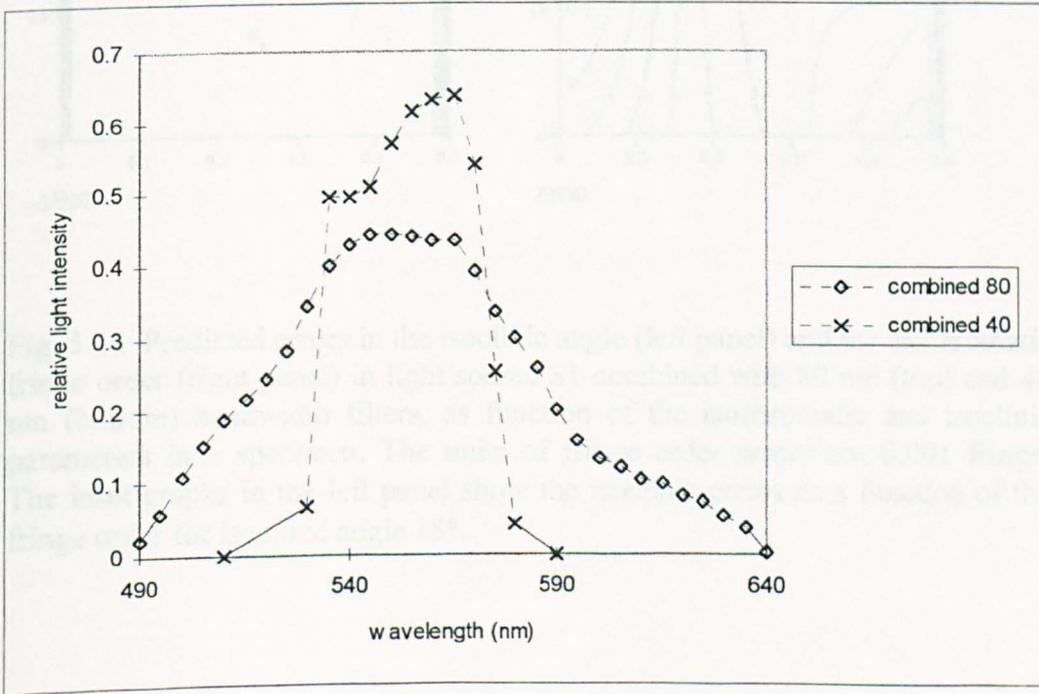


Fig. 5.13 Combined response curves of the medium-band filters and the source S1

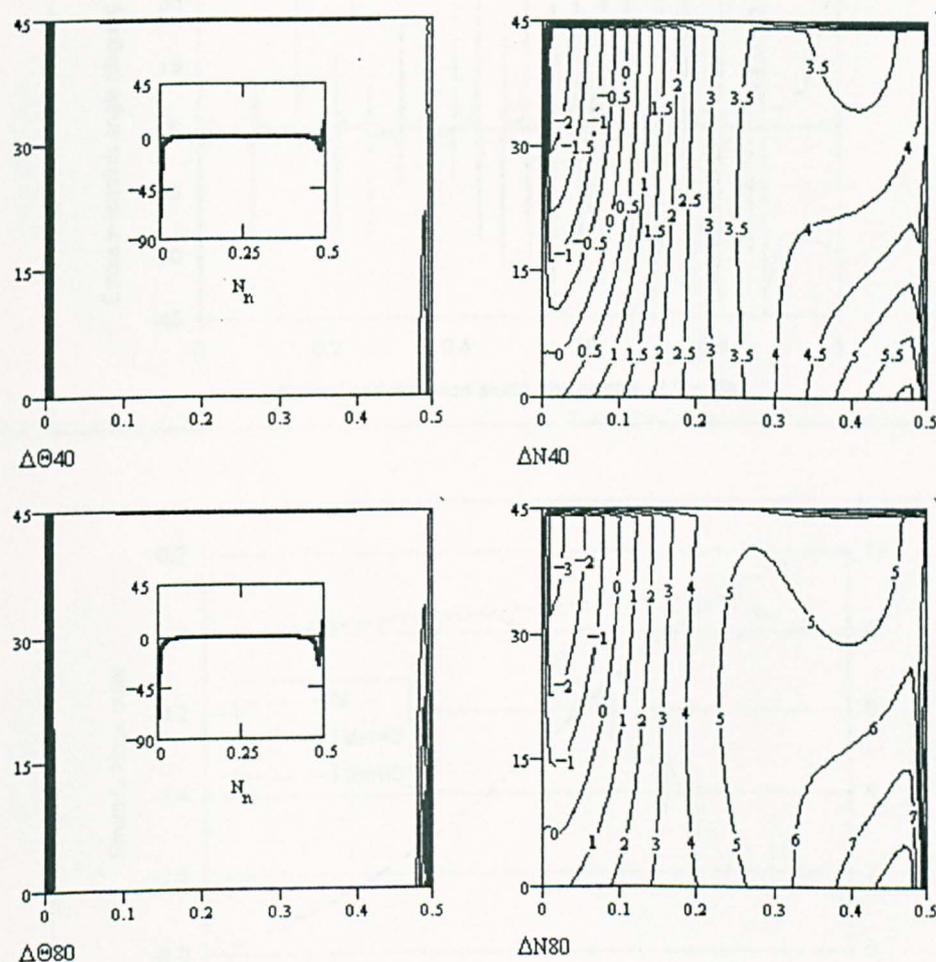


Fig. 5.14 Predicted errors in the isoclinic angle (left panel) and the isochromatic fringe order (right panel) in light source S1 combined with 80 nm (top) and 40 nm (bottom) bandwidth filters, as function of the isochromatic and isoclinic parameters in a specimen. The units of fringe order errors are 0.001 fringe. The inset graphs in the left panel show the isoclinic errors as a function of the fringe order for isoclinic angle 15° .

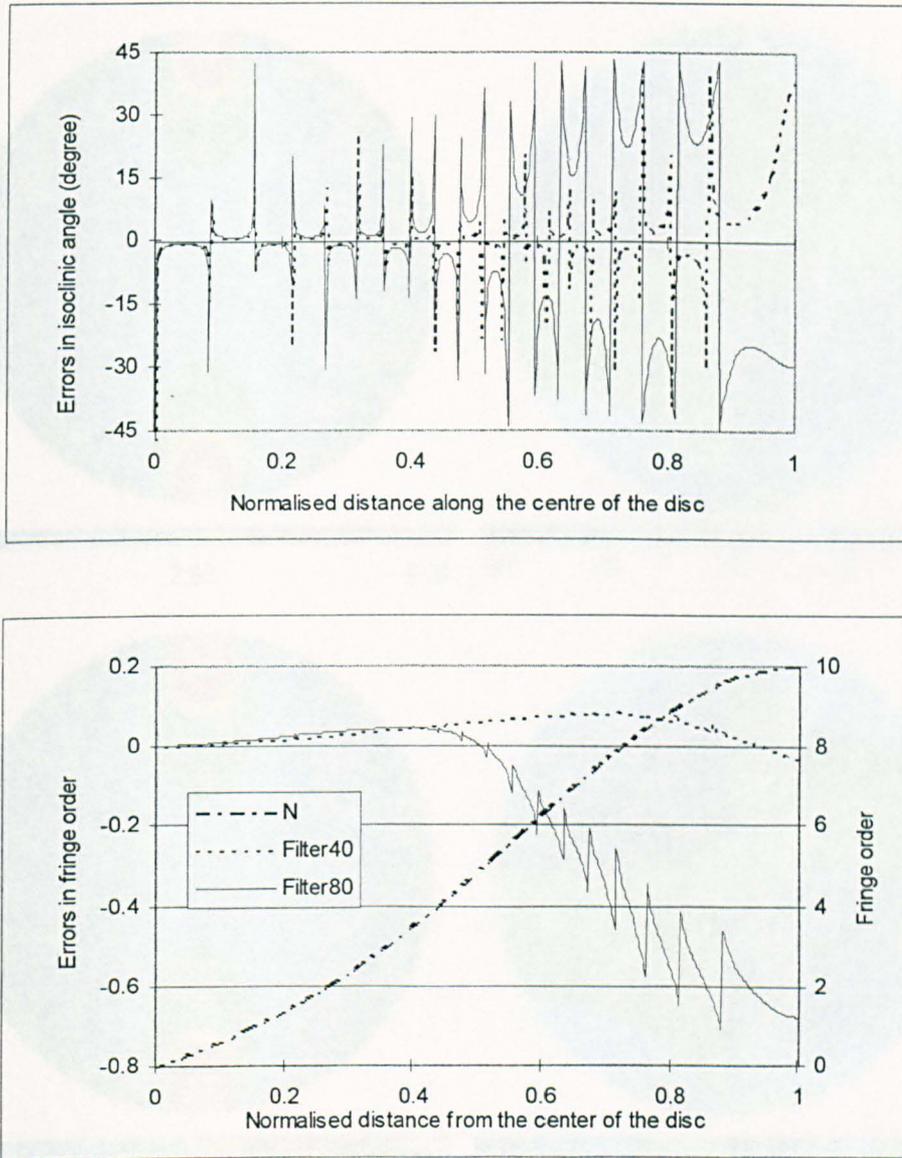


Fig. 5.15 Predicted errors in the isoclinic angle and isochromatic fringe order along the centre line of a disc subject to diametral compression ($N_{\max} = 10$). The worst light source S is combined with the use of medium-band filters.

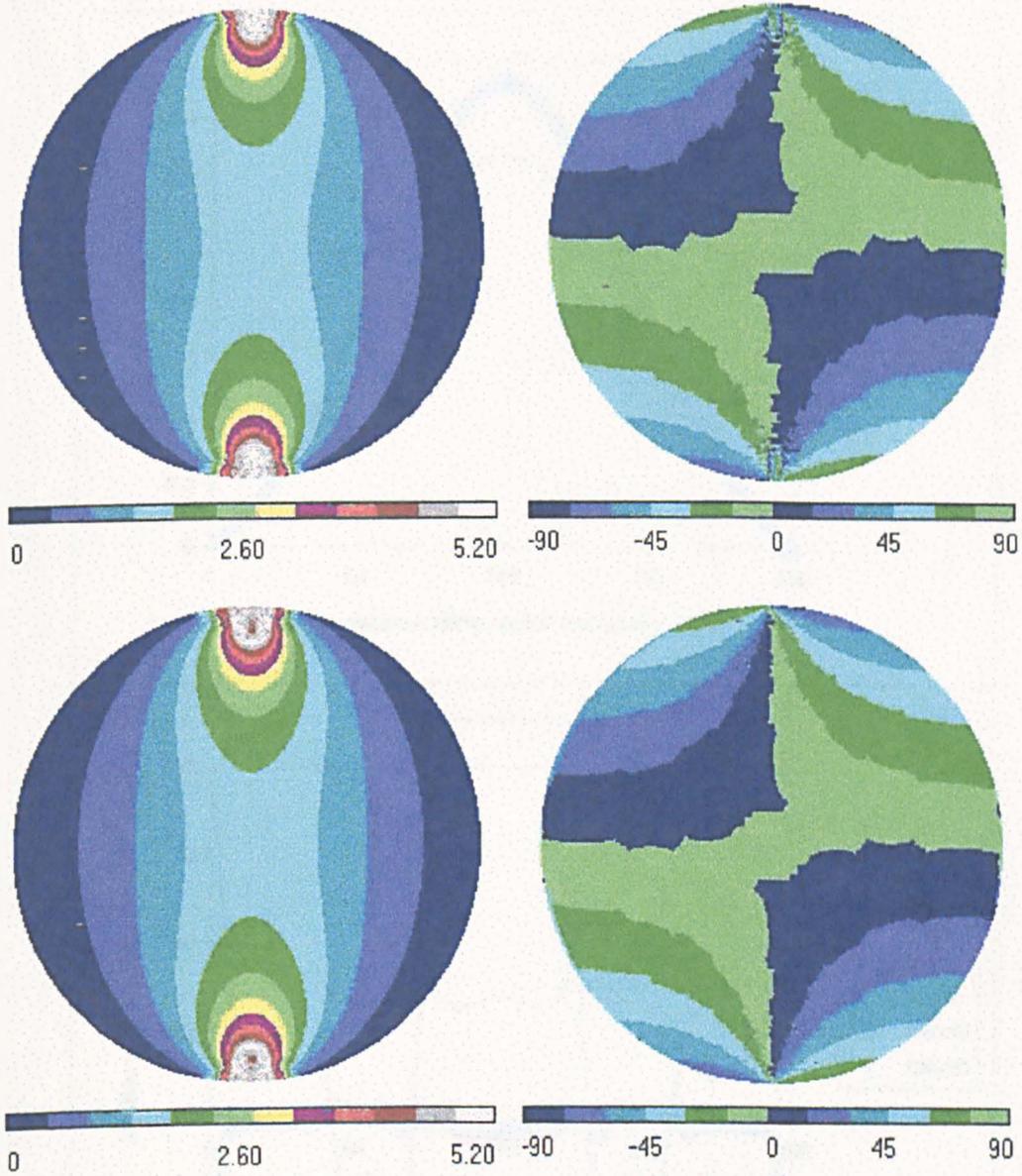


Fig. 5.16 Theoretical modelling for a circular disc subject to diametral compression with two medium band filters and the light source S1. The isochromatic fringe order and isoclinic angle are illustrated. Top panel for filter with bandwidth 80 nm, and bottom panel for filter with bandwidth 40 nm.

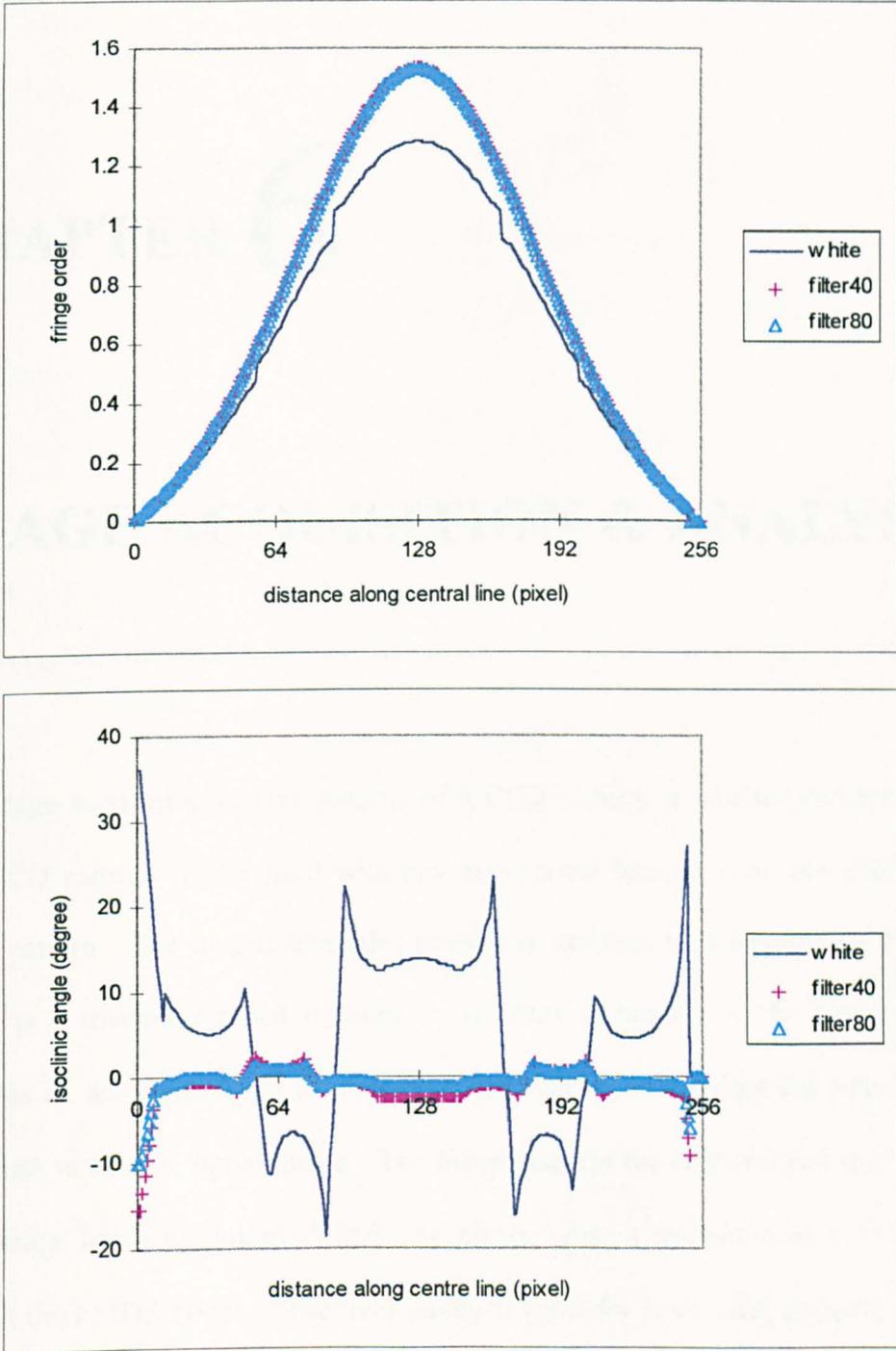


Fig. 5.17 The comparison of the fringe order and the isoclinic angle along the central line of the disc in light source S1 and with two medium-band filters. Data obtained from the simulation discussed in Section 5.4.

CHAPTER 6

IMAGE ACQUISITION & ANALYSIS

The image acquisition system consists of a CCD camera, a monitor and a digitiser. The CCD camera can be used with any appropriate lens to view the photoelastic fringe pattern. The output from this camera is supplied to a Micro Eye TM card which is a transputer-based digitiser. The card is housed in the computer and provides an analogue signal to a monitor for display, and a digitised signal to the computer to form a digital image. The image used in the conventional transmission polariscope had a resolution of 256×256 pixels, while a resolution of 512×512 was used in the PSIOS system. The light intensity value for each pixel occupies an 8-bit word which can express 256 distinct grey levels. The raw images are obtained by image acquisition and the processing of raw images through to the photoelastic results has been termed image analysis. This chapter describes an investigation of the errors in both the image acquisition and the image analysis system.

6.1 Introduction

The image acquisition system consists of a CCD camera, a monitor and a digitiser. The digitiser is divided into two parts: an amplifier and an A/D converter. The fringe pattern is generated by the optical system, and then collected using a standard monochromatic CCD camera. The output of the camera goes to the amplifier and is divided in two; one part supplies the monitor for a visual display and the other part is converted to a digital signal by the A/D converter and supplied to the computer for further digital analysis. The amplifier also provides the functions of brightness and contrast so that the signal intensity can be controlled from the computer keyboard. This chapter contains a discussion of the errors caused by the camera, amplifier and the A/D converter.

An investigation of the errors in the image analysis procedure is also described. A Gaussian smoothing filter is applied to improve the image quality.

6.2 Camera

The concept of the CCD camera was first proposed in 1970 [Amelio *et. al.* 1970 and Boyle *et. al.* 1970]. Nowadays it is widely used in laboratories, research and development work. This popularity is a result of its high resolution, high quantum efficiency, wide spectral response, low noise, linearity, geometric fidelity, fast response, small size, low power consumption and durability.

Due to these advantages the CCD camera was chosen for use in the phase-stepping system to record the images. Errors from the CCD camera are divided into two categories, the errors in the brightness and in the geometry. The former directly

influences the accuracy of isochromatic and isoclinic data, while the latter effects the shape of isochromatic and isoclinic data.

6.2.1 The Linearity of the CCD Camera in the Brightness Level

For the CCD camera, the errors in the brightness may be analysed from two aspects. Firstly, considering the linearity of the relationship between the output voltage and the light intensity; and secondly, considering the sensitivity of the individual CCD cells to the same brightness level.

The experimental configuration used for estimating errors in brightness is shown in Fig. 6.1. In order to reduce the effect of environmental noise, a black tube was used in the experiment to enclose the light path. To minimise the noise, the tube was designed as,

$$L > 10d \quad (6-1)$$

Where L is the length of tube and d is the length of each side of square tube. The Neutral Density Filter (26-5868, Ealing Electro-optics., Inc.) was used as a full-field filter to attenuate the light intensity. The density is defined as the common logarithm of the ratio of incident to transmitted light,

$$D = \log_{10} \frac{I_o}{I_t} \quad (6-2)$$

where I_o and I_t denote the intensity of incident and transmitted light, respectively. The transmission ratio is defined as

$$\text{Transmission Ratio} = \frac{I_t}{I_o} * 100\% \quad (6-3)$$

Eleven different filters were used in sampling the images. The corresponding relationship between the density and transmission ratio is given in Table 6.1.

Table 6.1 Transmission ratios of the Neutral Density filter

Density	2.0	1.0	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1
Transmission %	1	10	13	16	20	25	32	40	50	63	79

The linearity of the camera was measured using the mean value of 256×256 pixels to represent the output grey level, the experimental results are shown in Fig. 6.2. The output of the camera is expected to be linear with the input, but from Fig. 6.2, the linearity only existed for the medium to high intensity range but not in the low intensity range. This response curve is not expected. But it can be explained by comparing the photosensitivity of the camera to the photographic film [Gonzalez, *et al.*, 1993].

The exposure E to which a film is subjected is defined as *energy per unit area* at each point on the photosensitive area. Exposure depends on the incident intensity I and the duration of the exposure T . These quantities are related by the expression

$$E = I T \quad (6-4)$$

The most widely used description of photosensitive properties of photographic film is a plot of the density versus the logarithm of E . These curves are called characteristic curves, D - \log - E curves and H & D curves (after Hurter and Driffield, who developed the method [Longford 1984]). Fig. 6.3 shows a typical H & D curve for a

photographic negative. When the exposure is below a certain level, the density is independent of exposure and equal to a minimum value called the *gross fog*. In the *toe* of the curve, density begins to increase with increasing exposure. Next is a region of the curve in which density is linearly proportional to logarithmic exposure. Finally, the curve saturates in a region called the *shoulder*, and again density does not change with increasing exposure.

In the CCD camera, the duration of the exposure is fixed as 1/24 seconds so the exposure E is equal to the incident intensity I . The density D is then converted to an electronic signal and transferred to the digitiser for further analysis. Comparing Fig. 6.2 with the $H \& D$ curve, the curve shown in Fig. 6.2 is part of the $H \& D$ curve because no saturation region is observed in the experiment. Further explanation can be found in some monographs about CCDs [e. g. Howes 1979 and Barbe 1980].

In practice, the saturation region can be avoided by adjusting the aperture of the lens but the *gross fog*, since it is a minimum density value, always exists.

6.2.2 Sensitivity of the Individual Cell to the Light Intensity

Based on the working principle of the CCD camera, the outputs of the individual cells depend on their own properties and are independent of each other, so it is possible that every cell may have a different sensitivity to the same light intensity. The sensitivities of the individual cells can be investigated by considering the variance of the brightness of the light around its mean. The camera was used to sample a scene on a diffuser, its light distribution along an arbitrary line is shown in Fig. 6.4. The small amount of scatter in Fig. 6.4 may indicate the different sensitivities to the same

light level of the individual cells but the scatter may also be from three other sources, i.e., the light source, the digitiser or the ND filters.

The noise from the light source may have two causes: a) the distribution of the brightness of the light source is not perfectly uniform; b) the light source is not absolutely stable with time. (The results from time-varying sampling of the light source were shown in Fig. 4.14.)

The scatter in Fig. 6.4 may also have come from the digitiser. Because the output from the digitiser is a digital signal, quantisation errors are introduced during the analogue to digital conversion. Additional errors due to the unstable quantisation process may be also introduced, these lie in the range -0.5 to +0.5 grey levels. The details of this effect will be discussed in **Section 6.4**.

The variance varies with density and the maximum variance occurred when using the maximum density filter ($\sigma_{\max}^2 = 7.35$) with 79% transmission. Hence, the scatter may have come from the inhomogeneities in the ND filters.

Since the light passed through the diffuser, the light intensity along an arbitrary line should be a smooth distribution. The scatter in Fig 6.4 means that individual CCD cells have different sensitivities to the same light level.

Since the scatter occurred in Fig. 6.4 is the combined effect of several sources, it is impossible to distinguish the individual errors caused by the cells of the CCD camera. The total effect on the light intensity will be discussed in **Section 7.2.1**.

6.2.3 Errors in the Geometry

In the phase-stepping system, the optical axis of the camera should be perpendicular to the plane of the optical elements and of the specimen for an accurate recording of the results. In practice, this is achieved by manual inspection rather than by any specific equipment or procedure and so errors due to non-parallelism may be introduced, which have been discussed in a previous chapter. By comparison with the geometrical errors from the CCD camera itself, such as have been reported in reference [Tsai 1986 and Healey, 1994], the geometrical errors in the polariscope are much greater than the camera's errors. It is reasonable to assume that there are no geometrical errors introduced by the CCD camera.

