3049

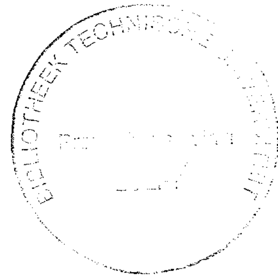
TR3051

Two-camera shearing speckle interferometer for real-time deformation measurements

Cover design by Bob Veldt

Two-camera shearing speckle interferometer for real-time deformation measurements

Proefschrift



ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus Prof. dr ir J. Blaauwendraad,
in het openbaar te verdedigen ten overstaan van een commissie,
door het College voor Promoties aangewezen,
op dinsdag 9 december 1997 te 10:30 uur
door

Aldegonda Lucia WEIJERS

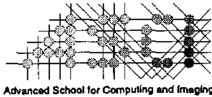
natuurkundig ingenieur
geboren te Bergh

Dit proefschrift is goedgekeurd door de promotor:

Prof. dr ir H.J. Frankena

Samenstelling promotiecommissie:

Rector Magnificus,	voorzitter
Prof. dr ir H.J. Frankena,	Technische Universiteit Delft, promotor
Dr ir H. van Brug,	Technische Universiteit Delft, toegevoegd promotor
Prof. dr ir P.G. Bakker,	Technische Universiteit Delft
Prof. dr ir J.J.M. Braat,	Technische Universiteit Delft & Philips Research Laboratorium
Prof. dr J. Greve,	Universiteit Twente
Prof. dr I.T. Young,	Technische Universiteit Delft
Dr ir A.A.M. Maas,	Technisch Physische Dienst TNO-TU Delft



This work was carried out in graduate school ASCI.
ASCI dissertation series number 26.

The work described in this thesis was financially supported by the commission of the EC under the BRITE-EURAM-II BRE2-CT92-0313 project and EUCLID-CEPA3-RTP3.1. project.

ISBN 90-9011166-2

Copyright © 1997, by A.L. Weijers.
All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means: electronic, mechanical, photocopying, recording or otherwise, without the prior written permission of the author A.L. Weijers, Faculty of Applied Physics, Delft University of Technology, P.O. Box 5046, 2600 GA Delft, The Netherlands.

Summary

Introduction

Composite materials are increasingly used as lightweight construction materials. They allow the fabrication of much lighter constructions than with conventional materials (e.g. aluminum), while matching their strength and stiffness. Moreover, composites show improved properties such as low corrosion and low fatigue. This latter is especially important for the aircraft industry because aluminum in aircraft degrades due to the strong load changes and temperature changes during take-off and landing. Besides offering many attractive properties, composites also introduce new problems. During fabrication or use, defects like cracks, inhomogeneities and delaminations may occur and grow. Therefore, it is important to detect these defects in an early stage, before they affect the performance of the structure.

Since the aircraft industry was one of the first to apply composites, it now faces the problem of aging composite parts. Consequently, it needs to inspect its aircraft fleet frequently to operate it safely. Key to a continuing and expanding use of composites in the (aerospace) industry is the ability to detect surface and sub-surface defects in these structures. Because currently available inspection methods are unsatisfactory. Therefore, there is a growing interest in finding new methods that can reveal possible defects quickly and reliably in industrial environments.

This thesis

The real-time phase-stepped shearing interferometer described in this thesis can contribute to a safe operation of the aircraft fleet. The development of this system was therefore directed primarily towards the use in aircraft industry. The newly developed shearography system allows fast, non-destructive inspection of aircraft structures. It is much less sensitive to external disturbances than more conventional interferometers, making the system more suitable for in situ measurements. The system has been successfully used for defect detection in a relatively hostile environment and can therefore be used to supplement the existing non-destructive testing (NDT) techniques.

The developed system measures the deformation of diffusely reflecting objects at video speed, i.e. 25 times per second. It uses two CCD cameras to record four interferograms simultaneously. The four interferograms are generated in separate branches of the interferometer by two specially designed Savart elements. In each of the two branches, two interferograms are recorded side by side on a single camera. This way, a compact and rigid

mechanical set-up could be realized. The Savart elements are simultaneously used as beam splitters and analyzers. Moreover, they are used to align the interferograms on the CCD cameras.

Defect detection is made possible by applying forces to the object under inspection (e.g. by heating it), such that the defects show up in the form of a deformation irregularity at the object surface. Because the optical set-up is connected to a real-time image processing system, phase changes (representing surface deformations) of an object can be monitored in real time. The resultant phase change pattern approximately describes the slope change of the surface in the shear direction.

Thanks to the used shearing technique and the short recording time (40 ms) of a phase image, the system is less sensitive for measurement errors that are introduced by time-dependent perturbations (like air flow, temperature changes or vibrations) than temporal phase-stepped shearing interferometers, which record the phase-stepped patterns successively in time. Another advantage of shearing interferometry concerns the requirements on the laser's coherence length. As compared to a conventional two-beam interferometer, the demands on the coherence length of the light source are relieved, thus enabling the use of semiconductor lasers. The developed system employs such a semiconductor laser, which makes the system easy to handle and to transport. The power of the used laser is limited to 1 W, which restricts the inspection area to approximately 25×35 cm^2. A 1 W laser should normally be sufficient to inspect a 0.25 m^2 area in a single recording, but due to aberrations, generated by imaging through the Savart elements, the inspection area is decreased. Other error sources in the system are decorrelation and incorrect positions or characteristics of the polarizing optical elements. The total measurement accuracy is maximally $\lambda/22$, which leads to a smallest measurable deformation of 22 nm.

Delft, A.L. Weijers

Contents

Summary v

Contents vii

1 Introduction 1
1.1 Background 1
1.2 Non destructive testing techniques 3
1.3 Purpose of the research 9
1.4 Outline of this thesis 10
1.5 References 10

2 Theoretical aspects of speckle interferometry 13
2.1 Basic principles of interferometry 13
2.2 Speckles 14
2.2.1 Speckle statistics 15
2.2.2 Speckle size 18
2.2.3 Optimal speckle size for speckle interferometry 20
2.3 Holographic interferometry 23
2.4 Electronic speckle interferometry 23
2.5 Shearing interferometry 25
2.6 Quantitative analysis of the interference phase distribution 27
2.6.1 Intensity based analysis methods 28
2.6.2 Phase sampling analysis methods 29
2.6.2.1 Time domain methods 31
2.6.2.2 Space domain methods 32
2.7 Summary 35
2.8 References 35

3 Description of the system 39
3.1 Description of the setup 39
3.2 The illumination and recording systems 49
3.2.1 Semiconductor laser 50
3.2.2 CCD cameras 51
3.3 The shearing head 55
3.4 The imaging system 58
3.4.1 The objective lens 59
3.4.2 The Kepler telescope 61
3.4.3 The numerical apertures 66
3.5 The Savart elements 69
3.5.1 Light propagation in calcite 70
3.5.2 Beam splitting by Savart elements 71
3.6 Mechanical design criteria of the prototype 73
3.7 References 75

4 Image processing **77**
4.1 Introduction 77
4.2 Real time image processing 79
4.2.1 Image splitting 80
4.2.2 Image reversal 81
4.2.3 Phase computation 82
4.2.4 Spatial filtering 83
4.3 Image post processing 87
4.3.1 Phase unwrapping 87
4.4 References 88

5 Accuracy analysis **89**
5.1 Error classification and analysis 89
5.1.1 Errors caused by optical elements 90
5.1.2 Errors caused by decorrelation 100
5.1.3 Errors caused by recording and computation system 100
5.1.4 Ambient disturbances 101
5.2 Polarization ray tracing 102
5.2.1 Refractive index 103
5.2.2 Maxwell theory applied to uniaxial media 104
5.2.3 Algorithms 105
5.3 The camera alignment procedure 107
5.3.1 Using Savart elements for alignment 109
5.3.2 Alignment procedure 113
5.4 Image formation through Savart elements 116
5.4.1 Polarization ray tracing for imaging analyis 117
5.4.2 Unresolved speckle patterns 119
5.5 Measured accuracy 123
5.6 References 125

6 Application of the system for defect detection **127**
6.1 Shearographic detection of defects 127
6.2 Defect detection under laboratory conditions 128
6.3 Application of shearography in hostile environments 132
6.4 References 137

7 Conclusions and suggestions **139**
7.1 Conclusions 139
7.2 Suggestions 140

Symbol list **143**

Acknowledgments **147**

Shearing spikkelinterferometer met twee camera's voor vervormingsmetingen met videosnelheid (samenvatting) **149**

Biography **151**

Chapter 1

Introduction

This chapter describes the background of this thesis. Section 1.2 gives a brief introduction into the current methods of non-destructive testing and their characteristics. The new interferometric system, that has been developed during the research described in this thesis, can supplement the existing methods in the field of non-destructive testing. Next, in Section 1.3, the choice for developing a real-time phase-stepped shearing speckle interferometer is motivated. Finally, the structure of this thesis will be outlined in Section 1.4.

1.1 Background

For many decades, aluminum has been one of the most popular material to build lightweight constructions. Since the sixties, fiber reinforced plastics, such as polyester and carbon composites, have obtained a position as construction materials. By applying fiber reinforced plastics, constructions can be made much lighter than with conventional materials, matching its strength and stiffness. Moreover, fiber reinforced plastics show improved properties (e.g. low corrosion and less fatigue)[1]. This latter is especially important for the aircraft industry because the aluminum in aircraft degrades (small cracks appear) due to the strong load changes and temperature changes during take off and landing.
Composites are mostly used in the aeronautical and space industry, but nowadays they are also being applied in sports articles (golf clubs, tennis rackets, bicycles, skis etc.), naval vessels, tooling, cars, medical devices (prosthetic and orthopedic devices), LPG tanks and artwork[2,3].

Despite the many attractive properties of composites, their use also generates new problems. During fabrication or use, defects like cracks, inhomogeneities and delaminations may occur and grow. These defects can upset the mechanical balance of the composite structures. Therefore, it is important to detect these defects in an early stage, before they affect the performance of the structure.
Since the aircraft industry was one of the first to apply composites, it now faces the problem of aging composite parts. Consequently, they need to inspect their aircraft fleets frequently to operate it safely. The work presented in this thesis can contribute to such a safe operation and it is therefore primarily directed towards its use in aircraft industry. However, the applicability of real-time phase-stepped shearography is in no way limited to aerospace constructions or even composites.

Inspection programs are well established for aircraft fabricated from conventional materials. The pertinent techniques are known as Non-Destructive Testing (NDT) techniques, also indicated as Non-Destructive Inspection (NDI) or Non-Destructive Evaluation (NDE) techniques. Inspecting a typical metal aircraft structure is essentially a matter of searching for cracks or flaws, which, if left undetected or permitted to grow, could result in structural failure. By way of contrast, inspection of composite structures is a much more complex process. The major problem with the laminated composite construction materials is that some defects (e.g. delaminations, see Fig. 1.1) can cause a strong reduction in stiffness and strength, but are often invisible from the outside.

Fig. 1.1 *Schematical representation of two types of delamination in a honeycomb sandwich: (a) a delamination in the skin and (b) a delamination of honeycomb material.*

Furthermore, laminated structures are more sensitive to impact damages (see Fig. 1.2) than conventional materials. The impacts can damage the surface but the internal damage can be even bigger since impacts generally induce delaminations. Moreover, initial delaminations tend to grow. Delaminations can arise at different stages of the life of an airplane: during manufacturing (incomplete bonds or inclusions), during assembling (a dropped tool) and during servicing (caused by structural overload or impacts from objects like birds, (hail) stones or lightning).

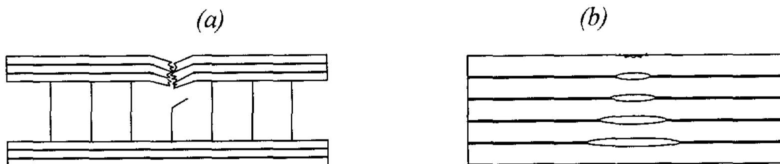


Fig.1.2 *Schematical representation of (a) an impact damage and (b) an impact damage which induced delaminations.*

Key to a continuing and expanding use of composites in the (aerospace) industry is the ability to detect surface and sub-surface defects in the structures, both during manufacturing and in-service inspection. Unfortunately, currently available inspection methods are unsatisfactory. They can not detect sub-surface defects or are time-consuming and therefore expensive. Consequently, there is a growing interest in finding new methods that can reveal possible defects quickly and reliably.

In Section 1.2, a short overview of the most important NDT techniques[4-6] is given, including their advantages and disadvantages.

1.2 Non-destructive testing techniques

The elementary method of NDT is **visual inspection**, with or without optical aids. Surface-breaking defects can be detected by careful direct visual inspection. Only relatively inexpensive equipment is required. Optical aids include magnifiers, microscopes, telescopes and endoscopes. Visual inspection with cameras or photography is also being used. The success of visual inspection strongly depends on the surface condition, the lighting arrangement and the operator. Surface preparation, such as cleaning and etching, is often applied.

One step up from visual inspection is **liquid penetrant testing**. The principle of the method is that the liquid penetrant is drawn into the surface-breaking defect by capillary action and excess surface penetrant is then removed. Again, defects are detected by viewing. This method solely works on flaws like cracks, laps and porosity, which must reach the surface to be detected. Also, it can not be used on dirty or very rough surfaces and surface cleaning is a vital part of the penetrant testing technique.

D-sight is another useful method for detecting surface defects. It uses the reflection of the surface to find dents or bumps. With a camera and a lamp above or below it, the surface is inspected under nearly grazing angles (typically 30 degrees). The light is reflected with a retroreflective screen that is positioned symmetrically to the camera and the lamp. A schematic drawing of a typical D-sight set-up is given in Fig. 1.3.

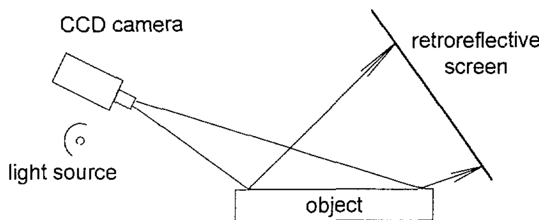


Fig. 1.3 *A typical D-sight set-up.*

The light from the source is reflected by the object's surface towards a screen, which is not perferctly retroreflective. Each ray is broadened into a small cone of light. Upon reflection from the surface for the second time, the light is recorded with the CCD camera. Dents or bumps change the reflection direction, thus causing light and dark areas in the image. With this technique, no image processing is required. However, the method is only suitable on shiny surfaces (which is no problem for civil aircraft, but already more difficult with military aircraft). If surfaces are not shiny, special treatment of the surface is necessary. The main disadvantage, is the limited types of defects that can be detected with this method. It is only suitable for surface defects, as is the case with visual inspection.

Since detecting sub-surface damages is very important with the composite materials, the inspection methods mentioned above are only partly appropriate for non-destructive testing of composites.

Tap testing is performed by tapping on the structure with a blunt instrument, often a large coin, and to listen. In current practice, tap testing is used to locate possible disbonds, delaminations and water in honeycomb core cells, but tap testing has its disadvantages. It only detects certain types of defects close to the surface. In addition, it is difficult to apply in noisy environments and it requires careful surface preparation. Certain coatings, aluminum foils, pressure sensitive tapes and similar films have to be removed from a surface before a tap test can be performed. Finally, it is even more operator sensitive than visual inspection.

X-ray inspection can provide a high-resolution measurement of density variations within a material. The technique is based upon showing changes in the material's transmission characteristics, which indicate possible anomalies in the internal configuration of the part. It may not detect a delamination, crack or debond if no (or only a very small) gap is formed. In addition, the sensitivity of the systems is influenced by the presence of other high-density material in front or behind the region of interest. Moreover, the systems tend to be large, cumbersome and have no real-time performance. The technique is not commonly used, mainly because it is impractical and causes safety risks for the operators due to radiation hazard.

Eddy current techniques[6] fall in the category of electrical methods and is only suitable for conducting materials. In eddy current testing, a coil carrying an AC current is placed close to the specimen surface. The current in the coil generates circulating eddy currents in the specimen close to the surface and these in turn affect the current in the detecting coil by mutual magnetic induction. In many cases the same coil is used both to excite the eddy currents and to detect the corresponding magnetic field.
Flaws and material variations in the specimen affect the strength of the eddy currents. The presence of flaws, etc. is therefore measured by electrical signal changes (voltage or phase changes) in the detecting coil. Since the eddy currents are concentrated at the outer surface of the specimen, only surface or near-surface defects can be detected. Most eddy current instruments require a calibration on a set of test specimens. The output signals are very complicated. Equipment varies from simple portable meter-read-out instruments, to more complex oscilloscope read-out instruments displaying both phase and voltage. Recently, the outputs have been digitized to produce fully-automated computer programmed equipment. The systems have a small inspection area (typically with a 20 mm diameter) and are mainly used to inspect areas around rivets for cracks and corrosion in metals. Two-dimensional scanning is required to inspect large areas, making this technique time-consuming.

Another traditional NDT technique is the **ultrasonic measurement technique,** which uses ultrasonic waves, transmitted from a small probe and detected by the same or other probes (see Fig. 1.4).

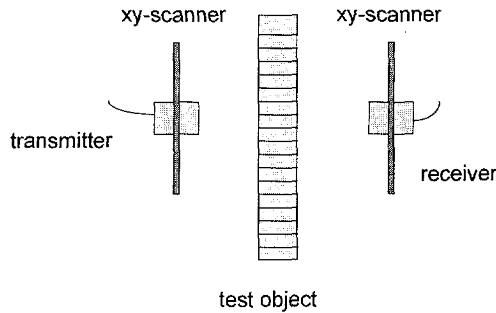


Fig. 1.4 *An ultrasonic measurement set-up.*

Usually, pulsed beams of ultrasound are used and in the simplest instruments a single, hand-held probe is placed upon the specimen surface. An oscilloscope display (amplitude versus time) shows the time that it takes for an ultrasonic pulse to travel to a reflector in terms of distance across the oscilloscope screen - the so-called A-scan display. Many sequentially recorded A-scans can be combined to make a so-called C-scan, using (automated) mechanical scanners which plot the position of the probe and the ultrasonic information (e.g. time of flight or ultrasonic attenuation, see Fig. 1.5).

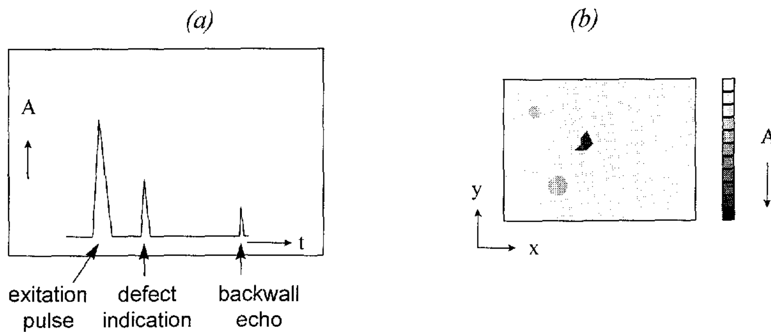


Fig. 1.5 *Ultrasonic data representations (a) A-scan and (b) C-scan.*

Transmission ultrasonic techniques require access from both sides of the part. The double transmission technique, also called the reflector plate technique, and pulsed ultrasonic inspection have the advantage of being single contact methods: they need access from one side only. The technique requires experienced operators and a liquid couplant, such as glycerin or water, between source and probe surface. Also the part must be clean and smooth. Rough surfaces tend to interfere with the contact probes, degrading the signal transmission. A common technique with ultrasonic measurements is to immerse the specimen in a large tank and use a stand-off probe with a mechanized movement. This method requires dismantling of the test part. New developments towards air-coupled ultrasonic measurements are being developed. However, they remain time-consuming methods, since using a point detector and two-dimensional scanning of the surface is required.

Thermography is a technique of obtaining an image of the heat distribution over the surface of an object. Usually, a special camera with an infrared sensitive detector and a lens which transmits infrared radiation are used (see Fig. 1.6) to display temperature variations in the object.

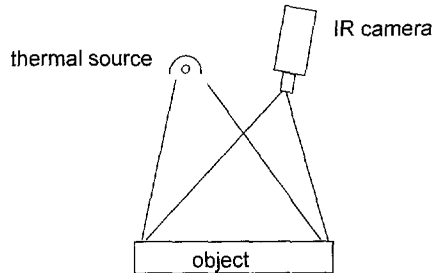


Fig. 1.6 *Thermography set-up.*

Defect detection relies on the differences between the thermal conductivity of defect free areas and those with defects. Temperature variations as small as 0.2° C can be detected. The method is e.g. used to examine Carbon Fiber Re-enforced Plastic (CFRP) laminates and honeycomb-bonded panels. One limitation of thermography is that the infrared emission depends also on the surface condition of the specimen (surface emissivity), e.g. paint influences the signal. New techniques, like lock-in thermography, are being developed to reduce this problem. Lock-in thermography can locate defects such as delaminations or inclusions up to several millimeters below a sample surface. It uses a harmonically modulated infrared source. Defects generate reflections that interfere with the incident radiation to produce a kind of standing wave pattern. This technique can be used to inspect large surfaces at once, but it is a rather slow method.

Alternative techniques for performing the much needed defect detections that have major advantages over the above listed more conventional approaches are **interferometric techniques**. Interferometric techniques produce three-dimensional patterns of fringes, whose density and form depend on the tested structure, the wavelength of the used light and the geometry of the interferometer. Interferometric methods visualize or measure object deformations in terms of changes in optical path lengths (or, equivalently, the phase) of the interfering beams on its way from the light source via the object surface to an appropriate detector. Detection of optical path lengths differences much smaller than the wavelength of the light has become possible. The sensitivity gained by comparing the optical path length changes with the wavelength of the laser light used makes interferometry an ideal measuring tool for a manifold of physical quantities, such as strain, stress, deformations, refractive index variations in solids, fluids gases, etc. Also, full-field, non-contact, non-destructive tests are possible. Both surface and sub-surface defects can be detected. An important advantage, shared with many other optical techniques, is that it allows large test areas to be examined in a single measurement using an imaging system and a powerful laser. There is no need for direct contact between the object and the inspection instrument.

Non-destructive inspection using interferometry is based upon the fact that, after loading, components with faults show different deformations than unflawed parts. Thus, defects can be identified by anomalies in the fringe pattern. Although the technique measures surface deformation, internal damage can also be detected because defects, depending on their proximity to the surface, can influence the surface deformation and hence can be detected. The principle of defect detection is sketched in Fig. 1.7.

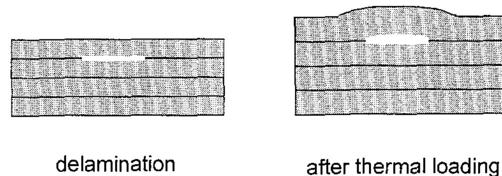


Fig. 1.7 *Detection of a delamination between the first and second layer in a laminate structure. The expansion of the material is exaggerated.*

A fault specific deformation in the sub-micrometer range is sufficient to be detected by the high accuracy of interferometric methods. The required amount of loading is small, e.g. with thermal stressing typically a few degrees change in temperature of the surface of the inspected sample is sufficient. The success of interferometry as a non-destructive inspection technique depends on the effectiveness of the chosen loading technique. Usually one or more of the following loading techniques are used in interferometric inspections:

- thermal loading
- pressure or vacuum loading
- mechanical loading
- vibrational loading

The type of loading to be employed is generally dictated by the geometry of the components and the type of defect being examined. The most practiced methods for in-service testing are thermal and vacuum loading. Vacuum loading usually incorporates small inspection areas (typically in the range of 0.03 m^2), where thermal loading is easily applied to larger areas (typically up to 0.4 m^2). In production testing of a single type of object, larger areas can be covered with both loading techniques if dedicated systems are used.

The oldest technique in the interferometric category is **holographic interferometry**[6-8], based upon holographic recording of interference patterns. Holography can provide a very high resolution but its use is limited to laboratory environments. Thanks to the development of electronic cameras, **electronic speckle pattern interferometry**[8-11] (ESPI) could be developed. This technique has a lower spatial resolution than holographic methods, but ESPI made inspections much easier and faster than the time-consuming process of developing and positioning holographic plates. ESPI can be used for all kinds of samples with a rough surface. If the surface is too shiny, applying a thin layer of removable diffusing powder on the surface eliminates this problem. The speckle pattern resulting from the illumination of a rough surface with laser light is the basis of this technique, that generates maps of deformations with (sub)micrometer-scale amplitudes. Both holographic interferometry and ESPI can measure

object deformations in real time in the form of fringe pattern deformations. For some time now, both techniques have been used as non-destructive inspection techniques at many manufacturing facilities. However, their full potential as control tools cannot be realized totally because of their high sensitivities. Environmental disturbances degrade the quality of the images. The entire system, including the test object, must be isolated from ambient disturbances, such as vibrations and air flow. Such isolation is impractical in large scale testing. Aircraft wings, for example, need to be tested at the aircraft or otherwise the major benefits of this NDT technique are lost.

To overcome this problem another technique, called shearography, is explored. **Shearography**[8,12-14] is less sensitive to external disturbances than ESPI, which makes it more useful in a factory or airport environment. Instead of using a separate reference beam, it splits the light beam and recombines the two parts such that they create overlapping and (laterally) displaced images on a camera, thus recording the resulting interferogram. A typical shearing interferometer is shown in Fig. 1.8.

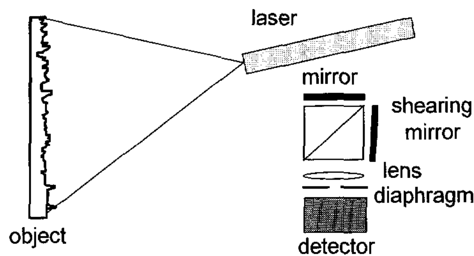


Fig. 1.8 *A typical shearing speckle interferometer.*

Shearography provides information about deformation gradients, whereas holographic interferometry and ESPI measure deformation amplitudes. Yet, shearography also reveals surface and sub-surface defects. In Fig. 1.9b, a typical shearographic result of an inclusion in the laminate skin of a composite sandwich structure (Fig. 1.9a) is depicted.

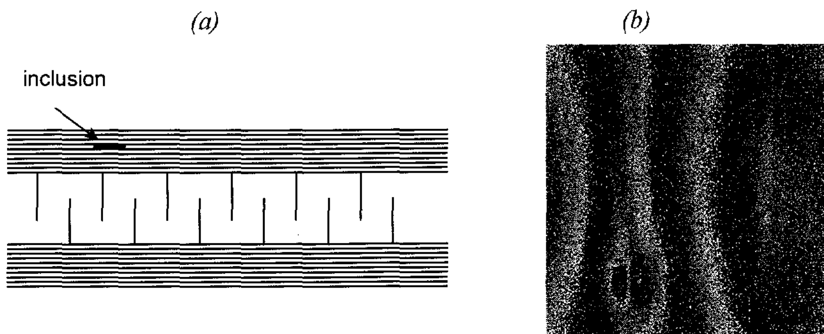


Fig. 1.9 *(a) A 15 mm inclusion in a sandwich structure and (b) the pertinent shearographic result.*

The size of the damage can usually be estimated from the size of the anomalous shearographic fringe pattern. Care must be taken to account for the effect that larger shears give a larger apparent defect because the defects expand in the shearing direction as the shear is increased. The minimal defective area discernible by this technique is typically in the order of a few millimeters.

The first industrial applications of shearography show the need for a better quality of the test results. For this reason, phase stepping techniques[15-17] have been added, greatly increasing the sensitivity of the system, and therefore increasing the detectability of defects. Phase stepping is a powerful technique for accurate quantitative determination of surface deformations. Instead of a single interference pattern, phase stepping interferometry uses at least three registrations of these patterns. These patterns are usually recorded successively in time with a CCD camera and between the recordings the phase difference between the object and reference beams is changed. Unfortunately, the advantage of the improved quality of the test results is partly canceled in industrial applications due to the loss of real-time operation. The presently used phase stepping systems have a high sensitivity for environmental disturbances. Time-dependent perturbations, e.g. air flow, temperature currents and vibrations, disturb the measurements since subsequent recordings are affected differently. Besides this, the usual phase stepping shearography systems are unsuitable to monitor time-dependent phenomena.

In order to fully utilize the benefits of phase stepping without introducing problems associated with loss of real-time operation, we propose a real-time phase-stepped shearography system[18] based upon the simultaneous recording of the phase-stepped interference patterns. This system has the potential to result in a fast, reliable and high-quality (sub)surface defect detection system, capable of performing large area inspections in noisy industrial ambiances.

1.3 Purpose of the research

The purpose of the research described in this thesis was to develop a real-time phase-stepped shearing speckle interferometer with reduced sensitivity for external disturbances and to realize a mobile, robust system for use as an NDT tool in an industrial environment. The goal of the project was to inspect an area of 0.25 m^2 in one view and to detect defects with a minimum size of 490 mm^2. Preliminary investigations[18] with a laboratory system demonstrated the applicability of shearography combined with real-time phase stepping for defect detection.

Real-time phase stepping can be performed by changing the polarization properties of light and recording several different interference patterns simultaneously by means of introducing several branches in an interferometer. The preferred number of phase-stepped interferograms depends on the desired accuracy. As a rule of thumb it can be stated that the larger the number of interferograms, the more accurate the measurements will be. On the other hand, if more interferograms are used, more branches and hence more cameras are required in the optical setup and thus the optical alignment becomes more complicated.

The initial goal of the research described here was to arrive at a four-camera phase-stepped shearing interferometer[18]. Only three cameras are needed to calculate the phase distributions, but by using four cameras the system is less sensitive to source light fluctuations. During the project, a two-camera system recording four phase-stepped images promised to be advantageous, since the requirements on the mechanical built-up of the system could be strongly reduced. Therefore, we focused our attention on developing such a system. Much effort was put into designing and analyzing the functional component of the interferometer, capable of generating two phase-stepped images side by side on a single CCD camera.
The accuracy of the interferometer is not only determined by the number of interferograms, but also by the degree of correlation between the interferograms. Therefore, the alignment of the images is extremely important. The development of alignment procedures for two interferograms on a single CCD camera made up a considerable part of the work performed. Furthermore, a mechanical system had to be designed and realized in order to obtain a robust system for use in industrial environments. This system must also enable the alignment procedures.
Besides the development of a compact and robust system, much attention was given to select a small, powerful and portable laser. Other main tasks were the development of a real-time image processing system to calculate and display deformations of objects with video speed and the verification of the improved performance of the system under vibrational disturbances.

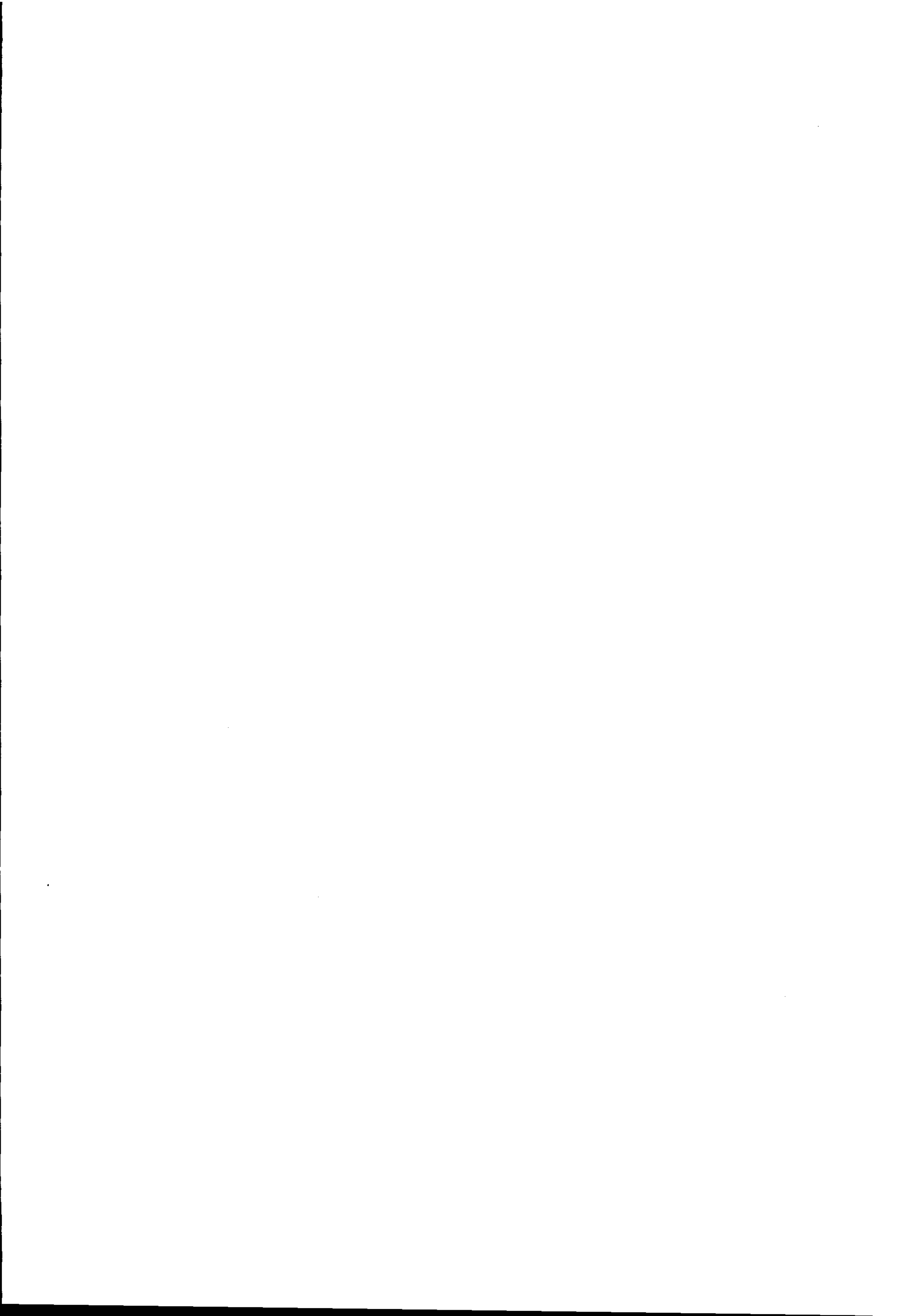
1.4 Outline of this thesis

This thesis is organized into seven chapters. The purpose of Chapter 2 is to provide the basic theory of speckle interferometry, as it is needed in the context of this thesis. Some important principles, methods and their (dis)advantages are described. Next, the optical system that has been developed is described in Chapter 3. This system comprises specially designed elements for generating two phase-stepped interference patterns side by side on a single CCD camera; the design of these elements is discussed in detail. Chapter 4 gives a general description of the image processing system, which includes the real-time image processing devices and some additional hard- and software. Subsequently, the accuracy of the system is analyzed in Chapter 5. Chapter 6 is devoted to the application of the system in a laboratory environment as well as in one with vibrational disturbances. Finally, Chapter 7 presents the conclusions and suggestions for further research.

1.5 References

1. L.A. Pilato, M.J. Michno, *Advanced Composite Materials*, (Springer Verlag, Berlin, 1994).
2. W.J. McGonnagle, *International Advances in Nondestructive Testing Vol. 9*, (Gordon and Breach Science Publishers, 1983).

3. L.A. Pilato, M.J. Michno, *Advanced composite materials*, (Springer Verlag, Berlin, 1994), pp. 157-185.
4. The British Institute of Non-Destructive Testing, *The NDT Yearbook 1989/90*, (Chartwell Press Ltd., Leicester, 1989).
5. T. Jones, H. Berger, *Nondestructive inspection of composite materials*, (Industrial Quality Inc., Gaithersburg, 1995).
6. M.J. Golis, *An introduction to nondestructive testing*, (The American Society for nondestructive testing Inc., Columbus, 1991).
7. R.K. Erf, J.P. Waters, R.M. Gagosz, F. Michael, G. Whitey, *Nondestructive holographic techniques for structures inspection*, (National Technical Information service, Springfield, 1972).
8. L.A. Pilato, M.J. Michno, *Advanced composite materials*, (Springer Verlag, Berlin, 1994), pp. 105-106.
9. J.N. Butters, J.A. Leendertz, "Speckle pattern and holographic techniques engineering metrology", Optics and Laser Technology **3**(1), pp. 26-30, (1971).
10. "Non-destructive testing of overlap shear joints using electronic speckle pattern interferometry", Optics and Lasers in Engineering **6**, pp. 79-90 (1985).
11. T. Kreis, "Speckle Metrology" in *Holographic Interferometry: Principles and methods*, (Akademie Verlag, Berlin, 1996).
12. S. Nakadate, T. Yatagai, H. Saito, "Digital speckle-pattern shearing interferometry", Applied Optics **19**(24), pp. 4241-4246 (1980).
13. A.R. Ganesan, D.K. Sharma, M.P. Kothiyal, "Universal digital speckle shearing interferometer", Applied Optics **27**(22), pp. 4731-4734 (1988).
14. M. Owner-Petersen, "Digital speckle pattern shearing interferometry: limitations and prospects", Applied Optics **30**(19), pp. 2730-2738 (1991).
15. S. Nakadate, H. Saito, "Fringe scanning speckle-pattern interferometry", Applied Optics **24**(14), pp. 2172-2180 (1985).
16. K. Creath, "Comparison of phase-measurement algorithms" in *Surface Characterization and Testing*, Proc. SPIE Vol. 680, pp. 19-28 (1986).
17. H.Z. Hu, "Polarization heterodyne interferometry using a simple rotating analyzer. 1: Theory and error analysis", Applied Optics **22**(13), pp. 2052-2056, (1983).
18. A.J.P. van Haasteren, *Real-time phase stepped speckle interferometry*, (Ph.D. Thesis, Delft university of Technology, Delft, 1994).



Chapter 2

Theoretical aspects of speckle interferometry

This chapter discusses some theoretical aspects of interferometry, as well as some basic principles regarding speckles (Section 2.2). Since there are excellent books on the principles and methods in speckle interferometry[1-5], the theory is presented only as far as it concerns the work described in this thesis. The emphasis in this chapter is on phase measurements. Therefore, holographic interferometry (Section 2.3), electronic speckle pattern interferometry (Section 2.4) and shearing interferometry (Section 2.5) are described concisely. Section 2.6 discusses several techniques for quantitative phase measurements. The advantages and limitations of the described methods are given. A summary of the most important principles and methods closes this chapter.

2.1 Basic principles of interferometry

Interference occurs if two or more coherent light waves are superimposed. Consider two planar waves, emitted by the same source, which differ in the directions $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$ and the phases ϕ_1 and ϕ_2. Both waves, having the same frequency ω and amplitude A_x, are linearly polarized in the x-direction. For the electric fields of these waves we can write

$$\boldsymbol{E}_1(\mathbf{r},t) = \begin{pmatrix} A_x \\ 0 \end{pmatrix} e^{i(\boldsymbol{k}_1\cdot \boldsymbol{r} - \omega t + \phi_1)}, \tag{2.1}$$

$$\boldsymbol{E}_2(\mathbf{r},t) = \begin{pmatrix} A_x \\ 0 \end{pmatrix} e^{i(\boldsymbol{k}_2\cdot \boldsymbol{r} - \omega t + \phi_2)}. \tag{2.2}$$

Superposition of these fields gives

$$\begin{aligned} \boldsymbol{E}_1(\boldsymbol{r},t) + \boldsymbol{E}_2(\boldsymbol{r},t) &= \begin{pmatrix} A_x \\ 0 \end{pmatrix} e^{i(\boldsymbol{k}'\cdot \boldsymbol{r} - \omega t + \phi')} \left\{ e^{i(\boldsymbol{k}''\cdot \boldsymbol{r} + \phi'')} + e^{i(-\boldsymbol{k}''\cdot \boldsymbol{r} - \phi'')} \right\} \\ &= 2\begin{pmatrix} A_x \\ 0 \end{pmatrix} e^{i(\boldsymbol{k}'\cdot \boldsymbol{r} - \omega t + \phi')} \cos(\boldsymbol{k}''\cdot \boldsymbol{r} + \phi''), \end{aligned} \tag{2.3}$$

in which $\boldsymbol{k}''=(\boldsymbol{k}_1-\boldsymbol{k}_2)/2$, $\boldsymbol{k}'=(\boldsymbol{k}_1+\boldsymbol{k}_2)/2$, $\phi''=(\phi_1-\phi_2)/2$ and $\phi'=(\phi_1+\phi_2)/2$. The exponential factor in Eq. (2.3) describes a temporally varying phase, whereas the cosine term is time independent. Thus, by taking the time average over an interval that is large as compared to the period $2\pi/\omega$, we arrive at the temporally constant intensity:

$$I(\boldsymbol{r})=\left\langle\left(\boldsymbol{E}_1(\boldsymbol{r},t)+\boldsymbol{E}_2(\boldsymbol{r},t)\right)\left(\boldsymbol{E}_1(\boldsymbol{r},t)+\boldsymbol{E}_2(\boldsymbol{r},t)\right)^*\right\rangle_t=4A_x^{\,2}\cos^2(\boldsymbol{k}''\cdot\boldsymbol{r}+\phi''), \tag{2.4}$$

in which * denotes the complex conjugate. The resulting position-dependent pattern $I(\boldsymbol{r})$ is called the interference pattern and the lines of constant intensity (or phase) are called interference fringes. From Eq. (2.4) it can be seen that the intensity is minimal where the cosine factor equals zero. At these positions, the wavefronts are said to be in anti-phase and destructive interference occurs, resulting in a dark fringe. The intensity is maximal where the cosine factor equals unity. Here, the wavefronts are in phase and we have constructive interference, resulting in a bright fringe.

Waves can only interfere if their polarizations are parallel. If the polarization directions of the two waves are orthogonal, no interference occurs and their superposition only consist of an addition of the intensities

$$I(\boldsymbol{r})=I_1(\boldsymbol{r})+I_2(\boldsymbol{r}). \tag{2.5}$$

For other orientations of the polarization directions, the electric field vector of the second wave has to be decomposed into two components: one parallel and one orthogonal to that of the first wave. The result contains an interference part (from the parallel components) as well as an addition of intensities (from the orthogonal components). Also if the light is only partially coherent such a combination of interference and intensity addition occurs. As a result of these intensity additions (and of other noise contributions) in interference patterns, the contrast in the pattern, defined by the fringe visibility[6] V

$$V=\frac{I_{\max}-I_{\min}}{I_{\max}+I_{\min}} \tag{2.6}$$

is reduced. Here, $I_{\max}$ and $I_{\min}$ correspond to the maximum and adjacent minimum intensities in the pattern. The visibility is maximal (V=1) for two coherent, parallel polarized light waves of equal intensity and minimal (V=0) for incoherent light or orthogonally polarized light.

2.2 Speckles

If a randomly scattering object like a diffusely reflecting surface is illuminated with coherent light, all illuminated object points emit waves that can interfere. The resulting field in space is

called a speckle field or speckle pattern. In Fig. 2.1, a speckle field recorded by a CCD camera is depicted.

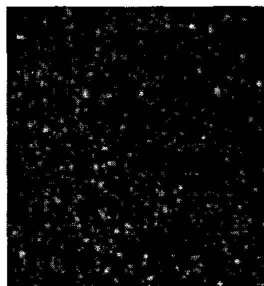


Fig. 2.1 *A typical speckle pattern.*

The speckles form a random pattern in space which is stationary in time but fluctuating from point to point. Their appearance depends strongly on the optical properties of the viewing system. In many applications, speckles are a troublesome effect since they influence the achievable resolution and accuracy of the measurement. On the other hand, a number of methods, such as speckle photography and speckle interferometry[7,8], employ the speckle effect for solving measurement problems. Speckles are the fundamental carriers of wavefront phase information and thus can be used in specific measurement techniques.

2.2.1 Speckle statistics

If a rough surface is illuminated with coherent light, a speckle pattern occurs in the observation plane (see Fig 2.2).

