# The Development and Application of Optically Generated Spatial Carrier Fringes for the Quantitative Measurement of Flowfields and Solid Surfaces

*Ping Hai Chan*
Department of Engineering,
University of Warwick

This thesis is submitted for the degree of
Doctor of Philosophy

March 11, 1996

## Declaration

This thesis is presented in accordance with the regulations for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any degree. The work described in this thesis has been done by myself except where references are made to the work of others.

Ping Hai Chan

## Acknowledgements

I would like to give my special thanks to Dr. P. J. Bryanston-Cross, my supervisor who has made this research possible, for his precious guidance and constant support. I wish to express my sincere gratitude to Prof. J. O. Flower for his time and effort to solve my personal problem during my research.

I would like to acknowledge the help and support of all members of our research team, especially Dr. T. R. Judge and Mr. Z. Zhang.

I would also like to think those who aided in proof reading this thesis; especially Mr. Geoff Davis. Finally, my special thanks go to my family and my friends, Miss Yong-Ying Xiao and Mr. Sai-Tim Leung. Without whose support and encouragement, I could not have completed my thesis.

Ping Hai Chan

## Abstract

There are two main approaches to fringe pattern analysis: spatial carrier and temporal carrier. The phase-stepping (PS) method plays a prominent part in the temporal carrier approach. The principal technique of the spatial carrier approach is the discrete Fourier transform (DFT) method. In this thesis, both methods are reviewed and illustrated with examples. The problems associated with these methods are discussed. The effect of weighting function and filtering window on the accuracy of the DFT method is investigated. A review of the minimum spanning tree approach to phase unwrapping is presented.

A new fringe evaluation technique has been explored. The technique combines the computational simplicity of the PS method with the single-image analysis capability of the DFT method. This technique is implemented over two steps. Firstly, a fringe pattern with spatial carrier is subdivided into three component images. Then, the phase is calculated by a three phase step algorithm from these images. A computer simulation has been undertaken to demonstrate the theory and to analyse the systematic errors. Applications in the analysis of the transonic flow-fields are presented.

The Fourier Transform Profilometry (FTP) decodes the 3-D shape information from the phase stored in a 2-D image of an object onto which a Ronchi grating is projected. Two different optical geometries used in the FTP have been compared. The phase information can be separated from the image signal by either the phase subtraction method or the spectrum shift method. The result of an experimental comparison between two phase extraction methods is also presented.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

Optical metrology is a technology which covers all measurement techniques using an optical element or aspect. There are several reasons for the use of optical metrology. Firstly, Optical measurements are non-contact. This property is important for the application where a measurement would be upset by touch, or the instrument cannot reach the object. Secondly, highly accurate measurement is possible because of the properties of light. Finally, optical techniques are non-destructive.

Optical metrology has experienced rapid development in recent years. This is due to advances in laser technology coupled with the development of electronic technology. Currently, research into optical metrology is mainly focused on the development of low-cost high-speed analysis systems. There is an interdependence between the design of the system and analysis techniques used in the system. One of the system designer's aims is to simplify and to automate the analysis techniques. The development of the analysis technique requires an understanding of the theoretical aspects and the electronic devices used in the system, as well as the algorithms employed in analysis. This is a multidisciplinary subject, involving Physics, Electronic Engineering and Computer Science.

In this project, several analysis techniques are studied. A new analysis method has been developed. The main area of this study is concerned with the method for measurement of 3-D object profile and fluid flow using spatial carrier.

## 1.1 An Overview of Fringe Pattern Analysis Techniques

Fringe patterns can be produced by a variety of optical systems, for example interference of light and Moire. Reid [1] suggests that the field of fringe analysis is now regarded by many as an independent subject. If Reid's suggestion holds water, the source of the fringe pattern has little influence on the analysis techniques. In early days, fringe analysis was carried out manually. The processing speed of computers has been significantly improved in the past few years. In conjunction with the development of photoelectrical detector technology, this has stimulated researchers to explore automatic fringe analysis techniques.

There are two approaches to the automatic analysis of fringe. The first one is the intensity-based method relying on a single fringe pattern image. Fringe tracking [2, 3, 4] is based on this approach. The fringe tracking method usually relies on the following procedures [5]:

1. Filter the image [6].
2. Either (a) fit curves to the intensity data with a view to interpolating between fringe centres [7] or (b) identify and track the intensity maxima and/or minima with a view to skeletonising the pattern and thereby minimising the amount of data which must subsequently be processed [8].
3. Number the fringe either interactively [9] or automatically [10].
4. Calculate the measurement parameter from the fringe pattern data.

The major problem associated with fringe tracking is that the information on the sign of the retrieved phase is lost. Secondly, it is difficult to identify the fringe orders when the fringe pattern is broken or wandered over to another fringe pattern. Furthermore, fringe tracking also suffers from low data resolution as well as bad data distribution. However, the intensity-based method may be the only applicable automatic technique for some interferograms, for example; the fringe pattern obtained form time-average holographic vibration analysis of which the intensity is described by a Bessel function.

The second method of automatic fringe analysis is to apply various concepts of phase measuring [11, 12]. This approach can be classified into two categories which are the temporal carrier approach and the spatial carrier approach. The phase measuring method gives several advantages over the intensity method. These include reduced sensitivity to the background and contrast variations in the fringe pattern, automatic peak and valley detection and reduced or eliminated operator interaction [13]. The detail of different phase measuring methods will be discussed in later chapters.

## 1.2 Brief Review of Optical Principles

The optical principles are well-established. This section will summarize the theories which are the key elements in optical metrology. The aim of this section is to outline the notation and approach used in this thesis. A more detailed treatment of optical principles will be found in any standard optics textbook- for example references [14] - [16]. The following discussion is derived from "Optics" edited by E. Hecht [15].

### 1.2.1 Harmonic Wave Motion and Light

A wave form for which the profile is a sine curve is known as harmonic wave. This form of wave motion is the foundation to all shapes of wave motion. The harmonic wave can be represented concisely by using complex notation. For example, if $A$ is the amplitude, $k$ is the propagation constant and $\theta$ is the phase constant;

$$U(x,t)|_{t=0} = A\sin(kx+\theta) \tag{1.1}$$

may be expressed as

$$U(x,t)|_{t=0} = \Re[Ae^{i(kx+\theta)}] \tag{1.2}$$

It is usually assumed that the real part is taken so Eq. 1.2 is simply written as

$$U(x,t)|_{t=0} = Ae^{i(kx+\theta)} \tag{1.3}$$

By replacing $x$ by $(x-vt)$, Eq. 1.1 is transformed into a progressive wave

travelling at speed $v$ in the positive x-direction and can be expressed as

$$U(x,t) = A\sin[k(x-vt)+\theta] \tag{1.4}$$

Light behaves like a transverse wave, as a wave on a string. A transverse wave arises when the disturbance is perpendicular to the propagation direction. The transverse wave motion is confined to a spatially fixed plane called the plane of vibration as illustrated in Figure 1.1.

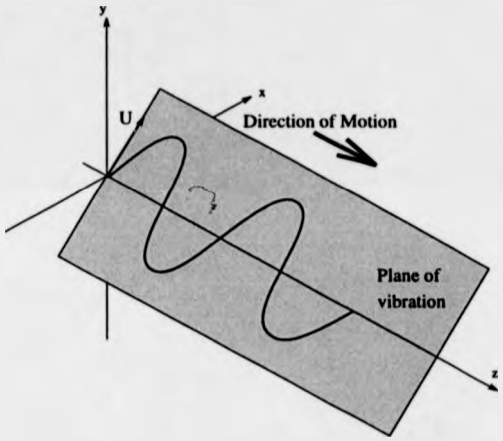



Figure 1.1: A transverse wave travelling in the Z direction

To determine the wave completely, the orientation of the plane of vibration must be specified, as well as the propagation direction. This is equivalent to resolving the disturbance into components along two mutually perpendicular axes, both normal to $z$, see Figure 1.2.

The angle at which the plane of vibration is inclined is a constant, so that at any time $U_x$ and $U_y$ differ from $\bar{\mathbf{U}}$ by a multiplicative constant. The transverse wave function behaves like a vector quantity. With the wave travelling along the $z$-axis, the transverse wave can be written as

$$\bar{\mathbf{U}}(z,t) = U_x(z,t)\bar{\mathbf{i}} + U_y(z,t)\bar{\mathbf{j}} \tag{1.5}$$

where, $\bar{\mathbf{i}}$, $\bar{\mathbf{j}}$ are the unit base vectors in Cartesian coordinates.

An appreciation of the vectorial nature of light is important in certain

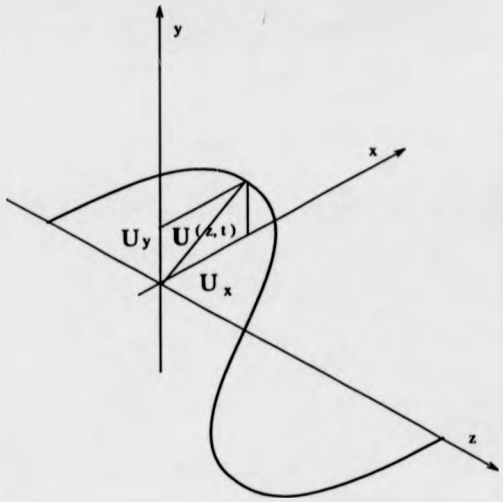



Figure 1.2: The wave resolved into two mutually perpendicular components

applications, such as the polariscope. However, there are many instances in which it is not necessary to be concerned with the vector nature of light. For example, if all the lightwaves propagate along the same line and share a common constant plane of vibration, they may each be described in terms of one electric-field component. This leads to a simple and very useful scalar theory. A scalar harmonic plane wave, moving in the z direction, is expressed as

$$U(z,t) = Ae^{i(kz \pm \omega t)} \tag{1.6}$$

where $\omega$ is the angular frequency.

## 1.3 Interference of Light

The principle of superposition of waves states that if two or more wavefronts are travelling past a given point, the total amplitude of the displacement at that point is given by the vector sum of the individual constituent disturbances. It is the basis of the light interference theory because optical interference may be thought of as an interaction of two or more lightwaves.

Given two scalar light waves described by

$$U_1 = A_1 \sin(\omega t + \alpha_1) \tag{1.7}$$

and

$$U_2 = A_2 \sin(\omega t + \alpha_2) \tag{1.8}$$

overlapping in space, the resultant disturbance is the linear superposition of the waves. Therefore

$$U = U_1 + U_2$$

The intensity, which is the observable quantity, becomes

$$\begin{aligned} I &= |U|^2 \\ &= |U_1 + U_2|^2 \\ &= A_1^2 + A_2^2 + 2A_1A_2\cos(\alpha_1 - \alpha_2) \\ &= I_1 + I_2 + 2\sqrt{I_1 I_2}\cos\delta \end{aligned} \tag{1.9}$$

where $\delta$, which is equal to $(\alpha_1 - \alpha_2)$, is the phase difference arising from a combined path-length and initial phase-angle difference. The $2\sqrt{I_1 I_2}\cos\delta$ is known as the interference term. At various points in space, the resultant intensity varies and depends on the value of the interference term. When

$$\delta = (2n+1)\pi, \quad n = 0, 1, 2, \cdots$$

$\cos\delta = -1$ and the resultant intensity reaches its minima. The two waves are $180^0$ out of phase which is referred to as total destructive interference. When

$$\delta = 2n\pi, \quad n = 0, 1, 2, \cdots$$

$\cos\delta = 1$ and the resultant intensity reaches its maxima. The two waves are in phase which is referred to as total constructive interference. For two waves of equal intensity, $I_0 = I_1 = I_2$, Equation 1.9 becomes

$$I = 4I_0\cos^2\left(\frac{\delta}{2}\right) \tag{1.10}$$

so that the resultant intensity varies between 0 and $4I_0$.

In developing Eq. 1.9, It is assumed that the phase difference, $\delta$, is constant in time. This means that $U_1$ and $U_2$ are assumed to have the same fixed frequency. Ordinary light sources produce light that is a mix of photon wavetrains. At each illuminated point in space there is a net field that oscillates, for less than 10ns or so, before it changes phase. As a result, the interference pattern produced by two ordinary light sources will randomly shift around in space at an exceedingly rapid rate, averaging out and making it quite impractical to observe.

The interval over which the lightwave resembles a sinusoid is a measure of what is called its temporal coherence. The spatial extent over which the light wave oscillates in a regular, predictable way is called the coherence length. To produce a stable interference pattern, the light sources must be coherent and have the same frequency. Until the advent of the laser it was a working principle that no individual sources could ever produce an observable interference pattern.

## 1.4 Polarization of Light

In the previous section, light is treated as a scalar quantity. However, light is a vector quantity which is perpendicular to the direction of propagation and with a defined orientation in space. This property is known as the polarization of light. When two harmonic light waves of the same frequency wave through the same region of space in the same direction, if their respective planes of vibration are mutually perpendicular, the orientation of the resultant light wave may or may not be constant. The exact form of the resultant light wave, i.e. its state of polarization, will be considered in this chapter. Given two orthogonal light waves described by

$$\vec{\mathbf{U}}_{\mathbf{x}}(z,t) = \vec{\mathbf{i}} A_x e^{i(kz-\omega t+\alpha_x)} \tag{1.11}$$

and

$$\vec{\mathbf{U}}_{\mathbf{y}}(z,t) = \vec{\mathbf{j}} A_y e^{i(kz-\omega t+\alpha_y)} \tag{1.12}$$

The resultant light wave is the vector sum of these two orthogonal light waves:

$$\begin{aligned}\bar{\mathbf{U}}(z,t) &= \bar{\mathbf{U}}_x(z,t) + \bar{\mathbf{U}}_y(z,t) \\ &= \bar{\mathbf{i}}A_x e^{i(kz-\omega t)}e^{\alpha_x} + \bar{\mathbf{j}}A_y e^{i(kz-\omega t)}e^{\alpha_y} \\ &= (\bar{\mathbf{i}}A_x e^{-i\frac{\delta}{2}} + \bar{\mathbf{j}}A_y e^{i\frac{\delta}{2}})e^{i\frac{\beta}{2}}e^{i(kz-\omega t)} \qquad (1.13)\end{aligned}$$

where

$$\beta = \alpha_x + \alpha_y$$

$$\delta = \alpha_y - \alpha_x$$

In Eq. 1.13, the factor $e^{i(kz-\omega t)}$ only gives the direction of propagation. $e^{i\frac{\beta}{2}}$ is a common phase factor and a constant. Therefore, the orientation of $\bar{\mathbf{U}}(z,t)$ is completely determined by three independent quantities, i.e. $A_x$, $A_y$ and $\delta$.

In the general case, when $\bar{\mathbf{U}}(z,t)$ passes a plane of observation perpendicular to z-axis, the tip of $\bar{\mathbf{U}}(z,t)$ will trace out an ellipse in that plane. A general state of polarization is therefore called elliptically polarized light. Hecht [15] gives a mathematical derivation of the general state of polarization. There are two cases of particular of interest. When $A_x = A_y$ and $\delta = \pm\frac{\pi}{2} + 2m\pi$, where $m = 0, \pm1, \pm2, \ldots$, the ellipse degenerates to a circle. This special case represents circular polarized light. If $\delta$ is zero or an integral multiple of $\pm2\pi$, the ellipse becomes a straight line. In this case, the resultant light wave is called plane (or linear) polarized light.

# Bibliography


[1] G. T. Reid, "Image Processing Techniques for Fringe Pattern Analysis", Proceedings of the 1st International Workshop on Automatic Processing of Fringe Patterns, Berlin, P. 13-20, 1989.

[2] B. L. Button, J. Cutts, B. N. Dobbins, C. J. Moxon, and C. Wykes, "The Identification of Fringe Positions in Speckle Patterns," Optics And Laser Technology, vol. 17, p. 189-192, 1985.

[3] W. R. J. Funnell, "Image Processing Applied to the Interactive Analysis of Interferometric Fringes," Applied Optics, vol. 20, p. 3245-3249, 1981.

[4] S. Nakadate and H. Saito, "Fringe Scanning Speckle-pattern Interferometry", Applied Optics, vol. 24(14), p. 2172-2180, 1985.

[5] G. T. Reid, "Automatic Fringe Pattern Analysis: A Review," Optics and Lasers in Engineering, vol. 7, p. 37-68, 1986/7.

[6] D. A. Chambless and J. A. Broadway, "Digital Filtering of Speckle Photography Data," Experimental Mechanics, vol. 19, p. 286-289, 1979.

[7] J. B. Schemm and C. M. Vest, "Fringe Pattern Recognition and Interpolation Using Non-Linear Regression Analysis," Applied Optics, vol. 22, p. 2850-2853, 1983.

[8] T. Yatagai, S. Nakadate, M. Idesawa, and H. Saito, "Automatic Fringe Analysis Using Digital Image Processing Techniques," Optical Engineering, vol. 21, p. 432-435, 1982.

[9] T. Yatagai, M. Idesawa, Y. Yamaashi, and M. Suzuki, "Interactive Fringe Analysis System: Applications to Moire Contourgram and Interferogram," Optical Engineering, vol. 21, p. 901-906, 1982.

[10] H. E. Cline, W. E. Lorensen, and A. S. Holik, "Automatic Moire Contouring," Applied Optics, vol. 23, p. 1454-1459, 1984.

[11] K. Creath, "Phase Measurement Interferometry Techniques", Progress in Optics, p. 350-393, 1988.

[12] J. Schwider, "Advanced Evaluation Techniques in Interferometry", Progress in Optics, p. 272-359, 1990.

[13] R. Jozwicki, "New Contra Old Wavefront Measurement Concepts for Interferometric Optical Testing", Optical Engineering, Vol. 31, p. 422-433, 1992.

[14] M. Born, and E. Wolf, "Principles of Optics", Pergamon Press, 1970.

[15] E. Hecht, "Optics", Addison-Wealey, 1987.

[16] F. T. S. Yu, and I. C. Khoo, "Principles of Optical Engineering", John Wiley & Sons, 1990.

# Chapter 2

# Fringe Pattern Analysis by Discrete Fourier Transform Techniques

## 2.1 Introduction

The discrete Fourier transform (DFT) technique is classified into the spatial carrier approach category. This analysis method, which was proposed by Takeda et al [1], uses a single fringe image of the form

$$g(x,y) = a(x,y) + b(x,y)\cos[2\pi f_0 x + \phi(x,y)] \tag{2.1}$$

Where $f_0$ is the spatial-carrier frequency, $a(x,y)$ is the background intensity, $b(x,y)$ is the fringe amplitude and $\phi(x,y)$ is the phase to be determined. In the case of interferometry, these carrier fringes may be introduced by a relatively large tilt to the reference mirror. Alternatively, a sine profile grating may be projected onto an object to generate the carrier fringes. A Ronchi grating may also be used to generate the carrier fringes because a Ronchi grating can be considered as a sum of a series of sinusoidal gratings.

The main advantage of the DFT technique is that it requires a single frame image. This is essential for the measurements of dynamic systems. Another advantage of these techniques is that a special phase-shifting device is not needed. On the another hand, the requirement of the detector array are

more stringent in the DFT technique. This is because the spatial distribution of detector sensitivities must be uniform over the array [2].

## 2.2 Discrete Fourier Transform Techniques

### 2.2.1 1-D Discrete Fourier Transform Technique

One-dimensional DFT technique is proposed by Takeda et al [1]. This technique was extended by Macy [3] to handle two-dimensional data. Eq. 2.1 can be rewritten in the following form for convenience of explanation :

$$g(x,y) = a(x,y) + c(x,y)\exp(2\pi i f_0 x) + c^*(x,y)\exp(-2\pi i f_0 x) \qquad (2.2)$$

with

$$c(x,y) = \frac{b(x,y)}{2}\exp[i\phi(x,y)]$$

where * denotes a complex conjugate. Figure 2.1, which shows a computer generated interferogram of the form described by Eq. 2.1, is included here as an example. It could represent an interferogram of a vibratory annular disk. Figure 2.2 shows a computer generated tilt-free interferogram of the same object in the Figure 2.1.

Then $g(x,y)$ is Fourier transformed with respect to x by the use of 1-D fast Fourier transform algorithm. The Fourier transform of Eq. 2.2 with respect to x is given by

$$G(f_x,y) = A(f_x,y) + C(f_x - f_0,y) + C^*(-f_x - f_0,y) \qquad (2.3)$$

where the upper-case letters denote the Fourier spectra and $f_0$ is the spatial carrier frequency in the $x$ direction. Since $a(x,y)$, $b(x,y)$ and $\phi(x,y)$ have frequencies which are much lower than the spatial carrier frequency $f_0$, the Fourier spectra in Eq. 2.3 are separated by the carrier frequency $f_0$. Figure 2.3 shows the spatial frequency spectrum of raster 256 of Figure 2.1. By using a filter function, one of the two side lobes, say $C(f_x - f_0,y)$, is extracted, and then is translated by $f_0$ toward the origin to obtain $C(f_x,y)$. Figure 2.4(a) shows the side lobe extracted from Figure 2.3. The side lobe translated by $f_0$

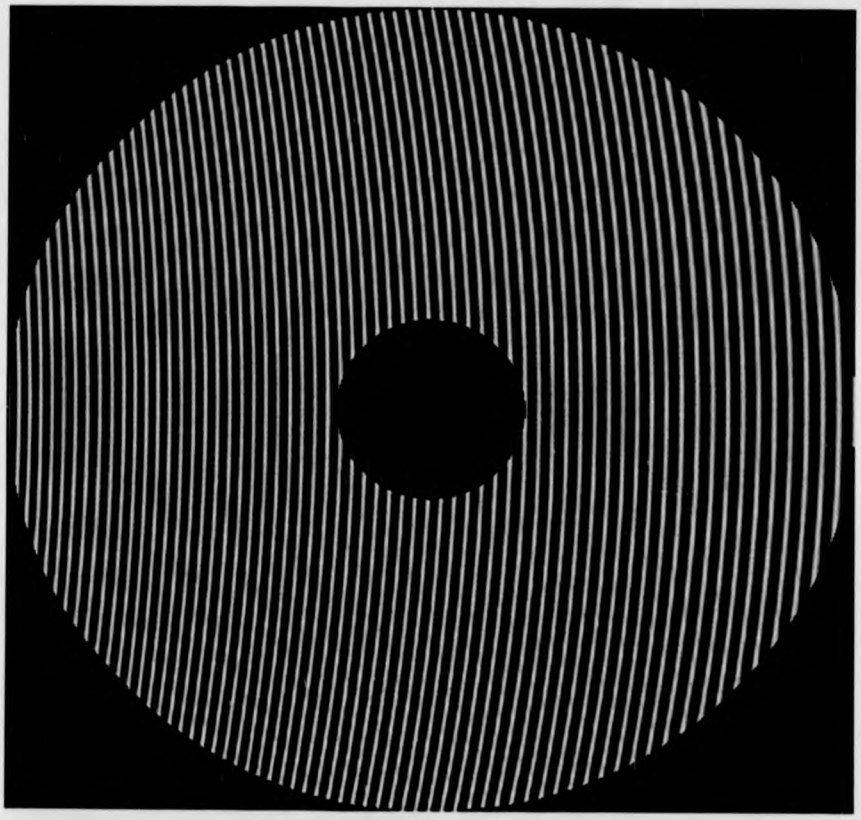


Figure 2.1: The computer generated interferogram of a vibratory annular disk with carrier fringe

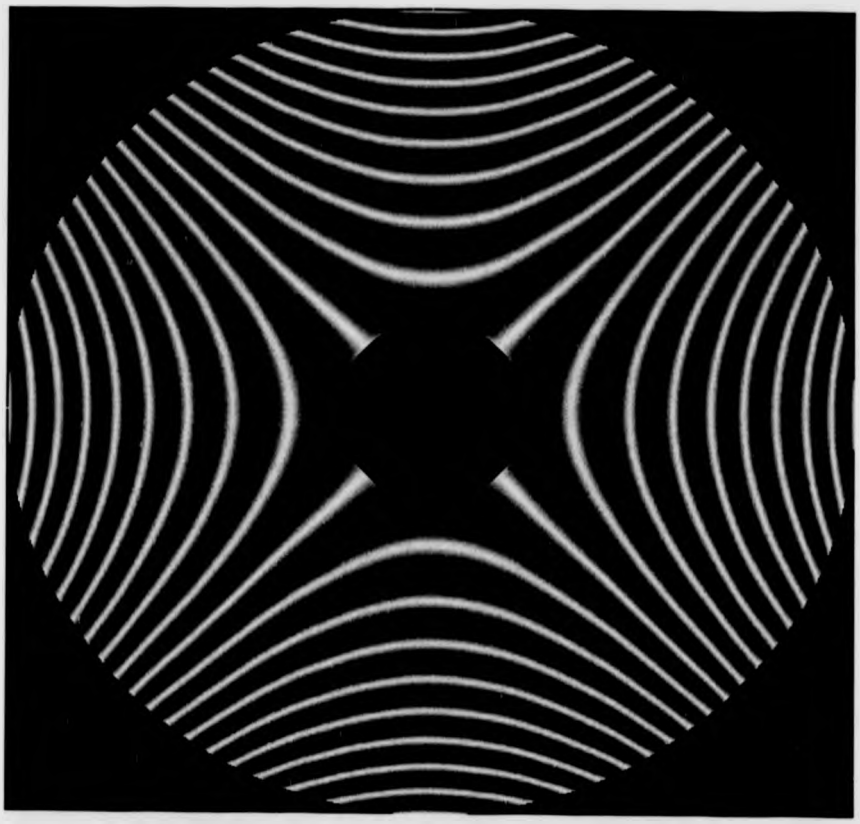


Figure 2.2: The computer generated tilt-free interferogram of a vibratory annular disk

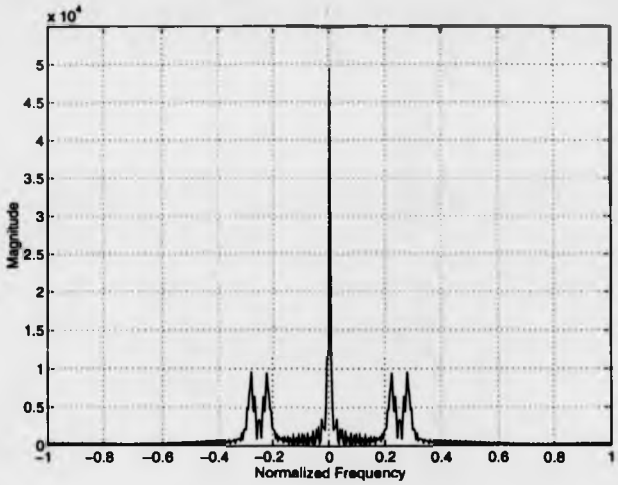



Figure 2.3: A spatial frequency spectrum of raster 256 of the computer generated interferogram with carrier fringe

is shown in Figure 2.4(b). It is noted that the background variation, $A(x,y)$, and another side lobe, $C^*(-f_x - f_0, y)$, have been filtered out in this stage. The inverse Fourier transform of $C(f,y)$ with respect to $f_x$ is computed to obtain $c(x,y)$. The phase may then be calculated from Eq. 2.4

$$\log[c(x,y)] = \log[\frac{b(x,y)}{2}] + i\phi(x,y) \tag{2.4}$$

The phase, $\phi(x,y)$ in the imaginary part completely separated from the amplitude variation, $b(x,y)$, in the real part. The phase can also be obtained from the equivalent operation

$$\phi(x,y) = \arctan\left[\frac{\Im[c(x,y)]}{\Re[c(x,y)]}\right] \tag{2.5}$$

Since the phase calculated from Eq. 2.4 or Eq. 2.5 gives principal values ranging from $-\pi$ to $\pi$, $\phi(x,y)$ is wrapped into this range. Consequently, $\phi(x,y)$ has discontinuities with $2\pi$ phase jumps for variations more than $2\pi$. $\phi(x,y)$ is referred to as a wrapped phase map.

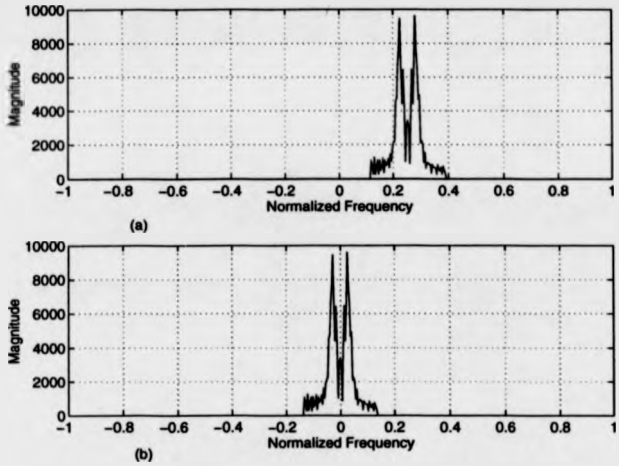



Figure 2.4: (a) The extracted side lobe; (b) The extracted side lobe translated by the carrier frequency to the origin position

### 2.2.2 2-D Discrete Fourier Transform Technique

The 1-D DFT technique requires that the carrier fringes are incremental along the horizontal axis of the frame. If the carrier fringes are not vertical, as shown in Figure 2.5, and the extracted side lobe is translated by a fixed carrier frequency, then the algorithm operates on the the horizontal component of the carrier and leaves a ramp in the solution running from the top to bottom of the frame. The gradient of the ramp corresponds to the vertical component of the carrier fringe pattern. This is because only one component of the heterodyning frequency can be removed by 1-D technique.

Although the 1-D DFT technique was extended by Macy [3] to handle 2-D data, this extension is still essentially the 1-D technique which analyses one slice of data at a time. Bone et al [4] generalized the technique to use 2-D discrete Fourier transform to analyse the fringe pattern. In order to facilitate the discussion, Eq. 2.2 is rewritten in vector form as following

$$g(\bar{\mathbf{r}}) = a(\bar{\mathbf{r}}) + c(\bar{\mathbf{r}})\exp[2\pi i(\bar{\mathbf{v}}_0 \bullet \bar{\mathbf{r}})] + c^*(\bar{\mathbf{r}})\exp[-2\pi i(\bar{\mathbf{v}}_0 \bullet \bar{\mathbf{r}})] \qquad (2.6)$$

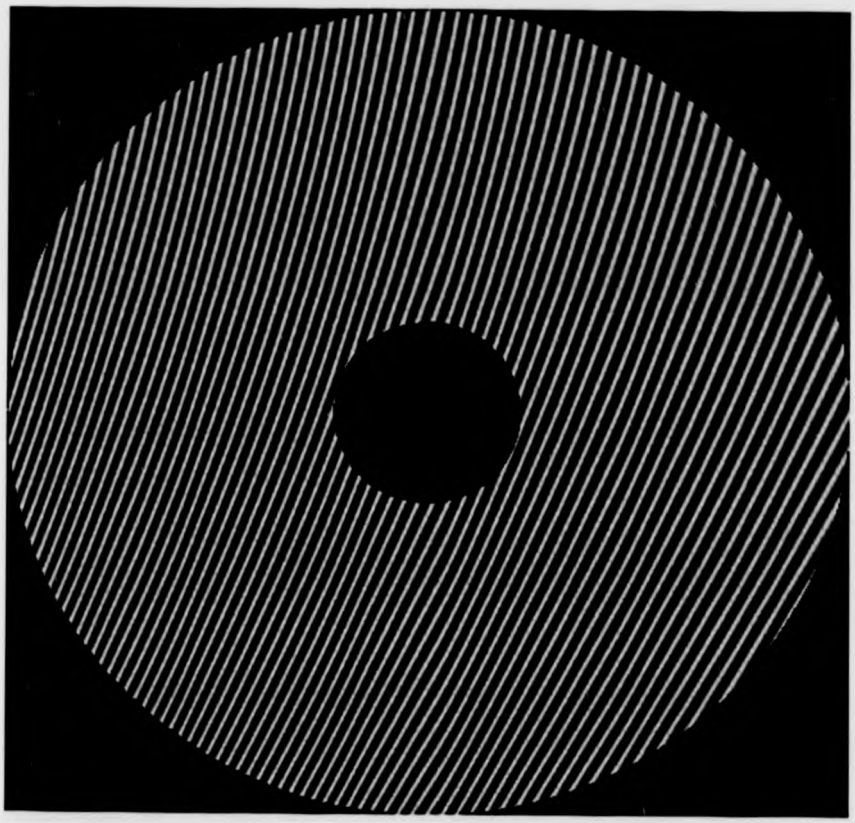


Figure 2.5: The computer generated interferogram of a vibratory annular disk with non-vertical carrier fringe

with

$$c(\bar{\mathbf{r}}) = \frac{b(\bar{\mathbf{r}})}{2} \exp[i\phi(\bar{\mathbf{r}})]$$

In Eq. 2.6, vector $\bar{\mathbf{r}}$ with Cartesian coordinates $(x, y)$ refers to a point on the fringe image and $\bar{\mathbf{v}}_0$, which is equal to $(f_{x0}, f_{y0})$, is the 2-D heterodyning frequency. The 2-D Fourier transform of $g(\bar{\mathbf{r}})$ becomes

$$G(\bar{\mathbf{v}}) = A(\bar{\mathbf{v}}) + C(\bar{\mathbf{v}} - \bar{\mathbf{v}}_0) + C^*(-\bar{\mathbf{v}} - \bar{\mathbf{v}}_0) \tag{2.7}$$

where the upper-case functions are the 2-D Fourier spectra of the lower-case functions. $\bar{\mathbf{v}}$ is the vector frequency with component $(f_x, f_y)$.