6.2.4 Simulation the Effect of the Camera Nonlinearity

The output of the camera is not perfectly linear with the inputs of the light intensity in the extremely low region. Two lines, for simplicity, were used to fit the curve in Fig. 6.2 by the least square method as,

$$I_{out} = I_{in} \times 0.75 \quad \text{if } I_{in} \leq 40; \quad (6-5a)$$

$$I_{out} = I_{in} \times \frac{45}{43} - \frac{510}{43} \quad \text{otherwise} \quad (6-5b)$$

Assuming for the case of the conventional polariscope with six images, if the input intensities of the images i_1 to i_6 are less than 40 (grey level), then the output of the light intensity can be simulated by Eq. 6-5a; otherwise, the output is simulated by Eq. 6-5b. In this case, the image 1 and image 2 of the disc model are shown in Fig. 6.5. The sum of image 1 and image 2 should be a constant according to the

theoretical formula, but now fringe pattern can be observed from their summation. Similar phenomenon can be observed in the experimental data, as shown in Fig. 6.6.

The nonlinearity of the CCD camera to the input light intensity would introduce errors in the isoclinic and isochromatic data since it changes the values of the light intensities i_m and i_v , which are assumed to be constant in the whole analysis procedure. Six images can be obtained using the same way as described for Fig. 6.5, then the modelling results for the isochromatic and isoclinic data are shown in Fig. 6.7. Comparisons were made for the disc with the linear response along the central line and are shown in Fig. 6.8. No obvious influence of the non-linearity can be observed in these two figures and this is due to the nonlinearity only influencing the darkest region, at which integer and half fringes are located and the isoclinic angles are undefined, and the values at this region are replaced by a use of spline function.

6.2.5 Discussion on the Nonlinearity of the Camera

The nonlinearity of the camera was not expected but can not be avoided in practice. This nonlinearity only influences the darkest region. Because the isoclinic angles at integer and half fringes are undefined and splined by the neighbouring points, this phenomenon makes splining even more difficult while having almost no influence on the other regions.

6.3 Amplifier

6.3.1 Introduction

The output of the camera is supplied to the Micro Eye TM card, which is a transputer-based digitiser. The card is divided into two parts, one is an amplifier and the other is an A/D converter. The signal from the camera is transferred via the amplifier and divided into two parts again; one part goes to the monitor for a visual display and the other part to the A/D converter for conversion from an analogue signal to a digital signal. The amplifier also provides a function to allow users to adjust the brightness and contrast of the output by moving the cursor keys on the computer keyboard. Users could only adjust the brightness and contrast parameters based on their own experience which is difficult for novices. It is desirable to provide principles for adjusting the brightness and contrast parameters.

6.3.2 Function of Brightness and Contrast Control

In order to understand how the brightness and contrast adjustment function works, the following experiments were carried out. A stress-frozen thread made from MY750 [Kenny, 1968] was used as a specimen and viewed with a 15 inch diameter diffuse light polariscope in which was installed the sodium light source (Sharples Stress Engineers Ltd, Preston, England). One set of image data has been obtained with only the brightness parameter changed while another set with only the contrast changed. Both the brightness and contrast parameters varied in the range from 0 to 255. The corresponding value of brightness and contrast are given in Table 6.2.

Table 6.2 Value of the brightness and contrast parameters

1	File name	Brightness parameter	Contrast parameter	2	File name	Brightness parameter	Contrast parameter
	B_1	127	127		C_1	127	127
	B_2	132	127		C_2	127	132
	B_3	137	127		C_3	127	137
	B_4	142	127		C_4	127	142

6.3.2.1 Change of the Brightness Parameter

First, the contrast parameter was fixed at the median value and the brightness parameter was increased by increments of 5 from its median value. Four sets of light intensity data were obtained as $B_1(x, y)$, $B_2(x, y)$, $B_3(x, y)$ and $B_4(x, y)$, where (x, y) are the co-ordinates of the pixel. The light intensities along an arbitrary line (e.g., $x = 57$) are shown in Fig. 6.9. By observation, the light intensity was clearly increasing with the increments of the brightness parameter, as shown in Fig. 6.10, for which:

$$\Delta B_j(x, y) = B_{j+1}(x, y) - B_j(x, y), \text{ where subscript } j=1, 2, 3, \quad (6-6)$$

The average value $\overline{\Delta B}$ at each point and its standard deviation $\sigma \Delta B_j$ for all points were then obtained by:

$$\overline{\Delta B} = \left[\sum_{j=1}^3 \Delta B_j(x, y) \right] / 3 \quad (6-7)$$

$$\sigma \Delta B_j^2 = \left[\sum_{y=0}^{255} (\Delta B_j(y) - \overline{\Delta B})^2 \right] / 256 \quad (6-8)$$

The average value of $\overline{\Delta B}$ and the standard deviation $\sigma\Delta B_j$ are shown in Fig. 6.11 and Table 6.3.

Table 6.3 The average value of $\overline{\Delta B}$ and the standard deviation $\sigma\Delta B_j$ due to the change in the brightness parameter

$\overline{\Delta B}$	$\sigma\Delta B_1$	$\sigma\Delta B_2$	$\sigma\Delta B_3$
7.354	1.173	1.107	1.119

Because the standard deviation is only about 1 and relatively small compared to the average value of $\overline{\Delta B}$, ΔB is then considered as a constant. Assuming that the light intensity before the brightness parameter changes is denoted by I , and after the brightness parameter changes is denoted by I' , then a relationship between I and I' can be expressed as,

$$I' = I + b \quad (6-9)$$

where b is the brightness coefficient. Changing the brightness parameter is therefore only increasing or decreasing the light intensity value.

6.3.2.2 Change of the Contrast Parameter

In this section the brightness parameter was fixed at its median value and the contrast value was increased in increments of 5 from its median value. Four sets of light intensity data were obtained as $C_1(x, y)$, $C_2(x, y)$, $C_3(x, y)$ and $C_4(x, y)$, where (x, y) are the co-ordinates of the pixels. The light intensity along an arbitrary line ($x = 57$) is shown in Fig 6.12. The difference in the light intensity is similar obtained as,

$$\Delta C_j(x, y) = C_{j+1}(x, y) - C_j(x, y), \text{ where subscript } j = 1, 2, 3 \quad (6-10)$$

Clearly in this case, as shown in Fig. 6.13, the difference in the light intensity $\Delta C_j(x, y)$ was no longer a constant with increasing the contrast parameter. The difference of light intensity $\Delta C_j(x, y)$ is a function related to the original image if the noise is ignored. So the adjustment of the contrast is more like to stretch the amplitude of the fringe pattern, i.e.,

$$\Delta C_j(x, y) = C_j(x, y) \times \Delta c_j(x, y) \text{ when only the contrast level is changed;} \quad (6-11)$$

Hence,

$$\Delta c_j(x, y) = \Delta C_j(x, y) / C_j(x, y) \quad (6-12)$$

The contrast coefficient Δc was obtained for each increment applied and is shown in Fig. 6.14. The mean value $\overline{\Delta c}$ at each point and its standard deviation $\sigma \Delta c_j$ for all points were found as:

$$\overline{\Delta c} = \left[\sum_{j=1}^3 \Delta c_j(y) \right] / 3 \quad (6-13)$$

$$\sigma \Delta c_j^2 = \left[\sum_{y=0}^{255} (\Delta c_j(y) - \overline{\Delta c})^2 \right] / 256 \quad (6-14)$$

The average value of $\overline{\Delta c}$ and the standard deviation $\sigma \Delta c_j$ due to the change of the contrast parameter are plotted in Fig. 6.15 and the mean values over 256 points are listed in Table 6.4. It is clear that the change in the contrast parameter is equal to the light intensity multiplied by a constant. Again, I and I' are used to denote the light intensity before contrast parameter changes and after the contrast parameter changes, then the relationship between I and I' is,

$$I' = I \times c \quad (6-15)$$

Table 6.4 The average value of $\overline{\Delta c}$ and the standard deviation $\sigma\Delta c_j$ due to the change in the contrast parameter

$\overline{\Delta c}$	$\sigma\Delta c_1$	$\sigma\Delta c_2$	$\sigma\Delta c_3$
1.178	0.032	0.023	0.021

6.3.2.3 Function of Brightness and Contrast Control

Because the standard deviations of $\sigma\Delta B_j$ and $\sigma\Delta c_j$ are relatively small compared with their mean values, the relationship between the light intensity after brightness and contrast changes, I' , and before changes, I , can be expressed by combining Eqs. 6-9 and 6-15 as,

$$I' = I \times c + b \quad (6-16)$$

The above conclusion was drawn from an arbitrary set-up of the polariscope. Hence, it can be concluded that the change in the brightness and contrast parameters is equal to adding or multiplying by a coefficient to the light intensity.

6.3.3 Effect of Brightness and Contrast

6.3.3.1 Overview on Adjusting Brightness and Contrast Parameters

In the automated phase stepping system, the image acquisition is a very important step as all the following analysis is based on the six (or four) images and therefore the

quality of the images directly influences the results of the isochromatic and isoclinic data. Based on experience a good image is expected to have an adequate range of intensity without any saturation. All of the six (or four) images used for analysis should meet this requirement with the same brightness and contrast parameters. So basically an experienced user would try to adjust the brightness and contrast level until the intensity range of all images is as large as possible.

The following analysis is using the conventional polariscope as an example, but the principle for the PSIOS system is the same.

6.3.3.2 Effect in theory

The function representing the adjustment of the brightness and contrast levels was obtained as shown in Eq. 6-16, and hence the isoclinic angle and relative retardation may be found by substituting Eq. 6-16 into Eqs. 3-16 and 3-17,

$$\begin{aligned}\theta' &= \frac{1}{2} \arctan \frac{i_3' - i_5'}{i_6' - i_4'} = \frac{1}{2} \arctan \frac{(i_3 \times c + b) - (i_5 \times c + b)}{(i_6 \times c + b) - (i_4 \times c + b)} \\ &= \frac{1}{2} \arctan \frac{i_3 - i_5}{i_6 - i_4} = \theta\end{aligned}\quad (6-17)$$

$$\begin{aligned}\alpha' &= \arctan \frac{i_6' - i_4'}{(i_2' - i_1') \times \cos 2\theta'} = \arctan \frac{(i_6 \times c + b) - (i_4 \times c + b)}{[(i_2 \times c + b) - (i_1 \times c + b)] \times \cos 2\theta} \\ &= \arctan \frac{i_6 - i_4}{(i_2 - i_1) \times \cos 2\theta} = \alpha\end{aligned}$$

(6-18)

So from Eqs. 6-17 and 6-18, it can be observed that the change in brightness and contrast parameters does not influence the isoclinic and the isochromatic data.

Therefore adjustment of brightness and contrast is not necessary based on theoretical analysis.

6.3.3.3 Truncating error

The adjusting of the brightness and contrast would not influence the photoelastic results from the point of view of the theoretical analysis. This conclusion assumes that there are no truncating errors in raw images. However in practice, since all the raw image data is expressed by 256 integer values, a low level of the brightness or contrast parameters would produce a small range of light intensity which would introduce a truncating error.

The following simulation shows the effect of the truncating error. Three sets of data were used for comparison. The first set used the data shown in Fig. 3.8a, which is an ideal circular disc without any error; the second set was obtained by multiplying the first set of images by 0.5 while the third set was obtained by multiplying the second set by 0.5. Referring to Eq. 6-16, assuming that the light intensity is I in the first set of images, then the brightness coefficient is 0 and the contrast coefficient is 0.5 for the second set; and 0 and 0.25 for the third set. The raw image comparison is shown in Fig. 6.16a for image 1. The processing results of the three sets of data by PECAN are shown in Fig. 6.16b.

Although a large difference exists in the intensity values in the raw images, there is no difference between the first set and second set of results and only a slight difference in the results for the third set. This later difference is caused by the truncating error and mainly expressed as the wrapping error in the high fringe gradient region. So it can be concluded that the adjustment of the brightness and contrast

parameters would not have any influence on the isochromatic and the isoclinic data in most cases, but the truncating error would have an effect when the brightness and contrast values are extremely low.

6.3.3.4 Conclusion on the Effect of the Brightness and Contrast

The low values of the brightness and contrast parameters would introduce a truncating error. This will increase the undefined zones of the isoclinic angle and make wrapping errors occurred in high fringe gradient region. But generally the adjustment of the brightness and contrast parameters would not influence the isochromatic and isoclinic data.

6.3.4. Saturation

Saturation is another important factor to influence the analysis. In the above analysis, the assumption is made that saturation does not occur during the adjustment of the brightness and contrast parameters. Actually, the phenomenon of the saturation is closely connected with the function of the brightness and contrast control.

6.3.4.1 Introduction

The chip of the CCD camera collects an optical signal and outputs an electronic signal. In the meantime, the electronic analogue signal is converted to an electronic digital signal by the A/D converter. An 8 bit word was used to express the light intensity from 0 to 255. If the input is larger than the value 255 or less than 0, the computer would give the output as 255 or 0. It is called saturation in both cases.

Normally, saturation caused by the signal being over 255 occurred more often than the other therefore more attention was paid for this case. The following analysis only discusses this case.

The saturation is caused mainly by two factors. First if the lens aperture is open too wide, e.g., F2.8 or F4, then saturation occurred. In this case, decreasing the brightness and contrast level would cause the saturated value to become less than 255. Secondly, increasing the brightness and contrast level excessively would also cause saturation. Once a saturation value is acquired, it is not possible to obtain accurate photoelastic data.

6.3.4.2 Detection of Saturation

Images 1 and 2 are much more important than the other four images for the judgement of the saturation level according to the experience. Proof can be found from the formulae of the phase-stepping system. Referring to Eqs. 3-14a to 3-14f, the fractional isochromatic and isoclinic angle can vary from (-90) to $(+90)$ degree, so the value of $(\cos\alpha)$ can vary from (-1) to $(+1)$. Hence, in images 1 and 2 in the conventional polariscope, the light intensities i_1 and i_2 must satisfy the following expression,

$$i_m - i_v \leq i_1, i_2 \leq i_m + i_v \quad (6-19)$$

The detection of the saturation is actually detecting whether the value of $(i_m + i_v)$ is less than 255. If not, the saturation must be occurring at the point where the isochromatic retardation is zero. Considering the image 1 and image 2, there must be some areas where the light intensity is equal to 255.

On the other hand, if the saturation is not observed in images 1 and 2, it will not occur in images 3 to 6. Since the value of $(\sin\alpha)$ is in the range $(-1, +1)$ and this term is further multiplied by $(\sin 2\theta)$ or $(\cos 2\theta)$ in images 3 to 6, which is not larger than $(-1, +1)$ again, hence the light intensity of i_3 and i_6 must have,

$$i_m - i_v \leq i_3, i_4, i_5, i_6 \leq i_m + i_v \quad (6-20)$$

Therefore, the detection of the saturation is only necessary in images 1 and 2.

6.3.5 Conclusions on Amplifier Control

The amplifier provides the function of brightness and contrast control. The adjustment of the brightness parameter increases or decreases the light intensity value of the whole image, while the adjustment of the contrast parameter is a multiplicative coefficient that acts on the whole image. The control of the brightness and contrast parameters does not have any effect on the isochromatic and isoclinic results if the truncating error is ignored.

Saturation is caused by too large a lens aperture or the parameters of brightness and contrast being too high. Saturation should not be allowed to occur during sampling of the image. The detection of saturation is only necessary in image 1 and image 2 in the conventional polariscope or the corresponding ones in the PSIOS system.

6.4 A/D Converter

When an image is taken, two kinds of digitisation occur. One is the spatial digitisation, the other is the digitisation of the brightness level. Typically, a scene is mapped onto an image, an m by n array of points. Such an array is used to represent the world in front of the camera. As a result of the resolution of this discretisation, an error is introduced into computations that involve the locations of image points and features. Because the number of the levels of brightness level that the receiver can distinguish is limited, the brightness at a given pixel in the image plane is also digitised into, say, κ values. The digitisation of the brightness level is a drawback for reliable computations. Unlike most other types of errors, the digitisation error can not be reduced by performing a more careful experiment. Rather it is caused by the inherent limitations of the digitiser.

6.4.1 Spatial Quantisation

In the phase-stepping system, a scene in the specimen was recorded by the camera and transferred via the digitiser to the computer. The compression of the scene into an image, a 256×256 array of points, means that a spatial quantisation error is introduced. The error in spatial quantisation depends on the size of the specimen and the number of fringes. The simulation results show that when the resolution of the image is larger than 0.26 fringes / pixel (**Section 3.4.2**), then a wrapping error would occur. So the error in spatial quantisation is negligible if the resolution is less than 0.26 fringes per pixel.

6.4.2 Digitisation of the Brightness Level

Evaluation of digitisation error in computer vision was studied by Kamar-parsi *et. al.* [1988] and an analytic expression for the probability density of error distribution due to a large number of independently digitised variables was also derived. In the phase-stepping system, the light intensity was represented by an 8-bit word length and so the brightness was divided into 256 grey levels. The maximum digitisation error in light intensity, i , is half the size of the digitisation unit, $\gamma = 2^{-8} = 1/256$. That is to say, the actual value, i_a , of the light intensity could be anywhere in the interval $I = [i_d - \gamma/2, i_d + \gamma/2]$, where i_d is the value of i after digitisation. Furthermore, the likelihood of the actual value of i , i.e., i_a , being at a certain place within the interval follows a uniform probability density, as shown in Fig. 6.17. In other words, the probability of lying inside a small interval di around the point within I is independent of i and equal to di/γ . Thus, the mean of the digitisation error in the light intensity is given by:

$$m_e = \int_{-\gamma/2}^{\gamma/2} x \cdot p(x) dx = \int_{-\gamma/2}^{\gamma/2} x \cdot \frac{1}{\gamma} dx = 0 \quad (6-21)$$

For a symmetric uniform distribution, the mean of the digitisation error is always zero, therefore an average error, which is the average of the absolute value of the digitisation error a_e , is introduced.

$$a_e = \int_{-\gamma/2}^{\gamma/2} |x| \cdot p(x) dx = \frac{1}{\gamma} \int_{-\gamma/2}^{\gamma/2} |x| dx = \frac{\gamma}{4}, \quad (6-22)$$

The variance of the digitisation error in the light intensity is given by

$$\sigma_e^2 = \int_{-\gamma/2}^{\gamma/2} (x - m_e)^2 \cdot p(x) dx = \int_{-\gamma/2}^{\gamma/2} x^2 \cdot \frac{1}{\gamma} dx = \frac{\gamma^2}{12} \quad (6-23)$$

The isochromatic data is a function of six light intensities for the conventional polariscope. Each light intensity is digitised, the actual value, i_{ja} , of the quantity $i_j \mid j = 1, 2, \dots, 6$ has equal probability within the interval $I_j = [i_{jd} - \gamma/2, i_{jd} + \gamma/2]$. As before, i_{jd} is the value of i_j after the digitisation. To obtain the error in α due to digitisation errors in i_j values, consider the Taylor Series expansion of α , retaining only the first order terms in the digitisation error, $\Delta i_j = i_j - i_{jd}$:

$$\Delta \alpha \approx \frac{\partial \alpha}{\partial i_1} \Delta i_1 + \frac{\partial \alpha}{\partial i_2} \Delta i_2 + \dots + \frac{\partial \alpha}{\partial i_6} \Delta i_6 \quad (6-24)$$

This expansion is valid when Δi_j is small compared to i_{jd} and it is true for most points of the phase-stepping. Then the maximum error in α , i.e. $E_{\alpha \max}$, is given by:

$$E_{\alpha \max} = \left| \frac{\partial \alpha}{\partial i_1} E_{1 \max} \right| + \left| \frac{\partial \alpha}{\partial i_2} E_{2 \max} \right| + \dots + \left| \frac{\partial \alpha}{\partial i_6} E_{6 \max} \right| = \frac{\gamma}{2} \cdot \sum_{j=1}^6 |\alpha_j|, \quad (6-25)$$

where $E_{j \max} = \gamma/2$ is the maximum error in the six light intensities i_j , and the coefficients $\alpha_j = \frac{\partial \alpha}{\partial i_j}$ are evaluated at i_{1d}, \dots, i_{6d} .

In the same way, the isoclinic is a function of four light intensities, thus the maximum error in θ , i.e., $E_{\theta \max}$, is given by:

$$E_{\theta \max} = \left| \frac{\partial \theta}{\partial i_3} E_{3 \max} \right| + \left| \frac{\partial \theta}{\partial i_4} E_{4 \max} \right| + \dots + \left| \frac{\partial \theta}{\partial i_6} E_{6 \max} \right| = \frac{\gamma}{2} \cdot \sum_{j=3}^6 |\theta_j|, \quad (6-26)$$

The most realistic measure of the digitisation error for the isochromatic data is the average error $\overline{E_\alpha}$ which is given by:

$$\overline{E_\alpha} = \frac{1}{\gamma_1 \dots \gamma_6} \int_{-\gamma/2}^{\gamma/2} di_1 \dots \int_{-\gamma/2}^{\gamma/2} di_6 \left| \sum_{j=1}^6 \alpha_j i_j \right| \quad (6-27)$$

Because the evaluation of the average error involves the integral of the absolute value, it is more complicated for most cases, and often, for convenience the standard deviation, $\Delta\sigma_\alpha^2$, is used as a measure of the error. The standard deviation of digitisation is defined through [Bevington, 1969],

$$\begin{aligned}\Delta\sigma_\alpha^2 &= \frac{1}{\gamma_1 \cdots \gamma_6} \int_{\gamma_1/2}^{\gamma_1/2} di_1 \cdots \int_{\gamma_6/2}^{\gamma_6/2} di_6 (\Delta\alpha)^2 \\ &= \frac{1}{12} (\alpha_1^2 \gamma_1^2 + \cdots + \alpha_6^2 \gamma_6^2) = \frac{\gamma^2}{12} (\alpha_1^2 + \cdots + \alpha_6^2)\end{aligned}\quad (6-28)$$

The standard deviation of digitisation for the isoclinic data is given by:

$$\begin{aligned}\Delta\sigma_\theta^2 &= \frac{1}{\gamma_3 \cdots \gamma_6} \int_{\gamma_3/2}^{\gamma_3/2} di_3 \cdots \int_{\gamma_6/2}^{\gamma_6/2} di_6 (\Delta\theta)^2 \\ &= \frac{1}{12} (\theta_3^2 \gamma_3^2 + \cdots + \theta_6^2 \gamma_6^2) = \frac{\gamma^2}{12} (\theta_3^2 + \cdots + \theta_6^2)\end{aligned}\quad (6-29)$$

The digitisation errors in the isochromatic and isoclinic data have been deduced. A contribution of the digitisation error to the total effect on the light intensity will be discussed in Section 7.2.2.1.

6.4.3 Conclusions

Converting from the analogue signal to the digital signal means a spatial quantisation, and a digitisation of the brightness level. The errors due to the conversion are inherent limitations of the digitiser and, unlike most types of error, they can not be reduced by performing a more careful experiment. The spatial quantisation error is negligible if the resolution of the image is less than 0.26 fringes per pixel. The digitisation error of the brightness level has been analysed and its influence on the isochromatic fringe order and the isoclinic angle has been deduced.

6.5 Image Analysis and Image Pre-processing

In the previous discussion, the fringe pattern was generated by the optical system and collected by the image acquisition system. The raw images were acquired for analysis. The procedure obtaining the photoelastic results from the raw images is termed “image analysis”. In this section an investigation into the errors in the image analysis procedure is described.

6.5.1 Accuracy of Image Analysis

To verify the accuracy of the image analysis, the model for the compressive disc was used again. The reason for using the simulation rather than experimental data is the accuracy of the experimental results is effected by all the sources of noise in the experiment rather than just the analysis procedure itself. Image analysis using six images and four images has been investigated.

Referring to Fig. 3.7, six images were generated from the known fringe order and the isoclinic angle, found from the theory of elasticity combined with the stress-optic law. The fringe order and the isoclinic angle can be obtained by the analysis of six images. By comparing the results obtained with the original photoelastic information, the accuracy of the image analysis procedure can be evaluated.

The comparison of the two sets of results clearly shows that there is no difference between the two sets except in the region with a high fringe gradient, which exceeds the resolution requirement.

Four images were used to simulate the analysis used in the PSIOS system and the same conclusion was drawn.

6.5.2 Image Noise

From previous discussions, the errors from the image analysis procedure itself are negligible. In a real experiment there is always a random noise existing in the images. Any deviation in the phase-stepping system will result in the degeneration of the observed phase-stepped images in the sense that the observed grey level in the images can only represent an approximation of light intensity which should be observed from a properly set-up system, as shown in Fig. 6.18. The possible errors due to misrepresentation of the light intensity in the grey level of the images will be discussed. The effect of the noise in the images on further analysis will be shown through several examples. In order to improve the accuracy of the analysis, several image processing techniques were examined and a Gaussian smoothing filter was found to be suitable to remove the noise without losing the phase-stepping information in the images. The following analysis is using the conventional polariscope as an example, but the principle for the PSIOS system is the same.