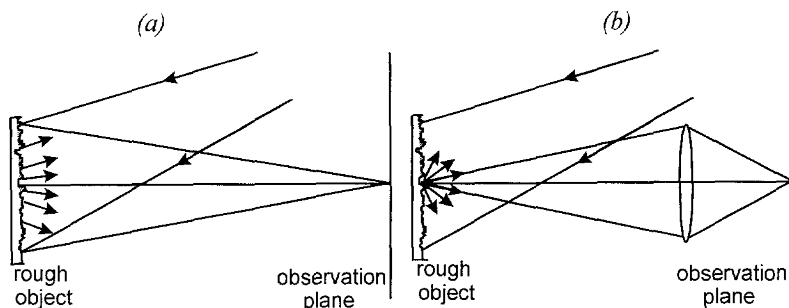


Fig. 2.2 *Formation of speckles: (a) free space propagation and (b) imaging geometry.*

According to the Huygens principle, each point of the coherently illuminated rough surface, if surrounded by an isotropic medium, emits a spherical secondary wavelet, described by the electric field

$$\boldsymbol{E}_n(r_n) = \frac{\boldsymbol{A}_n}{r_n} \mathrm{e}^{i(k \cdot r_n + \phi_n)} \qquad (n = 1,2..N), \tag{2.7}$$

in which the subscript n indicates the number of the wavelet and N is the total number of different scattering regions of the surface. The time dependence is omitted in this equation, since the mutual phases of the wavelets are temporally constant. The phases ϕ_n of the wavelets vary strongly with the emission point. Thus, at an observation point (x,y) in space all the individual field strengths sum up to

$$\boldsymbol{E}(x,y) = \sum_{n=1}^{N} \frac{\boldsymbol{A}_n}{r_n} \mathrm{e}^{i(k \cdot r_n + \phi_n)}, \tag{2.8}$$

where r_n is the distance from the pertinent surface point to (x,y). The summation in Eq. (2.8) is analogous to the random walk problem for two dimensions in probability theory. Each spherical wave can be represented as a vector in the complex plane[9], where the resulting field strength is given by the complex vector sum. Since the elementary scattering areas of the surface are unrelated and the strength of a given scattered contribution bears no relation to its phase we can write for the magnitude[9]

$$E_n(r_n) = \frac{1}{\sqrt{N}} |A_n| e^{i\phi_n}. \tag{2.9}$$

Moreover, for such rough objects, the amplitude $|A_n|/\sqrt{N}$ and the phase ϕ_n of each wavelet are statistically independent from each other and from the amplitudes and phases of all other wavelets and the phases ϕ_n are uniformly distributed[9] in the interval $\langle -\pi, \pi]$. This uniform probability density distribution of the phase p_ϕ can be described as

$$p_\phi(\phi) = \begin{cases} \dfrac{1}{2\pi} & (-\pi < \phi \le \pi), \\ 0 & (otherwise). \end{cases} \tag{2.10}$$

as is depicted in Fig. 2.3a). The intensity of a speckle pattern obeys a negative exponential probability distribution (see Fig. 2.3b), described by

$$p_I(I) = \begin{cases} \dfrac{1}{2\sigma^2} \mathrm{e}^{-\frac{I}{2\sigma^2}} & (I \ge 0), \\ 0 & (otherwise), \end{cases} \tag{2.11}$$

with σ^2 being the variance of the intensity.

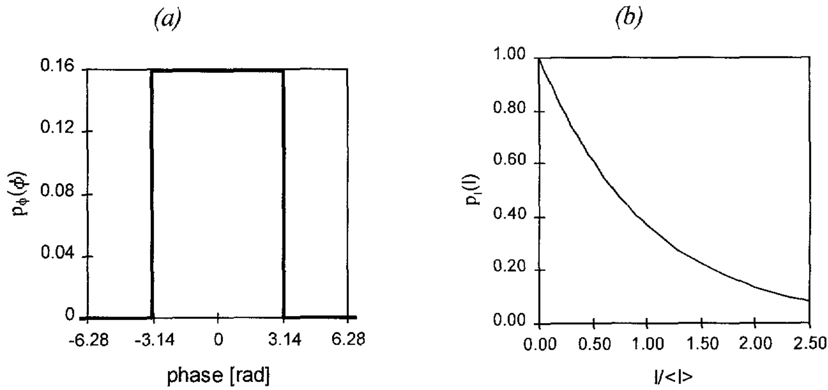


***Fig. 2.3** Probability density function of (a) the phase of a speckle pattern and (b) the intensity of a speckle pattern.*

The average value of the intensity $\langle I \rangle$ of a speckle pattern is $2\sigma^2$. Calculating the variance of the intensity

$$\sigma^2 = \langle I^2 \rangle - \langle I \rangle^2 = \langle I \rangle^2 \ , \tag{2.12}$$

we arrive at the result that the standard deviation σ of a speckle pattern equals the average intensity. Thus, the probability density function for speckle intensity, given by (2.11), can be rewritten as

$$p_I(I) = \frac{1}{\langle I \rangle} \exp\left(-\frac{I}{\langle I \rangle} \right) . \tag{2.13}$$

Some applications of speckle interferometry are concerned with rough metallic surfaces. When illuminated by linearly polarized light such surfaces diffract a strongly linear polarized wave and the intensity probability function satisfies Eq. (2.13). However, many surfaces show a depolarizing effect. Moreover, to obtain a more luminous image it may sometimes be necessary to deposit a strongly diffusing paint on the surface, which in general works depolarizing. The speckle field so obtained is the sum of two orthogonally polarized and uncorrelated speckle fields. The corresponding intensity distribution has the same statistical properties as the sum in intensity of the two incoherent speckle patterns and, assuming equal intensities for the two speckle fields, is given by[9]

$$p_I(I) = \frac{4I}{\langle I \rangle^2} \exp\left(-\frac{2I}{\langle I \rangle} \right) . \tag{2.14}$$

in which I represents the intensity sum of the two speckle fields. In Fig. 2.4, the probability density function of a single speckle pattern and that of two incoherent speckle patterns are displayed.

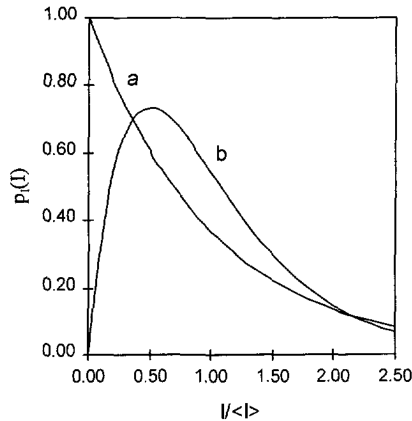


Fig. 2.4 *Probability density function of (a) a single speckle pattern and (b) incoherent combination of two speckle fields.*

This figure shows that the most probable brightness for a speckle in a speckle pattern is zero, i.e. there are more dark speckles in the field than speckles of any other brightness. If two independent speckle fields are superimposed, there is a vanishing probability for a dark speckle.

2.2.2 Speckle size

As was indicated in Section 2.2.1, two types of speckle patterns can be distinguished: free space speckle patterns and imaged speckle patterns. Depending on the geometry of the system, the speckles show different characteristics. Since imaging lenses are commonly used in speckle interferometry, only imaging geometry speckle patterns will be considered in this section. In the case of imaging geometry speckle formation, the spatial distribution of the speckles is determined by the scattering surface and the plane of observation, as well as by the diffraction limit of the imaging system. The average speckle size can be derived from the autocorrelation $R_a(x_1,y_1;x_2,y_2)$ of the intensity in the observation plane. The autocorrelation is defined as

$$R_a(x_1,y_1;x_2,y_2) = \langle I(x_1,y_1)I(x_2,y_2)\rangle , \tag{2.15}$$

in which $I(x_2,y_2)$ is a shifted copy of $I(x_1,y_1)$, the shift being $(\Delta x,\Delta y) = (x_2 - x_1, y_2 - y_1)$. For an imaging geometry, assuming a uniformly illuminated and uniformly scattering object area that is much larger than the spatial resolution of the used lens, the autocorrelation function of the intensity in the image plane is found to be[9]

$$R_a(r) = \langle I \rangle^2 \left(1 + \left| \frac{2J_1\left(\frac{\pi D_l r}{\lambda z}\right)}{\frac{\pi D_l r}{\lambda z}} \right|^2 \right), \tag{2.16}$$

where D_l is the diameter of the circular lens pupil, J_1 is the first order Bessel function, λ is the wavelength of the used light, z is the distance of the image plane from the lens pupil plane and $r = (\Delta x^2 + \Delta y^2)^{1/2}$. The average speckle size d_s is taken as the first minimum of J_1. Since $J_1(x) = 0$ at $x = 1.22\pi$, we obtain

$$d_s = 1.22 \frac{\lambda z}{D_l}. \tag{2.17}$$

The speckle size in the image plane can be related to the numerical aperture *NA*, which is given by

$$NA = \sin(\theta_i) \approx \tan(\theta_i) = \frac{D_l}{2f}, \tag{2.18}$$

where θ_i indicates the angle of convergence in the image space and f is the focal length of the lens. Since $z = f$ in the image plane, this results in

$$d_s = 0.61 \frac{\lambda}{NA}. \tag{2.19}$$

The size of the image plane speckles thus depends on the numerical aperture (*NA*) of the imaging system. If the aperture of the viewing lens is decreased, the speckle size will increase and vice versa. The maximum spatial frequency $f_{\max}$ in the speckle pattern is given by the size of the lens aperture and the distance of the lens from the observation plane by[1]

$$\frac{1}{f_{\max}} = \frac{\lambda z}{D}. \tag{2.20}$$

Similar to the average speckle diameter, the average speckle length l_s can be found[10] from

$$l_s = 2 \frac{\lambda}{NA^2}. \tag{2.21}$$

This shows that for a usual speckle interferometer, which has an *NA* smaller than 0.1, the speckle length is much larger than the speckle diameter.

2.2.3 Optimal speckle size for speckle interferometry

In the framework of speckle interferometry, the average speckle size is very important. The restriction on the speckle size originates from decorrelation of speckle patterns. The correlation coefficient $\mu(I_1, I_2)$ between two interferograms I_1 and I_2 can be determined using[9]

$$|\mu(I_1, I_2)|^2 = \frac{\langle I_1 I_2 \rangle - \langle I_1 \rangle \langle I_2 \rangle}{\sqrt{\langle I_1^2 \rangle - \langle I_1 \rangle^2} \sqrt{\langle I_2^2 \rangle - \langle I_2 \rangle^2}}. \tag{2.22}$$

Decorrelation is one of the fundamental limits of electronic speckle pattern interferometry[11,12]. Speckle pattern decorrelation will cause poor visibility of the fringes. Therefore, decorrelation between the different interferograms is a major error contribution. Decorrelation can be caused by the load applied on the object under inspection between two recordings (the so-called object-induced speckle decorrelation). A second cause of speckle decorrelation effect is only present in multichannel interferometers (see Section 2.6), where the detection planes in the separate channels of the interferometer can be decorrelated due to misalignment.

Many workers[13-17] in the field of speckle interferometry have emphasized the need to link the average speckle size to the spatial resolution of the device recording the speckle pattern. The basic idea behind this was that if speckles are too small to be resolved, the resulting pixel signal will be the average of many independent interferograms. Thus, the total pixel intensity modulation would be lower than the average speckle modulation. Therefore, some propose to fulfill the Nyquist sampling criterion, which states that the average speckle size should be twice the size of the picture elements in order to fully resolve the speckle pattern, whereas others suggest the use of smaller speckles since smaller speckles reduce the speckle noise and improve the fringe contrast. This is especially important in low light level applications, where larger apertures generally produce higher modulation intensities, such that the modulation loss associated with averaging speckles is more than compensated by several effects, like higher average intensities. Moreover, undersampling of the speckle pattern is allowed since, in principle, an accurate reproduction of the speckle pattern is not of interest; resolving the speckle pattern completely is in general not necessary. The most important factors determining the optimal speckle size in a speckle interferometer are object induced decorrelation and, in the case of a multichannel interferometer, also by camera decorrelation. The influence of the speckle size on these two decorrelation types are described in the remainder of this subsection.

Object decorrelation

The principle behind decorrelation of the speckle patterns between the exposures before and after deformation is illustrated in Fig. 2.5, with object displacement and object tilt as examples of deformation.

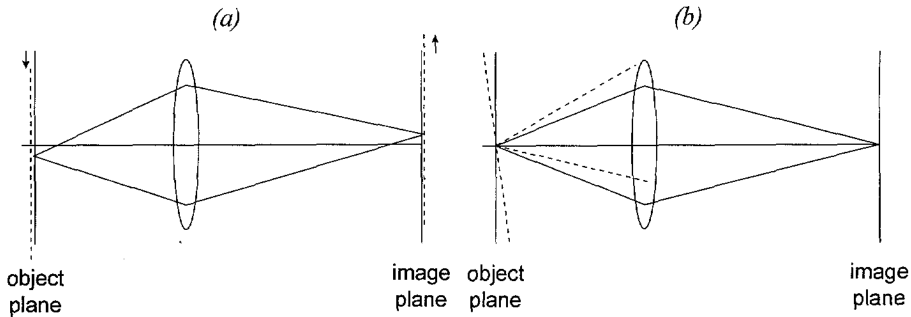


Fig. 2.5 *Speckle decorrelation (a) due to in-plane translation and (b) due to out-of-plane rotation of the object.*

In Fig. 2.5a, image plane decorrelation is depicted. Due to the image (or equivalently object) plane translation, an object point is imaged onto another picture element. Thus, speckle decorrelation deteriorates a recorded interferogram in two ways: firstly, by a variation of the phase in the interference pattern across an individual speckle and secondly, by the displacement of the speckles. Due to a small relative displacement between the object and image planes, the picture elements record partly the same field as before decorrelation and partly a new contribution. The fraction of the picture element area which is replaced gives the amount of decorrelation for unresolved speckle patterns, which can be defined by

$$1-\mu=\frac{\Delta p_a}{p_a}, \tag{2.23}$$

where μ represents the amplitude correlation between the speckle fields recorded before and after loading, Δp_a is the area of the picture element that records a different part of the speckle pattern and p_a the total area of the picture element. Thus, in-plane decorrelation is reduced by using larger picture elements, except for well-resolved speckles. The number of speckles per picture element does not affect the decorrelation induced phase error distribution over the whole image[17]. The only important parameters determining the phase error induced by the in-plane decorrelation are the amount of decorrelation and the picture element modulation.
Although for well-resolved speckle patterns the amount of decorrelation is related to the displacement relative to the speckle size, even for this category it can be advantageous to use large apertures. As Eq. (2.19) shows, increasing the aperture decreases the average speckle size. At the same time, a large value for the numerical aperture gives a high intensity in the detection plane. Therefore, using larger apertures can increase the modulation intensity of interferograms, especially for interferometers using two speckle fields[18], since they usually suffer from low intensities.

Due to the out-of-plane tilt of the object, as depicted in Fig. 2.5b, an object point is imaged onto the same picture element as before the rotation, but a different part of the field is collected

by the aperture. The amount of decorrelation is given by the fraction of the aperture surface that changes its contribution as given by

$$1-\mu=\frac{\Delta A}{A}, \tag{2.24}$$

in which ΔA is the area of the aperture surface which receives a different part of the speckle pattern, and A is the total area of the aperture. Because of the decorrelation, the picture elements no longer collect a slightly shifted region of the same speckle field but the speckle field amplitude itself slightly changes in each point. Both intensity and phase of the imaged speckle will be different. This type of decorrelation is also called pupil plane decorrelation, since the amount of decorrelation depends upon the fraction of the pupil collecting different contributions. Increasing the aperture of the imaging system reduces pupil plane decorrelation.

Concludingly, since the decorrelation induced phase errors are minimized by maximizing the modulation intensity of the interferograms and minimizing the amount of decorrelation by choosing a large aperture (and large picture elements if possible), small speckles are generally beneficial to use in electronic speckle pattern interferometry.

Camera decorrelation

If the cameras (or other recording devices) of a multichannel speckle interferometer are not recording the same parts of the speckle pattern, decorrelation of the speckle patterns occurs. The decorrelation in multichannel interferometers is predominantly caused by relative in-plane translations or rotations between the cameras[10]. The correlation between two corresponding detection points on the cameras recording the interferograms which have a relative shift r can be deduced from the crosscorrelation function R_c of two speckle patterns I_1 and I_2, defined as

$$R_c(x_1,y_1;x_2,y_2)=\langle I_1(x_1,y_1)I_2(x_2,y_2)\rangle, \tag{2.25}$$

and using Eq. (2.19). This result in an intensity correlation factor $|\mu(I_1,I_2)|^2$, described by

$$|\mu(I_1,I_2)|^2=2\frac{J_1\left(1.22\pi\left(\frac{r}{d_s}\right)\right)}{1.22\pi\left(\frac{r}{d_s}\right)}. \tag{2.26}$$

A displacement that is small as compared to the speckle size, changes the speckle pattern at a given point in the image plane only by a small amount. Equation (2.26) also shows that the speckle pattern at a point is totally decorrelated if the object is translated by an amount Δx or Δy equal to the speckle size[9,19]. Thus, large speckles are advantageous in a multicamera speckle interferometer. Also, using large picture elements reduces camera decorrelation for the same reason as for -plane object translation.

Combining the considerations from object induced decorrelation and camera decorrelation, it can be concluded that there is an optimum speckle size, but its size depends on which type of decorrelation is dominant, and this in turn is determined by the type of the speckle interferometer and the measurement geometry. Large speckles reduce camera decorrelation. However, decreasing the aperture also increases the object induced decorrelation. In general, for multichannel interferometers it is advantageous to use large speckles while for other systems small speckles are preferred.

2.3 Holographic interferometry

Holography is a method for recording and reconstructing optical waves. Holography is essentially a clever combination of interference and diffraction, two phenomena based on the wave nature of light. A pattern recorded in a holographic set-up is called a 'hologram'[20]. As the word hologram (the word originates from the Greek: "holos" meaning "whole" or "entire" and "graphein" meaning "to write") indicates, it deals with the amplitude as well as with the phase of the light. All detectors or recording materials like photographic film only register intensities: the phase is lost in the registration process. In holography the phase information is usually coded by interference into a recordable intensity. Interferometry and holography can be combined to obtain the method of holographic interferometry[1,21]. Holographically reconstructed waves can be compared interferometrically, either with a wave scattered directly by the object or with another holographically reconstructed wave.

Holographic interferometry is a useful tool in metrology which is capable of precise measurement of surface shape or deformation. It can yield a high spatial resolution. Unfortunately, due to some inherent application difficulties like photographic plate development, kinematic relocation, complex analysis and sensitivity to external vibrations, this technique's full potential cannot be realized in industrial applications.

2.4 Electronic speckle interferometry

Electronic speckle pattern interferometry[22-25], also called ESPI, was introduced to overcome the time consuming and cumbersome recording and wet-chemical processing of holograms. ESPI is characterized by electronic detection and processing of the speckle patterns. Integration of computers into the measuring system facilitated computer aided evaluation. The disadvantages associated with ESPI are the highly grainy fringes and strong decorrelation effects which limit the range of applications. This results in noisy interference patterns, making a (quantitative) evaluation of the interferograms more difficult.

We consider a speckle interferometer with the image plane (x, y) coinciding with the detection plane. The speckled object wave $\boldsymbol{E}_o(x, y)$ is described by

$$\boldsymbol{E}_o(x,y) = \begin{pmatrix} A_x(x,y) \\ 0 \end{pmatrix} \mathrm{e}^{i\varphi_o(x,y)}, \tag{2.27}$$

where $A_x(x,y)$ is the random amplitude distribution and $\varphi_o(x,y) = \boldsymbol{k}_o \cdot \boldsymbol{r}_o + \phi_o$ is the random phase due to the surface roughness. Since we are only interested in the spatial distribution of the field, the time dependent part, i.e. the factor $e^{-i\omega t}$ (see Section 2.1) which is known for each frequency, is omitted from our expressions. This notation will be used throughout this thesis.

The reference wave may have an arbitrary form and is described by

$$\boldsymbol{E}_r(x,y) = \begin{pmatrix} A_x(x,y) \\ 0 \end{pmatrix} \mathrm{e}^{i\varphi_r(x,y)}. \tag{2.28}$$

Only intensities are recorded by the camera, resulting in the interference pattern that is described by the intensity distribution

$$\begin{aligned} I_1(x,y) &= \left|\boldsymbol{E}_o(x,y) + \boldsymbol{E}_r(x,y)\right|^2 \\ &= I_o(x,y) + I_r(x,y) + 2\sqrt{I_o(x,y)I_r(x,y)}\cos(\varphi(x,y)), \end{aligned} \tag{2.29}$$

in which $\varphi(x,y)$ equals $\varphi_o(x,y) - \varphi_r(x,y)$. The sum of the first and second terms in Eq. (2.29) describes the background intensity $I_b(x,y)$ and the factor before the cosine function is the modulation intensity $I_m(x,y)$. This interference pattern is recorded, digitized and stored. A deformation changes the phase of $\varphi_o(x,y)$ at each point by $\Delta\varphi(x,y)$. The wave after deformation is

$$\boldsymbol{E}_o'(x,y) = \begin{pmatrix} A_x'(x,y) \\ 0 \end{pmatrix} \mathrm{e}^{i(\varphi_o(x,y)+\Delta\varphi(x,y))}. \tag{2.30}$$

Superposition of this new object wave with the reference wave leads to

$$I_2(x,y) = I_o'(x,y) + I_r(x,y) + 2\sqrt{I_o'(x,y)I_r(x,y)}\cos(\varphi(x,y) + \Delta\varphi(x,y)). \tag{2.31}$$

In a digital image processing system, this second pattern is subtracted pixel by pixel from the stored $I_1(x,y)$ and, assuming that the deformation changes the phase but not the amplitude (i.e. $A_x'(x,y) = A_x(x,y)$), it results in[1]

$$\begin{aligned} I_1(x,y) - I_2(x,y) &= 2\sqrt{I_o(x,y)I_r(x,y)} \times \\ &\quad \left\{\cos(\varphi(x,y)) - \cos(\varphi(x,y))\cos(\Delta\varphi(x,y)) + \sin(\varphi(x,y))\sin(\Delta\varphi(x,y))\right\} \\ &= 4\sqrt{I_o(x,y)I_r(x,y)}\sin\left(\varphi(x,y) + \frac{\Delta\varphi(x,y)}{2}\right)\sin\left(\frac{\Delta\varphi(x,y)}{2}\right). \end{aligned} \tag{2.32}$$

The square root factor in Eq. (2.32) describes the modulation in the intensity pattern. The first sine factor gives the stochastic speckle noise, which varies randomly from pixel to pixel. This noise is modulated by the sine of half the phase difference induced by the deformation. This low-frequency modulation of the high-frequency speckle noise is recognizable as the interference pattern. The relation between the displacement vector $\boldsymbol{d}(x,y)$ of the object and the phase difference $\Delta\varphi(x,y)$ is[1]

$$\Delta\varphi(x,y) = \frac{2\pi}{\lambda}\boldsymbol{d}(x,y)\cdot\left[\hat{\boldsymbol{b}}(x,y)-\hat{\boldsymbol{s}}(x,y)\right] = \boldsymbol{d}(x,y)\cdot\boldsymbol{e}(x,y), \qquad (2.33)$$

in which $\hat{\boldsymbol{s}}$ is the unit vector in the illumination direction, $\hat{\boldsymbol{b}}$ is the unit vector in the observation direction and $\boldsymbol{e}$ is the so-called sensitivity vector.

With this technique, as is also the case in holographic interferometry, the reference beam is generated somewhere within the interferometer using a reference surface. Therefore, this method is also extremely sensitive to external disturbances like air flow or vibrations. Consequently, the generation of a fringe pattern requires many precautions concerning vibration isolation and stability during the recording process.

2.5 Shearing interferometry

The requirement of vibration isolation can be dropped to a good extent when using shearing interferometry. The reference wave used in shearography takes nearly the same optical path as the object beam. In fact, it cannot be said which of both is the reference and which is the object beam.

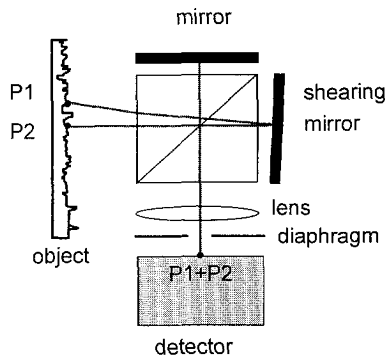


Fig. 2.6 *Schematic representation of a lateral shearing interferometer.*

In Fig. 2.6 the basic principle of shearing interferometry, alternatively called shearography, is depicted. Two laterally shifted (sheared) images of the object are generated in the detector plane by means of an optical device (e.g. as depicted in Fig. 2.6). In the overlapping regions of the two images interference will occur. This way, the need for a separate reference beam is

avoided. This technique has the properties of a self-reference technique and Since in this set-up the object and reference beams travel along neighboring paths from the object to the interferometer, object motions affect both beams in similar ways. Therefore, this method is rather insensitive to rigid body motions of the object. Also the disturbing influence of air turbulence, which changes the refractive index distribution of air and thus the optical path lengths randomly, is reduced strongly.

In a lateral shearography setup, as is depicted in Fig. 2.6, the interfering waves are two laterally, slightly shifted fields resulting from reflection at the rough object surface. In this case, the lateral shift is performed by a Michelson interferometer-like arrangement with one mirror slightly tilted. A shearing element can also be implemented in other ways, e.g. a wedge in front of one half of the imaging lens, a biprism or two tilted glass plates. Besides this lateral (or translational) shearing, also radial shearing, rotational shearing, inversion shearing or reversal shearing can be performed[26].

Let the two interfering waves be separated in the x-direction by the mutual shearing distance Δx, then the two nearly collinear waves are described by

$$\boldsymbol{E}_r(x,y) = \begin{pmatrix} A_x(x,y) \\ 0 \end{pmatrix} \mathrm{e}^{i\varphi_r(x,y)}, \tag{2.34}$$

$$\boldsymbol{E}_o(x,y) = \begin{pmatrix} A_x(x,y) \\ 0 \end{pmatrix} \mathrm{e}^{i\varphi_o(x,y)} = \begin{pmatrix} A_x(x+\Delta x,y) \\ 0 \end{pmatrix} \mathrm{e}^{i\varphi_r(x+\Delta x,y)}, \tag{2.35}$$

in which the object beam is (arbitrarily) chosen as the shifted one. Their interference produces a pattern described by

$$\begin{aligned} I_1(x,y) &= \left|\boldsymbol{E}_o(x,y) + \boldsymbol{E}_r(x,y)\right|^2 \\ &= I_o(x,y) + I_r(x,y) + 2\sqrt{I_1(x,y) I_2(x,y)} \cos\left(\varphi(x,y)\right), \end{aligned} \tag{2.36}$$

where the randomly distributed phase difference $\varphi(x,y)$ satisfies

$$\varphi(x,y) = \varphi_r(x,y) - \varphi_r(x+\Delta x,y). \tag{2.37}$$

Deformation of the object leads to a change in the waves, now indicated with primes, described by

$$\boldsymbol{E}_r'(x,y) = \begin{pmatrix} A_x(x,y) \\ 0 \end{pmatrix} \mathrm{e}^{i(\varphi_r(x,y)+\Delta\varphi(x,y))}, \tag{2.38}$$

$$\boldsymbol{E}_o'(x,y) = \begin{pmatrix} A_x(x,y) \\ 0 \end{pmatrix} \mathrm{e}^{i(\varphi_r(x+\Delta x,y)+\Delta\varphi(x+\Delta x,y))}, \tag{2.39}$$

their superposition yields the intensity distribution

$$I_2(x,y) = I_o(x,y) + I_r(x,y) + 2\sqrt{I_o(x,y)I_r(x,y)}\cos(\varphi(x,y) + \Delta\varphi(x,y) - \Delta\varphi(x+\Delta x,y)). \tag{2.40}$$

Both interference patterns $I_1(x,y)$ and $I_2(x,y)$ are digitally recorded and stored. Their pixelwise subtraction gives

$$\begin{aligned} &I_1(x,y) - I_2(x,y) \\ &= 2\sqrt{I_o(x,y)I_r(x,y)}\left\{\cos(\varphi(x,y)) - \cos(\varphi(x,y) + \Delta\varphi(x,y) - \Delta\varphi(x+\Delta x,y))\right\} \\ &= 4\sqrt{I_o(x,y)I_r(x,y)}\sin\left(\varphi(x,y) + \frac{\Delta\varphi(x,y) - \Delta\varphi(x+\Delta x,y)}{2}\right) \\ &\quad \times \sin\left(\frac{\Delta\varphi(x,y) - \Delta\varphi(x+\Delta x,y)}{2}\right). \end{aligned} \tag{2.41}$$

The square root factor describes the modulation of the intensity pattern, the first sine factor is the stochastic speckle noise, modulated by the second sine factor which stems entirely from the deformation. The displacement vector field $\boldsymbol{d}(x,y)$ occurs in the argument of the second sine via

$$\begin{aligned} &\frac{\Delta\varphi(x,y) - \Delta\varphi(x+\Delta x,y)}{2} \\ &\quad = \frac{\pi}{\lambda}\left\{\boldsymbol{d}(x,y)\left[\hat{\boldsymbol{b}}(x,y) - \hat{\boldsymbol{s}}(x,y)\right] - \boldsymbol{d}(x+\Delta x,y)\cdot\left[\hat{\boldsymbol{b}}(x+\Delta x,y) - \hat{\boldsymbol{s}}(x+\Delta x,y)\right]\right\} \\ &\quad \approx \frac{\pi}{\lambda}\left[\boldsymbol{d}(x,y) - \boldsymbol{d}(x+\Delta x,y)\right]\cdot\left[\hat{\boldsymbol{b}}(x,y) - \hat{\boldsymbol{s}}(x,y)\right] \\ &\quad \approx \frac{\partial \boldsymbol{d}(x,y)}{\partial x}\cdot\frac{\pi\Delta x}{\lambda}\left[\boldsymbol{b}(x,y) - \boldsymbol{s}(x,y)\right]. \end{aligned} \tag{2.42}$$

Here, the sensitivity and the illumination directions are assumed to be practically equal for two sheared object points. From this expression we can see that the interference pattern is in a good approximation directly related to the derivative of the displacement field in the shearing direction, here the *x*-direction. For rigid body motions, $\boldsymbol{d}(x,y)$=constant, implying $\partial \boldsymbol{d}(x,y)/\partial x = 0$. The sensitivity of the method can be adjusted by controlling the shearing distance.

2.6 Quantitative analysis of the interference phase distribution

Quantitative evaluation of ESPI interference patterns includes the pixelwise determination of the numerical value of the physical quantity which caused the optical path length change at each point (x,y) of the interferogram. This physical quantity is related to the phase distribution $\varphi(x,y)$, whereas the intensity distribution is recorded in interferometry.

The basic information in an interferogram is its intensity distribution, described by

$$I(x,y) = I_b(x,y) + I_m(x,y)\cos(\varphi(x,y)), \tag{2.43}$$

from which $\varphi(x,y)$ is to be computed. When one tries to extract the interference phase $\varphi(x,y)$ from the intensity distribution $I(x,y)$, the problem arises that the cosine is an even and periodic function. Moreover, interferograms suffer from noise (such as intensity noise, electronic noise, changing surroundings, etc.), degrading the pattern. Especially since ESPI uses speckle fields as information carriers, a high noise level is present, and fringe analysis is quite a difficult problem. Therefore, the intensity in a real interferogram is much more complicated than given in Eq. (2.43). Taking the noise contributions into account, an interferogram can be described as

$$I(x,y) = a(x,y) + b(x,y)\cos(\varphi(x,y)), \tag{2.44}$$

in which the additive and multiplicative noise contributions are contained in the term $a(x,y)$ and $b(x,y)$ respectively. In this section, several phase calculation methods will be presented. In Fig. 2.7 an overview of these methods is shown. Some techniques can reconstruct the phase using a single interference pattern, whereas others require several interference patterns. The basic properties, advantages and limitations of each technique will be discussed.

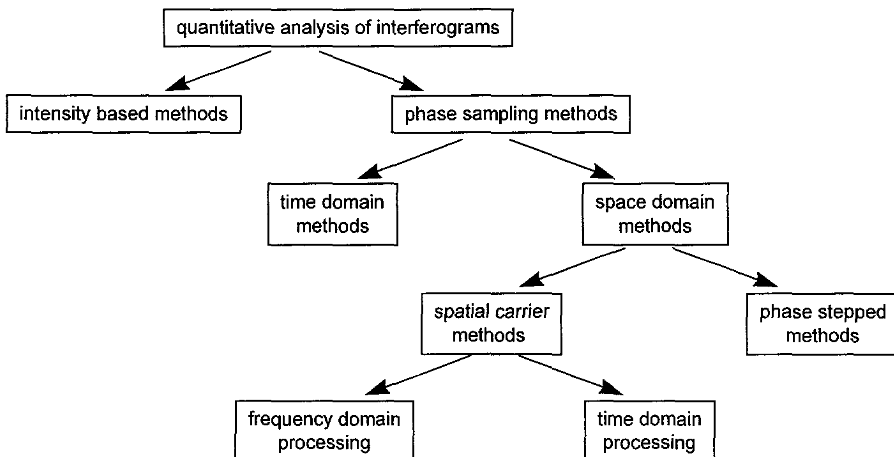


Fig. 2.7 *Categories of quantitative interferogram analysis methods.*

2.6.1 Intensity based analysis methods

Single interferograms can be analyzed using digital image processing techniques. The objective of intensity based image processing methods is to estimate the interference phase distribution $\varphi(x,y)$ from the more or less disturbed intensity distribution $I(x,y)$. The methods are based on the fact that constant values of $\varphi(x,y)$ define fringe positions on the object's surface.

Fringe extrema or other loci of equal phase values are determined, using e.g. spatial filtering, fringe tracking, skeletonizing or Fourier techniques. Then the local phase distribution $\varphi(x,y)$ between fringe extrema is approximated using interpolation techniques, or by fitting a polynomial in x and y to the data.

Experimentally, intensity based fringe analysis methods are the most simple, not requiring a complicated optical set-up. However, quantitative image analysis of a single interferogram often produces a number of problems and ambiguities whose solutions require time-consuming computations.[27] Due to the encryption of the phase distribution by a cosine function (see Eq. (2.43)), the interference phase distribution determined from a single intensity distribution remains undefined to an additive integer multiple of 2π and to the sign. Fringe localization techniques are rather cumbersome and lack sufficient accuracy, due to the high noise levels present in the interferograms. The phase is determined indirectly at points between fringe extrema by interpolation techniques, reducing the spatial resolution. It also yields a low resolution in phase because of the noise contributions. It needs expertise to process the images and automated calculation of $\varphi(x,y)$ is difficult.

2.6.2 Phase sampling analysis methods

Introducing phase sampling techniques for the evaluation of speckle interference patterns is an essential improvement. With these techniques, the phase distribution is calculated within a high accuracy and without sign ambiguity. Under equal conditions, phase sampling interferometers have a higher sensitivity than systems without phase sampling.

Phase sampling methods consist of recording multiple phase-stepped or phase shifted interferograms. The technique is based upon the modulation of the relative phase shift between the object and reference beams in the interferometer between the recordings. Two types of phase sampling methods can be recognized: phase shifting and phase stepping techniques. In phase shifting interferometry, the interference patterns are recorded while the phase is shifting (see Fig. 2.8a), whereas phase stepping interferometry records the interferogram after the phase has been shifted (see Fig. 2.8b).

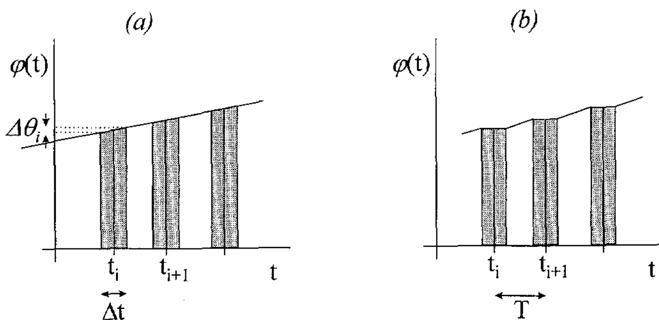


Fig. 2.8 *Diagrams describing (a) phase shifting and (b) phase stepping.*

For phase stepping interferometry, the intensity distributions $I_i(x,y)$ are expressed by[1]

$$I_i(x,y) = a(x,y) + b(x,y)\cos\left(\varphi(x,y) + \theta_i\right) \quad (i=1,...N;\ N\geq 3), \tag{2.45}$$

where θ_i is the stepped reference phase belonging to the i$^{\text{th}}$ intensity distribution $I_i(x,y)$. The integer N indicates the number of interferogram recordings (the so-called 'buckets').

For phase shifting interferometry, assuming a linear phase shift $\theta_i(t)$ with time, the interferograms can be described as

$$\begin{aligned} I_i(x,y) &= \frac{1}{\Delta t}\int_{t_i-\Delta t/2}^{t_i+\Delta t/2}\left[a(x,y) + b(x,y)\cos\left(\varphi(x,y) + \theta_i(t)\right)\right]dt \\ &= a(x,y) + \operatorname{sinc}\left(\frac{\Delta\theta_i}{2}\right)b(x,y)\cos\left(\varphi(x,y) + \theta_i\right). \end{aligned} \tag{2.46}$$

The difference with phase stepping is that the contrast term $b(x,y)$ is modified by the constant factor $\operatorname{sinc}(\Delta\theta_i)/2$, $\Delta\theta_i$ denotes the phase shift during the time interval Δt. For phase stepping, $\Delta\theta_i$ vanishes and the sinc function equals unity. Phase stepping is superior to phase shifting with respect to the achievable accuracy because of the reduction of the modulation intensity by the sinc function in phase shifting interferometry.

In order to determine the phase $\varphi(x,y)$ from Eq. (2.45) for each point (x,y), a nonlinear system of equations has to be solved. If the phase shift θ_i is known, three unknowns remain in Eq. (2.45), namely $a(x,y)$, $b(x,y)$ and the desired interference phase $\varphi(x,y)$. Hence, at least three equations are necessary, meaning that at least three phase-stepped or shifted interferograms have to be recorded. The interference phase is determined modulo 2π from[28]

$$\varphi(x,y) = \arctan\left(\frac{-\sum_{i=1}^{N} I_i \sin(\theta_i)}{\sum_{i=1}^{N} I_i \cos(\theta_i)}\right), \tag{2.47}$$

if $\theta_i = i2\pi/N$ and $N\geq 3$.

Simple solutions of Eq. (2.47) have been published[28-30] for phase steps which are constant and have special values like π/6, π/4, π/3, π/2 or 2π/3 rad. If the phase step is π/2 rad, N=3 and the start value θ_1=0, the phase can be calculated using

$$\varphi(x,y) = \arctan\left(\frac{I_1(x,y) - 2I_2(x,y) + I_3(x,y)}{I_1(x,y) - I_3(x,y)}\right). \tag{2.48}$$

In the case of N=4 and using the same phase step, the phase is calculated using

$$\varphi(x,y) = \arctan\left(\frac{I_4(x,y) - I_2(x,y)}{I_1(x,y) - I_3(x,y)}\right). \tag{2.49}$$

Using a 3-bucket method with a $2\pi/3$ phase step and a starting value of $\theta_1=0$, we obtain

$$\varphi(x,y) = \arctan\left(\sqrt{3}\,\frac{I_3(x,y) - I_2(x,y)}{2I_1(x,y) - I_3(x,y) - I_2(x,y)}\right). \tag{2.50}$$

If the phase shift is unknown but constant, an additional unknown is introduced, then at least four phase shifted or stepped interferograms are required.

The main advantage of phase sampling methods is that the phase is measured directly, so that these methods have a high accuracy. The additive and multiplicative noise components are automatically compensated for in the phase calculation process. The resolution and accuracy of the determined interference phase using a phase sampling method[30] can reach $2\pi/100$. Also, the spatial resolution is higher than with an intensity based analysis since no interpolation between skeleton lines is necessary. Furthermore, the sign ambiguity is resolved automatically. Finally, the calculation of $\varphi(x,y)$ can be fully automated and masks can be used to exclude bad pixels and pixels with low modulation from the evaluation process[27]. Some disadvantages associated with phase sampling involve the additional technical effort required to perform the phase shifts accurately and the need for additional computer storage capacity.

An alternative way of categorizing phase sampling methods is to divide the methods based upon the way the phase steps are implemented. Two categories can be recognized: time domain methods, taking the phase data sequentially and space domain methods, recording the phase data simultaneously.

2.6.2.1 Time domain methods

Time domain methods are also known as temporal phase sampling methods. The optical path length difference between the object and reference beams is made time-dependent. There are many variations of the temporal phase measurement technique. A frequently used option is a mirror, which reflects the reference wave, mounted on a piezo crystal that shifts the mirror over fractions of the used wavelength[28,31]. Other methods[23] are based on moving a grating, tilting a glass plate, rotating a half wave plate or analyzer, using an acousto-optic or electro-optic modulator or using object-induced phase shifts[32].

The major disadvantage of time domain methods is the lack of real-time performance. The phase-stepped or shifted interferograms are generated and recorded sequentially. Therefore, it requires absolute stability of the object during the generation and storage of the (at least) three

images. This strongly reduces the applicability of these methods in hostile environments. Vibrations, acoustic noise or air turbulence modify the interference phase distribution to be measured. Moreover, investigations of dynamic processes, such as deformations caused by creep, shrinkage, thermal loading, propagation of shock waves, or high-frequency vibrational states, are nearly impossible.

2.6.2.2 Space domain methods

Space domain methods are also called spatial phase stepping methods. In the spatial methods, the phase information is distributed in space (across the image plane) and is recorded in a single frame time. The spatial techniques can be further subdivided into spatial carrier methods and phase-stepped methods. The phase-stepped methods are based on the same principle as the phase sampling methods described above. However, as the word 'stepped' implies, the phase modulation is a single step process, as opposed to the multi-step process in phase sampling methods, yielding the required interference patterns simultaneously. Therefore, these methods are also often called real-time phase-stepped methods.

Spatial carrier phase measurement methods

Spatial carrier phase measurement methods are based upon a large tilt between the reference and object wavefronts. A fringe pattern with a spatial carrier frequency f_0 in the x-direction is given by

$$I(x,y) = a(x,y) + b(x,y)\cos\left(2\pi f_o x + \varphi(x,y)\right), \tag{2.51}$$

where $a(x,y)$ and $b(x,y)$ are again the background and the contrast functions.

In order to calculate the phase distribution without ambiguity, the spatial frequency has to obey several conditions[33]. Firstly, the largest gradient of the phase $\varphi(x,y)$ has to be less than the spatial carrier phase $2\pi f_0$. Secondly, the background $a(x,y)$ and the contrast function $b(x,y)$ are slowly varying compared with the carrier frequency. Thirdly, the frequency f_0 has to be less than half of the sampling frequency (Nyquist condition). There are two approaches to calculate the phase distribution from the spatial carrier interferogram. One is the Fourier transform method[33-36], in which the processing is performed in the frequency domain. The input fringe pattern is Fourier transformed with respect to x by an FFT algorithm. If the spatial variations of $a(x,y)$, $b(x,y)$ and $\varphi(x,y)$ are slow as compared to the spatial frequency f_0, the Fourier spectrum will exhibit three distinct parts; one DC term and two AC terms, located around the carrier frequency $\pm f_0$. One of these two spectra located around the carrier is separated from the rest and translated by $\pm f_0$ along the frequency axis towards the origin. This way, the unwanted background variation $a(x,y)$ has been filtered out. Next, using the inverse Fourier transform with respect to f_x, a complex function is obtained whose phase angle is the interference phase function $\varphi(x,y)$ of the original interferogram. This phase can be calculated

by taking the arctangent of the ratio of the imaginary and real parts of the complex function. The phase obtained in this way is inconclusive to a factor of 2π and the profile can be determined uniquely (without ambiguity about the sign of the phase).

Using Fourier methods, some knowledge of the experiment is required[37]. Therefore, the processing of arbitrary interference patterns can not be fully automated for arbitrary interference patterns. It requires operating experience and information of additional observations or à priori knowledge. Also, not all noise contributions can be filtered out, namely those with variations which lie in the same frequency bands as the desired interference patterns. Moreover, filtering is time-consuming. In a recent paper[36] it is expected that the phase could be displayed within seconds after the recording of the interferogram (using a resident FFT processor board). The phase distribution is not available in real time which classifies this technique as outside the real-time ones.

The second method processes the interferogram in space domain[38]. It involves introducing a spatial frequency carrier f_0 to the fringe pattern such that neighboring pixels have approximately $2\pi/3$ or $\pi/2$ rad phase difference. The interference pattern using a $\pi/2$ phase shift is described by

$$I_n(x,y) = a(x,y) + b(x,y)\cos\left[2\pi f_o x + \frac{n\pi}{2} + \varphi(x,y)\right] \quad (n=1..N). \tag{2.52}$$

The phase in a pixel (x,y) can be calculated by use of N subsequent pixels, in the same manner as a temporally phase shifted interferogram. After shifting one pixel to the right, the phase calculation is repeated for the next set of N pixels. The resulting phase map contains the desired phase plus a ramp with $\pi/2$ rad phase shift per pixel. This tilt is subtracted to produce the raw phase map. Using this method, the phase can be determined in the range $\langle -\pi, \pi]$ without sign ambiguity. The main disadvantage of space domain processing such spatial carrier frequency fringe patterns is associated with the phase variations that are sought. This implies frequency variations away from the spatial carrier frequency, and so the accuracy of the phase measurement will be perturbed. The resulting phase interpretation errors will be functions of both the spatial frequency and the phase of the actual fringes. Accurate alignment of the spatial carrier frequency is crucial for accurate results in this technique.

The main advantage of the spatial carrier techniques (using Fourier processing or space domain processing methods) is that phase distributions can be measured using a single image frame. Unfortunately, there are some disadvantages associated with these methods. They require a slowly varying background, modulation intensity and/or phase distribution, which is not generally true for diffusely reflecting objects. Moreover, spatial carrier techniques belong to the category of phase shifting techniques, where the phase measurement is less accurate because of the encryption of the phase with the sinc function (see Eq. (2.46)). Furthermore, spatial carrier techniques generally require much more sophisticated processing and more computing time to retrieve the phase from a fringe pattern than the phase-stepped methods[33].