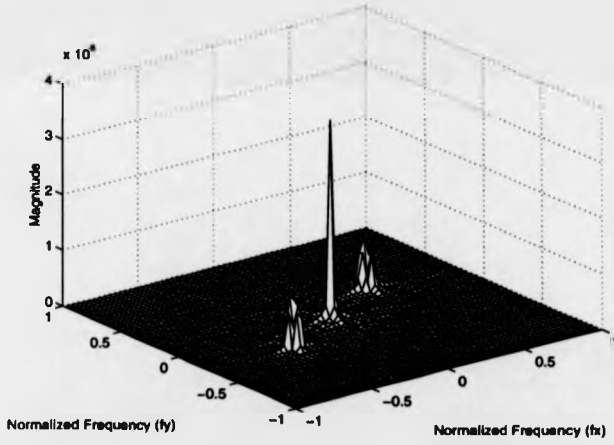



Figure 2.6: The Fourier transform plane of the interferogram with non-vertical carrier fringe

In the Fourier transform plane, all terms in Eq. 2.7 are found as strongly peaked lobes centered at the origin and $\pm\bar{\mathbf{v}}_0$. Figure 2.6 shows the Fourier transform plane of Figure 2.5. Then, one of the two side lobes, say $C(\bar{\mathbf{v}} - \bar{\mathbf{v}}_0)$, is isolated in Fourier plane from the other terms in $G(\bar{\mathbf{v}})$ by using a 2-D filtering window, and then is translated by $\bar{\mathbf{v}}_0$ toward the origin to obtain $C(\bar{\mathbf{v}})$. Figure 2.7 and Figure 2.8 shows the isolated side lobe and the translated side lobe in the Fourier transform plane, respectively. The inverse 2-D Fourier transform of $C(\bar{\mathbf{v}})$ is calculated to obtain $c(\bar{\mathbf{r}})$. The phase may then be computed from either Eq. 2.4 or Eq. 2.5.

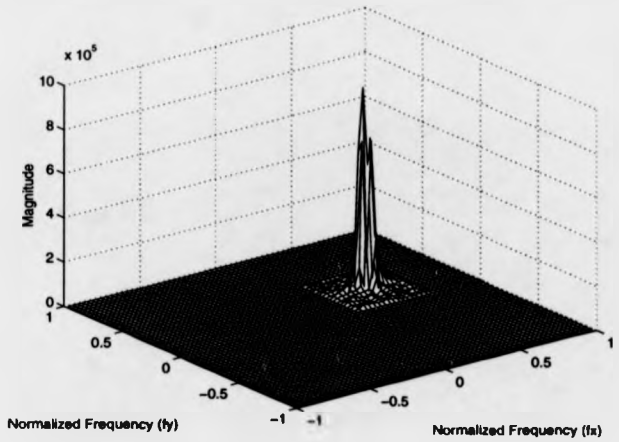



Figure 2.7: The side lobe isolated by a 2-D filter in Fourier transform plane

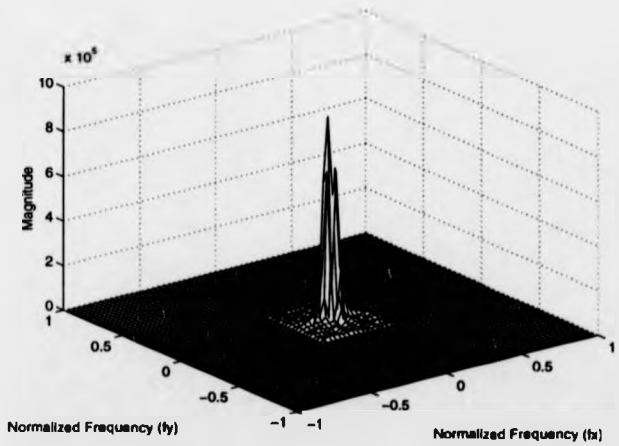



Figure 2.8: The isolated side lobe translated by the 2-D carrier frequency to the origin position

## 2.3 Other Fourier Transform Techniques

The frequency domain is an efficient way to store information in an image. This fact has been exploited in the image compression technique [6]. In a fringe pattern image, as revealed in Figure 2.6, the required information is gathered round the carrier frequency and occupies a relatively small number of pixels. There are mainly two techniques, the multi-channel Fourier fringe analysis and the spatio-temporal frequency multiplex, to utilize the redundant space in the frequency domain so that more information can be collected in a single fringe image.

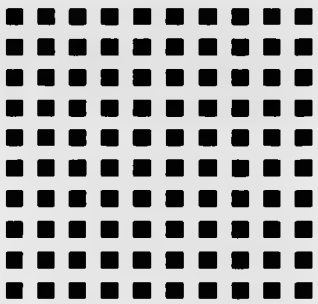


Figure 2.9: The cross grating for simultaneously measuring two directional deformation

Multi-channel Fourier fringe analysis [7, 8, 9] is a technique to process an image with more than one fringe pattern of which the spatial frequency bandwidths must separate completely in the frequency domain. Morimoto et al [7] made use of the multi-channel Fourier fringe analysis to measure simultaneously the x and y directional displacement. A cross grating as shown in Figure 2.9 is printed on the end surface of a specimen. The intensity function of the cross grating can be expressed as the sum of two single grating intensity functions. One grating is normal to the x axis, and another grating is normal to the y axis. The digital image of the deformed grating is transformed into the frequency domain by 2-D Fourier transform. The transform now consists of four peaks due to the fringe and a dc peak. The information of x directional and y directional deformation can then be partitioned by filtering out the appropriate peak before inverse Fourier transform.

Burton et al [9] applied the multi-channel Fourier fringe analysis to the problem of automatic phase unwrapping in the presence of surface discontinuities. The technique is based on the fact that if the fringe moves then all the phase wraps will move with them, but the physical discontinuities introduced by the surface will not move. An image that is the sum of two fringe pattern images on the same surface is analyzed by Fourier transform fringe analysis. This yields two wrapped-phase maps. Both wrapped phase map would show the technique-induced discontinuities as well as the genuine physical discontinuities. The discontinuities that remain stationary in both wrapped phase maps are detected by image processing techniques, and hence should be ignored during phase unwrapping.

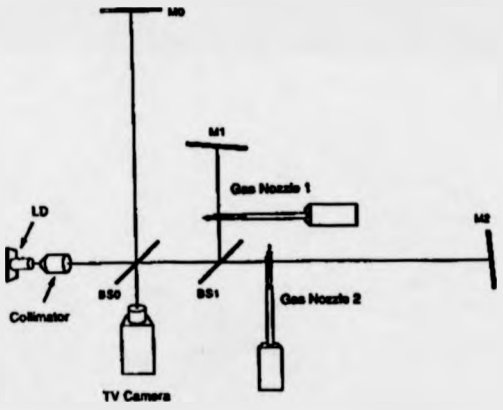



Figure 2.10: Michelson-type two-channel interferometer [10]

Spatio-temporal frequency multiplex [10, 11] uses both spatial and temporal carrier frequencies. This technique enables the measurement of a phase distribution with a wider spatial and temporal bandwidth when it combined with the Fourier fringe analysis method. The experiment reported in Ref. [10] is depicted here to demonstrate the principle of the technique. A two-channel Michelson-type interferometer, which is shown in Figure 2.10 , was used in the experiments. Two gas flow fields with different flow rate were introduced into each of the channels. The spatial carrier frequencies were introduced by tilting the mirrors M1 and M2. A ramp current was applied to a laser

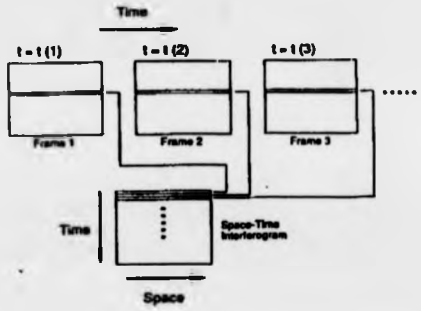



Figure 2.11: Data acquisition process [10]

diode to produce temporal carrier frequencies. Only one set of 1-D fringe data along a fixed scanning line was taken from each frame and piled sequentially in a 2-D memory array to form a 2-D spatio-temporal interferogram. Figure 2.11 explains graphically the process of data acquisition. Figure 2.12 shows a spatio-temporal interferogram with the present of two gas flow fields. The correlative spatio-temporal frequency spectra are shown in Figure 2.13. The spectrum corresponding to the two different gas flow fields are indicated by the arrow A and B. Two phase distributions retrieved from the single spatio-temporal interferogram are presented in Figure 2.14.

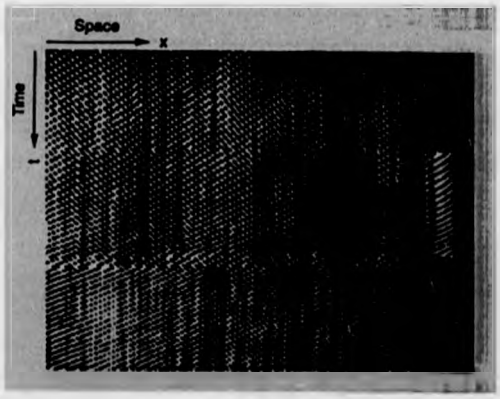



Figure 2.12: The Spatio-temporal interferogram [10]

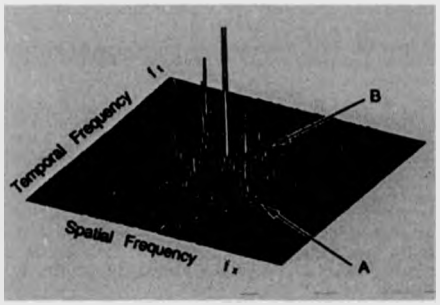



Figure 2.13: Spatio-temporal frequency spectra of Fig. 2.12 [10]

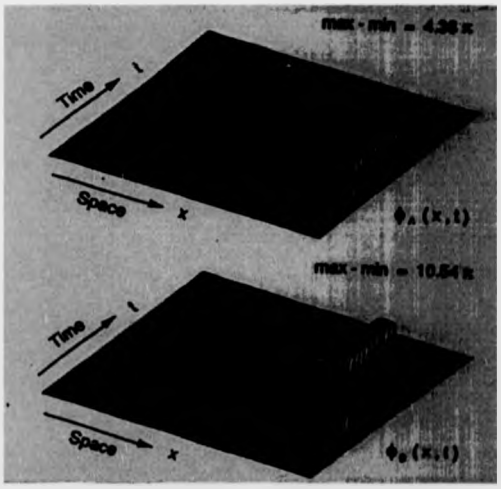



Figure 2.14: The phase distributions of two different gas flow fields [10]

## 2.4 The Problems Associated with Discrete Fourier Transform Techniques

### 2.4.1 Influence of Noise

For a small additive white noise, the rms phase errors in heterodyne techniques can be calculated in term of signal-to-noise ratios (SNR). In the case of the 1-D DFT technique, Takeda [12] gave a corresponding rms phase error for the Fourier transform technique as

$$\sqrt{\langle[\Delta\phi(x,y)]^2\rangle} = \frac{\left(\frac{n}{N}\right)^{\frac{1}{2}}}{S(x,y)} \tag{2.8}$$

$S(x,y)$ is a SNR defined by

$$S(x,y) = \frac{b(x,y)}{\sqrt{2}\sigma}$$

where $n$ is the number of spectral sample points within the filter passband in the frequency domain, $b(x,y)$ is the fringe amplitude and $\sigma$ is the rms value of the noise. For the two-dimensional case, Bone et al [4] gave a similar formula for the expectation value of the rms phase error as

$$E(\Delta\phi_{rms}) = \frac{\left(\frac{n}{2N}\right)^{\frac{1}{2}}\sigma}{\pi\bar{m}} \tag{2.9}$$

where $\bar{m}$ is the mean modulation amplitude.

These formulas can be interpreted as follows. In the Fourier transform technique, the phases is obtained from the $n$ frequency spectra spreading over the filter passband. The spectra passed by the filter include $n/N$ of the total noise power. Therefore, the phase error is proportional to the factor, $n/N$, under the small-noise assumption.

### 2.4.2 Fringe Discontinuities

The fringe discontinuities are an important source of error if they are not properly taken into account. In some applications, such as in the test of an object with a hole, the fringe abruptly disappears at the boundaries.

These discontinuities cause the fringe spectra to spread out. This causes problems not only in the Fourier transform technique, but also in the phase unwrapping.

A way to deal with this problem is to extrapolate the fringes image. Roddier et al [13] developed a systematic method of fringe extrapolation, which is based on Gerchberg's iterative algorithm for analytic continuation [14]. As an example, Figure 2.1 is extrapolated as in Figure 2.15 by adding a carrier fringe into the boundaries.

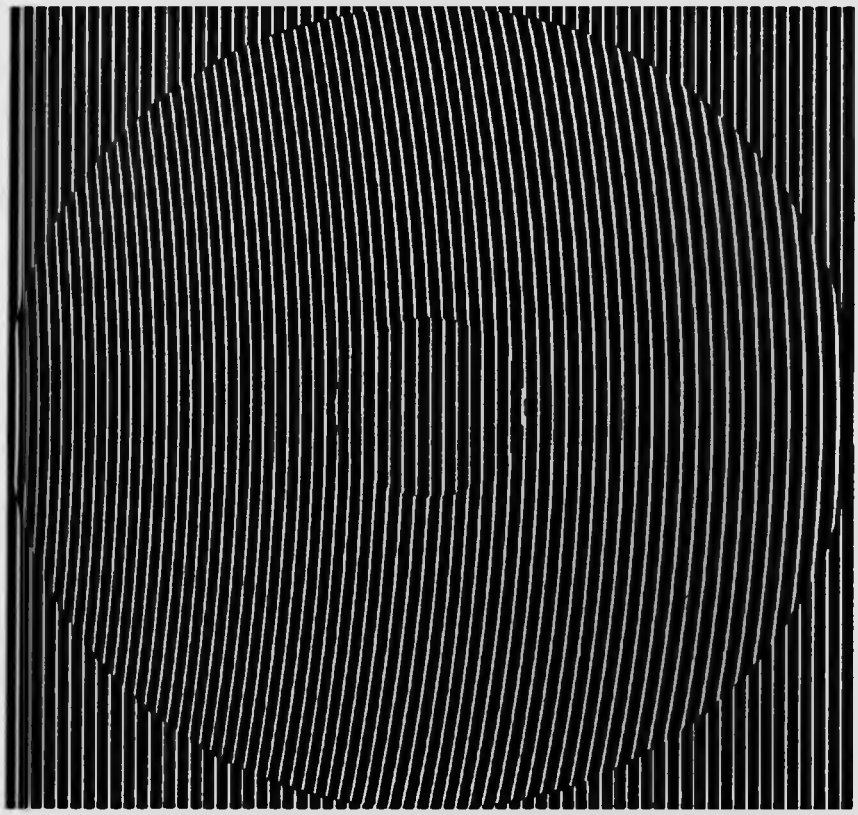


Figure 2.15: The extrapolated interferogram

### 2.4.3 Other Error Sources

When the frequency of detector nonlinearity is approximately equal to the carrier frequency, detector nonlinearity is one of the error sources. The response of CCD array is linear within its operation region for most CCD cameras in the market. Therefore, detector nonlinearity is usually not a problem for measurement using CCD cameras. However, if the image is captured by film, the error caused by non-linear response of the recording medium should be taken into account. The Fourier transform techniques will interpret the fringe distortion which is caused by the film as fringe distortion due to the presence of the physical quantity to be measured. There are many well-established techniques for calibrating film. Nugent [5] proposed a method to determine the film characteristics of interferogram by measuring the deviation of the fringe in object-free areas from pure sinusoids. This method utilizes the fact that fringe recorded using a perfect recording medium will be sinusoidal.

There are some errors associated with an FFT algorithm used in the discrete Fourier transform technique. These include :

1. Aliasing.
   Aliasing will occure, if the sampling frequency is lower then the Nyquist frequency which is twice the highest frequency component of signal to be sampled.

2. Leakage of energy due to inappropriate truncation of data.
   This will be discused in the next two chapter.

3. The picket fence effect.
   The DFT is a mechanism by which the observed singal is decomposed into an orthognal vector space. The sines and cosines with frequency equal to an integer multiple of $1/NT$ form an othogonal basis set, where $T$ is the sampling interval and $NT$ is the observation interval. From the continuum of possible frequencies, only those which coincide with the basis vector will project onto a single basis vector; all other frequencies will exhibit non-zero projections on the entire basis set. This effect is referred to as the picket fence effect. If the carrier frequency includes

a frequency which is not one of the basis frequency, the picket fence effect will occure.

## 2.5 Application Example of Discrete Fourier Transform Techniques

Two examples are presented in this section to demonstrate the capacity of the 1-D and 2-D DFT technique. The first example is to measure the change of air density profile caused by a heat source. In this case, the heat source is a soldering iron. In this experiment, the Twyman-Green interferometer was used because of the availability of the equipment. As the light wave pass through the flow field twice, the result obtained from the Twyman-Green interferometer should be divided by two. Figure 2.16 shows the schematic diagram of experiment set-up. R. Jones [15] had manifested that the visibility of the fringe will decrease, if the ratio of the fringe spacing to the mean speckle size approaches unity. As a result, no fringe is observed when this ratio is approximately equal to unity. The size of the subjective speckle is given by [16]

$$d_{sp} = \frac{2.4\lambda d}{a} \tag{2.10}$$

where

$d_{sp}$ = the subjective speckle size

$\lambda$ = wave length

$d$ = the distance from the lens to the image plane

$a$ = the diameter of the viewing lens aperture

Thus, the viewing lens aperture should be as large as possible. The neutral density filter is used to attenuate the light intensity evenly so that the lens aperture can be turned to its maximum. Figure 2.17 shows the linear fringe pattern modulated by the air density change that is caused by the heat of the soldering iron. The interferogram is analysed by 1-D DFT technique since the carrier fringe is vertical. The Figure 2.18 shows the wrapped phase map after Figure 2.17 is demodulated. The wrapped phase map is unwrapped by the Minimum Spanning Tree approach that will be discussed later. The

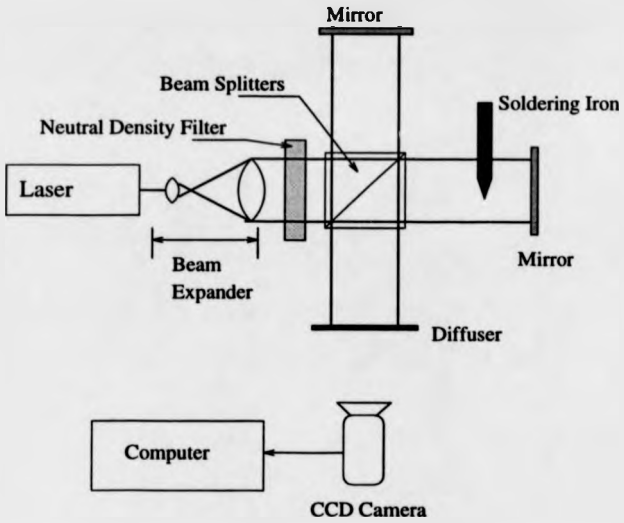



Figure 2.16: The schematic diagram of the experiment set-up

3-D mesh plot of the unwrapped solution is shown in Figure 2.19.

In the second example [19], a shearing interferogram with non-vertical carrier fringe is analysed by 2-D DFT technique. Figure 2.20 shows the experimental set-up to visualise the transonic jet wake flow using a lateral-shear interferometer. The collimated beam of light passes through the flow field and is then incident on a wedged shearing plate. The two reflected beams are sheared laterally by the a shearing plate. The fringe pattern in the region of overlap is formed on the screen by the interference of the wavefront with its sheared duplicate. A spatial carrier fringe pattern is introduced by the wedge angle between the two faces of the shearing plate.

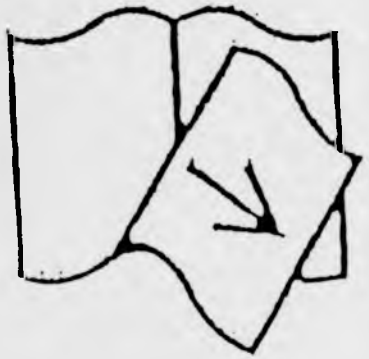
In lateral shearing interferogram with spatial carrier, the interference fringes can be described by

$$g(x,y) = a(x,y) + b(x,y)\cos[2\pi(xf_{x0}, yf_{y0}) + \phi_x(x,y) + \phi^{'}(x,y)] \quad (2.11)$$

where $a(x,y)$ and $b(x,y)$ represent background and modulation terms, respectively, $(f_{x0}, f_{y0})$ is the 2-D spatial carrier frequency introduced by wedged shearing plate, and $\phi^{'}(x,y)$ is the initial phase shift caused by optical geom-

# PAGES MISSING
# NOT AVAILABLE

fig 2.17 Page 29.



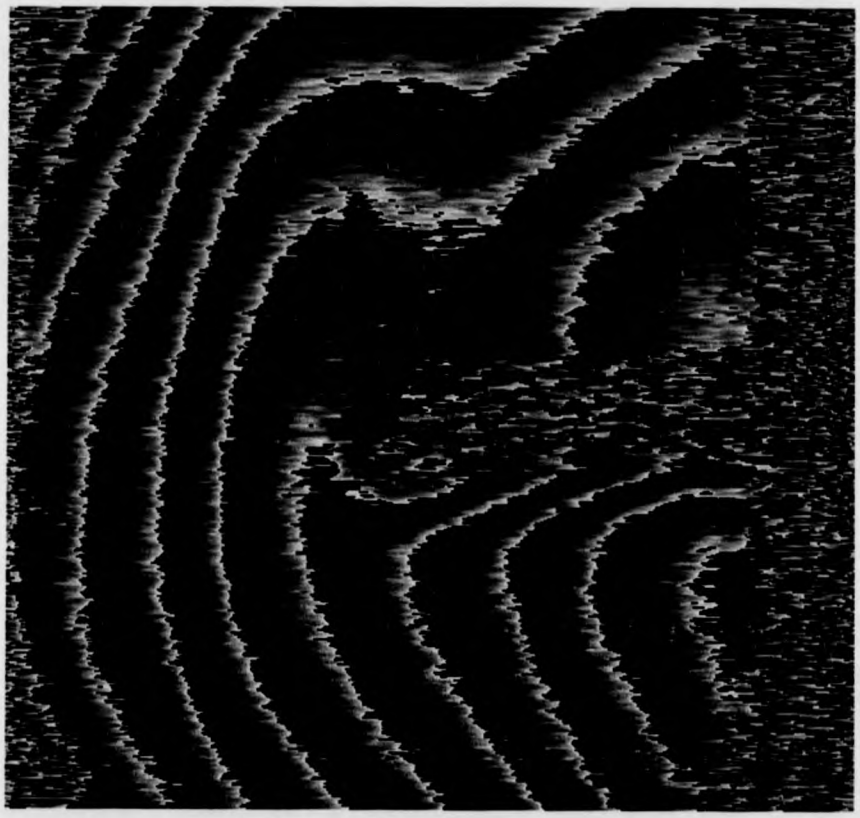


Figure 2.18: The wrapped phase map of the hot soldering iron

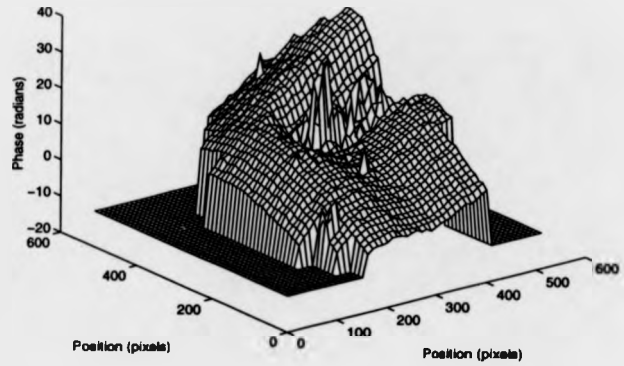



Figure 2.19: The 3-D mesh plot solution of the hot soldering iron

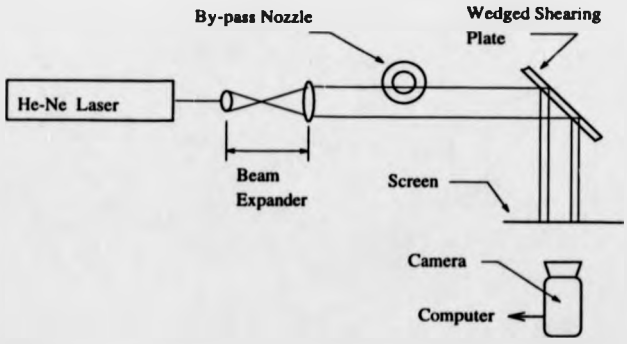



Figure 2.20: The experiment set-up to visualise the tranic jet wake flow

etry and distortion of the lenses etc. As the shear is a small fraction of the diameter of the wavefront, $\phi_x(x,y)$ corresponds to the derivative of the phase change caused by the flow field along the direction of shear [17] Eq. 2.11 can be rewritten in following form :

$$g(x,y) = a(x,y) + c(x,y)exp(2\pi j(xf_{x0}, yf_{y0})) + c^* exp(-2\pi j(xf_{x0}, yf_{y0})) \tag{2.12}$$

where $c(x,y) = (1/2)b(x,y)exp[j(\phi'(x,y)+\phi_x(x,y))]$ and * denotes a complex conjugate.

Figure 2.21 shows the interferogram with the flow field. In order to cancel the initial phase shift, a reference image is also taken prior to the presence of the flow field. Figure 2.21 is transformed into Fourier domain by using 2-D DFT. In the Fourier Transform Plane, which is shown in Figure 2.22, all terms in Eq. 2.12 are found as strongly peaked lobes centred at the origin and the spatial carrier. One of the two side lobes is isolated in Fourier plane from the other terms by using a 2-D filtering window. The same filtering operation was performed on both the sample and the reference images. The phase subtraction method [18] is then employed to retrieve $\phi_x(x,y)$, which is shown in Figure 2.23. It has been found that the retrieved phase is within the $2\pi$ range, and that noise has been added to the image as result of the poor image contrast in the the shadow region of the by-pass nozzle. Figure 2.24 shows the contour plot of the retrieved phase after the noise is filtered out.

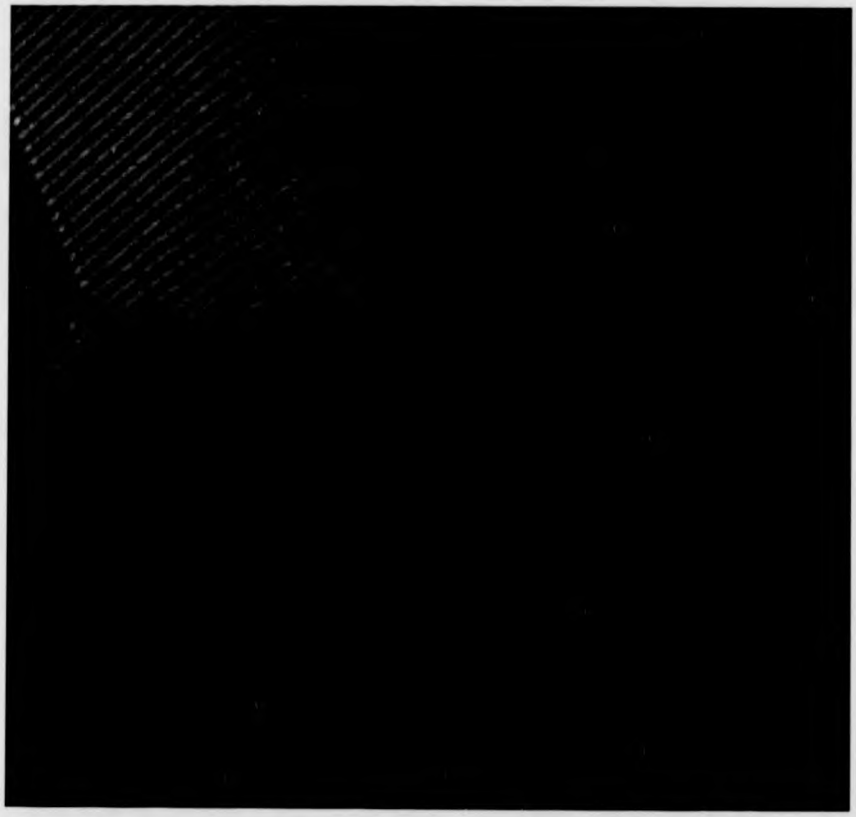


Figure 2.21: The lateral shearing interferogram with flow field

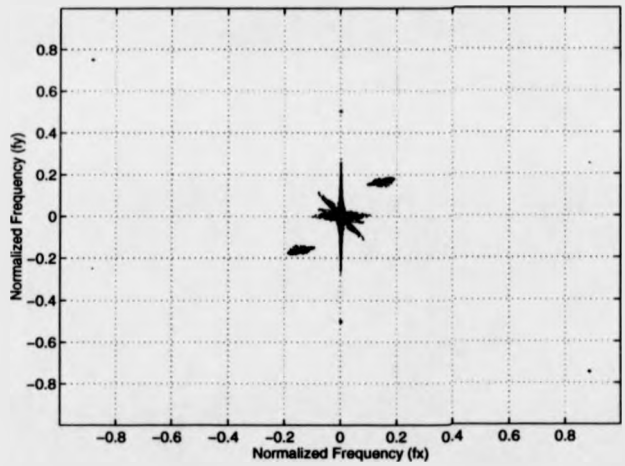



Figure 2.22: The contour plot of 2-D DFT of Figure 2.21

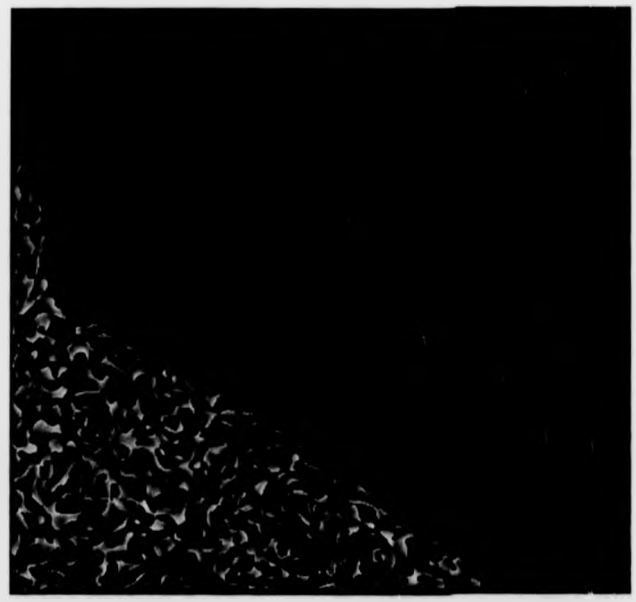


Figure 2.23: The phase map of the transonic flow field

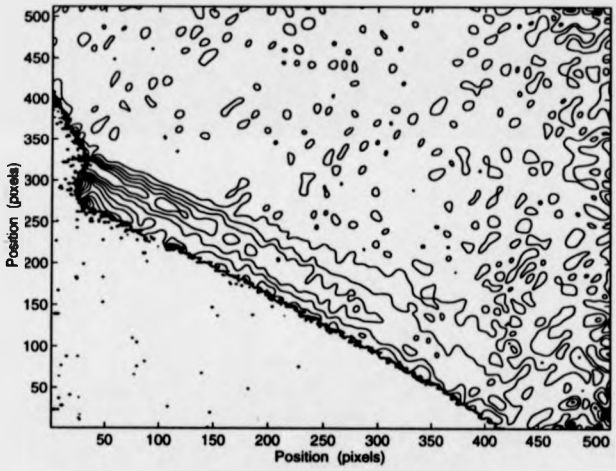



Figure 2.24: The contour plot of retrieved phase

# Bibliography


[1] M. Takeda, H. Ina and K. Kobayashi, "Fourier-transform Method of Fringe Pattern Analysis for Computer-based Topography and Interferometry", J. Opt. Soc. Am., Vol. 72, p.156-160, 1982.

[2] M. Takeda, "Spatial-carrier Fringe-pattern Analysis and its Application to Precision Interferometry and Profilometry: An Overview", Industrial Metrology, Vol. 1, p. 79-99, 1990.

[3] W. W. Macy, Jr., "Two-dimensional Fringe-pattern Analysis", Appl. Opt., Vol. 22, p. 3898-3901, 1983.