6.5.3 Analytical Analysis on the Error due to the Image Noise

The image noise appears as a random variation of the image grey level in a small range around the value which the grey level should be. Mathematically, this could be expressed as,

$$I(x, y) = I_E(x, y) + \eta(x, y) \quad (6-30)$$

where $I(x, y)$ is an observed image, $I_E(x, y)$ is the image that should be observed according to the physical law and $\eta(x, y)$ is the noise which should have two properties,

1. At every pair of co-ordinates (x, y) the noise is uncorrelated;
2. The noise has a zero average value.

So Eq. 6-30 could be rewritten as [Papoulis, 1965],

$$I(x, y) = I_E(x, y) + \Delta \quad (6-31)$$

Obviously, all six images could be affected by noise, and the error in the isochromatic and the isoclinic data due to the combined noise in all images could be expressed as,

$$d\theta = \frac{\partial\theta}{\partial\alpha_3} di_3 + \frac{\partial\theta}{\partial\alpha_4} di_4 + \frac{\partial\theta}{\partial\alpha_5} di_5 + \frac{\partial\theta}{\partial\alpha_6} di_6 = \sum_{n=3}^6 \frac{\partial\theta}{\partial\alpha_n} di_n, \quad (6-32)$$

$$d\alpha = \frac{\partial\alpha}{\partial\alpha_1} di_1 + \frac{\partial\alpha}{\partial\alpha_2} di_2 + \dots + \frac{\partial\alpha}{\partial\alpha_6} di_6 = \sum_{n=1}^6 \frac{\partial\alpha}{\partial\alpha_n} di_n, \quad (6-33)$$

6.5.4 Example of the Noise Effect

In the previous sections, the possible error due to the image noise was presented mathematically, which was on a single arbitrary image position. Here an example is presented in Fig. 6.19. A blade in air compressive engine with reflection coating was captured by the PSIOS system. Four raw images and the light intensity profiles along an arbitrary line are shown in Fig. 6.19a. Due to poor quality of the raw image, the wrapping error is clearly observed in Fig. 6.19b.

6.5.5 Image Pre-processing [Gonzalez *et. al.*, 1993]

The principal objective of pre-processing is to process the images so that the results are more suitable than the original images for further analysis. The noise in images

observed from phase-stepping system could be considered as a zero mean white noise which was characterised in **Section 6.5.3**

6.5.5.1 Image Average

Because of the character of the noise in the images, the averaging of the large samplings over the same view (same phase-step) would be an efficient way to reduce the noise. If $i_p(x, y)$ are a set of noisy images, and $\bar{i}(x, y)$ is formed by averaging M different noisy images, then the procedure could be expressed as,

$$\bar{i}(x, y) = \frac{1}{M} \sum_{p=1}^M i_p(x, y) \quad (6-34)$$

and it follows that,

$$E\{\bar{i}(x, y)\} = I_E(x, y) \quad (6-35)$$

and

$$\sigma^2[\bar{i}(x, y)] = \frac{1}{M} \sigma^2[i(x, y)] \quad (6-36)$$

where $E\{\bar{i}(x, y)\}$ denotes the expected value of \bar{i} and $\sigma^2[\bar{i}(x, y)]$ and $\sigma^2[i(x, y)]$ are the variances of \bar{i} and i , all at co-ordinates (x, y) .

The noise in the images, however, was not only due to the imaging part of the system, and hence the averaging procedure does not produce a satisfactory result.

6.5.5.2 Gaussian Smoothing Filter

A phase-stepped image appears to be a smoothly changing fringe image that means the grey level at any point should not have a sudden change from its neighbour's grey level. The observed image gives the approximation of the light intensity at every point. One can get an accurate estimation at each position by taking its neighbour's grey level into account. A Gaussian filter is a weighted averaging procedure over a certain area around every pixel. The implementation of a Gaussian filter is a convolution procedure using a Gaussian Mask [Nalwa 1993].

The Gaussian convolution mask is a circularly symmetrical (or isotropic) mask, such that any cross-section through its centre yields a weight profile that has the form of a Gaussian or normal curve. The Gaussian convolution mask has a formula as,

$$G(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (6-37)$$

where σ determines the width of the filter, and the factor $\frac{1}{2\pi\sigma^2}$ ensures that the integral of $G(x, y)$ is unity. The convolve of the Gaussian mask with a local signal will give the weighted averaging value over the point (x, y) . Since the noise has a higher spatial frequency compared with the local signal, which should be a smoothly changing signal and then has a lower spatial frequency, Gaussian smoothing filter will remove high spatial frequencies without distorting lower spatial frequencies. Varying the width of the Gaussian with which a signal is smoothed varies the level of discernible detail in the smoothed signal: the larger the width of the Gaussian, the greater the smoothing of the original signal.

6.5.6 The Effect of Gaussian Filter

Wrapping errors sometimes occurred in real experiments due to the influence of the noise. In order to express the effect of the Gaussian filter, the simulation of image 1 for the disc without any noise, with noise and noise reduction by Gaussian filter is compared in the left column of Fig. 6.20, and their profile along an arbitrary line are shown in the right column. The image without noise was generated in the same way as described in Section 3.4.2. Random noise can be produced within the defined range by an Excel spreadsheet (Excel 5.0, Microsoft Corporation). For example, *RAND(0, 1)* command will return random values between (0, 1). An array with 256×256 data can be produced and added to the image 1 to produce a noisy image. In Fig. 6.20, the range was defined as (0, 25) so 10% noise compared with the raw image (250 grey levels available in this image) was added. $\sigma = 3$ was used to smooth the noisy image by a Gaussian filter. The analysis results for the isochromatic fringe order and isoclinic angle for the comparison of three sets are shown in Fig. 6.21. The use of the Gaussian filter has removed much of the influence of the noise and overcome wrapping errors that the noise brings, in the meantime there is some loss of information in the region with high fringe gradient.

The effect of the Gaussian filter was also examined by the experimental data. The raw images of a blade showed in Fig. 6.19a were smoothed by the Gaussian filter and the filter size was chosen as $\sigma = 5$. The correspondent images after processing are shown in the top panel of Fig. 6.22a. The profiles along the same line are shown in the bottom panel. The isochromatic and isoclinic data are shown in Fig. 6.22b. The comparison between Fig. 6.22 and Fig. 6.19 clearly shows significant

improvements of the image quality by the Gaussian Filter. The wrapping errors in the analysis results was also removed.

The implementation of the Gaussian filter can significantly improve the image quality and further improve the analysis results. However, the selection of the width of the filter is the key point in practice. The larger the width of the Gaussian, the greater the smoothing of the original signal. However, the more information is lost in the high fringes gradient. Conversely, for a smaller width of the Gaussian filter, less information is lost in the region of high fringe gradients but less noise is removed. The selection of the width should take these factors into account. Also, the magnitude of the noise and the frequency of the fringe should be considered. Further study is needed to improve the system accuracy.

6.5.7 Discussion and Conclusions

The six and four phase-stepped images, which correspond to the conventional polariscope and the PSIOS system, were generated from the known isochromatic fringe order and isoclinic angle data. The analysis results show there is no difference compared to the known parameters. So no errors are introduced during the image analysis stage.

However, the analysis results may express some errors due to the accumulation of the errors in other stages. A Gaussian smoothing filter was examined and found suitable to remove the noise without losing the phase-stepping information in the images.

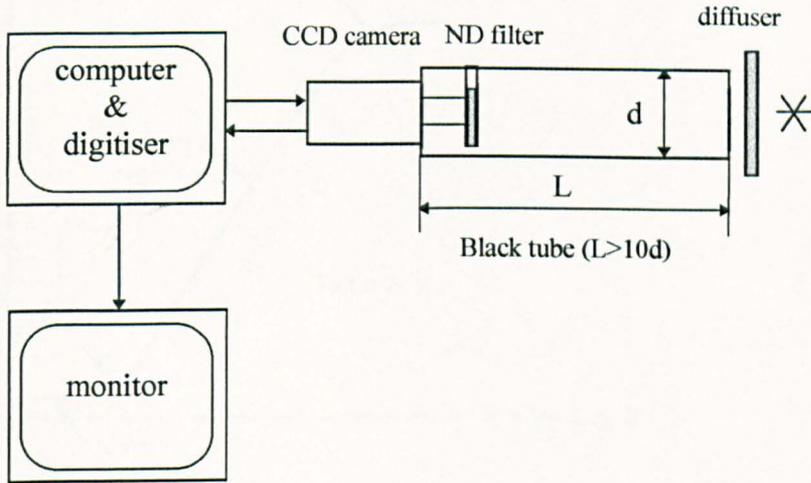


Fig. 6.1 The experimental configuration for testing the transmission ratio of the CCD camera.

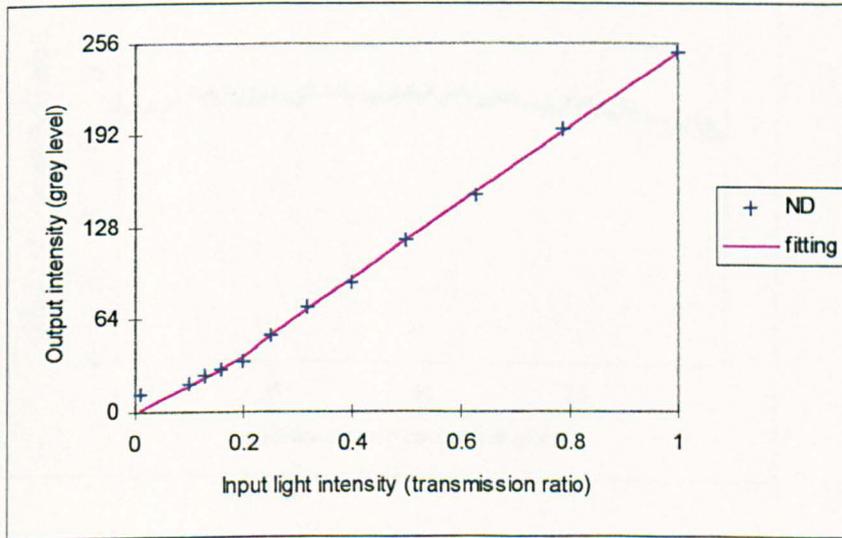


Fig. 6.2 The linearity of the CCD camera tested by Neutral Density filters

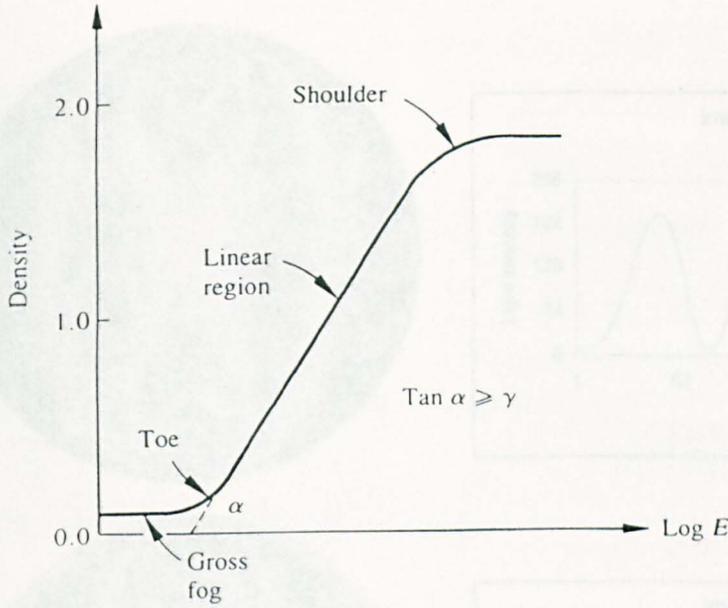


Fig. 6.3 A typical H&D curve (reproduced from Gonzalez, 1993)

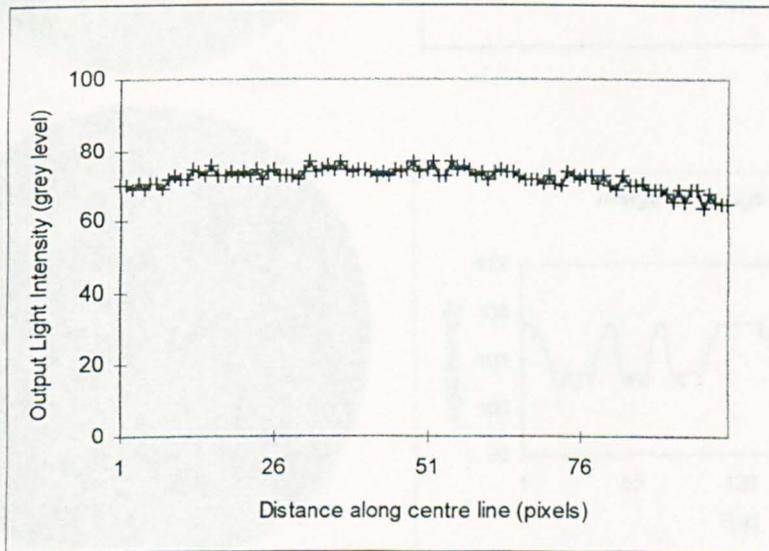


Fig. 6.4 Distribution of the light source along an arbitrary line

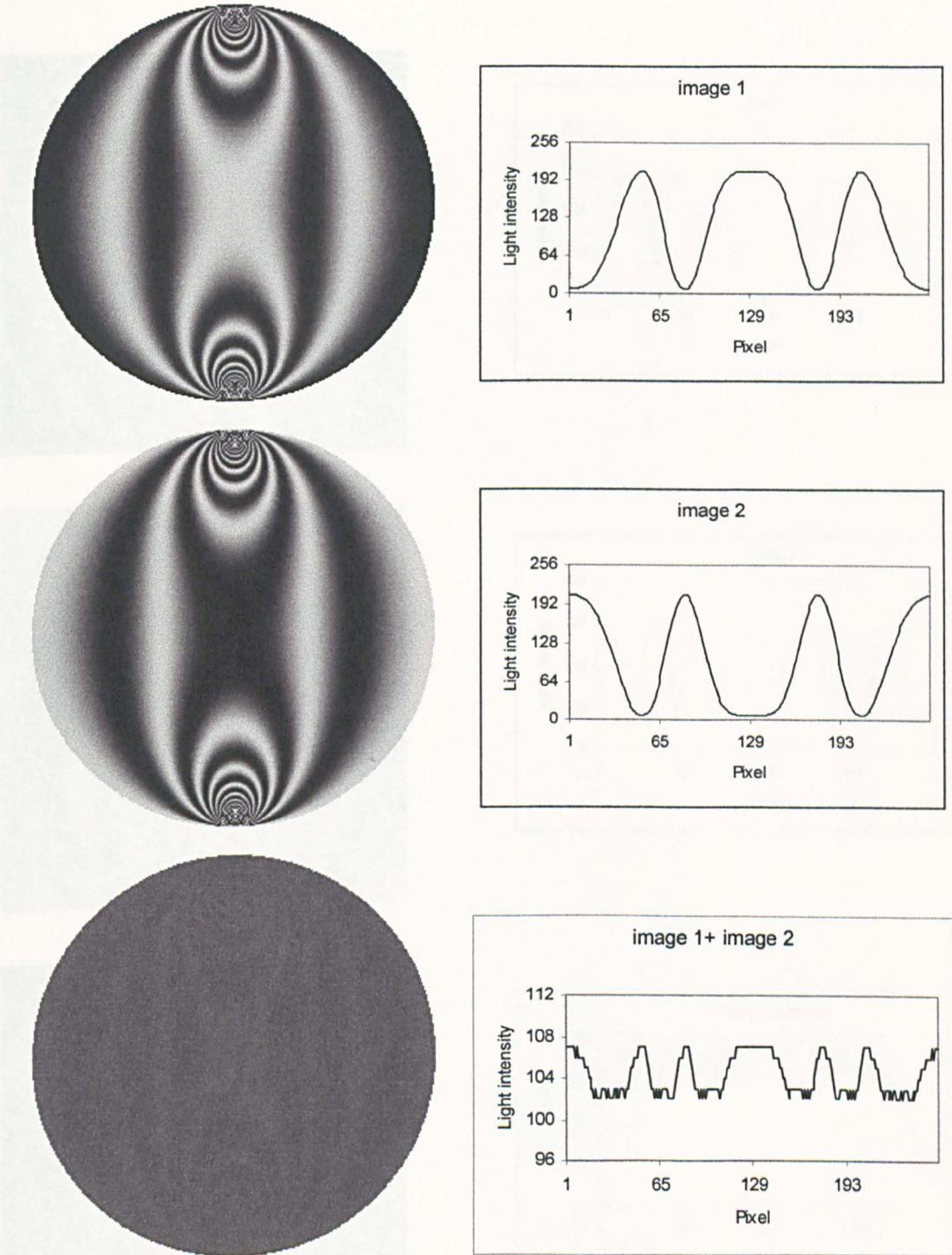


Fig. 6.5 The raw image 1 and image 2 and their sum for the disc model obtained from simulation due to nonlinearity of the CCD camera are illustrated in the left panel. The right panel shows their profiles along the line of Y=127.

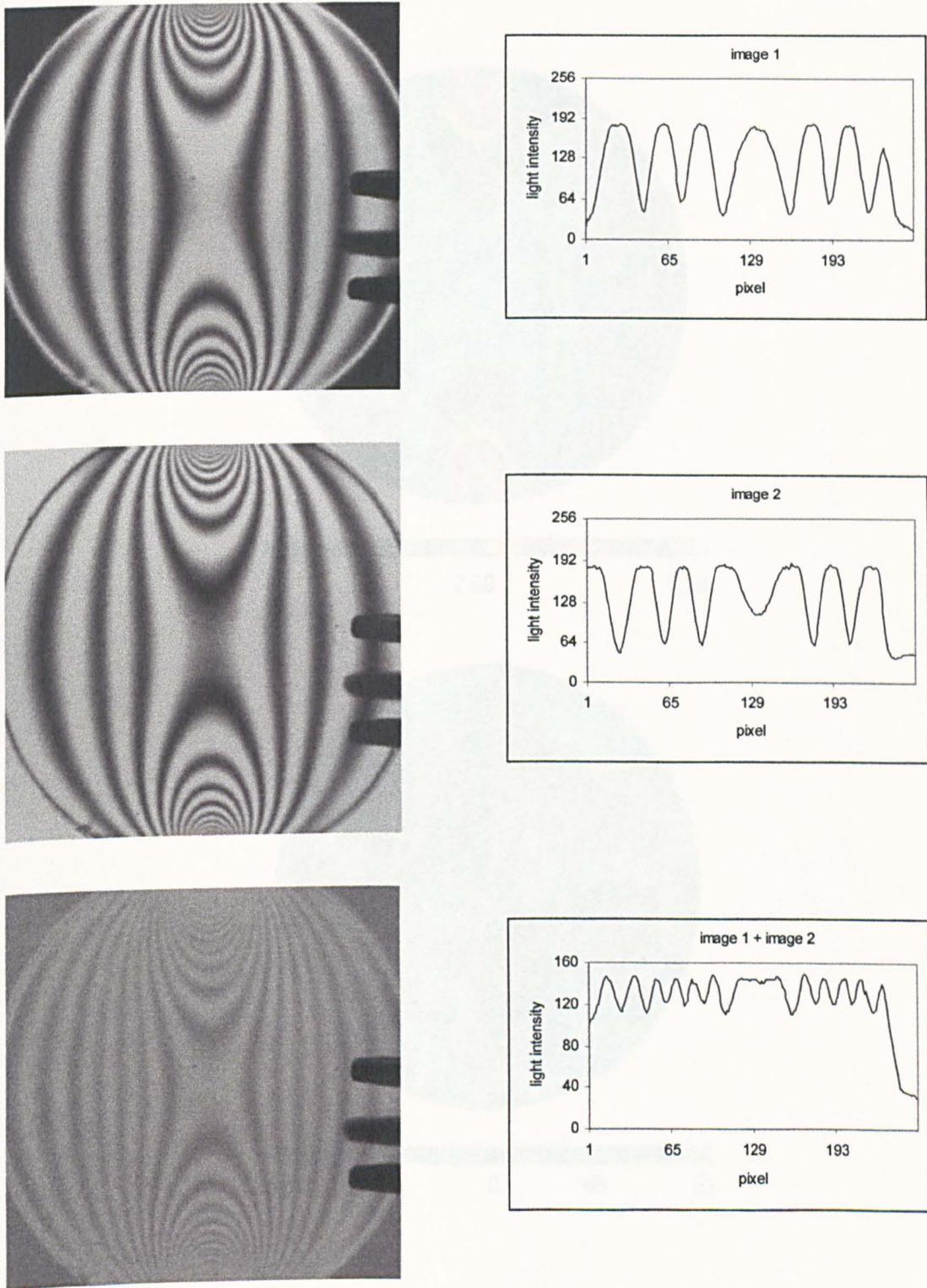


Fig. 6.6 The raw image 1 and image 2 and their sum for the disc from experiments are illustrated in the left panel. The right panel shows their profiles along the line of $Y=127$.

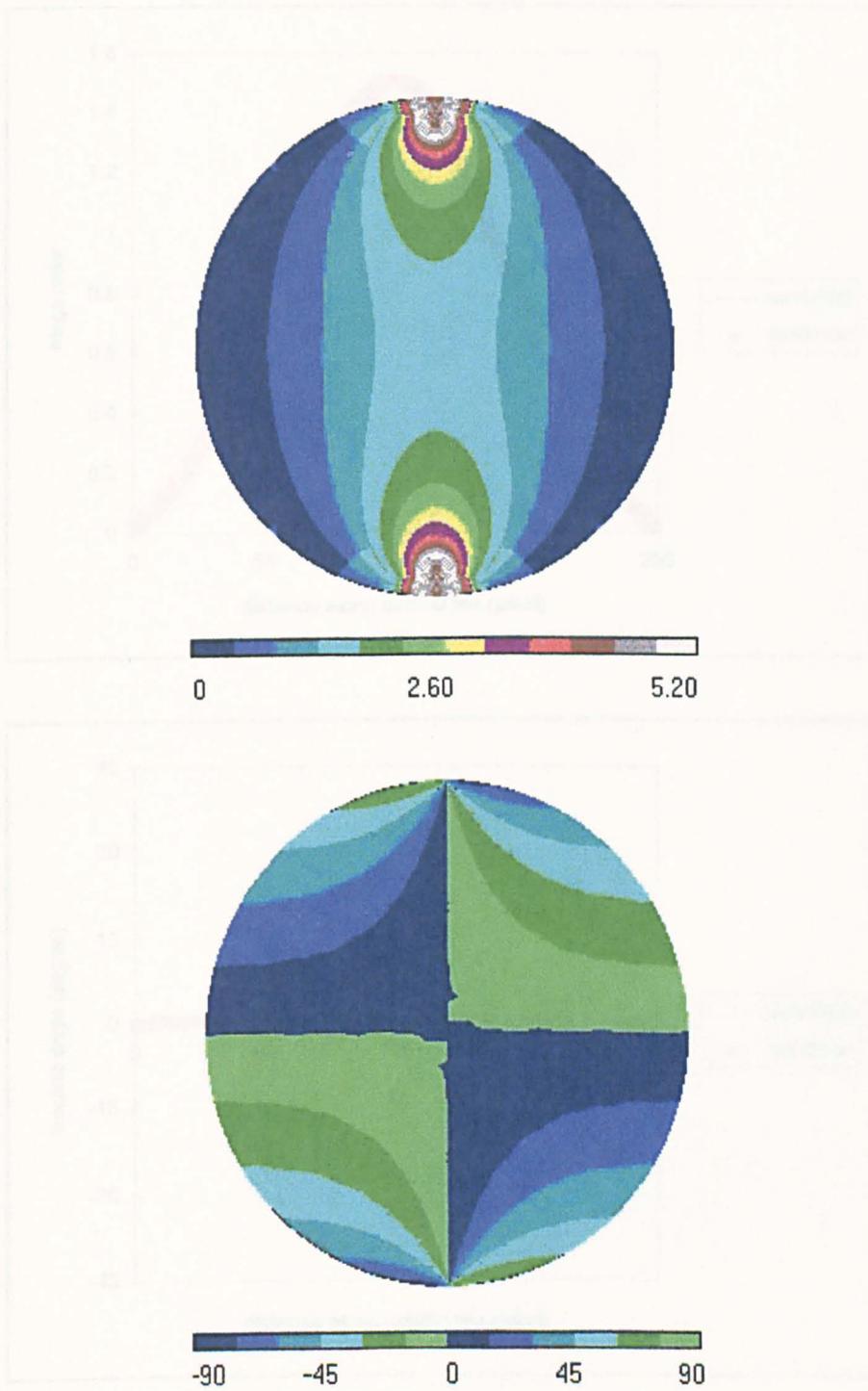


Fig. 6.7 The Modelling results for the isochromatic fringe order and the isoclinic angle due to nonlinearity of the CCD camera