Phase-stepped methods

Some of the methods used in temporal phase stepping can be used for phase-stepped interferometry. Optical methods to generate phase-stepped interference patterns simultaneously are based on polarization optics[10] or diffraction gratings[33]. In a polarization type of interferometer the interferometer branches are traversed by orthogonally polarized waves and interference is obtained with the aid of wave retarders and a polarizer[39]. In many cases this method is preferred since it offers the best signal to noise ratio and eliminates systematic errors.[40] In contrast to temporal phase stepping techniques, the three or more phase-stepped interferograms are generated and recorded simultaneously. Therefore, the interferometer has to be adapted by introducing the phase shifts in parallel channels. Such an interferometer is called a multichannel interferometer. The interfering object and reference beams are divided into several sets of interfering beams, which are mutually phase-stepped, and the information is recorded with several CCD cameras. At least three parallel channels have to be introduced, which simultaneously produce and record separate interferograms with the required phase shifts. From these phase-stepped patterns, the phase at each image point (x, y) is calculated.

Phase-stepped interferometry is a robust technique for phase measurement. No a priori information of the interferograms is required and no requirements are imposed on the spatial variations of the background, visibility or the phase distribution. Since it is a real-time technique, it has a reduced sensitivity to external disturbances. Also, it is less sensitive to temporal fluctuations of intensity and phase. However, in laboratory environments, the overall accuracy in spatial phase-stepped techniques are usually lower than for the temporal methods. The major disadvantage of multichannel interferometers is their complexity. Special care has to be taken to avoid errors (e.g. detector non-linearity and non-homogeneity, pixel mismatch) and to ensure the stability of the interferometer. Nevertheless, when these problems are mastered, an accurate interferometer can be obtained, which has a highly stable phase step that can be fixed during the assembling of the system. Such an interferometer is capable of measuring almost arbitrary phase patterns in real time. Therefore, this phase stepping technique was employed in the real-time phase-stepped shearing speckle interferometer, described in this thesis.

Although the purpose of the research is the development of a system which can be used for defect detection in industrial environments, real-time shearing interferometry can also be used in various other disciplines of mechanical engineering for deformation measurement and determination of material properties. The possibility to monitor time-dependent processes allows dynamic testing of materials, monitoring transient events or vibration analysis (using stroboscopic illumination to study dynamic behavior of harmonic oscillating objects, e.g. for sound reduction studies).

2.7 Summary

Holographic interferometry and (shearing) speckle interferometry have unique measurement capabilities. Waves are compared interferometrically with sub-wavelength accuracy. Holography can provide a very high resolution but is sensitive to external perturbations. Electronic speckle interferometry has a lower resolution capability but is simpler to implement and use. Shearography makes the interferometric system nearly insensitive to rigid body motion of the object and strongly reduces other external disturbances since only the spatial variations of the displacement in a predetermined direction are measured.

The basic images in speckle interferometry are intensity images, whereas the physical quantity of interest is related to the phase of the interferogram. The phase can be calculated from an interferogram using several techniques. Intensity based analysis methods require only a single interferogram. More accurate results can be obtained using phase sampling techniques. The phase distribution is calculated from at least three phase stepped interferograms. In temporal phase stepping techniques, the phase stepped images are produced and recorded successively in time. In general, this means that one detector will perform several subsequent recordings that have to be stored in order to enable processing of the data. Therefore, it requires absolute stability of the object during the generation and storage of the (at least) three images. In spatial phase stepping methods, the different phase stepped images are generated at the same time. Therefore, spatial techniques are best suited for measuring dynamic events and measurements under hostile conditions. Multichannel interferometers, e.g. based upon polarization phase stepped methods, combine the benefits of phase stepping interferometry with real time operation.

2.8 References

1. T. Kreis, *Holographic interferometry: principles and methods*, (Akademie Verlag, Berlin, 1996).
2. D.W. Robinson, G.T. Reid, *Interferogram analysis*, (IOP Publishing Ltd., London, 1993).
3. R.S. Sirohi, *Speckle metrology*, (Marcel Dekker Inc., New York, 1993).
4. J.C. Dainty, *Laser speckle and related phenomena*, (Springer-Verlag, Berlin, 1975).
5. A.E. Ennos, "Speckle interferometry" in *Progress in Optics* **26**, E. Wolf, ed., (North-Holland Publishing Company, Amsterdam, 1978), pp.235-288.
6. M. Born, E. Wolf, *Principles of optics*, (Pergamon Press, Oxford, 1980), p. 267
7. J.N. Butters and J.A. Leendertz. "Holographic and video techniques applied to engineering measurement", Measurement and control **4**(12), pp. 349-354 (1971).
8. A. Macovski, D. Ramsey, and L. F. Schaefer, Applied Optics **10**(12), pp. 2722-2727, (1971).
9. J.W. Goodman, "Statistical Properties of laser speckle patterns" in *Laser Speckle and related Phenomena*, J.C. Daintly, ed., (Springer-Verlag, Berlin, 1975).

10. A.J.P van Haasteren, *Real-time phase stepped speckle interferometry*, (Ph.D. Thesis, Delft University of Technology, Delft, 1994).
11. R. Jones, C. Wykes, "De-correlation effects in speckle-pattern interferometry 2. Displacement dependent de-correlation and applications to the observation of machine-induced strain", Optica Acta **24**(5), pp. 533-550 (1977).
12. K. Creath, "Phase-shifting speckle interferometry", Applied Optics **24**(18), pp. 3053-3058, (1985).
13. T. Yoshimura, M. Zhou, K. Yamahai Z. Liyan, "Optimum determination of speckle size to be used in electronic speckle pattern interferometry", Applied Optics **34**(1), pp. 87-91 (1995).
14. M.Owner-Petersen, "Decorrelation and fringe visibility: on the limiting behavior of various electronic speckle-pattern correlation interferometers" JOSA A **8**(7), pp. 1082-1089 (1991).
15. R. Spooren, "Standard charge-coupled device cameras for video speckle interferometry" Optical Engineering **33**(3), pp. 889-896, (1994).
16. T. Yoshimura, "Statistical properties of dynamic speckles". JOSA A **3**, pp. 1032-1054 (1986).
17. M. Lehman, "Decorrelation-induced phase errors in phase-shifting speckle interferometry", Applied Optics **36**(16), pp. 3657-3667, (1997).
18. M. Lehmann, "Phase-shifting speckle interferometry with unresolved speckles: A theoretical investigation", Optical Communications **128**, pp. 325-340, (1996).
19. T. Kreis, "Quantitative evalutation of the interference phase" in *Holographic interferometry: principles and methods*, W. Jüptner, W. Osten, eds., (Akademie Verlag, Berlin, 1996) p. 107.
20. D. Gabor, "A new microscopic principle", Nature, **161**, pp. 777-778, (1948).
21. C. M. Vest, *Holographic interferometry*, (John Wiley and Sons, New York, 1979).
22. *Selected papers on Electronic speckle pattern interferometry : principles and practice*, P. Meinlschmidt, K.D. Hinsch, R.S. Sirohi, eds., SPIE Milestone series v. MS132, (SPIE, Washington, 1996).
23. *Interferogram analysis: digital fringe pattern measurement techniques*, D.W. Robinson and G.T. Reid, eds., (IOP Puplishing, Bristol, 1993).
24. K. Creath, J.C. Wyant, "Holographic and speckle test" in *Optical Shop Testing*, D. Malacara, ed., (John Wiley and Sons, New York, 1992), pp. 599-652.
25. O.J. Løkberg, "Recent developments in video speckle interferometry" in *Speckle Metrology*, R.S. Sirohi, ed., (Marcel Dekker Inc., New York, 1993), pp. 157-194.
26. A.R. Ganesan, D.K. Sharma, M.P. Kothiyal, "Universal digital speckle shearing interferometer", Applied Optics **27**(22), pp. 4731-4734 (1988).
27. H.A. Vrooman, *Quantitative analysis of interferograms*, (Delft University Press, Delft, 1991).
28. J. E. Greivenkamp, J.H. Bruning, "Phase shifting interferometry" in *Optical Shop Testing*, D. Malacara, ed., (John Wiley & Sons, Inc. New York, 1992), pp. 501-598.
29. K. Creath, "Phase shifting speckle interferometry", Applied Optics **24**, pp. 3053-3058 (1985).
30. K. Creath, "Temporal phase measurement methods", in *Interferogram analysis: digital fringe pattern measurement techniques*, D. W. Robinson & G.T. Reid, eds., (IOP Publishing Ltd., 1993), pp. 94-140.
31. A.A.M. Maas, *Phase shifting speckle interferometry*, (Ph.D. Thesis, Delft University of technology, Delft, 1991).

32. X. Colonna de Lega and P. Jacquot, "Deformation measurement with object-induced dynamic phase shifting", Applied Optics **35**(25), pp. 5115-5121, (1996).
33. M. Kujawinska, "Spatial Phase Measurement Methods", in *Interferogram analysis: digital fringe pattern measurement techniques*, (IOP Publishing, Bristol , 1993), pp. 141-193.
34. Y.Y. Hung and C. Y. Liang, "Image-shearing camera for direct measurement of surface strains", Applied Optics **18** (7), pp. 1046-1051, (1979).
35. M. Tadeka, H. Ina, S. Kobayashi, "Fourier-transform method of fringe-pattern analysis for computer-based topography and interferometry", J. Opt. Soc. Am. **72**(1), pp. 156-160, (1982).
36. H. O. Saldner, N-E. Molin, K. A. Stetson, "Fourier-transform evaluation of phase data in spatially phase biased TV holograms", Applied Optics **35**(2), pp. 332-336, (1996).
37. J. D. Trolinger, J. C. Hsu, "Flowfield Diagnostics by Holographic Interferometry and Tomography", in *Physical Review Vol. 19*, W. Jüptner, W. Osten, eds., (Akademie Verlag, Berlin, 1993), pp. 423-438.
38. M. Pirga, M. Kujawinska, "Two directional spatial-carrier phase shifting method for analysis of crossed and closed fringe patterns", Optical engineering **34**(8), pp. 2459-2466, (1995).
39. H.Z. Hu, "Polarization heterodyne interferometry using a simple rotation analyzer: theory and error analysis", Applied Optics **22**, pp. 2052-2056, (1983).
40. J. Schwider, "State of the art in evaluation techniques for interferometry" in *Physical Review Vol. 19*, W. Jüptner, W. Osten, eds., (Akademie Verlag GmbH, Berlin 1993), pp. 27-36.

Chapter 3

Description of the system

This chapter presents details of the configuration of the real-time phase-stepped shearing speckle interferometer. In Section 3.1, the two-camera system is described and functional parts are defined. These are worked out in separate sections. The requirements and selection criteria for the recording system and the illumination system are summarized in Section 3.2. In Section 3.3 the shearing head is detailed and in Section 3.4 the imaging system is described. The properties of the Savart elements will be considered in Section 3.5. Finally, Section 3.6 describes the mechanical construction.

3.1 Description of the setup

The real-time phase-stepped shearing speckle interferometer described in this thesis is a four-channel set-up. Four phase-stepped interference patterns are generated and recorded simultaneously. Multi-channel interferometers are very sensitive to accurate positioning and stability of the recording devices. Therefore, strong requirements are imposed on the mechanical construction of such interferometers. To reduce these requirements, our interferometer uses two cameras to record four phase-stepped interferograms. For this purpose, two Savart elements are used as beam splitters. A Savart element splits an incoming beam with random polarization into two mutually shifted beams with mutually orthogonal linear polarization states. Applying retarder plates in front of these elements, the interferograms impinging on the cameras can be given the required phase step. The two interferograms are recorded side by side on the sensitive area of a single CCD camera. Thus, by means of two such cameras a four bucket system is created. A schematical representation of the system is presented in Fig. 3.1.

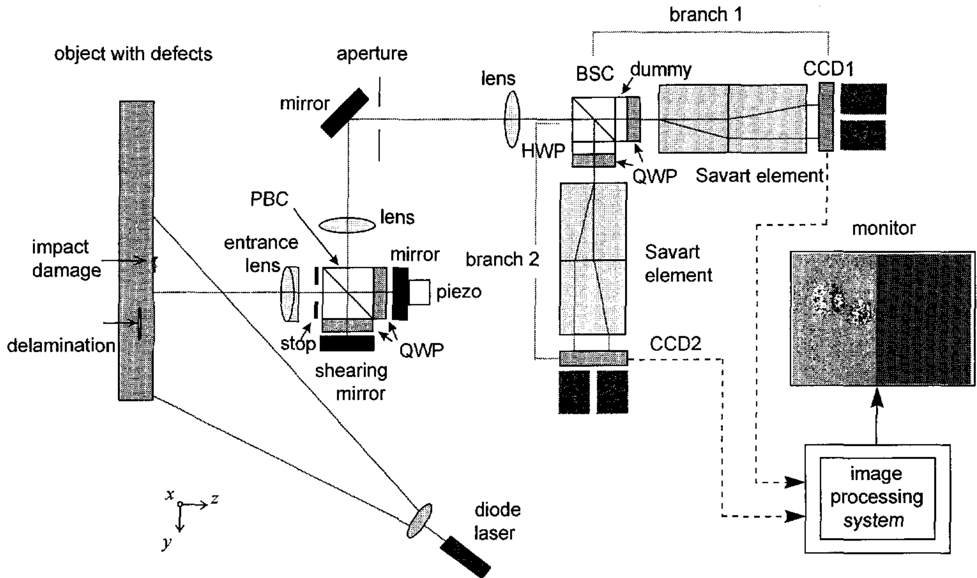


Fig. 3.1 The real-time phase-stepped shearing speckle interferometer

Using two CCD cameras to record four interference patterns reduces the number of picture elements available for each pattern with (at least) a factor two. This reduction of spatial resolution of the interferograms is acceptable, since the goal of our project is to measure defects with a minimum size of 490 mm^2 in an area of 0.25 m^2. This goal can be met using half the detection plane for a single interferogram.

System description using Jones matrices

Since our optical system is non-depolarizing, it can be described using the Jones calculus[1,2,3] and the complex wave representation[4]. By using Jones vectors, a plane parallel wave approximation is introduced, leaving the lenses in our system out of consideration. This is acceptable here, since in this section the polarization properties of the beams are of greater interest than the imaging properties of the optical system and because the numerical aperture of the system is small (usually below 0.06, see Section 3.4). The performance of the system will be somewhat degraded by the converging beams in our interferometer. This is analyzed and described in Chapter 5 of this thesis.

In this description of the system, a new coordinate system is introduced, the right-handed *uvw*-coordinate system. Since the beams change directions several times with respect to the *xyz*-coordinate system, it is more convenient to adopt a coordinate system that travels along with the beams inside the interferometer. Therefore, the light traveling in the system will be described by a two-dimensional vector, propagating in the *w*-direction and with its wavefront in the *uv*-plane. The *uvw*-system is coupled to the *xyz*-system by the *u*-direction, which is parallel to the *x*-direction.

The Jones matrix $\boldsymbol{M}$, defined as

$$\boldsymbol{M} \equiv \begin{pmatrix} m_1 & m_2 \\ m_3 & m_4 \end{pmatrix}, \tag{3.1}$$

represents an optical element. The matrix $\boldsymbol{M}^{(q)}$ of a complete optical system, containing q elements, can be found by multiplying the q separate matrices describing the subsequent elements

$$\boldsymbol{M}^{(q)} = \boldsymbol{M}_q \boldsymbol{M}_{q-1} \ldots \boldsymbol{M}_1 . \tag{3.2}$$

If the wave front of a monochromatic light beam lies in the u,v plane and is traveling in the positive w-direction (i.e. the planar wave approximation), the electric field vector $\boldsymbol{E}(u,v)$ can be expressed as a combination of the two basic components E_u and E_v into the Jones vector

$$\boldsymbol{E}(u,v) = \begin{pmatrix} E_u \\ E_v \end{pmatrix} = \begin{pmatrix} A_u \\ A_v \end{pmatrix} e^{i(kw-\omega t)}, \tag{3.3}$$

where ω is the angular frequency, t represents the time and k is the wave number. If the wave passes through an optical element represented by the matrix $\boldsymbol{M}_q$, then the relation between the incident and transmitted electric field vectors and $\boldsymbol{E}$ and $\boldsymbol{E}'$ of the wave becomes:

$$\boldsymbol{E}' = \boldsymbol{M}_q \boldsymbol{E} . \tag{3.4}$$

The coherent, diverging beam emitted by a semiconductor laser illuminates the object. This light wave is monochromatic and linearly polarized. The object reflects the light diffusely and part of this light is collected by the system and directed towards the CCD cameras, forming a speckle pattern on the cameras. Since we use the planar wave approximation, all waves will undergo the same transformations in the system and hence only a single wave will be considered in this system description. Omitting the time dependency (see Section 2.4), this wave can be described as

$$\boldsymbol{E}(u,v) = \begin{pmatrix} A_u \\ A_v \end{pmatrix} e^{i\varphi(u,v)} . \tag{3.5}$$

The diffusely reflected light is gathered by the entrance lens of the system. Next, this light passes through a polarizing beam splitting cube (PBC), which splits the incoming light into two parts with mutual orthogonal linear polarization directions. The component in the u-direction (also called s-polarized light) is reflected and the one in the v-direction (p-polarized light) is transmitted. For transmission, the transformation matrix $\boldsymbol{M}_{PBCt}$ of the PBC equals

$$\boldsymbol{M}_{PBCt} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \tag{3.6}$$

and for reflection

$$\boldsymbol{M}_{PBCr} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \tag{3.7}$$

Both the transmitted and reflected parts travel through quarter wave plates before impinging on a mirror. These stationary quarter wave plates (QWP), both with their fast axis oriented at -π/4 rad with respect to the u-axis, transform the incoming beam according to

$$\boldsymbol{M}_{QWP,-\pi/4} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix}, \tag{3.8}$$

yielding light that is right-handed circularly polarized for the reflected part and left-handed polarized for the other. Both parts reflect at the mirrors. One of the mirrors can be rotated around the u- and the v-axes to vary the shear in direction and distance. The part of the beam reflecting from this mirror surface is (arbitrarily) regarded as the object beam. The other part of the beam, the reference beam, impinges on the mirror mounted on a piezoelectric element. This mirror can be translated perpendicularly to the incoming beam by means of a piezoelectric element (see Sections 3.3 and 4.1 for the function of this piezoelectric element). As a result of a rotation of the mirror in the object beam path, its phase distribution $\varphi_o(u,v)$ is shifted over (Δu, Δv) with respect to that of the reference beam, as is described by

$$\varphi_o(u,v) = \varphi_r(u + \Delta u, v + \Delta v). \tag{3.9}$$

The reflection of the beams on the mirrors changes the coordinate system u,v,w of the beams, such that during the second pass of the beams through the quarter wave plates, the fast axes are positioned under π/4 rad with respect to the u-axis. The corresponding transformation matrix of the mirror is

$$\boldsymbol{M}_{mirror} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{3.10}$$

and of the QWP at π/4 rad with respect to the u-axis

$$\boldsymbol{M}_{QWP,\pi/4} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}. \tag{3.11}$$

At passing the quarter wave plates for the second time, both the circularly polarized beams are transformed into linear ones, such that the previously reflected beam is now transmitted by the PBC and vice versa. The reference beam now is described by

$$\boldsymbol{E}_r(u,v) = -i\begin{pmatrix} A_v \\ 0 \end{pmatrix} e^{i\varphi_r(u,v)} \tag{3.12}$$

after leaving the PBC. The object beam is described by

$$\boldsymbol{E}_o(u,v) = -i\begin{pmatrix} 0 \\ A_u \end{pmatrix} e^{i\varphi_o(u,v)}. \tag{3.13}$$

Next, supposing that the intermediate lenses and mirror do not alter the polarization states, both beams enter the second beam splitter, which is non-polarizing. This beam splitter generates the fields in the two branches of the interferometer. The path towards CCD1 will be called Branch 1 and the path towards CCD2 is Branch 2. For a non-polarizing beam splitting cube (BSC) the transmission matrix is:

$$\boldsymbol{M}_{BSC,t} = \begin{pmatrix} c_{tu} \\ c_{tv} \end{pmatrix}, \tag{3.14}$$

in which c_{tu} and c_{tv} represent the amplitude transmission coefficients for u- and v- components, respectively. For reflection, the transformation matrix is:

$$\boldsymbol{M}_{BSC,r} = \begin{pmatrix} c_{ru} \\ c_{rv} \end{pmatrix}, \tag{3.15}$$

where c_{ru} and c_{rv} represent the amplitude reflection coefficients for u- and v- component, respectively. If there is no absorption, $|c_{tu} + c_{ru}|^2$ equals unity and the same will be valid for the v-components. Behind the BSC, we have four beams, two in each branch. In Branch 1 we have the contributions from the reference beam

$$\boldsymbol{E}_{r,1}(u,v) = -i\begin{pmatrix} c_{tu} A_v \\ 0 \end{pmatrix} e^{i\varphi_r(u,v)} \tag{3.16}$$

and from the object beam

$$\boldsymbol{E}_{o,1}(u,v) = -i\begin{pmatrix} 0 \\ c_{tv} A_u \end{pmatrix} e^{i\varphi_o(u,v)}. \tag{3.17}$$

In Branch 2 the contributions from the reference and object beams reflected by the BSC occur, described by

$$\boldsymbol{E}_{r,2}(u,v) = -i\begin{pmatrix} c_{ru} A_v \\ 0 \end{pmatrix} e^{i\varphi_r(u,v)}, \tag{3.18}$$

$$E_{o,2}(u,v) = i\begin{pmatrix} 0 \\ c_{rv} A_u \end{pmatrix} e^{i\varphi_o(u,v)}. \tag{3.19}$$

In Branch1, a quarter wave plate is oriented such that its fast axis makes an angle of π/4 rad with both impinging polarization states. As a result it transforms the two linear states into two circular states with opposite rotation directions. These beams pass through a Savart element. Owing to its birefringent nature it simultaneously can be used to split an incoming beam into two components (an ordinary and an extraordinary component) and to set the phase steps. The Savart elements, as used in our set-up, consist of two segments of calcite; each segment is 29.76 mm long. The projection of the optical axes of the segments on the uv-plane are mutually orthogonal, making an angle of π/4 and -π/4 rad with the u-axis (see Fig. 3.2). These angles are determined by the requirements on the image displacement of the element (see Section 3.5).

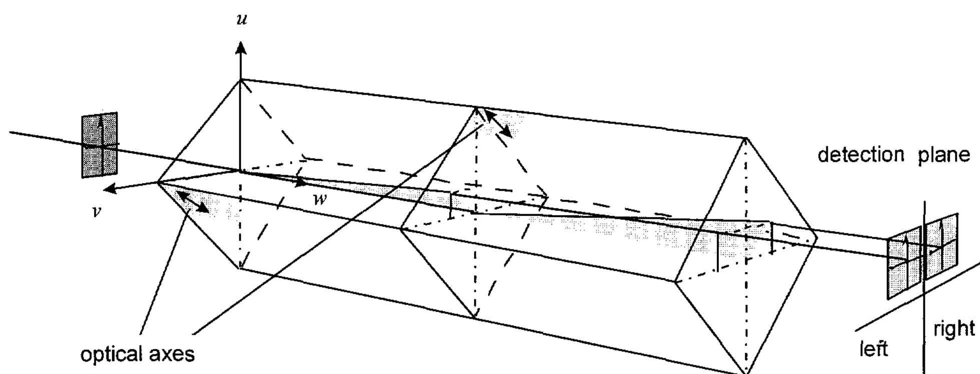


Fig. 3.2 *Savart element.*

A Savart element can be regarded as a linear polarizer with two transmission axes, one for the ordinary beam and one for the extraordinary beam. These transmission axes are mutually perpendicular and make angles of π/4 and -π/4 rad with the u-axis. For a linear polarizer with its axis at π/4 rad to the x-axis, the Jones matrix equals

$$M_{LP,\pi/4} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \tag{3.20}$$

and for -π/4 radians:

$$M_{LP,-\pi/4} = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}. \tag{3.21}$$

Calculating the fields of those parts of the reference and object beams that arrive at the left-hand side of the CCD camera (see Fig. 3.2), we arrive at:

$$\boldsymbol{E}_{r,1}^{left}(u,v) = \frac{(-i-1)c_{tu}A_v}{2\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}e^{i\varphi_r(u,v)} \tag{3.22}$$

and

$$\boldsymbol{E}_{o,1}^{left}(u,v) = \frac{(-i-1)c_{tv}A_u}{2\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}e^{i\varphi_o(u,v)}. \tag{3.23}$$

From these two equations it can be seen that, assuming c_{tu}, c_{tv}, A_u and A_v to be real, the phase step between the object beam and reference beam is zero. On the right-hand side of the CCD camera the object and reference beams are described by

$$\boldsymbol{E}_{r,1}^{right}(u,v) = \frac{(1-i)c_{tu}A_v}{2\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}e^{i\varphi_r(u,v)}, \tag{3.24}$$

$$\boldsymbol{E}_{o,1}^{right}(u,v) = \frac{(1-i)c_{tv}A_u}{2\sqrt{2}}\begin{pmatrix}-1\\1\end{pmatrix}e^{i\varphi_o(u,v)}, \tag{3.25}$$

which show a π rad phase step between the object and reference beam.

In Branch 2, an additional half wave plate is inserted before the quarter wave plate and analyzer. A half wave plate flips the polarization state around the fast axis of the retarder plate. The transformation matrix for a half wave plate with its fast axis at an angle ψ with respect to the u-axis is[1]

$$\boldsymbol{M}_{HWP} = -i\begin{pmatrix}\cos 2\psi & \sin 2\psi\\ \sin 2\psi & -\cos 2\psi\end{pmatrix}. \tag{3.26}$$

The half wave plate is such oriented that its fast axis makes an angle of π/8 rad with the u-axis. This gives the polarization direction an additional tilt of π/4 rad. Behind the HWP, the object and reference beams are described by

$$\boldsymbol{E}_{r,2}(u,v) = \frac{-c_{ru}A_v}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}e^{i\varphi_r(u,v)}, \tag{3.27}$$

$$\boldsymbol{E}_{o,2}(u,v) = \frac{c_{rv}A_u}{\sqrt{2}}\begin{pmatrix}-1\\-1\end{pmatrix}e^{i\varphi_o(u,v)}. \tag{3.28}$$

These beams pass through a quarter wave plate with its fast axis under π/4 with these polarization directions, i.e. it is parallel to the u-axis. The Jones transformation matrix of this element is:

$$\boldsymbol{M}_{QWP,0} = e^{i\pi/4}\begin{pmatrix}1 & 0\\ 0 & -i\end{pmatrix}. \quad (3.29)$$

As a result, two circularly polarized beams enter the Savart element, which is identical to the one in Branch 1 of the interferometer. The object and reference beams impinging on the left-hand side of CCD2 are described by

$$\boldsymbol{E}_{r,2}^{left}(u,v) = \frac{-c_{ru}A_v(1+i)}{2\sqrt{2}}e^{i\pi/4}\begin{pmatrix}1\\1\end{pmatrix}e^{i\varphi_r(u,v)}, \quad (3.30)$$

$$\boldsymbol{E}_{o,2}^{left}(u,v) = \frac{c_{rv}A_u(i-1)}{2\sqrt{2}}e^{i\pi/4}\begin{pmatrix}1\\1\end{pmatrix}e^{i\varphi_o(u,v)}, \quad (3.31)$$

and the beams at the right-hand side of CCD2 are

$$\boldsymbol{E}_{r,2}^{right}(u,v) = \frac{-c_{ru}A_v(i-1)}{2\sqrt{2}}e^{i\pi/4}\begin{pmatrix}-1\\1\end{pmatrix}e^{i\varphi_r(u,v)} \quad (3.32)$$

and

$$\boldsymbol{E}_{o,2}^{right}(u,v) = \frac{c_{rv}A_{ux}(i+1)}{2\sqrt{2}}e^{i\pi/4}\begin{pmatrix}-1\\1\end{pmatrix}e^{i\varphi_{rr}(u,v)}. \quad (3.33)$$

The interference pattern recorded by the left- or right-hand side of the considered CCD camera is calculated using

$$\begin{aligned}I_{CCDi}^{side}(u,v) &= (\boldsymbol{E}_{oi}^{side} + \boldsymbol{E}_{r,i}^{side})\cdot(\boldsymbol{E}_{o,i}^{side} + \boldsymbol{E}_{r,i}^{side})*\\ &= \boldsymbol{E}_{o,i}^{side}\cdot\boldsymbol{E}_{o,i}^{side}* + \boldsymbol{E}_{o,i}^{side}\cdot\boldsymbol{E}_{r,i}^{side}* + \boldsymbol{E}_{r,i}^{side}\cdot\boldsymbol{E}_{o,i}^{side}* + \boldsymbol{E}_{r,i}^{side}\cdot\boldsymbol{E}_{r,i}^{side}* \quad (i=1,2),\end{aligned} \quad (3.34)$$

in which i indicates the camera number, the superscript *side* the side of the camera and * the complex conjugates of the field vectors (of which the dependency on u,v is temporarily omitted). For the left hand side of CCD1, equations (3.22), (3.23) and (3.34) result in

$$I_{CCD1}^{left}(u,v) = \frac{1}{2}c_{tv}^2A_u^2 + \frac{1}{2}c_{tu}^2A_v^2 + c_{tu}c_{tv}A_uA_v\cos(\varphi(u,v)), \quad (3.35)$$

in which $\varphi(u,v)$ equals

$$\varphi(u,v) = \varphi_r(u,v) - \varphi_o(u,v). \quad (3.36)$$

The first term in Eq. (3.35) is the object intensity I_o, the second term the reference intensity I_r and the third term is the interference term. As we can see, this interference pattern has a zero

phase step between the object and reference beams. Equivalent to the calculation for the left-hand side, the interference pattern for the right-hand side of CCD1 can be derived (using Eq. (3.24) and (3.25)) as:

$$I_{CCD1}^{right}(u,v) = \frac{1}{2}c_{lv}^2 A_u^2 + \frac{1}{2}c_{lu}^2 A_v^2 + c_{lu}c_{lv}A_u A_v \cos(\varphi(u,v)+\pi). \tag{3.37}$$

The phase difference between the two interferograms recorded by CCD2 also have a phase shift of π rad between the object and reference beams. As a result of the half wave plate, the interferograms on CCD2 differ π/2 rad with the corresponding interferograms on CCD1, leading to

$$I_{CCD2}^{left}(u,v) = = \frac{1}{2}c_{rv}^2 A_u^2 + \frac{1}{2}c_{ru}^2 A_v^2 + c_{ru}c_{rv}A_u A_v \cos\left(\varphi(u,v)-\frac{\pi}{2}\right) \tag{3.38}$$

and

$$I_{CCD2}^{right}(u,v) = = \frac{1}{2}c_{rv}^2 A_u^2 + \frac{1}{2}c_{ru}^2 A_v^2 + c_{ru}c_{rv}A_u A_v \cos\left(\varphi(u,v)+\frac{\pi}{2}\right). \tag{3.39}$$

Thus, interpreting the intensities at the left- and right-hand sides of CCD1 as $I_1(u,v)$ and $I_3(u,v)$, respectively, and those of CCD2 as $I_4(u,v)$ and $I_2(u,v)$, we obtain the correct phase steps for a four bucket algorithm[5] (phase step values of 0, π/2, π and 3π/2) in which the phase is calculated using

$$\varphi(u,v) = \arctan\left(\frac{I_4(u,v)-I_2(u,v)}{I_1(u,v)-I_3(u,v)}\right). \tag{3.40}$$

The CCD cameras are connected to a pipeline image processing system (Datacube, MaxVideo 200) which calculates the phase at video speed (i.e. 25 times per second). This image processing system will be described in Chapter 4. In the following sections of this chapter, the optical set-up is characterized in more detail. For that purpose, the optical system is divided into several functional components: the illumination system, the shearing head, the imaging system, the phase stepping system and the recording system (see Fig. 3.3).

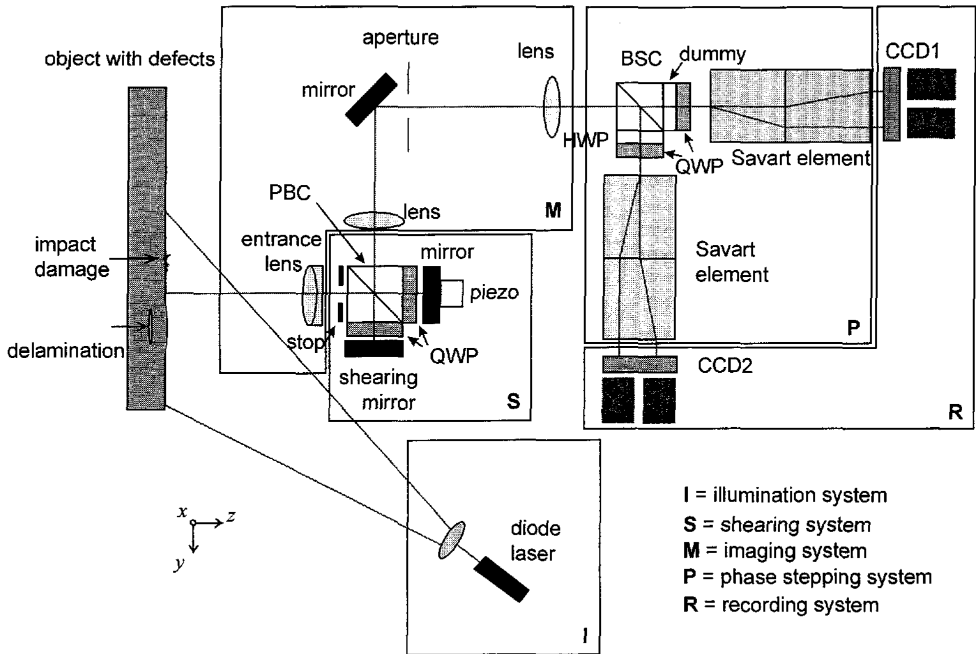


Fig. 3.3 *The functional components of the real-time phase-stepped shearing speckle interferometer.*

Section 3.2 treats the illumination and recording systems. The requirements for these systems are strongly related. Therefore, these systems could not be designed sequentially but they had to be considered together. The shearing device contains the polarizing beam splitting cube (PBC), the two quarter wave plates (QWP) and the two mirrors. It facilitates the relative displacement (shear) between the object and reference beams. The description of this shearing head is given in Section 3.3. The object is imaged onto the detection plane of the cameras, using the imaging system. It embodies an objective lens and a Kepler type telescope, consisting of two identical lenses. As a result, every point in the image plane is directly related to a small area on the object. In the common foci of the telescope lenses a diaphragm is positioned, which can be used to adjust the speckle size in the detection planes of the cameras. Section 3.4 details the properties of the imaging system. The phase stepping system consists of the two quarter wave plates, the half wave plate (HWP) and the two Savart elements. Section 3.5 will describe the Savart elements and the propagation of light through these birefringent elements.

3.2 The illumination and recording systems

To make the set-up transportable, a small laser, such as an air cooled argon ion laser, frequency doubled Nd-YAG laser or a semiconductor laser, should be applied. The use of a semiconductor laser (diode laser) is most favorable, mainly because of its small size, rugged construction, ease of use, low voltage and the absence of a forced cooling. Also, the possibility to use diode lasers in either CW, pulsed or modulated mode make them more versatile than other lasers.

CCD cameras (Charge Coupled Device) are used to record the interference patterns because of their linear response and compatibility with computers. Especially for phase stepping systems, a linear camera response is necessary. Usually, the linearity of the camera is specified by a quantity (gamma) that is defined as the slope of the signal voltage as a function of exposure in a double logarithmic plot. For linear response, gamma equals unity. However, for the best visual contrast on a monitor, a gamma of 0.45 is often applied for standard television cameras. Typically, the output of a CCD camera deviates not more than 5% from the linear response (until the saturation level is reached), if gamma is set to unity.

For optimal performance of the interferometer, the laser and CCD cameras have to be matched. The minimum required laser power is directly related to the sensitivity of the CCD chips, the wavelength of the light and the speckle size. Clearly, a high laser output power P_L per object area A, a high camera sensitivity (i.e. low minimum required illumination S) and a large aperture in the interferometer increase the intensity of the light in the detection plane of the interferometer. Since this intensity is quadratically dependent on the numerical aperture NA, which in turn is proportional to λ / d_s, we introduce a factor R to indicate usable laser and camera combinations, defined as

$$R = \frac{P_L S_\lambda}{AS} \left(\frac{\lambda}{d_s} \right)^2 . \tag{3.41}$$

Here, P_L represents the optical output power of the laser, λ the wavelength of the laser, S is the sensitivity (or minimum required illumination) of the CCD, d_s denotes the average speckle diameter, S_λ indicates the relative sensitivity (with respect to the maximal one) at the laser's wavelength and A the inspection area. The objective of the system is to inspect an area of 0.25 m^2 in a single recording. The factor R can be used to discriminate between useful and useless combinations of laser and camera.

Test measurements with a three camera shearing phase stepping speckle interferometer[6] showed a minimally required value for R of 24. Using Eq. (3.41), a quantitative evaluation of a laser and camera combination can be made (based on specifications), separating the useful combinations ($R > 24$) from the useless ones ($R < 24$). Since these specifications are not always comparable, care has to be taken using this R-factor.

3.2.1 Semiconductor laser

Despite its many attractive properties, a (diode) laser is not an ideal coherent light source. Apart from technical imperfections, the coherence of such a laser is affected by its inherent properties. The broadening of the emission linewidth is a fundamental effect present in any laser. The smaller the laser cavity, the larger this effect. Typically, diode lasers have a cavity length of several 100 micrometers. In lasers with a long cavity, the output beam is strongly enhanced at the resonant wavelength of the cavity. Compared to e.g. gas lasers, which have cavities in the range of tenth of centimeters to more than a meter, the coherence length of diode lasers will be much shorter. Fortunately, requirements on the coherence length are relatively weak for shearing interferometers. In a shearing interferometer, the electric fields originating from two closely positioned object points are combined in the detection plane. To obtain an interference pattern with an optimal modulation depth, the interfering fields must be spatially and temporally coherent: it requires a constant phase relation over at least several centimeters on the object's surface and a sufficient coherence length. The coherence length L_c can be defined as

$$L_c = c\Delta t_c = \frac{c}{\Delta \nu_c} = \frac{\lambda^2}{\Delta \lambda_c}, \tag{3.42}$$

in which c denotes the speed of light, $\Delta \nu_c$ and $\Delta\lambda_c$ represent the spectral bandwidth and the linewidth of the laser and t_c the coherence time. The laser's coherence length should be much larger than the difference in optical path length between the object beam and the reference beam. In shearography, this is generally in the order of microns or millimeters. Diode lasers with a single spectral mode will meet this criterion. The separation $\Delta\lambda$ between adjacent longitudinal modes is determined by the frequency spacing $\Delta \nu$ between the nth and the $(n+1)^{th}$ longitudinal mode in the laser cavity, as given by

$$\Delta\lambda = -\frac{\lambda^2}{c}\Delta \nu, \tag{3.43}$$

in which the frequency spacing $\Delta \nu$ is the spectral distance between two maxima in a Fabry-Perot interferometer with a length L, and cavity medium with a refractive index of n:

$$\Delta \nu = \frac{c}{2nL}. \tag{3.44}$$

Typically, the wavelength separation $\Delta\lambda$ is 0.3 nm for wavelengths around 800 nm. Using this value, Eq. (3.42) shows that the coherence length L_C is in the order of a mm even if two longitudinal modes are present. Therefore, single longitudinal mode operation is required.

The spatial coherence of laser light is associated with the transverse mode structure in the resonance cavity. For lasers resonating in the TEM_{00} mode, all points on the illuminated object

have extremely good spatial coherence. If several transverse modes are present, a non-uniform intensity pattern arises over the illuminated object, i.e. there are regions with low and high intensities. Since the measurement accuracy is directly related to the intensity, the defect detectability in the dark regions is strongly reduced. Therefore, it is benificial to use a laser that emits the fundamental mode only.

At the time of this research, it was not possible to use a single mode diode laser emitting light in the visible range, because of its limited output power. With single mode diode lasers emitting longer wavelengths, higher optical powers can be obtained. Near-infrared diode lasers (i.e. around 1000 nm) offer the highest output powers. We used a 1Watt single mode diode laser emitting in the near infrared at 984 nm*.

3.2.2 CCD cameras

Because of the rather low emission power of diode lasers and the restricted transmittance of the optical system (see Section 3.4), the sensitivity of the CCD cameras is of prime importance in the camera selection[7]. Spectral response is another important camera characteristic. In the optimal case, the laser wavelength coincides with the peak of the spectral response of the photosensitive material of the CCD camera.

The basic building block of a CCD camera is a metal-oxide-semiconductor (MOS) capacitor with silicon as the photosensitive material[6]. The responsitivity of silicon starts at about 300 nm, has a maximum around 900 nm and tends to zero just below 1200 nm (see Fig. 3.4a). The average absorption depth of light in silicon increases with the wavelength of the light. This reduces the spatial resolution of the CCD camera slightly. To avoid this, the infrared response of many cameras is reduced by placing an IR filter in front of the chip. Further, the CCD cameras are often processed to change the spectral response to approximate the human eye (see Fig. 3.4.b). The chip then has a maximum spectral response at around 550 nm and decays to zero just below 1100 nm (see Fig. 3.4b). The red and infrared responses are strongly suppressed, which make these cameras ineffective in combination with the above mentioned 1 Watt diode laser.

---

* Spectra Diode Labs, 5761-A6 MOPA laser, 1 Watt at 984 nm

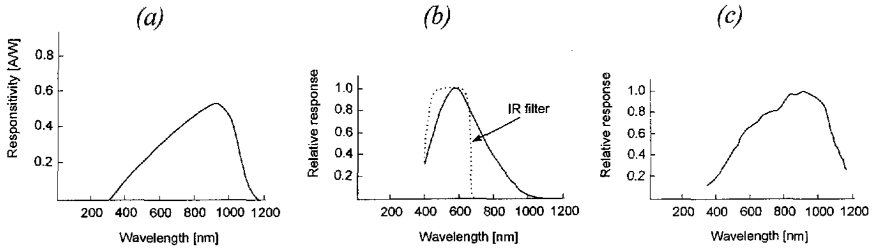


Fig. 3.4 Spectral response of (a) silicon, (b) a CCD camera matched to the human eye and (c) of the Kappa CF7a CCD camera.

Kappa CF7a CCD cameras have been used in our system. This camera type has no infrared reduction features (see Fig. 3.4c). The sensitivity S is 0.2 lux (i.e. 0.3 mW/m^2) and it has a maximal spectral response at 900 nm and about 90% at 984 nm. It accommodates 760×576 picture elements of 11 μm (H) × 11.3 μm (V) μm in a 8.8 × 6.6 mm^2 area. For a usual case in which the speckle size equals the picture element size, the value for R equals 64, which exceeds the critical value of 24, such that this is an usefull combination of camera and laser.

The Kappa CF7a camera is a 50 Hz frame transfer camera. The transfer type of a camera is an important property, since it has a large influence on its performance. There are two types of two-dimensional CCD sensors: interline transfer (IT) CCD's and frame transfer (FT) CCD's. Figure 3.5 shows schematically the configurations of the two types of sensors.

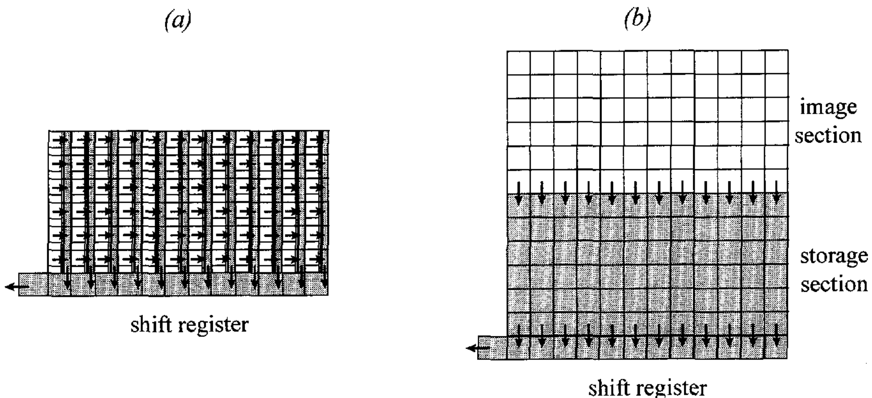


Fig. 3.5 Architecture of (a) interline transfer and (b) frame transfer CCD cameras. The shaded areas indicate optically shielded areas.

The difference between the two types lies primary the in the way the picture elements are read. With an interline transfer sensor, the photon-generated electrons are, after a certain integration period, rapidly transferred from the photocells into the vertical shift registers. These shift registers are optically shielded (indicated as shaded areas in Fig. 3.5). Horizontal rows of

charge packets are then transported simultaneously and eventually end up in a horizontal shift register, which transports the charge towards the sensor output line by line. The even and odd rows of picture elements are read out alternately, each after an integration time of 40 ms (CCIR standard). While the even lines are transported, the integration of the odd lines continues and vice versa. This read-out technique is called 'interlacing'. The two half images, also called TV fields, are combined to form a full image, the so-called TV frame. This read out structure is used to avoid flicker in the monitor image. A frame transfer sensor consists of an image section and an optically shielded storage section. After an integration period of 20 ms (CCIR), all charge packets are rapidly transferred line by line, over the image section to the storage section, after which they are read out through a horizontal shift register. All charge packets are transferred to the output register for each TV field. Therefore, the integration time is halved with respect to the interline transfer CCD. Interlacing is achieved by alternatingly collecting charge packets under different electrodes as depicted in Fig. 3.6.