[4] D. J. Bone, H. A. Bachor, and R. J. Sandeman, "Fringe-pattern Analysis Using a 2-D Fourier Transform", Appl. Opt., Vol. 25, p. 1653-1659, 1986.

[5] K. A. Nugent, "Interferogram Analysis Using an Accurate Fully Automatic Algorithm", Appl. Opt., Vol. 24, p. 3101 - 3105, 1985.

[6] A. Low, "Introductory Computer Vision and Image Processing", McGraw-Hill, ISBN 0-07-707403-3

[7] Y. Morimoto, Y. Seguchi and T. Higashi, "Two-dimensional Moire Method and Brid Method Using Fourier Transform", Experimental Mechanics, Vol. 29, p. 339 - 404, 1989.

[8] J. M. Huntley and J. E. Field, "High Resolution Moire Photography : Application to Dynamic Stress Analysis", Opt. Eng., Vol. 28, p. 926 - 933, 1989.

[9] D. R. Burton and M. J. Lalor, "Multichannel Fourier Fringe Analysis as an Aid to Automatic Phase Unwrapping", Appl. Opt., Vol. 33, p. 2939 - 2948, 1994.

[10] M. Takeda and M. Kitoh, "Spatio-temporal Frequency-multiplex Heterodyne Interferometry", Proc. SPIE 1553, p. 66 - 76, 1991.

[11] M. Takeda and M. Kitoh, "Spatiotemporal Frequency Multiplex Heterodyne Interferometry", J. Opt. Soc. Am. A, Vol. 9, p. 1607 - p.1614, 1992.

[12] M. Takeda, "Temporal Versus Spatial Carrier Techniques for Heterodyne Interferometry", Proc. SPIE 814, p. 329-330, 1987.

[13] C. Roddier and F. Roddier,"Interferogram Analysis Using Fourier Transform Techniques", Appl. Opt., Vol. 26, p. 1668-1673, 1987.

[14] R. W. Gerchberg, "Super-Resolution Through Error Energy Reduction", Opt. Acta Vol. 21, p.709,1874.

[15] R. Jones and C. Wykes, "Holographic and Speckle Interferometry", Camberidge University Press, 1989.

[16] J.W. Goodman, "Laser Speckle and Related Phenomena", Springer-Verlag, 1975.

[17] P. Hariharan,"Basics of Interferometry" Academic Press Inc, 1992.

[18] M. Takeda and K. Motoh, "Fourier Transform Profilometry for Automatic Measurement of 3-D Object Shapes", Appl. Opt., Vol. 22, p. 3977-3982, 1983.

[19] P.J. Bryanston-Cross, P.H. Chan and T.R. Judge, "Analysis of Lateral Shearing Interferogram Using 2-D DFT Techniques", Proc. of Engineering Applications of Optical Diagnostic Techniques, 7th December 1994, Cranfield University.

# Chapter 3

# The Effect of Weighting Function and Filtering Window on Accuracy of 1-D Fourier Fringe Analysis

## 3.1 Introduction

When a finite-extent sampled signal is transformed into the frequency domain by the discrete Fourier transform (DFT), it is assumed that the N sample values of the finite-extent signal represent exactly one period of a periodic function. The DFT implies periodicity in both the time and frequency domains. If the signal is not a periodic function or if the observation interval is not equal to a integer multiple of the period of a periodic signal, the periodic extension of the observed signal will exhibit discontinuities at the boundaries of the observation. The discontinuities cause a phenomena that is called spectral leakage. The form of this discontinuity is demonstrated in Fig 3.1.

In order to alleviate the spectral leakage, a weighting function is applied to the data prior to the DFT. This chapter will address the effect of weighting function and filtering window on the performance of fringe analysis using one dimensional (1-D) DFT.

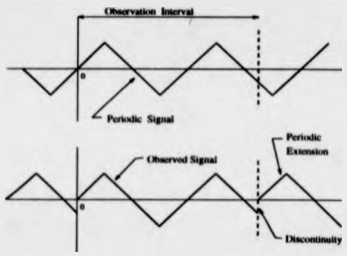



Figure 3.1: Periodic extension of observed one dimensional signal with discontinuity

## 3.2 The Characteristics of 1-D Weighting Function Used in the Experiment

It is often desirable to change the basic DFT output to meet the objectives of a particular application [1] [2] [3]. The modification can be accomplished by applying a weighting function to the data.

A number of popular weighting functions are employed in this study. The following sub-section, which is derived from Elliott et al [4] and Harris [5], presents the characteristics of these weighting functions. Firstly, the major parameters, that will allow performance comparison between different weighting functions, are identified. A short description of each of the parameters follow.

**Highest Sidelobe Level** It is the peak ripple value of the sidelobes. The amount of leakage into the sidelobes is directly proportional to the height of the highest sidelobe.

**Sidelobe Fall-off** This parameter describes the amplitude fall off of sidelobe peaks. This value is inverse proportional to the amount of leakage into the sidelones.

**Coherent Gain** This is a measure of the weighting function gain. For a rectangle weighting function, the gain is equal to one. For any other window, the gain is less than one due to the weighting function smoothly going to zero near the boundaries.

**3.0 dB Bandwidth** The point at which the gain of the shaped

DFT output is down 3.0 dB (half-power point), measured in DFT frequency bin widths, i.e. the sampling interval in frequency domain. This parameter indicates the minimum separation between two equal strength frequency such that their respective main lobes can be resolved.

### 3.2.1 Summary of 1-D Weighting Functions

All weighting functions have sidelobes of lower amplitude in the frequency domain than those of rectangular function and hence produce less leakage. On the other hand, all of the weighting functions have a broader main lobe. The trade-off between leakage (side-lobe level) and resolution (main-lobe bandwidth) is the main consideration to select the suitable weighting function for particular application. The summary of weighting functions used in this study are presented here. Each weighting function will firstly be defined and then its parameters are identitied.

**Rectangle weighting function** The Rectangle weighting function is unity over the observation interval and can be thought of as a grating sequence applied to the data so that they are of finite extent. The weighting function for DFT is defined as

$$w(n) = 1.0, \quad n = 0, 1, \cdots, N-1 \tag{3.1}$$

The performance parameters follow.

| | |
|---|---|
| Highest sidelobe level | = -13 dB |
| Sidelobe fall-off | = -6 dB/Oct |
| Coherent gain | = 1.0 |
| 3.0 dB bandwidth | = 0.89 Bins |

**Hanning (raised cosine) weighting function** The Hanning weighting function is a member of weighting function family, $\cos^{\alpha}(X)$. When $\alpha = 2$, it is the Hanning weighting function. The definition of Hanning weighting function and its parameters follow.

$$w(n) = 0.5 - 0.5\cos\left(\frac{2n}{N}\pi\right), \quad n = 0, 1, \cdots, N-1 \tag{3.2}$$

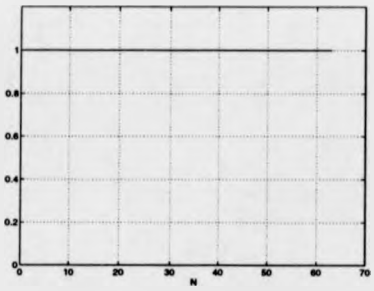



Figure 3.2: The Rectangle weighting function

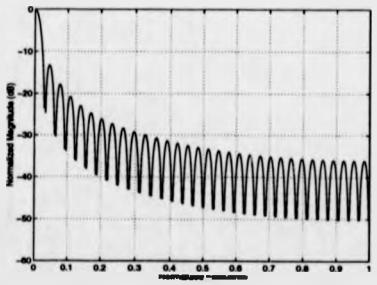



Figure 3.3: The frequency response of the Rectangle weighting function

| | |
|---|---|
| Highest sidelobe level | = -32 dB |
| Sidelobe fall-off | = -18 dB/Oct |
| Coherent gain | = 0.5 |
| 3.0 dB bandwidth | = 1.44 Bins |

Attractions of the $\cos^{\alpha}(X)$ family include the ease with which the terms can be generated, and the easily identified properties of the transform of the cosine function.

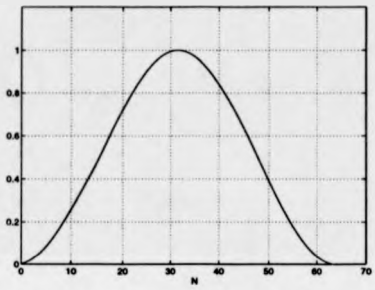



Figure 3.4: The Hanning weighting function

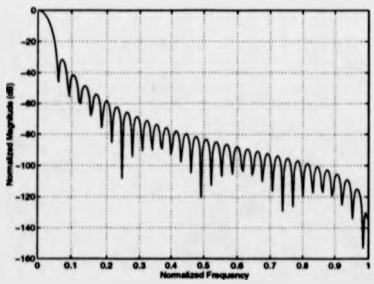



Figure 3.5: The frequency response of the Hanning weighting function

**Papoulis (sine lobe) weighting function** The Papoulis weighting function is another member of $\cos^{\alpha}(X)$ family. In this case, $\alpha$ is equal to one. This weighting function had been used in FRANSYS which is a automatic fringe analysis software package. The definition of the Papoulis weighting

function and its parameters follow.

$$w(n) = \sin\left(\frac{n}{N}\pi\right), \quad n = 0, 1, \cdots, N-1 \tag{3.3}$$

| | |
|---|---|
| Highest sidelobe level | = -23 dB |
| Sidelobe fall-off | = -12 dB/Oct |
| Coherent gain | = 0.64 |
| 3.0 dB bandwidth | = 1.23 Bins |

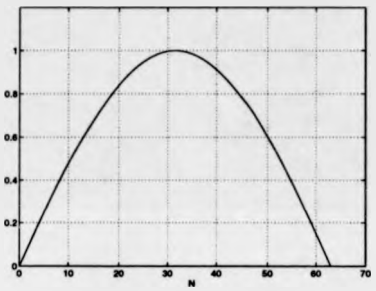



Figure 3.6: The Papoulis weighting function

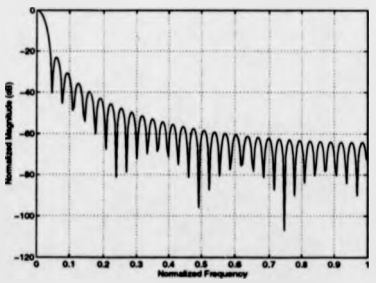



Figure 3.7: The frequency response of the Papoulis weighting function

**Hamming weighting function** The Hamming window can be thought of as a modified Hanning window. The Hamming weighting function for DFT is identified by

$$w(n) = 0.54 - 0.46\cos\left(\frac{2n}{N}\pi\right), \quad n = 0, 1, \cdots, N-1 \tag{3.4}$$

From Eq. 3.4, It is easy to spot the similarity between Hamming and Hanning weighting function. The Hamming weighting function and its frequency response is shown in Figure 3.8 and 3.9, respectively. It is noted the deep attenuation at the missing sidelobe position. The following is its performance parameters.

| | |
|---|---|
| Highest sidelobe level | = -43 dB |
| Sidelobe fall-off | = -6 dB/Oct |
| Coherent gain | = 0.54 |
| 3.0 dB bandwidth | = 1.5 Bins |

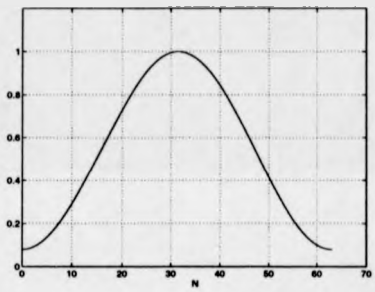



Figure 3.8: The Hamming weighting function

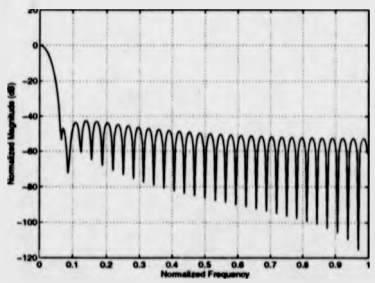



Figure 3.9: The frequency response of the Hamming weighting function

**Blackman weighting function** Generalizations of Hanning and Hamming weighting function for DFT sequence yield

$$w(n) = \sum_{m=0}^{K} (-1)^m a_m \cos\left(\frac{2\pi mn}{N}\right), \quad n = 0, 1, \cdots, N-1 \qquad (3.5)$$

Blackman used $K = 2$ to null the shaped DFT output at f = 3.5/NT and 4.5/NT by using $a_0 = 0.42$, $a_1 = 0.5$ and $a_2 = 0.08$. Figure 3.10 and 3.11 shows the Blackman weighting function and its frequency response, respectively. Its parameters follow.

| | |
|---|---|
| Highest sidelobe level | = -58 dB |
| Sidelobe fall-off | = -18 dB/Oct |
| Coherent gain | = 0.42 |
| 3.0 dB bandwidth | = 1.68 Bins |

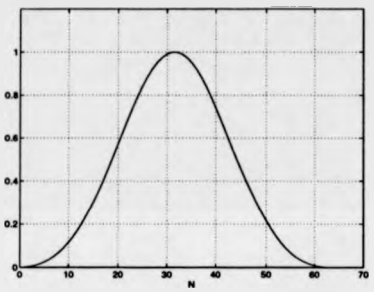



Figure 3.10: The Blackman weighting function

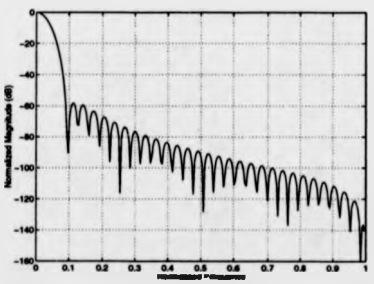



Figure 3.11: The frequency response of the Blackman weighting function

## 3.3 The Joint Effect of 1-D Weighting Function and Filtering Window

Previous study [6] only concerned the effect of the weighting function, and the rectangle filter window is used in the study. Although J. Schmit et. al. [7] have pointed out that the filter window influences the accuracy of Fourier fringe analysis, this study was only focussed on the effect of the size of the rectangular window. A computer simulation has been conducted to investigate the combined influence of various weighting function and filtering window on the result of Fourier fringe analysis. The filtering window investigated include the Rectangle, Hanning, Papoulis, Hamming and Blackman filtering window. This section is presented the result of the investigation.

When DFT is applied for fringe analysis in practical application, there are several sources of error which cannot be avoided. Some of these errors include :

- non-linearity of recording media
- errors introduced by optical geometry
- the influence of random noise

In order to be free from these errors, a computer generated interferogram is used in this experiment. The phase distribution of the computer generated interferogram is expressed as

$$\phi(x,y) = \begin{cases} f(x,y) & \text{if } f(x,y) \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{3.6}$$

where

$$f(x,y) = 0.8\pi \left[1 - \left(\frac{x-256}{256}\right)^2 + \left(\frac{y-256}{256}\right)^2\right]$$

Figure 3.12 shows the three dimensional plot of $\phi(x,y)$. The computer generated interferogram used in this experiment is depicted in Figure 3.13. The wave length of carrier frequency in this interferogram is six pixels.

In each computer simulation, the weighted data is transformed into the frequency domain by DFT. Then, various filtering windows with same size are applied to one of the sidelobes. It means that there are twenty-five results

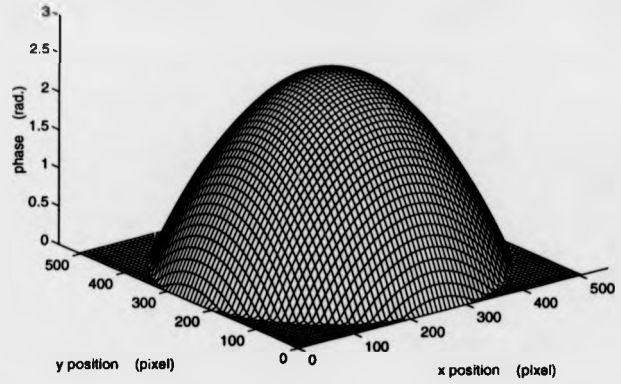



Figure 3.12: The 3-D plot of the computer generated interferogram's phase distribution

for five weighting functions and five filtering windows. The heterodyning was removed by subtracting the term $2\pi x f_0$, where $f_0$ is carrier frequency, after the whole phase is retrieved. This procedure tends to increase the accuracy [8]. The error values are obtained by subtracting the calculated phase value form the theoretical one. A statistical analysis of absolute value of the error is performed. Figure 3.14 and 3.15 show the analysis result of the full field errors

After the profile of full field errors is inspected, it is found that the phase error at the rim region has a high amplittude. Obviously, the error in the rim region produces a misleading analysis result. Futhermore, the area of interest is usually in the center of the image in many partical application, for example Juge et al [9]. Two further statistical analysis of absolute value of the error in the central regional are undertaken. In the first analysis, five pixels are masked out on both side of each raster line. The result is shown in Figure 3.16 and 3.17. Another is undertaken after ten pixels are masked out on both side of each raster line. The results of the second analysis are shown in Figure 3.18 and 3.19.

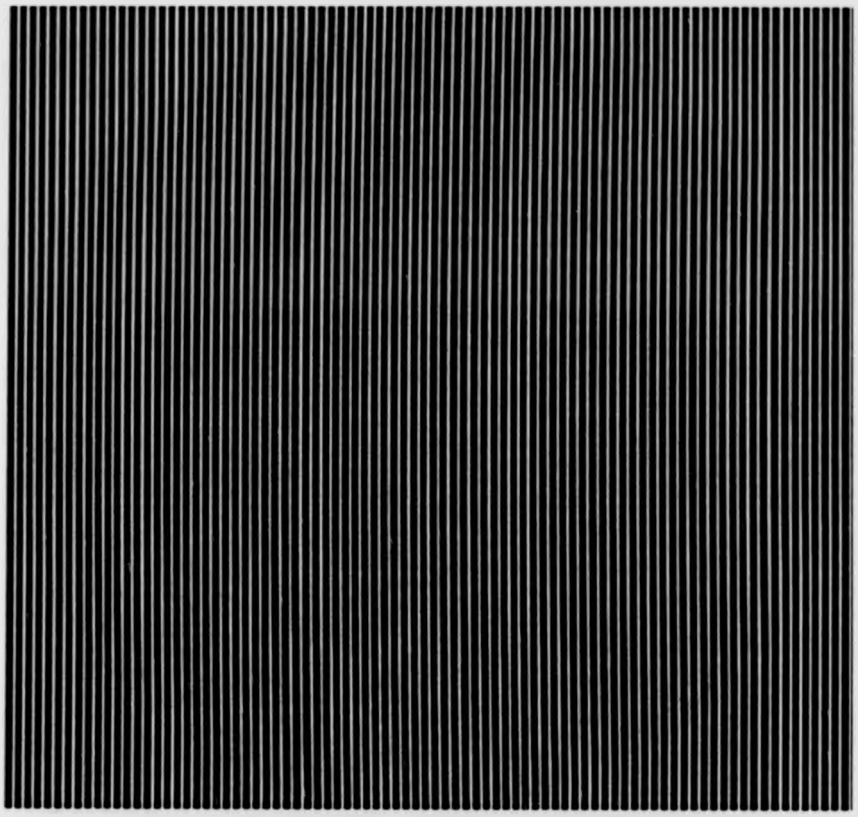


Figure 3.13: The computer generated interferogram

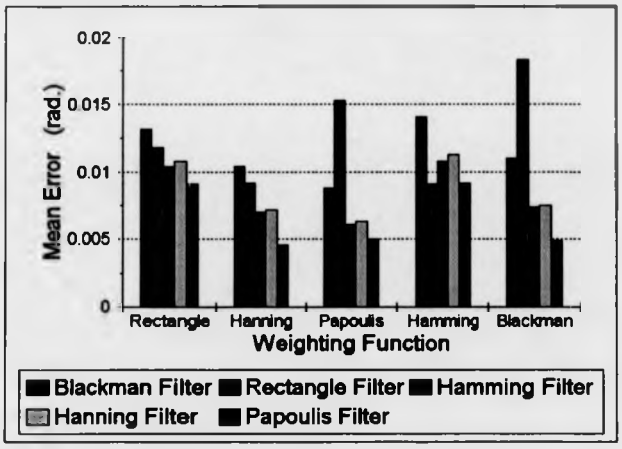



Figure 3.14: The mean absolute value of the full field errors

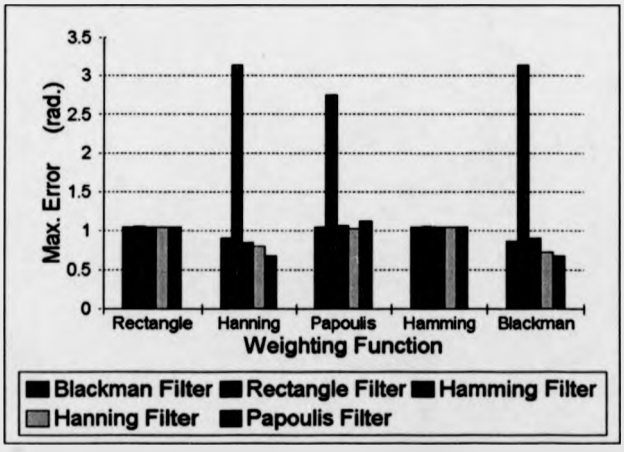



Figure 3.15: The maximum absolute value of the full field errors

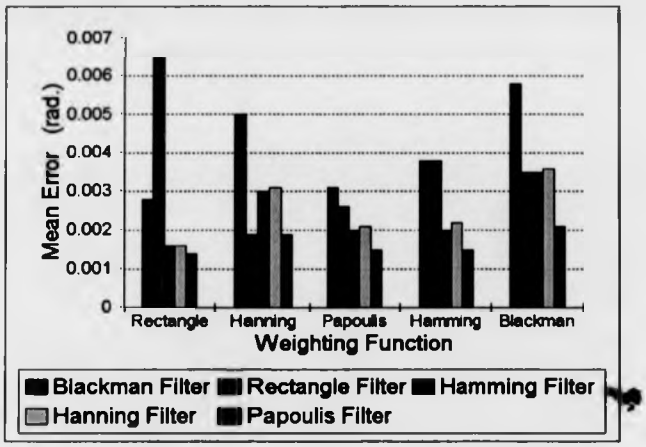



Figure 3.16: The mean absolute value of the errors after five pixels are masked out

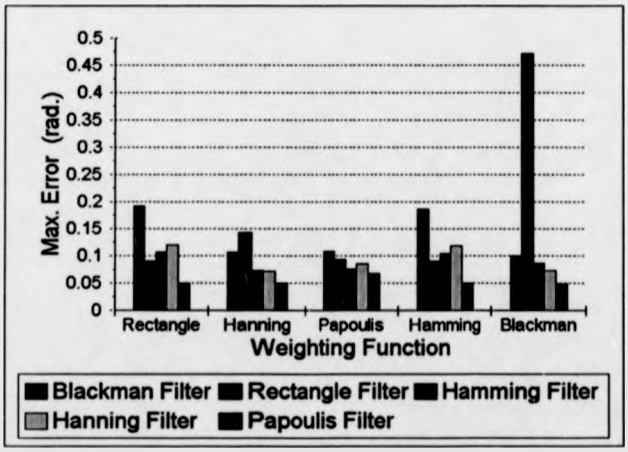



Figure 3.17: The maximum absolute value of the errors after five pixels are masked out

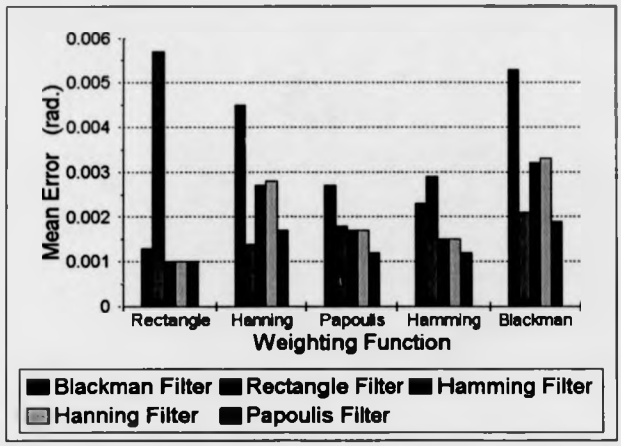



Figure 3.18: The mean absolute value of the errors after ten pixels are masked out

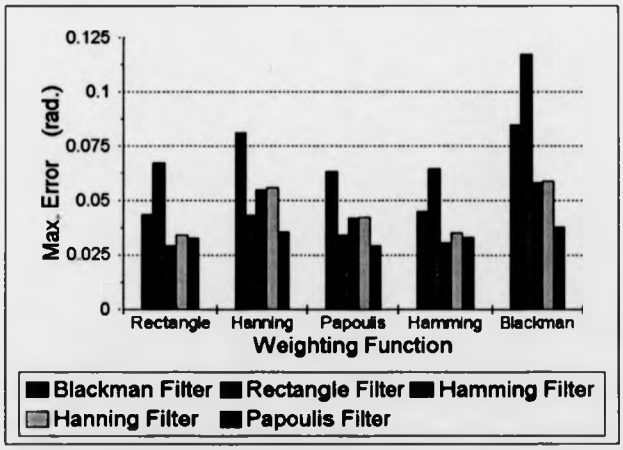



Figure 3.19: The maximum absolute value of the errors after ten pixels are masked out

## 3.4 Discussion

A computer generated interferogram was employed in all experiments so that the results are free from all error except those created by the spectral leakage. According to the analysis result of full field errors, it indicates that the Hamming funcion performes better than other functions on average when the Rectangle filter is employed. It also shows that, when the Rectangle filter is used, the lowest maximum error is produced by the Hamming function. This explains why the Hamming function is widely recommended to alleviate the spectral leakage problem. However, the highest mean accuracy, which is $\pm 0.0046$ rad. over the full field, can be achieved by the Hanning function and the Papoulis filter. The lowest maximum full field error produced by Hanning function and Papoulis filter is $\pm 0.6792$ rad..

Referring to partial field analysis, the mean error is significantly reduced by masking out 5 pixels on both sides of each raster line. On another hand, the reduction of mean error will be modest if further 5 pixels on both sides of each raster line are masked out. The lowest partial field mean error can be achieved by the Papoulis filter combined with the Hamming, the Rectangle or the Papoulis function. Further the amplitude of maximum error can be reduced significantly by masking more pixels in the rim region.

From Figure 3.14 to Figure 3.19, it is found that the Papoulis filter's performance outvies all other filters under all condition in the experiment. It can be concluded that, based on this experiment, the Papoulis filter should be used for high accuracy 1-D Fourier fringe analysis. If the mean error is the most important factor, the Papoulis function should be used. However, the Hanning function should be employed, if the maximum error is critical. Apart from combining with the Rectangle filter, the performance of the Rectangle weighting function is not worse than other weighting functions if ten pixels are masked out. The Rectangle weighting function should be used when the computing time is a critical factor and the area of interest is in the central region.

# Bibliography


[1] A. Papoulis, "Minimum Bias Windows for High-resolution Spectral Estimates", IEEE Trans. Informat. Theory IT-19 (1973), 9-12.

[2] L. R. Rabiner and B. Gold, "Theory and Application of Digital Signal Processing", Prentice-Hall, 1975.

[3] R. K. Otnes and L. Enochson, "Digital Time Series Analysis" Wiley, 1972.

[4] D. F. Elliott and K. R. Rao, "Fast Transforms Algorithms, Analyses, Application" Academic Press, 1982, ISBN 0-12-237080-5.

[5] F. J. Harris, "On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform", Proceedings of IEEE, Vol. 66 No. 1 (1978), 51-83.

[6] P. J. Bryanston-Cross, C. Quan and T. R. Judge, "Application of the FFT Method for the Quantitative Extraction of Information from High-resolution Interferometric and Photoslastic Data", Optics & Laser Technology, Vol. 26 No. 3 (1993), 147-155

[7] J. Schmit, K. Creath and M. Kujawinska,"Spatial and Temporal Phase-measurement Techniques: A Comparision of Major Error Sources in One-dimension", SPIE Vol. 1755 (1992), P. 202-211.

[8] M. Kujawinska and Wojciak, "High Accuracy Fourier Transform Fringe Pattern Analysis", Optics and Lasers in Eng., Vol 14 (1991), P. 325-339.

[9] T. R. Judge, C. Quan and P. J. Bryanston-Cross "Holographic Deformation Measurements by Fourier Transform Technique with Automatic Phase Unwrapping", Optical Eng., Vol.31 No.3 (1992), P. 533-543.

# Chapter 4

# The Effect of Weighting Function and Filtering Window on Accuracy of 2-D Fourier Fringe Analysis

## 4.1 Introduction

In the previous chapter, it is indicated that periodicity in both the time and frequency domain is implied by one dimensional (1-D) discrete Fourier transform (DFT). This implication is also true for the two dimensional (2-D) discrete Fourier transform. Figure 4.1 and 4.2 illustrates the implication. The 3x3 matrix, shown in Figure 4.1, is the sampled 2-D signal. When the signal is transformed into frequency domain by DFT, the signal must be interpreted as one period of a periodic 2-D function. Figure 4.2 shows the periodically extended function of the 3x3 matrix. The original signal is enclosed by the dotted line.

The negative effect of the periodicity is to introduce a discontinuity in the data between periods. As a consequence, spectral leakage occurrs. As in the one-dimensional case, a weighting function can be utilized to reduce the spectral leakage. In this chapter, the formulation and property of 2-D weighting functions is firstly discussed. Then, the result of a computer simulation is presented to indicate the combined effect of weighting function and

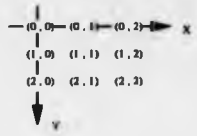


Figure 4.1: Two dimensional discrete signal

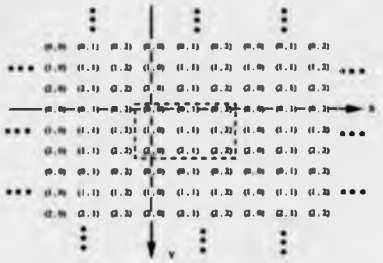


Figure 4.2: Two dimensional periodic extension of signal shown in figure 4.1

filtering window on the performance of fringe analysis using two dimensional DFT.

## 4.2 Formation of 2-D Weighting Function

The 1-D weighting window is often employed as a prototype for generating a 2-D weighting function. There are two dominating approaches. The first method [1] is to form a 2-D window, $w_r(n_1,\ n_2)$, by

$$w_r(n_1,\ n_2) = w_x(n_1)w_y(n_2) \tag{4.1}$$

The functions $w_x(n_1)$ and $w_y(n_2)$ is a 1-D window sequences. It can be thought of that $w_r(n_1,\ n_2)$ is obtained by taking the outer product of $w_x(n_1)$ and $w_y(n_2)$. The second method [2] is to generate a 2-D window, $w_c(n_1,\ n_2)$, by sampling a circularly rotated 1-D continuous window function.

$$w_c(n_1,\ n_2) = w_s(t) \tag{4.2}$$

where $t = \sqrt{n_1^2 + n_2^2}$. The function $w_s(t)$ in Eq. 4.2 is a 1-D analogy window.

Two well-behaved and popular 1-D windows , Hamming and Papoulis, are used as prototype in this experiment. Figure 4.3 shows the contour plot of the frequency responses of the Hamming outer-product window. Its first quadrant perspective plot of the frequency responses is shown in Figure 4.4. Figure 4.5 and 4.6 show the contour plot and first quadrant perspective plot of the frequency responses of the Hamming circularly rotated window. Figure 4.7, 4.8, 4.9 and 4.10 show a similar sequence of plot for the Papoulis outer-product window and the Papoulis circularly rotated window respectively.

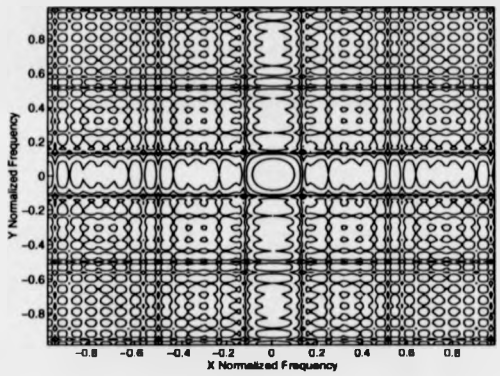



Figure 4.3: The contour plot of the frequency responses of the Hamming outer-product window

For a comparison, the contour plot and the first quadrant perspective plot of the frequency responses of the Rectangle window, i.e. no modification, is shown in Figure 4.11 and 4.12.