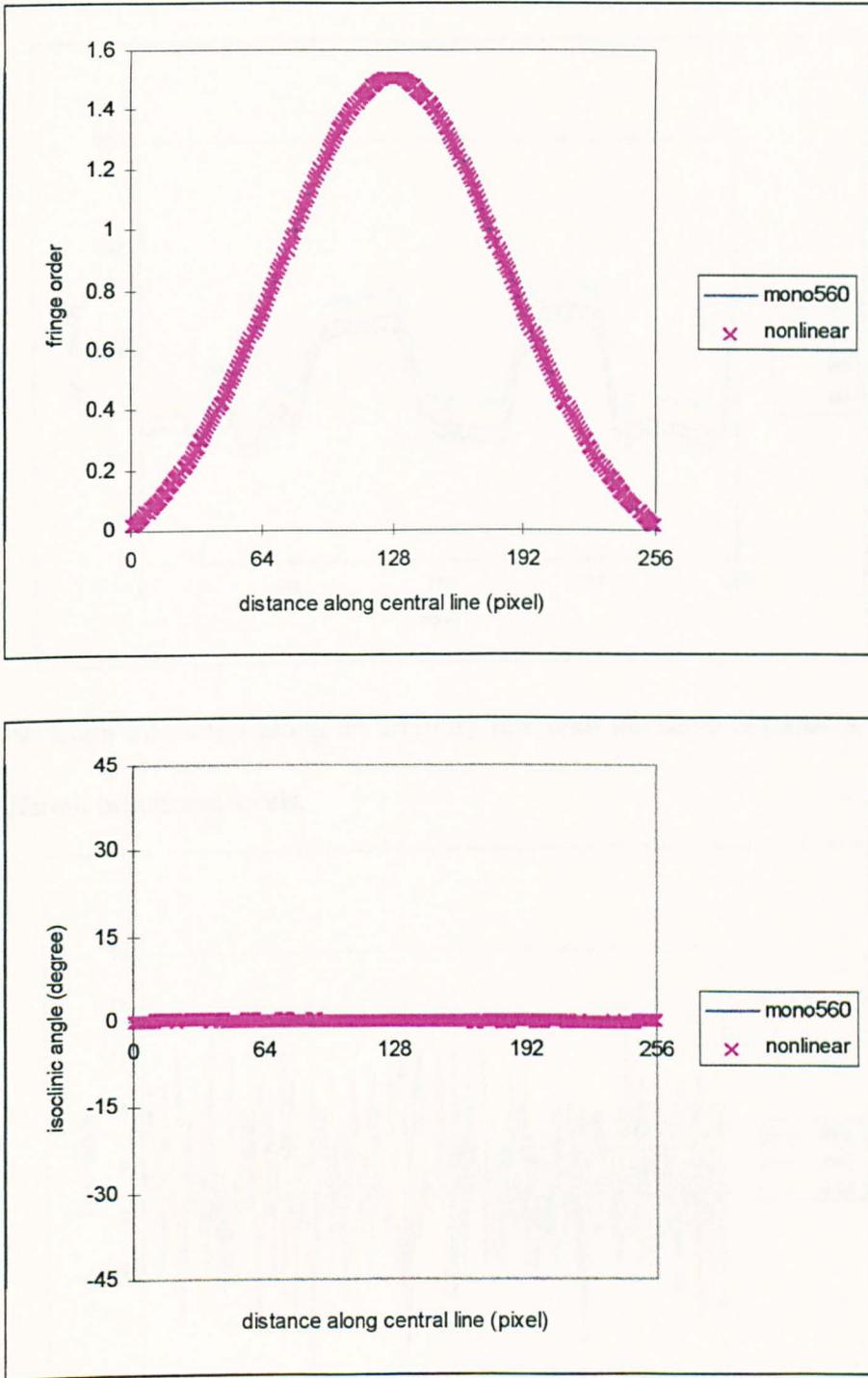


Fig. 6.8 The comparison of the modelling results on the isochromatic and isoclinic data due to the nonlinearity of the camera. The data are shown along the central line of the circular disc

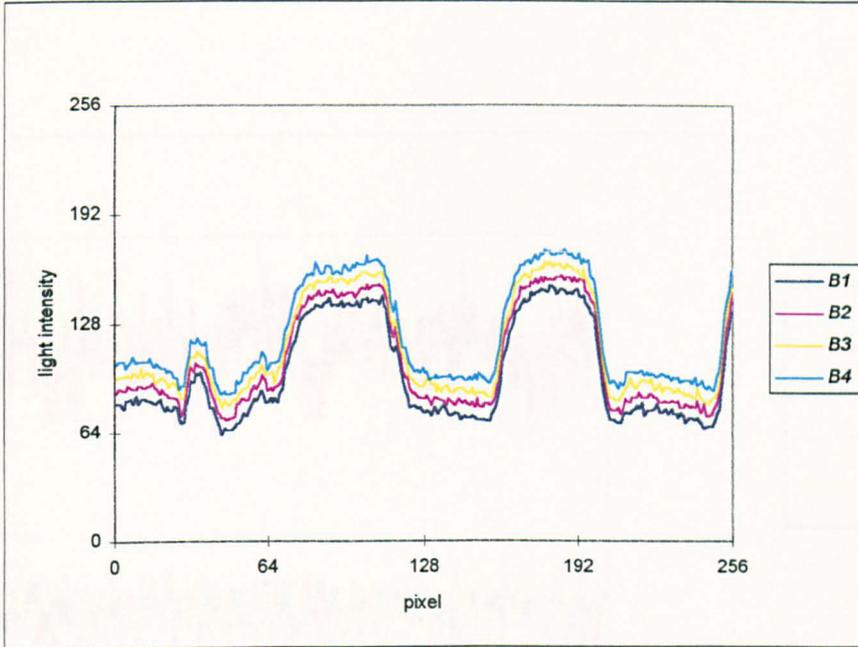


Fig. 6.9 Light intensities along an arbitrary line with the same contrast level (127) and different brightness levels.

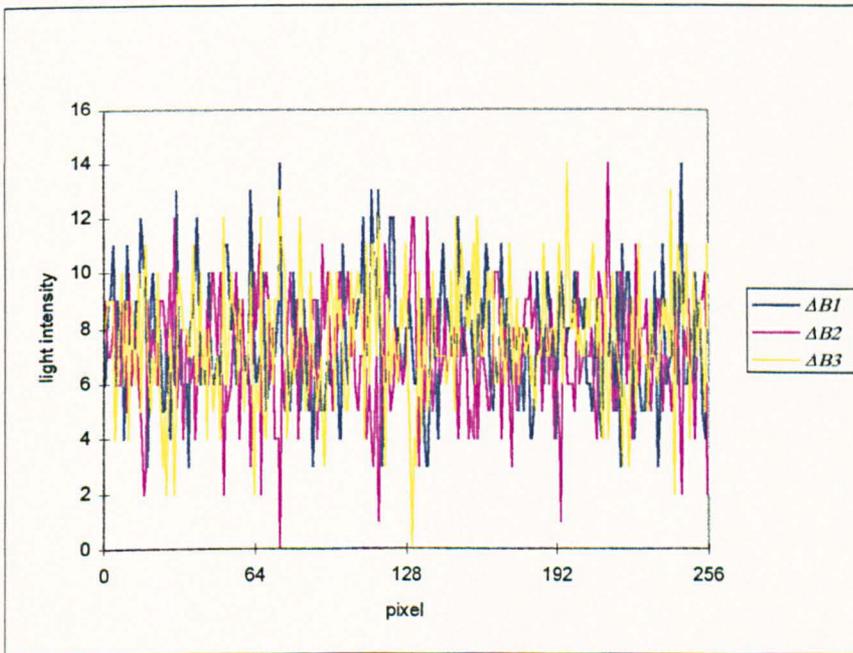


Fig. 6.10 The difference of the light intensity for an increment of 5 brightness levels

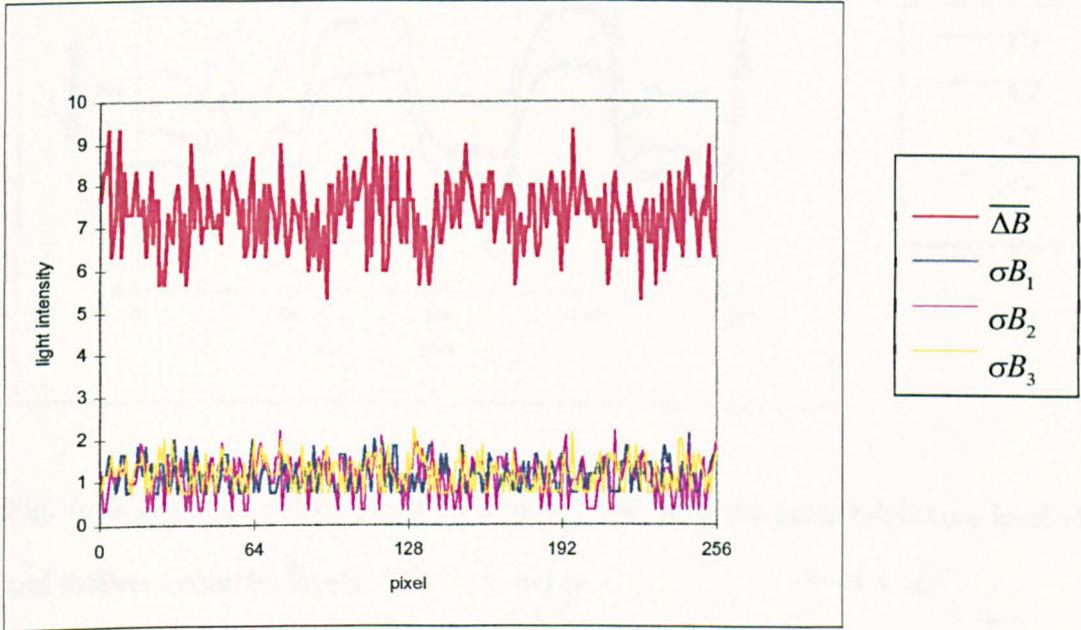


Fig. 6.11 The average value and the variance for each of 5 increments in brightness level

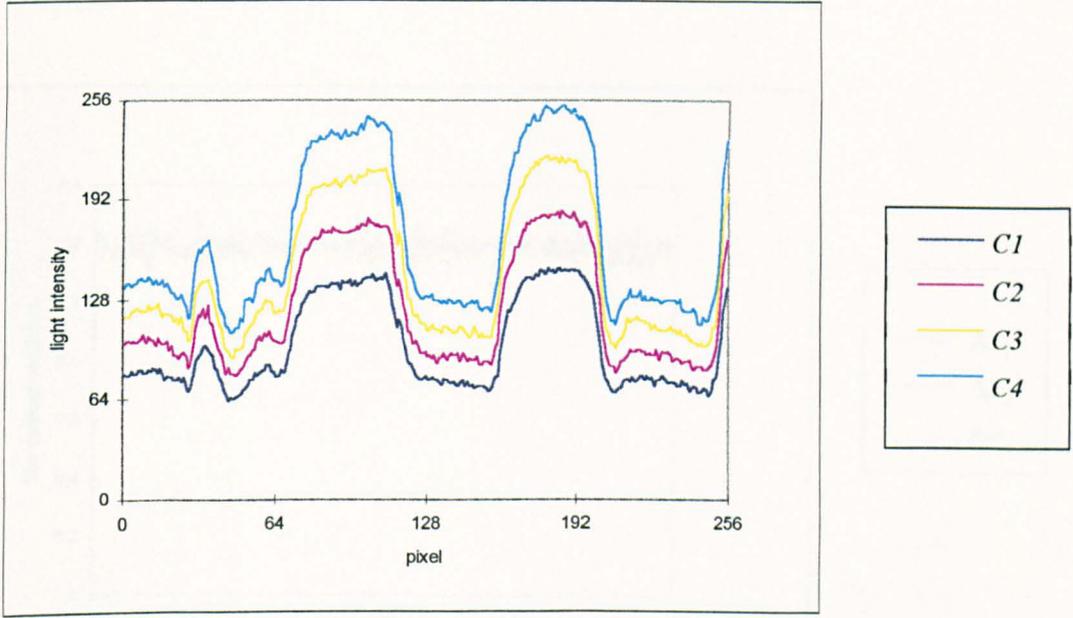


Fig. 6.12 Light intensities along an arbitrary line with the same brightness level (127) and different contrast levels.

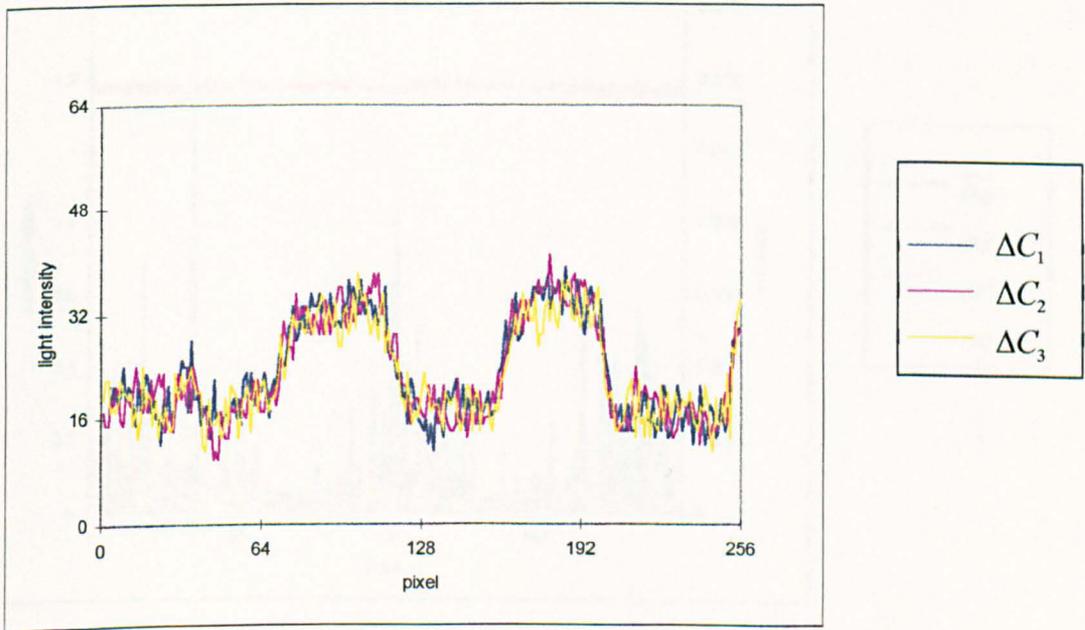


Fig. 6.13 The difference of the light intensity for an increment of 5 contrast levels

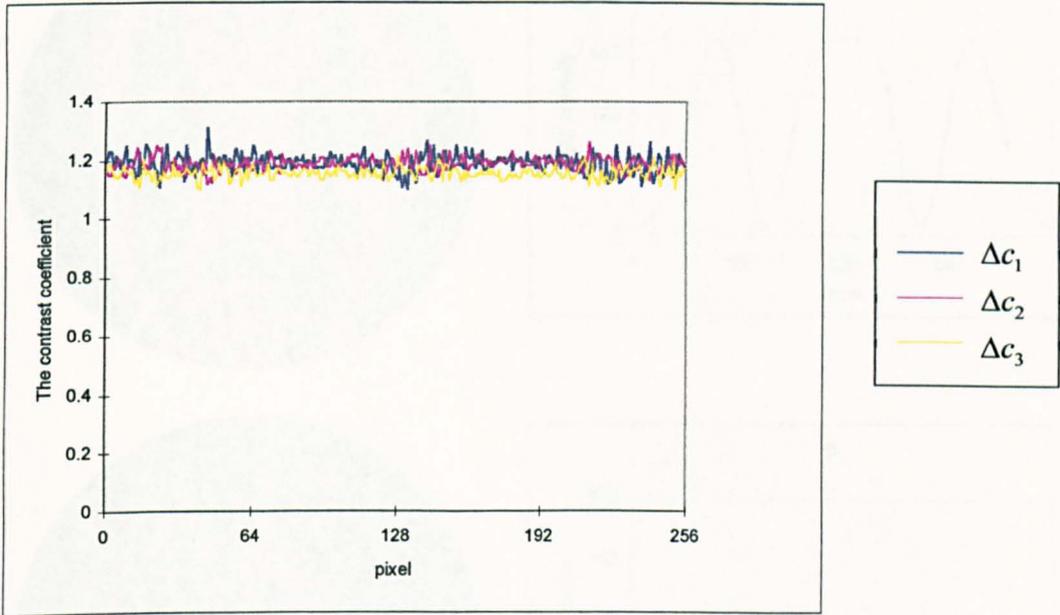


Fig. 6.14 The contrast coefficient for each increment of 5 levels

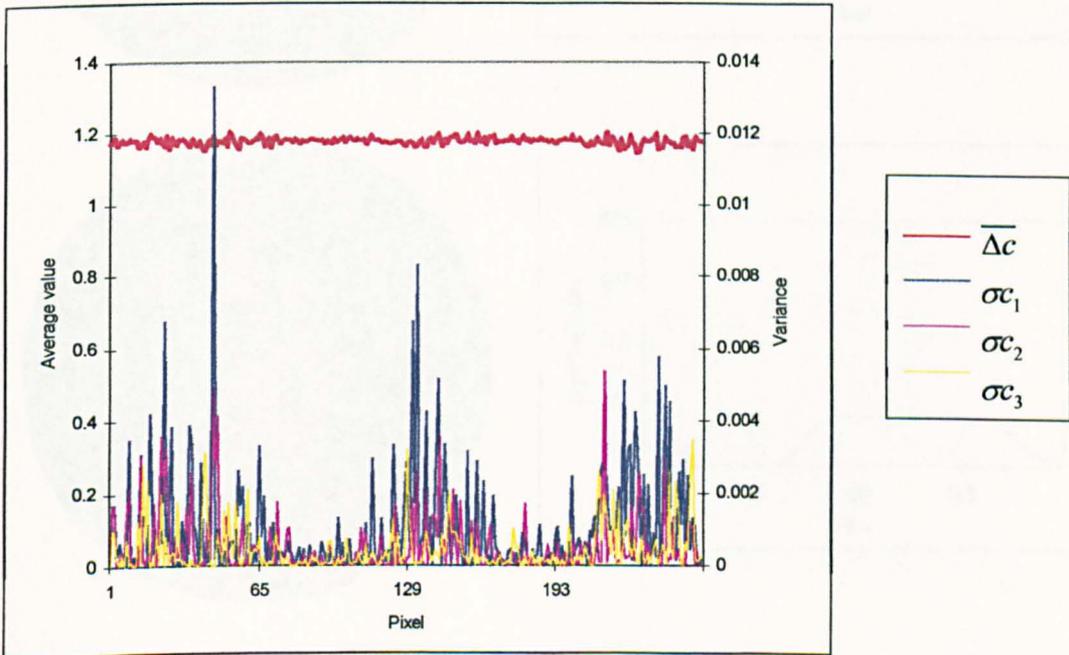


Fig. 6.15 The average value and the variance for each of 5 increments in contrast levels

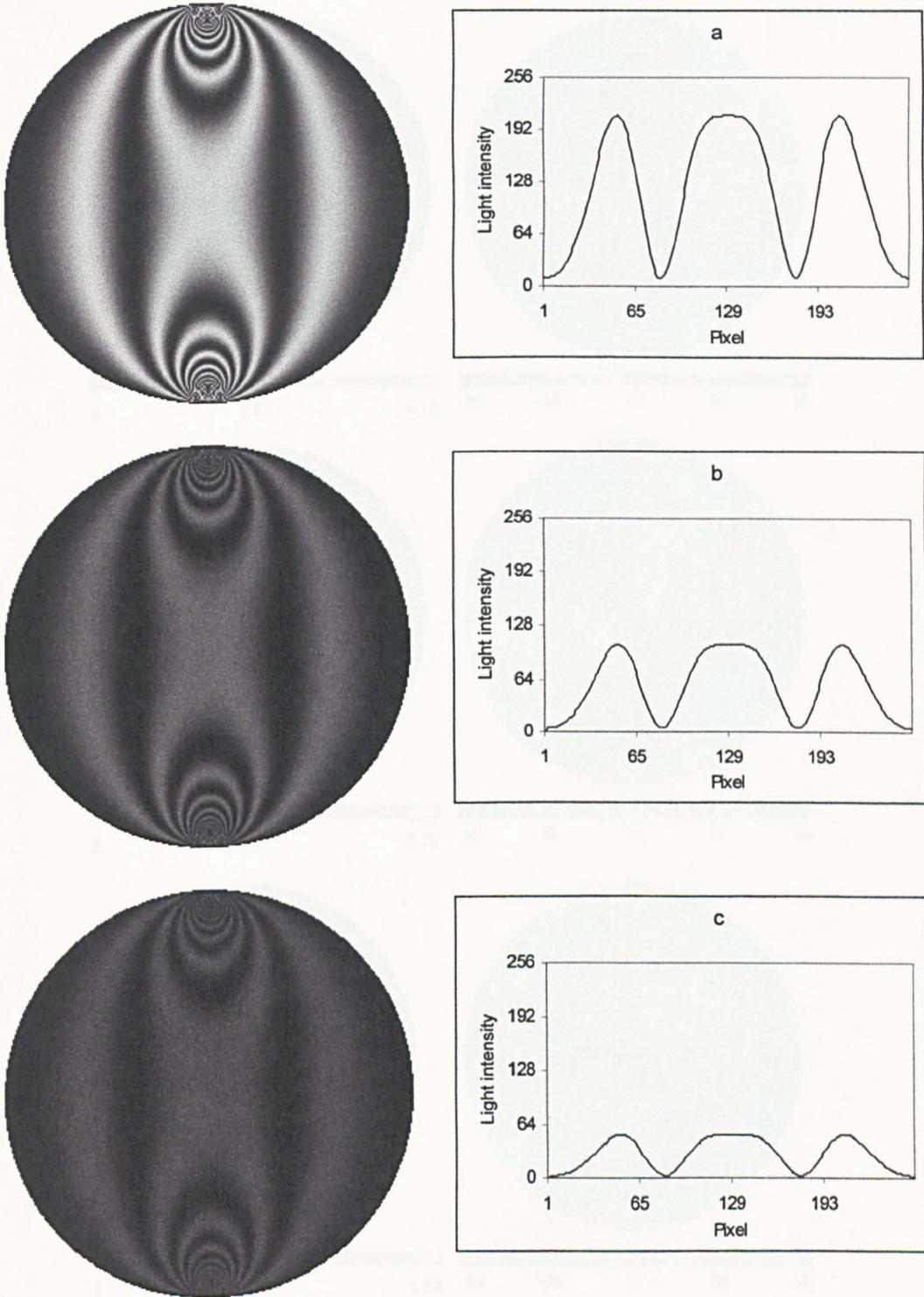


Fig. 6.16a The modelling of a circular disc with different light intensities. Raw images (image 1) are illustrated in left panel and their profiles along the line $Y=127$ are shown in right panel.

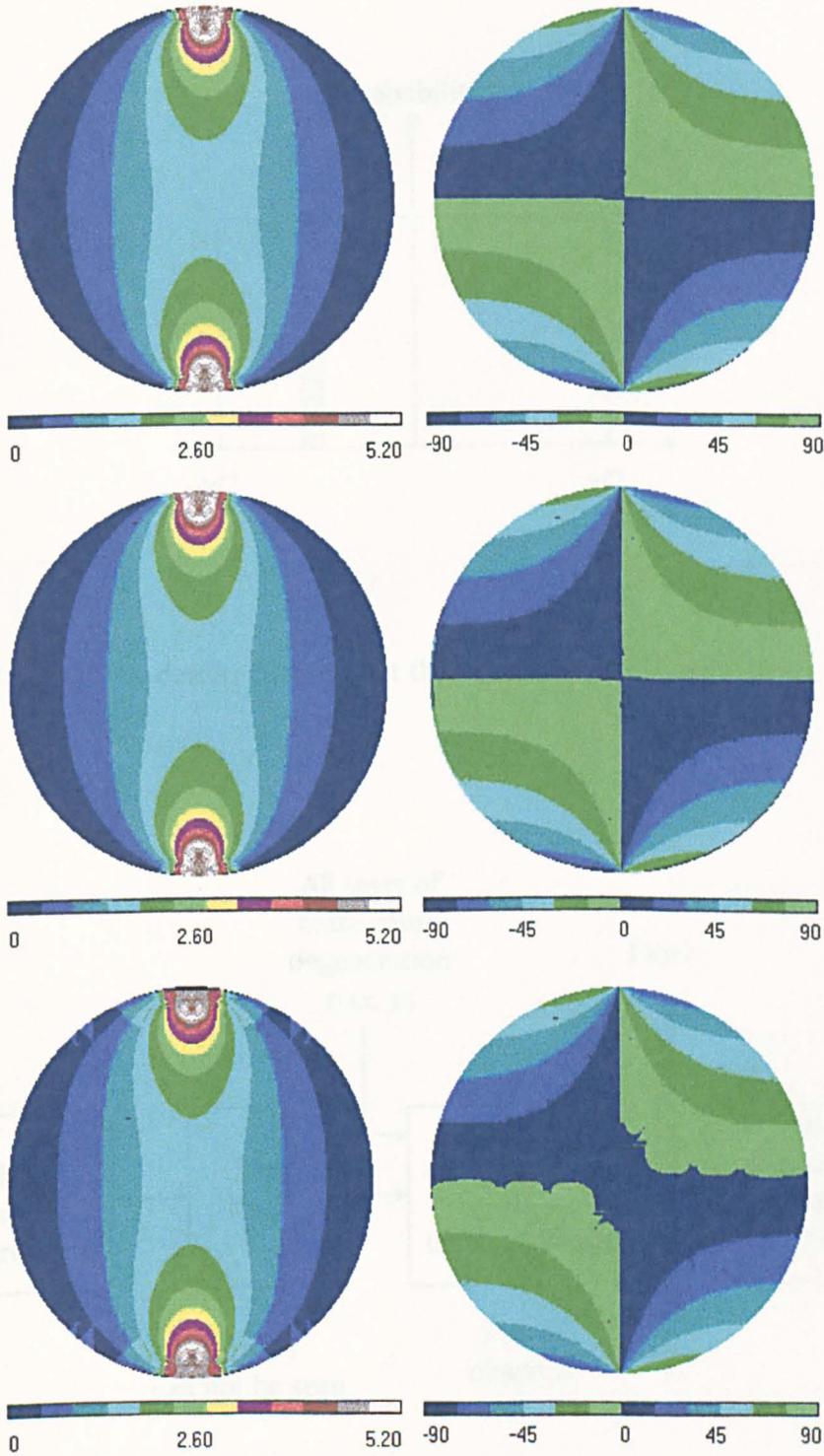


Fig. 6.16b The isochromatic fringe order (left panel) and the isoclinic angle (right panel) for a modelling disc with different light intensities.