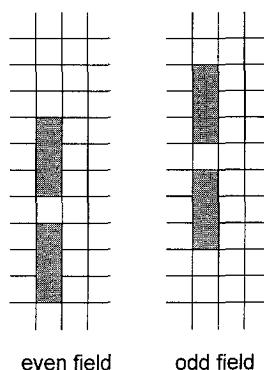


Fig. 3.6 *Picture elements of interlaced fields in a frame transfer CCD camera.*

As a consequence of the difference in transfer techniques, the performance characteristics differ. There is a remarkable difference in fill factor, defined as the fraction of light-sensitive area per total area in a picture element, between interline and frame transfer type CCD cameras. The interline transfer sensor has a fill factor of only 20 to 40%. In a frame transfer device, nearly the whole surface area is photosensitive, resulting in a fill factor of almost 100%. Thus, frame transfer CCD sensors usually provide a better sensitivity than interline transfer CCD sensors when the same photosensitive material is applied. Moreover, thanks to the large integration area per pixel, frame transfer CCD's usually offer a larger dynamic range than interline transfer CCD's. The dynamic range is mainly determined by the maximum number of electrons a picture element can hold, which is set by the volume of the depletion regions. The main disadvantage of frame transfer cameras is associated with pulsed laser sources. To analyze high speed transient phenomena, pulsed laser sources are needed. If a frame transfer CCD camera is used together with a pulsed laser, an interferogram cannot be recorded in the even and uneven lines simultaneously. Therefore, only half of the spatial resolution is obtained.

Generally, the readout of a CCD sensor results in video jitter, i.e. pixels do not have the same position in subsequent digitized frames. This can cause considerable errors. Besides this, it is necessary to synchronize the exposure of the cameras. The Kappa CF7a camera can fulfill these synchronization requirements demanded by the computational system (see also Chapter 4). Besides the video signal it also exports the pixel clock signal. This enables pixel synchronous digitization, thus avoiding video jitter. Moreover, both CCD cameras are locked on the same synchronization signals.

The Kappa CF7a camera consists of three parts: a small head (containing only the sensitive chip), mounted on a Printed Circuit Board (see Fig. 3.7), a small electronic unit (connected to the remote head with a 30 cm flat cable), containing the electronics for processing the sensor signals and a larger electronic unit containing the power supply for the two cameras. The latter also generates the H- and V- synchronization signals and the pixel clock signal on which the two cameras are locked and it facilitates the transfer of the video signal and the synchronization signals to the pipeline image processing devices (see also Chapter 4).

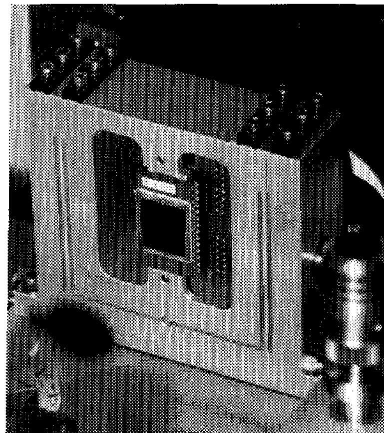


Fig. 3.7 *The CCD chip mounted on a printed circuit board in a translation stage.*

The remote head version of the camera system is necessary in order to enable integration of the cameras in a stable and compact mechanical construction. The small electronic units of both cameras are positioned underneath the base plate of the interferometer. In order to keep the electronic noise under acceptable values, the distance between the head and this electronic unit could not exceed 30 cm. The power supply can be positioned away from the interferometer (in our case at a distance of maximally 5 meter).

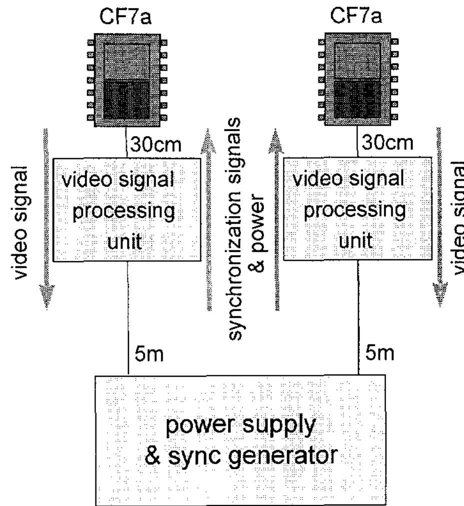


Fig. 3.8 *Schematical presentation of the recording system.*

3.3 The shearing head

In shearing interferometry, the wave front under test interferes with a sheared version of itself. There are several ways to introduce shear in an interferometer, namely by lateral, radial, rotational shifting, by inversion or by folding[8]. In lateral shearing interferometers, the interference pattern resulting from interference between two mutually displaced (or sheared) identical wavefronts is recorded (see Section 2.5 for a mathematical description). Orienting the shear such that it is in the same direction as the periodic structure of an object and adjusting the shear distance such that it is equal to an integer multiple of the structure's period, the object's structure becomes invisible in the measured deformation, which leads to an improved defect detectability[6]. For this reason, we prefer lateral shearing in our defect detection system, which is, among others, applied to (periodic) honeycomb structures.

Our shearing device has the configuration of a Michelson interferometer. The primary advantage of this type of set-up is the capability to vary both the direction and the magnitude of the shear (and hence the sensitivity of the system) over a continuous range. They can be varied easily and independently from each other by rotating the shearing mirror around the x- and y-axes. If a large shear is used, the measurement's sensitivity is increased. But it also results in a larger region in the image where no interference occurs. In Fig. 3.9, the lateral shearing device as it is implemented in our system is depicted. It consists of a polarizing beam splitting cube (PBC) , two quarter wave plates (QWP) and two mirrors.

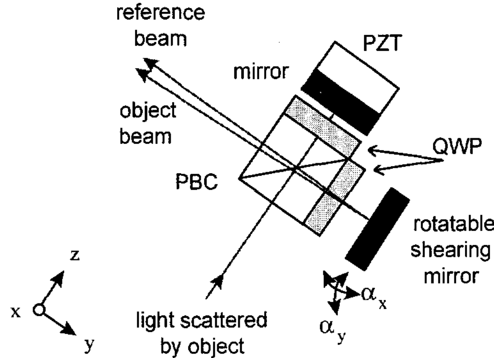


Fig. 3.9 *The shearing device.*

Using this combination of quarter wave plates and a PBC enables the use of nearly 100% of the light that passes the field stop that is positioned just in front of the shearing device. No light is reflected in the direction of the object, as is the case in a usual Michelson interferometer. Thus, except for the reflection losses and possible imperfect alignment or characteristics of the quarter wave plates or PBC, no light is lost, which is an important feature in low light level applications. At the same time, the polarization states of the two beams leaving this head are determined in accordance with the requirements of the phase-stepped interferometer.

The PBC in the shearing device splits the light scattered by the object into two linear polarized parts with mutually orthogonal polarization directions, as is described in Section 3.1. The beam travelling to the adjustable mirror is arbitrarily called the object beam, the other is the reference beam. The reference mirror is mounted on the piezoelectric element and positioned accurately perpendicular to the incoming beam. This element is used for initialization and testing purposes (see Section 4.2) and remains in a fixed position during measurements. By tilting the shearing mirror, a shear is introduced. For a small shear angle, the phase pattern, measured by this interferometer, approximately describes the deformation gradient of the object in the shear direction (see also Section 2.5). Using larger shear angles, the device gradually deviates from a common-path mode of operation. This effect is a potential error contributor since it reduces the tolerance to environmental disturbances.

We use a polarizing beam splitting cube, constructed from a pair of matched right angle prisms cemented together, with a multilayer dielectric coating at the interface. This polarizer transmits linearly polarized light with the plane of the electric field vector parallel to the plane of incidence (p-polarized) and reflects the orthogonally polarized (s-polarized) light. Cube polarizers are more rigid in construction than plate polarizers. In addition, the quarter wave plates can easily be mounted onto the polarizing cube's surfaces (see Fig. 3.10). This enables the set-up to be compact and reduces its sensitivity to external variations, such as vibrations, temperature changes etc.

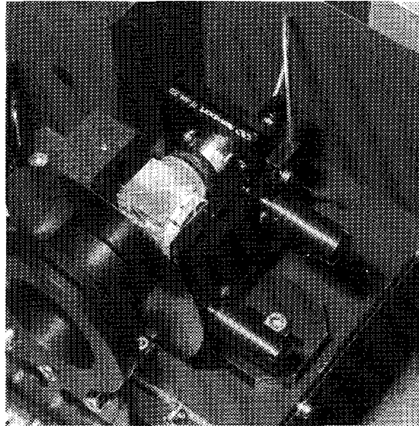


Fig. 3.10 *The shearing device.*

Since the PBC is used in a convergent beam, it will generate spherical aberration. As we are using small numerical apertures and the aberration of the object beam is nearly the same as that of the reference beam since they travel along comparable paths, this will only have a small effect on the measurement accuracy. Apart from imaging aberration, also polarization aberration[2] is present. There are two major contributions to polarization aberration. The first contribution is due to the intrinsic geometry of the polarizing beam splitting cube and the non-collimated beam. The *s*- and *p*-polarizations of a non-perpendicular incoming beam do not coincide with the polarizing axes of the beam splitter. As a result, some *p*-polarization is reflected and *s*-polarization is transmitted. This coupling of the *s*- and *p*- polarizations is usually low in applications with small angular deviation from the optimal value. Also, nearly 100% of this unwanted light propagates in the direction of the object and not further into the system, since at the second pass of the PBC, again only a small amount is reflected where it should have been transmitted and vice versa. Since the coupling between *s*- and *p*-polarization is stronger for angles in the vertical plane than in the horizontal plane[9], this effect will be reduced if we use a horizontal shear. The second contribution to the polarization aberration is caused by the angular dependence of the beam-splitting coating, which can differ between different coatings. How these two effects influence the accuracy of the system is discussed in Section 5.1.

We implemented zero order (narrow bandwidth) retarding plates, consisting of oriented polymers cemented between fused silica windows. The polymers used in these retarders have a birefringence that is nearly constant with wavelength. It is much less dependent on wavelength and temperature than multiple order retarders and the retardance is less dependent of tilt around its fast or slow axis[3]. It also has a higher tolerance on the deviation from perpendicular incidence of the incoming beam.

The mirrors are mounted about 5 mm behind the quarter wave plates. We have used 'hot mirrors', i.e. mirrors reflecting only the near infrared and infrared light, and therefore prevent the visible ambient light to enter the system, which otherwise might reach the sensitive area of

the CCD cameras. The mirrors are glued upon strongly absorbing glass filters that will absorb the light as far as it is transmitted by the mirrors.

3.4 The imaging system

The imaging system of our real-time phase-stepped shearing speckle interferometer comprises an objective lens and a Kepler type telescope. The objective lens generates an intermediate image of the object, which is duplicated by the telescope. The image plane of the objective lens is approximately 10 mm behind its mounting, which is insufficient to place the required optical elements for shearing and phase stepping. Therefore, a Kepler type telescope is used to shift the image plane, thus providing the necessary space. The image plane of the objective lens coincides with the object plane of the telescope. The object distance of the telescopic system must be large enough to provide the necessary space for the shearing head, whereas the image distance should be large enough to accommodate the phase stepping elements (see Fig. 3.11).

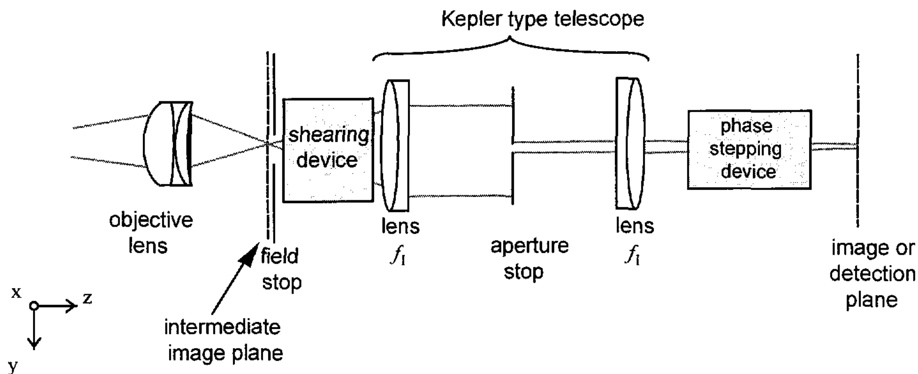


Fig. 3.11 The imaging system.

In each branch of the interferometer, two images have to be recorded, which are projected side by side onto a single CCD surface. The sensitive area of the CCD chip is 6.6 × 8.8 mm², which results in an upper limit of the image size at the detection plane of 6.6 × 4.4 mm². Since the system has to be able to inspect a 0.25 m² area in a single recording, the total magnification of the objective lens in combination with the telescopic system has to satisfy

$$\left|M_{obj} M_{tel}\right| = 10^{-2} . \tag{3.45}$$

Here, M_{obj} and M_{tel} are the lateral magnifications of the objective lens and of the telescope, respectively. The intermediate image, created by the objective lens, is inversely reproduced at the photosensitive surfaces of the CCD cameras by the telescopic system. Since the intermediate image itself is already a reversed reproduction of the object, produced by the

objective lens, the orientation of the resulting image in the image plane is the same as the object.

The astronomical or Kepler telescope[11] is normally used to view at objects at infinity. The back focal plane of the first lens f_1 coincides with the front focal plane of the second lens f_2. If the aperture stop, which is required to control the speckle size, is placed in the common focal planes of the lenses f_1 and f_2, the exit and entrance pupils of the telescope are at infinity, i.e. the telescope is telecentric. This is advantageous since this way the aperture of the system can be altered without changing the field of view. There is no vignetting by the aperture stop. Moreover, the principal rays in the image space are parallel to the z-axis, leading to an optimal performance of the elements in the phase stepping system.

Using this imaging system with its intermediate image has another advantage. The objective lens can be changed without affecting the imaging beyond the intermediate image plane. Furthermore, to avoid overlap of the images on the CCD surfaces, a field stop is used. This field stop must be sharply imaged on the CCD cameras and is therefore placed in the intermediate image plane.

Since we expect a low light level in our system, a very important factor in the design of the imaging system is its light gathering power. Moreover, to obtain a stable and rigid system, the total system should be as small as possible. How this affects the design of the imaging system is analyzed in the next subsections.

3.4.1 The objective lens

In Fig. 3.12 some parameters are depicted which will be used in the subsequent calculations. Ω_0 and Ω_1 indicate the solid angles subtended by the entrance pupil from the axial object point and the exit pupil from the axial image point, whereas θ_0 and θ_1 represent the convergence angle in the object and image space respectively. ϕ_0 and ϕ_1 are the angles between z-axis and the principal rays through the object point (x_0, y_0) and the image point (x_1, y_1) respectively.

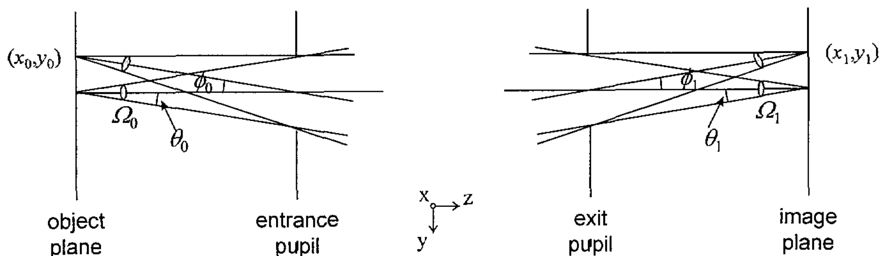


Fig. 3.12 *Parameters in the object and image spaces.*

We assume that the object with a surface S_0 is uniformly illuminated with a total energy flux per unit time Φ and is radiating according to the Lambert's law (i.e. its brightness B_0 is independent of the propagation direction). Then, the radiant flux F_1 at a point of the intermediate image plane satisfies[12]:

$$F_1 = B_0 \pi \sin^2(\theta_1) \cos^4(\phi_0), \tag{3.46}$$

where the brightness B_0 of the object is given by

$$B_0 = \frac{\Phi}{\pi S_0}, \tag{3.47}$$

if the object is radiating in all directions in an infinite half-space between the object and the image. The interferometer is located relatively far away from the object, such that the angle θ_0 is small and the solid angle in the object space Ω_0 can be approximated with

$$\Omega_0 = \pi \sin^2(\theta_0), \tag{3.48}$$

and in the image space Ω_1 is approximately equal to

$$\Omega_1 = \pi \sin^2(\theta_1). \tag{3.49}$$

This changes Eq. (3.46) into

$$F_1 = B_0 \Omega_1 \cos^4(\phi_0). \tag{3.50}$$

Combining (3.47) and (3.50) and using

$$\Omega_0 = M_{obj}^2 \Omega_1, \tag{3.51}$$

the radiant flux F_1 at a point of the intermediate image plane satisfies

$$F_1 = \frac{\Phi \Omega_0 \cos^4(\phi_0)}{\pi S_0 M_{obj}^2}. \tag{3.52}$$

From this it is seen that a larger value of ϕ_0 gives a lower radiant flux, which means that the average intensity reduces towards the edges of the detection plane. Choosing an objective lens with a long focal length reduces this problem since then ϕ_0 will be small. However, a long distance between object and measuring system is impractical and it reduces the light gathering power of the imaging system since Ω_O decreases. As a trade-off, we have chosen to use a compound objective lens with a (fixed) focal length of 10.5 mm for our system. Using the 10.5 mm objective lens, a surface of 0.25 m^2 can be imaged in the intermediate image plane at a

distance of 1 m. The diameter of the lens is 10 mm, which is sufficiently large not to limit the light gathering power of the complete optical system. This is namely determined by the aperture stop in the telescope. To add flexibility to the system, a zoom lens would have been preferred. Unfortunately, standard camera (zoom) objectives show a low transmittance for near infrared light due to the anti-reflection coatings for visible light on the surfaces. As an example, the transmission of a Fujinon C0305081 camera objective is 60% and for a Micro-Nikkor 55 mm camera objective the transmission is only 1% using 984 nm light.

3.4.2 The Kepler telescope

The Kepler type telescope is used to shift the intermediate image plane to the detection plane, thus providing the necessary space for the interferometer's optical elements. The object distance of the telescopic system must be large enough to provide the necessary space for the shearing head. The shearing head is composed of a polarizing beam splitting cube, with edges of 25 mm, two quarter wave plates of 7 mm thickness, glued onto the sides of the beam splitting cube, and two mirrors. These elements require a minimum object distance of 64 mm. Some extra space is required in between the wave plates and the mirrors, since the shearing mirror is rotatable and the other mirror is translatable. Actually, an object distance of 93 mm was chosen. Since the optical path length of a ray in a medium of refractive index n is equal to n times its geometrical path length, the minimum optical path length from the object plane to the first lens of the telescope is 70 mm.
The image distance of the telescopic system must offer enough space for positioning the elements of the phase stepping system. A non-polarizing beam splitting cube, sized 25.4 mm is part of it. In Branch 2 of the interferometer (see Fig. 3.1), the quarter- and half-wave plates are attached to this cube and in Branch 1 a quarter wave plate and a dummy plate are fixed. This dummy plate is made from the same material (fused silica) as the windows of the wave plates and is equal in thickness. It is implemented to avoid an optical path length difference in the two branches. The wave plates and the dummy each have a thickness of 7 mm. Next, in both branches the Savart elements of 59.5 mm are positioned. In total, the required minimum image distance of the telescope is 100 mm. Including some extra space for mounting, the minimum image distance is 143 mm. Again taking the refractive indices of the materials into account, the resulting minimum optical path length from the second lens of the telescope to its image plane becomes 110 mm.

With the industrial applications in mind, the portability of the interferometer head was a key factor in its design. In order to be mounted on a small base plate and to obtain a compact system, the path lengths in the imaging system should be minimized. The total length of the telescope, d_{tel}, is equal to the sum of their focal lengths f_1 and f_2, as is given by

$$d_{tel} = f_1 + f_2 . \tag{3.53}$$

Using the Gaussian lens formula[13], (i.e. a paraxial approximation for thin lenses), the following relation can be deduced for a Kepler type refracting telescope with a magnification M_{tel}

$$f_2 = \frac{z_0 + \left(\frac{1}{|M_{tel}|}\right)^2 z_i}{\left(\frac{1}{|M_{tel}|}\right)^2 + \frac{1}{|M_{tel}|}}, \tag{3.54}$$

in which the lateral telescope magnification satisfies

$$|M_{tel}| = \frac{f_2}{f_1}. \tag{3.55}$$

The distance between the first lens of the telescope and the object plane is represented by z_o and z_i represents the distance between the second lens and the image plane. Thus, the total length d_{tel} of the telescope can be written as

$$d_{tel} = z_0 |M_{tel}| + \frac{z_i}{|M_{tel}|}. \tag{3.56}$$

Fig. 3.13 displays a graphical representation of Eq. (3.56), using the minimum required values of 70 mm for z_o and 110 mm for z_i.

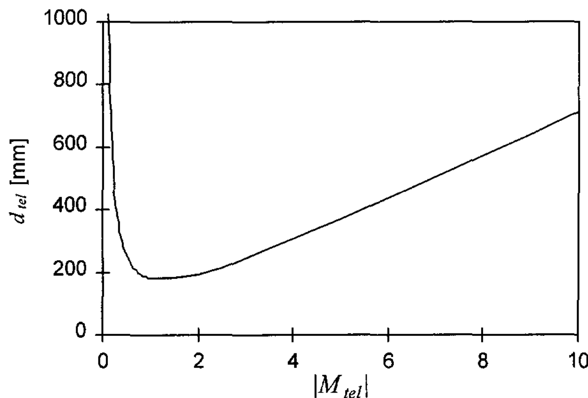


***Fig. 3.13** Length of the telescope for different magnifications.*

The telescope has a minimum length if

$$|M_{tel}| = \sqrt{\frac{z_i}{z_o}}, \tag{3.57}$$

which occurs in our case at $|M_{tel}|$=1.25. If $|M_{tel}|$=1 is chosen, the telescope length is nearly minimal (d_{tel}=180 mm instead of 175.5 mm, i.e. 2.6% larger) and has the advantage that the focal distances of the lenses can have available values.

Focal length of the lenses

The focal lengths of the two lenses in the Kepler type telescope have to be equal since we have chosen for minus unit magnification. The requirements for the minimum object distance and image distance determine the required focal length of the lenses in the telescopic system. Satisfying these requirements, the focal lengths of the telescope lenses should be at least 110 mm, if the object and image distances are to be equal. However, a more compact system can be obtained using unequal object and image distances.

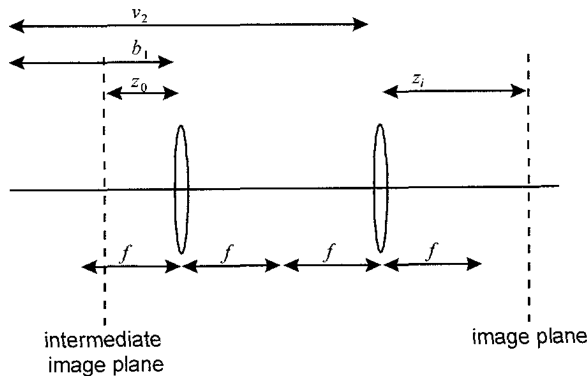


Fig. 3.14 *Definition of distances.*

Choosing for minus unit magnification, we obtain from the Newtonian paraxial imaging formula for thin lenses

$$z_o b_1 = f^2, \tag{3.58}$$

for the first telescope lens while for the second telescope lens we have

$$v_2 z_i = f^2, \tag{3.59}$$

in which b_1 indicates the image distance of the first lens with a focal length f, and v_2 the object distance of the second lens, also having a focal length of f. It can easily be shown that

$$\Delta z_o = -\Delta z_i, \tag{3.60}$$

if the telescope is correctly aligned such that the back focal plane of the first lens coincides with the front focal plane of the second lens. This relation shows that if the object distance is reduced, the image distance is increased by the same amount. In our case, where the minimum required image distance is considerably larger than the object distance, it is advantageous to decrease the object distance. Telescope lenses with a much smaller focal length can be chosen (90 mm instead of 110 mm), which decreases the total size of the optical system (i.e. the distance between the intermediate image plane and the detection plane, see Fig. 3.11) by 80 mm.

Light gathering power

Using a telescope with minus unit magnification, the intermediate image is inversely reproduced at the detection plane. For a lossless telescopic system, the brightnesses in the intermediate image plane and the detection plane are equal. Due to the presence of a circular aperture stop with diameter D in the focal plane between the two telescope lenses (see Fig. 3.15) the brightness of the images in the detection plane, represented by B_2, is reduced by a factor Q according to

$$B_2 = QB_1 . \tag{3.61}$$

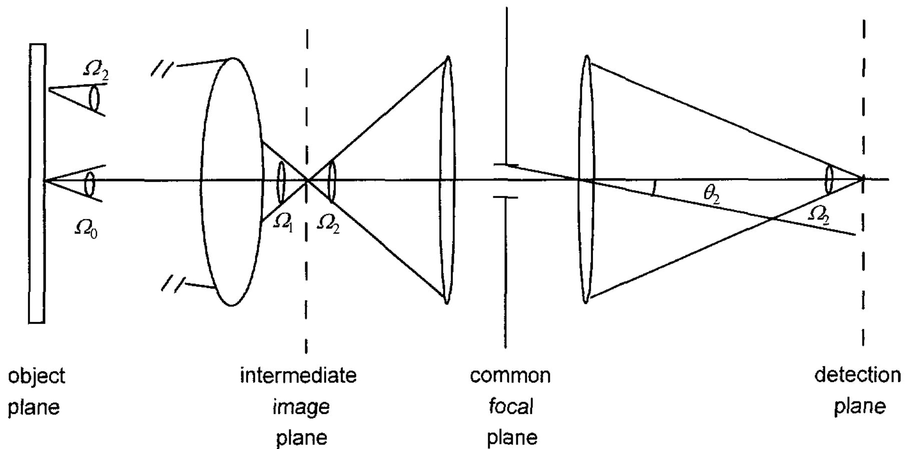


Fig. 3.15 *Additional parameters in the imaging system.*

The reduction factor Q is determined by the ratio of the solid angle with and without the aperture stop in the telescope

$$Q = \frac{\Omega_2}{\Omega_2'} , \tag{3.62}$$

in which the $\Omega_2^{'}$ prime indicates the solid angle for the case that no aperture stop is present in the system. The illumination F_2 in the detection plane follows from

$$F_2 = \frac{Q\Omega_0 \cos^4(\phi_0)}{\pi S_0 M_{obj}^2 M_{tel}^2}\Phi. \tag{3.63}$$

Without the aperture stop, the relation between the solid angles Ω_1 and $\Omega_2^{'}$ is determined by the magnification of the telescope and objective lens. In this particular case, the solid angle at the exit pupil of the telescope $\Omega_2^{'}$ satisfies

$$\Omega_2^{'} = \frac{\Omega_1}{M_{tel}^{\;2}} = \frac{\Omega_0}{M_{tel}^{\;2} M_{obj}^{\;2}}. \tag{3.64}$$

By placing the aperture stop in the common focal plane of the lenses of the telescope, the solid angle at the exit pupil Ω_2 is decreased to

$$\Omega_2 = \pi \sin^2(\theta_2) \approx \pi \tan^2(\theta_2) = \pi\left(\frac{D}{2f_2}\right)^2. \tag{3.65}$$

Combining Eq. (3.62), (3.64) and (3.65) yields

$$Q = \frac{\pi\left(\frac{D}{2f_2}\right)^2 M_{tel}^{\;2} M_{obj}^{\;2}}{\Omega_0}. \tag{3.66}$$

Using this relation, Equation (3.63) becomes

$$F_2 = \left(\frac{D}{2f_2}\right)^2 M_{tel}^{\;2} M_{obj}^{\;2} \frac{\cos^4(\phi_0)}{S_2}\Phi, \tag{3.67}$$

in which S_2 represents the size of the detection plane. From Eq. (3.67) it can be seen that inspecting larger objects, i.e. if $M_{tel}^2 M_{obj}^2$ decreases, the illumination in the detection plane reduces. Moreover, since the relation between the telescope and objective lens magnifications is fixed (see Eq. (3.45)), the light gathering power of the imaging system is determined by the aperture's diameter *D*. A telescope magnification of minus unity is preferred since this yields a telescope length close to smallest value, while identical lenses can be used.

Lateral sizes of the lenses

The diameter of the lenses are determined by the required image size (i.e. 6.6 × 4.4 mm^2), the image distance and the maximal numerical aperture. Assuming a maximal numerical aperture

NA_{max} in the telescope of 0.06 (see next section), the minimal lateral size D_l of the lenses is 21 mm. We have chosen achromatic doublets, with a diameter of 25 mm.

3.4.3 The numerical apertures

The numerical aperture in the telescope NA_2 is determined by the diameter of the aperture stop D and the focal distance of the second telescope lens f_2 (see also Eq. (3.65))

$$NA_2 = \sin(\theta_2) \approx \tan(\theta_2) = \frac{D}{2f_2}. \qquad (3.68)$$

As can be seen from Eq. (3.67), a large value for D / f_2 gives a high intensity in the detection plane. However, this value is restricted. The maximal numerical aperture is determined by the minimum allowable speckle size, the size of the Savart elements and its performance for non-perpendicular incoming rays.

The average speckle diameter d_s at the detector plane and the numerical aperture NA_2 in the image space of the telescope are related according to[14]

$$d_s \approx 0.61 \frac{\lambda}{NA_2}. \qquad (3.69)$$

Thus, increasing the aperture, the average speckle size decreases. In Section 2.2, the influence of the speckle size on the performance of a speckle interferometer is discussed. It showed that the optimum speckle depends on the type of interferometer used and the geometry of the measurement. Large speckles reduce camera decorrelation in a multichannel interferometer, whereas small speckles decrease the object induced decorrelation. As a trade off, we have chosen a maximum numerical aperture NA_2 of 0.06, such that the average speckle size can be chosen equal to that of the picture elements.

The maximum numerical aperture is also affected by possible vignetting by optical elements. The Savart elements are the most restrictive elements in this respect. The Savart elements are positioned in the image space of the telescope. Converging beams travel through these elements towards the detection plane. Therefore, the size of the entrance face of the Savart elements limits the maximally allowable convergence of the beams. Due to the size of the entrance face, i.e. with sides of 15.0×15.0 mm^2 (see also Section 3.5), 60 mrad is the largest convergence angle before vignetting occurs, when purely the size of the Savart elements is taken into account. This allows us to chose an average speckle size equal to a picture element (11 μm), without vignetting. If smaller speckles are used, some vignetting will occur.

An additional effect related to the Savart elements is present in our system, which influences the optimum numerical aperture size. The Savart elements are designed for collimated beams

under normal incidence. However, in our setup, the beams are converging inside the Savart elements. In Fig. 3.18, the maximum convergence is indicated with the angle θ_2. Moreover, since the Savart element is used for image position manipulations by rotating the element around the v-axis and w-axes (see also Section 5.2), non-perpendicular incidence occurs. The maximum angle between the w-axis and the entrance plane of the Savart element is indicated by γ in Fig. 3.16.

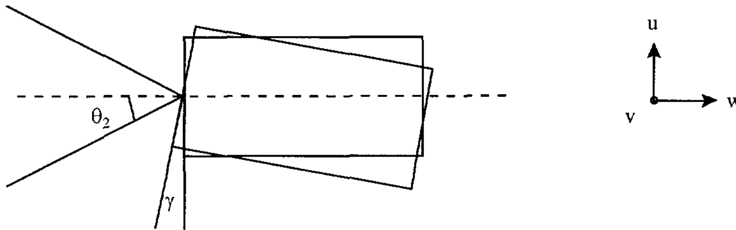


Fig. 3.16 *Maximum deviation from perpendicular incidence at the Savart element.*

The non-perpendicular rays are refracted differently from the perpendicular ones. Therefore, a reduced performance of the Savart elements can be expected, and this performance reduction is related to the used aperture. In Section 5.4 a detailed analysis of the performance of the Savart elements at larger apertures will be given. However, measurements have shown (see Section 5.4) that large apertures are benificial in our low light level applications, despite the performance degradation of the Savart elements, since this degradation is more than compensated for by the higher modulation intensities at the detection plane, resulting in higher measurement accuracies.

In conclusion, we have chosen a maximum numerical aperture of the beams of 0.06, such that speckles equally sized as the picture elements can be used. Some vignetting will occur due to the finite dimension of the Savart elements when a larger aperture than 0.06 is chosen.

3.5 The Savart elements

The Savart elements in the system have three functions. First, they split the incoming set of an object and a reference beam into two laterally shifted sets of combined contributions from the object and reference beams. Secondly, they introduce a phase shift difference of π radians between the combined beams of one set as compared to the other (see also Section 3.1). Thirdly, the Savart elements are used in the process of aligning the interference patterns on the CCD cameras. The latter function will be discussed in detail in Section 5.2. The present section discusses the optical properties of the Savart elements and its constituent material as far as it is relevant for our system.

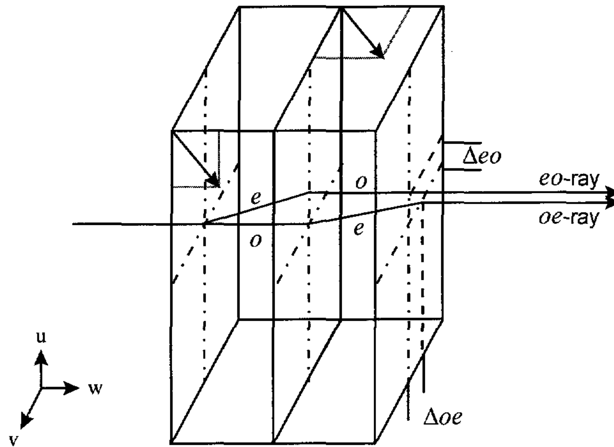


Fig. 3.17 *Savart polariscope.*

In Fig. 3.17, a Savart element, or Savart polariscope[15], is depicted. It consists of two identical segments of uniaxial birefringent material which are cemented together. The optical axes of the segments make an angle of π/4 rad to the entrance face normal and their projections on this entrance face are mutually perpendicular. In Fig. 3.17, the optical axes are indicated by the arrows and lie in the vertical and horizontal facets of the segments. The main property of a Savart element is that it separates an incoming ray with arbitrary polarization into two separate rays with orthogonal linear polarization states: the ordinary and the extraordinary ray. The polarization direction of the ordinary ray in one segment is that of the extraordinary ray in the other. This arrangement ensures that the phase difference vanishes for normal incidence, assuming that the segments are perfectly aligned and have equal lengths. In Fig. 3.17 the splitting of an incoming ray into an ordinary (indicated as *o*) and an extraordinary ray (indicated as *e*) is depicted as well as their interchange at the interface between the two segments of the Savart elements. The *eo*-ray is displaced vertically over a distance Δu and the *oe*-ray horizontally over Δv.

Savart elements can be made from, e.g., quartz (SiO_2), calcite ($CaCO_3$) or sodium nitrate ($NaNO_3$). Our Savart elements have been made from calcite, which is a highly birefringent material for visible wavelengths. Some other materials exhibit even larger birefringence (e.g. sodium nitrate). However, calcite is preferred because of its relatively high optical quality, availability in larger sizes and lack of problems[16,17], like being overly fragile, poisonous or optically unprocessable. As an example, sodium nitrate shows a larger birefringence (45% at 589.2 nm) than calcite but is severely hygroscopic, which would make its application more complicated. Optically pure calcite is found naturally and is mined from various locations throughout the world. It is also grown artificially[18], but then the maximum dimensions are limited to 3 to 4 mm.

The Sellmeyer dispersion formulas for the *o*- and the *e*-ray in calcite are as follows[19]

$$n_o^2 - 1 = \frac{0.8559\lambda^2}{\lambda^2 - (0.0588)^2} + \frac{0.8391\lambda^2}{\lambda^2 - (0.141)^2} + \frac{0.0009\lambda^2}{\lambda^2 - (0.197)^2} + \frac{0.6845\lambda^2}{\lambda^2 - (7.005)^2} \tag{3.70}$$

$$n_e^2 - 1 = \frac{1.085\lambda^2}{\lambda^2 - (0.07897)^2} + \frac{0.09881\lambda^2}{\lambda^2 - (0.142)^2} + \frac{0.317\lambda^2}{\lambda^2 - (11.468)^2}. \tag{3.71}$$

These equations are valid from 200 to 2200 nm and the obtained refractive indices are accurate within a few parts in the 5th decimal. In Fig. 3.18 the refractive indices as a function of the wavelength are depicted as well as the birefringence Δn, defined as

$$\Delta n = n_e - n_o. \tag{3.72}$$

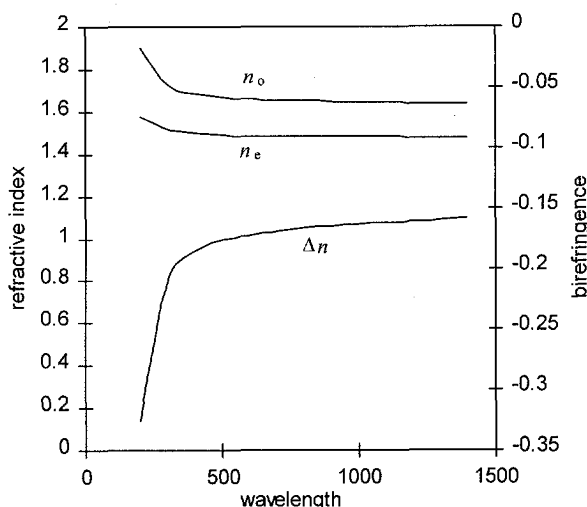


Fig. 3.18 *Refractive indices as a function of the wavelength (at $T = 18°C$).*

For 984 nm, the values for the refractive indices are 1.6440 and 1.4803 respectively for the ordinary and extraordinary components and the birefringence is -0.1637. Calcite is optically transparent in a wide wavelength range, with an absorption coefficient for both the *e*-ray and the *o*-rays of less than 0.1 m^{-1} from 260 to 1700 nm. The *o*-ray has a somewhat wider transmission range (especially in the infrared).

3.5.1 Light propagation in calcite

$CaCO_3$ has three different crystal configurations: aragonite and vaterite, which are rhombic and calcite, which has a hexagonal crystal structure[20,21]. Crystals can be regarded as more or less ordered, systematic arrangements of dipoles which conform to the requirements of internal symmetry. In isotropic materials, the refractive index is the same for light traveling in all

directions through the substance and for all polarization directions. However, calcite has a non-cubic crystal structure and therefore it is optically anisotropic. Its cleavage form (the form in which calcite tends to split along well defined smooth planes) is a rhombohedron[22] with each face a parallelogram (see Fig. 3.19). The optical axis is the axis of symmetry, passing through a blunt corner of the cleavage form such that is makes equal angles with each face.

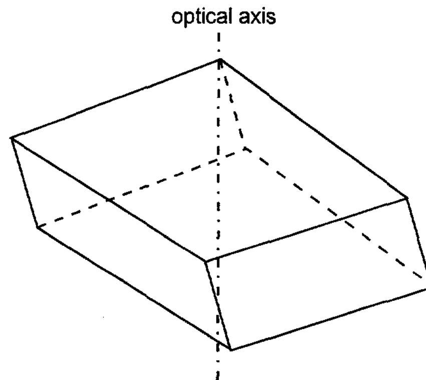


Fig. 3.19 *Crystal structure of calcite and the location of the optical axis.*

The refractive indices of an anisotropic material are represented by the indicatrix or index ellipsoid[22,23]. It is a curved, three-dimensional body and its dimensions, form and orientation relative to the crystal directions are the important properties. Calcite is uniaxial and has thus two principal refractive indices: n_o and n_e. Since n_e is smaller than n_o in the ultraviolet, visible and infrared regions, calcite is labeled as a negative uniaxial crystal. The short axis defining the index ellipsoid is proportional to the refractive index of the extraordinary ray propagating perpendicular to the optical axis and the long axis to that of the ordinary ray. A plane, containing the optical axis and the propagation vector of a wave, is called the principal plane. Randomly polarized light traveling in the principal plane is divided into two contributions that are linearly polarized at right angles (see Fig. 3.20). One of these, the ordinary ray, has its plane of polarization perpendicular to its principal plane. The second, the extraordinary ray, has its plane of polarization parallel to its principal plane.

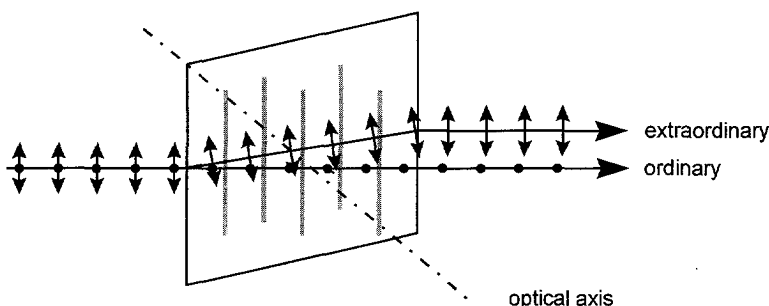


Fig. 3.20 *An incident plane wave is split in two orthogonal polarizations.*

The extraordinary ray is refracted, even for normal incidence. The refraction of the extraordinary ray generally violates Snell's law in its simple form. The only exception occurs if the optical axis is parallel to the propagation direction of the extraordinary wave inside calcite. The anomalous deflection of the *e*-ray is caused by the wavefont becoming ellipsoidal (Huygens' construction[23]), such that the direction of propagation of the light (indicated with the solid arrowed line in Fig. 3.20) is not along the wave vector $\hat{\boldsymbol{k}}$ (perpendicular to the gray lines in Fig. 3.20). The wave vector $\hat{\boldsymbol{k}}_e$ (but not the propagation direction $\hat{\rho}_e$) of the *e*-ray obeys Snell's law, provided that the index of refraction n_ϕ of this wave is given by

$$\frac{1}{n^2{}_\phi} = \frac{\sin^2\phi}{n^2{}_e} + \frac{\cos^2\phi}{n^2{}_o} \ , \tag{3.73}$$

where ϕ is the angle between the direction of the wave vector $\hat{\boldsymbol{k}}_e$ and the optical axis. Note that for ϕ equals zero, $n_\phi = n_o$ and that $n_\phi = n_e$ when ϕ reaches π/2. In Fig. 3.21 the ordinary and extraordinary rays and their field vectors for perpendicular incidence on the Savart elements are depicted. $\boldsymbol{D}$ is the electric displacement vector, $\boldsymbol{E}$ is the electric field vector and the subscripts *e* and *o* refer to the *e*-ray and *o*-ray respectively. For the ordinary ray, $\boldsymbol{D}$ and $\boldsymbol{E}$ are parallel, but for the extraordinary one this is not the case.

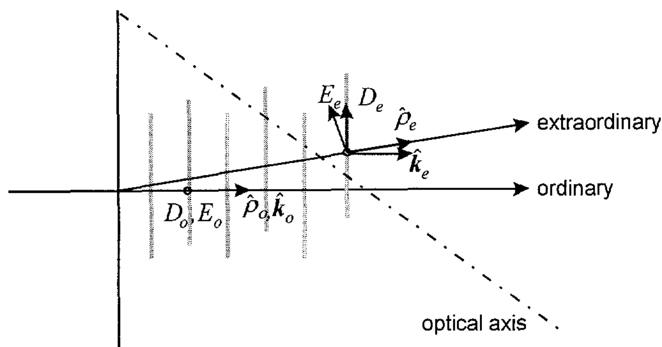


Fig. 3.21 *Field vectors of ordinary and extraordinary rays in calcite.*

Thus, for the extraordinary wave, the direction of the energy flow (indicated by the unit vector $\hat{\rho}_e$ along the Poynting vector) is not the same as that of the wave vector $\hat{\boldsymbol{k}}_e$.

3.5.2 Beam splitting by Savart elements

In Fig. 3.22 the Savart element, two of which are included in our system, is depicted. The optical axes in the Savart elements are orientated so their projections on the *uv*-plane make an angle of π/4 rad to the vertical axis. In that case the beams that are generated by the element, are separated horizontally. This is the optimum orientation of the Savart element, because this

way the camera's detection plane, which is rectangularly shaped with its long side in the horizontal direction, is used as efficiently as possible. Since we used commercially available calcite blocks with the optical axis parallel to one of its sides, the Savart element is rotated such that the projection of the optical axes on the uv-plane are positioned under an angle of $\pi/4$ rad with the u-axis.

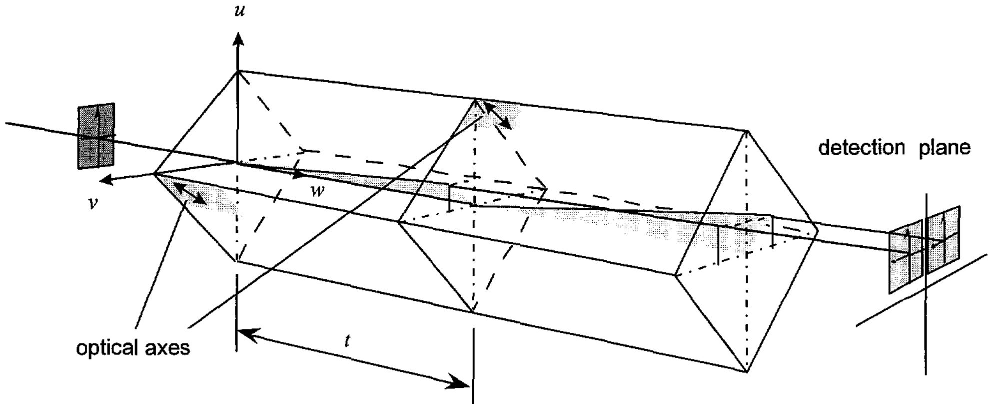


Fig. 3.22 The Savart element.