## 4.3 The Joint Effect of 2-D Weighting Function and Filtering Window

Studies conducted previously only address either the weighting function or the filter window [3] [4]. In this section, a number of weighting function and filter window are employed to examine the combined effect on the accuracy

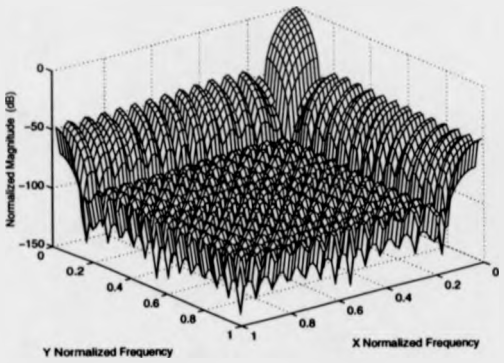



Figure 4.4: The first quadrant mesh plot of the frequency responses of the Hamming outer-product window

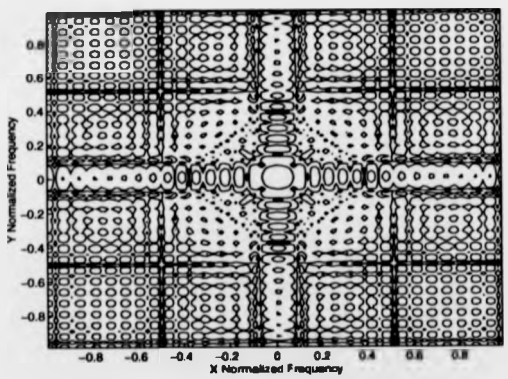



Figure 4.5: The contour plot of the frequency responses of the Hamming circularly rotated window

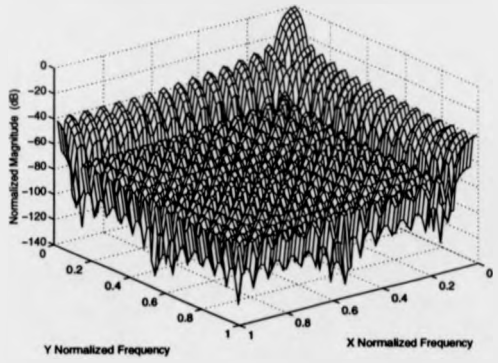



Figure 4.6: The first quadrant mesh plot of the frequency responses of the Hamming circularly rotated window

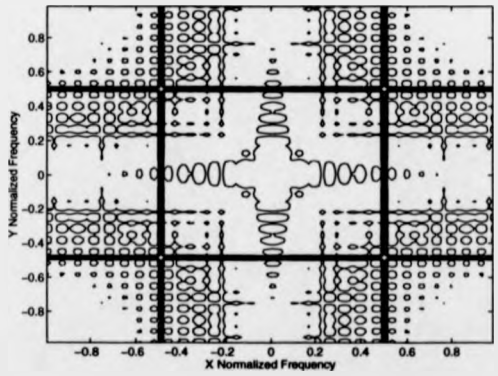



Figure 4.7: The contour plot of the frequency responses of the Papoulis outer-product window

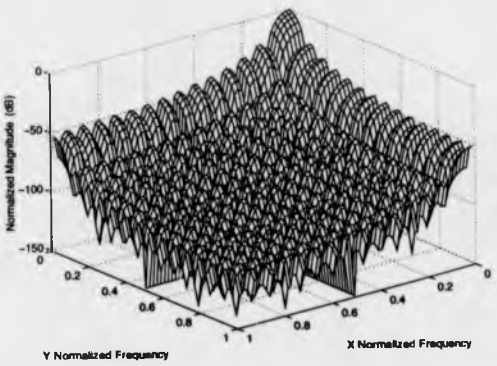



Figure 4.8: The first quadrant mesh plot of the frequency responses of the Papoulis outer-product window

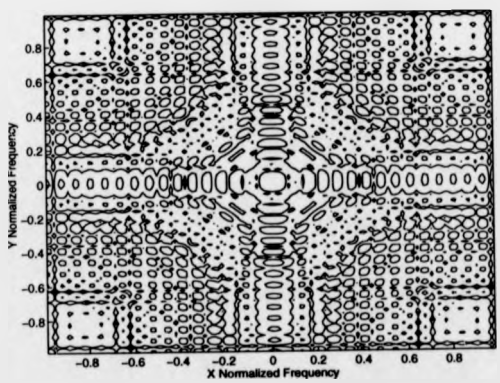



Figure 4.9: The contour plot of the frequency responses of the Papoulis circularly rotated window

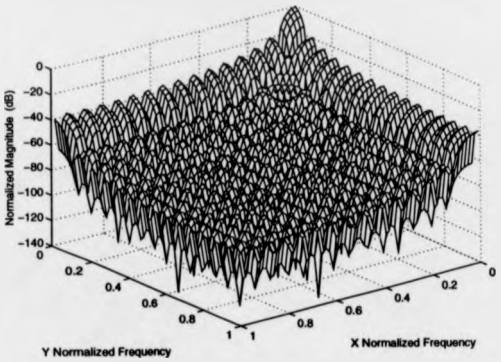



Figure 4.10: The first quadrant mesh plot of the frequency responses of the Papoulis circularly rotated window

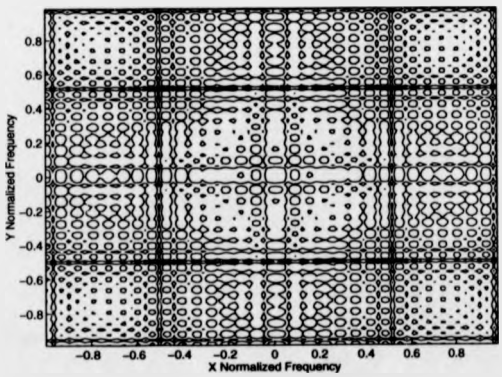



Figure 4.11: The contour plot of the frequency responses of the Rectangle window

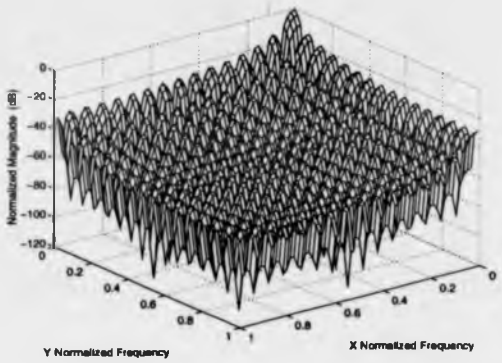



Figure 4.12: The first quadrant mesh plot of the frequency responses of the Rectangle window

of 2-D Fourier fringe analysis. The one dimensional Rectangle, Papoulis and Hamming windows are used as prototypes to generate 2-D filter windows as well as weighting functions by both the outer product method and the circularly rotated method. These windows are described as well-behaved windows in the digital signal processing text books.

A computer generated interferogram also is used in this experiment. The phase distribution of the computer generated interferogram is expressed as

$$\phi(x,y) = \begin{cases} f(x,y) & \text{if } f(x,y) \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{4.3}$$

where

$$f(x,y) = 0.8\pi \left[1 - \left(\frac{x-256}{256}\right)^2 + \left(\frac{y-256}{256}\right)^2\right]$$

A interferogram with non-vertical fringe, which is depicted in Figure 4.13, is used in the experiment. It is because the 2-D Fourier fringe analysis is able to cope with non-vertical carrier fringes. In order to increase the accuracy, the heterodyning was remove by subtracting the term $2\pi \mathbf{r}\mathbf{f_0}$, where $\mathbf{r}$ and $\mathbf{f_0}$ is the position vector and the carrier frequency vector respectively, from the retrieved phase. The absolute error value is the absolute value of the difference between the theoretical phase value and the calculated one. Figure 4.14 and 4.15 show the analysis result of the full field absolute error.

It is similar with the 1-D Fourier fringe analysis result that the rim region

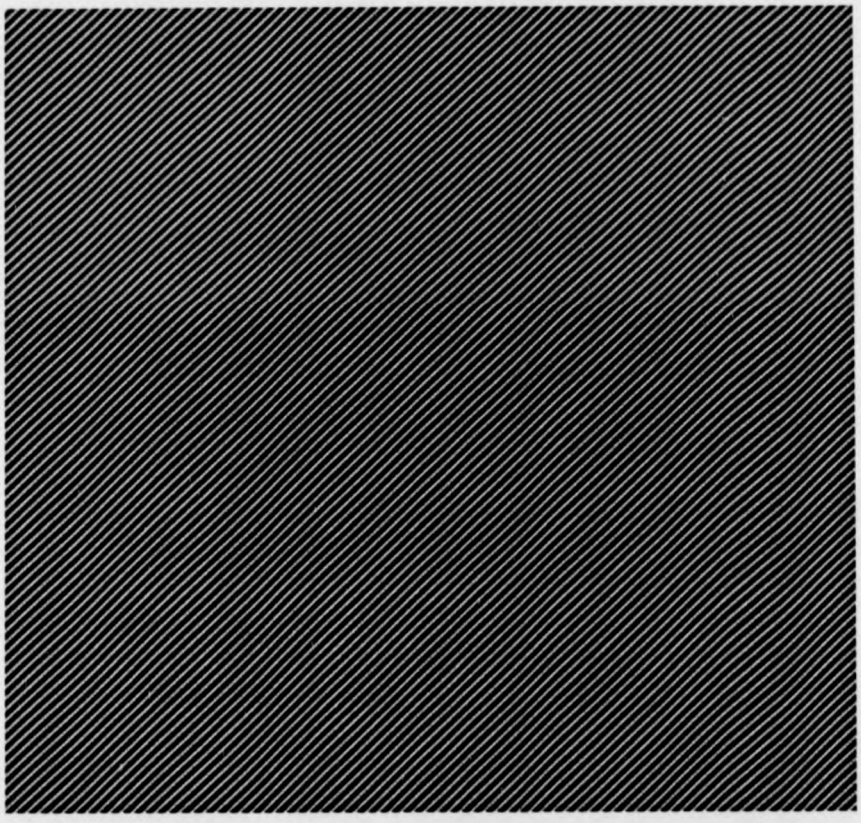


Figure 4.13: The computer generated interferogram with non-vertical fringe

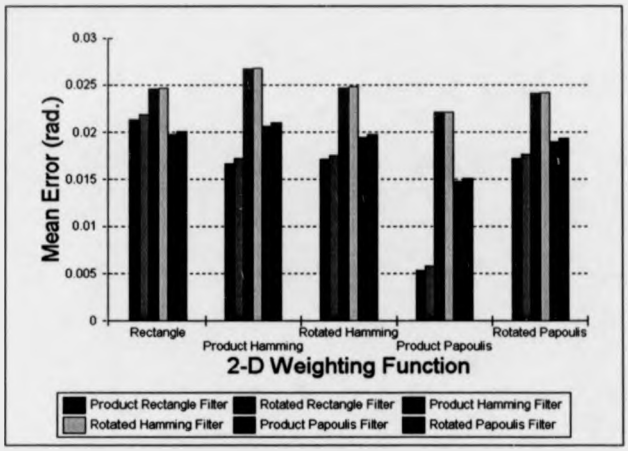



Figure 4.14: The mean absolute value of the full field errors

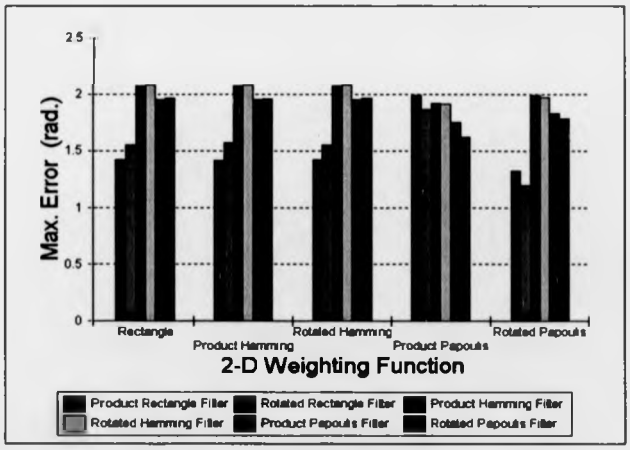



Figure 4.15: The maximum absolute value of the full field errors

has a high phase error. However, there is a distinction between them. In the two dimensional case, there are four rim regions on each edge of the interferogram in stead of two rim regions for one dimensional case. In order to avoid the misleading data, two further statistical analysis of absolute error in the central region are performed. The first analysis is carried out after five pixels are masked out on each side of interferogram. The result are shown in Figure 4.16 and 4.17. Figure 4.18 and 4.19 show the another analysis result of absolute error excluding ten pixels on each side of interferogram.

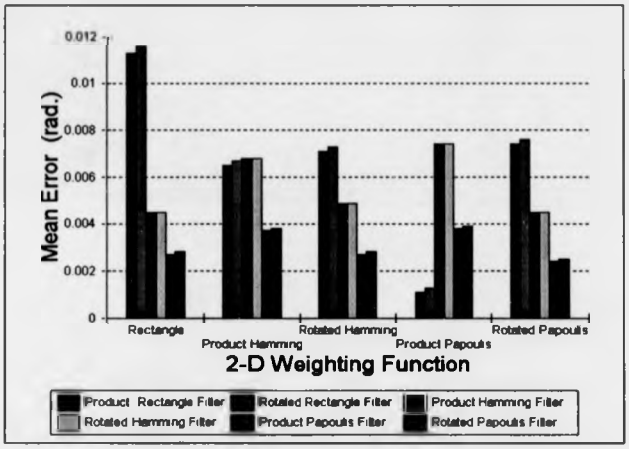



Figure 4.16: The mean absolute value of the errors after five pixels are masked out

## 4.4 Discussion

A computer generated interferogram with non-vertical carrier fringe is used to reflect a generic case. The result of full field error indicates that the Papoulis outer-product weighting function and the Rectangle outer-product filter produce the lowest mean error, which is equal to $\pm 0.0053 rad.$. The lowest maximum error is achieved by the combined effect of the Papoulis circularly rotated weight function and the Rectangle circularly rotated filter.

The lowest maximum error is reduced to $\pm 0.0397 rad.$ and to $\pm 0.0189 rad.$ when 5 and 10 pixels are masked out on each side of interferogram. The lowest

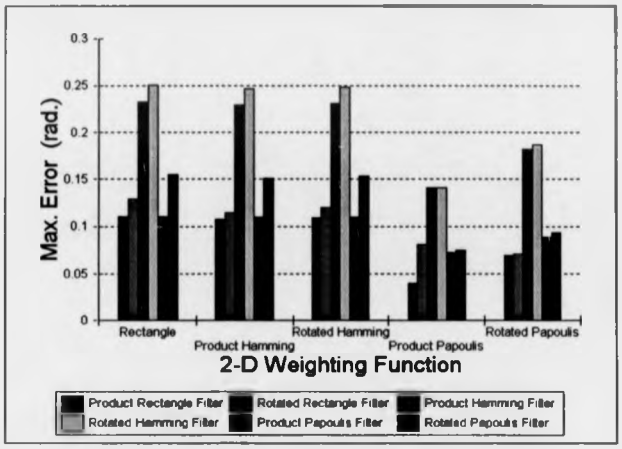



Figure 4.17: The maximum absolute value of the errors after five pixels are masked out

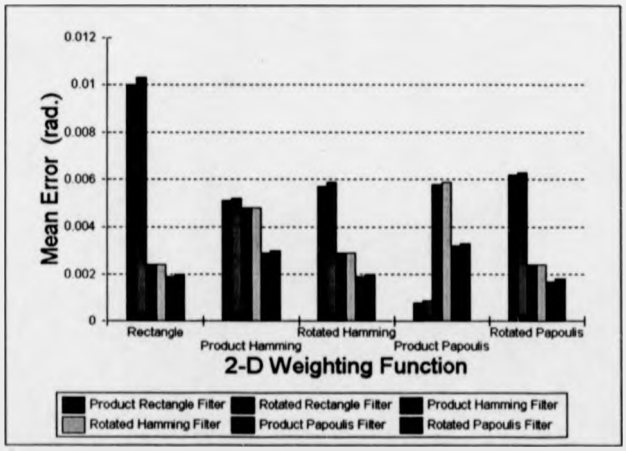



Figure 4.18: The mean absolute value of the errors after ten pixels are masked out

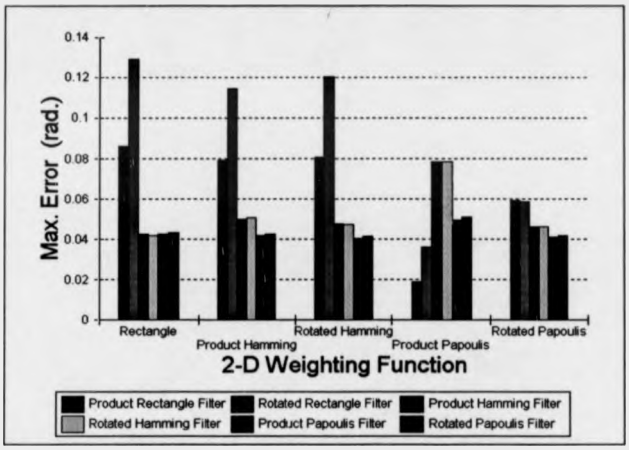



Figure 4.19: The maximum absolute value of the errors after ten pixels are masked out

mean error is $\pm 0.0011 rad.$ and $\pm 0.0008 rad.$ corresponding to 5 and 10 pixels being masked out. From this experiment result, it is speculated that the Papoulis outer-product weighting function and the Rectangle outer-product filter is the best choice to alleviate the spectral leakage problem caused by 2-D Fourier fringe analysis. Finally, the computer simulation reveals that the outer-product method is superior to the circularly rotated method when the same prototype window is employed to form a filter window.

# Bibliography

[1] D.E. Dudgeon and R.M. Mersereau, "Multimensional Signals and System". Prentice-Hall, 1984.

[2] T.S. Huang, "Two-dimensional Windows", IEEE Trans. Audio Electroacoust, Vol. AU-20, p. 89-90, 1972.

[3] P.R. Stephenson, "Evaluation and Solutions of Key Problems in Fourier Fringe Analysis", PhD Thesis, Engineering and Technology Management, Liverpool J. M. University.

[4] M. Kujawinska and J. Wojciak, "High Accuracy Fourier Transform Fringe Pattern Analysis",Optics and Lasers in Engineering, Vol. 14, p. 325-339, 1991.

# Chapter 5

# Fringe Pattern Analysis by Phase-stepping Techniques

## 5.1 Introduction

The Phase-stepping techniques are developed from the integrating-bucket technique. The integrating-bucket technique [1] uses a time-varying fringe images of the form

$$\begin{aligned} I(x,y;t) &= I_0(x,y)\{1+\gamma\cos[\alpha(t)+\phi(x,y)]\} \\ &= I_0(x,y)\{1+\gamma\cos[2\pi f_o t+\phi(x,y)]\} \end{aligned} \tag{5.1}$$

where $I_0(x,y)$ is the average intensity at each detector point, $\gamma$ is the modulation of the fringe pattern, $\phi(x,y)$ is a phase to be determined, and $f_o$ is the temporal-carrier frequency introduced by a frequency-shifting device [2]. The phase-stepping techniques [3] can be considered as a discrete-time version of the integrating-bucket technique where the time-varying interferogram is sampled over a single period at time intervals $\Delta t = 1/Nf_o$, so that

$$\begin{aligned} I_n(x,y) &= I_0(x,y)\{1+\gamma\cos[\frac{2\pi n}{N}+\phi(x,y)]\} \\ &= I_0(x,y)\{1+\gamma\cos[\alpha_n+\phi(x,y)]\} \end{aligned} \tag{5.2}$$

Where $I_n$ represents the nth data frame of the phase-stepping technique, $\alpha_n$ is the phase shift of the nth frame of fringe images and $n = 1, 2, \cdots, N$.

Since there are so many possible algorithms, a general phase measurement problem is firstly addressed. Then, the specific equations for the most used algorithms is presented.

Normally, N measurements of the intensity are recorded as the phase is shifted. For the general technique the phase shift is assumed to change during the detectors' integration time, and this change is the same from data frame to data frame. The detector array will integrate the fringe intensity data over a change in relative phase of $\Delta$. One set of recorded fringe pattern will be written as [4]:

$$I_i(x,y) = \frac{1}{\Delta}\int_{\alpha_i-\frac{\Delta}{2}}^{\alpha_i+\frac{\Delta}{2}} I_o(x,y)\{1+\gamma\cos[\phi(x,y)+\alpha(t)]\}d\alpha(t) \qquad (5.3)$$

where $\alpha_i$ is the average value of the relative phase shift for the $i$th exposure. The integration over a phase shift $\Delta$ makes Eq. 5.3 applicable for any phase-shifting technique. After integrating Eq 5.3, the recorded intensity is

$$I_i(x,y) = I_o(x,y)\{1+\gamma sinc(\frac{\Delta}{2})cos[\phi(x,y)+\alpha_i]\} \qquad (5.4)$$

where

$$sinc(\frac{\Delta}{2}) = \frac{\sin(\frac{\Delta}{2})}{\frac{\Delta}{2}}$$

If the phase shifts are stepped, i.e. $\Delta = 0$, the *sinc* function will have a value of unity. Therefore, phase stepping is a special case of the integrating-bucket method.

For a total of $N$ recorded intensity measurements the phase can be calculated using a least-squares technique. Eq. 5.4 is first rewritten in the form

$$I_i(x,y) = a_0(x,y) + a_1(x,y)\cos\alpha_i + a_2(x,y)\sin\alpha_i \qquad (5.5)$$

where

$$\begin{aligned} a_0(x,y) &= I_o(x,y), \\ a_1(x,y) &= I_o(x,y)\gamma sinc(\frac{\Delta}{2})\cos[\phi(x,y)], \\ a_2(x,y) &= I_o(x,y)\gamma sinc(\frac{\Delta}{2})\sin[\phi(x,y)], \end{aligned}$$

The unknowns of this set of equations are $I_o(x,y)$, $\gamma$ and $\phi(x,y)$, which are represented by $a_0(x,y)$, $a_1(x,y)$ and $a_2(x,y)$. The least-squares solution to these equations is

$$\begin{bmatrix} a_0(x,y) \\ a_1(x,y) \\ a_2(x,y) \end{bmatrix} = A^{-1}(\alpha_i)\, B(x,y,\alpha_i) \tag{5.6}$$

where

$$A(\alpha_i) = \begin{bmatrix} N & \sum\cos(\alpha_i) & \sum\sin(\alpha_i) \\ \sum\cos(\alpha_i) & \sum\cos^2(\alpha_i) & \sum[\cos(\alpha_i)]\sin(\alpha_i) \\ \sum\sin(\alpha_i) & \sum[\cos(\alpha_i)]\sin(\alpha_i) & \sum\sin^2(\alpha_i) \end{bmatrix}$$

and

$$B(x,y,\alpha_i) = \begin{bmatrix} \sum I_i(x,y) \\ \sum I_i(x,y)\cos(\alpha_i) \\ \sum I_i(x,y)\sin(\alpha_i) \end{bmatrix}$$

Eq. 5.6 must be solved at each point in the fringe image for $a_1(x,y)$ and $a_2(x,y)$. The matrix $A$ needs to be calculated and inverted just once because it is dependent only on the phase shift. The phase map of the unknown wavefront can now be determined easily by taking the ratio of $a_2(x,y)$ to $a_1(x,y)$:

$$\phi(x,y) = \arctan\left[\frac{a_2(x,y)}{a_1(x,y)}\right] \tag{5.7}$$

This phase calculation assumes that the phase shifts between measurement are known and that the integration period is constant for every measurement.

## 5.2 Phase-stepping Measurement Algorithms

### 5.2.1 Three Phase Step Algorithm

Three fringe images of different reference beam phase are required since the minimum number of equation required to solve Eq. 5.3 is three. In the case of three fringe images at a phase step of $\alpha$, the phase value at a given pixel

may be calculated from

$$\phi(x,y) = \arctan\left[\frac{(I_3 - I_2)\cos\beta + (I_1 - I_3)\cos(\beta+\alpha) + (I_2 - I_1)\cos(\beta+2\alpha)}{(I_3 - I_2)\sin\beta + (I_1 - I_3)\sin(\beta+\alpha) + (I_2 - I_1)\sin(\beta+2\alpha)}\right] \tag{5.8}$$

where $\beta$ is the initial phase difference.

In the case of a phase shift of $\frac{\pi}{2}$ with initial phase difference which is equal to $\frac{\pi}{4}$, i.e. $\beta = \frac{\pi}{4}$. The three intensity measurement can be expressed as

$$\begin{aligned} I_1(x,y) &= I_0(x,y)\{1 + \gamma\cos[\phi(x,y) + \frac{\pi}{4}]\} \\ I_2(x,y) &= I_0(x,y)\{1 + \gamma\cos[\phi(x,y) + \frac{3\pi}{4}]\} \\ I_3(x,y) &= I_0(x,y)\{1 + \gamma\cos[\phi(x,y) + \frac{5\pi}{4}]\} \end{aligned} \tag{5.9}$$

The phase at each point in the fringe image is then simplified as

$$\phi(x,y) = \arctan\left[\frac{I_3 - I_2}{I_1 - I_2}\right] \tag{5.10}$$

If a phase shift of $\frac{2\pi}{3}$ is used and $\beta = \frac{-2\pi}{3}$, the three intensity measurements become

$$\begin{aligned} I_1(x,y) &= I_0(x,y)\{1 + \gamma\cos[\phi(x,y) - \frac{2\pi}{3}]\} \\ I_2(x,y) &= I_0(x,y)\{1 + \gamma\cos[\phi(x,y)]\} \\ I_3(x,y) &= I_0(x,y)\{1 + \gamma\cos[\phi(x,y) + \frac{2\pi}{3}]\} \end{aligned} \tag{5.11}$$

The phase value is computed using the formula

$$\phi(x,y) = \arctan\left[\frac{\sqrt{3}(I_3 - I_2)}{2I_1 - I_2 - I_3}\right] \tag{5.12}$$

### 5.2.2 Four Phase Step Algorithm

The four step algorithm utilises four successive fringe images with the reference phase shifts between each frame. For phase-shift interferometry, the phase steps of multiples of $\frac{\pi}{2}$ is convenient to achieve. Furthermore, with frame processing and lookup-table manipulations, the phase value could be rapidly calculated using four step algorithm [5]. Usually, a price is paid in

accuracy if a lookup-table is used. When the phase shifts are introduced as $0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}$ within one period of the fringe, four sets of intensity measurement can be written as

$$\begin{aligned} I_1(x,y) &= I_0(x,y)\{1+\gamma\cos[\phi(x,y)]\} \\ I_2(x,y) &= I_0(x,y)\{1+\gamma\cos[\phi(x,y)+\frac{\pi}{2}]\} \\ I_3(x,y) &= I_0(x,y)\{1+\gamma\cos[\phi(x,y)+\pi]\} \\ I_4(x,y) &= I_0(x,y)\{1+\gamma\cos[\phi(x,y)+\frac{3\pi}{2}]\} \end{aligned} \tag{5.13}$$

The phase value at each point in the fringe image can be obtained by

$$\phi(x,y) = \arctan\left[\frac{I_4 - I_2}{I_1 - I_3}\right] \tag{5.14}$$

### 5.2.3 Carre's Algorithm

In the algorithm described in section 5.2.1 and 5.2.2, the phase shift is known either by calibrating the phase shifter or by measuring the amount of phase shift each time it is moved. Carre [6] introduced a technique of phase measurement that is independent of the amount of phase shift. It is assumed that the phase is linearly shifted by $2\alpha$ between consecutive intensity measurements. Four frames of intensity data are recorded in this manner:

$$\begin{aligned} I_1(x,y) &= I_0(x,y)\{1+\gamma\cos[\phi(x,y)-3\alpha]\} \\ I_2(x,y) &= I_0(x,y)\{1+\gamma\cos[\phi(x,y)-\alpha]\} \\ I_3(x,y) &= I_0(x,y)\{1+\gamma\cos[\phi(x,y)+\alpha]\} \\ I_4(x,y) &= I_0(x,y)\{1+\gamma\cos[\phi(x,y)+3\alpha]\} \end{aligned} \tag{5.15}$$

The phase calculated at each detected point $(x,y)$ in the fringe image is

$$\phi(x,y) = \arctan\left[\frac{\sqrt{|[3(I_2-I_3)-(I_1-I_4)][(I_2-I_3)(I_1-I_4)]|}}{|(I_2+I_3)-(I_1+I_4)|}\right] \tag{5.16}$$

where the phase falls between 0 and $\frac{\pi}{2}$. The signs of the quantities

$$I_2 - I_3 = (2I_0\gamma\sin\alpha)\sin[\phi(x,y)]$$

$$(I_2 + I_3) - (I_1 + I_4) \;=\; (2I_0\gamma \sin^2\alpha \cos\alpha)\cos[\phi(x,y)]$$

determine in which quadrant to place the phase because these quantities are proportional to $\sin[\phi(x,y)]$ and $\cos[\phi(x,y)]$. Difficulties can arise with Carre's algorithm, if the original phase difference between the two beams is close to $m\pi$, where $m$ is an integer. In these conditions, the numerator and denominator of Eq 5.16 tend to zero. As a result, the uncertainty in the values of $\tan[\phi(x,y)]$ incresaes.

### 5.2.4 Five Phase Step Algorithm

Schwider et al [7] proposed a algorithm with five intensity measurements to diminish the errors caused by the reference phase deviations from ideal. Hariharan et al [8] refined this idea and developed the five phase step algorithm. The algorithm uses five measurements of the intensity corresponding to additional phase steps of $-2\alpha$, $-\alpha$, 0, $+\alpha$ and $+2\alpha$. The five intensity measurements can be expressed as

$$\begin{aligned}
I_1(x,y) &= I_0(x,y)\{1 + \gamma\cos[\phi(x,y) - 2\alpha]\} \\
I_2(x,y) &= I_0(x,y)\{1 + \gamma\cos[\phi(x,y) - \alpha]\} \\
I_3(x,y) &= I_0(x,y)\{1 + \gamma\cos[\phi(x,y)]\} \\
I_4(x,y) &= I_0(x,y)\{1 + \gamma\cos[\phi(x,y) + \alpha]\} \\
I_5(x,y) &= I_0(x,y)\{1 + \gamma\cos[\phi(x,y) + 2\alpha]\}
\end{aligned} \tag{5.17}$$

These equations yield the result

$$\frac{I_2 - I_4}{2I_3 - I_5 - I_1} = \frac{\sin\alpha \sin[\phi(x,y)]}{(1 - \cos 2\alpha)\cos[\phi(x,y)]} \tag{5.18}$$

When $\alpha = \frac{\pi}{2}$, the right hand side of Eq. 5.18 reduces to

$$\frac{1}{2\tan\phi(x,y)}$$

Therefore, the phase value at each point in the fringe image can be cal-

culated by

$$\phi(x,y) = \arctan\left[\frac{2(I_2 - I_4)}{2I_3 - I_5 - I_1}\right] \tag{5.19}$$

Since it is not possible for the numerator and denominator of Eq. 5.19 to go simultaneously to zero. The problem associated with Carre's algorithm is avoided.

Hariharan et al [8] showed that the five step algorithm gives very small errors, $\pm 0.02°$, for 2.2% deviations of the phase step $\alpha$ from a nominal value of $\frac{\pi}{2}$. Furthermore, the use of the five step algorithm also reduces substantially the effects of deviations from linearity of phase shifting device.

### 5.2.5 N+1 Phase Step Algorithms

Larkin et al [9] generalized the work of Hariharan et al [8], and introduced a class of phase-stepping algorithms that are called N+1 phase step algorithms. The algorithms have low sensitivity to error caused by phase shifter miscalibration. It can also be used in measurement system with non-sinusoidal periodic fringe. This means that the error introduced by detector nonlinearities is eliminated. The algorithms are characterized by having N+1 samples separated by N equal intervals over whole period. In other words, This algorithms have a phase shift of $2\pi/N$, which is similar to the algorithms described previously, with an extra data having a $2\pi$ phase shift relative to the first data. The N+1 phase step algorithms can be represented by the following equations

$$\phi(x,y) = \arctan\left[\frac{c_0(I_{N+1} - I_1) + \sum_{n=1}^{N-1} \sin\left(\frac{2\pi n}{N}\right) I_{n+1}}{-\frac{I_1 + I_{N+1}}{2} - \sum_{n=1}^{N-1} \cos\left(\frac{2\pi n}{N}\right) I_{n+1}}\right] \tag{5.20}$$

when N is even

$$c_0 = \frac{1}{N} \sum_{n=1}^{(N/2)-1} (N - 2n) \sin\left(\frac{4\pi n}{N}\right) \tag{5.21}$$

and when N is odd

$$c_0 = \frac{1}{N} \sum_{n=1}^{(N-1)/2} (N - 2n) \sin\left(\frac{4\pi n}{N}\right) \tag{5.22}$$

If N = 6, the relative phase shift is equal to $\pi/3$. From Eq. 5.21

$$c_0 = \frac{1}{2\sqrt{3}}$$

Eq. 5.20 therefore becomes to the following.

$$\phi(x,y) = \arctan\left[\frac{\sqrt{3}(I_2 + I_3 - I_5 - I_6) + (I_7 - I_1)/\sqrt{3}}{-I_1 - I_2 + I_3 + 2I_4 + I_5 - I_6 - I_7}\right]$$

### 5.2.6 "2+1" Algorithm

When a large optic is tested, it is sometimes difficult or impossible to move the optic to a test environment. If the optic is tested in a workshop environment, the vibration will be likely to impair the algorithm described previously in this chapter. Wizinowich [10] invented a "2+1" algorithm to alleviate the problem caused by the presence of vibration and air turbulence. For this algorithm, two intensity measurements, $I_1$ and $I_2$, are taken very quickly with a $\pi/2$ phase shift between them. The two intensity measurement can be expressed as

$$\begin{aligned} I_1(x,y) &= I_0(x,y)\{1 + \gamma\cos[\phi(x,y)]\} \\ I_2(x,y) &= I_0(x,y)\{1 + \gamma\cos[\phi(x,y) - \frac{\pi}{2}]\} \end{aligned} \tag{5.23}$$

A third intensity measurement, $I_3$, that is the average of two frames with a $\pi$ phase shift between them provides the dc intensity.

$$\begin{aligned} I_3(x,y) &= \frac{1}{2}\{I_0(x,y)\{1 + \gamma\cos[\phi(x,y)]\} + I_0(x,y)\{1 + \gamma\cos[\phi(x,y) + \pi]\}\} \\ &= I_0(x,y) \end{aligned} \tag{5.24}$$

The phase is then calculated from the following equation.

$$\phi(x,y) = \arctan\left[\frac{I_2 - I_3}{I_1 - I_3}\right]$$

Since $I_3$ can be acquired at any time, only $I_1$ and $I_2$ are time critical. If a CCD video camera is used with a synchronized shutter, two frames with $1ms$ or less exposures can be recorded as quickly as $1\mu s$ apart on either field of

the interline transfer.