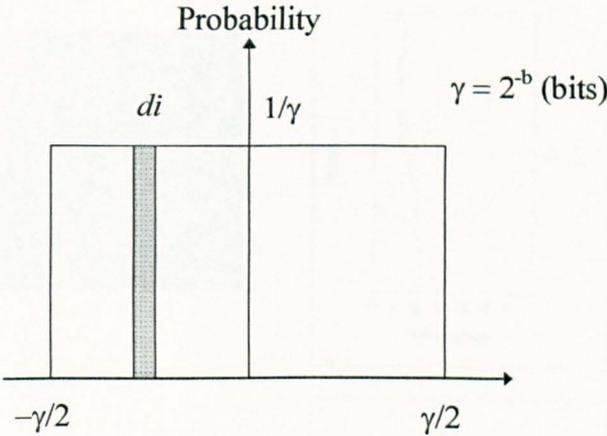


Fig. 6.17 Probability density function for the digitisation.

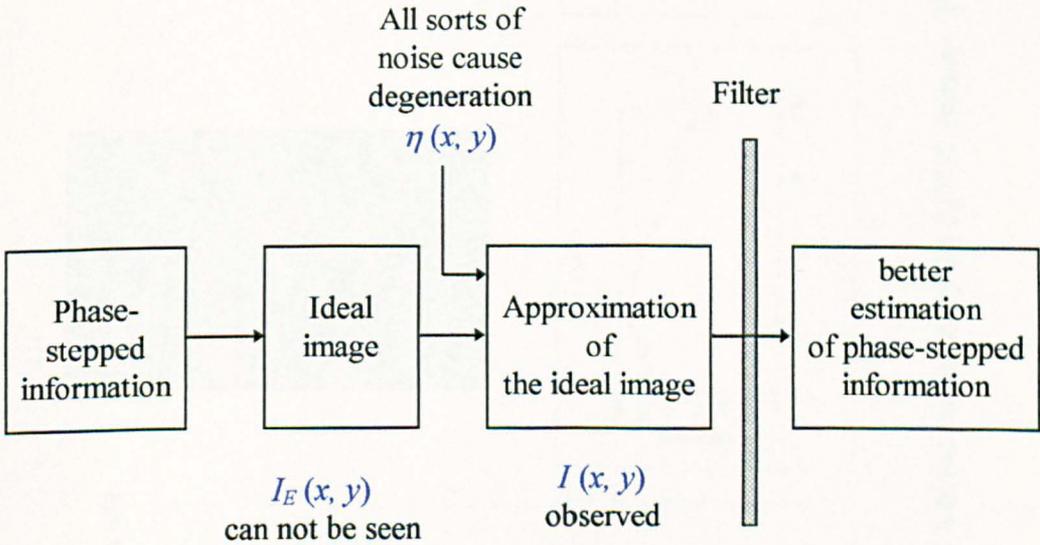


Fig. 6.18 Flow diagram shows the effect of the Gaussian filter in the phase-stepping system.

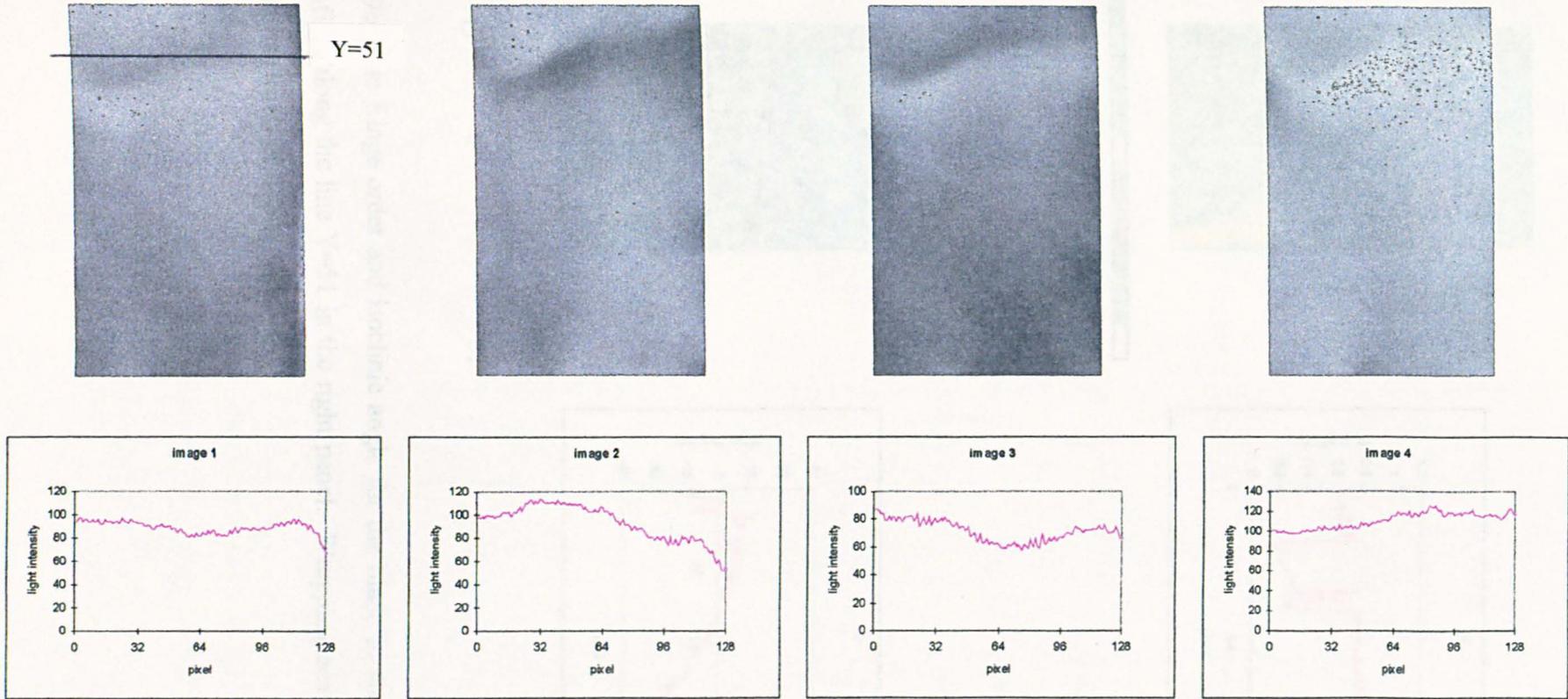


Fig. 6.19a The raw images for a blade captured by the PSIOS system. The bottom panel shows the profile along the line $Y=51$.

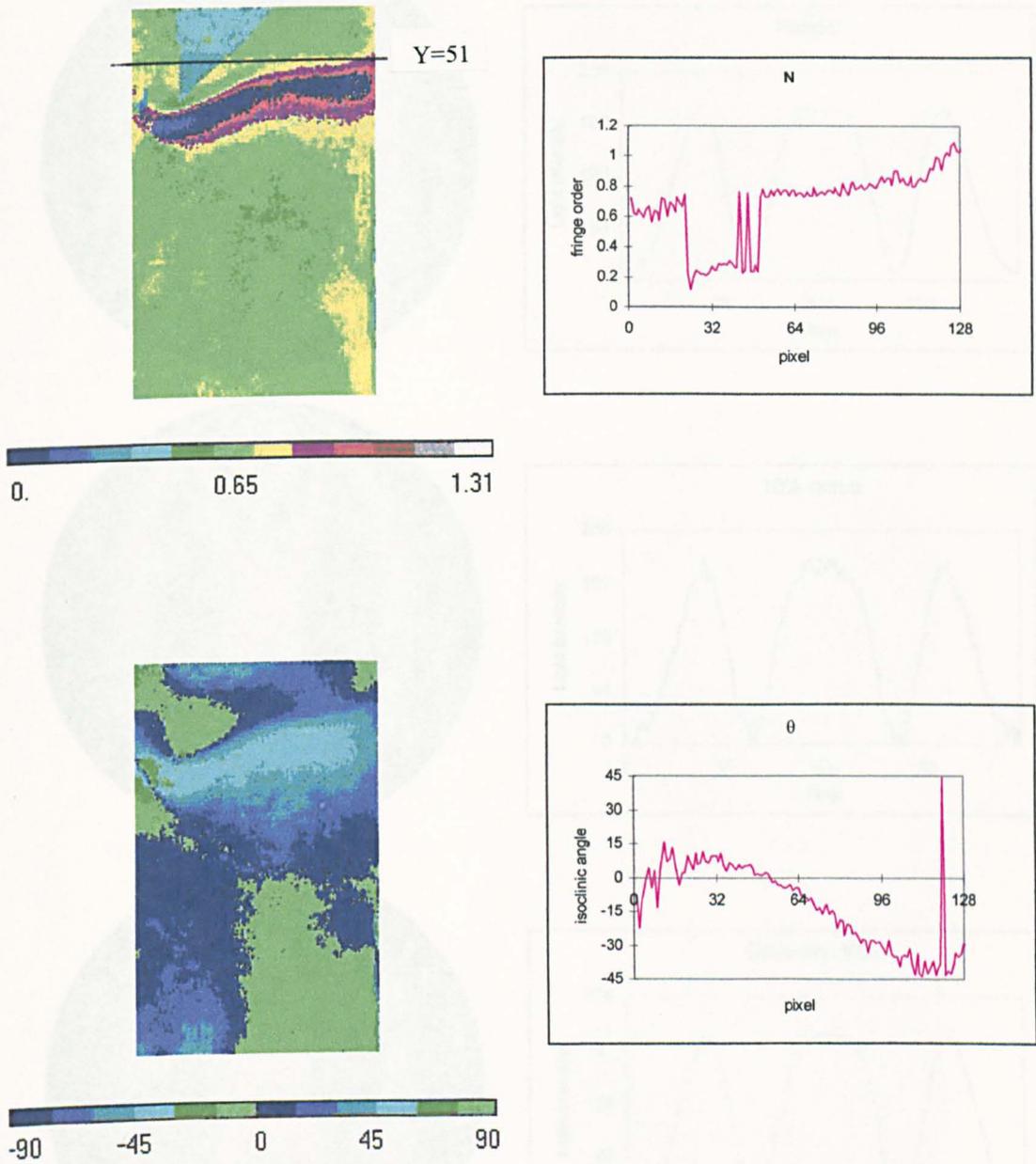


Fig. 6.19b The fringe order and isoclinic angle for the blade in the left panel and their profiles along the line $Y=51$ in the right panel. Wrapping error was observed due to various kinds of noise.

Fig. 6.20 Effect of the Gaussian filtering. Perfect data, left with 10% noise and data after processing, are illustrated in the left panel. Their light intensities along the line of $Y=127$ are shown in the right panel.

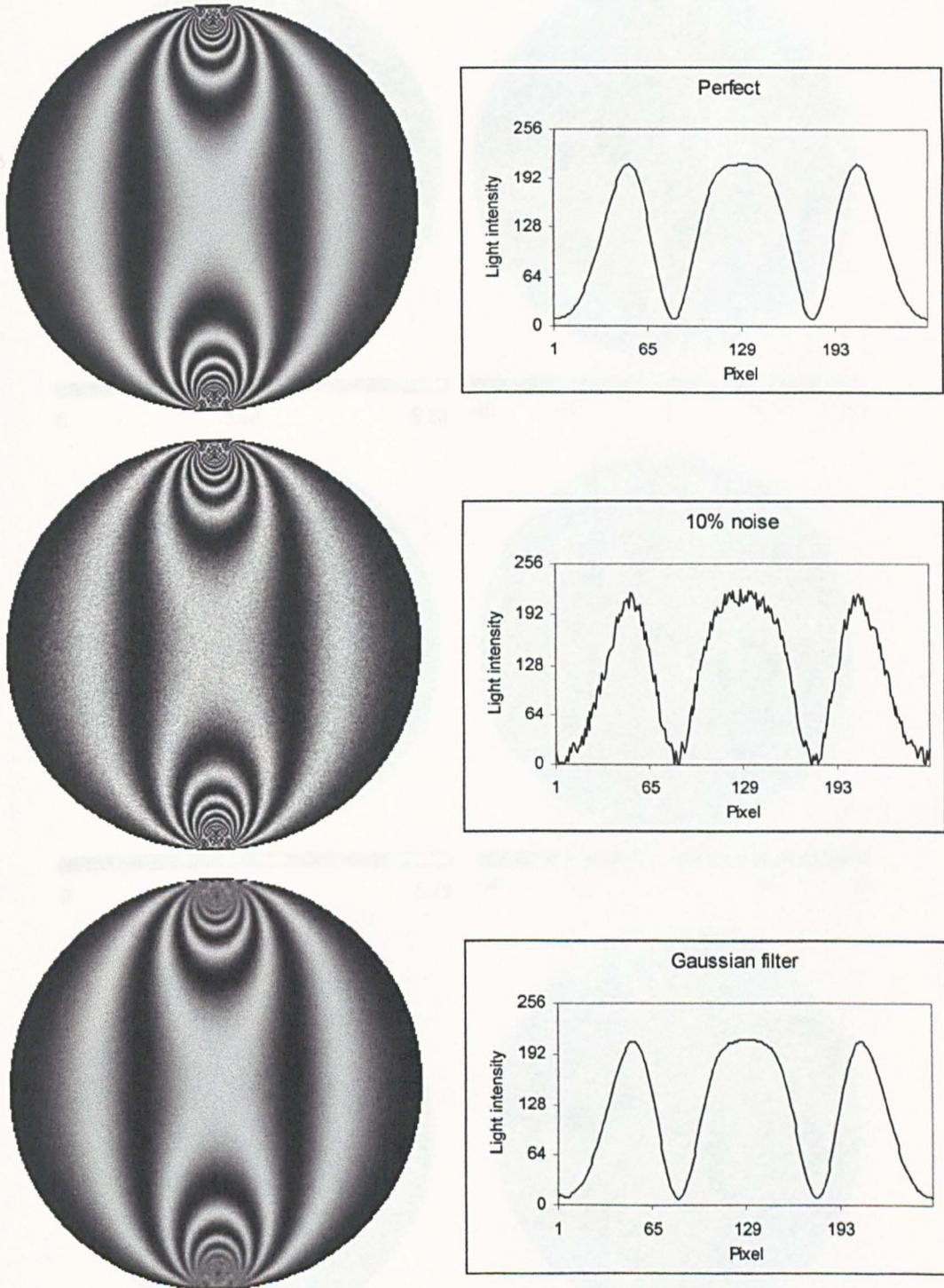


Fig. 6.20 Effect of the Gaussian filtering. Perfect data, data with 10% noise and data after processing are illustrated in the left panel. Their light intensities along the line of $Y=127$ are shown in the right panel.

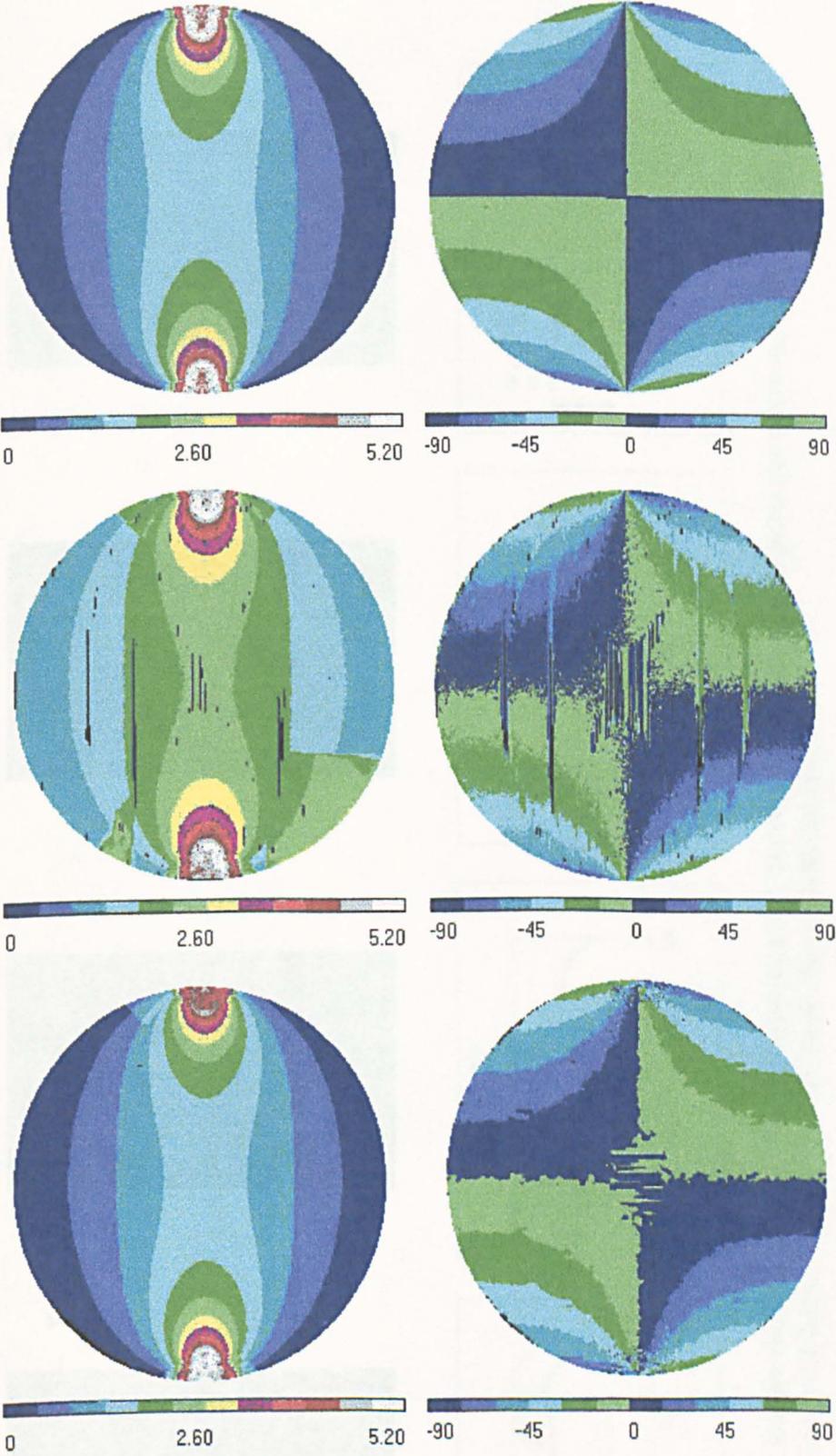


Fig. 6.21 The isochromatic fringe order (left) and isoclinic angle (right) from perfect data (top), data with 10% noise (middle) and Gaussian filtered noisy data (bottom).

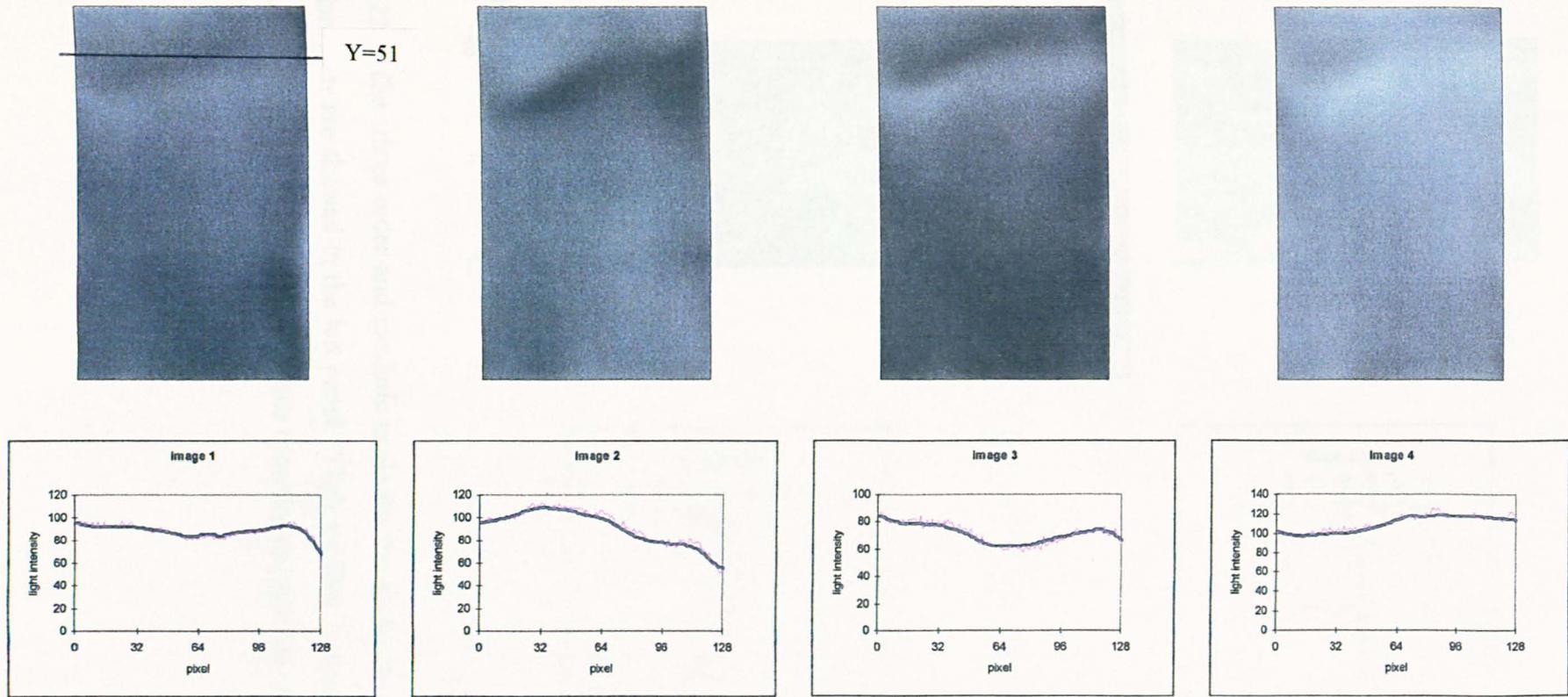


Fig. 6.22a The raw images showed in Fig. 6.19a was smoothed by Gaussian Filter. The bottom panel shows the comparison of the light intensity along the line $Y=51$ before smoothing (cyan) and after smoothing (dark blue).