In Fig. 3.23 the position of the images in the detection plane of the camera is indicated. It shows that the images are shifted both horizontally and vertically. The input image at the entrance plane of the Savart element is depicted as an unhatched box.

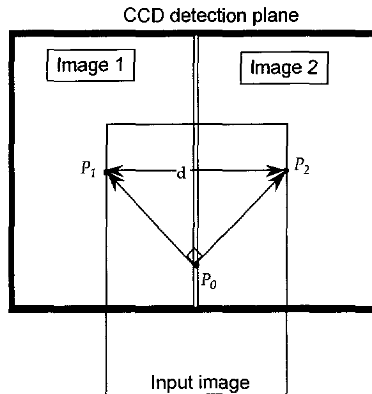


Fig. 3.23 The orientation of the images in the detection plane.

From Fig. 3.23 it is seen that the size of the detection plane determines the dimensions of the images and therefore the dimensions of the Savart element. The horizontal distance d between two corresponding points of the images is defined by the following relationship[20]

$$d = \sqrt{2}\frac{\left(\frac{1}{n_e}\right)^2 - \left(\frac{1}{n_o}\right)^2}{\left(\frac{1}{n_e}\right)^2 + \left(\frac{1}{n_o}\right)^2}t, \tag{3.74}$$

in which $2t$ is the total length of the Savart element. From Eq. (3.74) it is clear that it is advantageous to employ a birefringent material which exhibit a strong birefringence.

Using calcite for the Savart elements, the required length of the elements can be calculated using Eq. (3.74), using the refractive indices at a wavelength of 984 nm and given the size of the detection plane of our CCD cameras. The latter is 8.8×6.6 mm^2, which leads to a required d of 4.4 mm and the required length of one calcite element of the Savart element becomes 29.76 mm. The dimensions of the entrance face of the Savart element are 15×15 mm, as is required by the size of the beams at the entrance plane of the Savart elements.
As can be seen from Eq. (3.74), the image displacement depends linearly on the lengths of the two segments constituting each Savart element. Therefore, the two elements have to be exactly equal. To prepare two identical Savart elements with identical segments, four pieces of calcite have been processed simultaneously.

In conclusion, the Savart elements achieve a splitting of an incoming set of an object and a reference beam into two laterally shifted sets of combined contributions of object and reference beams. The phase difference between the object and reference beams in one set differs by π radians from that of the set in the other branch, as was shown in Section 3.1

3.6 Mechanical design criteria of the prototype

The real-time phase-stepped shearing speckle interferometer has to operate in an industrial environment. Therefore, special care has to be taken with regard to the stability of the system. Since it is a multichannel system, the mechanical stability of the Savart elements and the CCD cameras is most important. From a design viewpoint it would be preferable to fix as many degrees of freedom of the optical elements as possible. However, rotations of the Savart elements and translations of the CCD cameras are needed in the alignment procedure (see Section 5.3). Therefore, the Savart elements are mounted in specially designed housings, which facilitate the rotations of the elements (two rotations for each Savart element). The CCD chip has been removed from its original camera housing. Thereupon, these chips have been mounted on specially designed translation stages and the housing containing the electronics of the camera have been positioned directly under the base plate of the interferometer. After aligning the setup, the rotational settings of the Savart elements and the translational settings of the CCD chips have been fixed and the adjustment screws taken off, to improve the compactness of the system.

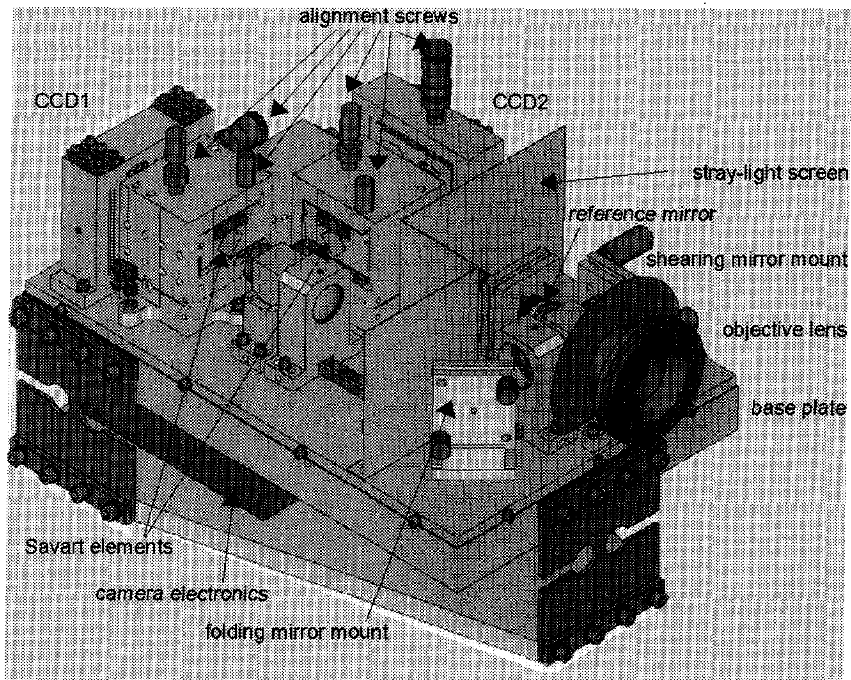


Fig. 3.24 *Birds-eye view from the set-up.*

Besides mechanical disturbances, also thermal changes can deteriorate the performance of the system, because temperature changes will alter the dimensions of components used in the set-up. In optical set-ups, aluminum is the most commonly used construction material. As compared to steel, it has a relatively high coefficient of thermal linear expansion (24×10^{-6} $^{\circ}C^{-1}$), but this is compensated by a low relative distortion figure of the set-up. As a consequence, the deformation due to temperature changes is smaller for aluminum than for steel. A material with a particularly low thermal expansion figure is Invar, but it has other less desirable properties, such as machinability and cost. In the mechanical design of the two-camera system aluminum has been chosen.

With industrial applications of the interferometer in mind, the portability of the interferometer head was a key factor in the design of the system. The inspection head (consisting of the interferometer and its laser) should be compact and should be easy to position in front of the inspection area. Therefore, the interferometer was mounted on a base plate. The laser and the interferometer have been attached to a common base plate; the interferometer is covered by a metal box, the size of which is $38\times24\times25$cm^3 (l×w×h). The weight of the optical head is ca. 10 kg. The entire system has been fastened to a mobile tripod, which is adjustable in height and can be tilted in multiple directions. The final result is an easy to position, rigid, compact and relatively portable shearographic instrument. Fig. 3.24 shows the real-time system and its diode laser mounted on the tripod.

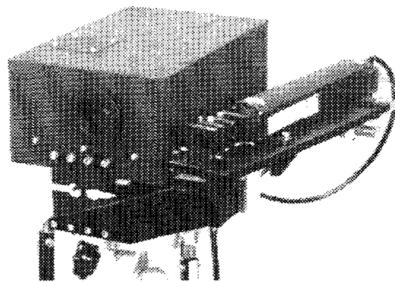


Fig. 3.24 *The mounting of the real time system and diode laser.*

3.7 References

1. G. E. Sommargren, "Up/down frequency shifter for optical heterodyne interferometry", JOSA **65**, pp. 960-961 (1975).
2. R. C. Jones, JOSA **31**, pp. 488-493 (1941).
3. W. Swindell, "Polarized light" in *The Jones calculus*, (Dowden, Halsted Press, 1975), pp. 186-241.
4. M. Born, E. Wolf, *Principles of Optics*, (Pergamon Press, Oxford, 1980), p. 17 et seq.
5. J.E. Greivenkamp, J.H. Bruning, "Phase shifting interferometry" in *Optical Shop Testing*, D. Malacara, ed., (John Wiley & Sons, Inc., New York, 1992), pp. 501-598.
6. A.J.P. van Haasteren, *Real-time phase-stepped speckle interferometry*, (Ph.D. Thesis, Delft University of Technology, Delft, 1994).
7. R. Spooren, "Standard charge-coupled device cameras for video speckle interferometry", Optical Engineering **33**(3), pp. 889-896 (1994).
8. A.R. Ganesan, D.K. Sharma, M.P. Kothiyal, "Universal digital speckle shearing interferometer", Applied Optics **27**, pp. 4731-4734, (1988).
9. J.L. Pezzaniti, R. A. Chipman, "Angular dependence of polarizing beam-splitter cubes", Applied Optics **33**(10), pp. 1916-1929, (1994).
10. Polarization optics catalog and handbook, Meadowlark Optics, 1992, p. 3.
11. E. Hecht, A. Zajac, *Optics*, (Addison-Wesley, Massachusetts, 1974) p. 152.
12. M. Born, E. Wolf, *Principles of Optics*, (Pergamon Press, Oxford, 1980), pp. 188-190.
13. E. Hecht, A. Zajac, *Optics*, (Addison-Wesley, Massachusetts, 1974), p. 108.
14. J. W. Goodman, "Statistical properties of laser speckle patterns" in *Laser Speckle and Related Phenomena*, J. C. Dainty, ed., (Springer Verlag, Berlin, 1975).
15. M. Françon, *Polarization interferometers*, (John Wiley & Sons, London, 1971), pp. 19-24.
16. *Handbook of Chemistry & Physics*, (CRC Press, Boca Raton, 1988).

17. *Landolt Börnstein* III7c3, p. 97.
18. H.J. Nickl and H. K. Henisch, J. Electrochem. Soc. **116**, pp. 1258-1260, (1969).
19. *Handbook of optics*, part 2, second edition , M. Bass, ed., (MacGraw-Hill Inc., New York, 1995), p. 33.62.
20. W.L. Wolfe, S.S. Ballard, K.A. Mc.Carthy, *American Institute of Physics Handbook*, (Mc Graw-Hill, New York).
21. J.M. Bennet, "Polarizers" in *Handbook of Optics, part 2*, M. Bass, ed., (McGraw-Hill Inc., New York, 1995).
22. E.E. Wahlstrom, *Optical Crystallography*, (John Wiley and Sons Inc. New York, 1969).
23. Cooke, Throughton, Simms, *Introduction to crystal optic*, (Cooke, Throughton & Simms, Ltd., York, ca. 1946).
24. M. Born, E. Wolf, *Principles of Optics*, (Pergamon Press, Oxford, 1980), p. 685.

Chapter 4

Image processing

This chapter presents the image processing system, as it is used in our configuration. First, some consequences of using CCD cameras for recording the interferograms are discussed (Section 4.1). Secondly, the separate steps of the phase calculation scheme are described (Section 4.2) after which (in Section 4.3) some important properties of image post processing, such as phase unwrapping, are briefly mentioned.

4.1 Introduction

The optical system simultaneously generates four phase-stepped interference patterns. These patterns are recorded on two CCD cameras; two (side by side) on each camera. It is the primary task of the image processing system to calculate the phase distribution from these recorded interferograms.

To enable computer processing of the recorded images, they are digitized into an array of numbers. First, the picture elements of the CCD cameras sample the incident intensity pattern on an $m \times n$ rectangular grid (see Fig. 4.1), leading to a finite number of sample points. Secondly, each camera is connected to its individual analog to digital converter (ADC), which digitizes the output of the CCD cameras, leading to a finite number (usually 256) of gray values. The result is a 8 bit stream that represents the intensity at the pertinent picture element.

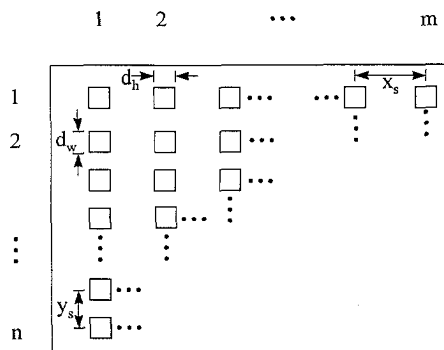


Fig. 4.1 *CCD sensor geometry: The picture elements with sizes $a \times b$ are spaced over x_s, y_s in the horizontal and vertical direction, respectively.*

Using CCD cameras and ADC's involves an integration of the intensity pattern in space and time: it is integrated over the picture element and during the recording time of the CCD camera. Thus, the recorded and digitized intensity I is a discrete representation of the interference pattern $I(x,y,t)$ and can be written as

$$I(m,n) = C_{R,D} \int_{t}^{t+T} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I(x,y,t) \mathrm{rect}\left(\frac{x - mx_s}{a}, \frac{y - ny_s}{b} \right) dxdydt \quad , \tag{4.1}$$

in which $C_{R,D}$ is a constant factor depending on the camera and ADC characteristics, T is the recording time of the camera and the rect function limits the integration over the rectangular picture elements.

The electronic equipment (see Fig. 4.2) comprises a pipeline image processing system (Datacube, MaxVideo200, indicated as MV200) and a workstation (Sun, Sparc Classic). The pipeline system is applied for digitizing the recorded images and calculating the phase from them at video speed (i.e. 25 times per second). The video signals from the two CCD cameras are digitized with two separate A/D converters, the so-called Analog Scanners (AS), provided by two MaxVideo200 modules. The phase (difference) calculation is performed using the Arithmetic Units (AU) and the Advanced Pipeline Processor (AP). The resulting phase or phase difference pattern is displayed on an RGB video monitor using the Analog Generator (AG). The workstation is employed to create and control the data paths of the pipeline system and to fine-tune the responses of the CCD cameras and their pertinent frame grabbers. It can also be used for post processing of the phase change images (using Scil-Image). The pipeline image processing devices, internally connected by a Maxbus (40MHz), are based on a VME-bus and are coupled to the workstation via a VME-bus to S-bus coupler.

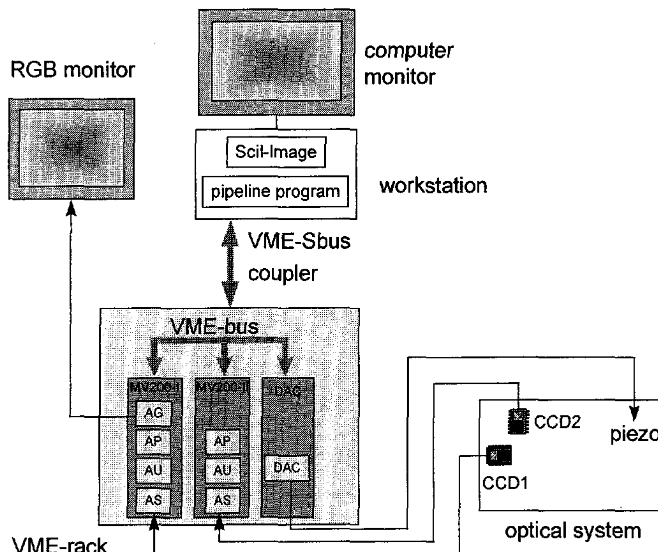


Fig. 4.2 The image processing configuration.

4.2 Real-time image processing

The entrance data of the real-time image processing system consist of the video signal of the two CCD cameras together with the synchronization signals and the pixel clock signal. The pixel clock signal is required for pixel-synchronous read out of the cameras. Further, the common electronic unit of the two cameras (see Fig. 3.8) acts as a 'master' that generates the synchronization signals upon which the CCD cameras, the 'slaves', are locked. This way, the noise due to video jitter is strongly reduced.

After the images are grabbed from the two CCD cameras, the phase distribution is calculated using

$$\varphi(m,n) = \arctan\left(\frac{I_4(m,n) - I_2(m,n)}{I_1(m,n) - I_3(m,n)}\right). \tag{4.2}$$

This equation is valid under the assumption of equal modulation intensities and background intensities of the four recorded and digitized interferograms (see also Section 2.6). The modulation and background intensities of the ith interferogram (i=1..4) are indicated with $I_{m,i}$ and $I_{b,i}$ respectively. Since in general the characteristics of the cameras and frame grabbers are different, they have to be matched to minimize the errors introduced by such differences. The gains of the cameras can be adjusted manually. These are set to their maximum values to increase the signal level. The fine tuning of the gain is performed in software, where the gain of the A/D converter can be changed until

$$\langle I_{m,1}(m,n)\rangle_{m,n} = \langle I_{m,3}(m,n)\rangle_{m,n} = \langle I_{m,2}(m,n)\rangle_{m,n} = \langle I_{m,4}(m,n)\rangle_{m,n} \tag{4.3}$$

is reached. In Eq. (4.3), the angled bracket indicates the average over all pixels *m,n* in the interferogram. In the remainder of this thesis, these subscripts will be omitted. The modulation intensity of each interferogram can be measured by recording four phase-stepped images at the pertinent detector areas. Phase stepping is performed by computer controlled (via the D/A converter DAC in Fig. 4.2) λ/8 displacement of the mirror mounted on the piezo element in the shearing head. The average modulation intensity of each interferogram can be calculated using

$$\langle I_{m,i}(m,n)\rangle = \frac{1}{2}\left\langle\sqrt{\left(I_i^0(m,n) - I_i^\pi(m,n)\right)^2 + \left(I_i^{\pi/2}(m,n) - I_i^{3\pi/2}(m,n)\right)^2}\right\rangle. \tag{4.4}$$

In Eq. (4.4), the superscripts at the recorded intensities represent the phase step. In our image processing system it is impossible to adjust the gains separately of the interference patterns (half frames) recorded by the same CCD camera. Therefore, the gain is adjusted such that

$$\langle I_{m,1}(m,n)\rangle + \langle I_{m,3}(m,n)\rangle = \langle I_{m,2}(m,n)\rangle + \langle I_{m,4}(m,n)\rangle. \tag{4.5}$$

The background intensities of the cameras can be adequately matched using the offset adjustments of the image processing software such that

$$\left\langle I_{b,1}(m,n)\right\rangle + \left\langle I_{b,3}(m,n)\right\rangle = \left\langle I_{b,2}(m,n)\right\rangle + \left\langle I_{b,4}(m,n)\right\rangle . \tag{4.6}$$

The average background intensity of each interferogram $I_{b,i}(m,n)$ can be measured using

$$\left\langle I_{b,i}(m,n)\right\rangle = \frac{1}{4}\left\langle I_i^0(m,n) + I_i^{\pi}(m,n) + I_i^{\pi/2}(m,n) + I_i^{3\pi/2}(m,n)\right\rangle . \tag{4.7}$$

As a result, we correct for the major differences in camera and frame grabber response as well as for the imperfect characteristics of the non-polarizing beam splitter. However, spatially inhomogeneous characteristics of the optical elements or of the camera response as well as imperfect characteristics of the Savart elements are not corrected.
After the correct gains and offsets are fixed, the phase distribution can be calculated using Eq. (4.2). The separate steps in the calculation scheme are schematically depicted in Fig. 4.3 and subsequently described in the following subsections.

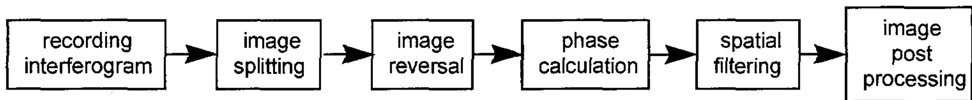


Fig. 4.3 *Steps in the phase calculation scheme.*

4.2.1 Image splitting

In each CCD detection plane, two phase-stepped interference patterns are recorded side by side, as is indicated in Fig. 4.4. Since the used CCD cameras are frame transfer cameras, the lower half of the CCD chip is masked and functions as image storage area (see also Section 3.2).The two interferograms recorded on the same CCD camera have to be separated, in order to be able to treat them as separate interferograms in the phase calculation process.

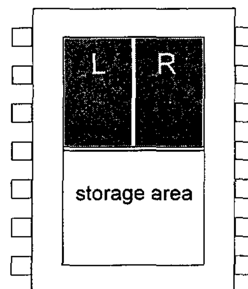


Fig. 4.4 *The location of the interferograms in the detection plane of CCD2.*

To separate the interferograms, the data stream of each CCD camera is duplicated to two memory modules (the so-called Virtual Surface Memory Modules or VSIMMs). Then the left- or right-hand parts of the images are selected as Region Of Interest (ROI) and only these parts of the images are exported. These four parts, two from each CCD camera, are treated as separate data streams in the remainder of the image processing system. These operations are depicted in Fig. 4.5.

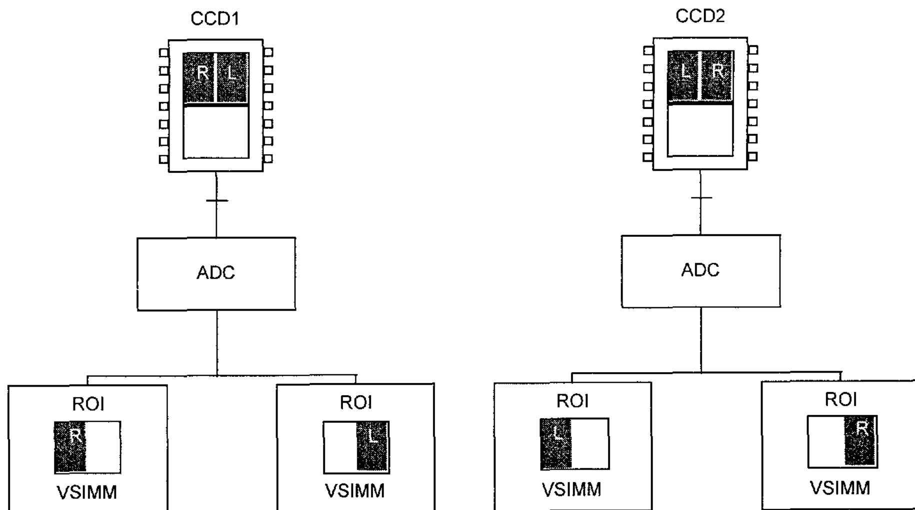


Fig. 4.5 *Image splitting.*

4.2.2 Image reversal

The interferograms recorded with CCD2 are mirrored with respect to the ones recorded with CCD1, since the number of reflections in the optical set-up is even for CCD2 and odd for CCD1. For proper processing of the images, the two interferograms recorded by one of the cameras should undergo an additional mirroring process. This could be achieved optically by an additional mirror in Branch 2 of the interferometer (see Fig. 3.1). However, the alignment of such a mirror is complicated and, by introducing another optical element, the size and complexity of the set-up are increased. Another possibility is an electronic reversal of one of the images. The advantage of this method is that no additional optical elements are required to reverse the images, but at the cost of more complicated hard- and software. Using the reversed read-out option of the VSIMMs, the reversal process can be performed at the full 40 MHz pipe rate, thus enabling real-time image reversal. Fig. 4.6 shows the reversed read out of the two memory modules associated with CCD2.

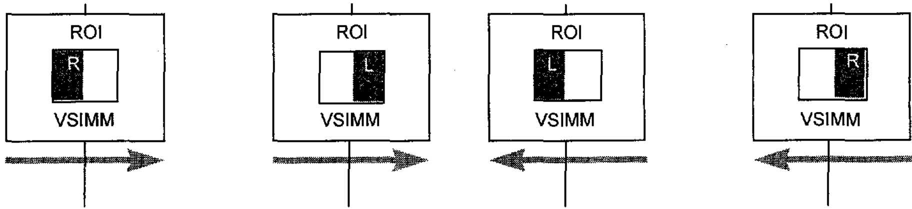


Fig. 4.6 Reversed read-out of CCD2

4.2.3 Phase computation

The phase pattern is calculated using Eq. (4.2). The relevant parts of the images are exported to two separate Arithmetic Units (AU). With the aid of these devices, the numerator and denominator of the phase equation are computed. The calculated values are the input of a look up table (LUT) on a MaxVideo200 module which outputs the phase $\varphi(x,y)$ modulo 2π with 8 bits accuracy. In Fig. 4.7 the arctangent LUT is displayed.

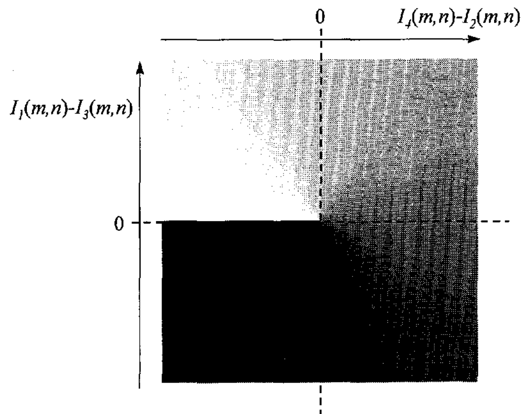


Fig. 4.7 Arctangent Look Up Table.

This 256×256 table (LUT1), that renders the phase value modulo 2π, is used instead of a actual arctangent computation, to speed up the calculation. Since the numerator and denominator can have values between [-255, 255], they are divided by two before entering the 256×256 LUT. Fig. 4.8 shows the image processing flow chart.

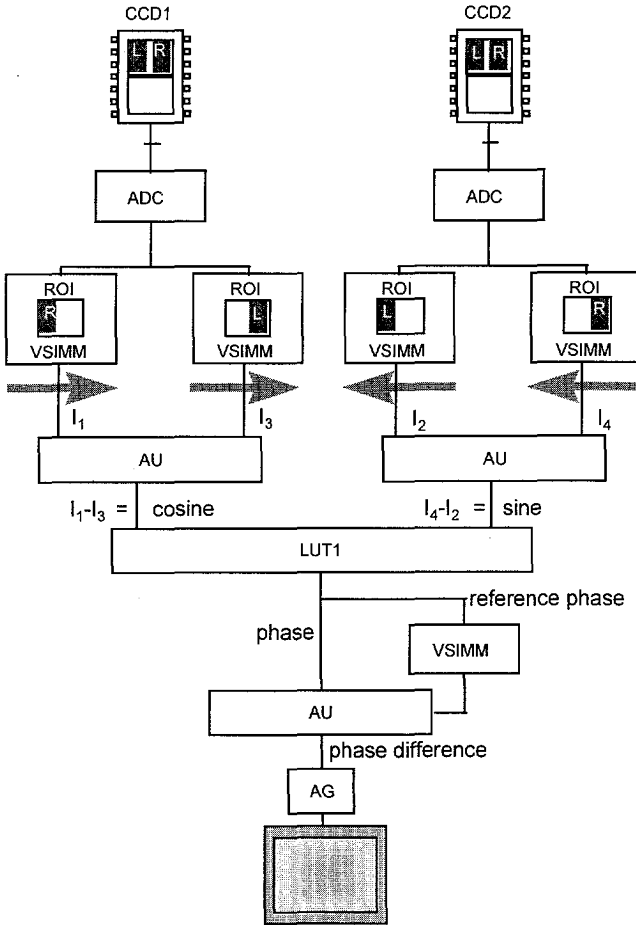


Fig. 4.8 *Real-time phase calculation.*

Performing a measurement, first the phase calculated for an unloaded object is stored in a VSIMM, residing on the MV-200 board. Then, all consecutive phase images are subtracted from the stored reference image by a third Arithmetic Unit, also residing on this MaxVideo200 module. The resulting image represents the phase change caused by loading and is converted to an analog video signal by an Analog Generator (AG). This signal is displayed on an RGB monitor.

4.2.4 Spatial filtering

Even though the speckle pattern is the information carrier, in the phase (difference) pattern the speckles introduce noise. This is most noted in regions where the calculated phase is around $-\pi$ or π rad, where already small noise contributions can cause the phase to jump from $-\pi$ to π rad

and vice versa. The noise can be caused by saturated picture elements, picture elements with low modulation intensities and can also originate from decorrelation of the interferograms before and after loading the object. To reduce the noise, a special filtering technique is applied. The sine and cosine of the phase distribution is filtered instead of the phase itself. If the phase varies continuously across the detection plane, the sine and cosine of the phase are continuous, too. In this case, the noise in the sine and cosine patterns can be reduced by filtering with a *moving mask operation*[1]. The central pixel value is replaced by a value calculated from the (weighted) average of the surrounding pixels in the filter mask. Low resolution features are unaffected by the filtering process, but higher resolution features (such as speckle noise) are removed or reduced. Thus, uniform filtering results in low pass filtering the image. Usually, the filter mask is rectangular and it can have any size $s_n \times s_n$, although the size is usually an odd number of pixels, such that the central pixel is unambiguously determined. In general, the larger the filter dimensions, the smoother the result. For example, uniform filtering reduces the noise proportionally to the square root of the number of pixels within the filter mask[2]. Directional filtering is being performed when $s_n \neq s_n$. This can be advantageous when straight fringes are present. In our applications, this is generally not the case and therefore we use square filter masks.

The sine/cosine filter is implemented in the real-time pipeline image processing system, which facilitates an 8×8 filter mask as a maximum, but odd filter sizes (e.g. 7×7) are usually chosen. First, the sine and cosine of the measured phase (or phase difference) are determined using a second look-up table (LUT2). Secondly, the sine and cosine are filtered. Thirdly, the phase pattern is reconstructed using the filtered sine and cosine values as inputs for the arctangent look-up table (LUT1). A schedule of this filtering technique is depicted in Fig. 4.9.

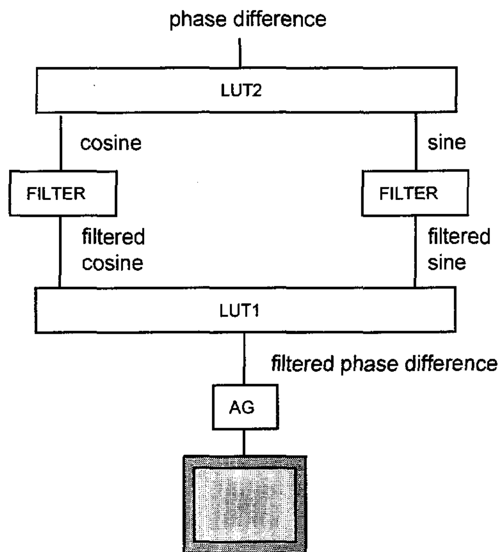


Fig. 4.9 *Real-time noise reduction filtering process.*

The ratio of filter size to speckle size and spatial period of the fringe pattern influences the resulting improvement. When the speckles are small as compared to the filter size, the noise in the phase pattern will be strongly reduced. Also, the smallest spatial period of the fringes must be larger than the filter size in order not to be filtered out.

Since this filtering technique reduces the spatial resolution of the images, a compromise has to be found between filter size and the required accuracy of the phase calculation for each measurement. Therefore, the type and the size of the filter can be altered in our system. Uniform and gaussian filtering in a 3×3, 5×5 or 7×7 filter mask has been implemented. In uniform filtering, all neighboring pixels are equally important in calculating the central pixel value, i.e. they have the same filter weight. However, using a gaussian filtering technique, the picture elements closest to the central one have the largest influence on the filtered value. The weight function $W(\Delta m, \Delta n)$ of the gaussian filtering for a picture element at a distance $(\Delta m, \Delta n)$ of the central pixel in the square filter mask is given by

$$W(\Delta m, \Delta n) = C_n e^{-\frac{1}{2}\left(C_\delta \sqrt{\Delta m^2 + \Delta n^2}\right)^2}, \tag{4.8}$$

in which C_n is the normalization constant and C_δ a constant determining the decay speed of the gaussian function. The latter sets the gaussian coefficient $W(\Delta m, \Delta n)$ equal to δC_n at picture elements with coordinates $(-½(s-1),0)$, $(½(s-1),0)$, $(0,-½(s-1))$ and $(0, ½(s-1))$ in a $s \times s$ filter mask. Thus, using this filter, C_δ is given by

$$C_\delta = \frac{\sqrt{-2\ln(\delta)}}{(s-1)/2} \quad (\delta \leq 1). \tag{4.9}$$

Combining Eq. (4.8) and (4.9), the gaussian coefficient can be written as

$$W(\Delta m, \Delta n) = C_n e^{4\ln(\delta)(\Delta m^2 + \Delta n^2)/(s-1)^2}. \tag{4.10}$$

As an example, the 7×7 gaussian coefficients (with δ=0.1) are given in Table 4.1.

Table 4.1 *Gaussian coefficients in a 7×7 neighborhood.*

$C_n/100$	$C_n/27.8$	$C_n/12.9$	$C_n/10$	$C_n/12.9$	$C_n/27.8$	$C_n/100$
$C_n/27.8$	$C_n/7.74$	$C_n/3.59$	$C_n/2.78$	$C_n/3.59$	$C_n/7.74$	$C_n/27.8$
$C_n/12.9$	$C_n/3.59$	$C_n/1.67$	$C_n/1.29$	$C_n/1.67$	$C_n/3.59$	$C_n/12.9$
$C_n/10$	$C_n/2.78$	$C_n/1.29$	C_n	$C_n/1.29$	$C_n/2.78$	$C_n/10$
$C_n/12.9$	$C_n/3.59$	$C_n/1.67$	$C_n/1.29$	$C_n/1.67$	$C_n/3.59$	$C_n/12.9$
$C_n/27.8$	$C_n/7.74$	$C_n/3.59$	$C_n/2.78$	$C_n/3.59$	$C_n/7.74$	$C_n/27.8$
$C_n/100$	$C_n/27.8$	$C_n/12.9$	$C_n/10$	$C_n/12.9$	$C_n/27.8$	$C_n/100$

To avoid a gain factor when applying the filter that would increase all intensities, the central pixel values are normalized using C_n. The value for C_n can be found using

$$\sum_{\Delta m=-\frac{s-1}{2}}^{\frac{s-1}{2}} \sum_{\Delta n=-\frac{s-1}{2}}^{\frac{s-1}{2}} W(\Delta m, \Delta n) = 1. \tag{4.11}$$

In the example mentioned above, where a 7×7 gaussian filter is used while δ=0.1, the value for C_n is 1/12.03. We have implemented this Gaussian filter and the uniform filter. During measurements the filter type or filter size can be altered instantly. In Fig. 4.10 an overview of the complete real-time pipeline image processing scheme is depicted.

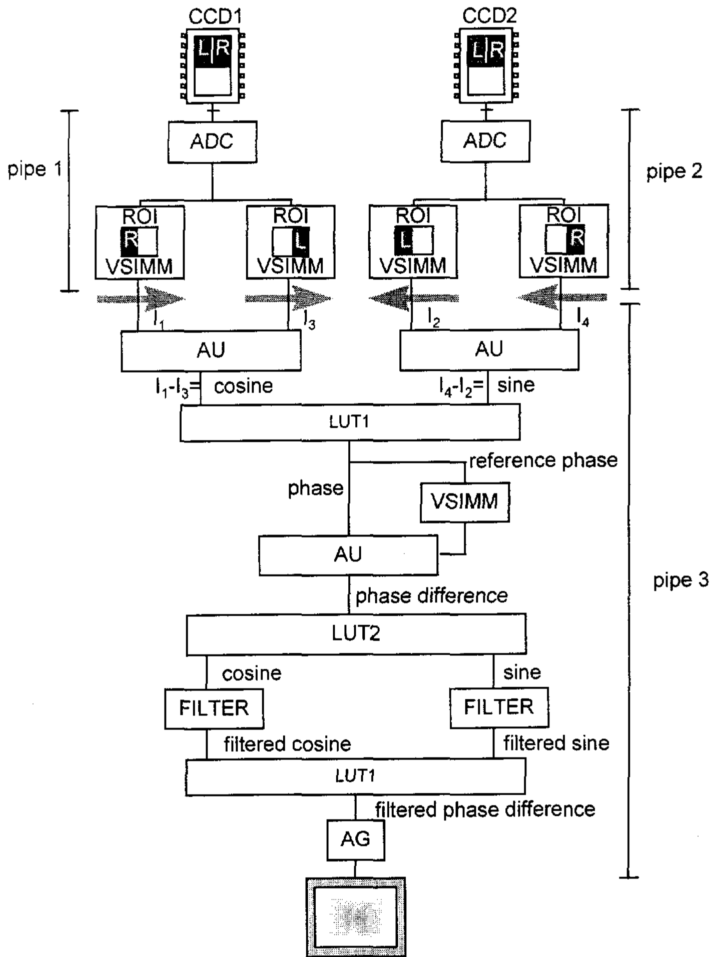


Fig. 4.10 *Data pipes in the pipeline image processing system.*

4.3 Image post processing

The final image displayed on the video monitor represents the phase change between the moment that the reference phase distribution was stored and the actual phase distribution. Because the phase is determined modulo 2π, artificial 2π phase steps may occur in the phase change image. The removal of these steps, called phase unwrapping[3], is necessary to make the phase data continuous. Also, the noise in the unwrapped image can be further reduced using filtering techniques (e.g. uniform, median or percentile filtering). There are many filtering algorithms available. It depends on the measured phase difference which filter routine or filter size performs best. In general, good results can be obtained using a sine/cosine filtering before unwrapping, after which other filtering routines, background subtraction, thresholding, contrast stretching, etc. are performed. In our system, this image post processing is performed off-line on the workstation that is coupled to the real-time image processing system. For this purpose, we use a commercial image processing package (Scil-Image) together with in-house written software in C.

In our system, the two-dimensional phase pattern is presented as gray value (or false color) images. The scaling factor between the gray values and the actual data is defined such that the total gray value range of 8 bits (256 values) represents the total range of 2π phase difference. Using post processing, the scaling factor can be changed to improve the visibility of the defects, but then care should be taken when interpreting the gray value images quantitatively.

4.3.1 Phase unwrapping

The simplest phase unwrapping method consists of checking the difference between the wrapped phase values at adjacent pixels of the discrete lattice. If the absolute value of this difference exceeds π, an integer multiple of 2π is added or subtracted to make it less than π. This process is carried out for each row and for each column of the phase map $\varphi(m,n)$ to yield the unwrapped phase map $\varphi_u(m,n)$

$$\varphi_u(m,n) = \varphi(m,n) + 2\pi g(m,n), \tag{4.12}$$

in which $g(m,n)$ is an integer valued field.

Thus, for the reliable removal of discontinuities, the phase must not change by more than π between adjacent pixels. Moreover, phase unwrapping algorithms based on a 2-D spatial comparison of phase values at neighboring pixels can not be used when the phase is not continuous, e.g. with an impact damage which breaks the surface of the object. A novel algorithm has been proposed in literature[4,5] to perform unwrapping along the time axis instead of in space coordinates. The main advantage of this temporal phase unwrapping technique is the possibility to unwrap phase maps from objects with physical discontinuities. This technique

can be used for applications where phase changes are of interest, rather than absolute phase values.

4.4 References

1. H.A. Vrooman, *Quantitative analysis of interferograms*, (Delft University Press, Delft, 1991).
2. A.A.M. Maas, *Phase shifting speckle interferometry*, (Ph.D. Thesis, Delft University of Technology, Delft, 1991).
3. D.W. Robinson, "Phase unwrapping methods", in *Interferogram analysis: Digital fringe pattern measurement techniques* (Institute of Physics, Bristol, UK, 1993), pp. 194-229.
4. J.M. Huntley, H. Saldner, "Temporal phase-unwrapping algorithm for automated interferogram analysis", Applied Optics **32**(17), pp.3047-3052 (1993).
5. H.O. Saldner, J.M. Huntley, "Temporal phase unwrapping: application to surface profiling of discontinuous objects", Applied Optics **36**(13), pp. 2770-2775 (1997).

Chapter 5

Accuracy analysis

This chapter considers the accuracy of the system as well as its error sources. First, a classification of errors is made (Section 5.1) and the major error contributions are analyzed. Speckle decorrelation is one source of those major inaccuracies and, therefore, accurate alignment procedures are required. In Section 5.2 a polarization ray tracing algorithm is developed, which is used to design such an alignment procedure that is subsequently described in Section 5.3. The calculation scheme for ray propagation through calcite from Section 5.2 is also used for calculating the effects of image formation through Savart elements. Section 5.4 describes the performance deterioration caused by aberrations in the multichannel interferometer. Finally, in Section 5.5, measurements of the resulting accuracy of the system are presented.

5.1 Error classification and analysis

Interferograms usually suffer from a number of distortions degrading the interference pattern. In general, practical factors restrict the accuracy of the interferometer instead of fundamental limits[1]. Error sources can be related to the optical system (e.g. phase step errors, speckle noise, decorrelation, aberrations in the optics of the interferometer), or to the recording and computational system (digitization and quantization, integrated speckle patterns, electronic noise, detector nonlinearities). Environmental disturbances also degrade the performance of the system (e.g. mechanical vibrations, temperature changes, air flow). These errors are extensively described in literature[2,3,4,5]. Therefore only an overview of the dominant errors, limiting the accuracy of the interferometric measurement, is given in this section.

As presented in Section 2.6, the noise contributions to an interferogram can be divided into two categories: additive noise contributions, collected in $a(x,y)$ and a multiplicative noise contribution described by $b(x,y)$, resulting in an interferogram described by

$$I(x,y) = a(x,y) + b(x,y)\cos\big(\varphi(x,y)\big), \tag{5.1}$$

where $a(x,y)$ and $b(x,y)$ are the perturbed background and modulation intensities, respectively and $\varphi(x,y)$ is the phase distribution. Some errors (the errors that are equal for each phase stepped interferogram and for each picture element of the camera recording the interferogram) can be eliminated by the phase calculation scheme, especially with simultaneously phase stepped systems, such as our multichannel interferometer. Moreover, in a

multichannel interferometer several errors are avoided[6] (e.g. the phase step error caused by nonlinearities of the phase shifter (with PZT driven mirrors) and mechanical or thermal drift or vibrations). Most important, it has a reduced sensitivity to fluctuations of environmental parameters, which often limit phase stepped interferometers with sequential acquisition of the interferograms. However, some additional error sources appear in a multichannel interferometer. They are mainly due to the fact that the data are recorded by different cameras. Therefore, the overall accuracy of simultaneously phase stepped interferometers is in general not better than that of a sequential phase stepped system under laboratory environments. Possible additional errors are

- pixel mismatch,
- aberration of the imaging optics,
- different average intensity of an interferogram,
- different sensitivity of the relevant pixels.

The most prominent errors in our setup will be discussed in detail in the following subsections.

5.1.1 Errors caused by optical elements

Illumination optics

The object is illuminated by an expanded laser beam. For this reason the background intensity of the pattern varies according to the beam profile, in a good approximation a 2-dimensional Gaussian curve. Moreover, the scattered intensity distribution may vary considerably over the object surface due to, e.g., difference in diffuse scattering properties of different surface finishes. This spatial variation in intensity scattered from the object influences the intensity in the detection plane and therefore the dynamic range of the interferogram. This may reduce the accuracy of the phase calculation at points where the intensity is too low (low modulation depth) or too high (saturation). This error can locally be a problem, although it can easily be reduced using surface preparations (spraying diffusely reflecting powder on the surface) or decreasing the field of view.

Imaging optics

In interferometers, the requirements on the imaging elements are not very strict, unless extremely accurate results are needed. In general it is sufficient when the imaging elements do not smear out or distort the fringe pattern beyond the camera resolution. Besides, the imaging system influences the speckle decorrelation[6]. Telecentric and afocal imaging systems appear to have a smaller decorrelation effect than a single lens. In many interferometers, converging or diverging beams exist in the interferometer. In speckle interferometry, the numerical aperture of the imaging system is usually small and the induced aberrations can be ignored. However, in polarizing interferometers, such as our system, nonaxial beams pose a problem, since the performance of the polarizing optics generally depends on the angle of incidence. Moreover, in our system converging beams occur within thick pieces of birefringent material. The errors associated with converging beams inside birefringent material will be described in detail in Section 5.4.

Polarization optics

In polarization interferometry, the main potential error sources are imperfections in the polarization elements and their positions. Imperfections of the quarter wave plates and lack of a perfect extinction ratio in the polarizing beam splitter[7,8] produces optical cross talk and causes errors in the phase measurement[9]. An incorrect azimuth angle of the quarter wave plates introduces further phase errors. The analyzer in a system determines the phase step between object and reference beams. Incorrect angular position of its transmission axis causes a phase step error. A phase step error is a systematic error and is one of the most serious limitations[10] of phase stepping systems. Errors due to phase step inaccuracies can usually be eliminated by careful calibration of the system and choice of the phase calculation algorithm[11-13]. Generally, the choice of the algorithms can reduce one error, often increasing the effects of others[14].

An extensive description on imperfect properties or incorrect positioning of quarter wave plates, analyzers and beam combiners can be found in literature[15]. Here, we will describe the errors associated with incorrect positioning or non-ideal characteristics of the polarizing elements in our system, i.e. retardation plates, polarizing beam splitter and analyzer, which lead to errors in the calculation of the phase distribution. The goal of this analysis is to determine which element or setting is most limiting for the accuracy of the interferometer. Therefore, the influence of one inaccuracy on the effect of another, being of second order, are omitted here. We will treat the errors associated with the polarizing elements separately.

Polarizing beam splitting cube

The polarizing beam splitting cube (Technical Optics, P1153-000-000) is included in the shearing device of our system (see Fig. 5.1).