## 5.3 The Problem associated with Phase-stepping Techniques

### 5.3.1 Phase Shifter Errors

The basic equation for phase-stepping techniques was given by Bruning et al [3]:

$$\tan\phi(x,y) = \frac{\sum_{r=1}^{R} I_r(x,y)\sin\alpha_r}{\sum_{r=1}^{R} I_r(x,y)\cos\alpha_r} \tag{5.25}$$

where $\phi(x,y)$ is the phase to be measured, $\alpha_r$ is the phase shift, $R$ is the number of phase steps and $I_r(x,y)$ is the intensity distribution in the detector plane in the absence of errors. If a phase-shifter calibration error is present, the phase shift may be written as

$$\alpha_r^{'} = \alpha_r + \varepsilon_r \tag{5.26}$$

The intensity as the measured quantity can be written as

$$I_r^{'}(x,y) = I_0(x,y)\{1 + \gamma\cos[\phi(x,y) + \alpha_r + \varepsilon_r]\} \tag{5.27}$$

$I_r^{'}(x,y)$ is subjected to the measuring algorithm contained in Eq. 5.25. Therefore, the measuring phase $\phi^{'}(x,y)$ differs from the true phase $\phi(x,y)$ by a small deviation :

$$\triangle\phi(x,y) = \phi^{'}(x,y) - \phi(x,y)$$

Furthermore, the following holds

$$\tan\phi^{'}(x,y) = \frac{N}{D} = \frac{\sum_{r=1}^{R} I_r^{'}(x,y)\sin\alpha_r}{\sum_{r=1}^{R} I_r^{'}(x,y)\cos\alpha_r} \tag{5.28}$$

Finally, $\triangle\phi(x,y)$ can be expressed as

$$\begin{aligned} \triangle\phi(x,y) &= \arctan\frac{N}{D} - \arctan(\tan[\phi(x,y)]) \\ &= \arctan\frac{N - D\tan[\phi(x,y)]}{D + N\tan[\phi(x,y)]} \end{aligned} \tag{5.29}$$

Using the orthogonality relations of the sine and cosine function and assuming that $\varepsilon_r$ is small, we can put $\cos\varepsilon_r = 1$ and $\sin\varepsilon_r = \varepsilon_r$. Then the following expression can be derived

$$\Delta\phi(x,y) = \arctan\frac{\sum_{r=1}^{R}\varepsilon_r - C\cos[2\phi(x,y)] - S\sin[2\phi(x,y)]}{R - C\sin[2\phi(x,y)] + S\cos[2\phi(x,y)]} \tag{5.30}$$

where

$$C = \sum_{r=1}^{R}\varepsilon_r\cos 2\alpha_r$$

$$S = \sum_{r=1}^{R}\varepsilon_r\sin 2\alpha_r$$

In Eq. 5.30, the mean value of the phase error

$$\frac{1}{R}\sum_{r=1}^{R}\varepsilon_r$$

preponderates all other terms. It means that the errors in phase resulting from a zero mean calibration error or nonlinearity in the phase shifter will decrease as the number of measurements increases [7].

Nonlinear phase-shifter errors can be reduced by applying certain algorithms such as the Carre algorithm; however, they cannot be eliminated [1].

### 5.3.2 Detector Nonlinearities

A detector's nonlinear response can introduce phase errors. Recently, CCD detector arrays are usually used to recode the fringe image. The phase errors will be obvious, if the CCD detector array are not consistent from detector to detector, or the gains in the two set of shift registers are not equal and nonlinear.

When a second-order nonlinearity is present in the detected intensity, a minimum of four measurements is necessary to obtain an accurate phase calculation [1]. For higher-order nonlinearities, Table 5.1 shows which orders of detection errors will affect the measurement for small numbers of phase steps [11]. The dashes in Table 5.1 indicate which detection nonlinearity orders do not contribute to phase errors in the various algorithms.

| Number of steps | Harmonic order | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| 3 | 2 | - | 4 | 5 | - | 7 | 8 | - | 10 | 11 |
| 4 | - | 3 | - | 5 | - | 7 | - | 9 | - | 11 |
| 5 | - | - | 4 | - | 6 | - | - | 9 | - | 11 |

Table 5.1: The effect of harmonic order on various algorithms

### 5.3.3 Intensity Error

One of various error sources in phase-stepping techniques is intensity error; for example: statistically independent intensity fluctuation and intensity quantization error. Brophy [12] had established an algorithm-dependent relationship between the mean-square error of the computed phase and correlations of intensity error between frames.

If the computed phase, $\phi$, is expanded as a function of many intensity variables in a multivariate Taylor series, the $\phi$ can be expressed as

$$\phi' = \phi + \sum \left( \frac{\partial \phi}{\partial I_n} \right) \triangle I_n + O(\triangle I^2) \tag{5.31}$$

where $O(\triangle I^2)$ is all terms containing products of intensity deviations and $I_n$ is given by Eq. 5.2. If these intensity differences are small, all terms higher than the linear term in intensity difference may be neglected. Furthermore, if the errors in intensity are regarded as statistical quantities, manipulation of Eq. 5.31 gives

$$V_\phi \approx \sum_n \sum_m \left( \frac{\partial \phi}{\partial I_n} \right) \left( \frac{\partial \phi}{\partial I_m} \right) \langle \triangle I_n \triangle I_m \rangle \tag{5.32}$$

where $V_\phi$ is the mean-square error in the computed phase and $\langle \triangle I_n \triangle I_m \rangle$ is a general cross-correlation of intensity deviations among the various frames.

Many phase-stepping algorithms are of the form

$$\phi = \arctan \frac{\sum A_n I_n}{\sum B_n I_n} \tag{5.33}$$

where $I_n$ is the measured intensity distribution and $A_n$ and $B_n$ are coefficients of the specific algorithm. Using Eq. 5.2 for $I_n$ and Eq. 5.33 to evaluate the

derivatives in Eq. 5.32, the mean-square error is approximated by

$$V_\phi \approx \left[\frac{1}{(C\gamma I_0)^2}\right] \sum_n \sum_m (A_n \cos\phi - B_n \sin\phi) X (A_m \cos\phi - B_m \sin\phi) \langle \triangle I_n \triangle I_m \rangle \tag{5.34}$$

where

$$C = (\sum B_n I_n)^2 + (\sum A_n I_n)^2$$

is algorithm-dependent coefficient that is independent of $\phi$.

Brophy [12] also pointed out an interesting difference between the case of quantization error and the case of uncorrelated intensity fluctuation. In the case of uncorrelated intensity fluctuation, the more measurements that are made, the better the statistical averaging. On the other hand, for quantization error, the phase mean-square error is independent of the number of frames.

## 5.4 Application Example of Phase-stepping Techniques

This application example, which is stored in the example database of the optical engineering group at the University of Warwick, records the modes of vibration of a plane sheet caused to vibrate with an impinging air jet. This example was originally recorded in a dual reference beam holographic system. The experiment is fully described in reference [13].

The three step algorithm was applied with a phase step $\frac{2\pi}{3}$ between the images. Figure 5.1 - 5.3 show the three interferograms recorded in the experiment. In a three phase algorithm, Eq. 5.8 is employed to compute the phase, $\phi(x, y)$, at a given pixel. A low modulation point is identified when, for the quantity under the arctan, both the numerator and denominator of Eq. 5.8 are small. The critical size for the numerator and denominator is empirically determined. Figure 5.4 shows the detected low modulation points. Obviously, this is a successful method of detecting bad data points for phase-stepping techniques. Figure 5.5 shows the wrapped phase map for the vibrating board. The wrapped phase map is unwrapped by the minimum spanning tree unwrapping algorithm which is described in Chapter 7. The

Figure 5.1: First holographic interferogram of vibrating board

Figure 5.2: Second holographic interferogram of vibrating board

Figure 5.3: Third holographic interferogram of vibrating board

edge detection and tile connection tree is shown in Figure 5.6. Figure 5.7 shows the mesh plot of unwrapped phase for the vibrating board. Any tile with more than 25 % of their area consisting of low modulation points has been rejected during the analysis.

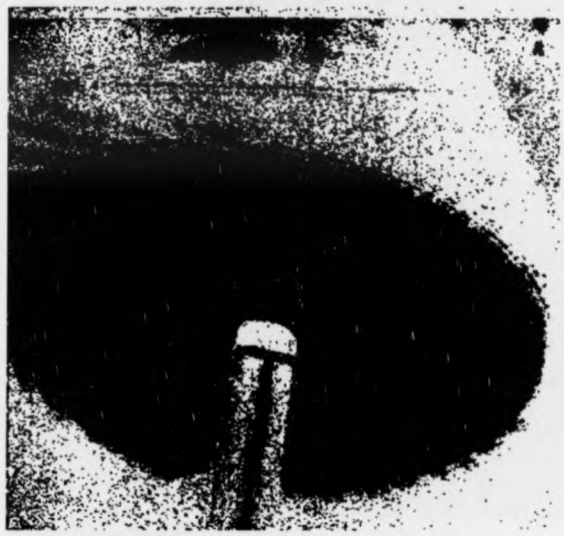


Figure 5.4: The detected low modulation points for vibrating board

Figure 5.5: The wrapped phase map for vibrating board

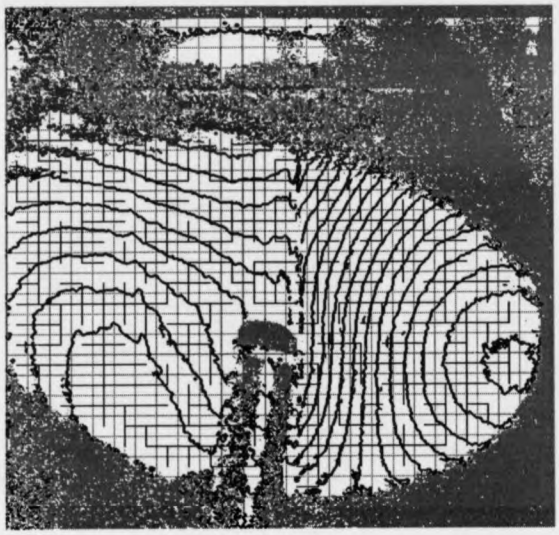


Figure 5.6: The tiles and connection of the wrapped phase map for vibrating board

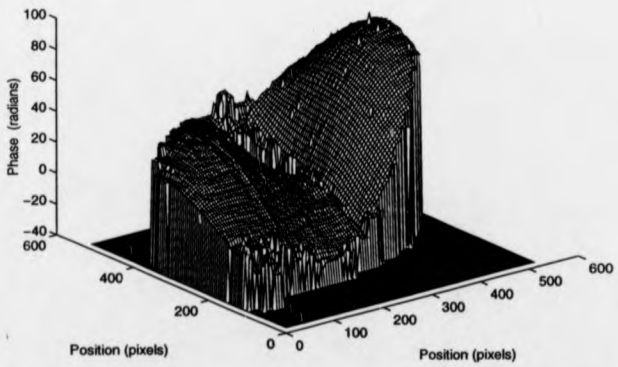



Figure 5.7: The mesh plot of unwrapped phase for vibrating board

# Bibliography


[1] K. Creath, "Phase-measurement Interferometry Techniques", Prog. Opt. 26, p. 351-393, 1988.

[2] N. A. Massie, R. D. Nelson and S. Holly, " High-performance Real-time Heterodyne Interferometry", Appl. Opt., Vol. 18, p. 1797 - 1803, 1979.

[3] J. H. Bruning, D. R. Herriott, J. E. Gallagher, D. P. Rosenfeld, A. D. White and D. J. Brangaccio, "Digital Wavefront Measuring Interferometer for Testing Optical Surfaces and Lenses", Appl. Opt.,Vol. 13, p. 2693-2703, 1974.

[4] J. E. Greivenkamp, "Generalized Data Reduction for Heterodyne Interferometry", Opt. Eng., Vol. 23, p. 350-352, 1984.

[5] G. O. Rosvold, "Fast Measurements of Phase Using a PC-based Frame Grabber and Phase Stepping Technique", Appl. Opt., Vol. 29, p. 237-241, 1990.

[6] K. Creath,"Phase-shifting Speckle Interferometry", Appl. Opt., Vol. 24, p. 3053-3058, 1985

[7] J. Schwider, R. Burow, K-H. Elssner, J. Grzanna, R. Spolaczyk and K. Merkel,"Digital Wave-front Measuring Interferometry: Some Systematic Error Sources", Appl. Opt., Vol 22, p. 3421 - 3432, 1983.

[8] P. Hariharan, B. F. Oreb and T. Eiju, "Digital Phase-shifting Interferometry: A Simple Error-compensating Phase Calculation Algorithm", Appl. Opt., Vol. 26, p. 2504 - 2506, 1987.

[9] K. G. Larkin and B. F. Oreb, "Design and Assessment of Symmetrical

Phase-Shifting Algorithms", J. Opt. Soc. Am. A, Vol. 9, p. 1740 -1748, 1992.

[10] P. L. Wizinowich, "Phase Shifting Interferometry in the Presence of Vibration : A New Algorithm and System", Appl. Opt., Vol. 29, p. 3271 - 3279, 1990.

[11] K. A. Stetson and W. R. Brohinsky,"Electrooptic Holography and its Application to Holographic Interferometry", Appl. Opt., Vol. 24, p. 3631-3637, 1985.

[12] C. P. Brophy,"Effect of Intensity Error Correlation on the Computed Phase of Phase-shifting Interferometry", J. Opt. Soc. Am. A, Vol 7, p. 537-541, 1990.

[13] D. P. Towers, T. R. Judge and P. J. Bryanton-Cross,"Automatic Interferogram Analysis Techniques Applied to Quasi-heterodyne Holography and ESPI", Opt. and Lasers Eng., Vol. 14, p. 239 - 281, 1991.

# Chapter 6

# Fringe Pattern Analysis by Spatial Phase-stepping Technique

## 6.1 Introduction

Phase-stepping (PS) techniques and discrete Fourier transform (DFT) techniques are two main groups of techniques for quantitative phase measurement from fringe patterns. Each technique has its inherent merits and drawbacks. Selection of the analytic technique should be depended on the circumstance of individual case.

The DFT techniques are operated globally and based on a spatial carrier approach. The main advantage of DFT techniques is that only one fringe image is required. This is essential for analysing dynamic events or performing measurement in adverse conditions. Another asset is that DFT techniques do not require a special phase-shifter. However, in interferometric measurement, additional systematic errors will be caused by unwanted aberrations because of the high frequency carrier fringe. The PS techniques which are based on a temporal carrier approach measure the phase of fringe patterns by using $N$ frames of images which is taken while the reference wavefront is shifted by $2\pi/N$ between successive frames. The main advantage of PS techniques is that its computing processing is simple and fast. This enables the analysis to be completed in one frame rate. There is also an image size advantage to

PS techniques. The PS techniques have a time complexity of $O(n)$; however, the DFT techniques have a time complexity of $O(n^2)$,where $n$ is the number of pixels in the fringe image. This implies that the PS technique is suitable for analysing high-resolution fringe images. Another advantage is that low modulation points can be easily detected in PS techniques. The disadvantages are that the PS techniques require additional calibration procedure and a stable experimental environment. A detailed comparison between the two techniques can be found in reference [1, 2, 3].

Recently, research work has focused on developing an analytical method which combines the computing simplicity of PS techniques with the dynamic event analysis capability of DFT techniques. Firstly, previous research work in this area is reviewed. Then, a novel Spatial Phase-stepping technique is presented. As an application example, the technique is employed to analyse an interferometric fluid flow measurement.

## 6.2 Previous Research Works

### 6.2.1 Spatial Version of the Phase-stepping Technique

Fig. 6.1 (a) exemplifies the temporal version of the phase-stepping technique described in Chapter 5. In this method, $N$ fringe images with a phase shift of $2\pi/N$ are separated at $N$ discrete time intervals. On the contrary, the spatial version of phase-stepping technique simultaneously acquires $N$ phase shifted fringe images in $N$ discrete spatial locations. Fig. 6.1 (b) illustrates this concept. There are three principal optical methods to generate simultaneous phase shifted fringe images: (i) polarisation techniques, (ii) colour separation techniques and (iii) diffraction grating techniques. These techniques are now described briefly.

In polarisation techniques [4, 5], fixed polarisation elements that are orientated properly are inserted into separate output channels to generate the required phase shift simultaneously. The three-channel polarisation interferometer [5] is described to explain the basis concept. Fig. 6.2 depicts the optical arrangement of the interferometer. The three-channel polarisation interferometer can be considered as an instrument that is composed of two

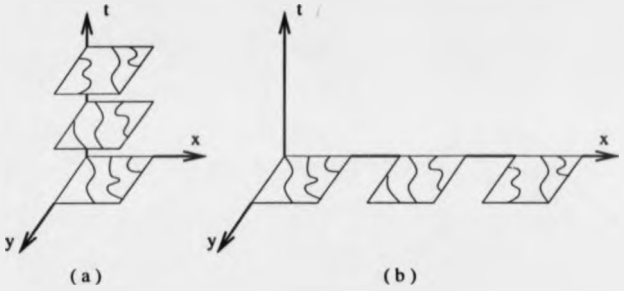



Figure 6.1: Schematic respresentation of temporal (a) and spatial (b) version of phase-stepping technique

parts: the polarisation interferometer and the interference channels. The polarisation interferometer, which is shown in the left part of Fig. 6.2, generates collinear reference and signal wavefronts of orthogonal polarisation states. Then, the light is steered to the interference channels. The input light is split by non-polarising beamsplitters into three channels. Quarter-wave retarders, that are oriented in parallel to the polarisation axis, shift the relative phase by $\pi/2$ and $\pi$ in channel C1 and C2, correspondingly. In each channel, a polariser oriented at 45 deg to the polarisation axis mixes the wavefront polarisations, and the resulting fringe pattern is detected by CCD cameras. The phase can be calculated by three phase step algorithm.

Harding et al [6] suggested a method to obtain simultaneously three images by the use of RGB colour camera viewing gratings with three different colours. Fig. 6.3 shows the colour projection moire system proposed by Harding et al [6]. The system is a modified version of the standard projection moire system. A three-colour grating is employed as a master grating while the submaster grating is an ordinary black and white one. This configuration produces a set of three phase-shifted moire patterns while a white light source is used. The moire fringe patterns are then separated and recorded simultaneously using a solid state RGB camera.

The diffraction grating technique utilizes the ability of a grating to diffract and phase-shift an illuminating beam. Combining with the principle of phase-stepping interferometry, the grating technique had been applied to different types of interferometer [7]. Recently this technique had also been used in

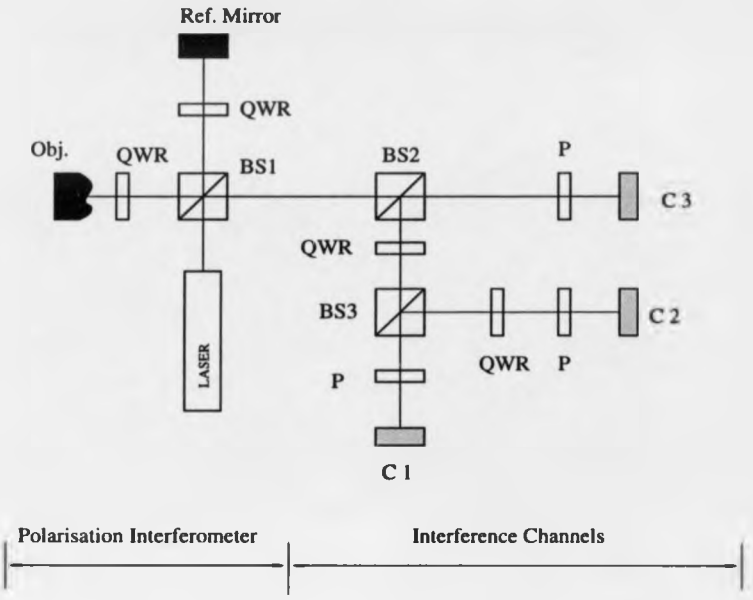



Figure 6.2: The three-channel polarisation interferometer; BS 1-3, beam-splitter; P, polarisator; QWR, quarter-wave retarder; C 1-3, CCD cameras

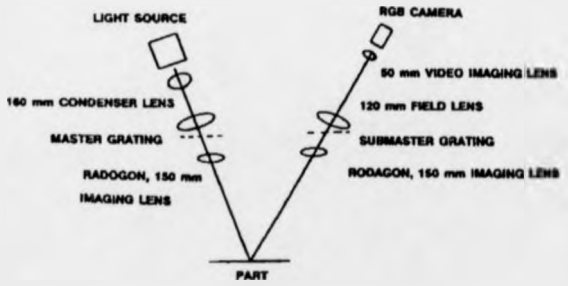



Figure 6.3: The colour projection moire system with RGB camera [6]

holographic interferometers [8]. Fig. 6.4 shows the schematic diagram of a multi-channel radial-shear interferometer [9]. A set of an afocal telescope and a diffraction grating is added to each arm of beam paths to generata multiple reference and test beams. The radial-shear ratio is determined by the combination of line frequencies of gratings and focal lengths of re-collimating lenses. Translating the grating 2 laterally by $x_g$ with respect to the optical axis, a phase-shift is introduced to each of the higher order diffracted beams. Three phase-shifted interferograms are simultaneously recorded, and then analyzed.

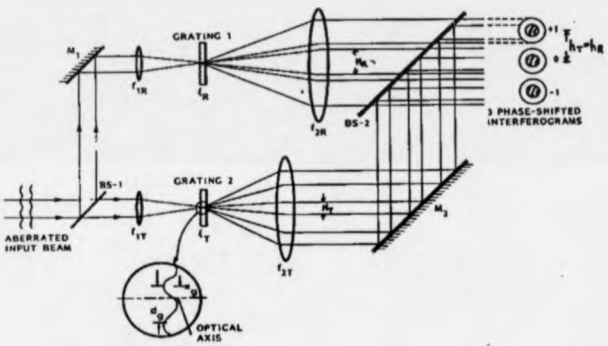



Figure 6.4: The multi-channel radial-shear interferometer [9]

### 6.2.2 Spatial-carrier Phase Stepping Technique

Shough et al [10] proposed a technique which is related to both the Fourier transform technique and phase stepping technique of computing the phase to analyse an aerodynamic phenomena, and called it spatial phase-shifting interferometry. Kujawinska et al [11] extended the phase-shifting interferometry to analyse the general fringe image and called Spatial-carrier phase shifting (SCPS) technique.

In Spatial-carrier phase shifting technique, a large amount of tilt is introduced to obtain approximately $\pi/2$ phase shift between successive pixels:

$$I_i(x,y) = a(x,y) + b(x,y)\cos\{2\pi f_0[x + \frac{i}{4f_0}] + \phi(x,y)\} \qquad (6.1)$$

where $f_0$ is the spatial carrier frequency in the fringe image. Then, three sequential pixels are used to calculate the phase according to the standard three phase-stepping formula:

$$\phi''(i,j) = \arctan\left[\frac{I_3 - I_2}{I_1 - I_2}\right] \tag{6.2}$$

where $(i,j)$ is the coordinate $(x,y)$ given in number of pixel and $I_1 = I(i-1,j)$, $I_2 = I(i,j)$ and $I_3 = I(i+1,j)$ are intensities in three sequential pixel in line.

Then, the previously introduced tilt is removed by

$$\phi'(i,j) = \phi''(i,j) - \frac{i\pi}{2} \tag{6.3}$$

In order to obtain the wrapped phase map, arctan has to be used again:

$$\phi(i,j) = \arctan\left[\frac{\sin[\phi'(i,j)]}{\cos[\phi'(i,j)]}\right] \tag{6.4}$$

Finally, the unwrapping procedure is used to obtain the contiguous phase function.

The basic assumptions made in this method are: the proper carrier-frequency, the sinusoidal profile of the analysed fringe and slowly varying phase, so that it may be assumed that its value is constant within three sequential pixel. Kujawinska et al [11] pointed out that the use of the three phase-stepping algorithm suffers from two types of systematic errors : one is that an envelope error function modulating with the double frequency of the phase fringe; another is that a high frequency error modulating with the frequency of the fringe pattern.

The high frequency error may be significantly removed by using a 3x3 median filter on an unwrapped phase or using a five phase-stepping algorithm [12]. The last solution is very efficient because the algorithm gives the phase value at the centre of the group of pixels and it is not significantly affected by moderate departures of the fringe spacing from its nominal value.

## 6.3 Spatial Phase-stepping Technique

In the SCPS technique, the amount of tilt has to be introduced very accurately so that the phase shift between successive pixels is equal to $\pi/2$. Otherwise, an additional linear phase term will be caused in the retrieved phase. Although the linear phase term can be cancelled by subtraction, the value of the linear term in the retrieved phase has to be found by a complicated curve fitting processing. This section presents a novel spatial phase-stepping (SPS) technique which can be applied to the fringe image with any spatial carrier frequency. Furthermore, this technique removes the introduced tilt automatically.

For a fringe image with spatial carrier, its intensity can be described as

$$I(x,y) = I_0(x,y)\left\{1 + \gamma\cos\left[\frac{2\pi x}{P} + \phi(x,y)\right]\right\} \tag{6.5}$$

Where $\phi(x,y)$ contains the desired information and $I_0(x,y)$ represents intensity caused by non-uniform reflection by the test object. $P$ is the wave length of the spatial carrier, which is independently determined. After the fringe image is digitized, its intensity can be described by the following discrete equation.

$$I(m,n) = I_0(m,n)\left\{1 + \gamma\cos\left[\frac{2\pi m}{P} + \phi(m,n)\right]\right\} \tag{6.6}$$

where

$$m = 0, 1, \cdots, N-1 \quad n = 0, 1, \cdots, N-1$$

The resolution of the digitized interferogram is $N$ by $N$ pixels. The interferogram may be split into three images in the following manner:

$$\begin{aligned} I_1(m,n) &= I\,(0 + mP, n) \\ &= I_{01}(m,n)\,\{1 + \gamma\cos\,[0 + \phi_1(m,n)]\} \qquad (6.7) \\ I_2(m,n) &= I\left[int\left(\frac{P}{3}\right) + mP, n\right] \\ &= I_{02}(m,n)\,\{1 + \gamma\cos\,[\beta + \phi_2(m,n)]\} \qquad (6.8) \\ I_3(m,n) &= I\left[int\left(\frac{2P}{3}\right) + mP, n\right] \\ &= I_{03}(m,n)\,\{1 + \gamma\cos\,[2\beta + \phi_3(m,n)]\} \qquad (6.9) \end{aligned}$$

where

$$m = 0, 1, \cdots, int\left(\frac{N}{P}\right) \quad n = 0, 1, \cdots, N-1;$$

$int(x)$ returns the integer value of $x$; and

$$\beta = \frac{2\pi}{P}\, int\left(\frac{P}{3}\right)$$

Based on equations 6.7 to 6.9, $I_1(m,n)$, $I_2(m,n)$ and $I_3(m,n)$ can be treated as three separate phase stepped fringe images of a test object, with phase distribution $\phi'(m.n)$ and phase shift $\beta$, provided :

$$\phi'(m.n) \approx \phi_1(m,n) \approx \phi_2(m,n) \approx \phi_3(m,n) \tag{6.10}$$

As a consequence, $\phi'(m.n)$ can be calculated by the three phase step algorithm. The SPS technique is based on the following assumptions:

1. The profile of the analysed fringe is sinusoidal.
2. The resolution of the fringe image is high and the phase varies slowly, so that Eq. 6.10 is valid.
3. The spatial carrier frequency is high. Usually, the wave length of the spatial carrier will not be greater than ten pixels.

The SPS technique utilizes a line of pixels, distributed over the whole wavelength of the spatial carrier, to calculate the phase value of the centre pixel. Obviously, the resolution of computed phase in x direction is reduced by a factor of the spatial carrier wavelength. Although the three-phase step algorithm is used to illustrate the SPS technique, any phase step algorithm can be employed to calculate the phase value.

## 6.4 The Error Analysis

Computer simulation experiments have been undertaken to demonstrate the theory and to analyse systematic errors. A number of computer generated interferograms with various phase profiles are used. All computer generated interferograms are sampled by 1024 X 1024 pixel and quantised by 8 bits.

For all error analyses in this section, it is assumed that the fringe visibility is 100%, i.e. $\gamma = 1$, the average intensity, $I_0$, is constant and the wavelength of the spatial carrier is equal to six pixel. The phase distribution of the computer generated interferogram is expressed as

$$\phi(x,y) = \begin{cases} f(x,y) & \text{if } f(x,y) \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{6.11}$$

where

$$f(x,y) = 2M\pi \left[1 - \left(\frac{x-1024}{1024}\right)^2 + \left(\frac{y-1024}{1024}\right)^2\right]$$

and M is a dimensionless constant.

The first analysis is tested on a computer generated interferogram with $M = 3$. The three split images are shown in Fig. 6.5. The split images are analysed by FRANSYS using phase-stepping option. The computed wrapped phase map is shown in Fig. 6.6. The error map was obtained by subtracting the retrieved phase from the theoretical phase value. The contour plot of phase errors is shown in Fig. 6.7. The Maximum phase error was found to be approximately $\pm 1/70th$ of a fringe. Fig. 6.11 and 6.13 show the theoretical phase value and error map of raster line 512 and 256 respectively.

In order to investigate the effect of phase profile, another analysis is conducted on a computer generated interferogram with maximum phase value of $3\pi$, i.e. $M = 1.5$. It means that the peak-to-valley (P-V) value of phase distribution is reduced to half of the previous analysis. The split images are shown in Fig. 6.8. The spilt images are also analysed by FRANSYS using phase-stepping option. The computed wrapped phase map is shown in Fig. 6.9. This calculation, with the peak-to-valley (P-V) value of phase error reduced by a factor of two, produces maximum phase errors of $\pm 1/140th$ of a fringe. Fig. 6.10 shows the contour plot of phase errors produced in the analysis. It is found that the profile of error map of this analysis is similar to that of the previous analysis. Fig. 6.12 and 6.14 show the theoretical phase value and error map of raster line 512 and 256 in this analysis respectively.