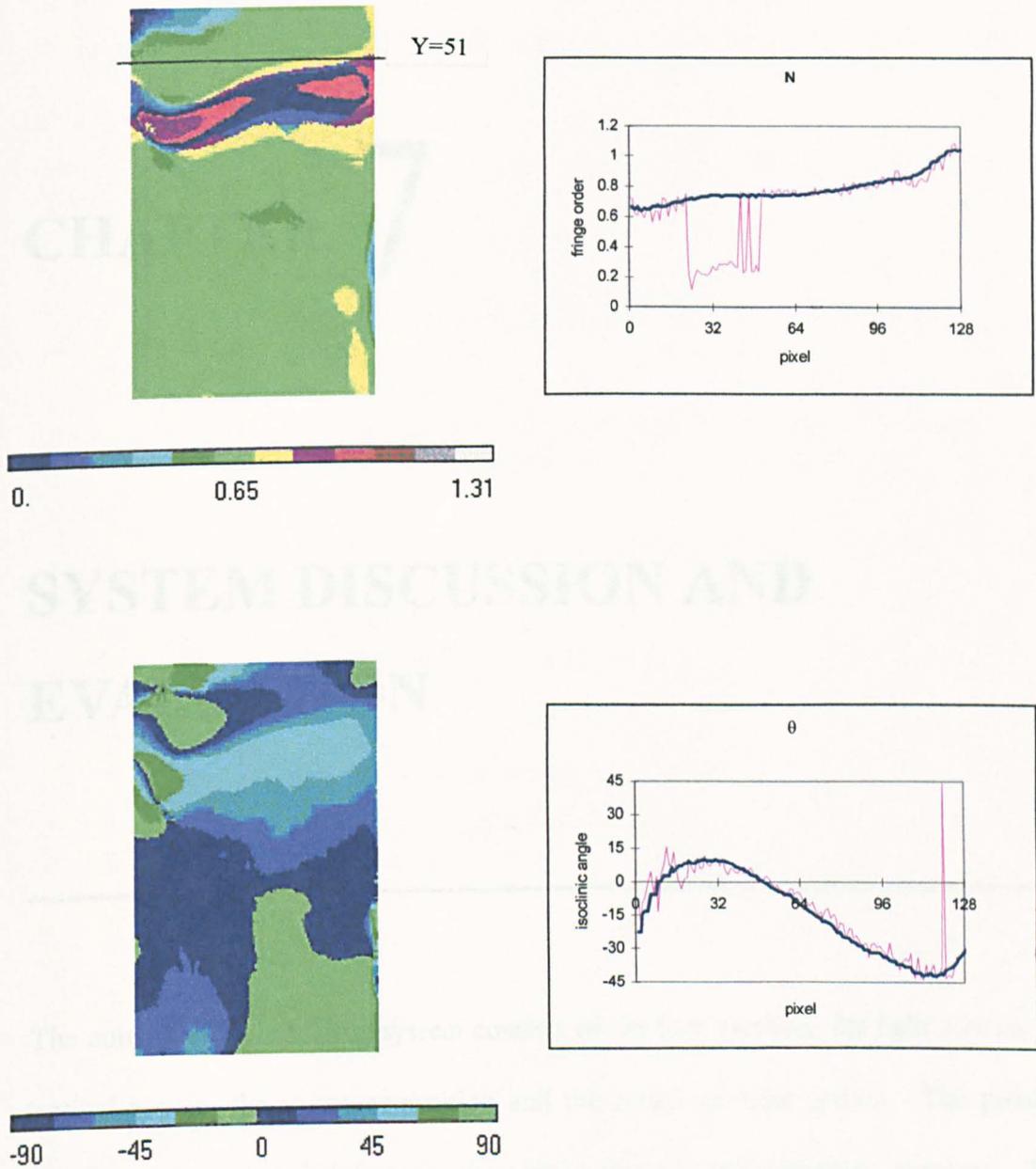


Fig. 6.22b The fringe order and isoclinic angle for the blade after processing by a Gaussian filter are showed in the left panel. Their profiles along the line Y=51 are showed in blue line in the right panel, while cyan line are profiles before processing.

CHAPTER 7

SYSTEM DISCUSSION AND EVALUATION

The automated photoelastic system consists of the four sections: the light source, the optical system, the image acquisition and the image analysis system. The possible error sources and their influence were investigated in the preceding chapters. This chapter will review all the errors and their contributions to the whole system and an evaluation of the accuracy of whole system will also be given.

7.1 Overview of Error Sources in Phase-stepping Systems

All the possible errors and their sources in the phase-stepping systems are listed in Table 7.1.

<p>Optical System</p> <p><i>Conventional Polariscopes</i></p> <ol style="list-style-type: none"> 1. Mismatch of the quarter-wave plates 2. Linear alignment 3. Angular alignment 4. Spatial and temporal variation of the light and environment 5. Material inhomogeneities <p><i>PSIOS</i></p> <ol style="list-style-type: none"> 1. All the errors in conventional polariscopes 2. Spatial calibration of the four images 3. Brightness calibration 4. Incident angle of the light source
<p>Light Source</p> <ol style="list-style-type: none"> 1. Mismatch of the quarter-wave plate 2. Dispersion of birefringence
<p>Image Acquisition</p> <p><i>CCD Camera</i></p> <ol style="list-style-type: none"> 1. The nonlinear relationship between the output and the input of the camera 2. Sensitivity of the individual camera cell 3. Geometric distortion <p><i>Amplifier</i></p> <ol style="list-style-type: none"> 1. Function of the brightness and contrast control 2. Saturation <p><i>A/D Converter</i></p> <ol style="list-style-type: none"> 1. Spatial quantisation 2. Digitisation of the brightness level
<p>Image Analysis</p>

Table 7.1 Overview of error sources in the phase-stepping systems

All of those error sources were analysed in the previous chapters and the results will be reviewed below for a clear presentation because the evaluation of the whole system is based on the knowledge of every individual system section.

7.1.1 Optical System

Two optical systems were investigated, one is the conventional phase-stepping system and the other is PSIOS. The same theoretical foundation are discussed for both systems and all the optical elements used in the conventional system are used in the PSIOS system, so that the errors that exist in the conventional system also exist in the PSIOS system. Some special errors were introduced for the PSIOS system due to the use of beamsplitters.

In the conventional phase-stepping system, a mismatch of the quarter-wave plates was considered to be the largest error source, and generates larger than 6% errors in the isochromatic fringe order and up to 31 degrees error in the isoclinic angle. Either the light source or the quarter-wave plates must be changed to remove this error. In practice both linear and angular alignment of the optical elements are commonly achieved by manual inspection rather than by any specific equipment or procedure, and the error in a carefully conducted experiment could be less than 1.3% in the isochromatic fringe order and 4 degrees in the isoclinic angle. Spatial and temporal variation of the light source and test environment can introduce the errors in the measured light intensities which appear in the brightness of the images and will be discussed later in this chapter. Finally, material inhomogeneities in the optical

elements and the specimen must be avoided when employing automated fringe analysis since they are difficult to identify unless they are totally opaque. All the errors in the optical system for the conventional polariscope are insignificant everywhere for the isochromatic fringe order, but large errors occur in the isoclinic angle at the location of the integer and half order fringes. This is undesirable because it makes replacement of data in this zone more difficult.

In the PSIOS system, four images were acquired simultaneously by four different cameras, so their images must correspond to the same scene spatially and with the same light intensity distribution before each beam arrives at the output quarter-wave plate and analyser. In order to let the same spatial position appear with the same co-ordinates in all images, a calibration of magnification, translation and rotation of the cameras should be taken. To guarantee that the same intensity is viewed by all the cameras requires the appropriate transmission / reflection ratios of the three beamsplitters. When the PSIOS system is used as a reflection polariscope, the incident angle of the light source to the specimen is no longer a right angle. The analysis of the errors in the PSIOS system is from these three aspects.

The spatial calibration for the PSIOS system is very important, because the analysis of the four images without an alignment makes no sense for the results. A target was designed for this purpose. At least two points are needed to identify the magnification in four images; then the first point is used to align the translation of the images and the second point is used to align the rotation of the images. Only after spatial calibration of the four images can the acquisition and the processing be carried on sensibly.

The same light intensities are required in each beam before they arrive at the output optical elements. The nominal transmission : reflection ratios of the three beamsplitters were calculated and used to modify the intensity value through Eqs. 3-23 and 3-24. The beamsplitter is allowed to have $\pm 1\%$ tolerance in the transmission : reflection ratios from the manufacturing process, and this results in about 5 degrees error in the isoclinic angle and no significant influence on the fringe order.

In the PSIOS system the position of the light source and the camera determines an angle between the incident and reflection light beams. Since this angle is less than 10 degrees, the error it introduced to the system is thought to be negligible [Patterson and Wang, 1997b].

7.1.2 Light Source

The optical system requires a matching light source with respect to the quarter-wave plates. The perfect light source should be a monochromatic light, with a wavelength matched for the requirement of the quarter-wave plates. However, the PSIOS instrument was designed for use in an industrial environment and so portability is important and makes high intensity monochromatic sources impractical. In addition, the potential combined use of Spectral Contents Analysis (SCA) with the phase-stepping system or the RGB method requires a broad band of wavelengths. In all of the above cases, a white light source is a better choice.

Several white light sources were available in the author's laboratory, namely S1, O1 and L1, and have been tested. The errors in the isoclinic angle are always large at the position of the integer and half order fringes and have a greater influence on higher fringe orders than lower orders. The errors in the isochromatic fringe order are generally small for the low fringe order case. The errors increased dramatically when higher fringe orders were considered. So the white light sources are limited to low fringe orders if reliable results are to be acquired, typically 1.5 fringes for $\pm 5\%$ errors.

The form of the spectrum of white light directly influences the errors in the isoclinic angle and isochromatic fringe order. The ideal spectrum should have high intensity in the area adjacent to the matching wavelength of the quarter-wave plates and low intensity elsewhere. In addition, the spectrum should be symmetrical about the matching wavelength.

The employment of the medium band filters reduces the errors both in the isochromatic fringe order and in the isoclinic angle significantly. If $\pm 5\%$ accuracy for fringe order were acceptable, then the maximum fringe order that could be measured would be 0.5, 3, and 2 using sources S1, O1 and L1 respectively and increase to 6 and 10 if the 80 and 40 nm filters are combined with the source S1.

7.1.3 Image Acquisition

Three parts were considered to be included in image acquisition, namely the CCD camera, the amplifier and the A/D converter. The fringe pattern was viewed by the

CCD camera with an appropriate lens, transferred and divided via the amplifier to the monitor for display and to the A/D converter for production of a digital signal for computer analysis.

7.1.3.1 CCD Camera

The errors in the sensor can be generally considered to arise from two aspects: one is the intensity and the other is geometry. The geometric distortion caused by the CCD camera can be ignored. The errors in intensity can be divided into two parts: one is the relationship between the output and the input of the camera, and the other is the sensitivity of the individual CCD cells to the same input. The output of the camera is expected to be linear with the input light intensity, however, the experiment using the Neutral Density filters showed that the linear relationship only existed in the medium to high intensity range but not in the extremely low range.

The output curve was fitted by the least square method and its influence on the isoclinic angle and the isochromatic fringe order was simulated for the compression disc. Since this nonlinearity only influences the darkest region, it makes spline even more difficult while having almost no influence on the other region.

It is hard to distinguish the sensitivity of the individual cell to the light intensity because the effect is combined with the influence of the light source, the digitiser and the ND filter. The combined effect is discussed later in this chapter.

7.1.3.2 Amplifier

The output of the camera is transferred via the amplifier to the computer. The brightness and the contrast of the signal can be adjusted by the amplifier. The experimental investigation showed that the control of the brightness is by adding a coefficient to the light intensity; while the adjustment of the contrast is by multiplying the intensity by a coefficient.

From a theoretical view, adjusting the brightness and contrast will not influence the isoclinic angle and the isochromatic fringe order because they will be cancelled by subtraction and division respectively in Eqs. 6-15 and 6-16. However, in practice all the raw image data are expressed by 256 grey levels and a truncation error would be introduced. The low level of the brightness and contrast results increases the size of the undefined zones of the isoclinic angle, and this potentially makes the replacement of undefined data more difficult.

Saturation can be caused by too large a lens aperture or by a high level of the brightness or contrast. Both cases would make the analysis results incorrect and should be avoided during the experiment. The detection of the saturation should be concentrated on image 1 and image 2 in the conventional polariscope or the corresponding ones in the PSIOS system.

7.1.3.3 A/D Converter

The conversion of the signal from analogue to digital is essential for computer analysis; and this conversion will introduce spatial quantisation and digitisation of the

brightness level. The former should not introduce any error as long as the resolution is less than 0.26 fringes per pixel (see **Section 3.4.2**). The latter will always bring some influence to the results. This is an inherent drawback of the digitiser and can not be deleted in anyway. Its contribution to the measured intensity will be discussed later.

7.1.4 Image Analysis

The analysis of the simulated model, which is noise free, showed that no errors were introduced during the procedure itself except in the region with high fringe gradient. However, noise in the images was inevitable in practice due to many kinds of error source and will cause some mistakes, such as wrapping errors, in the analysis. Use of a Gaussian smoothing filter, as a weighted average method, was found to be effective for improving the image quality and reducing the wrapping errors.

7.1.5 Discussion

As pointed out in **Chapter 3**, the techniques of automated analysis for photoelastic fringes was shown to be derived from the same principles. It would, therefore, be expected that the techniques would also share common sources of error when they are implemented in a similar way. The above analysis uses the phase-stepping systems proposed by Patterson and Wang as exemplars. All the error sources analysed can also be applied to other phase-stepping systems, such as those of Hecker and Morche [1986], and of Asundi [1993]. For the one proposed by Sarma *et. al.* [1992], since a

plane polariscope was employed, the errors associated with the quarter-wave plates and the light source are not applicable in this case. Object step-loading system proposed by Ekman and Nurse [1997] used the similar method as Patterson and Wang [1991] and therefore shared the same error sources.

Point-by-point spectral contents analysis requires the broad band of wavelengths to allow a comparison of a series of theoretical spectra with the experimental one in order to find a solution with the best fit. So the errors associated with the light source and the matching quarter-wave plates are not applicable. Also because a sensor other than a CCD camera was used to acquire the light intensity, the image acquisition part has different error sources.

For the Fast Fourier Transform (FFT) method [Morimoto *et. al.* 1994, Quan *et. al.* 1993], the error sources existed in the phase-stepping system also exist in FFT. They may bring different influence to the results due to a different implementation method used. Object step-loading system proposed by Ng [1997] also have the same error sources but may have different influences.

The RGB method [Ajovalasit *et. al.* 1995a, b, Ramesh and Deskmukh 1996] uses a circular polariscope with the RGB camera to collect the colour images in white light. All the analysis can be applied in this case except the part for the monochromatic camera.

7.2 System Evaluation

The influence of some of the error sources can be reduced by calibration and carefully conducting the experiment, or by using a high quality instrument, while others are inherent properties. The influence, however, will always result in the misrepresentation of the observed images. From a point of view of every image, the errors fall into two categories, one results in distortion of the image geometry and the other introduces a deviation in the measured light intensity. The former is normally negligible after calibration, while the latter can not be avoided in the whole system. This section will first consider this combined effect on the measured light intensity, then evaluate the accuracy of whole system.

7.2.1 Total Effect on the Measured Light Intensity

In the above discussion, the spatial and temporal variation of the light and environment, the different sensitivity of the individual camera cells, the digitisation errors in the brightness level and the inhomogeneities of the optical elements were all seen to contribute to the errors in the measured light intensity. It is difficult to identify their individual effects and to assign a magnitude to each of them. However the total effect on the light intensity can be estimated by studying a large number of samples [Kamgar-Parsi *et.al.*, 1993].

A standard conventional polariscope was used with a 15 inch field of view and a sodium light source. The data was collected using a monochromatic CCD camera (Panasonic WV-BP100), with a 25 mm lens. The compression disc was placed in the

field of view and the light intensity fields were collected 100 times for each of the phase-stepped images. The data along the diameter of the disc perpendicular to the load was examined. Fig. 7.1 shows the distributions of the variations in light intensity for all positions. From Fig. 7.1, the variation of the recorded data has an approximately normal distribution with a mean of zero and variance of $\sigma_i^2 = 3.51$. There was no significant difference between the variations for the six phase-steps, thus it can be concluded that the error sources are the same and the probabilities of the errors occurring in each step are equal.

7.2.2 Error Distribution for the Isochromatic and Isoclinic Data

In the conventional polariscope, the fractional isoclinic angle θ_f and the relative retardation α_f are determined by:

$$\theta_f = \frac{1}{2} \arctan \frac{i_s}{i_c} = \frac{1}{2} \arctan \frac{i_3 - i_5}{i_6 - i_4} \quad (3-16)$$

$$\alpha_f = \arctan \frac{i_3 - i_5}{(i_2 - i_1) \cdot \sin 2\theta} = \arctan \frac{i_6 - i_4}{(i_2 - i_1) \cdot \cos 2\theta} \quad (3-17)$$

Clearly, the errors in the recorded light intensity will influence the isoclinic angle and the isochromatic fringe order. If di_n is used to represent the errors in the n th image, then $di_1 = di_2 = \dots = di_6 = di$. The total effect on the isoclinic angle can be obtained from the differential form of Eq. 3-16, i.e.,

$$d\theta = \frac{\partial \theta}{\partial i_3} di_3 + \frac{\partial \theta}{\partial i_4} di_4 + \frac{\partial \theta}{\partial i_5} di_5 + \frac{\partial \theta}{\partial i_6} di_6 = \sum_{n=3}^6 \frac{\partial \theta}{\partial i_n} di, \quad (7-1)$$

This equation describes the error which would occur in the isoclinic angle, θ , when each image has a small error in the intensity. Since statistically the errors are the same for the six images, the errors in the light intensity have the same distribution as,

$$di_1 = di_2 = \dots = di_6 = di \sim N(0, \sigma_i^2) \quad (7-2)$$

According to **Appendix B** entitled "The Distribution of Multivariate Normal Density Function", the distribution of errors in the isoclinic data E_θ , has a mean of zero and a variance of σ_θ that will be given by [Brown, 1983],

$$E_\theta = \frac{1}{\sqrt{2\pi} \cdot \sigma_\theta} \cdot \exp\left(-\theta^2 / 2\sigma_\theta^2\right) \quad (7-4)$$

where

$$\sigma_\theta^2 = \sum_{n=3}^6 \left(\frac{\partial\theta}{\partial i_n}\right)^2 \cdot \sigma_n^2 = \frac{2\sigma_i^2}{(i_3 - i_5)^2 + (i_6 - i_4)^2}$$

Similar expressions can be derived for the isochromatic fringe order:

$$E_\alpha = \frac{1}{\sqrt{2\pi} \cdot \sigma_\alpha} \cdot \exp\left(-\alpha^2 / 2\sigma_\alpha^2\right) \quad (7-5)$$

$$\sigma_\alpha^2 = \sum_{n=1}^6 \left(\frac{\partial\phi}{\partial i_n}\right)^2 \cdot \sigma_n^2 = \frac{2\sigma_i^2}{(i_2 - i_1)^2 + (i_3 - i_5)^2 + (i_6 - i_4)^2}$$

where α and θ are the computed values for the relative retardation and isoclinic angle, and σ is the variance of the image intensity (Fig. 7.1) from the multi-sample experiments. So for an arbitrary point, the distribution of errors in the isoclinic and isochromatic data can be estimated accurately using the values of the six intensities.

For a disc subject to diametral compression, the ideal isoclinic angle along the central line perpendicular to the loading direction is 0 degree, i.e., $i_3 = i_4 = i_5 = i_6$, the denominator of the expression (7-4) is zero, hence there is no means to discuss the error distribution along this line. Line CD, which is 9 mm above the central line, the

diameter of the disc was 76 mm, was chosen. The resulting distributions of errors in the isochromatic and isoclinic parameters are shown in Fig. 7.2. For the fringe order of 1.5, the variance of the fringe order at this point was about 0.045. It can be clearly seen that the largest variance values occurred in the region with integer or half fringe orders and the magnitude of the variance depends on how near the fringe value is to a integer or half. The variance of the isoclinic angle is up to 536 degrees. Again, the variance in integer and half fringe orders is much larger than in other areas. The distribution for the isoclinic angle in the other regions can be seen after removing the two largest points (536 and 118 degrees).

In a similar way, the total effect of the measured light intensity on the isoclinic angle and isochromatic fringe order for the PSIOS system can be derived from the differential form of Eqs. 3-23 and 3-24.

7.2.2.1 The Digitisation Error in the Isochromatic and Isoclinic Data

In Section 6.4.3, the standard deviation caused by the digitisation error in the isochromatic and isoclinic data was deduced as

$$\Delta\sigma_a^2 = \frac{1}{12}(\alpha_1^2 + \dots + \alpha_6^2) \quad (6-26)$$

$$\Delta\sigma_\theta^2 = \frac{1}{12}(\theta_3^2 + \dots + \theta_6^2) \quad (6-27)$$

Compared with Eqs. 7-4 and 7-5, the deviation caused by the digitisation error in the total deviation of the light intensity is obtained as:

$$\frac{\Delta\sigma_a^2}{\sigma_a^2} = \frac{\Delta\sigma_\theta^2}{\sigma_\theta^2} = \frac{1}{12\sigma_i^2} \quad (7-6)$$

For the case of $\sigma_i^2 = 3.51$,

$$\frac{\Delta\sigma_a}{\sigma_a} = \frac{\Delta\sigma_\theta}{\sigma_\theta} = 15.41\% \quad (7-7)$$

So, for the isochromatic and isoclinic data the deviation caused by the digitisation error makes a 15.41% contribution to the whole intensity error. If it is possible to reduce the variance of the light intensity by improving other experimental conditions and performing a more careful experiment, then the digitiser will make more contributions to the total error.

7.2.3 System Evaluation in Current Conditions

The factors that influence the system accuracy can be identified from the discussions about the error sources and their contributions to the system. The accuracy of the conventional polariscope and the PSIOS system can then be evaluated through the use of disc model. The evaluation is based on the simulation of the disc, while the simulation is firmly based on the experimental work.

7.2.3.1 The Conventional Polariscope

Assuming that the operation is perfect with respect to the system's requirement or specification, i.e., no alignment error is introduced, then for the conventional polariscope, the mismatch of the quarter-wave plates to the sodium light source,

nonlinearity of the CCD camera output and their combined effect on the measured light intensity are considered as the main factors to influence the system accuracy. Because the errors in the light intensity have a normal distribution with a mean of zero, as stated previously, the combined effect on the isochromatic fringe order and the isoclinic angle follows a normal distribution with a mean of zero. Hence, the evaluation of system accuracy only takes account of the first two factors.

The evaluation is based on the simulation of the disc, while the simulation is firmly based on the experimental validation. Referring back to Fig. 3.7, two disc models, one with considering the mismatching of the quarter-wave plates to the sodium light source and nonlinearity of the CCD camera and the other without any error influence, were used to compare and to obtain the system accuracy. The analysis results of the first model for the isochromatic fringe order and isoclinic angle are shown in Fig. 7.3. The comparison of the two discs was made along the central line of the disc and is shown in Fig. 7.4. For a fringe order of 1.5, the model including the errors gives 1.408, and the accuracy achieved is 93.87%. If the fringe order at sodium light source is converted to the one at the 560 nm matching wavelength, then the accuracy increases to 99.23%. The errors in the isoclinic are less than 5 degrees if the undefined zones are ignored.

7.2.3.2 The PSIOS System

In the PSIOS system, mismatch of the quarter-wave plates to the white light source (named L1 in Chapter 5), nonlinearity of the CCD camera output, the allowed

tolerance in the cube beamsplitter, the oblique viewing of the specimen by the instrument and their combined effect on the measured light intensity are considered as the main factors to influence the system accuracy provided that the spatial position has been carefully calibrated. Again, the combined effect on the isochromatic and the isoclinic angle due to the errors in the light intensity follows a normal distribution with a mean of zero. So, the factors except the combined effect on the measured light intensity are used simulate the errors in the disc.

The analysis results for the disc model are shown in Fig. 7.5. The wrapping error is clearly observed in the high fringe order area. The comparison was again made along the central line of the disc to the simulated model without any error and is shown in Fig. 7.6. For a fringe order of 1.5, the model including errors gives 1.419, and the accuracy achieved is 94.60%. The errors in the isoclinic angle are less than 6 degrees if the undefined zones are ignored. Unlike the conventional polariscope, the errors in the fringe order can not be changed to the one at 560 nm, and the wrapping errors at high fringe order regions limit the analysis to low fringe orders, normally less than 1.5.

7.3 Conclusion

All of the error sources and their contribution to the system have been discussed. The errors in the measured light intensity were studied and it was concluded that the error distributions for the isochromatic and isoclinic data follow a normal distribution with a mean of zero and variances as expressed in Eqs. 7-4 and 7-5.

The accuracy of the conventional polariscope and the PSIOS system was evaluated through the simulation of the disc, while the simulation was firmly based on the previous experimental works.

For the conventional polariscope, the system accuracy can be 93.87% for the fringe order, but if the fringe order at sodium light source is converted to the one at the 560 nm matching wavelength, then the accuracy increases to 99.23%. The errors in the isoclinic angle are less than 5 degrees if the undefined zones are ignored.

For the PSIOS system, the system is limited to the low fringe order, normally less than 1.5 and the accuracy is 94.60% for the fringe order. The errors in the isoclinic angle are about 6 degrees if the undefined zones are ignored.

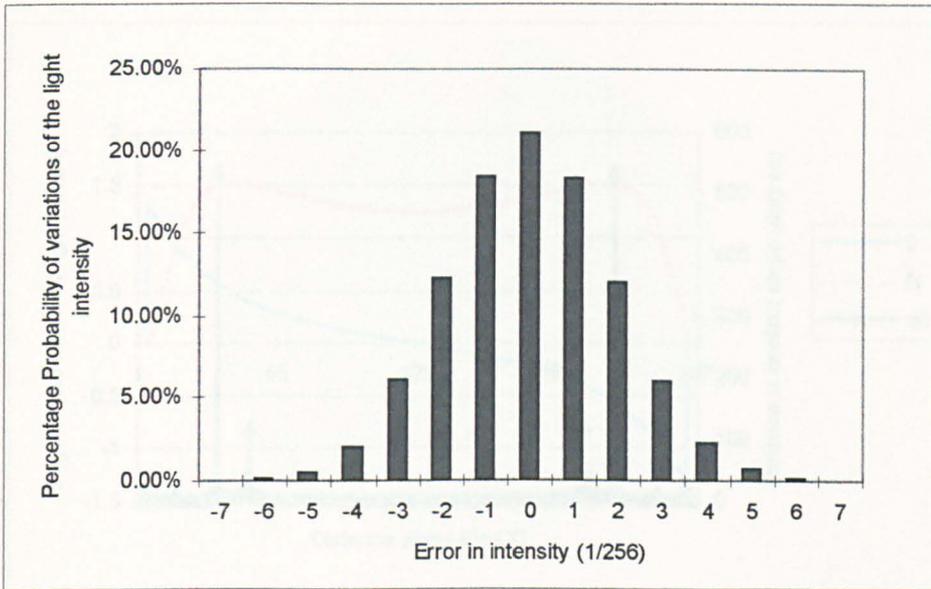


Fig. 7.1 A histogram showing the distribution of variation in measured light intensities as compared to the mean value for 100 samples. The data is aggregated in the graph. The variance of the data shown is 3.51.