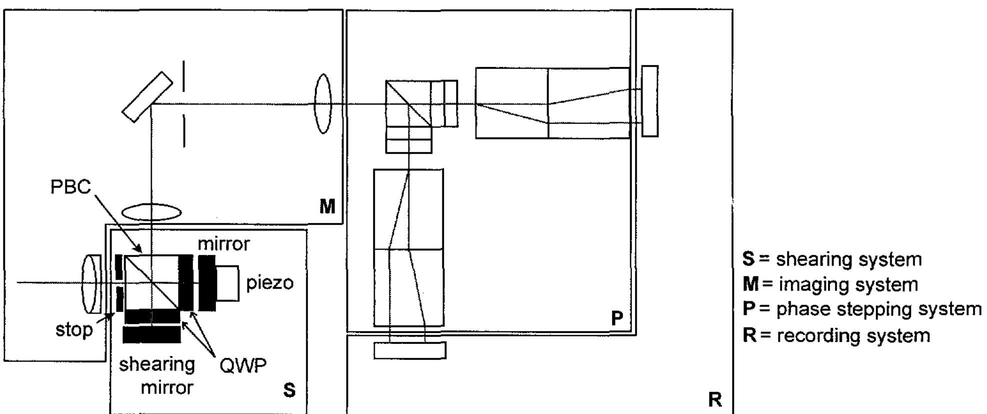


Fig. 5.1 *Position of shearing device in the interferometer.*

It transmits linearly polarized light with its electric field vector parallel to the plane of incidence (*p*-polarized light, with only the electric field component E_y) and reflects orthogonally polarized light (*s*-polarized light, with E_x). A perfect PBC would transmit 100% of the *p*-polarized light and reflect all the *s*-polarized light. However, due to the angular

dependence of the PBC, some p-polarized light is reflected and some s-polarized light is transmitted. Schematically, it can be depicted as in Fig. 5.2.

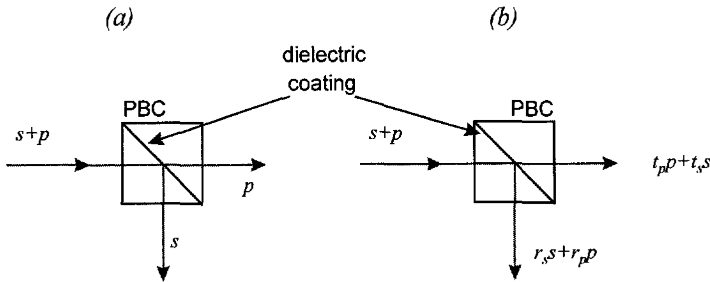


Fig. 5.2 *Splitting of polarization in (a) an ideal PBC and (b) imperfect PBC.*

The polarizing beam splitting cube (PBC) in our system shows only a small dependence on the angle of incidence for the s-polarized light. The unwanted power transmission coefficient for the s-polarized components, T_S, which is related to the amplitude transmission coefficient t_s by

$$T_s = t_s t_s^* , \tag{5.2}$$

where t_s^* is the complex conjugate of t_s, remains smaller than 0.2% for the pertinent convergence of the beams in our system. The dependence of the PBC on the angle of incidence for p-polarized light is much larger and shows asymmetry. For rays incident on the dielectric coating (see Fig. 5.1) under an angle of 0.75 rad (which would occur at a NA of 0.06), about 75% of the intensity is transmitted instead of the required value of 100%. For angles of incidence of 0.83 rad, the influence is much smaller: 98% of the intensity is transmitted. Assuming the amplitude transmission and reflection coefficients to be real, the amplitude transmission coefficients of the PBC for two relevant angles of incidence with respect to the entrance face of the PBC (indicated in the subscript) are

$$t_{p,0.83} = 0.99,\ t_{p,0.75} = 0.87,\ t_{s,0.83} = t_{s,0.75} = 0.045. \tag{5.3}$$

The power reflection coefficient can be calculated using R=1-T. After passing the quarter wave plates in the shearing head, reflecting from the mirror and passing the quarter wave plates a second time, again coupling of p- and s-polarization occurs, as is indicated in Fig. 5.3.

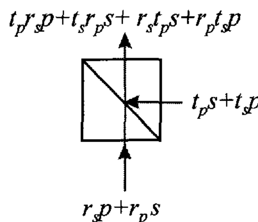


Fig. 5.3 *Coupling of polarization in an imperfect PBC.*

The PBC is designed such that the beam leaving the PBC has equal components of s- and p-polarization. The case is slightly different when shearing is taken into account. In this case the object beam is tilted, such that the angle of incidence on the PBC changes and the r and t are changed at the second passing through the PBC. The changed reflection and transmission coefficients associated with the tilted object beam are indicated with a prime in the superscript. The beam leaving the PBC can then be described by

$$t_p * r_s p + t_s * r_p s + r_s t_p s + r_p t_s p = p(t_p * r_s + r_p t_s) + s(t_s * r_p + r_s t_p) \tag{5.4}$$

For relevant shear angles (i.e. maximally 20 mrad, but generally around 5 to 10 mrad) only the t_p and r_p change notably (estimated $\Delta t_p = t_p - t_p* = 0.02$). Worst case, this results in a transmitted beam with different ratio of p- and s-contributions, given by

$$ratio = \frac{p}{s} = \frac{(t_p * r_s + r_p t_s)}{(t_s * r_p + r_s t_p)} = 0.98 \tag{5.5}$$

Thus, the ray with the worst performance in the shearing head shows a 2% coupling of s- and p-polarization. The average over all rays will show much less coupling. The coupling between s and p polarization also means a coupling between the object and reference beams. The object beam can be described by

$$\boldsymbol{E}_o(x,y) = t_p * r_s p + t_s * r_p s = t_p * r_s \begin{pmatrix} 0 \\ 1 \end{pmatrix} e^{i\varphi_o(x,y)} + t_s * r_p \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{i\varphi_o(x,y)} \tag{5.6}$$

and the reference beam as

$$\boldsymbol{E}_r(x,y) = r_s t_p s + r_p t_s p = r_s t_p \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{i\varphi_r(x,y)} + r_p t_s \begin{pmatrix} 0 \\ 1 \end{pmatrix} e^{i\varphi_r(x,y)} \tag{5.7}$$

Due to this coupling, the modulation in the interferogram is decreased and the phase measurement is corrupted. Besides imperfect characteristics, also an incorrect positioning of the element must be taken into account. The PBC was positioned within an accuracy of 1mrad, which leads to an negligible additional error, since the angular inaccuracy is much smaller than the shear angle and angle of convergence. We assume a 2% coupling of s- and p-polarization as the upper case. Using this value, the average phase error over the detection plane can be calculated. Since the system is used for phase change measurement, the constant phase error will be neglected. Using the results from literature[15], it is calculated that a 2% coupling result in an average phase error over the detection plane of approximately 15 mrad, which is the largest error due to the shearing head. The other error sources in the shearing device are due to imperfect characteristics and angular mispositioning of the quarter wave plates. These are described in the following subsection.

Quarter wave plate in shearing device

If the angle between the fast axis of the QWP and the x-axis deviates from the required value of $\pi/4$, the beams leaving the QWP's will not be circularly polarized. They can be described by using the transformation matrix of a QWP with a deviation from the $\pi/4$ position denoted by ε:

$$\boldsymbol{M}_{QWP,\varepsilon} = \frac{1}{2}\sqrt{2}\begin{bmatrix} 1+i\sin(2\varepsilon) & -i\cos(2\varepsilon) \\ -i\cos(2\varepsilon) & 1-i\sin(2\varepsilon) \end{bmatrix}, \tag{5.8}$$

The deviation angle ε is estimated from the alignment procedure to be within 0.02 radians. Assuming the incident object and reference beams to be linearly polarized, the incorrect position of the QWP leads to a reference wave, after passing this QWP twice, described by

$$\boldsymbol{E}_r = \begin{bmatrix} -i\cos(2\varepsilon) \\ i\sin(2\varepsilon) \end{bmatrix} = \begin{bmatrix} -0.9992i \\ 0.04i \end{bmatrix}, \tag{5.9}$$

and the object beam is given by

$$\boldsymbol{E}_o = \begin{bmatrix} i\sin(2\varepsilon) \\ i\cos(2\varepsilon) \end{bmatrix} = \begin{bmatrix} 0.04i \\ 0.9992i \end{bmatrix}. \tag{5.10}$$

Thus, both beams $\boldsymbol{E}_r$ and $\boldsymbol{E}_o$ are rotated over roughly 0.04 rad. This changes the transmission efficiency of the PBC (assumed perfect), since the efficiency depends on the projection of the polarization directions onto the polarization axes of the PBC. After passing the PBC, a negligible amplitude loss is noticed. If one of the QWP's is assumed to be perfect and perfectly aligned, a small decrease (<0.1%) of the modulation depth of the interferograms will occur. Since this will be the same for all four interferograms, the phase calculation will hardly be affected by this error.

The error associated with imperfect characteristics of the QWP is more important. At a wavelength of 984 *nm* the two QWP's in our shearing head have a 0.2491 and 0.2515 wave retardance for perpendicular incidence. The converging beams in the shearing head cause the retardation to deviate from this value. The acceptance angle of the quarter wave plates are ±0.17 rad (retardance error will be within 0.02 rad). As a worst case, combining the effect of the beam convergence with the imperfect properties of the QWP and the maximum shearing angle, the retardances can have values of 0.2461 and 0.2545, respectively. These values change the transformation matrix of a QWP which, under $\pi/4$ rad with the incident polarization directions, is given by

$$\boldsymbol{M}_{\mathrm{QWP}} = \frac{1}{\sqrt{2}}\begin{pmatrix} \cos(\frac{\beta}{2}) & -i\sin(\frac{\beta}{2}) \\ -i\sin(\frac{\beta}{2}) & \cos(\frac{\beta}{2}) \end{pmatrix} \tag{5.11}$$

where β is the retardance. For the worst case values of the retardance, the transformation matrix becomes

$$M_{\mathrm{QWP},0.2545} = \begin{pmatrix} 0.697 & -0.717i \\ -0.717i & 0.697 \end{pmatrix} \tag{5.12}$$

and

$$M_{\mathrm{QWP},0.2491} = \begin{pmatrix} 0.709 & -0.705i \\ -0.705i & 0.709 \end{pmatrix}. \tag{5.13}$$

The beams in the shearing head pass through the quarter wave plates twice, and as a result of the deviation from ideal retardance values, the polarizations of the beams are not linear, but elliptical. As a consequence, the object and reference beams are not linearly polarized anymore, which results in an object beam entering the PBC for the second time given by the Jones vector

$$E_o = \begin{bmatrix} 0.00565 \\ -0.99998i \end{bmatrix}, \tag{5.14}$$

while the reference beam is given by

$$E_r = \begin{bmatrix} -0.983i \\ 0.0287 \end{bmatrix}. \tag{5.15}$$

This shows that the object beam is slightly right-handed elliptically polarized and the reference beam is left-handed elliptically polarized. The polarizing beam splitter, which now is assumed to have perfect properties, will only transmit the p-component of the object beam and reflect the s-components of the reference beam. After passing the PBC, this only results in a loss of amplitude of the ray in the reference beam at the largest angle of incidence of almost 2%, the loss is negligible (<0.2%) for the object beam.

Quarter wave plate in interference section

In the interference section (alternatively called phase stepping system) of the set-up, two other QWP's are positioned, as indicated in Fig. 5.4.

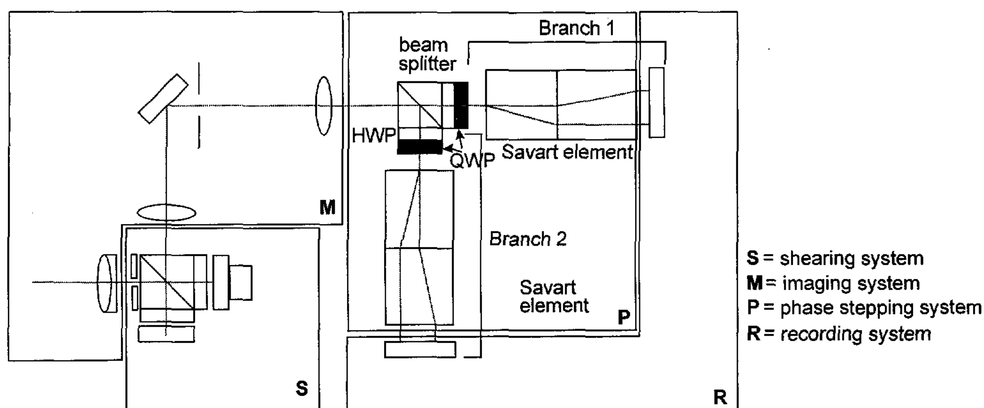


Fig. 5.4 *The interference section of the interferometer.*

The properties of our QWP's are slightly better in the interference section than in the shearing section, having a retardance of 0.2505 in Branch 1 and 0.2502 in Branch 2. As a result of the incorrect retardance value slightly elliptical light is produced instead of, ideally, circularly. As a result, after passing a perfect HWP and analyzer with an azimuth angle of $\pi/4$, the object and reference beams will have a very small intensities and phase difference deviations (several ‰) at the detection plane of the interferometer. The associated error in the phase measurement ($<10^{-6}$ rad) is much smaller than the error resulting from the possible incorrect azimuth angle of the QWP's. This latter error has been calculated, assuming again a possible inaccuracy of the azimuth angle of 0.02 rad. Using Eq.(5.8), the object beam becomes

$$\boldsymbol{E}_o = \begin{bmatrix} 1+0.04i \\ -0.9992i \end{bmatrix} \tag{5.16}$$

and the reference beam

$$\boldsymbol{E}_r = \begin{bmatrix} -0.9992i \\ 1-0.04i \end{bmatrix}. \tag{5.17}$$

Assuming a perfect analyzer under the correct angle with the x-axis, the average phase error in the detection plane due to this quarter wave plate misorientation is calculated to be nearly 0.07 rad[15].

Half wave plate in interference section

For the half wave plate in the system, the same effects are present as are for the QWP's in the interference section. Again, the misorientation of the HWP is dominant over the incorrect retardance value (i.e. 0.5021). As a result of the incorrect retardance value, both waves become slightly elliptical. This ellipticity leads to a phase error smaller than 10 mrad. The second error source associated with the HWP is the error introduced by its incorrect angular position. A half

wave plate flips the polarization state around the fast axis of the wave plate. The transformation matrix for a half wave plate with its fast axis at an angle ψ to the x-axis equals[16]

$$\boldsymbol{M}_{\text{HWP}} = -i\begin{pmatrix} \cos(2\psi) & \sin(2\psi) \\ \sin(2\psi) & -\cos(2\psi) \end{pmatrix}. \tag{5.18}$$

A possible misorientation of ε (in our case 0.02 rad) from the required value for ψ of $\pi/8$ rad leads to an object beam

$$\boldsymbol{E}_o = -i\begin{pmatrix} 0.678 \\ -0.735 \end{pmatrix}, \tag{5.19}$$

and a reference beam

$$\boldsymbol{E}_r = = -i\begin{pmatrix} -0.735 \\ -0.678 \end{pmatrix}. \tag{5.20}$$

Hence, both beams are rotated by an amount of 2ψ and thus the error in the alignment of the HWP is doubled in the error in the rotation of the object and reference beams. Consequently, the phase step between interferograms in Branch 1 and the corresponding interferograms in Branch 2 deviates from the desired value of $\pi/2$ rad, being $(\pi/2+0.04)$ rad. Since the correct position of the quarter wave plate is determined after the orientation of the HWP is fixed, the HWP's angular misorientation does not lead to an incorrect angular position of the quarter wave plate. However, the orientation of the Savart element with respect to the incoming polarization direction can have an additional 0.04 rad angular error. This error will be treated in the next paragraph.

Analyzer

The analyzers in our system are the Savart elements. The maximum mismatch of the angle between the optical axes and the sides of the calcite segments is specified by the manufacturer* as ± 15 arc minutes (i.e. 4.4 mrad). Since rotations of the Savart elements are used in the alignment procedure (see Section 5.2), the production inaccuracy can be compensated. If the Savart element is made from identical calcite segments (equal optical properties and equal length), and assuming a perfectly horizontally aligned CCD camera (picture element lines coincide with the v-axis and columns with the u-axis), the angular mismatch of the Savart elements vanishes. However, due to a possible slight mismatch between the calcite segments and a possible angular misorientation of the CCD cameras, an angular mismatch (with respect to the required value of $\psi = \pi/4$) of less than 2.5 mrad (following from the Savart elements production process and the alignment procedure) has to be assumed. Combining this value with the additional 0.04 rad angular mismatch of the HWP in Branch 2 and using the transformation

* Karl Lambrecht Corporation, Chicago, Illinois

matrix $\boldsymbol{M}_{LP}$ for a linear polarizer which transmission axis makes an angle of ψ rad with the x-axis,

$$\boldsymbol{M}_{\text{lp}} = \begin{pmatrix} \cos^2(\psi) & \sin(\psi)\cos(\psi) \\ \sin(\psi)\cos(\psi) & \sin^2(\psi) \end{pmatrix}, \tag{5.21}$$

it can be shown that the phase step in the recorded interferograms is slightly deteriorated. The interferogram recorded at the left-hand side of CCD2 is changed into

$$I_{CCD2}^{left}(u,v) = I_o(u,v) + I_r(u,v) + I'_m(u,v)\cos\left(\varphi(u,v) - \frac{\pi}{2}\right) + I_e(u,v) \tag{5.22}$$

and likewise, the interferogram recorded on the right-hand side of CCD2 can be written as

$$I_{CCD2}^{right}(u,v) = I_o(u,v) + I_r(u,v) + I'_m(u,v)\cos\left(\varphi(u,v) + \frac{\pi}{2}\right) + I_e(u,v). \tag{5.23}$$

In Eq. (5.22) and Eq. (5.23), the modulation intensity $I'_m(u,v)$ is now described with

$$I'_m(u,v) = c_{r,u}c_{r,v}A_uA_v\sin(2\psi), \tag{5.24}$$

in which $c_{r,u}$ and $c_{r,v}$ are amplitude reflection coefficients of the BSC and A_u and A_v are the electric field amplitudes in the u and v direction, respectively (see Section 3.1). The factor $\sin(2\psi)$ in Eq. (5.24) disappears in an undisturbed interferogram, since then $\psi = \pi/4$ and the sine equals unity. Now, however, also an additional term in the interferograms occurs, indicated with $I_e(u,v)$ and described by:

$$I_e(u,v) = c_{r,u}c_{r,v}A_uA_v(-\cos(2\psi))\cos(\varphi(u,v)). \tag{5.25}$$

The value $I_e(u,v)$ is added to the background intensity $I_o(u,v) + I_r(u,v)$. In the worst case, assuming ψ to deviate from the required value $\pi/4$ by 0.0425 rad (as described in the beginning of this subsection), this result in a 0.4% lower modulation intensity and in an additional noise contribution with an amplitude factor of 0.085 $c_{r,u}c_{r,v}A_uA_v$. As a consequence, the measured phase distribution is corrupted. The error in the phase calculation algorithm can directly be found by collecting the errors in the interferogram into the phase term as an additional term 2ψ, writing

$$I_{\text{CCD2}}^{left}(u,v) = I_o(u,v) + I_r(u,v) + I_m(u,v)\cos(\varphi(u,v) - 2\psi) \tag{5.26}$$

and

$$I_{\text{CCD2}}^{right}(u,v) = I_o(u,v) + I_r(u,v) + I_m(u,v)\cos(\varphi(u,v) + 2\psi). \tag{5.27}$$

and using the series expansion of the cosine. This results in a measured phase difference distribution $\varphi'(x,y)$ given by

$$\varphi'(x,y) = \arctan\left(\tan(\varphi(x,y)) - \frac{\xi^2}{2}\tan(\varphi(x,y))\right). \tag{5.28}$$

in which ξ represents the angular mismatch (i.e. maximally 0.0425 rad). In this analysis, the interferograms recorded by CCD1 are assumed to have the proper phase steps. As a result, a coherent sinusoidal phase error is introduced, as is depicted in Fig. 5.5.

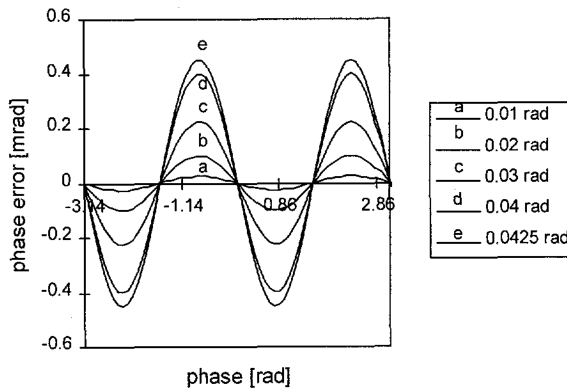


Fig. 5.5 *Calculated phase error for incorrect analyzer and HWP position.*

The larger the value for ξ, the more the form of this error deviates from a sinusoidal form towards a sawtooth-form. The actual phase difference $\varphi(u,v)$ determines the value of the error. The peak-to-valley phase error for $|\xi| <0.045$ is just below 1 mrad.

In conclusion, it can be stated that the incorrect azimuth angle of the QWP's in the interference section is the most dominant error source of the errors treated in this section, leading to a 70 mrad phase error. The phase error associated with the HWP and the analyzers remains smaller than 1 mrad. The performance deterioration in the shearing head is the second largest error source and is mainly due to the converging beams in the polarizing beam splitting cube PBC, leading to a 15 mrad phase error. The PBC tends to reduce the transmission of object waves more than that of the reference waves, whereas the quarter wave plates tend to reduce the reference wave transmission efficiency, nearly compensating the object wave reduction by the PBC.

One important error source associated with polarizing optical elements is not treated in this section. This error source was also mentioned in the description of errors associated with the imaging optics, since it is related to imaging through thick pieces of calcite. In calcite, the refractive index depends on the propagation direction of the light. Due to the converging beams inside the calcite elements, this variation in refractive index can deteriorate the interferogram. For the analysis of his error source we need the polarization ray tracing algorithm, as will be

developed in Section 5.2, while the influence of this error on the phase measurement will be described in Section 5.4.

5.1.2 Errors caused by decorrelation

Decorrelation between the various interferograms is one of the fundamental limits of ESPI[17,18]. In Section 2.2.3, it is discussed how the phase measurement is affected by speckle decorrelation. It was shown that two types of speckle decorrelation are present in our interferometer. One is caused by a displacement of the object due to the loading and the other occurs because of misalignment of the CCD cameras (pixel mismatch). Which type of decorrelation is dominant is determined by the actual alignment accuracy of the interferometer and the actual object displacement. The object decorrelation is reduced using a large aperture of the system, whereas the camera decorrelation reduces with smaller apertures. Depending on the application, the aperture of the multicamera system can be adjusted to obtain the optimal value. Object induced decorrelation can only be influenced by choosing the loading method but can not be avoided. The misalignment induced decorrelation can be minimized by optimizing the alignment of the system. How this is performed in our system will be described in Section 5.3. The alignment procedure is designed to facilitate an alignment accuracy of one tenth of a picture element. This way, a large aperture can be used, which increases the light gathering power of the system.

5.1.3 Errors caused by recording and computation system

Two types of errors can be ascribed to the optoelectronics, electronics and computational part of our system. The first one is caused by the recording device and its amplifiers: electronic noise and nonlinear response. The second type results from digitization and quantization of signals by the recording device and the finite accuracy of the phase calculation algorithm.

Electronic noise in the sensor causes a random fluctuation of the measured voltage or current. Thermal noise, shot noise, dark current noise, read-out noise and 1/f noise are contributions that occur even if no light impinges on the sensors. Additional noise is generated when photons are incident on the detector, mainly photon noise (shot noise as a result of the photon's statistical behavior). In applications of speckle interferometry, electronic noise plays a minor role as compared to the speckle noise.

For an accurate quantitative evaluation of results, a linear camera response is necessary. Non-linearity will cause phase errors. For commercial CCD cameras, the output deviates typically not more than 5% from the linear response. Depending on the used phase calculation algorithm, this leads to small phase errors, usually below 0.03 rad[19]. Some picture elements may show a deviation from a linear dependence larger than 5% due to fabrication errors. The corresponding

pixels are known as blemished pixels. Some camera manufacturers supply blemish compensation: the location of a bad pixel is stored in an extra chip (ROM) and its readout is replaced by the value of the previous pixel.
The phase calculation algorithms have a reducing effect on nonlinear camera response, with different efficiency for different algorithms. In general, the higher the number of buckets, the higher the efficiency will be. For a four bucket algorithm a 5% nonlinearity leads to phase errors in the order of 1 mrad. In multichannel interferometry the phase calculation is also corrupted by inhomogeneous camera responses or by inequality of the CCD cameras (i.e. the gain or the noise is different for the different cameras or picture elements). Adjusting the gain and bias of the cameras can compensate for inequalities between cameras but not between individual picture elements.

Digitization and quantization noise occur in interferometers using digital computer memory for storage and processing. The recorded intensity is digitized into an array of m×n picture elements and quantized into discrete gray values. Nowadays, most commercial cameras require a quantization into 8 bits, corresponding to 256 gray levels. Quantization noise is caused by this analog-to-digital conversion and is affected by the intensity of the signal[20]. The intensity of an interferogram should cover as much of the detector's dynamic range as possible. For an 8-bit (or more) digitization, the phase error caused by quantization is usually negligible[21]. Using a video amplifier can be of use when the quantization noise is larger than other noise sources, e.g. when there is very little light and the signal level has to be raised well above the quantization steps such that the signal can be represented with enough gray values. if the quantization noise is in the same range as other noise sources, further amplification of the signal is useless.
The effects of sampling the interferogram with a rectangular grid, has been described in Section 4.1.

5.1.4 Ambient disturbances

Ambient distortions may cause degradation of the interferograms. Vibration and air flow cause the most problematic errors[22] which reduce both the system measurement sensitivity and accuracy. Usual precautions are placing the interferometer on a vibration-isolated table, shielding the interferometers beam paths from air flow, and recording the interferograms in a short time. In our interferometer, however, the requirements of shielding are relieved to a great extend, since it is a real time shearing interferometer. For this system, the mechanical stability of the set-up itself is crucial, since the interfering beams do not coincide within the interferometer. This is successfully accomplished in our system, by using the mechanical set-up as is described in Section 3.6. Measurements with the system in industrial environments have shown that the real time shearing system is not limited by ambient disturbances other than vibration, since this induces object decorrelation.

5.2 Polarization ray tracing

In this section a method for polarization ray tracing, based upon the phase matching condition, will be presented. The algorithms derived in this section are restricted to uniaxial and linear anisotropic materials. The description of the performance of the Savart elements given in Section 3.5 is a simplified case, being restricted to a propagation direction of the incident light along the surface normal of the crystal. Here, the more general case of arbitrary incident beams will be treated. The derived polarization ray tracing equations enable the calculation of the effects for non-perpendicular incident beams. This is used for analyzing the interferogram displacements on the CCD cameras caused by rotations of the Savart elements, as it is used in the alignment procedure. Moreover, these polarization ray tracing equations are used for analyzing the aberrations induced by imaging through calcite, as will be described in Section 5.4.

Maxwell's equations govern the propagation of light through matter, but under the approximation that the wavelength is small, geometrical optics suffices to describe the propagating planar wave solutions to Maxwell's equations[23]. In geometrical optics, the wave front is considered to be locally planar. Speckled beams can be regarded as composed of a large collection of spherical wavelets, originating from the object under inspection, which is usually at a long distance from the interferometer. In the interferometer, only a small portion of each wavelet is present and in approximation it can be regarded as a planar wave. Therefore, geometrical optics will be used for ray tracing of speckled beams.
The direction of the energy flow of the local planar wave is indicated by the unit vector $\hat{\rho}$, called the ray direction, which is parallel to the Poynting vector. Generally, a ray changes its direction when it encounters an interface. Conventional geometric ray tracing calculates this change of direction as well as the optical path length. Polarization ray tracing supplements geometric ray tracing, taking the vectorial nature of the fields into account to provide the ray direction $\hat{\rho}$, the polarization direction of $\boldsymbol{E}$, and the optical path length *OPL* of rays as they propagate through an optical system that incorporates birefringent optical elements. It also provides information about the amplitudes of the refracted and reflected rays, using the Fresnel transmission and reflection equations. Since we are only interested in the directions of the ordinary and extraordinary beams as functions of the angle of incidence, unit vectors will be used (indicated by circumflexes).

For uniaxial crystals, like calcite, the response of the medium to an electromagnetic wave depends on the polarization state of the wave. The refractive index is a function of the wave direction and polarization state of the light[24]. The refractive index is not an input parameter to the optical system but has to be calculated for each ray. For each propagation direction, there are two refractive indices, associated with two orthogonal linear polarization states: the ordinary and the extraordinary beam. The ordinary rays behave as if the medium were isotropic and therefore they obey Snell's law of refraction in its simple form. In the case of extraordinary rays in a uniaxial refracting medium, Snell's law is not as easily applied because the refractive index of the medium depends on the direction of the refracted wave vector $\hat{\boldsymbol{k}}_e$ of the

extraordinary beam with respect to the optical axis $\hat{c}$. Thus, the refracted wave direction and the refractive index of the uniaxial medium are mutually dependent. This makes ray tracing more complicated than in isotropic media, in which the refractive index is a constant for all rays of a given wavelength, regardless of the propagation direction or polarization state.

5.2.1 Refractive index

Both wave vectors $\hat{\boldsymbol{k}}_o$ and $\hat{\boldsymbol{k}}_e$ obey Snell's law. For the *o*-wave this means:

$$n_i \sin(\theta_i) = n_o \sin(\theta_o) \tag{5.29}$$

and for the *e*-wave

$$n_i \sin(\theta_i) = n_\phi \sin(\theta_e), \tag{5.30}$$

in which n_i is the refractive index of the isotropic incident medium, n_o is a constant, whereas the refractive index n_ϕ is variable and depends on the direction of the refracted wave vector $\hat{\boldsymbol{k}}_e$. It can be calculated from[25]

$$\frac{1}{n_\phi^{\ 2}} = \frac{\sin^2(\phi_e)}{n_e^{\ 2}} + \frac{\cos^2(\phi_e)}{n_o^{\ 2}}, \tag{5.31}$$

where ϕ_e is the angle between the direction of $\hat{\boldsymbol{k}}_e$ and the optical axis. Note that if ϕ_e equals zero, $n_\phi = n_o$ and that $n_\phi = n_e$ if ϕ_e equals π/2, as can also be seen from Fig.5.6.

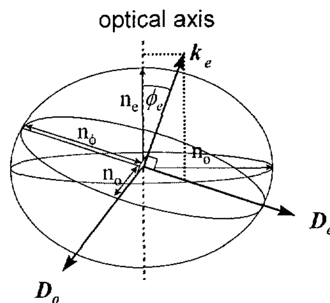


Fig. 5.6 *Indicatrix showing the dependency of the refractive index to the wave vector.*

The angle ϕ_e between the wave vector $\hat{\boldsymbol{k}}_e$ and the optical axis $\hat{c}$, can be obtained from

$$\cos(\phi_e) = \hat{\boldsymbol{k}}_e \cdot \hat{c}. \tag{5.32}$$

In the special case where the wave vectors and the optical axis are in the plane of incidence, (see Fig. 5.7) Snell's law for the extraordinary wave becomes

$$n_i \sin\theta_i = \frac{n_e n_o \sin(\phi_e - \beta)}{\sqrt{n_o^2 \sin^2\phi_e + n_e^2 \cos^2\phi_e}}, \tag{5.33}$$

where θ_i is the angle of incidence of light in the isotropic entrance medium with refractive index n_i, and β is the angle the normal $\hat{\eta}$ of the surface makes with the optical axis.

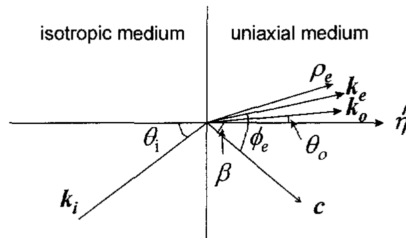


Fig. 5.7 *Ray doubling of a refracted wave at a birefringent interface.*

Since all other quantities in this equation are known, ϕ_e is uniquely determined but often must be solved for by iteration[24]. Such an iterative procedure, but then for the general case, for calculating ϕ_e and n_ϕ will be described in Subsection 5.2.2.

5.2.2 Maxwell theory applied to uniaxial media

The propagation of rays through a birefringent medium follows from the relation between the electric field vector $\boldsymbol{E}$, the electric displacement $\boldsymbol{D}$ and magnetic field $\boldsymbol{H}$. The Maxwell equations for planar monocromatic waves satisfying harmonic time dependence in a lossless, linear medium yield[26]

$$\boldsymbol{k} \times \boldsymbol{H} = -\omega \boldsymbol{D}, \tag{5.34}$$

$$\boldsymbol{k} \times \boldsymbol{E} = \mu_0 \omega \boldsymbol{H}, \tag{5.35}$$

$$\boldsymbol{k} \cdot \boldsymbol{D} = 0, \tag{5.36}$$

$$\boldsymbol{k} \cdot \boldsymbol{H} = 0. \tag{5.37}$$

Consider a planar wave incident on an uniaxial birefringent crystal. The refracted wave is in general split into two parts (ordinary and extraordinary). However, when the incident light propagation direction is along the surface normal, no change of direction will take place for the extraordinary ray if the optical axis is parallel or perpendicular to the refracting surface. In all other cases, the refracted wave is split into two parts, with wave vectors $\hat{\boldsymbol{k}}_o$ and $\hat{\boldsymbol{k}}_e$(see also Fig. 5.7). Uniaxial birefringent material, such as calcite, is characterized by a wavelength dependent tensor, the electric permittivity tensor ε, relating $\boldsymbol{D}$ to $\boldsymbol{E}$ via

$$\mathbf{D} = \varepsilon \mathbf{E}. \tag{5.38}$$

Generally, $\mathbf{D}$ is not parallel to $\mathbf{E}$ (for the extraordinary rays) as a result of the fact that ε is an arbitrary tensor. In a uniaxial material, the relation between $\mathbf{D}$ and $\mathbf{E}$ is invariant with respect to the rotation around the symmetry axis of the crystal, i.e. the optical axis $\hat{c}$. In an isotropic medium, according to the Maxwell equations, both $\mathbf{D}$ and $\mathbf{E}$ are perpendicular to $\mathbf{H}$, as well as the wave vector $\hat{\mathbf{k}}$ and the ray vector $\hat{\rho}$ for a planar wave in an anisotropic medium. Consequently, the vectors $\mathbf{D}$, $\mathbf{E}$, $\hat{\mathbf{k}}$ and $\hat{\rho}$ are coplanar. Furthermore, since $\hat{\rho}$ is perpendicular to $\mathbf{E}$ and $\hat{\mathbf{k}}$ to $\mathbf{D}$, the vectors $\mathbf{E}$ and $\mathbf{D}$ form the same angle with $\hat{\rho}$ and $\hat{\mathbf{k}}$. This means that for each ray, the $\hat{\mathbf{k}}$, $\hat{\rho}$ and $\hat{c}$ always ly in the same plane[27].

5.2.3 Algorithms

The first part of the calculation procedure determines the transmitted wave vector directions $\hat{\mathbf{k}}_o$ and $\hat{\mathbf{k}}_e$, and refractive indices n_o and n_ϕ for the ordinary and extraordinary rays, respectively. The ray directions $\hat{\rho}_e$ are used in transferring the rays to the next surface. However, $\hat{\mathbf{k}}_o$, $\hat{\mathbf{k}}_e$, n_o and n_ϕ are needed as an intermediate step in the determination of $\hat{\rho}_e$ and the *OPL*. For the calculation scheme two properties of refraction, as follow from the Maxwell equations, have been taken into account. First, the refracted wave-vector, $\hat{\mathbf{k}}_e$ lies in the plane defined by the incoming wave-vector $\hat{\mathbf{k}}_i$ and the surface normal $\hat{\eta}$ of the interface, and second, the ray vector $\hat{\rho}_e$ lies in the plane defined by the wave-vector $\hat{\mathbf{k}}_e$ and the optical axis $\hat{c}$.

Phase matching at the boundary between the birefringent medium and the surrounding medium leads to Snell's law, which determines the new wave direction. In vector form it states that[28] :

$$\hat{\eta} \times \left(n_i \hat{\mathbf{k}}_i\right) = \hat{\eta} \times \left(n_\phi \hat{\mathbf{k}}_e\right). \tag{5.39}$$

Since $\hat{\mathbf{k}}_e$ is coplanar with $\hat{\mathbf{k}}_i$ and $\hat{\eta}$, Snell's law may be written in another form[29]

$$n_e \hat{\mathbf{k}}_e = n_i \hat{\mathbf{k}}_i + \Gamma \hat{\eta}, \tag{5.40}$$

where Γ is called the astigmatic constant, which is determined by solving the quadratic equation that results from $\hat{\mathbf{k}}_e \cdot \hat{\mathbf{k}}_e = 1$:

$$\Gamma = -n_i \hat{\mathbf{k}}_i \cdot \hat{\eta} \pm \sqrt{n_i^2 \left(\hat{\mathbf{k}}_i \cdot \hat{\eta}\right)^2 + \left(n_\phi^2 - n_i^2\right)}. \tag{5.41}$$

The positive sign applies for refraction, whereas the negative sign is for reflection, which will be disregarded since we are not interested in reflected waves. If the transmitting medium is isotropic, then n_ϕ is a known quantity. In general, one needs an additional relation to find n_ϕ for both modes. Equation (5.31) relates the refractive index to the wave direction.

As pointed out previously, the wave direction $\hat{k}_e$ and the ray direction $\hat{\rho}_e$ generally do not coincide in anisotropic media. By definition, the direction in which a ray travels is the direction of its energy flow, which coincides with the direction of the Poynting vector. Consequently, the ray direction $\hat{\rho}_e$ can be determined using

$$\hat{\rho} = \frac{\Re\{\boldsymbol{E} \times \boldsymbol{H}\}}{|\Re\{\boldsymbol{E} \times \boldsymbol{H}\}|} . \tag{5.42}$$

To determine this, $\boldsymbol{E}$ and $\boldsymbol{H}$ have to be calculated. An alternative (and, in our case, simpler) way to calculate $\hat{\rho}_e$ is based upon using the relation between α and ϕ,

$$\tan(\alpha) = \frac{n_o^2}{n_e^2} \tan(\phi_e) . \tag{5.43}$$

The difference between the angles α and ϕ_e is called the dispersion angle.

The algorithm given below calculates the refractive index n, the wave vector direction $\hat{k}$ and ray direction $\hat{\rho}$ for each of the rays leaving the interface. The wave vector and refractive index for each exiting e-ray are determined through an iterative procedure. The steps for calculating $\hat{k}$ and n for each refracted and reflected ray are listed below. The subscript i refers to the incident ray, $\hat{\eta}$ is the surface normal at the intersection of the incident ray and the interface.

1. Start with an initial value for n_ϕ (e.g. n_o).
2. Calculate Γ from $\Gamma = -n_i \hat{k}_i \cdot \hat{\eta} + \sqrt{n_i^2 \left(\hat{k}_i \cdot \hat{\eta}\right)^2 + \left(n_\phi^2 - n_i^2\right)}$.
3. Calculate $\hat{k}_e$ from $\hat{k}_e = \dfrac{n_i \hat{k}_i + \Gamma \hat{\eta}}{\left|n_i \hat{k}_i + \Gamma \hat{\eta}\right|}$.
4. Calculate ϕ_e from $\phi_e = \arccos\left(\hat{k}_e \cdot \hat{c}\right)$.
5. Calculate a new value for n_ϕ using $n_\phi = \dfrac{n_o n_e}{\sqrt{n_o^2 \sin^2(\phi_e) + n_e^2 \cos^2(\phi_e)}}$.
6. Repeat steps 2-5 until the value of n_ϕ converges.
7. The final value of $\hat{k}_e$ from Step 3 and the final value of n_ϕ from Step 5 are the wave vector and the refractive index, respectively, for the refracted wave.

In practice, convergence of the algorithm is very rapid, taking no more than two or three iterations for n_ϕ to stabilize within 10^{-4}. Once the wave vector $\hat{k}_e$ and the refractive index n_ϕ are established, the ray direction can be calculated. Since the ray direction is within the plane defined by the optical axis and the wave vector, it can in general be written as

$$\hat{\rho}_e = \frac{\hat{\boldsymbol{k}}_e + \Lambda\hat{c}}{\left|\hat{\boldsymbol{k}}_e + \Lambda\hat{c}\right|}, \tag{5.44}$$

where the constant Λ has to be chosen such that the angle α between the ray direction and the optical axis satisfies Eq. (5.43). From this requirement it follows that Λ has to be equal to

$$\Lambda = \left(\frac{n_e^{\ 2}}{n_o^{\ 2}} - 1\right)\cos(\phi_e), \tag{5.45}$$

which gives the normalization factor for the ray direction via

$$\left|\hat{\boldsymbol{k}}_e + \Lambda c\right| = \sqrt{1 + \Lambda^2 + 2\Lambda\cos(\phi_e)} = \sqrt{1 + \left(\frac{n_e^{\ 4}}{n_o^{\ 4}} - 1\right)\cos^2(\phi_e)} = \frac{\sin(\phi_e)}{\sin(\alpha)}. \tag{5.46}$$

Therefore Eq.(5.44) results in

$$\hat{\rho} = \frac{\hat{\boldsymbol{k}}_e + \Lambda\hat{c}}{\left|\hat{\boldsymbol{k}}_e + \Lambda\hat{c}\right|} = \frac{\hat{\boldsymbol{k}}_e + \left(\frac{n_e^{\ 2}}{n_o^{\ 2}} - 1\right)\cos(\phi_e)\hat{c}}{\sin(\phi_e)/\sin(\alpha)}. \tag{5.47}$$

The optical path length between the entrance and exit surfaces is computed with the ray direction, the wave direction and the refractive index from

$$OPL = nl\hat{\boldsymbol{k}} \cdot \hat{\rho}, \tag{5.48}$$

in which l is the geometrical length of the ray path.

5.3 The camera alignment procedure

In a multichannel interferometer, such as our real time phase stepped shearing set-up, speckle decorrelation between the four simultaneously recorded interferograms can occur due to misaligned cameras. Instead of using four separate CCD cameras we have chosen to use two CCD cameras to record the four interferograms. The use of calcite analyzers in the polarization phase stepping system to determine the phase step led to the idea to use a Savart element as an analyzer which, owing to its birefringent nature, can simultaneously be used as beam splitter. The basic idea behind this choice is a reduced sensitivity to environmental variations and a more compact system since less cameras are needed. As we saw in Chapter 1, only three interferograms are needed to obtain the phase distribution of an interferogram. In principle, more buckets yield more accurate phase patterns, less a-priori knowledge of the measurement

(e.g. phase step) is necessary and more error sources can be eliminated. By recording two interferograms on a single camera, the system design can be more compact and more interferograms are available for calculating the phase distribution, making the phase calculation algorithm more robust. Moreover, the alignment of two interferograms on a single camera is expected to be mechanically more stable than if they were recorded on separate cameras. An obvious disadvantage is the reduction of the number of picture elements per interferogram. The interferograms on a single camera are used in calculating the denominator of the argument of the arctangent function used in calculating the phase distribution, and the two on the other camera for its numerator (See Section 3.1). Thus, even if there is a relative displacement between the two CCD cameras (e.g. due to vibrations), the calculation of the numerator and denominator of the arctangent function will be correct. This reduces the error in the phase calculation algorithm. However, the measuring accuracy of the interferometer not only depends on the number of interferograms. For accurate measurements, the interference patterns on the CCD cameras have to be mutually related. This requires a precise mutual alignment of the images on the cameras, which forms a complicated task. The alignment of separate CCD cameras is already intensely investigated and reported[15]. To achieve an accurate alignment of the two images on a single camera, rotations of the Savart elements can be used. The better the alignment, the smaller the speckle size (and thus the larger the aperture) can be. Larger apertures reduce the errors caused by object decorrelation. However, an excessively small speckle size puts stronger requirements on the rigidity of the mechanical construction in order to avoid an increase of the system's sensitivity for camera decorrelation. As a trade off, the goal of the alignment procedure is to obtain an alignment accuracy of one tenth of a picture element, with the picture element size being 11 μm.

Two types of alignment errors are possible, those that result in a relative displacement between the interferograms in the detection plane (in-plane misalignment) and the ones that result in an out-of-plane displacement (out-of-plane misalignment). The average length of speckles is much larger than their average diameter (see Section 2.2). These long speckles have a long correlation length with the corresponding speckles in other images. In-plane decorrelation is thus more critical and is reduced by using small apertures. The cameras have six degrees of freedom: 3 rotations and 3 translations. Since the in-plane decorrelation is much more critical than the out-of-plane decorrelation, our mounting of the CCD cameras is chosen such that it allows accurate adjustment of the two in-plane translations and of one in-plane rotation. Since only the relative positions of the cameras are of influence, the CCD camera in Branch 1 (CCD1) has only a horizontal adjustment possibility, whereas CCD2 can only be adjusted vertically. The coarse alignment of the in-plane rotation is performed using an alignment microscope, whereas the coarse out-of-plane alignment is accomplished using the accurately positioned fixation points on the base plate of the interferometer.