Referring to Fig. 6.13 (a) and 6.14 (a), the rate of phase change at the edge of the field-of-view of raster line 256 are zero, for both phase profiles. The errors, which are shown in Fig 6.13 (b) and 6.14 (b), correlated with the

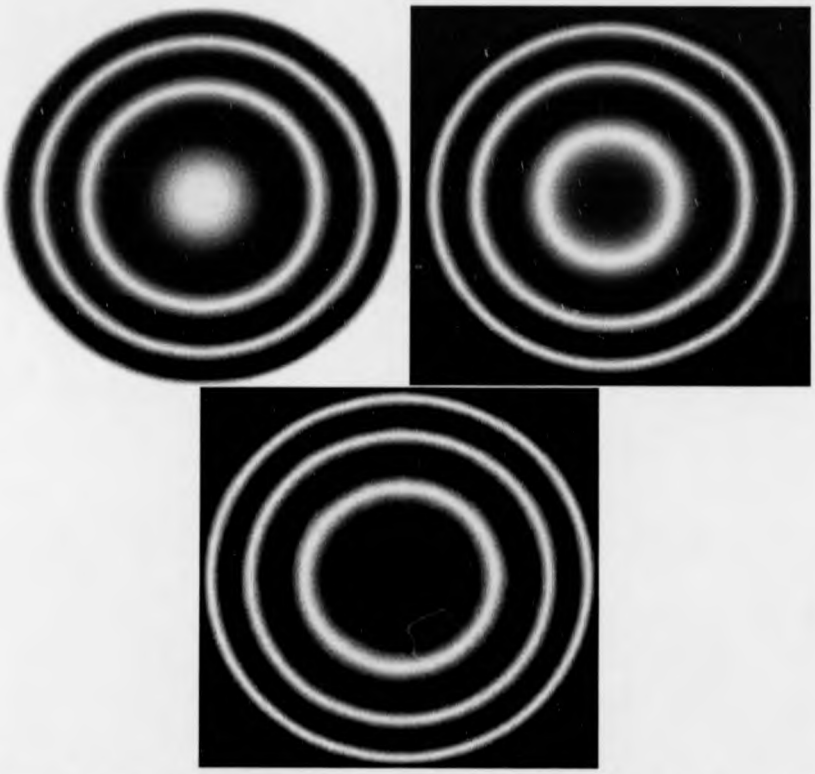


Figure 6.5: The three images split from the interferogram with M=3 (pixel samples in steps of 6 in y axis)

Figure 6.6: The wrapped phase map of the interferogram with M=3 (pixel samples in steps of 6 in y axis)

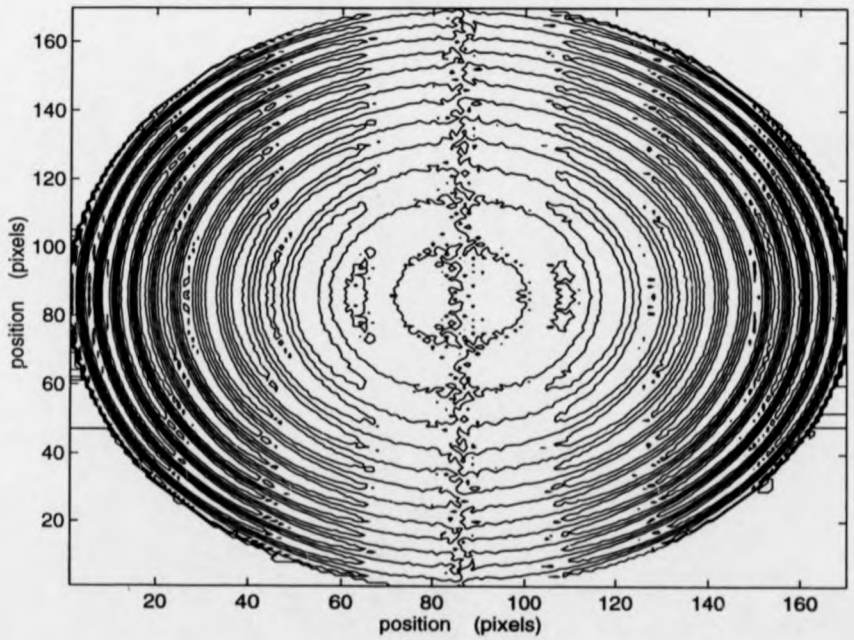



Figure 6.7: The contour plot of phase error of the interferogram with M=3

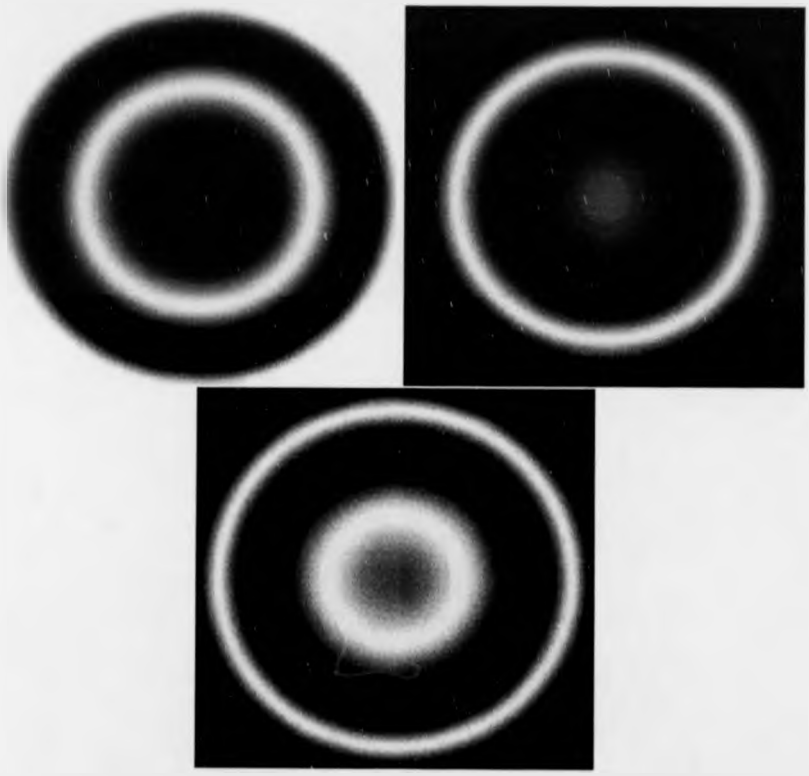


Figure 6.8: The three images split from the interferogram with M=1.5 (pixel samples in steps of 6 in y axis)

Figure 6.9: The wrapped phase map of the interferogram with M=1.5 (pixel samples in steps of 6 in y axis)

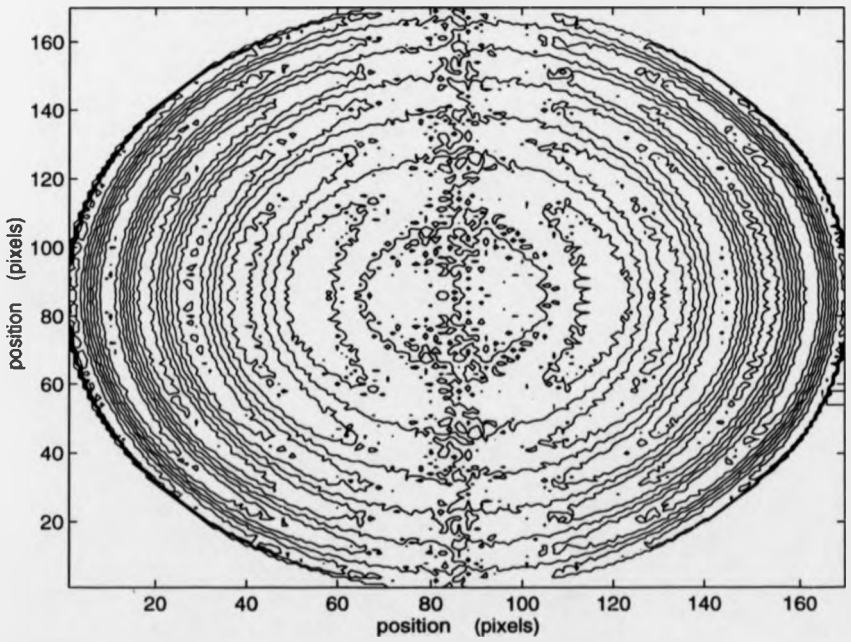



Figure 6.10: The contour plot of phase error of the interferogram with M=1.5

same region are also zero in both analysis. In addition, it is observed that the error associated with the central turning point of phase profile of each raster line is zero. It is known that the slope of the turning point of a curve is zero. A comparison of the absolute phase errors for both phase profiles shows that the error is doubled when the P-V phase excursion is doubled. Furthermore, the maximum error occurs in the region where the phase distribution varies rapidly. From these observations, it can be speculated that the absolute phase error is directly proportional to the gradient of the phase surface across the whole spatial carrier wave.

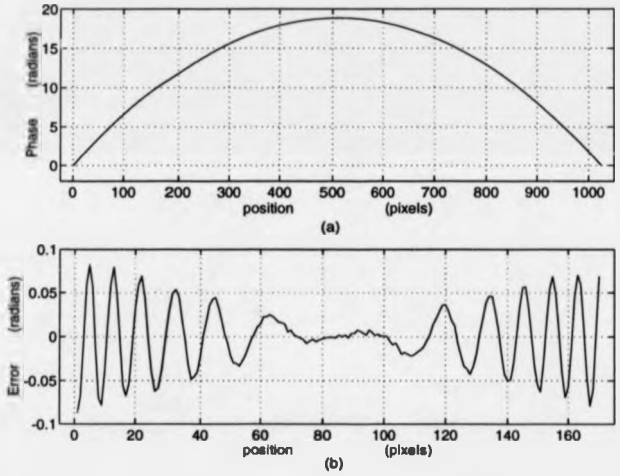



Figure 6.11: The error analysis of the interferogram with M=3 a) The theoretical phase value of raster line 512, b) The error map of raster line 512

## 6.5 Application Examples of the Spatial Phase-stepping Technique

Two examples are now presented to demonstrate the potential of spatial phase stepping in fluid flow measurement. Initially, the interferograms were recorded using double-exposure pulsed holography. The interferometric images were subsequently generated in a conventional holographic reconstruc-

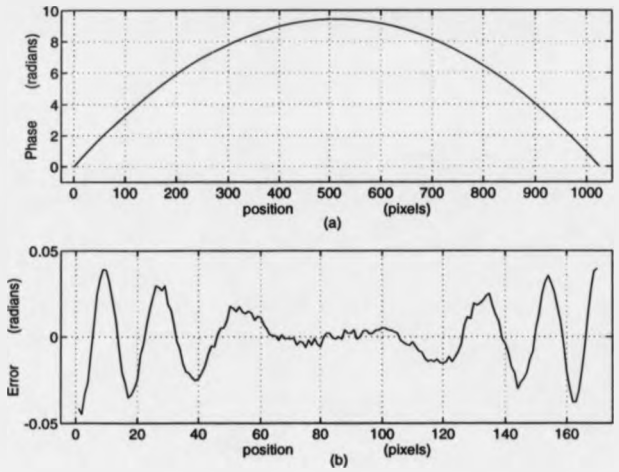



Figure 6.12: The error analysis of the interferogram with M=1.5 a) The theoretical phase value of raster line 512, b) The error map of raster line 512

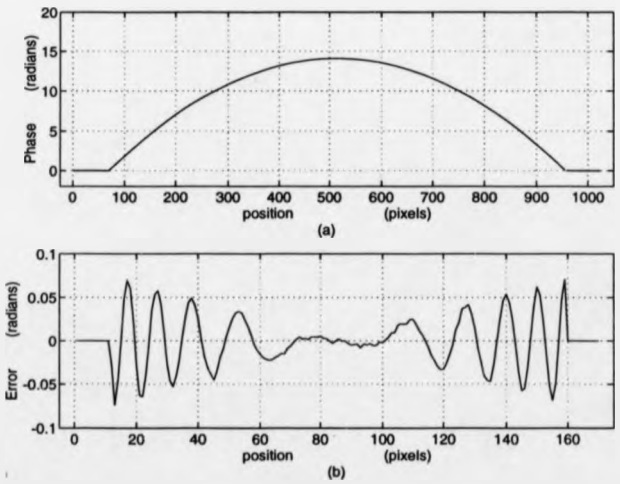



Figure 6.13: The error analysis of the interferogram with M=3 a) The theoretical phase value of raster line 256, b) The error map of raster line 256

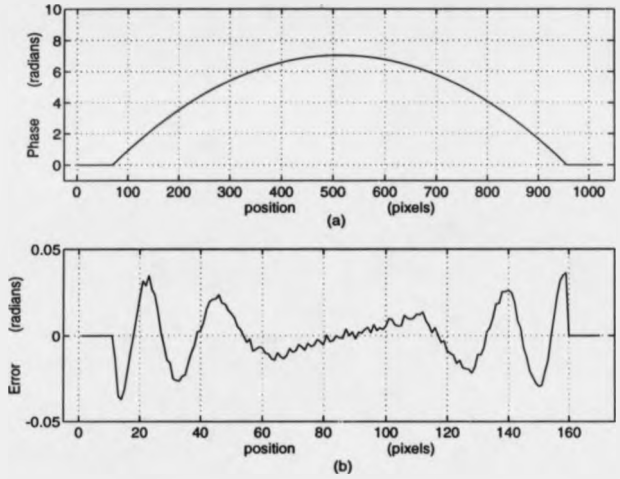



Figure 6.14: The error analysis of the interferogram with M=1.5 a) The theoretical phase value of raster line 256, b) The error map of raster line 256

tion system and imaged onto photographic film. In both cases, the spatial carrier frequency, upon which the data was encoded, was determined by measuring the frequency at which the power spectrum reached a maximum. The resolution in the y direction is also reduced by the same factor as in the x direction.

Fig. 6.15 shows an interferogram of two-dimensional transonic flow around a DCA-Profile (cord length = 75.7 mm). The original hologram was recorded at EPFL Lausanne. The experiment is described in detail in reference [13]. The photographic image was digitized to a resolution of 1762 × 1776 pixels and the wavelength of the spatial carrier wave was found to be seven pixels. The split phase stepped images, derived from the single input image, are shown in Fig. 6.16. Fig. 6.17 shows the wrapped phase map computed by spatial phase stepping. The SPS technique required 4 minutes 14 seconds CPU time to process on a Sun Sparc 1 workstation. Fig. 6.15 was also analysed by the DFT technique and, in comparison, the processing took 34 minutes 57 seconds. The wrapped phase map retrieved by the DFT technique is shown in Fig. 6.18. In order to facilitate the comparison, Fig. 6.19 and Fig. 6.20 show the contour plot of unwrapped phase retrieved by SPS and

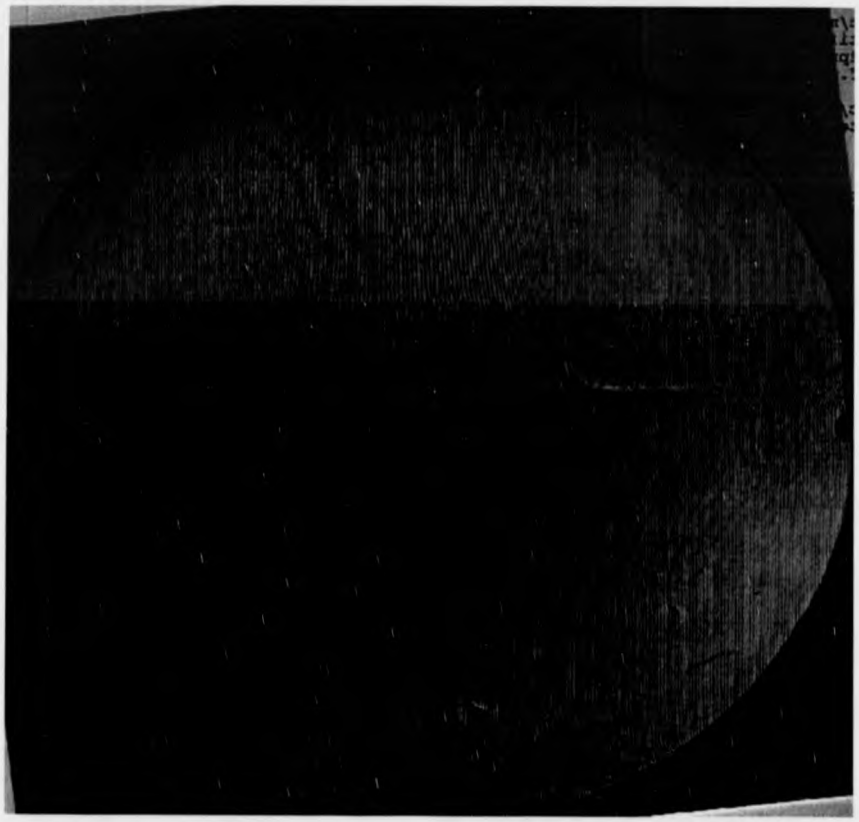


Figure 6.15: The interferogram of a 2-D transonic flow field

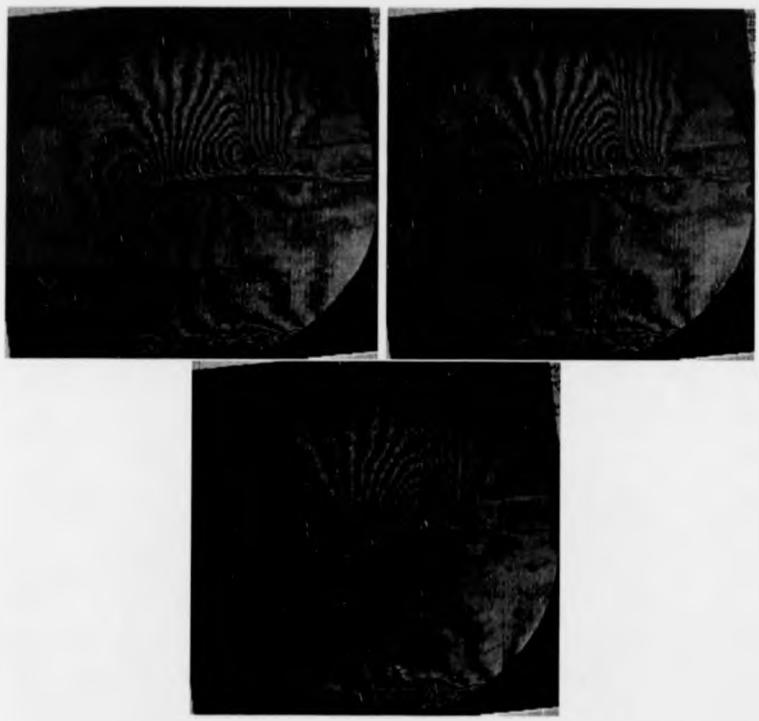


Figure 6.16: The three images split from the interferogram of a 2-D transonic flow field

DFT technique, correspondingly.

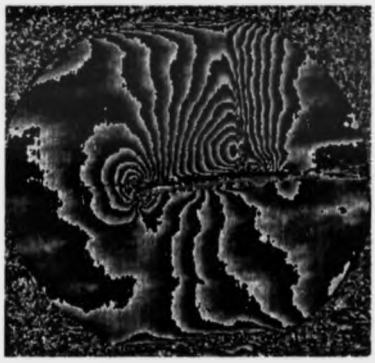


Figure 6.17: The wrapped phase map of 2-D transonic flow field computed by SPS technique

Figure 6.18: The wrapped phase map of 2-D transonic flow field computed by DFT technique

The second example concerns the analysis of transonic flow around a NACA 0012 aerofoil at a free stream Mach number of 0.80. The image, which is shown in Fig. 6.21, was recorded by the Sowerby Research Centre, British Aerospace PLC, and digitized at a resolution of 1500 dpi to form a 3852×3860 pixels image.

The wavelength of the spatial carrier wave was calculated as twelve pixels and the subdivided phase stepped images are shown in Fig. 6.22. Fig. 6.23

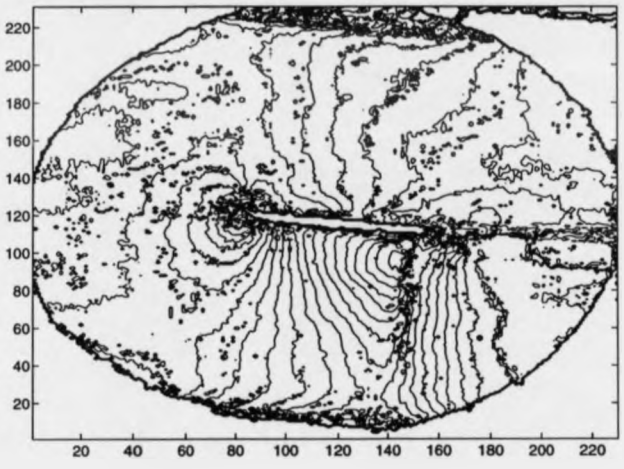



Figure 6.19: The contour plot of unwrapped phase of 2-D transonic flow field retrieved by SPS technique

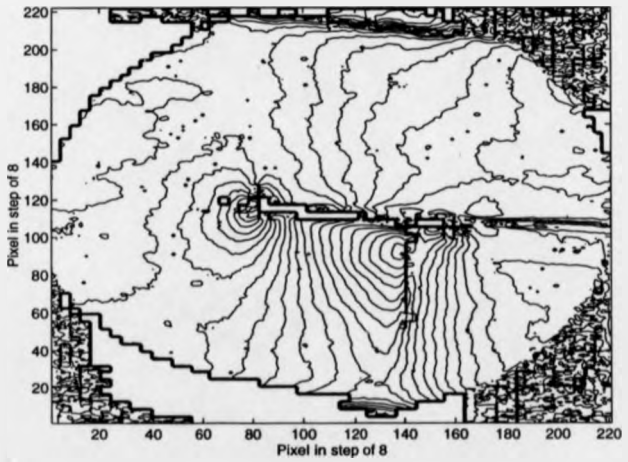



Figure 6.20: The contour plot of unwrapped phase of 2-D transonic flow field retrieved by DFT technique

shows the wrapped phase map computed by spatial phase stepping. The SPS technique and phase unwrapping required 36 seconds CPU time to process on a Sun Sparc 1 workstation. If the 3852x3860 pixels image is directly analysed by DFT method, the computing time will be excessively long. Furthermore, no information will be lost, if the image is digitized at a lower sampling rate on condition that the sampling rate is higher than the Nyquist rate. Therefore, in order to make a comparison between the SPS and DFT method, the image was again digitized at a lower resolution of 600 dpi, which is just above the minimum sampling rate. The low resolution image was then analysed by the DFT method. There are two main reasons for using a low resolution image for DFT analysis: firstly, the shortest computing time required by the DFT method can be obtained; secondly, the resolution difference between two methods is shortened so that the quality of the results retrieved by two methods can be fairly compared. The wrapped phase map of the low resolution image retrieved by the DFT technique is shown in Fig. 6.24. The processing, which includes phase unwrapping, took 13 minutes 55 seconds.

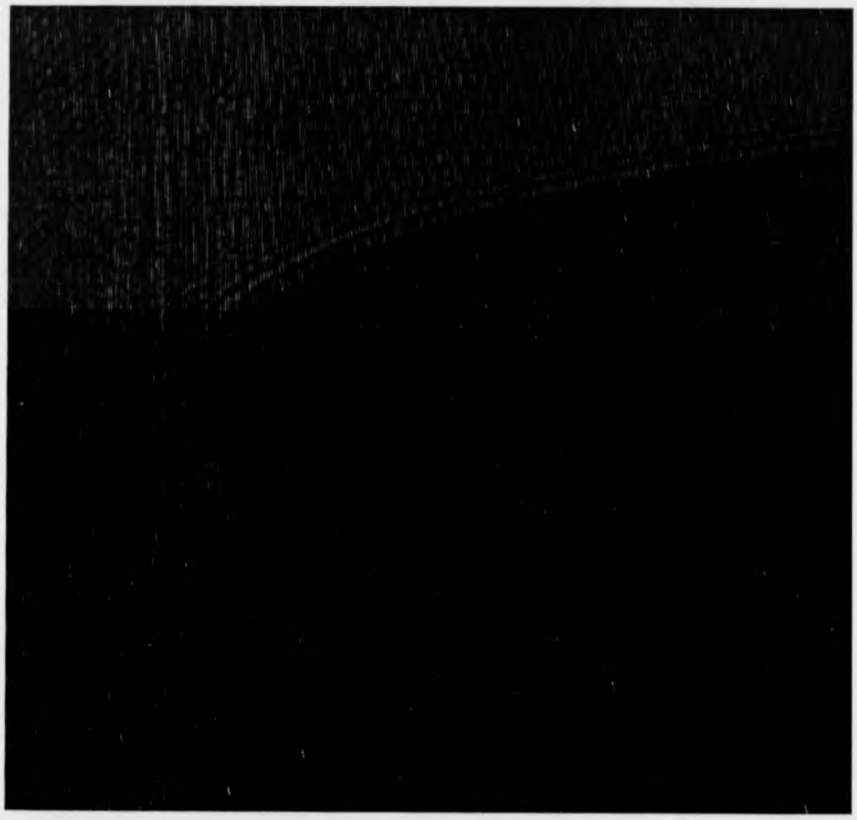


Figure 6.21: The interferogram of transonic flow around a NACA 0012 aerofoil at Mach 0.8

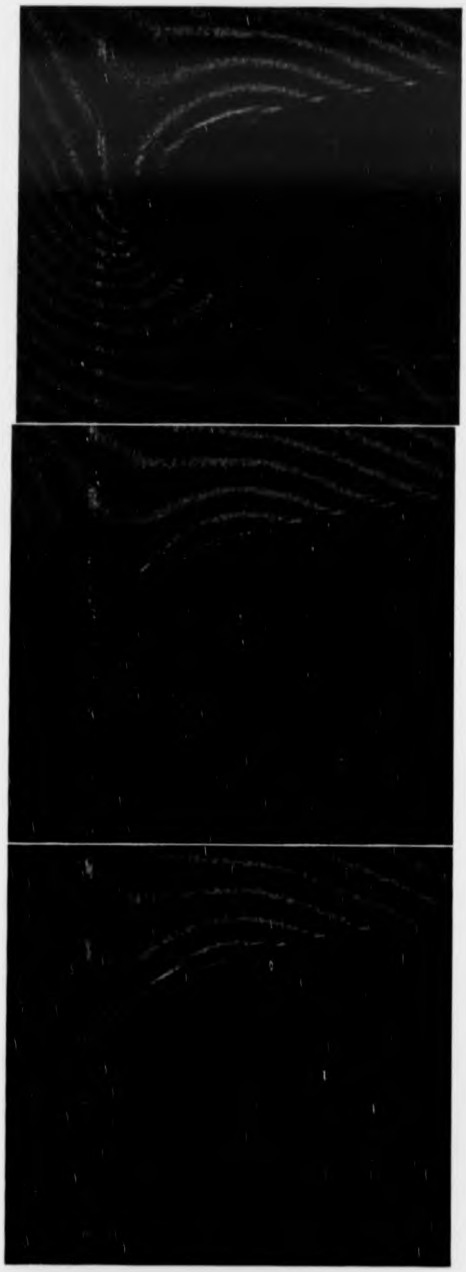


Figure 6.22: The three images subdivided from Figure 6.21

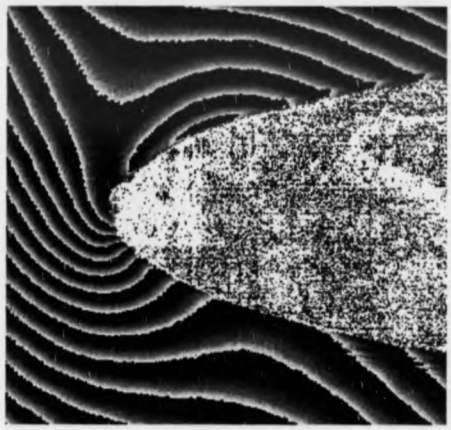


Figure 6.23: The wrapped phase map of Figure 6.21 computed by SPS technique

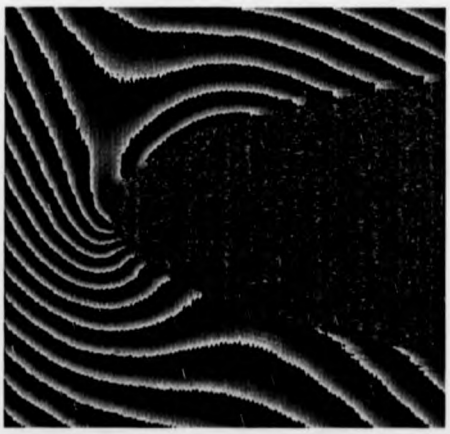


Figure 6.24: The wrapped phase map of the low resolution image computed by DFT technique

# Bibliography


[1] M. Takeda, "Temporal Versus Spatial Carrier Techniques for Heterodyne Interferometry", Proc. SPIE, Vol. 813, p.329-330, 1987.

[2] J. Schmit, K. Creath and M. Kujawinska, "Spatial and Temporal Phase-measurement Techniques: A Comparison of Major Error Sources in One-dimension", Proc. SPIE, Vol. 1755, p. 202-211, 1992

[3] R. Jozwicki, M. Kujawinska and L. Salbut, "New Contra Old Wavefront Measurement Concepts for Interferometric Optical Testing", Opt. Eng., Vol. 31, p. 422-433, 1992.

[4] R. Smythe and R. Moore, "Instantaneous Phase Measuring Interferometry", Opt. Eng., Vol. 23, p. 361 - 364, 1984.

[5] N. Bareket, "Three-channel Phase Detector for Pulsed Wavefront Sensing", Proc. SPIE, Vol. 551, p. 12 - 16,1985.

[6] K. G. Harding, M. P. Coletta and C. H. VanDommelen, "Color Encoded Moire Contouring", Proc. SPIE, Vol. 1005, p. 169 - 179, 1988.

[7] O.Y. Kwon and D.M. Shough, "Multichannel Grating Phase-shift Interferometers", Proc. SPIE, Vol. 599, p. 273 - 279, 1985.

[8] M. Kujawinska and W. Robinson, "Multichannel Phase-stepped Holographic Interferometry", App. Opt., Vol. 27, p. 312 - 320, 1988.

[9] O.Y. Kwon, "Real-time Radial Interferometer", Proc. SPIE, Vol. 551, p. 32 - 34, 1985.

[10] D. M. Shough, O. Y. Kwon and D. F. Leary,"High-speed Interferometric Measurement of Aerodynamic Phenomena", Proc. SPIE, Vol. 1221, p. 394-403, 1990.

[11] M. Kujawinska and J. Wojciak,"Spatial-carrier Phase Shifting Technique of Fringe Pattern Analysis", Proc. SPIE, Vol. 1508, p. 61-67,1991.

[12] D. C. Williams, N. S. Nassar, J. E. Banyard and M. S. Virdee,"Digital Phase-step Interferometry: a Simplified Approach", Optics & Laser Technology, Vol. 23, p. 147-150, 1991.

[13] P. Bryanston-Cross, C. Chenggen and M. Wehner, "The Automatic Numerical Processing of Interferometric Transonic Flow Measurements", Proc. SPIE, Vol. 2176 p.262-271 (1994)

# Chapter 7

# The Minimum Spanning Tree Approach to Phase Unwrapping

## 7.1 Introduction

All the techniques described in pervious chapters produce a result that is wrapped into the range $-\pi$ to $\pi$. The result is called a wrapped phase map. The wrapped phase map can be displayed by using a grey scale to represent phase. Fig. 7.1 shows an example of the wrapped phase map, in which black is $-\pi$ and white represents $\pi$. The aim of the phase unwrapping is to restore the unknown multiple of $2\pi$ to all pixels so that the physical quantity can be deduced from the phase distribution. In order to produce an automatic phase unwrapping system, the phase unwrapping strategy must be able to cope with the following serious problems.

1. Points of low modulation, which are produced from quantization of intensity measurement and may result in a large error in the measurement of phase.

2. Points of noise, caused by speckle or random noise.

3. Discontinuities, resulting from gaps in the fringe field, the edges of objects, or overlapping objects.

4. Ambiguous areas, resulting from an insufficient pixel resolution to resolve all of the fringe.

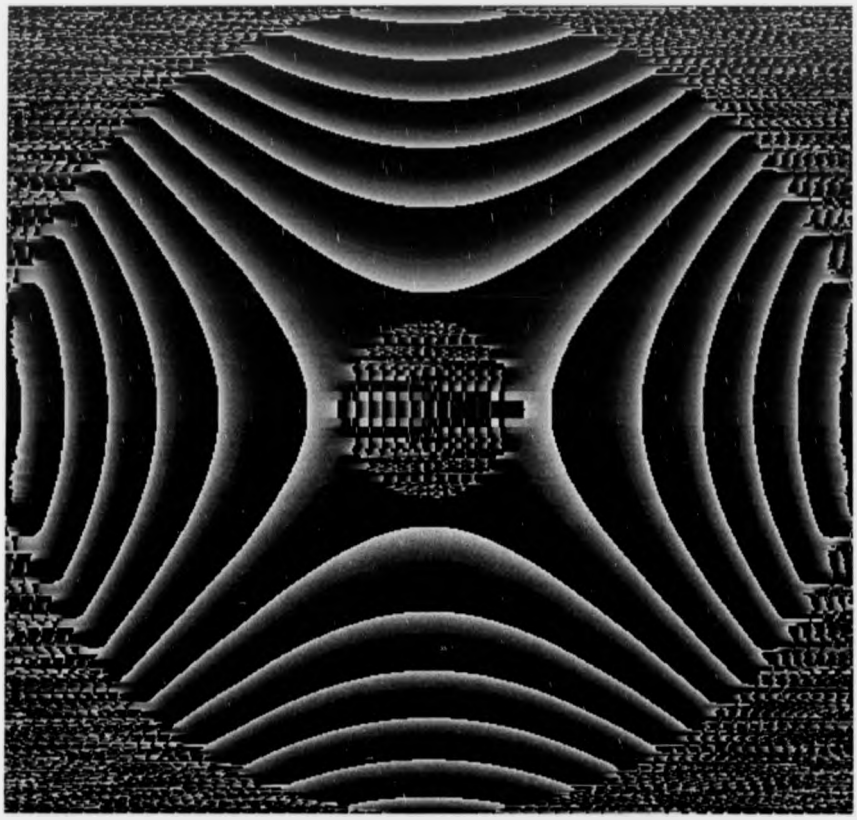


Figure 7.1: A wrapped phase map computed by DFT method

A brief overview is presented to outline the recent developement of the phase unwrapping algorithm. Then, the hierarchical phase unwrapping technique using the Minimum Spanning Tree approach [1] and the theory on which it is based will be discussed in detail.