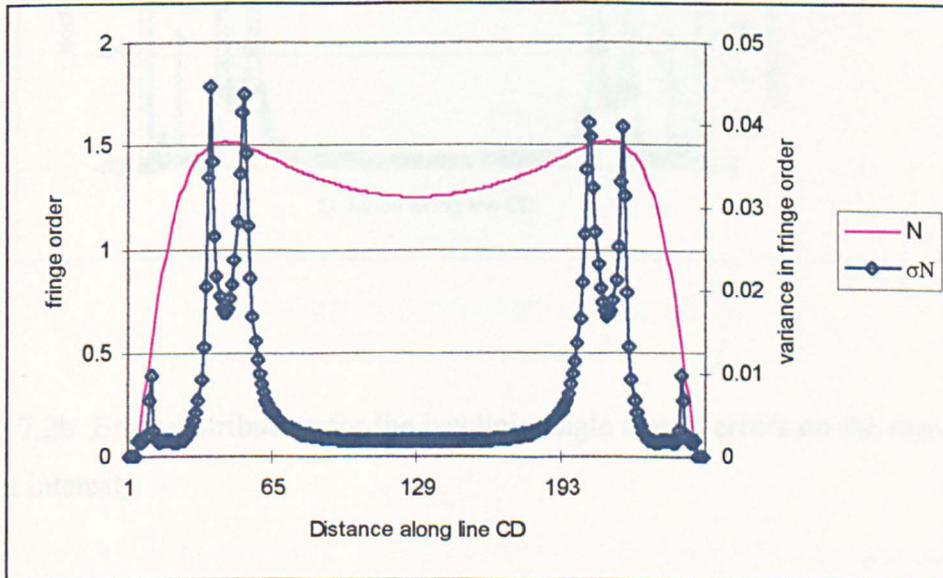
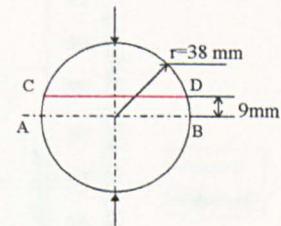


Fig. 7.2a Error distribution for the fringe order and isoclinic angle due to errors on the measured light intensity

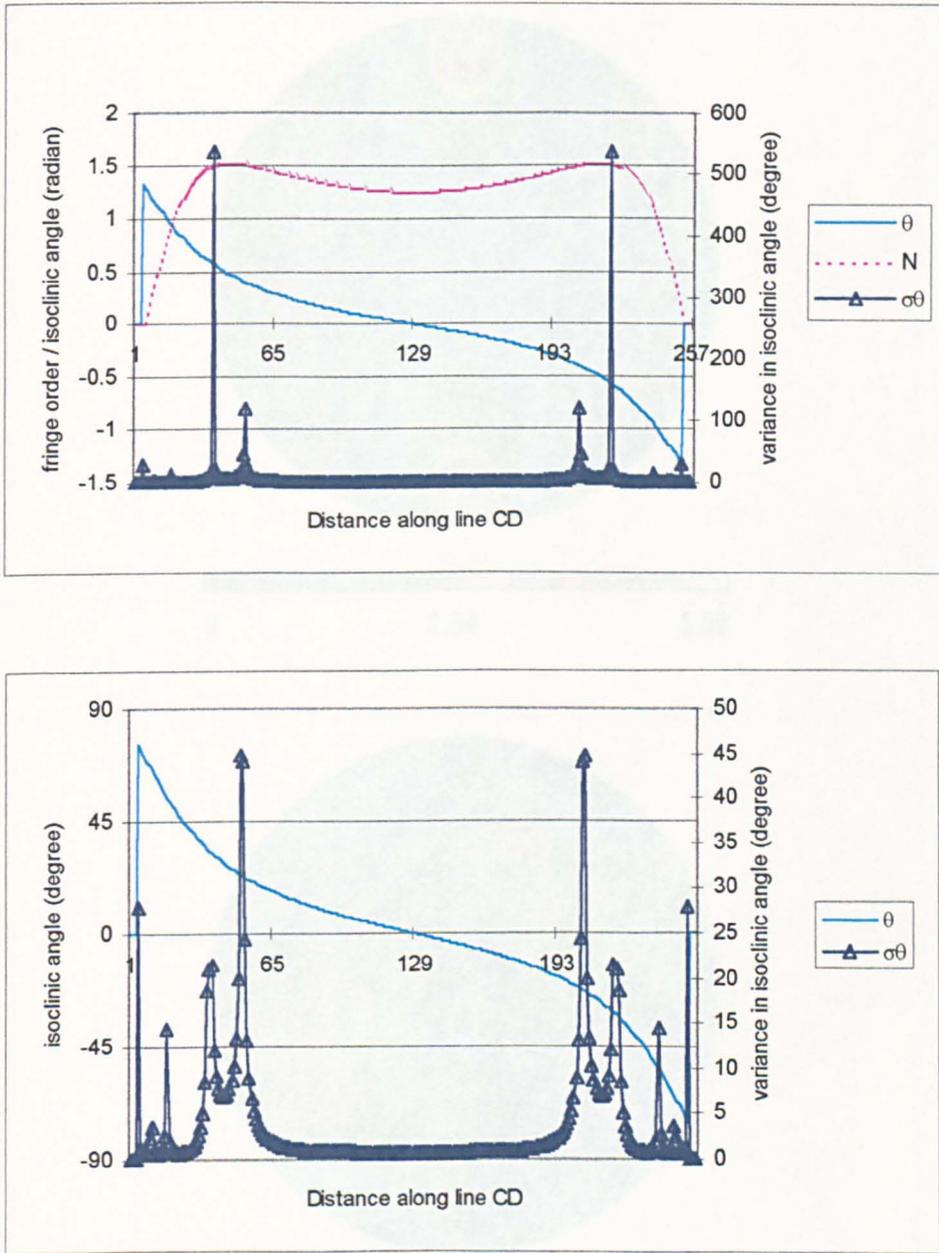


Fig. 7.2b Error distribution for the isoclinic angle due to errors on the measured light intensity

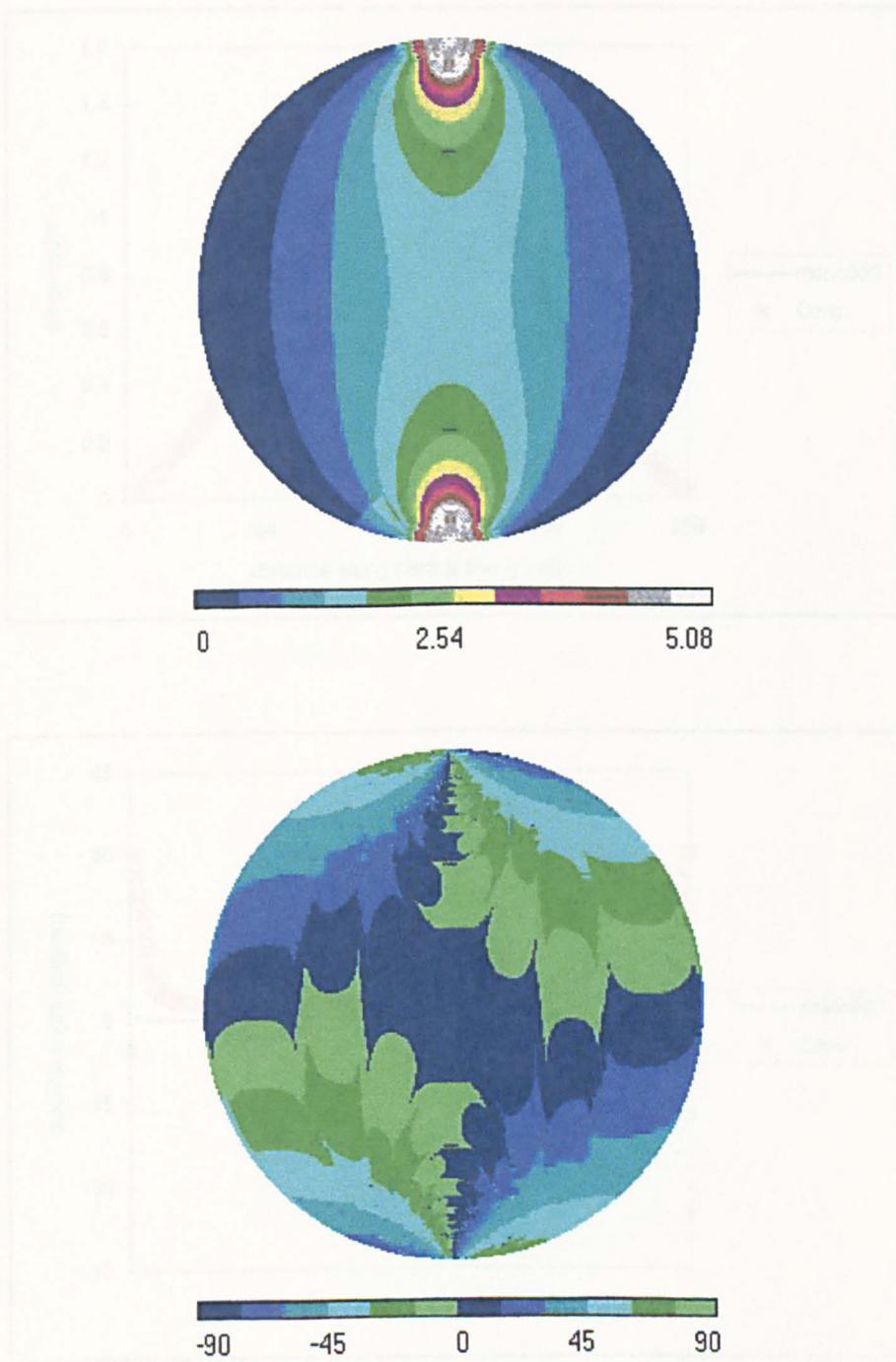


Fig. 7.4 The system evaluation on the image color and motion using the

Fig. 7.3 The disc model shows the system evaluation for the conventional polariscope

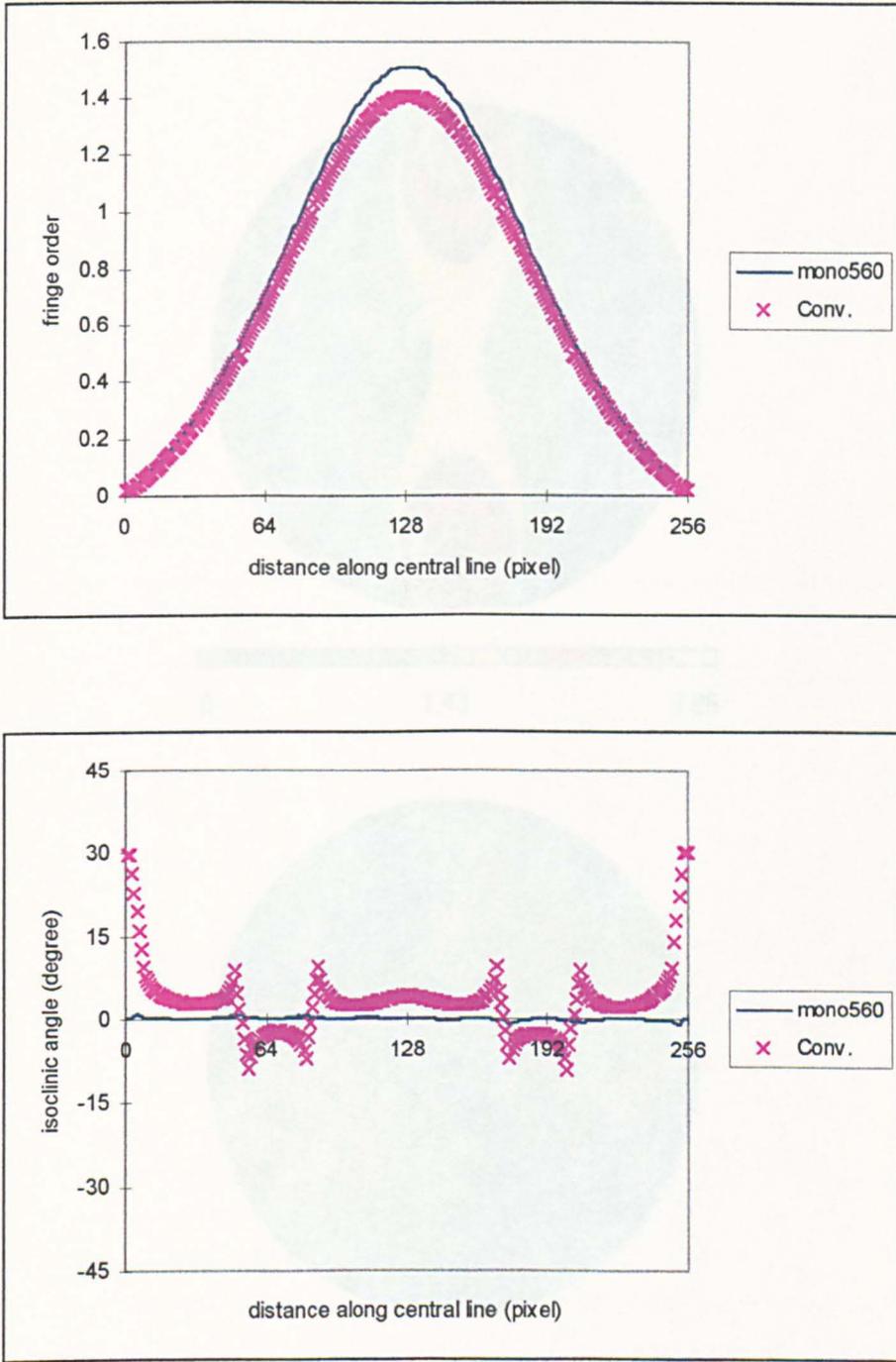


Fig. 7.4 The system evaluation on the fringe order and isoclinic angle along the central line of the disc for the conventional polariscope

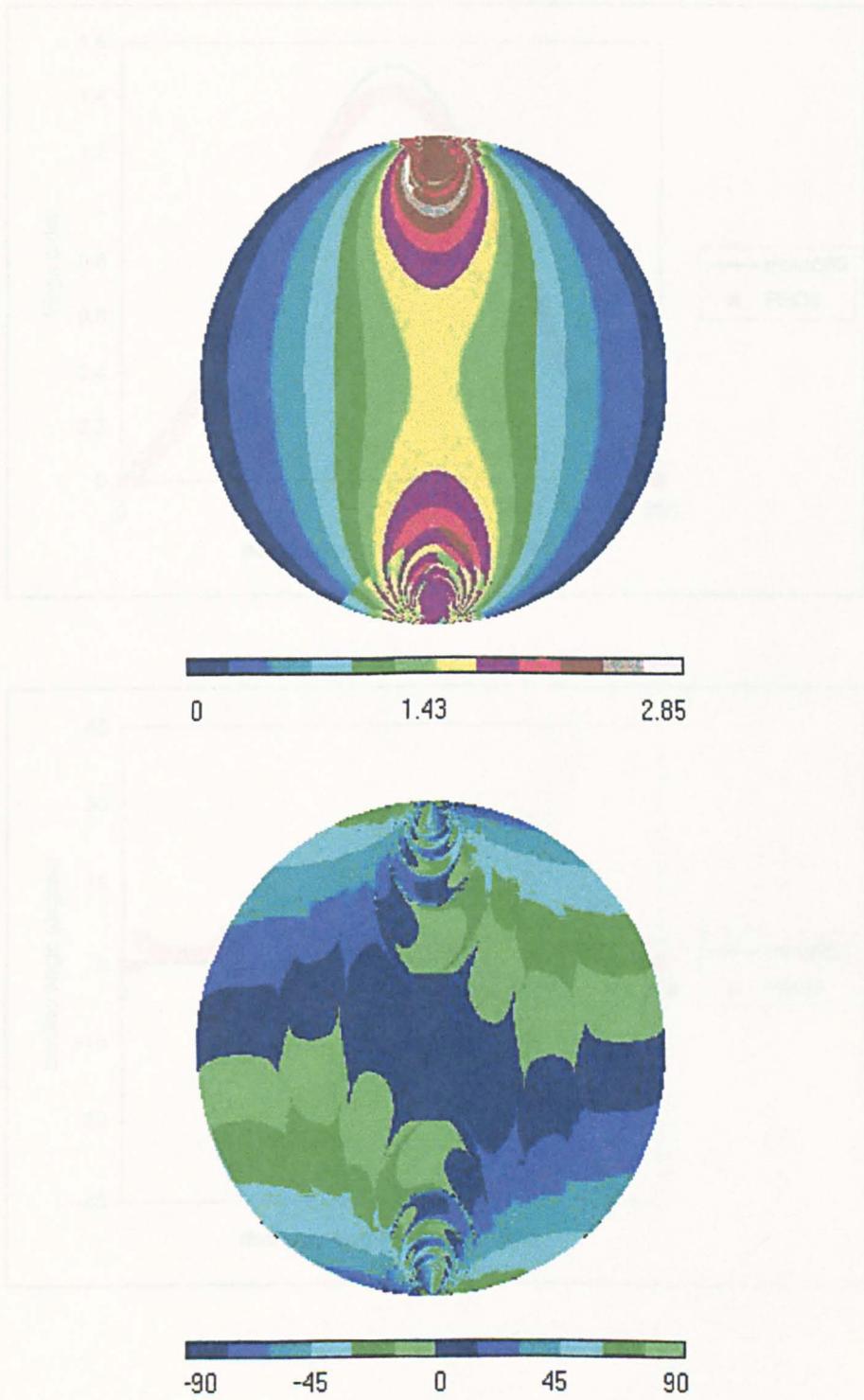


Fig. 7.5 The system evaluation on the two sides and central region along the

Fig. 7.5 The disc model shows the system evaluation for the PSIOS system

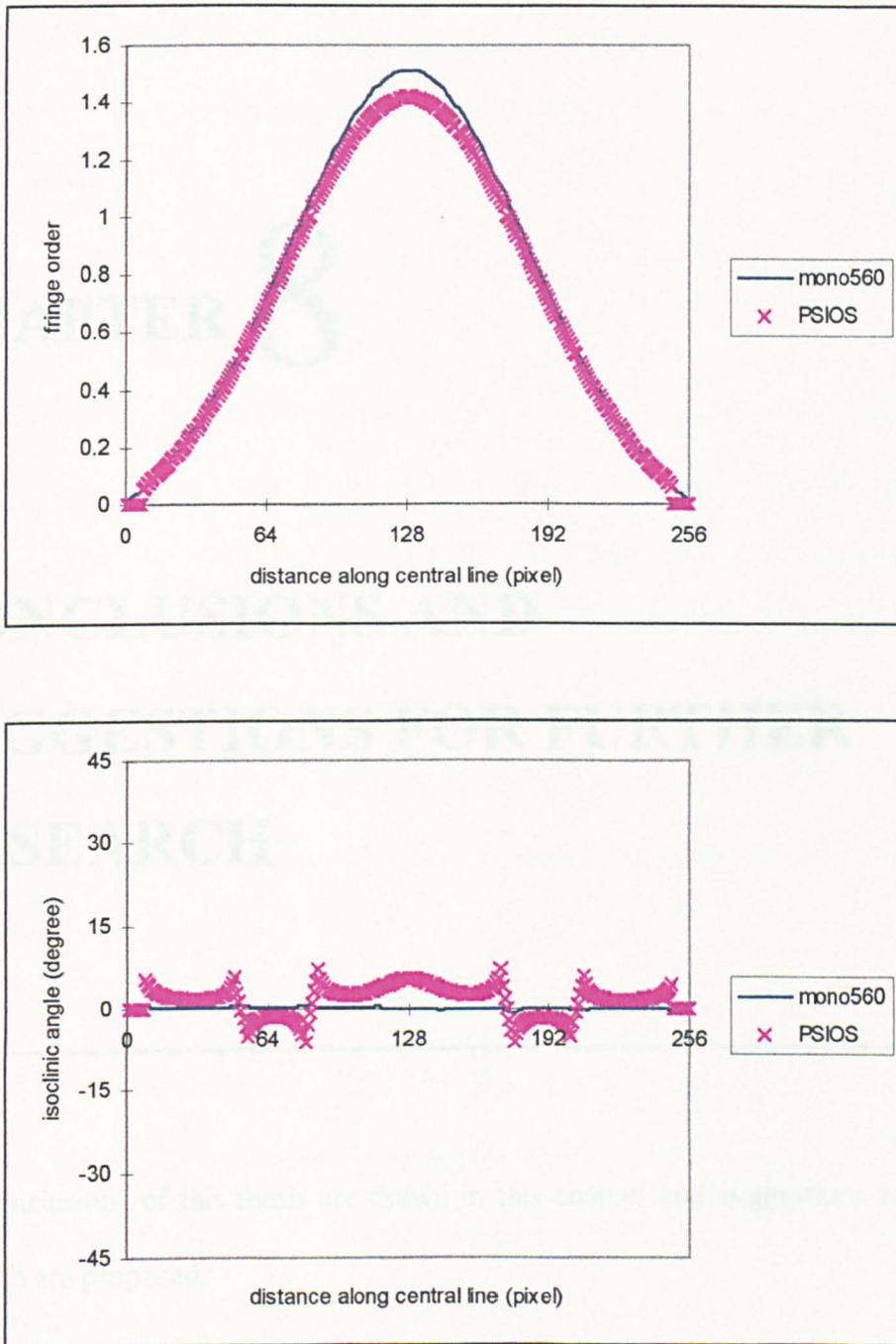


Fig. 7.6 The system evaluation on the fringe order and isoclinic angle along the central line of the disc for the PSIOS system

CHAPTER 8

CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH

The conclusions of this thesis are drawn in this chapter and suggestions for further research are proposed.

8.1 Conclusions

The error analysis and system improvements have been based on the following guiding principles:

- To separate the system concerned into consecutive sections according to their function;
- In order to investigate the nature of each section and the general relationship between the output of the system and the system components, the formulation of a general system set-up, within which the parameters of optical elements could change arbitrarily, was deduced;
- All sections of the system were analysed so that the possible contribution of each section to the errors in the final output of the system could be found;
- The analysis of each section was used to identify potential improvements to the accuracy of the output;
- Because the system acts as a whole, and there is always a possibility that an unexpected miscondition could be involved in every section of the system. The system evaluation was carried out based on the combined influence of all sections of the system on the accuracy of the system output;
- To identify the contribution of each possible error to the final system output, a model was used to simulate the errors and was employed throughout the research.

8.1.1 Fundamental Theory for Photoelasticity

The general relationship for a polariscope with all the optical elements in an arbitrary orientation was deduced using the method of Mueller Matrices. All the factors associated with an imperfect optical apparatus or an imperfect optical arrangement and which would result in errors in the isochromatic fringe order and isoclinic angle were included. The study of this relationship indicated the inherent connectivity among the optical elements and provided a knowledge of the errors.

The deduction of the general formulation also showed the common foundation of the automatic techniques in photoelasticity and, consequently, these techniques share many common error sources.

8.1.2 Error Analysis and Improvement in Each Stage

The automated phase-stepping system can be divided into four parts, namely the optical system, the light source, the image acquisition equipment and the image analysis software. All the possible error sources were investigated separately and the methods for reducing the influence of the errors and improving the system accuracy were presented.