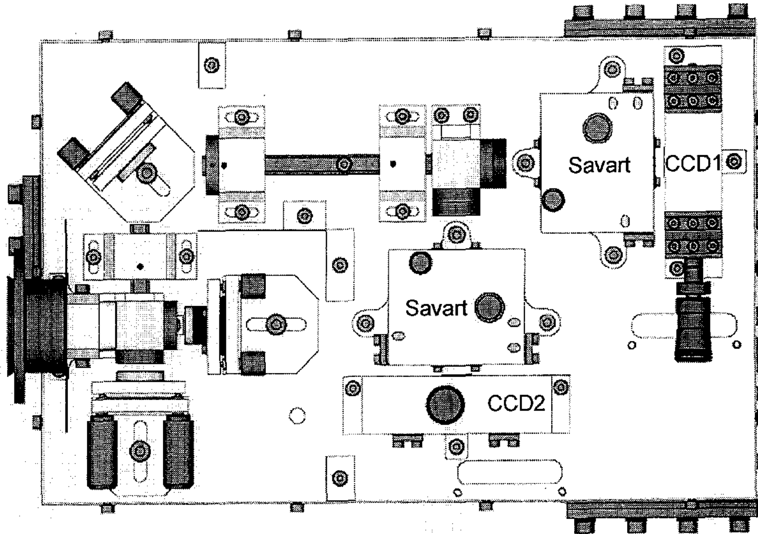


Fig. 5.8 *Top view of the set-up. The CCD camera mounts can be seen at the right-hand side (CCD1) and at the bottom (CCD2) of the image.*

5.3.1 Using Savart elements for alignment

An analysis was performed to study the operation of Savart elements. Special attention was paid to determine which rotations can be used to align the images with sub-pixel accuracy. Three rotation axes have been analyzed, the u-, v- and w-axis. In Fig. 5.9 the axes with respect to the Savart element are indicated.

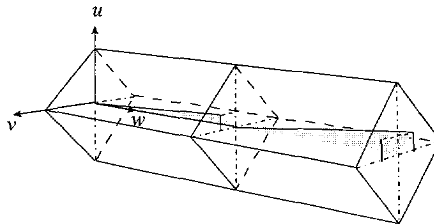


Fig. 5.9 *Orientation of the* u,v,w*-coordinate system.*

In this section the results of the analysis are described. We used polarization ray tracing to calculate the optical paths of the interferograms towards the detection plane. The optical paths of the beams in the Savart elements are described by the unit vector $\hat{\rho}$ parallel to the Poynting vectors in the Savart elements. The ray tracing algorithms have been described in Section 5.2.

From the computed orientations of the vectors $\hat{\rho}_0$ and the $\hat{\rho}_e$ for the ordinary and extraordinary waves, respectively, the positions of the images can be calculated. The position of the images is calculated in the detection plane, located 5 mm behind the back face of the Savart elements. Since we are working with slowly converging beams, the displacement of an image will be calculated using the chief ray. The influence of the convergence of the beams is further discussed in Section 5.4.

Owing to the birefringent nature of calcite, rotating a Savart element around the *u*- or *v*-axis does not only result in a plane-parallel plate shift of the interferograms on the cameras, but also to a relative shift in the vertical and horizontal directions, respectively. In Fig. 5.10, the image positions as a function of the rotation around the *u*-axis are depicted.

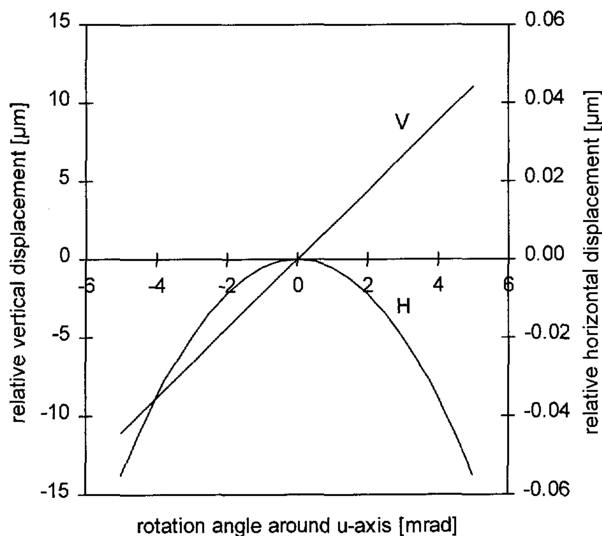


Fig. 5.10 *Horizontal (H) en vertical (V) relative displacement of the images for rotations around the u-axis.*

From Fig. 5.10 it is clear that if the Savart element is rotated around the *u*-axis, relative horizontal and vertical displacements between the two images occur. Beside these relative displacements, also a large horizontal displacement of both images is present. This is nearly equal for both images and is the parallel-plate shift. For a rotation of 5 mrad, this shift has been calculated to be almost 107 μm for both images. The small difference in this horizontal shift between the left-hand and the right-hand image is 0.055 μm, which is small as compared to the more interesting shift in this case: the relative vertical displacement of the images which amounts to 11 μm for 5 mrad rotation. As a result, the rotation around the *u*-axis can be used effectively to correct for vertical alignment errors.

In the same way the shifts of the images have been analyzed for rotations of the Savart elements around the *v*-axis. In Fig. 5.11, the relative image positions as a function of the rotation around the *v*-axis is given.

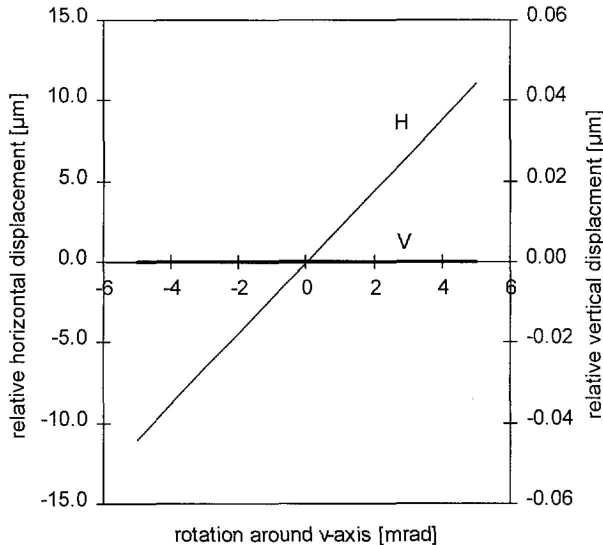


Fig. 5.11 *Horizontal (H) en vertical (V) relative displacement of the images for rotations around the v-axis.*

The analysis showed a 107 µm absolute shift of the images in the vertical direction for a rotation of 5 mrad. For both images this displacement proved to be equal (the relative vertical displacement was vanishing within the accuracy of the computer program). It appeared that a 5 mrad rotation gives a useful relative displacement in horizontal direction of 11 µm. Therefore, this rotation is suitable for correcting horizontal alignment errors.

The analysis of the image positions as a function of a rotation around the *w*-axis does not require complicated polarization ray tracing computations. The relative orientations of the principal planes in the element are not changed by *w*-axis rotations, as the orientation of the incident light wave remains perpendicular to the input face of the Savart element. Simple geometrical calculations show a relative vertical displacement of the images, as well as a small relative horizontal displacement. In Fig. 5.12, the calculated values of the displacements are given as a function of a rotation around the *w*-axis.

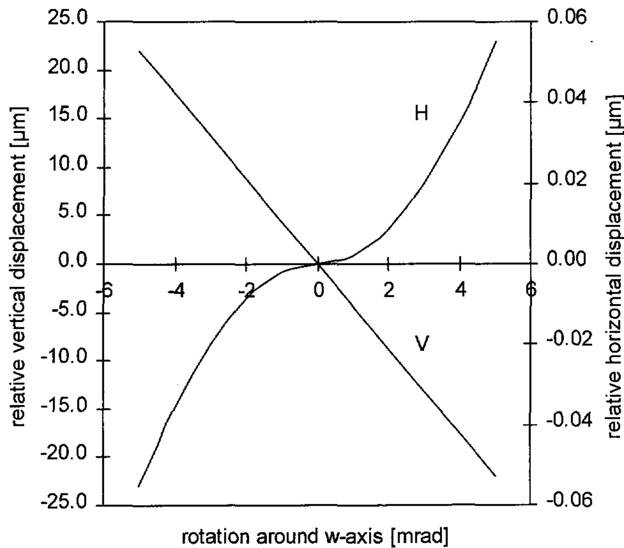


Fig. 5.12 *Horizontal (H) en vertical (V) relative displacement of the images versus rotation around the w-axis.*

Fig. 5.12 shows that a rotation of 2.5 mrad around the *w*-axis results in a relative vertical displacement of the images of 11 µm. The relative horizontal displacement for this rotation is only 0.011 µm.

For the alignment of the images, only two rotations are needed: one for the vertical alignment and one for the horizontal alignment. For horizontal alignment, only a rotation around the *v*-axis can be used. For the vertical alignment, a choice has to be made between a rotation around the *u*-axis and one around the *w*-axis. The main advantage of the rotation around the *w*-axis is that there is no absolute displacement ('parallel-plate shift') of both images. A disadvantage of this *w*-rotation is that it influences the correct phase step value. However, for a rotation of 2.5 mrad the effect on the phase step is small (see Section 5.1). Both the *w*- and *u*-axis rotation show a coupling of relative vertical and horizontal displacement, but for both rotations this coupling is extremely small (0.5% and 0.6% respectively) and can thus be ignored. A rotation around the *u*-axis has two additional minor disadvantages: a polarization interchange error of 0.01% and a small optical path length difference. Since the rotation around the *w*-axis does not have these disadvantages, and thanks to the absence of the parallel-plate shift, this rotation is preferred for use in the alignment procedure.

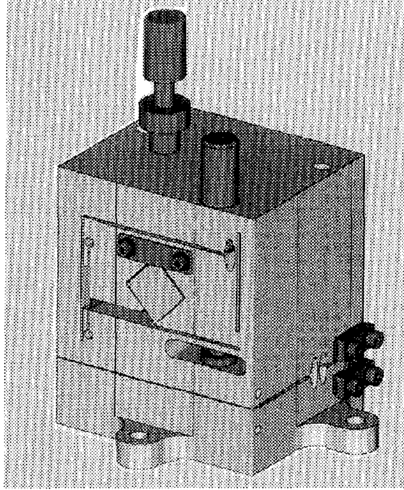


Fig. 5.13 *Special Savart element rotation stage. The Savart element is the tilted crystal in the center of the mount.*

5.3.2 Alignment procedure

For sub-pixel alignment of the two CCD cameras, the procedure has been described previously[8]. Using the rotations of the Savart elements around the *v*- and *w*-axis, the sub-pixel alignment of the two images on one CCD camera can be performed. Then, the alignment procedure is as follows:

1. Orientation of cameras around their *u*,*v*,*w*-axes using a moiré technique.
2. Orientation of the Savart elements around their *w*-axis using a cross-hair and gratings.
3. Relative positioning of the two cameras using a moiré technique.
4. Sub-pixel alignment of the images on CCD1 using correlation optimization.
5. Sub-pixel alignment of the images on CCD2 using correlation optimization.
6. Sub-pixel alignment of CCD1 and CCD2 using correlation optimization.

Steps 1 to 3 are coarse alignment steps, which are refined by Steps 4 to 6. The procedure starts by aligning the cameras when the Savart elements are not yet inserted. This way, the whole detection area of the cameras can be used in the first step of the alignment process, which maximizes its accuracy. In this first step, the angular orientation of the cameras can be aligned using a grating with vertical (V-grating) or horizontal (H-grating) lines (10 or 20 lines per mm), positioned near the intermediate image plane of the interferometer, such that the grating is sharply imaged on the detection planes of the cameras. The grating is used to correct for out-of-plane rotation differences between the two cameras by subtracting the images of the two cameras and displaying the resulting moiré pattern. The V-grating is used to correct the *u*-rotation and the H-grating to verify the *v*-rotation. Next, the in-plane orientation of the cameras

can be aligned by imaging a H or V grating by displaying the resulting moiré pattern and correcting for the low frequency lines perpendicular to the grating lines, as is depicted in Fig. 5.14.

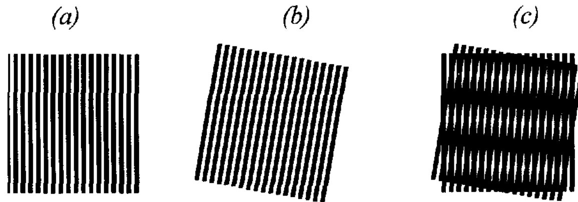


Fig. 5.14 *In-plane angular alignment of the cameras. Schematic representation of (a) vertical lines recorded by CCD1, (b) those on CCD2 and (c) the resulting moiré pattern.*

An additional method for coarse alignment of the cameras was found during the alignment process. The picture element grid of a camera appears to be working as a two-dimensional grating. The back reflection pattern of an incident collimated laser beam can be used for aligning the picture element lines in the horizontal direction and the columns in the vertical direction as well as minimizing the out-of-plane rotation of a camera.

After the angular alignment of the cameras, the Savart elements are inserted in the system. A cross-hair at the intermediate image plane is imaged as sharp as possible onto the cameras, by translating the cameras. For this, the aperture of the system is as large as possible, creating the smallest focal depth. Again the grating is used to check if an out-of-plane rotation of the cameras has occurred and if it is necessary to realign the cameras. To align the Savart element around the w-axis, first the cross-hair is used. If the Savart elements are not positioned with their optical axes making an angle of $\pi/4$ rad with the u-axis, the displacement of the two interferograms is not the same (see Fig. 5.12). As a result, the horizontal lines of the two cross hair images is mutually shifted, as is depicted in Fig. 5.15.

Fig. 5.15 *Positions of the cross-hairs in the detection plane (a) for incorrect angular position of the Savart element around the w-axis and (b) for a correct orientation.*

After correcting the orientation of the Savart elements, this alignment step is repeated but the cross-hair being replaced with a horizontal grating. This enables a further accuracy increase of the angular position, since the visibility of the grating is better than that of a cross hair.

The cameras are correctly oriented thanks to the alignment of Step 1, but there may be a relative horizontal or vertical shift of the cameras present. A shift equal to an integer multiple of the picture element size can be corrected in the software routines, but since it is a waste of picture elements, it is preferred to record corresponding parts of the interferograms by corresponding picture elements. This can be performed coarsely by overlapping the cross hair projections, recorded by the two cameras and refining this process with overlapping the horizontal and vertical grating lines. If the images of the two cameras are subtracted and the relative shift is zero, the resulting image should display an equal gray value for each picture element. An incorrect alignment will result in a display of the cross-hair or the grating structure. Using this method of translating the cameras until the resulting difference image is homogeneous, the cameras are aligned on the picture element level.

Steps 4 to 6 are used to refine the adjustments made in Steps 1 to 3. A speckle pattern is generated close to the first lens of the system. No illuminated object is imaged, since then the average intensity in the detection plane is near to the sensitivity threshold of the cameras for small apertures. Small apertures are preferred since this generates large speckles, making the alignment process easier since less noise is present in the calculated correlation. For the same reason, the shear mirror is covered, such that only contributions from the reference mirror proceed further into the system. The correlation is optimized by measuring the correlation coefficient, defined by

$$|\mu(I_1, I_2)|^2 = \frac{\langle I_1 I_2 \rangle - \langle I_1 \rangle \langle I_2 \rangle}{\sqrt{\langle I_1^2 \rangle - \langle I_1 \rangle^2} \sqrt{\langle I_2^2 \rangle - \langle I_2 \rangle^2}}, \tag{5.49}$$

between two interferograms I_1 and I_2 and rotating the Savart elements first around the w-axis and thereupon around the v-axis.

The last step is to optimize the alignment between the two cameras. For this purpose, one of the interferograms on CCD1 is taken as a reference and the correlation of this interferogram with one of the interferograms on CCD2 is optimized by using horizontal translations of CCD1 and vertical translations of CCD2.

At the end of the alignment procedure, a correlation of 0.91 has been obtained at an aperture of 2 mm. Furthermore it is noted that the alignment accuracy of interferograms recorded on the left-hand parts of the CCD is higher than the alignment accuracy of an interferogram recorded on the left-hand side and the other on the right-hand side, the difference being dependent on the used aperture (for an aperture of 2 mm: approximately 5%). This may be due to converging beams inside the calcite elements, as will be described in Section 5.4.

5.4 Image formation through Savart elements

As pointed out in Section 5.1, imaging through thick pieces of birefringent material generates errors, since the performance of Savart elements depends on the propagation direction inside the elements. Three effects occur by imaging through calcite, two of which are also present in ordinary plane parallel plates, made from isotropic material. The first effect is displacement of the focal plane. Secondly, Seidel aberrations (mainly spherical aberration for small angles of incidence) are introduced. These effects are relatively unimportant, since the focal plane shift can be compensated for (by shifting the detection planes accordingly) and the Seidel aberrations are relatively small due to the small aperture in the system and because they are equal for all four cameras. The third effect is the most important one, since it has a considerable influence on the performance of the system. It is generated by the birefringent character of the material and causes a decrease of the modulation intensity in the recorded interference patterns (see Fig. 5.16). The average modulation intensity of a measured phase distribution in our interferometer is calculated from the four interferograms using (see Section 4.2)

$$\langle I_m(x,y)\rangle = \left\langle \frac{1}{2}\sqrt{\left(I_1(x,y)-I_3(x,y)\right)^2+\left(I_4(x,y)-I_2(x,y)\right)^2}\right\rangle . \tag{5.50}$$

In ideal situations, the average modulation intensity increases linearly with increasing average intensity on the detection plane. Since the average intensity of the light at the detection plane increases quadratically with increasing aperture, an ideal modulation intensity also shows this quadratic behavior. The measured and expected modulation intensities, both as functions of the diaphragm diameter in our system, are depicted in Fig. 5.16.

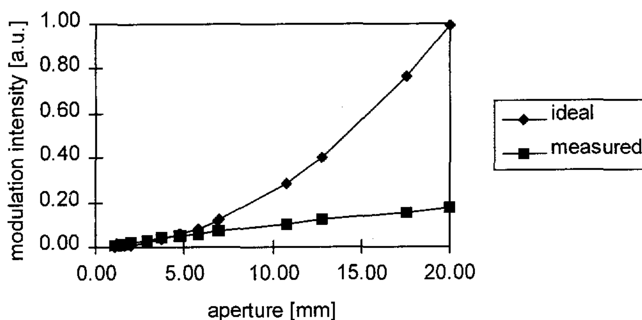


Fig. 5.16 *Modulation depth versus the aperture size.*

The modulation intensity shows a (nearly) linear increase with increasing aperture, instead of the expected quadratic behavior. This linear behavior can be understood from image formation analysis through calcite. In the following subsection, the effects of imaging through calcite will be analyzed by using the polarization ray tracing algorithm developed in Section 5.2.

5.4.1 Polarization ray tracing for imaging analysis

In this section beam propagation through the Savart elements will be analyzed. Our main interest is to determine the effect of the convergence of the beams on the image quality, since the Savart elements are designed for collimated beams. With the theory presented in Section 5.2, a computer program is developed, which traces rays through calcite elements, taking the polarization directions into account. The incident beam is split into two parts. One part travels as an ordinary beam in the first segment of the Savart elements and as an extraordinary one in the second segment. This beam will further be indicated as the *oe* beam. The second part starts as an extraordinary beam in the first segment and becomes ordinary in the second segment of the Savart element, this one will be indicated with *eo*. These beams will be recorded on different parts of the CCD camera.

Using the computer program, the deformation of a converging beam, with a circular cross-section at the entrance face of the calcite elements, is calculated. Polarization ray tracing is applied on rays at the outer boundary of a (virtual) circular aperture positioned just in front of the entrance face of the Savart elements. The radius of this aperture is called the semi-aperture *SA*. In Fig. 5.17 the endpoints of the marginal rays, coming from a two mm radius *SA*, after traveling through one of the Savart elements are displayed. The points in the center of the ellipses are the crossings of the chief ray.

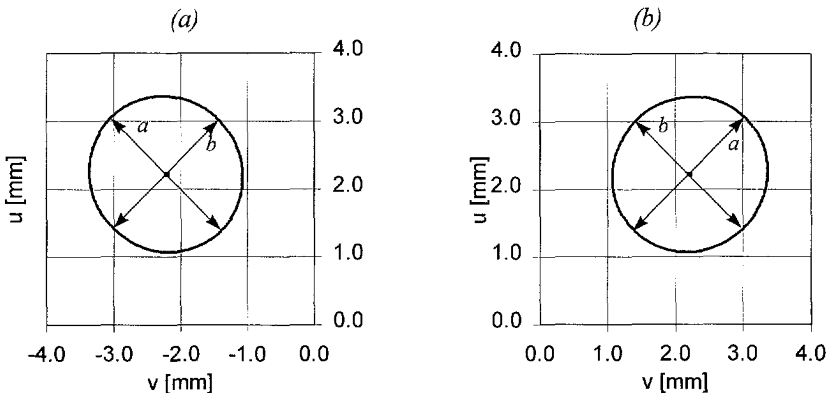


Fig. 5.17 *Deformation of a converging circular beam by imaging through calcite for (a) oe beam and (b) eo beam.*

The circular beams at the entrance face of the Savart elements are deformed into elliptical beams at the exit face of the elements. The long axes (indicated by a in Fig. 5.17) of the ellipses make an angle of $\pm\ \pi/4$ rad with the u-axis and the short axes (indicated with b) are perpendicular to this direction, namely being under an angle of $\pm\ 3\pi/4$ rad with the u-axis. This will lead to two foci, as is shown in Fig. 5.18.

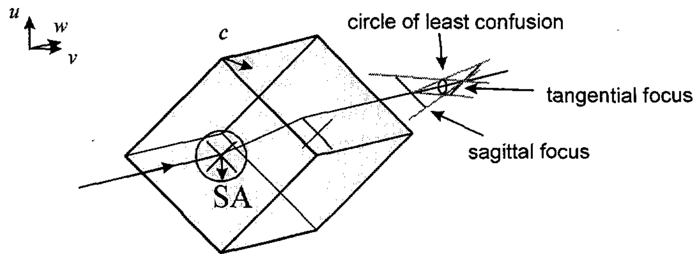


Fig. 5.18 Astigmatism in calcite.

The deformation of a circular beam into an elliptical one is due to astigmatism. The difference in length between the long axis a and short axis b of the ellipse at the exit face of the Savart element equals the length of the astigmatic lines. Calculation of the astigmatism due to imaging through Savart shows that the larger the SA (or equivalently the NA, since they are related due to the geometry of the interferometer by SA=NA*43mm), the more the astigmatism. Fig. 5.19 shows this linear relation between a-b and the NA for values of NA up to 0.12.

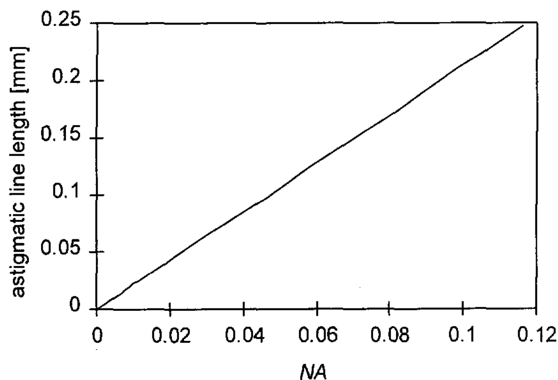


Fig. 5.19 Astigmatic line length (a-b) as function of the numerical aperture.

This relation between the lengths of the astigmatic lines and the aperture applies for both the *oe* and the *eo* beams. The distance between the astigmatic lines is determined by the angle of propagation in calcite, in our case nearly 6 degrees, and equals 2.2 mm. The diameter of the circle of least confusion is half the length of the astigmatic lines, i.e. (a-b)/2. The aberration increases linearly with the plate thickness t.

The presence of astigmatism is confirmed by experiments. A circular spot is projected on a planar diffusely reflecting object, which is in turn imaged on the detection plane. If the detection planes of the interferometer coincide with the circle of least confusion, the spot is circular. However, if the focusing is changed by displacing the objective lens of the interferometer, the circular spots change into lines. In Fig. 5.20, such a defocused image, recorded by CCD1, is depicted. Two sets of two lines are recorded: one set at the left-hand side

of the camera and the other at the right-hand side. The two lines in each set correspond to the object beam (upper one) and the reference beam (lower one), where the displacement is caused by tilting the mirror in the shearing head.

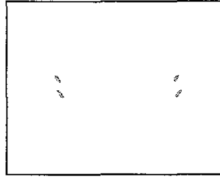


Fig. 5.20 *Deformation of a defocused circular spot (gray values of image are inverted).*

As a consequence of astigmatism, the speckles will spread over a larger part of the detection plane than they would do without astigmatism. This has been experimentally verified by measurement of the correlation between the speckle patterns recorded on either side of a CCD camera. It shows an increased correlation with increasing aperture, whereas in normal cases, a decrease is expected (see Fig. 5.21).

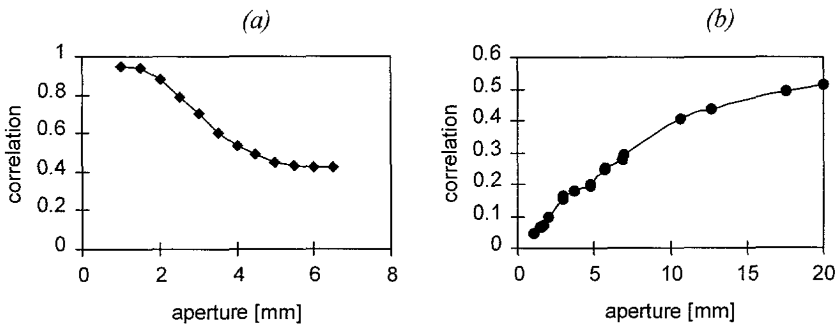


Fig. 5.21 *Correlation for (a) normal speckle field and (b) two astigmatic speckle fields.*

5.4.2 Unresolved speckle patterns

The effect of astigmatism on the performance of the interferometer can be understood using the theory of unresolved speckle patterns. In this theory, the speckle fields are assumed to be fully developed and linearly polarized, which means that the field intensity obeys negative exponential statistics and that their phase is uniformly distributed (see Section 2.2). If K speckles are recorded on a picture element, the total measured intensity I is given by

$$I = \frac{1}{K}\sum_{k=1}^{K}\left(I_{b,k} + I_{m,k}\cos\left(\varphi_k + \theta\right)\right), \tag{5.51}$$

where the index k refers to the kth speckle. The intensity I, which represents the incoherent sum of n independent interferograms, can again be written as

$$I = I_b + I_m \cos(\varphi + \theta), \tag{5.52}$$

but with new values of I_b and I_m and φ, satisfying

$$I_b = \frac{1}{K}\sum_{k=1}^{K}\left(I_{b,k}\right) \tag{5.53}$$

and

$$I_m e^{i\varphi} = \frac{1}{K}\sum_{k=1}^{K}\left(I_{m,k}e^{i\varphi_k}\right). \tag{5.54}$$

Equation (5.53) shows that the resulting background intensity corresponds to the average of the various background intensities. Equation (5.54) shows that a uniform distribution of the phases φ_k leads to a decrease of the resulting modulation. The average value $\langle I_m \rangle$ of the modulation intensity, which is the value of interest in our case, satisfies[30]

$$\langle I_m \rangle = \frac{(2K)!}{\left(2^K K!\right)^2}\pi\langle I \rangle. \tag{5.55}$$

where <I> is the average intensity of each speckle field. For large values of K, This equation can be approximated by

$$\langle I_m \rangle \approx \sqrt{\frac{\pi}{K}}\langle I \rangle, \tag{5.56}$$

which is accurate within 1% for K equal to 15. The modulation depth of an interferogram, defined as $\langle I_m \rangle / \langle I \rangle$, as a function of the number of speckles per pixel is depicted in Fig. 5.22.

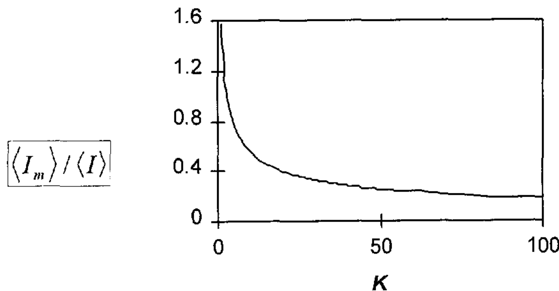


Fig. 5.22 Reduction of the modulation intensity due to unresolved speckles. The number n indicates the number of speckles per picture element.

For normal speckle patterns, the speckle diameter is determined by the aperture of the beams, according to

$$d_s = 0.61 \frac{\lambda}{NA}. \tag{5.57}$$

Thus, increasing the *NA* by a factor of *K*, the speckle diameter is decreased with the same factor *K*. As a result, the number of speckles per picture element increases by K^2. Thus

$$K \propto \sqrt{NA}. \tag{5.58}$$

This leads to a lower modulation depth $\langle I_m \rangle / \langle I \rangle$ of an interferogram, which is approximately described by

$$\langle I_m \rangle \propto \sqrt{\frac{\pi}{\sqrt{NA}}} \langle I \rangle. \tag{5.59}$$

For astigmatic speckle fields, the aperture has a larger influence on the modulation depth than for normal speckle fields. In our case, the imaging through calcite results in a spreading of the speckles. As an example, a numerical aperture of 0.06 would normally result in an average speckle size equal to the picture element's size, but now results in a much larger speckle, given by the diameter of the circle of least confusion, i.e. 63.8 μm. This means that a single speckle is smeared out over approximately 25 picture elements. It also means that each picture elements receives contributions of 25 speckles. This results in a lowering of the modulation depth with a factor

$$\frac{(2K)!}{\left(2^K K!\right)^2}, \tag{5.60}$$

which equals 0.35 for *K*=25.

The polarization ray tracing analysis has shown a linear relation (see Fig. 5.16) between diameter of circle of least confusion D_{clc} and the numerical aperture *NA* of the beams inside calcite:

$$D_{clc} \propto NA. \tag{5.61}$$

The number of speckles integrated on a single picture element depends quadratically on the diameter of the circle of least confusion:

$$K \propto D_{clc}^{\ 2}. \tag{5.62}$$

Thus, the number of overlapping speckles per picture element is proportional to the square of the numerical aperture as is described by

$$K = NA^2 . \tag{5.63}$$

Combining this with Eq. (5.56), the modulation depth is influenced by the numerical aperture as

$$\langle I_m \rangle \propto \sqrt{\frac{\pi}{NA^2}} \langle I \rangle = \frac{1}{NA} \sqrt{\pi} \langle I \rangle . \tag{5.64}$$

This equation shows that the modulation intensity of a astigmatic speckle field increases linearly with the aperture, since $\langle I \rangle$ increases quadratically with the aperture. This is in perfect agreement with our measurement (see Fig. 5.16).

Conclusion

Imaging through the Savart elements leads to astigmatism. This in turn leads to a lower modulation depth of the resulting interference patterns. This effect can be described by expanding the theory of unresolved speckle patterns to speckle patterns, in which several speckles overlap on a single picture element.

The modulation depth of aberration free speckle patterns is decreased for larger apertures due to the averaging of multiple speckles by a single picture element and it is inversely proportional to $\sqrt[4]{NA}$. Due to the imaging through the Savart elements, the modulation intensity decreases much faster and is inversely proportional to the aperture. This is summarized in Fig. 5.23.

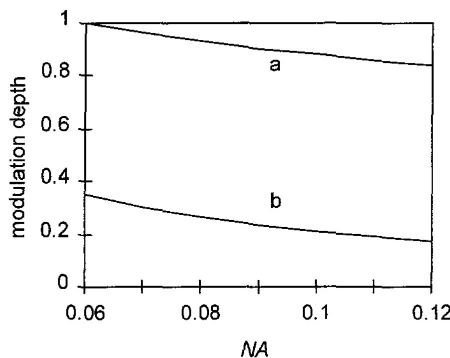


Fig. 5.23 Reduction of the modulation depth due to (a) unresolved speckles and (b) astigmatic speckles.

For optimal modulation depth, astigmatism should be avoided. Fig. 5.19 shows that astigmatism reduces to zero for a collimated incident beam (vanishing *NA*). Therefore, the Savart elements should preferably be positioned in a collimated beam.

Note

The quadratic dependence of the modulation intensity is only valid for resolved speckle patterns. For unresolved speckle patterns, the modulation intensity reduces slightly as a consequence of averaging multiple speckles in a single picture element (see also the discussion in Section 2.2). This effect is not taken into account in the 'ideal' curve in Fig. 5.10. If it would have been, the modulation intensity would show a somewhat slower increase than the indicated quadratic one, starting from values of the aperture of approximately 5mm and as a result the expected value at 20 mm would have been roughly 25% lower than the presented one.

5.5 Measured accuracy

From the analysis presented in the previous sections of this chapter, it is deducted that many noise sources reduce the accuracy of the system. Some errors vary with the actual application (e.g. object decorrelation) and some depend merely on the optical system (imperfect alignment or performance of optical elements). From the analysis it becomes clear which are the most important noise sources, but it is quite complicated to obtain a value for the overall accuracy of the system. It would be advantageous to use a single parameter to indicate the accuracy of the system. This can e.g. be done by measuring the smallest deformation of an object surface (expressed in parts of the wavelength of the used light) that exceeds the noise sufficiently to be measured. Using this definition, a measure of the system's accuracy can be given.

We determined the phase change of a diffusely reflecting object, which is composed of two parts, as depicted in Fig. 5.24. Using this test object, the result is independent of the shearing distance, which normally determines the measurement sensitivity of the shearing interferometer, since the deformation is sharply localized.

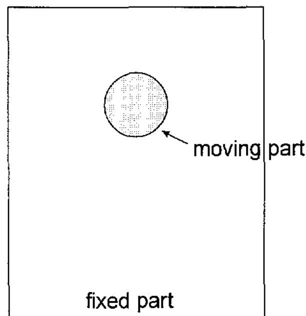


Fig. 5.24 *Test object.*

The moving part is mounted on a piezoelectric transducer, coupled to the DA-converter of a computer to control the out-of-plane displacement. Measurements show that a 22 nm (i.e. $\lambda/44$) displacement of the moving part can be measured. At this value for the displacement, the signal to noise ratio just exceeds the value of 1. Since we measure in reflection, the phase change associated with this translation is $2\pi/22 = 0.28$ rad. This value is considered as the measured phase error. The phase error depends on the modulation intensity of the interferograms, which ca be altered by adjusting the aperture of the interferometer.

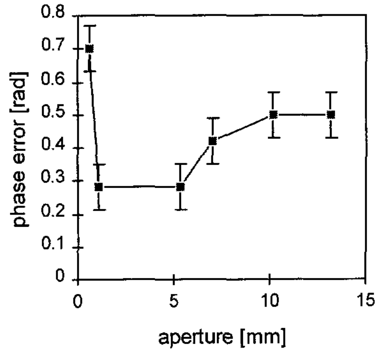


Fig. 5.25 Measured phase error versus aperture size.

This measurement shows an increase of the phase error for larger apertures, as could be expected since for larger apertures the convergence of the beams in the polarizing optical elements is larger, leading to larger phase errors. For smaller apertures, the intensity at the detection planes becomes lower but only for apertures below 1 mm it becomes such low that electronic noise and quantization noise become dominant. However, in normal applications, the phase error would increase at larger values of the aperture than the 1 mm measured here, since object decorrelation takes an important role. Only for this special test object, the measurement is not affected by object decorrelation, since the fixed part of the object is not affected by the displacement of the moving part. Therefore, this measurement yields the accuracy of the interferometer itself.

In conclusion, there is an optimal aperture of the interferometer, leading to a measured phase error of 0.28 rad, which corresponds to a $\lambda/22$ accuracy. Larger apertures lead to larger errors since the misalignment of the optical elements and the cameras and the latter's imperfections become dominant. Choosing smaller apertures than the optimal one, the accuracy is not influenced until the intensity at the detection plane becomes too low and electronic and quantization noise become dominant. In general, testing a sample containing a subsurface defect, object decorrelation will reduce the accuracy of the measurement at smaller apertures and the accuracy of the measurement will be lower than the one depicted in Fig. 5.25.

5.6 References

1. O. Kafri, "Fundamental limit on accuracy in interferometry", Optics Letters **14**(13), pp. 657-658 (1989).
2. J.E. Greivenkamp, J.H. Brunning, "Phase shifting interferometers", in *Optical Shop Testing*, D. Malacare, ed., (Wiley, New York, 1991), pp. 501-598.
3. K. Creath, "Temporal phase measurement methods", in *Interferogram Analysis: Digital fringe pattern measurement techniques*, D.W. Robinson, G. T. Reid, eds, (IOP Publishing, Bristol, 1993), pp. 94-140.
4. M. Kujawinska, "Spatial phase measurement methods", in *Interferogram Analysis: Digital fringe pattern measurement techniques*, D.W. Robinson, G. T. Reid, eds, (IOP Publishing, Bristol, 1993), pp. 141-193.
5. T. Kreis, "Quantitative evaluation of the interference phase", in *Holographic Interferometry, Principles and Methods*, W. Jüptner, W. Osten, eds, (Akademie Verlag, Berlin, 1996), pp.101-170.
6. M. Kujawinska, D.W. Robinson, "Multichannel phase-stepped holographic interferometry", Applied Optics, **27**(2), pp. 312-320, (1988).
7. R.A. Chipman, "Polarization analysis of optical systems", Optical Engineering **28**(2), pp. 90-99, (1989).
8. J.L. Pezzaniti, R.A. Chipman, "Angular dependence of polarizing beam-splitter cubes", Applied Optics **33**(10), pp. 1916-1929, (1994).
9. Hu, H.Z. "Polarization heterodyne interferometry using a simple rotating analyzer. 1 Theory and error analysis, Applied Optics **22**(13), pp. 2052-2056, (1983).
10. K. Kinnstaetter, A.W. Lohman, J. Schwider, N. Streibl, "Accuracy of phase shifting interferometry", Applied Optics **27**(24), pp. 5082-5089, (1988).
11. K. Hibino, B.F. Oreb, D.I. Farrant, K.G. Larkin, "Phase shifting for nonsinusoidal waveforms with phase-shift errors", JOSA A **12**(4), pp. 761-768, (1995).
12. K. Creath, "Comparison of phase-measurement algorithms", Surface Characterization and Testing, K. Creath, ed., Proc. SPIE Vol. 680, pp. 19-28, (1986).
13. Y. Surrel, "Phase stepping: a new self-calibrating algorithm", Applied Optics **32**(19), pp. 3598-3600, (1993).
14. J. Van Wingerden, H.J. Frankena, C. Smorenburg, "Linear approximation for measurement errors in phase shifting interferometry", Applied Optics **30**(19), pp. 2718-2729, (1991).
15. A.J.P. van Haasteren, *Real-time Phase Stepped Speckle Interferometry*, (Ph.D. thesis, Delft University of Technology, Delft, 1994).
16. G.E. Sommargren, "Up/down frequency shifter for optical heterodyne interferometry", JOSA. **65**, p. 960-961 (1975).
17. R. Jones, C. Wykes, "De-correlation effects in speckle-pattern interferometry 2. Displacement dependent de-correlation and applications to the observation of machine-induced strain", Optica Acta **24**(5), pp. 533-550, (1977).
18. K. Creath, "Phase-shifting speckle interferometry", Applied Optics **24**(18), pp. 3053-3058, (1985).

19. J. Schmit, K. Creath, “Extended averaging technique for derivaton of error-compensating algorithms in phase-shifting interferometry”, Applied Optics **34**(19), pp. 3610-3619, (1995).
20. C. P. Brophy, “Effect of intensity error correlation on the computed phase of phase-shifting interferometry”, JOSA A **7**(4), pp. 537-541, (1990).
21. B. Zhao, Y. Surrel, “Effect of quantization error on the computed phase of phase-shifting measurements”, Applied Optics **36**(10), pp. 2070-2075, (1997).
22. N. Bobroff, “Residal errors in laser interferometry from air turbulence and nonlinearity”, Applied Optics **26**(13), pp. 2676-2682, (1987).
23. M. Born an E. Wolf, *Principles of optics*, (Pergamon Press, Oxford, 1980).
24. S.C. McClain, L.W. Hillman, R.A. Chipman, “Polarization ray tracing in anisotropic optically active media. 1 Algorithms”, JOSA A **10**(11), pp. 2371-2382, (1993).
25. *Handbook of optics, part 2*, M. Bass, ed., (MacGraw-Hill Inc., New York, 1995).
26. M. C. Simon, “Ray tracing formulas for monoaxial optical components”, Applied Optics **22**(2), pp. 354-360, (1983).
27. Q-T. Liang, “Simple ray tracing formulas for uniaxial optical crystals”, Applied Optics **29**(7), pp. 1008-1010, (1990).
28. W,T, Welford, *Aberrations of Optical Systems*, (Hilger, Bristol, UK, 1986), p.12.
29. J.D. Trolinger, Jr., R. A. Chipman, and D.K Wilson, “Polarization ray tracing in birefringent media”, Optical Engineering **30**(4), pp. 461-466, (1991).
30. M. Lehmann,” Phase shifting speckle interferometry with unresolved speckles: A theoretical investigation”, Optics Communications, **128**, p.325-340, (1996).

Chapter 6

Application of the system for defect detection

Many measurements have been performed during this investigation. They can be divided into two categories, the first group aiming at testing and optimizing the performance of the system and the second one consisting of actual measurements to demonstrate the capabilities of the system. This chapter will present a selection from the latter category. The two-camera system is used for contactless inspection of a composite sandwich panel of a helicopter. The measurement procedure used for defect detection is elucidated in Section 6.1. The inspection is performed in a laboratory environment (Section 6.2) as well as under more severe circumstances, namely with a vibrating object (Section 6.3), simulating an industrial environment.

6.1 Shearographic detection of defects

Shearography reveals defects by measuring defect-induced out-of plane deformation anomalies (see also Section 1.2). It measures the deformation of the test object under different loads. Such a load can be induced by changing the air pressure around the object (pressure or vacuum loading), by heating or cooling it (thermal loading) or by applying mechanical vibration (mechanical or vibrational loading). Loading will cause a deviating displacement of surface points in an area with flaws, as compared to a flawless area. In our measurements we used thermal loading. Usually, the front side of the sample was heated using a heat gun for one or two seconds, increasing the surface temperature by a few degrees. Most defects were revealed during the subsequent cooling process instead of during heating. The former is a much slower process (duration in the order of minutes instead of one or two seconds), which makes the identification of defects in the shearography images easier.

The application of a white reflective powder on the surface of the inspected area can improve the scattering characteristics of the surface without generating considerable artifacts in the deformation measurement[1], e.g., by spraying a white powder upon the surface. This powder does not noticeably change the elastic properties of the surface. Applying the powder, the diffuse reflectivity of objects can in general be increased, especially of objects with a low reflectivity at the used wavelength. Moreover, it can overcome local saturation of the system due to possible specular reflections. For the measurements described in this chapter, the

object's surface was treated with such a white diffusely reflecting powder, namely Dubl-Chek developer D-100.

Using the developed interferometer for non-destructive inspection of materials, the following steps have been executed:

1. Record a reference phase distribution before loading the object.
2. Apply the loading technique.
3. Monitor the live phase difference image, representing the surface shape difference (i.e. deformation) between the moments the reference and actual images were recorded, on the video screen.
4. If the loading causes decorrelation or a (too) high fringe density, take a new reference phase distribution.
5. Adjust the filtering algorithm during the measurement if necessary to improve the distinctness of the defects.
6. Store the image when a defect is clearly visible.

Under laboratory conditions, the real time system is not expected to perform much different from sequential (also called temporal) phase-stepped systems. However, if the environmental conditions are hostile (e.g. vibrations, air current), our system will still be able to perform the phase measurement correctly, whereas the measurements with the temporal phase-stepped systems will be deteriorated by (irregular) changes of the phase between successive recordings due to environmental disturbances.
In this chapter, the measurements performed with the real-time system are described and results are discussed. First, measurements under laboratory conditions are described and results obtained with our system will be presented. Next, the application of the real-time shearing interferometer in more hostile environments will be compared with that of a conventional shearing interferometer.

6.2 Defect detection under laboratory conditions

In the framework of one of the projects* of the European Union, in which this research was performed, several specially prepared samples have been tested. Some of them are realistic aircraft structures (e.g. part of a helicopter fuselage and an engine tunnel sample) whereas others were manufactured as test specimen (e.g. a J-spar of a wing and a step wedge). In order to evaluate the performance of our real-time system under laboratory conditions, the results were compared with those from a more conventional temporal phase-stepped shearography system. This system used an Argon ion laser and the measurement set-up was positioned on an optical table. Our system is mounted on a tripod (see Section 3.6) and the sample is fixed in front of it on a normal table. In Fig. 6.1 a schematic drawing of a temporal phase-stepped system is given.