## 7.2 A Brief Overview of Phase Unwrapping Algorithm

Fringe counting or fringe scanning is the primitive algorithm for phase unwrapping. This algorithm involves traversing through each row of the data looking for the phase roll-over points. A count, $N$, of the roll-over point is kept. The count is increased or decreased dependent on the direction of roll-over. $2N\pi$ is then added to the phase value of each successive pixel. The relative phase between each row of data is determined by unwrapping through the centre column of pixels. Although a number of improvements has been suggested, this approach fails easily in the presence of noise. Ghiglia et al [2] applied cellular automata to phase unwrapping. This algorithm is capable of adjusting automatically to point like inconsistencies. The cellular automata treats each pixel in the image as a point in a discrete lattice of sites. The value of each site evolves in discrete time steps according to the rules that depend only on a local neighbourhood of sites around it. Many global iterations are required for convergence.

Gierloff [3] has suggested a method to confront the problem posed by large scale discontinuities. This method identifies regions within which the phase appears consistent. These regions are intelligently phase shifted with respect to one another. Later, a number of researchers employed the area based approach to phase unwrapping [4, 5]. A fundamentally different approach to phase unwrapping has been suggested by Green and Walker [6]. The proposed algorithm attempts to eliminate the step changes by examining the spectrum of the unwrapped phase. This algorithm has been tested on the one-dimensional case.

Goldstein et al [7] and Huntley [8] have suggested an algorithm based on placing cuts between discontinuity sources of opposite sign. The cut acts as a barrier to unwrapping. Bone [9] has described a fast flood fill algorithm to improve the speed of the cut algorithms. The use of a Hopfield-type recurrent neural network to construct the cut lines has been demonstrated by Takeda et al [10]. Su et al [11] has proposed the use of fringe modulation information to determine the location of the cut. Rencently, Cusack et al [12] compare

four methods for optimizing the placement of the cut; nearest neighbour, stable marriage algorithm, simulated annealing and modified nearest neighbour algorithm. It is found that the modified nearest neighbour approach is the most successful one. The modified nearest neighbour algorithm can reliably unwrap phase map with discontinuity source densities of 0.05 sources/pixel.

Huntley and Saldner [13] have proposed an alternative approach that is based on one-dimensional unwrapping along the time axis for each pixel independently. This method has the possibility of real-time phase unwrapping because it is inherently simple. Other advantages are that phase errors are constrained within the high-noise regions, and that phase maps containing global discontinuities will be unwrapped correctly, if the object boundaries do not change with time. However, this algorithm is only applicable to a subclass of phase unwrapping problem.

## 7.3 Graph Theory and Algorithms

The graph theory [14, 15, 16] is a powerful analytical tool in the understanding and solution of large complex problems that arise in the study of engineering, computer science and operations research. In the graph theory, graphs are used to represent many types of problems. A classic example is the travelling salesman problem. The following sections, which derived from N. Deo [14] and A. Gibbons [15], address the topics in the graph theory related to the phase unwrapping problem.

### 7.3.1 The Basic Concepts

A graph $G = (V, E)$ consists of a set of objects $V = v_1, v_2, v_3, \cdots$ called vertices, and another set $E = e_1, e_2, e_3, \cdots$, whose elements are called edges, such that each edge $e_k$ is identified with an unordered pair $(v_i, v_j)$ of vertices. The vertices $v_i$, $v_j$ associated with edge $e_k$ are called the end vertices of $e_k$. A graph can be represented by a diagram in which a vertex is represented by a dot and an edge is represented by a line segment connecting the dots,

which represent the end vertices of the edge. For example, if

$$V = v_1, v_2, v_3, v_4, v_5$$

and

$$E = e_1, e_2, e_3, e_4, e_5, e_6, e_7, e_8$$

such that

$$\begin{aligned} e_1 &= (v_2, v_2), \\ e_2 &= (v_2, v_4), \\ e_3 &= (v_1, v_2), \\ e_4 &= (v_1, v_3), \\ e_5 &= (v_2, v_3), \\ e_6 &= (v_3, v_4), \\ e_7 &= (v_4, v_5), \\ e_8 &= (v_1, v_3), \end{aligned}$$

then the graph $G = (V, E)$ is represented as in Figure 7.2.

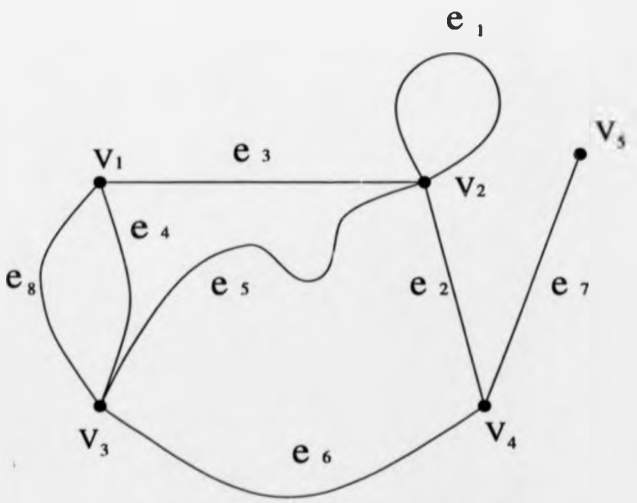



Figure 7.2: The graph G = (V, E)

A graph $G'$ is said to be a subgraph of a graph $G$, if all the vertices and all the edges of $G'$ are in $G$, and each edge of $G'$ has the same end vertices in $G'$ as in $G$. For instance, the graph in Figure 7.3 is a subgraph of the graph shown in Figure 7.2.

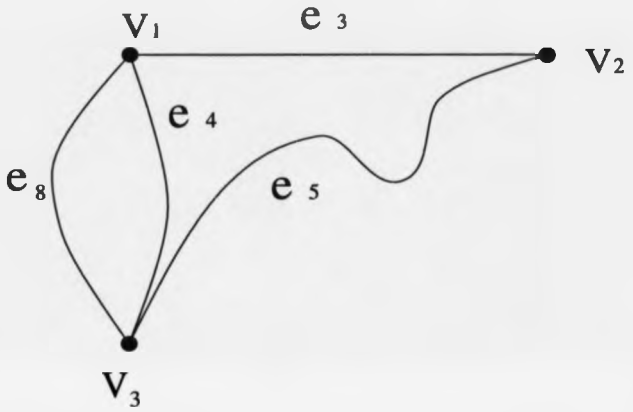



Figure 7.3: The subgraph of graph G = (V, E)

A walk is defined as a finite alternating sequence of vertices and edges, beginning and ending with vertices, such that each edge is incident with the vertices preceding and following it. No edge appears more than once in a walk. A walk is open if its end vertices are distinct; otherwise it is closed. A closed walk in which no vertex, except the initial and the final vertex, appears more than once is called a circuit. Four different circuits are shown in Figure 7.4

An important concept in the graph theory is that of connectedness. A graph $G$ is said to be connected if there is at least one walk between every pair of vertices in $G$; otherwise it is disconnected. The graph shown in Figure 7.2 is a connected graph.

### 7.3.2 Minimum Spanning Trees

A tree is a connected graph without any circuit. The graph shown in Figure 7.5 is a tree. A spanning tree $T$ of a connected graph $G$ is defined as that a tree $T$ is a subgraph of $G$ and contains all vertices of $G$. If a graph $G$

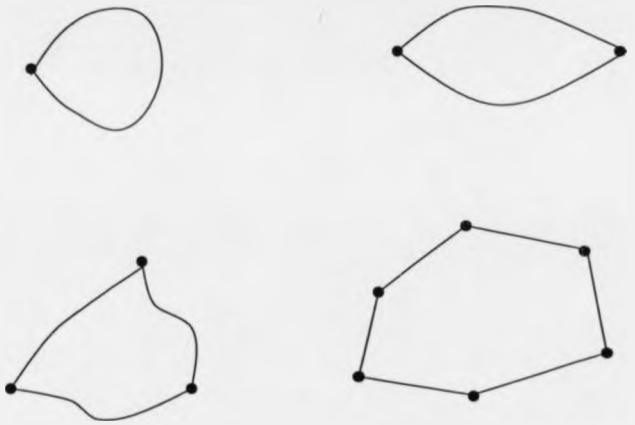


Figure 7.4: Four different circuits

is a weighted graph, then the weight of a spanning tree $T$ of $G$ is defined as the sum of the weights of all the branches in $T$. Among all of the spanning trees of $G$, that with the smallest weight is called a Minimum Spanning Tree (MST). There are a number of algorithms known to find a MST in a given graph. The following example illustrates the Prim's algorithm [17], which is used in FRANSYS.

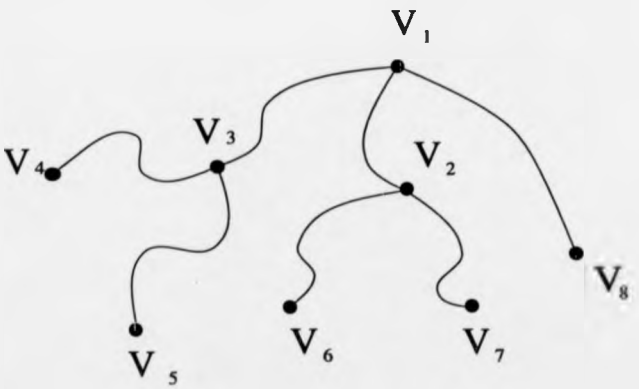
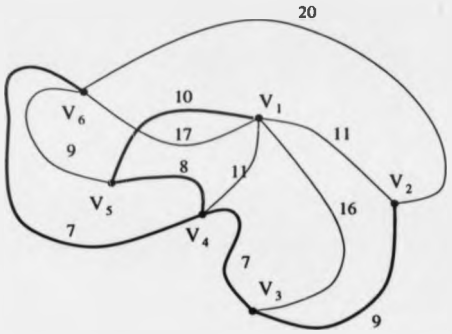



Figure 7.5: An example of a tree

A connected weighted graph with six vertices is shown in Figure 7.6(a). The weight of its edges is tabulated in Figure 7.6(b). The 'x' in the weight




| | V1 | V2 | V3 | V4 | V5 | V6 |
|---|---|---|---|---|---|---|
| V1 | - | 11 | 16 | 11 | 10 | 17 |
| V2 | 11 | - | 9 | x | x | 20 |
| V3 | 16 | 9 | - | 7 | x | x |
| V4 | 11 | x | 7 | - | 8 | 7 |
| V5 | 10 | x | x | 8 | - | 9 |
| V6 | 17 | 20 | x | 7 | 9 | - |

(a) (b)

Figure 7.6: A minimum spanning tree in a weighted graph

table means the weights of non-existent edge. We start with $v_1$ and pick the smallest entry in row 1, which is $(v_1, v_5)$. The starting vertex is selected arbitrarily. The closest neighbour of subgraph $(v_1, v_5)$, i.e. a vertex other than $v_1$ and $v_5$ that has the smallest weight among all entries in rows 1 and 5 of weight table, is $v_4$. The three remaining edges selected following the above procedure turn out to be $(v_4, v_6)$, $(v_4, v_3)$ and $(v_3, v_2)$ in that sequence. The resulting MST is shown in Figure 7.6(a) in heavy lines.

## 7.4 Hierarchical Phase Unwrapping Strategy Using MST

The phase unwrapping strategy is divided into two levels. The low level covers the procedures required to unwrap phase pixel to pixel, within tiles. The high level considers the assembly of the tiles into a whole field solution. In the operation, the low level processes are applied first. As a product of the low level processing the factors required at the higher level are extracted.

### 7.4.1 Pixel Level Phase Unwrapping

The pixel level phase unwrapping in a tile using MST has developed from a fringe counting method. The problem is to construct a path for unwrapping which maximises confidence. The weights of graph edges will be used to signal the confidence of a particular route, pixel to pixel. The initial stages of MST methods include:

1. Computation of wrapped phase map.
2. Fringe edge detection using Sobel operator.
3. To unwrap phase row by row.
4. To unwrap phase column by column

Then, every pixel is considered to be a vertex in a graph of confidence over the tile. Each pixel has four neighbours; north, south, east and west, and a corresponding edge in the graph for each pixel neighbour. The edge weights may be calculated from the horizontal and vertical scans. These contain partially unwrapped data, localised to either rows or columns. The local nature of the data means that the absolute offset of the data relative to the rest of the field is unknown. However, only relative phase changes are necessary for the purpose of weight computation.

In order to minimise phase unwrapping errors, each pixel effectively represents a crossroad at which the best direction to move must be determined. The vertical scan set is used to weigh the merits of the north/south options, while the horizontal scan set is used to weigh the merits of the east/west options. The further steps in the unwrapping procedure include:

1. Computation of east/west pixel edge weights.
2. Computation of north/south pixel edge weights.
3. Phase unwrap during calculation of MST.

The edge weights are computed as an average over a two pixel area, which is larger than the pixel sized spike noise. In reference to Figure 7.7, $X$ and $Y$ indicate a horizontal and vertical unwrap scan, respectively. Let the

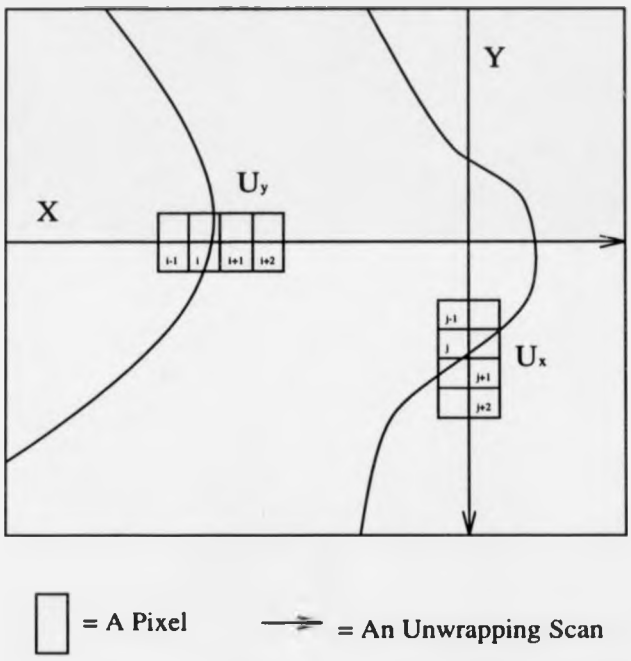



Figure 7.7: Computation of edge weights at pixel level

unwrapped phase of row scan y be denoted by $U_y(i)$, where $i$ is an index to the pixel of the scan, and similarly let $U_x(j)$ represent the unwrapped phase of column x at pixel $j$ of the scan. For a row pixel pair $i$ and $i+1$, the weight of the edge connecting the pixel vertices of row y is denoted by $w_y$ and calculated via Equ. 7.1 from the unwrapped rows:

$$w_y = |(U_y(i) + U_y(i-1)) - (U_y(i+1) + U_y(i+2))| \tag{7.1}$$

and similarly the north to south edge weights are calculated from the unwrapped columns in Equ. 7.2:

$$w_x = |(U_x(j) + U_x(j-1)) - (U_x(j+1) + U_y(j+2))| \tag{7.2}$$

The change in phase $\delta$ between the unwrapped pixel pairs is also recorded so that phase may be unwrapped during the formation of the MST:

$$\delta = U(i+1) - U(i) \tag{7.3}$$

Spike noise may be characterised by a rapid change in phase. If phase is unwrapped by a tree which seeks to minimise the change in phase at each step, noise points will be prevented from entering the solution until the tips of the tree branches are reached. As a result. the noise points do not affect the solution. The minimum spanning tree approach in the pixel level therefore presents a noise immune phase unwrapping strategy.

### 7.4.2 Tile Level Phase Unwrapping

Figure 7.8 shows a tiled section of the wrapped phase map. Firstly each tile is considered to be a vertex in a weighted connected graph $G$. Edges are added to the graph where tiles have a common boundary. That is each tile vertex is connected, by edges $e$, to the vertices of neighbouring tiles. The weights $w$, corresponding to the edges, are calculated to represent the validity of phase unwrapping across the boundaries. The factors include :

1. The agreement of solutions in neighbouring areas.

   This factor quantifies the fit of solutions in adjacent tiles. This is

calculated from a comparison of the profiles of the unwrapped solutions at the spatial boundaries of the tiles.

2. The extent of low modulation noise.
   This factor is used as an indicator for bed data. The low modulation factor is computed from the sum of the number of low modulation points found on either side of the tile boundaries.

3. The local fringe density.
   It is believed that in general the more fringe there are in a given area of an interforogram, the more probable an error is likely to result. The density of fringe, which has been estimated using the edge detection procedure, is the number of points found on fringe edges.

4. The extent of fringe terminations.
   This factor is obtained from a test for points where fringe edges terminate. That is where otherwise continuous fringes suddenly stop.

A normalisation process is applied to each factor discussed above to bring them into a domain where they may be directly compared. The combined weighting factor is obtained from the mean of the weighting factors after normalisation. A low combined weighting factor is defined to represent a good route. The assembly path is obtained by constructing the MST of the connected graph formed by the tile vertices.

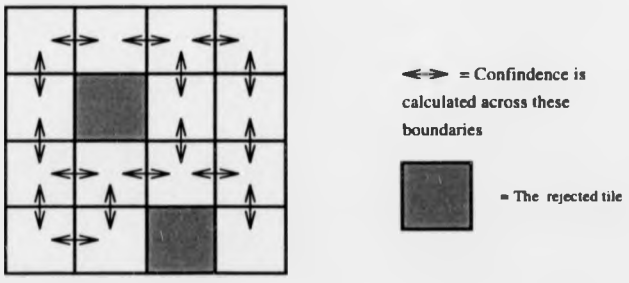



Figure 7.8: A tiled section of the wrapped phase map

The tiles are repositioned in height as the tree is computed, as shown in Figure 7.9. The height offset of each new tile added is calculated by summing

height across the tiles, from the root tile to the edge connecting the new tile to the tree. Defective tiles are forced to the tips of the tree branches and so distortions are reduced.

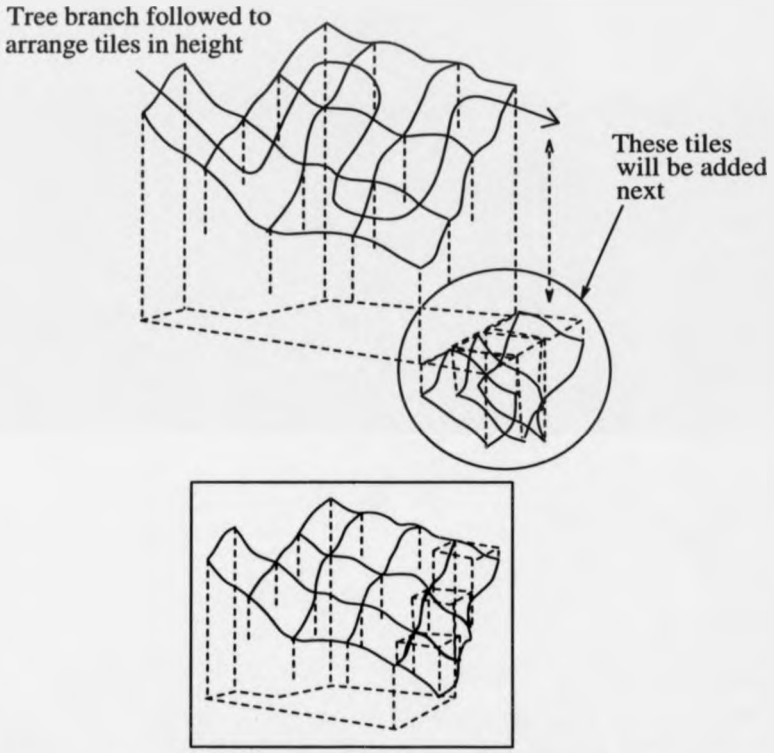



Figure 7.9: Tiles being arranged at their correct height offsets [1]

### 7.4.3 Selection of an Appropriate Tile Size

The confidence calculation is that there is a relationship between the best tile size and fringe density. To achieve the best unwrapped solution, the tile size should be comparable to the average fringe spacing, provided that any gaps in the fringe edges are not also of this order.

Figure 7.10 shows the plan view of a tile with a single broken fringe edge laying across it. The algorithm must decide whether or not a fringe edge is actually present. This is of particular importance for the connection of the

neighbouring tile along the line PQ.

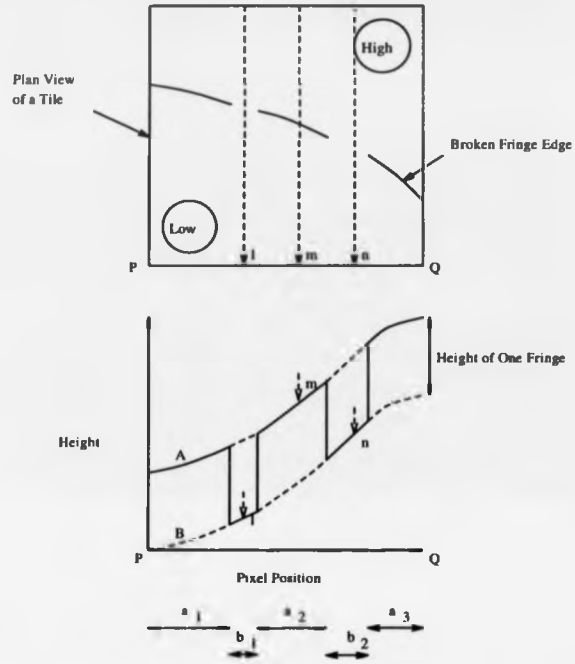



Figure 7.10: The tile with one badly disrupted fringe edge [1]

Figure 7.10 also shows the height profile at the bottom edge of the tile, along the line PQ. The profile consists of two interrupted curves differing by a single fringe height. The first curve A represents the cases where the fringe edge was detected, the second curve B represents those where it was not. The solution which is most likely to be correct is that of the curve with the longest stretch over the horizontal axis, when all of its component sections are summed. In this case, for example, the sum for the A solution is given by

$$l_A = a_1 + a_2 + a_3$$

and for the B solution is given by

$$l_B = b_1 + b_2$$

If $l_A$ is greater than $l_B$ then the A solution is selected, otherwise the B solution is chosen. This works well in terms of confidence as the algorithm has only a binary choice. The idea in the selection of tile size is to obtain a situation where at most one fringe crosses each tile. The level to which confidence may be determined if more fringes cross the tile decreases in a non-linear fashion.

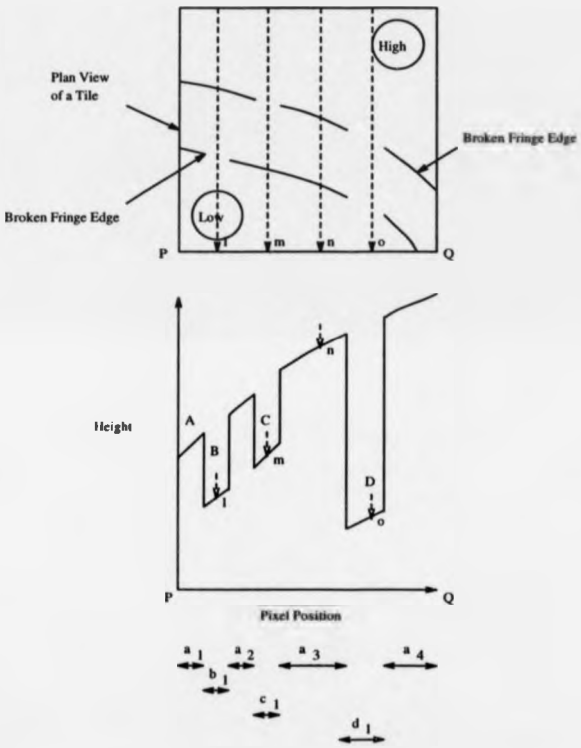



Figure 7.11: The tile with two badly disrupted fringe edge [1]

Consider Figure 7.11, where two disrupted fringes cross the tile. There are four possible solutions, A, B, C and D. The addition of one more fringe edge has doubled the number of solutions that must be considered. As the range of solutions is larger, it is more confusing for the algorithm as a binary decision no longer exists. If $m$ is the actual number of fringe edges which cross a tile, the number of solutions which must be distinguished between,

$n$, in the worst case, increases as a power of two;

$$m = 2^n$$

The discussion above deals with the upper limit of the tile size issue. On another hand, if the tiles are smaller than the fringe spacing, the connection flow would track the fringe. However, the binding force of a tile bridging the area of a broken boundary would cause a ridge to be formed along the length of the common edge. It is because the broken area is forced to the same height in the two fringes. This situation is avoided by scaling the tile size to exceed the size of the largest break in any fringe edge.

Therefore, there are two points to consider in the selection of the tile size. The lower limit is that the tiles must not be so small that they often fall between gaps in the edges of adjacent fringes and cause them to be merged into one. On the another limit, the tiles must not be so big that many edges fall across them, since a multitude of possible solution may complicate tile unwrapping.

## 7.5 Discussion

Tile size is a major parameter to achieve the best unwrapped solution. Normally, the MST unwrapping strategy has to run a number of times in order to determine the optimal tile size. This is undesirable for real time analysis.

In references to section 7.4.3, the lower limit of tile size is that the tile size should exceed the size of the largest break in any fringe edge. As a result, if there are some large fringe breaks in the wrapped phase map, a large tile size has to be selected. This will cause two main problems. Firstly, a large amount of information will be lost if a tile is rejected because of, for example, high low modulation noise. Secondly, large tile size will complicate the low level unwrapping, i.e. the lower limit will exceed the higher limit. An adaptive tile sizing approach is proposed and summarized as below :

1. The tile size is selected to exceed the largest break size in any fringe edge.

2. Wrapped phase map is segmented into tile.

3. The tile with high percentage of low modulation points is split into four equal subtiles. The subtile will be accepted as a valid tile if the area of low modulation points does not exceed the specified threshold; otherwise it is rejected.

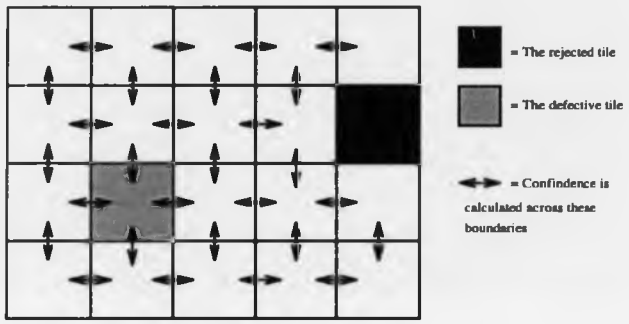
4. Fringe edges are found. Those tiles which contain fringe edges are unwrapped.

5. The combined weighting factor is computed for each edge, i.e. common boundary, of tile.

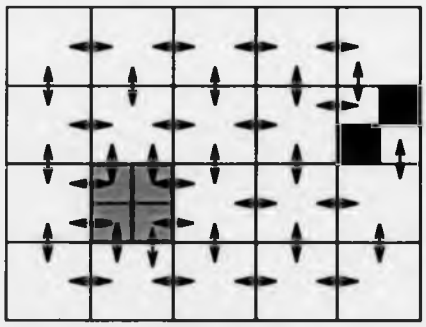
6. The defective tiles which have a high combined weighting factor are split into four equal subtiles. Each subtile, as if it was a normal tile, will go through the step 4 and 5 again.

7. The MST for the field gives the tile-to-tile unwrapping path.

Generally speaking, the above adaptive tile sizing approach is not necessary if the tile size determined at step one is less than 30 pixels. Figure 7.12 illustrates the adaptive tile sizing approach.




a) Before the tile size is adapted



b) After the tile size is adapted

Figure 7.12: The adaptive tile sizing approach

# Bibliography


[1] T. R. Judge,"Quantitative Digital Image Processing in Fringe Analysis and Particle Image Velocimetry", PhD Thesis, 1991.

[2] D.C. Ghiglia, G.A. Mastin and L.A. Romero,"Cellular-automata Method for Phase Unwrapping", J. Opt. Soc. Am. A, Vol 4, p. 267 - 280, 1987.

[3] J.J. Gierloff, "Phase Unwrapping by Regions", Proc. SPIE, Vol. 818, p. 2 - 9, 1987.

[4] D. Winter, R. Ritter and H. Sadewasser,"Evaluation of Modulo 2 $\pi$ Phase-images with Regional Discontinuities: Areas Based Unwrapping", Proc. 2nd Int. Workshop on Automatic Processing of Fringe Patterns, P. 157 - 159, 1993.

[5] G. Stoilov,"Method for Adaptive Unwrapping Using a Criterion for Maximum Probability of the Information Reliability",Proc. 2nd Int. Workshop on Automatic Processing of Fringe Patterns, P. 160 -162, 1993.

[6] R.J. Green and J.G. Walker,"Phase Unwrapping Using a Priori Knowledge About the Band Limits of a Function", Proc. SPIE, Vol. 1010, p. 36-43, 1988.

[7] R.M. Goldstein, H.A. Zebker and C.L. Werner,"Satellite Radar Interferometry: Two-dimensional Phase Unwrapping", Radio Sci., Vol 23, p.713-720, 1988.

[8] J.M. Huntley,"Noise-immune Phase Unwrapping Algorithm", Appl. Opt., Vol. 28, p. 3268 - 3270, 1989.

[9] D.J. Bone,"Fourier Fringe Analysis: The Two-dimensional Phase Unwrapping Problem", Appl. Opt., Vol. 30, p. 3627 - 3632, 1991.

[10] M. Takeda, K. Nagatome and Y. Watanable,"Phase Unwrapping by Neural Network", Proc. 2nd Int. Workshop on Automatic Processing of Fringe Patterns, P. 136 - 141, 1993.

[11] X.Y. Su, G.V. Bally and D. Vukicevic,"Phase-stepping Grating Profilometry: Utilization of Intensity Modulation Analysis in Complex Objects Evaluation", Opt. Comm., Vol. 98, p. 141 - 150, 1993.

[12] R. Cusack, J.M. Huntley and H.T. Goldrein,"Improved Noise-immune Phase-unwrapping Algorithm", App. Opt., Vol. 34, p. 781 - 789, 1995

[13] J.M. Huntley and H. Saldner,"Temporal Phase-unwrapping Algorithm for Automated Interferogram Analysis", App. Opt., Vol. 32,p. 3047 - 3052, 1993.

[14] N. Deo, "Graph Theory with Applications to Engineering and Computer Science", Prentic-Hall, 1974.

[15] A. Gibbons, "Algorithmic Graph Theory", Cambridge University Press, 1985.

[16] K. Thulasiraman and M. N. S. Sawamy, "Graphs: Theory and Algorithms", John Wiley and Sons, 1992.

[17] R. C. Prim, "Shortest Connection Networks and Some Generalizations", Bell Sys. Tech. J., Vol.36, 1389-1401, 1957.

# Chapter 8

# Theoretical Analysis of Fourier Transform Profilometry

## 8.1 Introduction

The measurement of 3-D object profile is important in 3-D solid modeling, robotic 3-D vision, industrial inspection and medical application. When measurement could be upset by touch or the instrument cannot reach the object, optical non-contact measurement is a promising choice.

Takeda et al [1] employed the discrete Fourier transform (DFT) to measure 3-D object shape and call this method as Fourier Transform Profilometry (FTP). In this method a Ronchi grating pattern is projected onto the object surface. The grating pattern is perturbed according to the topography of the object. Then the perturbed grating is Fourier transformed and processed in its spatial frequency domain to retrieve the object shape information from the fundamental frequency component in the Fourier spectra.

There are several reasons for the use of the Fourier Transform Profilometry. It can accomplish distinction between a depression and an evaluation of the object shape. Due to the DFT method Fourier Transform Profilometry permits fully automatic measurement. Furthermore, measurement is accomplished by using one image only. This is essential for measurement of instantaneous quality and dynamic system.

## 8.2 Optical Geometry of FTP

Optical geometry is similar to that of projection moire topography, but in FTP the perturbed grating is directly processed by computer without using the second grating to generate moire fringe. Two different optical geometries, crossed-optical-axis and parallel-optical-axis, have been proposed by Takeda et al [1]. The discussion of section 8.2.1 and 8.2.2 is derived from Takeda et al [1]. Each geometry has some merits over the other as well as some drawbacks. This will be discussed at the end of this section.