- **Optical system**

In the conventional polariscope, the requirement of matching the light source to the quarter-wave plate is very important, and a mismatch of 32 nm could cause 6% errors for a fringe order of 1.5 and a 31 degrees error on the isoclinic angle. For the case with less than 5 degrees linear misalignment and 0.1 degrees angular misalignment of

the optic components, the influence of the misalignments could be restricted to less than 0.02 on fringe order and 4 degrees on the isoclinic angle. (Section 4.1.7)

For the PSIOS system, the spatial alignment is very important because the system can be used only after calibration. An alignment target was designed and methods for calibration were provided. The transmission : reflection ratios for the system were obtained and used to modify the intensity coefficients. The maximum allowed tolerance of the cube beamsplitters have almost no influence on the fringe order and produce less than 5 degrees of error on the isoclinic angle. (Section 4.2.5)

- **Light source**

The effect of the light source on the analysis was revealed. An ideal light source should have a wavelength matched to the quarter-wave plate. However, if a white light source has to be used, the form of the spectrum should have a high intensity in the area adjacent to the matching wavelength and low intensity elsewhere and also should be symmetrical about the matching wavelength. White light sources are limited to low fringe orders if reliable results are to be acquired, typically 1.5 fringe orders for $\pm 5\%$ errors. The employment of medium-band filters could reduce the errors both in the isochromatic fringe order and isoclinic angle significantly. For 40 and 80 nm bandwidth filters, the maximum fringe order could be measured increased to 10 and 6 fringes with the same accuracy requirements. (Section 5.5)

- **Image acquisition**

The sensor of the CCD camera shows nonlinearity of its output with respect to the input light intensity and this results in the mismeasurement of the light intensity in the darkest regions. Although no obvious influence is observed, it increases the difficulty in applying interpolation on the undefined zones in the isoclinic parameter. (Section 6.2.5)

The amplifier has a function to control the brightness and contrast parameters. The effect of adjustments using this function was revealed and the best values for the parameters were provided. (Section 6.3.5)

The A/D converter limits the image resolution to less than 0.26 fringes per pixel and makes a 15.41% contribution to the total errors in the measured light intensity. (Sections 6.4.3 and 7.2.2.1)

- **Image analysis**

No errors were introduced by the image analysis procedure. However, the analysis results may include some errors due to the accumulation of the errors in other parts. A Gaussian smoothing filter was examined and found to be suitable to remove the noise without losing the phase-stepping information in the images. (Section 6.5.7)

8.1.3 System Evaluation

The error distribution for the isochromatic fringe order and isoclinic angle was deduced to be due to the errors in the light intensity measurements. The system accuracy can be as good as 93.87% for the fringe order and the error less than 5

degrees for the isoclinic angle in a conventional polariscope. The accuracy increases to 99.23% if the fringe order in sodium light source is converted to the one for the 560 nm matching wavelength.

For the PSIOS system, the system is limited to the low fringe orders. For a fringe order of less than 1.5, the accuracy is 94.60% for fringe orders and the error about 6 degrees for the isoclinic angle if the undefined zones are ignored. (Section 7.3)

8.1.4 A Simulation of Errors

A simulation of a circular disc under diametral compression was built and used throughout the investigation. This way of investigating the errors and their influence has the following advantages:

1. It can separate the errors and their influence and then identify the contributions of each error source to the whole system. This can reveal the intrinsic properties of the errors and provide the methodology for reducing the errors.
2. It is very helpful in testing the effectiveness of a proposed method for reducing the influence of an error.

8.2 Suggestions for the Further Research

The study of the error simulation indicates that even a tiny error will introduce a large influence on the isoclinic angle in the zones of integer and half fringes due to its undefined nature in this area. The errors in the isoclinic angle influence the fringe

order, and sometimes may even introduce wrapping errors. This is a big disadvantage of the phase-stepping method.

The employment of the Gaussian filter can improve the image quality effectively and reduce the incidents of wrapping error. The implementation of the image pre-processing technique to the automated photoelasticity would be very useful for improving the system accuracy.

More effort is needed to remove or minimise the interaction of the isoclinic angle and the isochromatic fringe order.

References

- [1] Ajovalasit, A, Barone, S., and Petrucci, G.,
"Towards RGB Photoelasticity: Full-Field Automated Photoelasticity in White Light"
Exp. Mech., 35 (1995a), 193-200.
- [2] Ajovalasit, A, Barone, S., and Petrucci, G.,
"Automated Photoelasticity In White Light: Influence of Quarter-wave Plates"
J. Strain Analysis, v30, n1 (1995b), p29-34.
- [3] Allison, I. M. and Nurse, P.
"The Benefits and Pitfalls of Automatic Processing for Photoelastic Test Data"
In "*Applied Solids Mechanics*" Ed. by Hyde, T. H. and Ollerton, E.
Elsevier Applied Science, London, p536-544 (1990)
- [4] Allison, I. M.
"Computation and Interpretation of Photoelastic Test Data"
In "*Applied Stress Analysis*" Ed. by Allison, I. M. and Ruiz, C.
Elsevier Applied Science, London, p258-268 (1989)
- [5] Allison, I. M. and Nurse, P.
"Automatic Acquisition of Photoelastic Data"
Proc. JBCSA Conf. on the Recording and Interpretation of Engineering Measurements Inst. Mar. Engrs., London, p203-207 (1972)

-
- [6] Amelio, G., Tompsett, M. and Smith, G.
“Experimental Verification of the Charge Coupled Device Concept”
Bell Syst. Tech. J., Vol. 49, p593-600, Apr. 1970
- [7] Asundi, A.
“Phase Shifting in Photoelasticity”
Exp. Tech., v17n1, p19-23 (1993)
- [8] Barbe, D. F.
“Charge-coupled Devices”
“*Topics in Applied Physics*” vol 38, Springer-Verlag, 1980
- [9] Barone, S., and Patterson, E. A.
“Full-field Separation of Principal Stress by Combined Thermo- and Photoelasticity”
Exp. Mech., v36n4, p318-324 (1996)
- [10] Bevington, F. R.
“Data Reduction and Error Analysis for the Physical Sciences”
New York : McGraw-Hill, 1969
- [11] Boyle, W. and Smith, G.
“Charge Coupled Semiconductor Devices”
Bell Syst. Tech. J., Vol. 49, p587-593, Apr. 1970
- [12] Brewster, D
“On the Communication of the Structure of Doubly Refracting Crystals to Glass, Muriate of Soda, Fluor Spar and Other Substances, by Mechanical Compression and Dilatation”
Philosophical Transactions of the Royal Society, p156-178, (1816)

-
- [13] Brown, G. M. and Sullivan, J. L.
"The Computer-aided Holophotoelastic Method"
Exp. Mech., v30, p135-144 (1990)
- [14] Brown, R. G.,
"Introduction to Random Signals and Applied Kalman Filtering"
John Wiley & Sons, Inc., New York (1983),
- [15] Carazo-Alvarez, J., Haake, S. J. and Patterson, E. A.
"Completely Automated Photoelastic Fringe Analysis"
Optical and Lasers in Engineering, 21,p133-149 (1994).
- [16] Chaudhari, U. M. and Godbole, P. B.
"Direct Oblique Incidence Method for Reflection Photoelasticity"
Exp. Tech., v14n1, p37-39 (1990)
- [17] Chen, T. Y. and Taylor, C. E.
"Computerized Fringe Analysis in Photomechanics"
Exp. Mech., v29, p323-329 (1989)
- [18] D'Agostino, J., Drucker, D. C., Liu, C. K. and Mylonas, C.
"An Analysis of Plastic Behaviour of Metals with Bonded Birefringent Plastic"
Proc. Soc. Exp. Stress Anal. 12, pp115-122 (1955)
- [19] Dally, J. W. and Riley, W. F.
"Experimental Stress Analysis" 3rd Edition ,McGraw-Hill, (1991)
- [20] Duffy, J.
"Effect of the Thickness of Birefringent Coatings"
Exp. Mech. v1n3, p74-82 (1961)

-
- [21] Ekman, M. J. and Nurse, A. D.
"On the Phase Unwrapping of Isochromatics"
Proc. 1997 SEM Spring Conf. on Exp. Mech., p317-318 (1997)
- [22] Fessler, H., Marston, R. E. and Ollerton, E.
"A Micropolariscope for Automatic Stress Analysis"
J. Strain Analysis, v22 n1, p25-35 (1983)
- [23] Frocht, M. M
"Photoelasticity"
John Wiley, New York, Vol1, 1941; Vol2, 1948.
- [24] Gonzalez, R. C. and Woods, R. E.
"Digital Image Processing"
Addison Wesley, Reading, MA, (1993)
- [25] Grant, I.
"Particle Image Velocimetry: A Review"
Proc. Institution of Mechanical Engineering, Vol 211 Part C, p 55-76
(1997)
- [26] Haake, S. J., Patterson, E. A.,
Photoelastic analysis using a full-field spectral analyser.
Proc. SEM Spring Conf: on Exp. Mech., p342-5 (1995).
- [27] Haake, S. J., Wang, Z. F. and Patterson, E. A.
"Evaluation of Full-field Automated Photoelastic Analysis Based on Phase-stepping System"
Exp. Tech. v17n6, pp19-25 (1993)
- [28] Haake, S. J. and Patterson, E. A.

-
- "Photoelastic Analysis of Frozen Stressed Specimens Using Spectral Contents Analysis"
Exp. Mech., v32, p266-272 (1992)
- [29] Healey, G. E.
"Radiometric CCD Camera Calibration and Noise Estimation"
IEEE Trans. on Pattern Analysis and Machine Intelligence, v16n3, p267-276, (1994)
- [30] Hecker, F. W. and Morche, B.
"Computer measurement of Relative retardations in Plane Photoelasticity"
Expt Stress Anal., (Ed) Wieringa, H., Martinus Nijhoff Publishers Dordrecht, The Netherlands, p532-543 (1986)
- [31] Howes, M. J. and Morgan, D. V.
"Charge-coupled Devices and Systems"
John Wiley & Sons, 1979
- [32] Hung, Y. Y. and Pottinger, M. G.
"An Improved Oblique Incidence Technique for Principal Strain Separation in Photoelastic Coatings"
Exp. Mech., v20n5, p170-173 (1989)
- [33] Kamar-Parsi, Behrooz and Kamar-Parsi, Behzad
"Evaluation of Quantization Error in Computer Vision"
Computer Vision and Pattern Recognition, June 1988, pp52-60
- [34] Kenny, B.
"The Casting of a Low Exotherm Resin"
J. Sci. Instruments, 42, pp719-720 (1965)

-
- [35] Kihara, T.,
“Automatic Whole-Field Measurement of Photoelasticity Using Linear Polarized Incident Light”
Proc. 9th Int. Conf. on Experimental Mechanics, Copenhagen, Denmark, vol.2, pp821-7 (1990)
- [36] Langford, M. J.
“The Darkroom Handbook”
Random House, New York, (1984)
- [37] Lee, T. C., Mylonas, C. and Duffy, J.
“Thickness Effects in Birefringent Coatings with Radial Symmetry”
Exp. Mech. v1, n10, p134-142 (1961)
- [38] Levi, L
“Applied Optics” Vol 1, “A Guide to Optical System Design”
John Wiley & Sons, New York, 1968
- [39] Mawatari, S., Takashi, M., Toyada, Y., & Junio, T.,
A single valued representative function for determination of principal stress direction in photo elastic analysis.
Proc. 9th Int. Conf: Exptl. Mech., 5. SEM, Copenhagen, Denmark, pp.2069-2078 (1990).
- [40] Maxwell, E. A.
“General Homogeneous Co-ordinates in Space of Three Dimensions”
Cambridge University Press, Cambridge, England, 1951
- [41] Mesnager, M.
“Sur la Determination Optique des Tensions Interieures dans les Solides a Trois Dimentsions” *C. R. Acad. Sci.* 190, 1249 (1930)

-
- [42] Morimoto, Y., Morimono, Y. Jr. and Hayashi, T.
"Separation of Isochromatics and Isoclinics Using Fourier transform"
Exp. Tech. v18n5, p13-17 (1994)
- [43] Morimoto, Y., Morimono, Y. Jr. and Hayashi, T.
"Separation of Isochromatics and Isoclinics Using Fourier Transform and Its Accuracy"
Proc. 1993 SEM Spring Conf. on Exp. Mech., p149-158 (1993)
- [44] Müller, R. K. and Saackel, L. R.
"Complete Automatic Analysis of Photoelastic Fringes"
Exp. Mech., v19, p245-251 (1979)
- [45] Nalwa, V. S.
"A Guided Tour of Computer Vision"
Addison-Wesley Publishing Company, 1993
- [46] Ng, T. W.
"Photoelastic Stress Analysis Using an Object Step-Loading Method"
Exp. Mech., v37n2, p137-141 (1997)
- [47] Oppel, G. U.
"Polarisationoptische Untersuchung, Raumlicher Spannungs - und Dehnungszustände"
Forsch, Geb. Ingenieurw., v7, pp240-248 (1936)
- [48] Otani, Y., Shimada, T., Yoshizawa, T., and Umeda, N.,
"Two-dimensional birefringence measurement using the phase shifting technique"
Optical Engineering, 33, p1604-1609 (1994)

-
- [49] Pacey, M. N. ,Wang, X. Z., Haake, S. J. and Patterson, E. A.
“The Application of Evolutionary and Entropic Algorithms to Photoelastic Spectral Analysis”
Submitted to Journal of *Experimental Mechanics*
- [50] Patterson, E. A., Ji, W. and Wang, Z. F.
“On Image Analysis for Birefringence Measurement in Photoelasticity”
Optical and Lasers in Engineering, 28, p17-36, (1997)
- [51] Patterson, E. A. and Wang, Z. F.
“Phase-stepping in Photoelasticity Using White Light”
Proc. 1997 SEM Spring Conf. on Exp. Mech., p366-367 (1997a)
- [52] Patterson, E. A. and Wang, Z. F.
“Simultaneous Observation of Phase-stepped Images for Automated Photoelasticity”
Provisional accepted for publication in journal of *Strain* (1997b)
- [53] Patterson, E. A. and Wang, Z. F.
"Towards Full Field Automated Photoelastic Analysis of Complex Components"
Strain, v27 p49-53 (1991)
- [54] Patterson, E. A.
“Automated Photoelastic Analysis”
Strain, v24n2, p15-20, (1988)
- [55] Post, D. and Zandman, F.
“Accuracy of birefringent-coating Method for Coatings of Arbitrary Thickness”
Exp. Mech. v1, n1, p21-32 (1961)

-
- [56] Quan, C., Brysanton-Cross, P. J. and Judge, T. R.
"Photoelasticity Stress Analysis Using Carrier Fringe and FFT Techniques"
Optics and Lasers in Eng., v18, p79-108, (1993)
- [57] Ramesh, K., and Deshmukh, S. S.
"Automation of White Light Photoelasticity by Phase-shifting Technique
Using Colour Image Processing Hardware"
Optics and Lasers in Eng., v28, p47-60, (1997)
- [58] Ramesh, K., and Deshmukh, S. S.
"Three Fringe Photoelasticity - Use of Colour Image Processing Hardware
to Automate Ordering of Isochromatics"
Strain., v32, p79-86, (1996)
- [59] Ramesh, K., and Ganapathy, V.
"Comparative Performance Evaluation of Various Phase-shifting Algorithm
in Photoelasticity"
Proc. ICAME, Bangalore, p1191-1204, (1995)
- [60] Ramesh, K., Ganesan, V. R. and Mullick, S. K.
"Digital Image Processing of Photoelastic Fringe--A New Approach"
Exp. Tech., v15n5, p41-46 (1991)
- [61] Ramesh, K., and Mangal, S. K.
"Automation of Data Acquisition in Reflection Photoelasticity by Phase-
shifting Methodology"
Strain., v33, p95-100, (1997)
- [62] Redner, A. S.
"Photoelastic Measurements by Means of Computer Assisted Spectral
Contents Analysis"

-
- Exp. Mech.*, v25, p148-153 (1985)
- [63] Redner, A. S.,
"Photoelastic measurements by means of computer assisted spectral contents analysis"
Proc. 5th Int. Conf. on Exp. Mech., Montreal, p 421-7 (1984)
- [64] Redner, A. S.
"A New Automatic Polariscopes System"
Exp. Mech., v14, p486-491 (1974)
- [65] Redner, S.
"New Oblique Incidence Method for Direct Photoelastic Measurement of Principal Strains"
Exp. Mech., v3n3, p67-71 (1963)
- [66] Robert, A. J.,
"New methods in photoelasticity"
Exp. Mech., v7, p224-232 (1967)
- [67] Sanford, R. J. and Iyengar, V.
"The Measurement of the Complete Photoelastic Fringe Order Using a Spectral Scanner"
Proc. 1985 SEM Spring Conf. on Exp. Mech., p160-168 (1985)
- [68] Sapaly, J.,
Contribution a l'etude de la photoextensometric statique et dynamique.
Th. Sc. Phys. series A3761 (1961), 4612.
- [69] Sarma, A.V.S.S.R., Pillai, S. A., Subramanian, G., & Varadan, T. K.,
"Computerised Image Processing for Whole-Field Determination of

- Isoclinics and Isochromatics”
Exp. Mech., v31, p24-29 (1992)
- [70] Seguchi, Y., Tomita, Y. and Watanabe, M.
"Computer-aided Fringe-pattern Analyzer--A Case of Photoelastic Fringe"
Exp. Mech., v19, p362-370 (1979)
- [71] Theocaris, P. S. and Dafermos, K.
"A Critical Review on Thickness Effect of Birefringent Coatings"
Exp. Mech. v4, n9, p271-276 (1964)
- [72] Theocaris, P. S. and Gdoutos, E. E.
"Matrix Theory of Photoelasticity"
Springer Series in Optical Sciences, Vol 11 (1979)
- [73] Tsai, R. Y.
"An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision"
CH2290-5-86, IEEE, p364-374, (1986)
- [74] Umezaki, E., Tamaki, T., and Takahashi, S.,
"Automatic stress analysis from photoelastic fringes due to image processing using a personal computer"
Proc. Soc. Photo., 504 (1984), 127-134.
- [75] Umezaki, E. Tamaki, T. and Takahashi, S.
"Automatic Stress Analysis of Photoelastic Experiment by Use of Image Processing"
Exp. Tech., v13n12, p22-27 (1989)
- [76] Voloshin, A. S. and Burger, C. P.

- "Half-fringe Photoelasticity - A New Approach to whole-field Stress Analysis"
Exp. Mech., v23n3, p304-313 (1983)
- [77] Voloshin, A. S. and Redner, A. S.
"Automated Measurement of Birefringence: Development and Experimental Evaluation of the Technique"
Exp. Mech., v28, p252-257 (1989)
- [78] Wang, Z. F., and Patterson, E. A.,
"Use of Phase-stepping with Demodulation and Fuzzy Sets for Birefringence Measurement"
Optics and Lasers in Engineering, 22, (1995): 91-104
- [79] Wang, Z. F.
"Development of an Automatic System for Photoelastic Analysis "
PhD Thesis, University of Sheffield, Nov. 1991
- [80] Zandman, F.
"Photoelastic Coatings"
Soc. Exp. Stress Analysis, Iowa State University Press, Iowa 1970
- [81] Zandman, F., Render, S. S. and Reigner, E. I.
"Reforcing Effects of Birefringent Coatings"
Exp. Mech., v2n2, p56-64 (1962)
- [82] Zandman, F. and Wood, M. R.
"Photostress"
Product Eng. (London) v27n9, pp167-178, (1956)

Appendices

Appendix A

Published Papers Arising from the Thesis

1. Patterson, E. A., Ji, W. and Wang, Z. F.
“On Image Analysis for Birefringence Measurement in Photoelasticity”
Optical and Lasers in Engineering, 28 (1997), p17-36
2. Ji, W and Patterson, E. A.
“Simulation of Errors in Automated Photoelasticity”
Accepted for publication in *Journal Exp. Mech.*, in press.
3. Patterson, E. A. and Ji, W
“Simulation of Errors in Automated Photoelasticity”
Proc. 1997 SEM Spring Conf. on Exp. Mech., p366-367 (1997)
4. Ji, W and Patterson, E. A.
“Hardware Error Analysis in Automated Phase-stepping System”
Proc. 1995 BSSM Conf. on Automated Strain Measurement & Analysis, p40-42
5. Ji, W and Patterson, E. A.
“A simulation of a Polariscope for Error Analysis in Automated Photoelasticity”
Abstract accepted by *11th International Conf. on Exp. Mech.*, August 1998,
Oxford.

Appendix B

Distribution of Multivariate Normal Density Function

The appendix describes the distribution of multivariate normal density function. First consider the bivariate case.

Let X and Y be independent normal random variables with zero mean and variance σ_x^2 and σ_y^2 . The probability density functions of X and Y are expressed as $f_X(x)$ and $f_Y(y)$, respectively.

$$f_X(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma_x} \exp\left(-\frac{x^2}{2\sigma_x^2}\right) \quad (\text{A-1})$$

$$f_Y(y) = \frac{1}{\sqrt{2\pi} \cdot \sigma_y} \exp\left(-\frac{y^2}{2\sigma_y^2}\right) \quad (\text{A-2})$$

Define another random variable Z as

$$Z = aX + bY \quad (\text{A-3})$$

where a and b are given constants.

Now for convenience, the two variables U and V are introduced, and defined as

$$U = aX \quad (\text{A-4})$$

$$V = bY \quad (\text{A-5})$$

Equation Eq.(A-4) can be reversed and X can be rewritten as a function of U

$$X = U/a \quad (\text{A-6})$$

The probabilities that X and U lie within corresponding differential regions must be equal. That is,

$$P(X \text{ is between } x \text{ and } x+dx) = P(U \text{ is between } u \text{ and } u+du)$$

or

$$f_X(x) = \begin{cases} \int_u^{u+du} f_U(t) dt, & \text{for } du \text{ positive} \\ -\int_u^{u+du} f_U(t) dt, & \text{for } du \text{ negative} \end{cases} \quad (\text{A-7})$$

The differential equivalent of Eq.(A-7) is

$$f_X(X)dX = f_U du |u| \quad (\text{A-8})$$

where dx is assumed to be positive. From Eq.(A-8),

$$f_U(u) = \left| \frac{dx}{du} \right| \cdot f_X(x) = \left| \frac{1}{a} \right| \cdot f_X\left(\frac{u}{a}\right) \quad (\text{A-9})$$

Substituting Eq.(A-1) into Eq.(A-9) and yields:

$$f_U(u) = \frac{1}{|a|} \cdot \frac{1}{\sqrt{2\pi} \cdot \sigma_x^2} \exp\left[-\frac{(u/a)^2}{2\sigma_x^2}\right] = \frac{1}{\sqrt{2\pi} \cdot (a\sigma_x)^2} \exp\left[-\frac{u^2}{2(a\sigma_x)^2}\right] \quad (\text{A-10})$$

It can now be seen that transforming a zero mean normal random variable with a simple scale factor yields another normal random variable with corresponding scale change in its standard deviation. That is,

$$\begin{aligned} U &\sim N(0, (a\sigma_x)^2) \text{ where } \sigma_U^2 = (a\sigma_x)^2 \\ V &\sim N(0, (b\sigma_y)^2) \text{ where } \sigma_V^2 = (b\sigma_y)^2 \end{aligned} \quad (\text{A-11})$$

Now consider the probability density function for the sum of U and V , Which will again be denoted by Z .

It is apparently that the probability of between z and $z+dz$ should be equal to that of U and V lying within the differential strip. Therefore,

$$P(z \leq Z \leq z + dz) = P(u \text{ and } v \text{ lie in differential strip}) = \iint_{\text{Diff. strip}} f_U(u) f_V(v) du dv \quad (\text{A-12})$$

But within differential strip, U is constrained to V according to

$$U = Z - V \quad (\text{A-13})$$

Also, since the strip is only of differential width, the double integral of Eq.(A-13) reduces to a single integral. Choosing U as the variable of integration and noting that $dU = dZ$ leads to

$$P(z \leq Z \leq z + dz) = \left[\int_{-\infty}^{+\infty} f_U(u) f_V(Z - u) du \right] dZ \quad (\text{A-14})$$

The quantity within the brackets is the desired probability density function for Z .

Thus,

$$f_z(Z) = \left[\int_{-\infty}^{+\infty} f_U(u) f_V(Z - u) du \right] dZ \quad (\text{A-15})$$

Noting that the integral on the right side of Eq.(A-15) is a convolution integral.

Thus, from Fourier transform theory, Eq.(A-15) can be rewritten as:

$$\mathfrak{R}(f_z) = \mathfrak{R}(f_U) \cdot \mathfrak{R}(f_V) \quad (\text{A-16})$$

Where $\mathfrak{R}(\ast)$ denotes the Fourier transform of (\ast) .

Using Fourier transform tables, the transforms of f_U and f_V are expressed as:

$$\begin{aligned} \mathfrak{R}(f_U) &= \exp(-\sigma_u^2 w^2 / 2) \\ \mathfrak{R}(f_V) &= \exp(-\sigma_v^2 w^2 / 2) \end{aligned} \quad (\text{A-17})$$

where w represents the frequency domain. Forming their product yields

$$\mathfrak{R}(f_U) \cdot \mathfrak{R}(f_V) = \exp[-(\sigma_u^2 + \sigma_v^2)w^2 / 2] \quad (\text{A-18})$$

Then the inverse gives the desired f_z :

$$f_z(z) = \mathfrak{R}^{-1} \exp[-(\sigma_u^2 + \sigma_v^2)w^2 / 2] = \frac{1}{\sqrt{2\pi \cdot (\sigma_u^2 + \sigma_v^2)}} \exp[-z^2 / 2(\sigma_u^2 + \sigma_v^2)] \quad (\text{A-19})$$

Note that the density function for Z is also normal in form and its variance is given by

$$\sigma_z^2 = \sigma_u^2 + \sigma_v^2 \quad (\text{A-20})$$

Substituting Eq.(A-11) into Eq.(A-20) and then

$$\sigma_z^2 = (a\sigma_x)^2 + (b\sigma_y)^2 \quad (\text{A-21})$$

Because the summation of any number of random variables can always be thought of as a sequence of summing operations on two variables: it should be clear that summing any number of independent normal random variables leads to a normal random variable.