---

* Brite Euram Project Number BE.5879

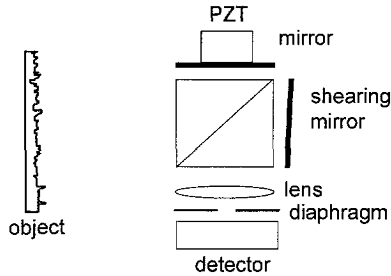


Fig. 6.1 *Temporal phase-stepped shearing speckle interferometer.*

The system depicted in Fig. 6.1 is a Michelson-type shearing interferometer. Phase stepping is introduced by computer controlled translation of the mirror mounted on a piezoelectric element (PZT). The same phase calculation algorithm is used for both the real-time and the conventional system, i.e. a four bucket method. From the comparison of the results it was concluded that the sensitivity of the real-time system under laboratory conditions is similar to that of the conventional system although the measurements with the real-time system have a considerably smaller field of view at the same laser power. Both systems detect the same type and size of the defects. The loading method appeared to be very important. Using thermal loading with a halogen lamp or heat gun usually resulted in the detection of a defect, although the sensitivity for defects, located deep within the material and in thick and solid samples, is lower and a stronger or longer load is needed. The dynamic character of the real-time system proved to be very valuable, especially with samples containing (for shearography) difficult defects (e.g. small defects, defects deep below the surface, etc.). Defects are more easily recognized since temporal changes of the signature of defects can be observed. Furthermore, the ability to observe object deformations in real time allows the operator to verify the correct loading duration and strength.

To illustrate the performance of the system, the results of an inspection of a part of the fuselage of a helicopter are given. It is a sandwich structure consisting of a honeycomb core and thin skin plates, as is depicted in Fig. 6.2. The skin thickness is 0.8 mm and the honeycomb core height is approximately 12.6 mm. The honeycomb core is tapered at the edges.

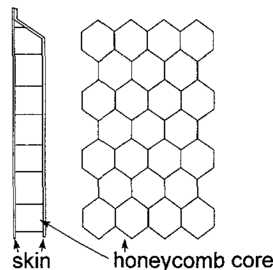


Fig. 6.2 *Sandwich construction with a honeycomb core between two composite skin parts.*

This sample contained 12 defects of two types. Eight impact damages and four artificial inserts between the honeycomb and the skin material were present (the artificial inserts simulating delaminations). In the area we considered here, two defects were present: a circular PTFE insert (with a 40 mm diameter) between the skin material and the core and a 20×20 mm^2 impact damage. The impact crushed the surface material of the sample locally and was classified as a Clearly Visible Impact Damage. The insert could not be seen by the eye. Fig. 6.3 schematically depicts the defects present in the inspected area.

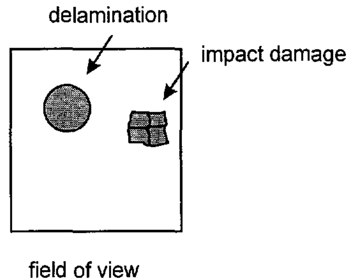


Fig. 6.3 *Defects in the inspected area.*

The settings used in the experiment are listed in Table 6.1.

Table 6.1 *Experimental parameters*

Distance between the measuring system and the object	0.4 m
Field of view	10×15 cm^2
Shearing distance and direction	15 mm horizontal
f/# (aperture size)	22 (4mm)
Output power of the laser	0.9 Watt
Real-time Filter	7×7 uniform

The impact damage is a surface defect, which shows up very easily using thermal loading with a heat gun for less than a tenth of a second. The other defect, the insert, is a sub-surface defect and it also shows up very easily, but with a somewhat longer loading duration (approximately 1 s). In order to measure both defects simultaneously and to enable the storage of the resulting phase image, we used thermal loading with the heat gun during a period of 10 seconds. This way, the heat is built up in the inspection area and it diffused further into the sample. In general, using this loading cycle, the defects could still be measured after several minutes. In the measurement, the result of which is displayed in Fig. 6.4, the reference image was taken several minutes after loading, when most of the heat was already diffused into the material. Roughly 30 seconds after recording the reference phase, the 'live' phase difference image was grabbed and stored onto the hard disk of the workstation. In Fig. 6.4 a gray scale representation of the measured phase difference image is displayed. The perception of the defects is easier when a color representation of the phase difference is used. Therefore, we use a false color

Look Up Table for displaying the phase measurement on a monitor in our system instead of a gray value representation.

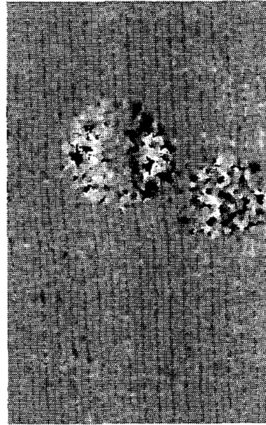


Fig.6.4 *The measured phase difference image, with real-time 7×7 filtering, no postprocessing.*

In the image, two defects can be identified clearly: the one at the right-hand side is the impact damage and the defect at the left of it the 40 mm insert. The form of the insert can be recognized in the measured phase difference. The area affected by the defect-induced deformation has a more or less circular shape, as could be expected from a circular defect. From Fig. 6.4 it can also be seen that approximately the gradient of the surface deformation is measured. The phase image shows a rising flank in the left-hand part of the insert (indicated with a higher average gray value than the background value) and a descending flank (indicated with a lower average gray value than in the background). The sharp transitions from white to black and vice versa of neighboring picture elements is an artefact of the phase calculation algorithm and is not related to sharply localized deformation changes. It is caused by phase wrapping in the $\langle -\pi, \pi]$ interval by the arctangent function.

The second defect, the impact damage, generates a more chaotic phase change pattern than the insert and the (more or less square) shape of the defect is less striking. The surface of the object is broken locally and the phase change can be discontinuous from one picture element to another. As a result, the phase discontinuities can either be actual surface discontinuities or phase wrapping discontinuities.

Using image post processing, the visibility of the defects can be further improved. Since the impact damage leads to real discontinuities in the phase change measurement, only the part of the phase change image containing information on the insert is processed. In Fig. 6.5a, the induced delamination is made more apparent by performing a sine-cosine filtering in a 31×31 neighborhood and stretching the contrast of the image to its full range [0-255].

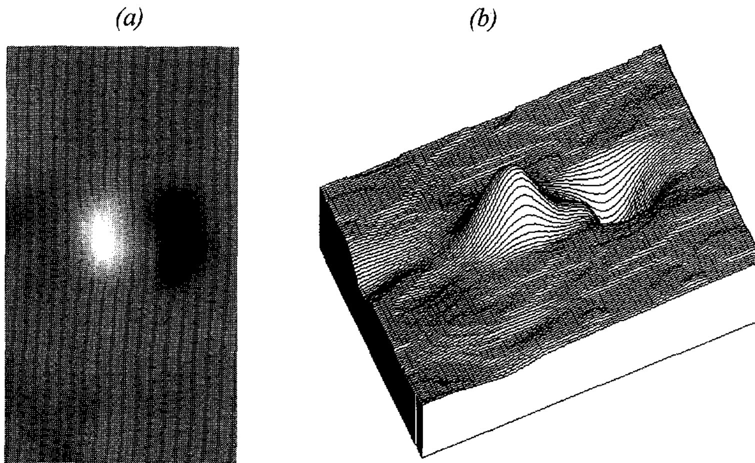


Fig. 6.5 Results of image post processing: (a) after filtering and contrast stretching and (b) the pertinent surface plot.

Fig. 6.5b shows a height profile of the processed image. This picture illustrates the fact that shearography approximately measures the deformation gradient.

6.3 Application of shearography in hostile environments

The goal of these experiments was to determine the applicability of the real-time phase-stepped shearography set-up under industrial environments. For this purpose, the experiments were focused on a comparison of the real-time phase-stepped shearography system with a more conventional temporal phase-stepped shearography system under industrial conditions. Therefore, these two systems were used under (one of) the most common disturbances for shearography, i.e. mechanical vibrations. The temporal phase-stepped system used in this measurement was similar to the one used in Section 6.2 (see Fig. 6.1) but now it used the same laser (SDL-5762-A6) as the real-time system.

To simulate vibrational disturbances, a shaker was used to vibrate the sample under inspection. The amplitude and the frequency of the vibration were varied during the measurements. The amplitude of the object vibration was measured at the rear side of the inspected area using a triangulation set-up. The experiments were conducted using the same test sample as in Section 6.2. The shaker was attached to the rear side of the inspected area containing the defects. As indicated in Fig. 6.6, the position of the sample was fixed to the shaker and two supports where used to suppress in-plane displacements of the sample when vibrating, since out-of-plane displacement was expected to be the dominant displacement in industrial environments.

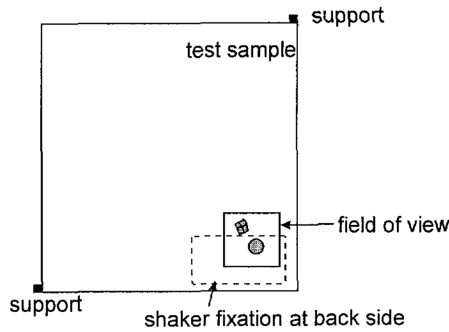


Fig.6.6 *Front view of the inspected sample.*

A schematical drawing of the measurement geometry is given in Fig. 6.7.

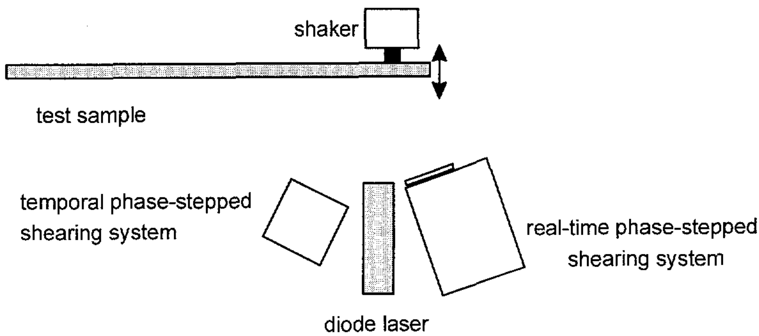


Fig. 6.7 *Top view of the experiment.*

In this set-up, both systems were used to inspect the same part of the object. The inspected part is the same part of the sample as discussed in Section 6.2. The shearing distance was chosen equal for both systems in order to have the same sensitivity. Also, the field of view was approximately equal. The settings of both systems are given in Table 6.2.

Table 6.2 *Experimental parameters of the real-time and the conventional system.*

Experimental parameter	real-time system	conventional system
Distance between the measuring system and the object	40 cm	80 cm
Field of view	10×15 cm^2	11×11 cm^2
Shearing distance	6 mm	6 mm
Shearing direction	$\pi/6$ rad off horizontal	$\pi/6$ rad off horizontal
f/# (aperture size)	11 (=8 mm aperture)	11
Output power of the laser	0.9 Watt	0.9 Watt
Real-time filter	5×5 or 7×7 uniform	none

The measurement procedure for the temporal phase-stepped system is a little different from the one for the real-time system. This is caused by the fact that no direct feedback can be applied on the load. A usual procedure for NDT using the conventional system is the following:

1. Apply the loading technique.
2. Record four phase-stepped images and compute a reference phase distribution.
3. Keep one of the phase-stepped interference patterns in the working memory of the computer.
4. Again record four phase-stepped images.
5. Subtract one of the interferograms of Step 4 from the corresponding one in Step 3.
6. Check if the loading was within the required limits using the intensity difference pattern resulting from Step 5. If not, repeat Step 4 to 6.
7. Compute a phase distribution using the interferograms recorded in Step 4 and calculate the phase difference distribution.
8. Perform image post processing (e.g. filtering and unwrapping) if a defect is visible or suspected.
9. Store the image if a defect is visible

In practice, it turned out to be quite difficult to tell from the intensity difference between two interferograms, recorded at different times, if the applied load was sufficient. Therefore, we used the real-time system to monitor the amount of load. Thanks to the dynamic character of the real-time system the correct stimulation duration could be verified and Steps 3 to 6 could be omitted.
With the conventional system the defects are best revealed during the relatively slow cooling down process instead of during the heating up phase since this system needs several seconds to record four phase-stepped images.
Using the real-time system, defects showed up various times just by changing the vibration frequency of the test sample. This can be understood from a different vibration behavior of a defected area than a defect free area and shows the applicability of vibrational loading for defect detection. However, the defects were much easier detected by thermal loading.

The measurements have been performed such that for each vibration frequency (steps of 1 Hz for the real-time system and steps of 0.1 Hz for the conventional system) the amplitude of vibration was varied in order to determine to what extend defect detection appeared possible. In interpreting the measurements, we assume that, if for some amplitude detection is possible, this also can be performed for lower amplitudes. Likewise, if for a certain frequency and amplitude no defects are detected, it will be assumed that larger amplitudes of vibration also don't reveal the defects. The measurement results can then be expressed in the form of a displacement versus frequency graph in which bands indicate if in a certain frequency-amplitude range, defect detection is possible. This way, the detection limits of both systems are obtained for this specific situation, giving a good indication of the allowed range of frequencies and amplitudes for both systems. Fig. 6.8 shows the results of measurement with the real-time system and those of the conventional system.

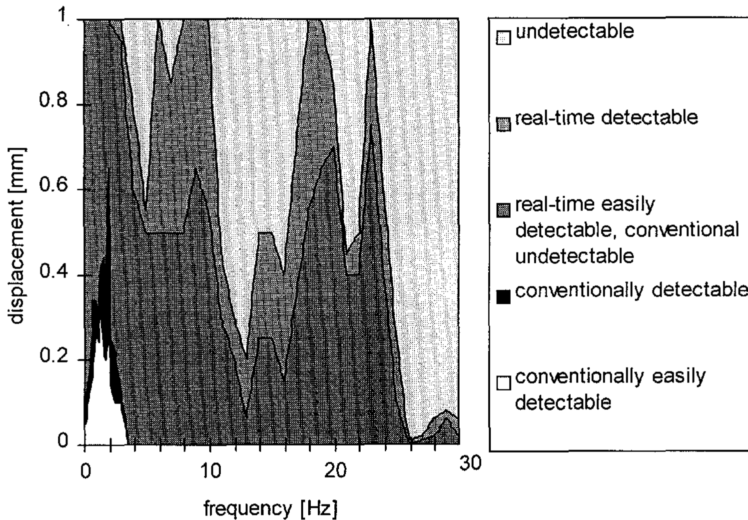


Fig. 6.8 *Defect detection on a vibrating sample with a real-time and a conventional system.*

Comparing the results of both systems, it is clear that the real-time system can perform defect detection for vibrations with higher amplitudes and higher frequencies than the conventional system. As an example, for object displacements of 0.1 mm, the real-time system is able to detect defects in the vibrating object up to ca. 25 Hz, whereas the conventional system is suitable for defect detection up to ca. 3 Hz. Even during vibrations with amplitudes which cause speckle decorrelation (e.g. at the turning points of the vibration) defects are still visible for some time during the period of the vibration.

From Fig. 6.8 a general tendency is observed which indicates that the higher the frequency of the vibration, the lower the allowed amplitude must be to detect the defects with the real-time system. However, from the measurements it is not easy to obtain a direct relation between the amplitude and the frequency of the vibration and the defect detection capabilities of the real-time system. The measurements show that some frequencies of vibrations are worse for the image quality than others, although they have the same amplitude. As an example, an object vibration of about 13 Hz with an amplitude of ca. 0.2 mm obscures the visibility of the defects completely, whereas at a 0.4 mm, 25 Hz vibration, both defects could still be detected. Observation showed that around 13 Hz the sample showed a complicated behavior, which appears to be caused by a resonance of the sample-shaker combination. This resulted in a combined in-plane and out-of-plane movement of the sample, leading to decorrelation between phase patterns recorded before and after loading. This is confirmed by measurement of the in-plane velocity, measured on the tripod on which the shaker is mounted (see Fig. 6.9).

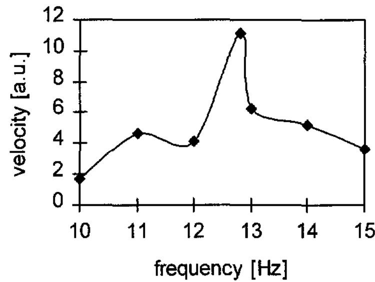


Fig. 6.9 In-plane velocity of the sample.

This measurements shows a strong increase of the in-plane velocity around 13 Hz.
The measurement was repeated for the real-time system using a twice as small aperture (4 mm), leading to lower modulation intensity of the interferograms but also to lower phase errors (see Section 5.5).

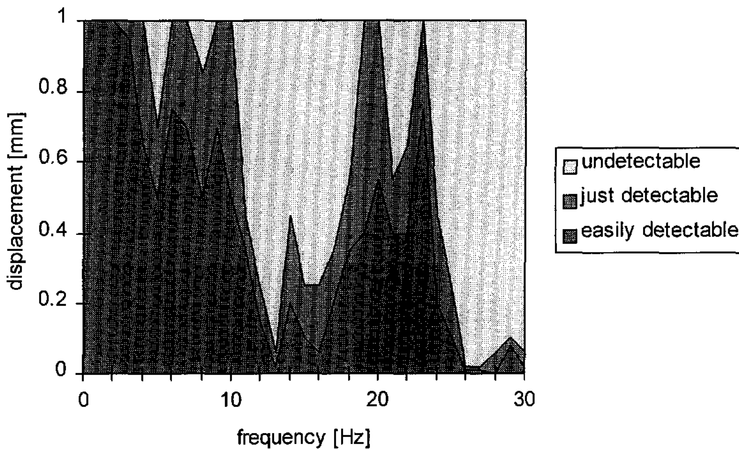


Fig. 6.10 Defect detection on a vibrating sample with the real-time system, using a 4 mm aperture.

This measurement does not show much difference with the one presented in Fig. 6.8, although around 13 Hz, it shows that defect detection is only possible at smaller vibrational amplitudes than when using a larger aperture (see Figs. 6.8 and 6.10). This, too, confirms the assumption that object decorrelation reduces the detectability of the defects around 13 Hz (and also around 26 Hz), since smaller apertures increase the sensitivity for object decorrelation (see Section 2.2).

The measurements with the real-time system were performed during renovation activities in the building, which introduced vibrations on top of those induced by the shaker. Therefore, the experimental conditions of the measurements with the real-time system where more severe than at the time the measurements with the conventional systems were performed. Another

measurement was performed in which both systems were used simultaneously and in absence of the vibrations induced by the reconstruction activities. Moreover, the camera of the conventional system was replaced by a Kappa CF7 camera, i.e. the same camera type as is present in the real-time system. Also, a stronger lens was used in the conventional system (see Fig. 6.1), such that it could be placed at similar distance from the object as the real-time system. This way, the environmental conditions were equal for both systems and the experimental parameters were similar (except for the used phase stepping technique) making the comparison more objective. The experimental parameters of the experiment are listed in Table 6.3.

Table 6.3 *Experimental parameters of the real-time system and the conventional system*

Experimental parameter	real-time system	conventional system
Distance between the measuring system and the object	40 cm	50 cm
Field of view	10×15 cm^2	10×10 cm^2
Shearing distance	10 mm	10 mm
Shearing direction	horizontal	horizontal
f/#	22 (= 4 mm aperture)	16
Output power of the laser	0.9 Watt	0.9 Watt
Real-time Filter size	5×5 or 7×7	none

This measurement showed a similar result as the one presented in Fig. 3.8 and Fig. 3.10, although the local minimum around 13 Hz appeared to be shifted to ca. 14 Hz. It was also observed that at higher frequencies and/or amplitudes, the impact damage was easier to detect than the artificial delamination. For very small amplitudes (<10 μm) neither of the systems are influenced by the vibration, regardless of the frequency (measured up to 200 Hz).

Summarizing the results, which show that the real-time system could detect defects in a composite panel, vibrating at higher frequencies and with larger amplitudes, we can conclude that the sensitivity of the real-time phase-stepped shearography system to object vibrations is significantly lower than that of the conventional phase-stepped shearography system.

6.4 References

1. J.J.J. Dirckx, W.F. Decraemer, "Coating techniques in optical interferometric metrology" Applied Optics **36**, pp. 2776-2782 (1997).

Chapter 7

Conclusions and suggestions

7.1 Conclusions

A system has been developed that measures the deformation of diffusely reflecting object surfaces with video speed, i.e. 25 times per second. It uses two CCD cameras to record four phase-stepped interferograms simultaneously. Phase stepping is performed using the polarization properties of light. The system has a higher measurement sensitivity than real-time speckle shearing interferometry without phase stepping. Thanks to the used shearing technique and the short recording time (40 ms) of a phase image, the system is less sensitive for measurement errors that are introduced by time-dependent perturbations, like air flow, temperature changes or slow vibrations, than sequentially phase-stepped shearing interferometers. The system is successfully used for defect detection in a relative hostile environment and can therefore be used as a supplement to existing non-destructive testing (NDT) techniques.

Savart elements have been used to generate two interferograms on each camera, which have a π radians phase step difference. This way, a compact and rigid mechanical set-up could be realized. The Savart elements are also used in the sub-pixel alignment routines, which are developed to align the interferograms on the pertinent CCD cameras. The stability of the system has proven to be sufficient for application in industrial environments. No adjustment to the system is required during measurements or after transportation.

From the four interferograms, two on each of the two cameras, the object deformation is calculated and displayed on a video monitor using a pipeline image processing system. This is a real-time technique, meaning that the processing speed of the system is high enough to instantly calculate the phase maps. As a result of this ability to observe changes in real time, defects are more easily recognized since temporal changes, carrying the signature of defects, can be observed. Furthermore, an operator can verify the correct stimulation duration and strength while performing the inspection.
In the pipeline image processing system a real-time filtering process was implemented, which simplifies and speeds up the process of finding defects because no (or less) post processing of the images is required to reduce the noise.

The real-time phase-stepped shearing interferometer employs a semiconductor laser as light source and is mounted on a tripod, which makes the system easy to handle. The dimensions and weight of the selected laser (SDL, type 5762-A) are much smaller than that of the formerly used Argon ion lasers. The diode laser can be operated from a standard 220V/50Hz power supply and does not need external cooling. A disadvantage is the limited power of the laser (i.e. 1 W), which restricts the inspection area to approximately 25×35 cm^2. A 1 W laser should normally be sufficient to inspect a 0.25 m^2 area in a single recording, but the modulation intensity in the interference patterns that our system records is rather low due to astigmatism, generated by imaging through the Savart elements. This invokes a 'smearing' of the speckle patterns, thus reducing the modulation depth. As a result, more laser power is needed to inspect a certain object area. The effect of aberrations can be reduced by limiting the aperture of the system, which leads to a lower intensity in the detection plane of the interferometer and thus yields a limited inspection area. Other main error sources in the system are incorrect positions or characteristics of the polarizing optical elements, mainly the quarter wave plates in the interference section, and incorrect camera position. As a result of these errors, a total measurement accuracy of 0.28 rad ($=\lambda/22$) can be obtained, if only the phase errors in the optical system are regarded. Since deformations are measured in reflection, a 22 nm deformation of an object can be measured. For usual defect detection applications, also object decorrelation has to be taken into account, reducing the resulting accuracy of the measurement.

7.2 Suggestions

The shearing interferometer, developed during this thesis, is designed for use in an industrial environment. The interferometer's sensitivity for external disturbances is already strongly reduced as compared to more conventional systems. It also enables the monitoring of the dynamic behavior of objects and transient phenomena. However, an even stronger reduction of the sensitivity for external disturbances can be obtained by using a pulsed laser source. Likewise, to analyze high speed events (e.g. in vibration analysis), pulsed (or modulated) laser sources are required. Thus, the application of a pulsed laser can extend the scope of application of shearography further, e.g. into the field of modal analysis for the study of dynamic behavior of structures, or for sound reduction studies. Certain risks must be assessed, mainly those related to the safety of the laser. Additionally, the type of CCD camera becomes critical. If a frame transfer CCD camera is used together with a pulsed laser, an interferogram cannot be recorded in the even and uneven lines simultaneously. Therefore, only one of the two TV fields contains information. This leads to a reduction of the vertical resolution for a frame transfer camera by a factor of two. Since an interline transfer camera fills a full TV frame in a single short exposure, this type of camera is preferable if used in combination with pulsed laser sources.

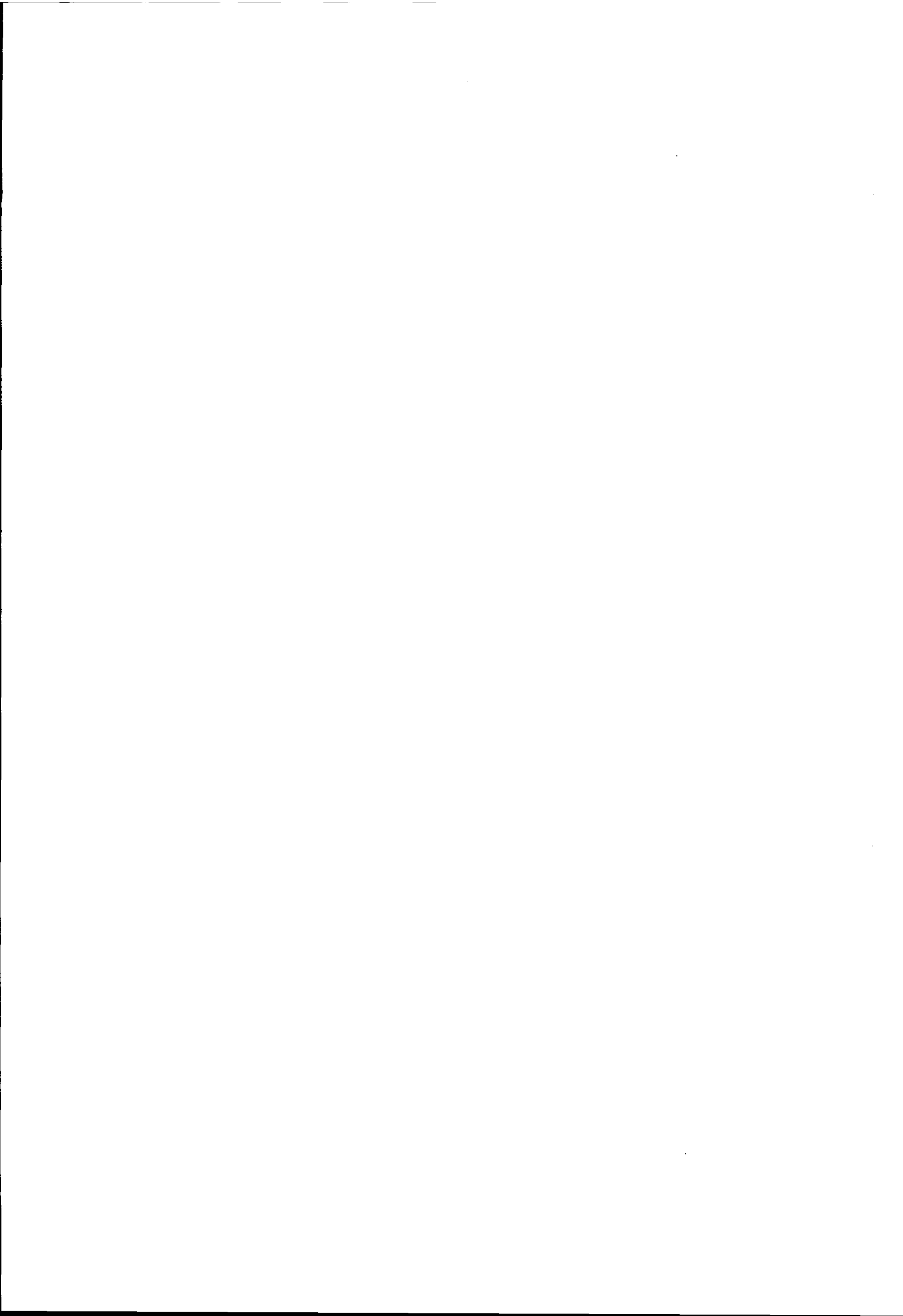
The performance of the developed interferometer on large inspection areas can be increased by decreasing or, even better, avoiding the astigmatism generated in the Savart elements by converging beams. This astigmatism vanishes for collimated beams. Therefore, the Savart

elements should preferably be positioned in a beam where it is collimated. This can be performed by modifying the imaging system, e.g. by removing the second lens of the telescopic system and placing several lenses behind the Savart elements, one for each interferogram. Strong position requirements will be imposed on these lenses. These lenses could also be used for aligning the interferograms on the CCD cameras, which would replace the present alignment procedure which uses rotations of the Savart elements. The maximum numerical aperture is then determined by the actual camera alignment accuracy and the performance of the polarizing elements for converging beams. Also, using a stronger laser (or cameras with higher sensitivities) would enable the use of small apertures in the system, which in turn decreases the errors associated with astigmatism. However, this is less preferable because a small aperture increases object decorrelation, which is a major error contribution in speckle interferometers.

The present system uses a near-infrared diode laser (984 nm). For safety reasons in industrial environments, the application of a laser with a visible wavelength is suggested. An additional advantage of a shorter wavelength is the improved spatial resolution of usual CCD cameras.

To obtain a transportable set-up, the optical elements have been fixed to a base plate. Because of the total weight (ca. 10 kg) and dimensions ($38\times24\times25cm^3$) of the optical head of the system and the semiconductor laser, a heavy duty rigid tripod is required to position the instrument in front of a test structure. The weight and dimensions are primarily determined by the required rigid mechanical construction. A system with a reduced weight and smaller dimensions is desirable. New concepts have been generated to achieve this. By reducing the number of cameras to a single one, recording two phase stepped interferograms simultaneously and adopting temporal phase unwrapping, a more compact and portable system should be feasible. Together with reducing the optical head of the system, reducing the size and costs of the image processing system would improve the acceptance of this new technique into the field of non-destructive testing. Using a PC for real-time image processing is within reach.
Also improved image processing or automated diagnosis would increase the acceptance of shearography as an NDT tool. The presently used procedure to detect defects relies on the interpretation of the images by the operator. Optimal performance is expected when automatic defect detection would be implemented.

Finally, improvement and standardization of the loading methods (e.g. thermal, vibrational or vacuum loading) and systems is required for full acceptance of shearing interferometry as an NDT technique. In order to increase reproducibility, computer controlled stimulation duration and strength is suggested. A loading system should permit extensive adjustments, such that the use of shearography is not restricted by the design of the loading system. The loading system should be easy to handle and is preferably integrated with the system on the tripod. In the presented research, thermal stimulation using a halogen lamp or a heat gun has been used, which are simple to use. But as different structure types and their typical flaws make it necessary to use different loading techniques, intensities or areas, it is necessary to investigate the field of applications and loading methods to find the most promising method and the necessary equipment for a given object.



Symbol list

Symbol	Description	Unit
A_n	electric field amplitude vector of wavelet at a distance of 1 m of its origin	V
A	amplitude of electric field	V m^{-1}
A	aperture area	m^2
ΔA	area of the aperture surface which receives a different part of the speckle pattern,	m^2
B	brightness	W m^{-2} sr^{-1}
C_n	normalization constant of filter routine	-
$C_{R,D}$	constant factor characterizing camera and AD converter	-
C_δ	decay speed constant	-
$\boldsymbol{D}$	electric displacement	C m^{-2}
D_l	diameter of lens pupil	m
$\boldsymbol{E}$	electric field vector	V m^{-1}
F	radiant flux	W m^{-2}
$\boldsymbol{H}$	magnetic field	A m^{-1}
I	intensity	W m^{-2}
I_b	background intensity	W m^{-2}
I_e	intensity noise term	W m^{-2}
I_m	modulation intensity	W m^{-2}
I'_m	changed modulation intensity	W m^{-2}
J_1	first order Bessel function	-
L	cavity length	m
L_c	coherence length	m
$\boldsymbol{M}$	Jones matrix	-
M_{obj}	magnification of objective lens	-
M_{tel}	magnification of telescopic system	-
N	number (usually large)	-
NA	numerical aperture	-
P_L	ptical laser output power	W
Q	brightness reduction factor	-
R	discrimination factor for laser/camera combination	-
R_a	autocorrelation function	W^2 m^{-4}
R_c	crosscorrelation function	W^2 m^{-4}
R	intensity reflection coefficient	-
S_λ	relative sensitivity	-
S	sensitivity of CCD	W m^{-2}

T	intensity transmission coefficient	-
T	recording time of camera	s
V	visibility	-
W	weight function of filter routine	-
a	background intensity containing additive noise	W m^{-2}
$\hat{\boldsymbol{b}}$	unit vector indicating observation direction	-
b	modulation intensity containing multiplicative noise	W m^{-2}
b_1	image distance of first telescope lens	m
c	speed of light	m s^{-1}
c_r	amplitude reflection coefficient of beam splitting cube	-
c_t	amplitude transmission coefficient of beam splitting cube	-
d	beam displacement by Savart elements	m
$\boldsymbol{d}$	displacement vector	m
d_{tel}	length of the telescope	m
d_h	height of picture element	m
d_w	width of picture element	m
d_s	average speckle diameter	m
$\boldsymbol{e}$	sensitivity vector	-
f_o	spatial carrier frequency	m^{-1}
f	focal length of the lens	m
g	integer valued field used for unwrapping	-
$\boldsymbol{k}$	wave vector	m^{-1}
$\boldsymbol{k}'$	scaled wave vector sum	m^{-1}
$\boldsymbol{k}''$	scaled wave vector difference	m^{-1}
l_s	average speckle length	m
l	geometric path length of a ray	m
n_i	refractive index in incident medium	-
n	refractive index	-
n_e	refractive index of extraordinary beam	-
n_o	refractive index of ordinary beam	-
n_ϕ	refractive index of extraordinary beam travelling under angle ϕ with z-axis	-
Δn	birefringence	-
p_a	the area of the picture element	m^2
Δp_a	area of picture element that records a different part of the speckle pattern	m^2
p_ϕ	probability density of phase	-
p_I	probability density of intensity	-
$\boldsymbol{r}$	position vector	m
r	relative displacement between cameras	m
$\hat{\boldsymbol{s}}$	unit vector indicating illumination direction	-
t	amplitude transmission coefficient	-
r	amplitude reflection coefficient	-

t_c	coherence time	s
t	time	s
v_2	object distance of second telescope lens	m
x_s	horizontal spacing of picture elements	m
y_s	vertical spacing of picture elements	m
$\Delta x, \Delta y$	shearing distance	m
z	distance	m
z_o	object distance	m
z_i	image distance	m
m,n	picure element coordinate system	
u,v,w	*uvw*-coordinate system	
x,y,z	*xyz*-coordinate system	
Φ	energy impinging on the object	W
Γ	astigmatic constant	-
Λ	constant	-
Ω_o	solid angle in object space	sr
Ω_i	solid angle in image space	sr
α	angle between optical axis and ray direction	rad
β	retardance of retardance plate	rad
δ	decay variable of filter routine	-
ε	electric permittivity tensor	F m^{-1}
ε	angle of deviation of fast axis from optimal position	rad
ϕ	phase	rad
ϕ'	scaled phase sum	rad
ϕ''	scaled phase difference	rad
ϕ_e	angle between $\hat{\boldsymbol{k}}$ and $\boldsymbol{c}$	rad
ϕ_0	angle between *z*-axis and the principal rays through an object point	rad
ϕ_1	angle between *z*-axis and the principal rays through an image point	rad
φ_o	phase of object beam	rad
φ_r	phase of reference beam	rad
φ	phase difference between object and reference beam	rad
φ_u	unwrapped phase difference	rad
$\Delta\varphi$	phase difference caused by object deformation	rad
γ	angle between Savart entrance face and *w*-axis	rad
η	surface normal	-
λ	wavelength	m
μ	amplitude correlation factor	-

$\Delta\nu_c$	bandwidth	s^{-1}
$\Delta\nu$	frequency spacing	s^{-1}
θ	phase step	rad
θ_i	angle of incidence	rad
θ_0	convergence angle in the object space	rad
θ_1	convergence angle in the intermediate image space	rad
θ_i	convergence angle in image space	rad
$\hat{\rho}$	normalized wave direction	-
σ^2	variance	$W^2\ m^{-4}$
ϖ	angular frequency	rad s^{-1}
ξ	angular mismatch of analyzer	rad
ψ	angle between optical axis and u-axis	rad

Acknowledgments

Many people have contributed to this thesis and I would like to thank them all. Taking the risk of missing someone, I wish to highlight several persons in particular. First of all, I would like to express my gratitude to Professor H. Frankena, my promotor, for giving me his confidence and for the fruitful and pleasant cooperation. I also would like to thank H. van Brug, or Eddy as we know him. He was always willing to listen and discuss problems. Especially during the writing of this thesis, both Eddy and professor Frankena were a great support. Furthermore, I would like to thank J. Braat, P. Bakker, J. Greve, A. Maas and I. Young for their willingness to be a member of the 'promotiecommissie'.

The research described in this thesis research could not have been performed without the financial support of two European Projects: Euclid RTP 3.1 and Brite Euram II. I wish to thank all project partners for the pleasant cooperation. I am also grateful to Arjan van Haasteren for giving me a good start with these projects and for introducing me into the world of speckles. I also owe a great deal to Michel Laurijs, the first TWAIO in our group. His work on the alignment procedures of the system contributes to an essential part of this thesis. Maybe even more memorable than the alignment procedures are the many laughs and great fun we had during so many occasions....

I am grateful to Roland Horsten for his indispensable contribution to the research. He made all the bits and bytes fall in the right place and solved many electronic problems just by bending his magic index finger. I also wish to thank Robert-Jaap van der Bijl and Rob Vink for their contribution to our interferometer. Moreover, the contributions of Inge, Indro, Rob, Bob, Rob, Adri, Ab (of the optical and mechanical workshop and thin film facility) and Aad proved invaluable to the system. Without their competent contribution, the system could not have been made the way it was designed. I thank Professor Braat, Cees and Fokke for their tips and advises during the last few years. I am also very much in debt with professor Braat for letting me use his super-PC. René, thanks for your help on mathematica and your introductory course "preparing your thesis for the printer". Peter Rennspies, thank you for convincing me not to give up teaching gymnastics. And Sehra, thank you for reminding me of so many things. In fact, I can not think of any member or ex-member of our group (Peter, Tammo, Li, Zhu, Maarten, Nigel, Jan-Jaap, to name a few more) who has not contributed in any way. Therefore, I would like to give a general thank-you to all members of the Optics Research Group, for their help and their company. I am sure I will miss the good times during our 'vakgroepsuitjes', TPKV, Wijnhaven, Delftse Hout or the beach. Special thanks go out to my roommates, Gerard,

Luigi and Oliver, who have made my stay in room E014 a very pleasant one. Thanks to them, the chronic lack of space and the sometimes barely endurable heat was hardly ever annoying. Please water our plants at least once a week. Gerard, thank you for being just ahead of me and Oliver and Luigi, I wish you lots of success in the last part of your Ph.D.

From the persons outside the Optics Research Group, I especially want to thank Staas Bruinsma from TPD/TNO-TUD for the pleasant cooperation and his practical way of thinking. I am also grateful to many other people from the TPD/TNO-TUD, in particular Willem van Werkhoven and his colleagues from the mechanical workshop. It took some time before we spoke the same language, but I think we are all proud of the result. Their help and expertise are greatly appreciated and priceless to the system. Also the kind help of Carel van der Poel of Philips Research with diode lasers is acknowledged. In conclusion, I would like to thank everybody who has contributed to this thesis in any way.

Finally, I would like to take this opportunity to thank my family and close friends for their support and for the many good times we had and hopefully will have even more in the future. I am most grateful to Bob Veldt. Thank you Bob, for your assistance, support, patience and, most importantly, your love and your faith in me.

Gon

Samenvatting

Shearing spikkelinterferometer met twee camera's voor vervormingsmetingen met videosnelheid

Introductie

Composietmaterialen worden in toenemende mate gebruikt als lichtgewicht constructiemateriaal. Ze maken de fabricage van veel lichtere constructies mogelijk dan met conventionele materialen (bijvoorbeeld aluminium) mogelijk zou zijn, terwijl sterkte en stijfheid van de constructie gelijk blijven. Daarnaast hebben composieten betere eigenschappen, waaronder een betere corrosiebestendigheid en verminderde vermoeiing. Dit laatste is vooral van belang voor de vliegtuigindustrie, omdat het aluminium van vliegtuigen degradeert ten gevolge van de grote belastingvariaties en temperatuurschommelingen tijdens het opstijgen en landen. Ondanks de aantrekkelijke eigenschappen van composieten, ontstaan er ook nieuwe problemen bij het toepassen ervan. Tijdens productie of gebruik van deze materialen kunnen defecten, zoals scheuren, inhomogeniteiten of laagloslatingen ontstaan en groeien. Het is belangrijk dat deze defecten in een vroeg stadium gedetecteerd worden, nog voordat ze van invloed worden op de prestatie van de structuur.

Aangezien de luchtvaartindustie als één van de eersten composieten toepaste, heeft ze nu te maken met het probleem van verouderde composietdelen. Hierdoor moet de luchtvloot om veiligheidsredenen regelmatig geïnspecteerd worden. De sleutel tot een aanhoudend en groeiend gebruik van composieten in de (vliegtuig)industrie is het vermogen om oppervlaktedefecten en inwendige defecten in constructies te detecteren. Hiervoor zijn de op dit moment beschikbare inspectiemethoden helaas niet toereikend. Derhalve is er een groeiende interesse voor nieuwe methoden die onder industriële omstandigheden mogelijke defecten snel en betrouwbaar kunnen detecteren.

Dit proefschrift

De 25 Hz fasegestapte shearing interferometer die in dit proefschrift wordt beschreven, kan bijdragen tot een veilig gebruik van de luchtvloot. Daarom is de ontwikkeling van dit systeem primair gericht op gebruik in de vliegtuigindustrie. Het nieuwe shearografiesysteem is geschikt voor non-destructieve inspecties van vliegtuigstructuren. Het is veel minder gevoelig voor externe verstoringen dan meer conventionele interferometers, waardoor het geschikter is voor gebruik in een hangar. Het systeem is reeds succesvol toegepast voor defectdetectie in een relatief vijandige omgeving en kan dus gebruikt worden als aanvulling op de bestaande non-destructieve testmethoden (NDT).

Het ontwikkelde systeem meet de vervorming van diffuus reflecterende voorwerpen met videosnelheid, oftewel 25 keer per seconde. Het gebruikt twee CCD camera's om vier interferogrammen gelijktijdig op te nemen. De vier interferogrammen worden gegenereerd in afzondelijke takken van de interferometer met behulp van speciaal hiervoor ontworpen Savart elementen. In elk van de takken worden twee interferogrammen naast elkaar opgenomen door één camera. Op deze wijze is een compacte en stijve mechanische opbouw gerealiseerd. De Savart elementen worden simultaan gebruikt als bundelsplitsers en als analisator. Daarnaast worden ze ook gebruikt bij de uitlijning van de interferogrammen op de CCD camera's.

Defecten kunnen gedetecteerd worden door krachten op het te inspecteren voorwerp uit te oefenen (bijvoorbeeld door verwarming ervan), zodanig dat defecten te zien zijn in de vorm van een vervormingsafwijking van het oppervlak. Omdat het optische systeem verbonden is met een snel beeldbewerkingssysteem, kunnen faseveranderingen van een voorwerp (die oppervlaktevervormingen representeren) bekeken worden met videosnelheid. Het resulterende faseverschilpatroon beschrijft bij benadering de afgeleide van de veranderingshoek in de richting van de shear, hetgeen overeenkomt met de vervorming.

Dankzij de gebruikte shearing techniek en de korte opnametijd (40 ms) van de vier interferogrammen, waaruit het fasepatroon wordt berekend, is het systeem minder gevoelig voor tijdsafhankelijke verstoringen (zoals luchtstromingen, temperatuurvariaties of trillingen) dan fasegestapte interferometers die gebaseerd zijn op het achtereenvolgens opnemen van de fasegestapte interferogrammen. Een ander voordeel van de shearing techniek betreft de eisen aan de coherentielengte van de laser. In vergelijking met conventionele interferometers, zijn de eisen aan de coherentielengte van de lichtbron minder hoog, waardoor halfgeleider lasers gebruikt kunnen worden. Het ontwikkelde systeem gebruikt zo'n halfgeleider laser, zodat het systeem makkelijker te hanteren en te vervoeren is. Het vermogen van de gebruikte laser is beperkt tot 1 Watt, waardoor het inspectiegebied beperkt is tot ongeveer 25×35 cm^2. Een 1 Watt laser zou normaal gesproken een gebied van 0.25 m^2 in één keer moeten kunnen inspecteren, maar door aberraties, ontstaan door het afbeelden door de Savart elementen, wordt het inspectiegebied verkleind. Andere foutbijdragen in het systeem zijn decorrelatie en fouten in de positionering van de polariserende optische elementen en fouten in de eigenschappen ervan. Deze fouten resulteren in een totale meetnauwkeurigheid van maximaal $\lambda/22$, wat leidt tot een kleinst meetbare vervorming van 22 nm.

Delft, A.L. Weijers

Biography

Gon Weijers was born on October 11, 1967 in Bergh, the Netherlands. She finished her secondary education (VWO) at the Gemeentelijke Scholen Gemeenschap Doetinchem in 1986. In that same year she started her Master of Science studies in Applied Physics at the University of Twente. She received her propadeutic certificate in 1987. Shortly after a practical training at the Research Center of ABB in Switzerland (1990) she started her masters project in the Applied Optics Group of the Faculty of Applied Phyiscs. The subject of this research was laser Doppler selfmixing velocimetry in blood flows. She graduated in 1992.
After her graduation she joined her partner on a trip around the world. In 1993 she started working towards her Ph.D. at the Optics Group of the Delft University of Technology under supervision of Professor H.J. Frankena. Her project concerned the development of a real-time phase-stepped shearing speckle interferometer for defect detection in composite materials.