### 8.2.1 Crossed-Optical-Axis Geometry

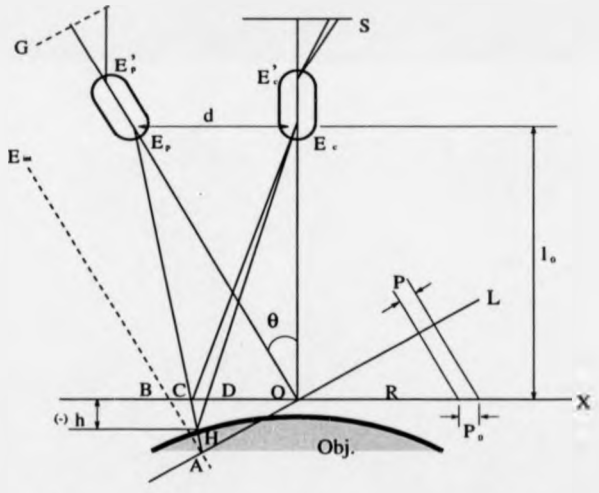



Figure 8.1: Crossed-optical-axis geometry

Figure 8.1 shows a geometry in which the optical axis $E_p'$ $E_p$ of a projector lens crosses the other optical axis $E_c'$ $E_c$ of a camera lens at point O on a reference plane R, which is a fictitious plane normal to $E_c'$ $E_c$ and serves as a reference from which object height h(x,y) is measured. Grating G has its lines normal to the plane of the figure, and its conjugate image with period $P$ is formed by the projector lens on plane L through point O; $E_p'$ and $E_p$ denote the centers of the entrance and the exit pupils of the projector lens,

respectively. The camera lens, with the centers of the entrance and the exit pupils at $E_c$ and $E_c'$, images reference plane R onto the image sensor planes S. $E_p$ and $E_c$ are located at the same distance $l_o$ from plane R. When the object is a flat and uniform plane on R and if $E_p$ is at infinity ( as denoted by $E_{int}$ for a telecentric projector), the grating image projected on the object surface and observed through $E_p$ is a regular grating pattern which can be expressed by a Fourier series expansion:

$$g_t(x,y) = \sum_{n=-\infty}^{\infty} A_n \exp(2\pi i n f_o x) \tag{8.1}$$

where

$$f_o = \frac{1}{P_o} = \frac{\cos\theta}{P} \tag{8.2}$$

is the fundamental frequency of the observed grating image. The x axis is chosen as in the Figure 8.1 and the y axis is normal to the plane of the Figure 8.1. If $E_p$ is at finite distance, we observe on the image sensor plane a deformed grating image with a pitch increasing with x, even for h(x,y) = 0. It is noted that the principal ray through a conjugate image point A strikes reference plane R at point B in the telecentric case and at point C in the nontelecentric case. We write the deformed grating image for h(x,y) = 0 as

$$g_o = \sum_{n=-\infty}^{\infty} A_n \exp\{2\pi i n f_o[x + s_o(x)]\} \tag{8.3}$$

where $s_o = \bar{BC}$ has a positive sign when C is to the right of B as in the Figure 8.1. For the convenience of later discussion, Equation 8.3 is expressed as a spatially phase-modulated signal

$$g_o(x,y) = \sum_{n=-\infty}^{\infty} A_n \exp\{i[2\pi n f_o x + n\phi_o(x)]\} \tag{8.4}$$

where

$$\phi_o(x) = 2\pi f_o s_o(x) = 2\pi f_o \bar{BC} \tag{8.5}$$

For a general object with varying h(x,y), the principal ray $E_pA$ strikes the object surface at point H, and point H will be seen as a point D on plane R when observed through $E_c$. Hence, the deformed grating image for a general

object is given by

$$g(x,y) = r(x,y) \sum_{n=-\infty}^{\infty} A_n \exp\{2\pi i n f_o[x + s(x,y)]\} \tag{8.6}$$

or

$$g(x,y) = r(x,y) \sum_{n=-\infty}^{\infty} A_n \exp\{i[2\pi n f_o x + n\phi(x,y)]\} \tag{8.7}$$

where

$$\phi(x,y) = 2\pi f_o s(x,y) = 2\pi f_o BD \tag{8.8}$$

and $r(x,y)$ is a nonuniform distribution of reflectivity on the object surface.

### 8.2.2 Parallel-Optical-Axis Geometry

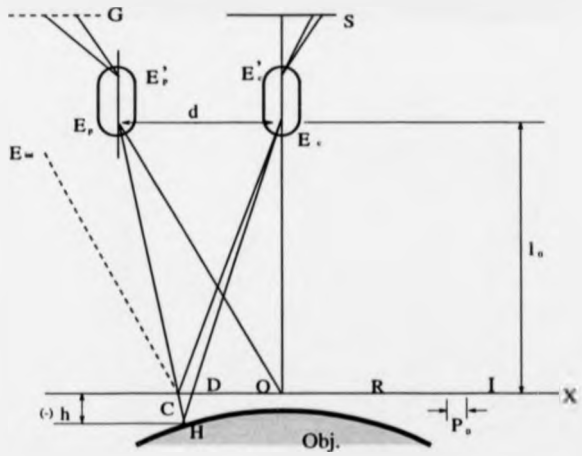



Figure 8.2: Parallel-optical-axis geometry

Figure 8.2 shows the parallel-optical-axis geometry. The optical axis $E_p'$ $E_p$ of a projector lens and that of a camera lens $E_c'$ $E_c$ are parallel and are normal to reference plane R. The conjugate image of grating G is formed on Plane R, and the three points A, B, and C in Figure 8.1 combine into point C in Figure 8.2, so that Equation 8.5 and 8.8 become

$$\phi_o(x) = 2\pi f_o s_o(x) = 2\pi f_o BC = 0 \tag{8.9}$$

$$\phi(x,y) = 2\pi f_o s(x,y) = 2\pi f_o \bar{CD} \tag{8.10}$$

Hence, the grating image projected on the plane h(x,y) = 0 remains as a regular grating pattern regardless of the position of the pupil of the projector.

## 8.3 Comparison of Different Optical Geometry

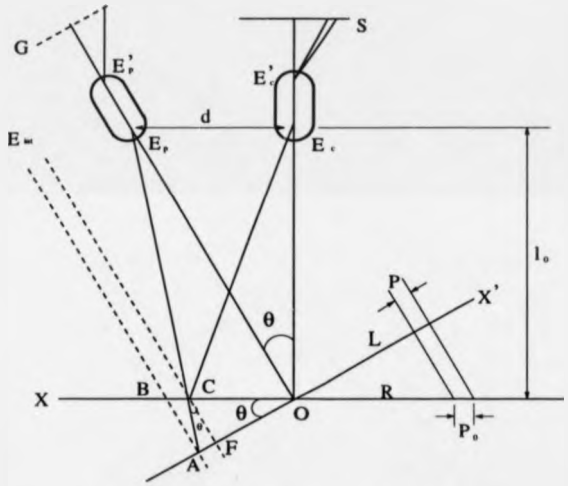



Figure 8.3: General optical geometry

Previous research works [1, 2] compared crossed-optical-axis geometry with parallel-optical-axis geometry based on the easiness of construction. Takeda et al [1] stated that :

> In crossed-optical-axis geometry, the optical axis of a projector and a camera lie in the same plane and intersect a point near the center of the object. This geometry is easy to construct because both a grating and an image sensor can be placed on the optical axis of the projector and the camera, respectively, but it gives planar contours only when the optics are telecentric. In parallel-optical-axis geometry, the optical axis of a projector and a camera

lie in the same plane and are parallel. This geometry gives planar contours but is somewhat awkward because the grating must be placed far off the optical axis of the projector to ensure that the grating image is formed within the field of view of the observation camera.

Apart from the merits and drawbacks mentioned above, in fact, parallel-optical-axis can provide higher range of measurement than crossed-optical-axis geometry. In order to explain this, a viewpoint which is different from Takeda et al [1] is taken.

Figure 8.1 is re-drawn as Figure 8.3. This is called general optical geometry because the parallel-optical-axis geometry is a special case of the crossed-optical-axis optical geometry. This will be explained later in the section. The $x$ and $x'$ are shown as in the Figure 8.3. If $E_p$ is at infinity, the grating image projected on plane R can be expressed as :

$$g_t(x,y) = \sum_{n=-\infty}^{\infty} A_n \exp(in\phi) \tag{8.11}$$

where

$$\phi = 2\pi x f_o \tag{8.12}$$

$$f_o = \frac{1}{P_o} = \frac{\cos\theta}{P} = f\cos\theta \tag{8.13}$$

$f$ and $f_o$ is fundamental frequency of conjugate image and that of grating G, respectively. Using formula of transformation of coordinate involving pure rotation, Equation 8.12 can be expressed as

$$\phi = 2\pi x' f \tag{8.14}$$

If $E_p$ is at finite distance, the pitch of grating image on plane R is increasing with x. It means that a ramp has been inserted into image by the optical geometry. In Figure 8.3, a principal ray through point A on the conjugate image strike reference plane R at point B in the telecentic case and at point C in the non-telecentric case. Furthermore, a principal ray through point C on the reference plane R strikes conjugate image at point E at telecentric

case. The deformed grating image for $h(x,y) = 0$ is expressed as

$$g_o = \sum_{n=-\infty}^{\infty} A_n \exp(in(\phi + \phi_0(x))) \tag{8.15}$$

where

$$\phi_0(x) = 2\pi f \bar{AF} = 2\pi f_o \bar{BC} \tag{8.16}$$

In Figure 8.3, $\angle COF = \angle ACF = \theta$ so that $\bar{CF} = \bar{OF}\tan\theta$ and

$$\begin{aligned} \bar{AF} &= \bar{CF}\tan\theta \\ &= \bar{OF}\tan^2\theta \\ &= x'\tan^2\theta \end{aligned}$$

Substituting $\bar{AF}$ into Equation 8.16, It is obtained that

$$\phi_0(x) = 2\pi f x' \tan^2\theta \tag{8.17}$$

Since

$$f = \frac{f_o}{\cos\theta}$$

and

$$x' = x\cos\theta$$

Equation 8.17 can also be expressed as

$$\phi_0(x) = 2\pi f x' \tan^2\theta = 2\pi f_o x \tan^2\theta \tag{8.18}$$

The $\phi_0(x)$ is a function of $x$ and $\theta$. It is not only a function of $x$ as stated in Takeda et al [1]. Section 9.1 will show that the maximum range of measurement is

$$\left|\frac{\partial h(x,y)}{\partial x}\right|_{max} < \left(\frac{1}{3} - \tan^2\theta\right)\frac{l_o}{d} \tag{8.19}$$

For parallel-optical-axis geometry, $\theta$ is equal to zero so Equation 8.19 become the same equation shown in Takeda et al [1]. $\tan^2\theta$ is alway positive. It is clear that parallel-optical-axis geometry can provide higher range of measurement than the crossed-optical-axis geometry if the value of $l_o$ and $d$ is same.

## 8.4 The Fourier Transform Profilometry

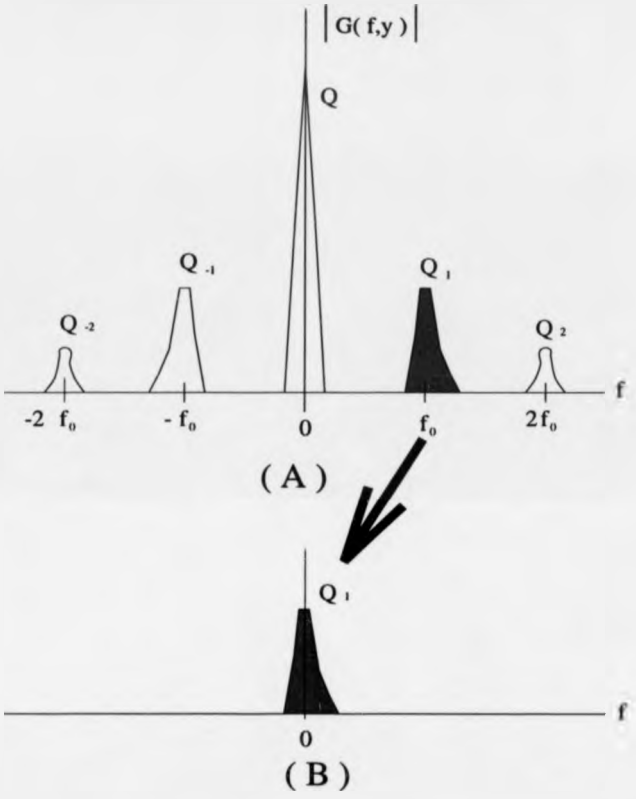



Figure 8.4: (A) Spatial frequency spectra of deformed grating image for a fixed y. (B) For spectrum shift method, only the fundamental frequency spectrum (shaded) is selected and translated to the origin.

The deformed grating image given by Equation 8.7 can be interpreted as multiple signals with spatial carrier frequencies $nf_o$ modulated both in phase $\phi(x,y)$ and amplitude $r(x,y)$. We rewrite Equation 8.7 as

$$g(x,y) = \sum_{n=-\infty}^{\infty} q_n(x,y)\exp(2\pi i n f_o x) \tag{8.20}$$

where

$$q_n(x,y) = A_n r(x,y) exp[in\phi(x,y)] \tag{8.21}$$

By using a FFT algorithm, we compute the 1-D Fourier transform of Equa-

tion 8.20 for the variable x only, with y being fixed:

$$\begin{aligned} G(f,y) &= \int_{-\infty}^{\infty} g(x,y)\exp(-2\pi fx)dx \\ &= \sum_{n=-\infty}^{\infty} Q_n(f-nf_o,y) \end{aligned} \tag{8.22}$$

where $G(f,y)$ and $Q_n(f,y)$ are the 1-D Fourier spectra of $g(x,y)$ and $q_n(x,y)$, respectively, computed with respect only to the variable x, and the other variable y being treated as a fixed parameter. Since in most cases $r(x,y)$ and $\phi(x,y)$ are slower compared with the frequency $f_o$ of the grating pattern, all the spectra $Q_n(f-nf_o,y)$ are separated from each other by the carrier frequency $f_o$ as shown in Figure 8.4(A). There are two methods to separate $\phi(x,y)$ from the unwanted $r(x,y)$ caused by nonuniform reflectivity on the object surface.

### 8.4.1 The Phase Subtraction Method

In this method [1], only one spectrum $Q_1(f-f_o,y)$, which is shadeed in the Figure 8.4A, is selected and computed its inverse Fourier transform to obtain a complex signal

$$\begin{aligned} \bar{g}(x,y) &= q_1(x,y)\exp(2\pi i f_o x) \\ &= A_1 r(x,y)\exp\{i[2\pi f_o x+\phi(x,y)]\} \end{aligned} \tag{8.23}$$

In the crossed-optical-axis case, the same filtering operation is done for Equation 8.4 to obtain

$$\bar{g}_o(x,y) = A_1\exp\{i[2\pi f_o x+\phi_o(x)]\} \tag{8.24}$$

From Equation 8.23 and 8.24, a new signal is generated

$$\bar{g}(x,y) * \bar{g}_o(x,y) = |A_1|^2 r(x,y)\exp\{i[\Delta\phi(x,y)]\} \tag{8.25}$$

where

$$\Delta\phi(x,y) = \phi(x,y)-\phi_o(x)$$

$$= 2\pi f_o(\bar{B}D - \bar{B}C)$$
$$= 2\pi f_o \bar{C}D \tag{8.26}$$

Since the initial phase modulation $\phi_o(x)$ for $h(x,y) = 0$ is now subtracted, $\triangle\phi(x,y)$ in Equation 8.25 gives the phase modulation due to the object-height distribution. If it is applied to parallel-optical-axis geometry, the errors caused by misalignment-alignment and/or distortion of the lenses will be cancelled by the phase subtraction method. A complex logarithm of Equation 8.25 is computed :

$$\log[\bar{g}(x,y) * \bar{g}_o(x,y)] = \log[|A_1|^2 r(x,y)] + i\triangle\phi(x,y) \tag{8.27}$$

The phase distribution $\triangle\phi(x,y)$ is obtained in the imaginary part of Equation 8.27.

### 8.4.2 The Spectrum Shift Method

In the spectrum shift method [3], either of the two fundamental spectra, say $Q_1(f - f_o, y)$, is used and translated by $f_o$ on the frequency axis towards the origin to obtain $Q_1(f,y)$, as shown in Figure 8.4B. The inverse Fourier transform of $Q_1(f,y)$ is computed with respect to f to obtain a complex signal

$$\bar{q}_1 = A_1 r(x,y) exp[i\phi(x,y)] \tag{8.28}$$

A complex logarithm of Equation 8.28 is calculated :

$$\log[\bar{q}_1] = \log[A_1 r(x,y)] + i\phi(x,y) \tag{8.29}$$

The phase distribution $\phi(x,y)$ is obtained in the imaginary part of Equation 8.29. Although the image of reference plane is not required in this method, the initial phase modulation $\phi_o(x)$ must be cancelled by another method.

### 8.4.3 Phase-Height Conversion

In this section, a formula for converting the measured phase distribution into the physical height distribution is derived. It is noted that $\triangle E_p H E_c$ is similar to $\triangle CHD$ in both Fig. 8.1 and 8.2. $\overline{CD}$ can be expressed as:

$$\overline{CD} = \frac{-dh(x,y)}{l_0 - h(x,y)} \tag{8.30}$$

where the object height, $h(x,y)$, is defined positive when measured upward from reference plane R. Substituting Eq. 8.30 into Eq. 8.26 and solving it for $h(x,y)$, the conversion formula is obtained as

$$h(x,y) = \frac{l_0 \triangle\phi(x,y)}{\triangle\phi(x,y) - 2\pi f_0 d} \tag{8.31}$$

# Bibliography

[1] M. Takeda and K. Motoh, "Fourier Transform Profilometry for the Automatic Measurement of 3-D Object Shapes", Applied Optics, Vol. 22 No. 24 (1983), 3977-3982.

[2] J. Li, X. Su and L. Gue, "Improved Fourier Transform Profilometry for the Automatic Measurement of Three-dimensional Object Shapes", Optical Engineering, Vol. 29 No. 12 (1990), 1439-1444.

[3] M. Takeda, H. Ina and S. Kobayashi, "Fourier-transform Method of Fringe-pattern Analysis for Computer-based Topography and Interferometry",J. Optical Society of America, Vol. 72 No. 1 (1982), 156-160.

# Chapter 9

# Limitations of Fourier Transform Profilometry

## 9.1 Maximum Range of Measurement

In this section, a generalized formula for calculating the maximum range of FTP measurement is derived. It is shown that the formula derived by Takeda et al [1] is a special case of the generalized formula.

The carrier frequency $f_o$ must separate the fundamental frequency spectrum from all other spectra because FTP is based on selecting only a single spectrum of the fundamental frequency component by filtering operation. This condition limits the maximum range measurable by FTP. Since $r(x,y)$ varies much slower than $f_o$, the $n^{th}$ spectrum component can be defined as a local spatial frequency $f_n$ analogous to an instantaneous frequency of FM signal :

$$\begin{aligned} f_n &= \frac{1}{2\pi}\frac{\partial}{\partial x}[2\pi n f_o x + n\phi(x,y)] \\ &= n f_o + \frac{n}{2\pi}\frac{\partial \phi(x,y)}{\partial x} \end{aligned} \tag{9.1}$$

For the fundamental spectrum to be separated from all other spectra, it is necessary that

$$(f_1)_{max} < (f_n)_{min} \quad (n = 2, 3, ...) \tag{9.2}$$

and that

$$f_b < (f_1)_{min} \tag{9.3}$$

where $f_b$, $(f_n)_{max}$ and $(f_n)_{min}$ are shown in Figure 9.1 for n = 1, 2 and 3. Substituting Equation 9.1 into Equation 9.2 and 9.3, we have

$$f_o + \frac{1}{2\pi}\left(\frac{\partial\phi}{\partial x}\right)_{max} < nf_o + \frac{n}{2\pi}\left(\frac{\partial\phi}{\partial x}\right)_{min} \qquad (n = 2, 3, \ldots) \tag{9.4}$$

$$f_b < f_o + \frac{1}{2\pi}\left(\frac{\partial\phi}{\partial x}\right)_{min} \tag{9.5}$$

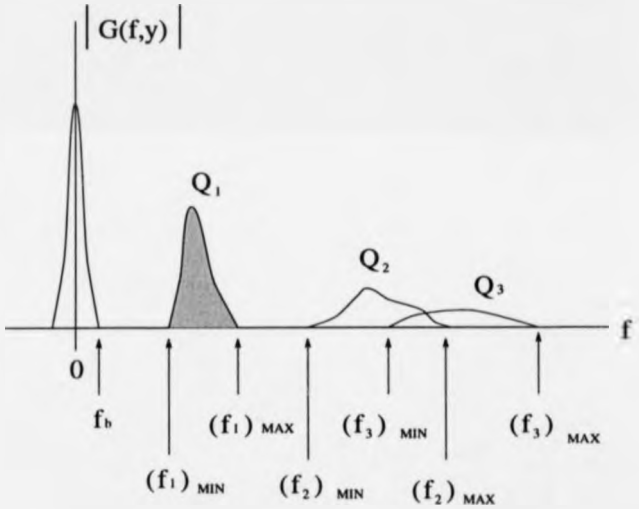



Figure 9.1: Condition for separating fundamental frequency spectrum $Q_1$ (shaded) from other spectra

A safer and more practical condition can be set by

$$f_o + \frac{1}{2\pi}\left|\frac{\partial\phi}{\partial x}\right|_{max} < nf_o + \frac{n}{2\pi}\left|\frac{\partial\phi}{\partial x}\right|_{max} \qquad (n = 2, 3, \ldots) \tag{9.6}$$

$$f_b < f_o + \frac{1}{2\pi}\left|\frac{\partial\phi}{\partial x}\right|_{max} \tag{9.7}$$

where $|(\partial\phi)/(\partial x)|_{max}$ denotes the maximum absolute value which is a larger

value of $|[(\partial\phi)/(\partial x)]_{max}|$ and $|[(\partial\phi)/(\partial x)]_{min}|$. From Equation 9.6 and 9.7,

$$\left|\frac{\partial\phi}{\partial x}\right|_{max} < \left(\frac{n-1}{n+1}\right) 2\pi f_o \quad (n = 2, 3, ...) \tag{9.8}$$

$$\left|\frac{\partial\phi}{\partial x}\right|_{max} < 2\pi(f_0 - f_b) \tag{9.9}$$

Since in most cases $f_b$ is much smaller than $f_o/2$ and $(n-1)/(n+1)$ increases monotonically with n, the limit is set by Equation 9.8 for n=2:

$$\left|\frac{\partial\phi}{\partial x}\right|_{max} < \frac{2\pi f_o}{3} \tag{9.10}$$

Section 8.4.3 has proven that the phase-to-height conversion formula is :

$$h(x,y) = \frac{l_o \triangle\phi(x,y)}{\triangle\phi(x,y) - 2\pi f_o d} \tag{9.11}$$

Assuming $l_o \gg h(x,y)$, Equation 9.11 can be written as:

$$\triangle\phi(x,y) \simeq \frac{-2\pi f_o d}{l_o} h(x,y) \tag{9.12}$$

From Equation 8.18 and 9.12,

$$\begin{aligned} \phi(x,y) &= \triangle\phi(x,y) + \phi_0(x) \\ &\simeq \frac{-2\pi f_o d}{l_o} h(x,y) + 2\pi f_o x \tan^2\theta \end{aligned} \tag{9.13}$$

Substituting Equation 9.13 into Equation 9.10, we finally obtain

$$\left|\frac{\partial h(x,y)}{\partial x}\right|_{max} < \left(\frac{1}{3} - \tan^2\theta\right) \frac{l_o}{d} \tag{9.14}$$

Equation 9.14 states that the maximum range of measurement is not limited by the height distribution $h(x,y)$ itself but by its derivative in the direction normal to the line of the grating. In other words, the range of measurement is governed by the band-width of the height distribution. Furthermore, when $\theta$ is equal to zero, i.e. in the case of parallel-optical-axis geometry, Eq. 9.14 becomes the same equation shown in Takeda et al [1]. This implies that Takeda et al [1] is a special case of Eq. 9.14.

## 9.2 Object Size and Depth of Field

The maximum area which can be inspected in one view by FTP is not limited by the available power of light source. However, it is governed by the area of grating image which is formed within the field of view of the camera. In the case of crossed-optical-axis geometry, the object size can be equal to the full view field of the observation camera, if the angle between the optical axis of the projector lens and that of the camera lens is large. Of course, the maximum range of measurement will be reduced. Generally speaking, the object size is not a restriction of FTP.

Due to the speckle effect, the viewing lens aperture used for interferometry should be as large as possible in order to increase the visibility of the fringe. As a result, the depth of field is limited by the lens aperture size [2]. Since an incoherent light of which phase changes rapidly, and randomly, in time is used for FTP, the speckle patterns are washed out. Therefore, the FTP's depth of field is limited by the cut-off frequency of the viewing lens aperture and the frequency of Ronchi grating. If a CCD camera is used, the spatial resolution of CCD sensor array is usually lower than the cut-off frequency of the viewing lens aperture. It means that the cut-off of the viewing lens aperture is not a limitation of FTP. As a result, the depth of field is only limited by the highest F-number of the camera.

## 9.3 Measurement Sensitivity

For the real time measurement, a charge-coupled devices ( CCD ) camera is used to capture the image. Then, the digital image signal is sent to computer for analysing. Since the spatial resolution of the CCD camera is considerably less than that of the photographic emulsion, the measurement sensitivity of FTP is governed by the structure and characteristic of CCD camera. Therefore, the principles of CCD camera is reviewed before the formula for measurement sensitivity of FTP is derived.

### 9.3.1 Charge-coupled Devices Camera

The CCD is a semiconductor device which converts the incident energy, e.g. light energy, into electrical charge. A basic CCD cell is an enhancement mode MOS capacitor. Figure 9.2 shows a cross-sectional diagram of a three-phase, two-bit, n-channel CCD. The six MOS capacitors connected to $C_1$, $C_2$ and $C_3$ clock lines form the main body of CCD. The input diode, the input grate, the output diode and the output grate are the input and output structure that injects and detects charge packets to and from the main CCD body.

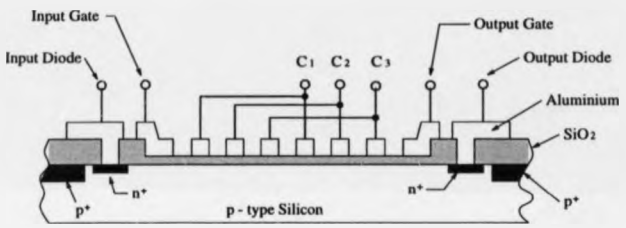



Figure 9.2: The cross-section diagram of a three-phase, two-bit, n-channel CCD

The operation of the device as an image sensor can be explained with the aid of Figure 9.3, which shows the various clock wave forms. During the time $T_{capt}$, $C_1$ is high while $C_2$ and $C_3$ are low. This creates two energy wells under the two $C_1$ electrodes. It is supposed that an image is shown on the device with high light intensity around the second $C_1$ electrode and low intensity around the first $C_1$ electrode. The photon which falls onto the device will generate an electron-hole pair in the semiconductor. The generated electrons will be collected in the energy wells under the $C_1$ electrodes. More electrons will be collected under the second $C_1$ electrode because the light intensity is higher around there. During $T_{scan}$, the charge packets are transferred to the output. The output signal is similar to the one shown in Figure 9.3. The magnitude of output signal will be proportional to the light intensity falling on the corresponding area of the device. [3]

The key component of the CCD camera is the sensor array. The array is usually organized as a rectangular grid as shown in Figure 9.4. This discrete sensing array involves a finite sensing area under each potential well that

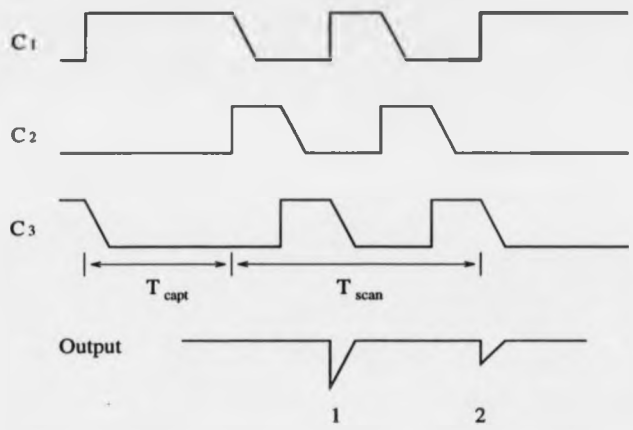



Figure 9.3: The clock wave forms and output signal for CCD as an image sensor

converts incident illumination in this area into an electrical signal. There is a wide variety of CCD cameras available in the market. At the present time, most image system digitize the output of CCD camera to 8-bits of intensity resolution. The standard CCD camera has a spatial resolution of the order of 500 by 500 pixel. For example, the Cohu solid state camera, which is used in the experiments, consists of 512 by 512 pixel within a 6.4 by 4.8 mm active area.

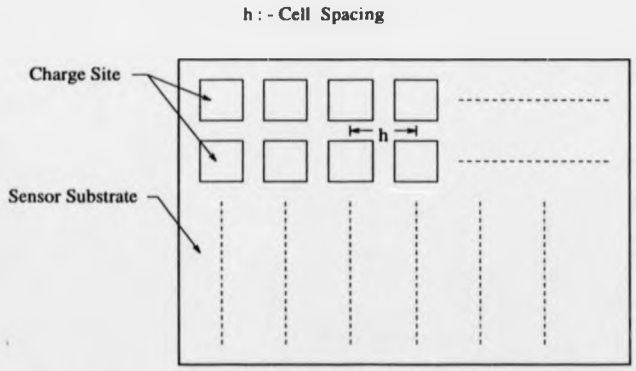



Figure 9.4: CCD cells array arranged on rectangular grid

### 9.3.2 Measurement Sensitivity of FTP Using CCD Camera

The formula for measurement sensitivity of FTP using CCD camera is derived in this section. Figure 9.5 shows the parallel-optical-axis FTP. In Figure 9.5, different scales are used in different side of camera lens. For the convenience of discussion, a sinusoidal grating is used in Figure 9.5. The derived formula is also valid for Ronchi gratings because it can be considered as a sum of sinusoidal gratings.

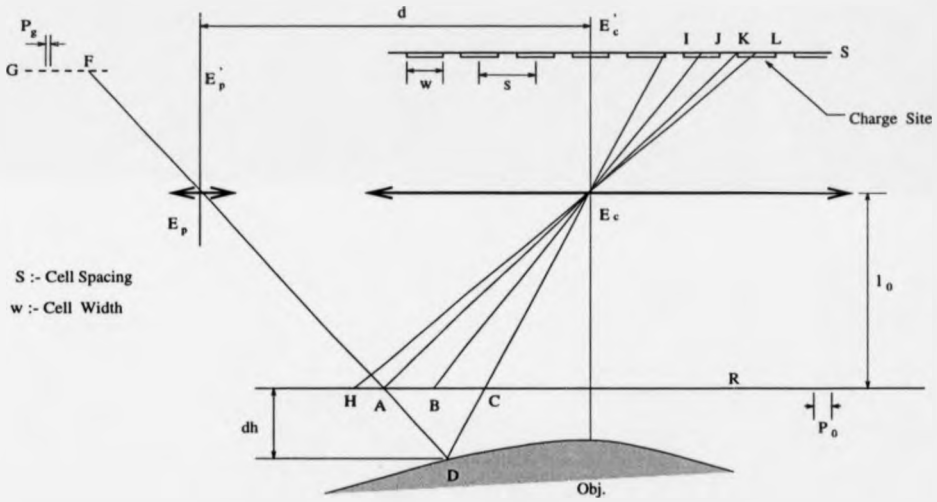



Figure 9.5: The parallel-optical-axis FTP using CCD camera

The grating G, of period equal to $P_g$, has its lines normal to the plane of the figure. The conjugate image of grating G is formed by the projector lens, with the optical axis at $E_p'$ $E_p$, on the reference plane R. When the object is a flat and uniform plane on R, the grating image projected on the object surface and observed through the camera lens, with the optical axis at $E_c'$ $E_c$, is a regular grating pattern which can be expressed by

$$g_t(x, y) = A \sin(2\pi f_0 x) \tag{9.15}$$

where

$$f_o = \frac{1}{P_o} = \frac{1}{M_p P_g}$$

is the frequency of the observed grating image and $M_p$ is the magnification factor of projector lens. The x axis is parallel to the plane of figure. If the sensor plane S is a the CCD array, the output of charge site IJ, which is an arbitrary charge site in the sensor array, in Figure 9.5 is the average illumination in the object area ABC since CCD camera converts incident illumination in the finite sensing area under each charge site into an electrical signal. The average illumination in the object area ABC is equal to the illumination at point B which is the centre of ABC. Similarly, the output of adjacent charge site KL is equal to the illumination at point H. The discrete output of the CCD camera can be expressed by

$$\begin{aligned} g_t(m,n) &= A\sin\left(2\pi f_0 m \bar{HB}\right) \\ &= A\sin\left(2\pi f_0 m \frac{\bar{h}}{M_c}\right) \end{aligned} \tag{9.16}$$

where

$$m = 0, 1, \cdots, N-1 \quad n = 0, 1, \cdots, N-1$$

and $M_c$ is the magnification factor of camera lens. The resolution of the CCD array is $N$ x $N$ pixel.

For a general object with varying $h(x,y)$, the principal ray FAD strikes the object surface at point D and strikes the reference plane at point A. Point D will be seen to be point C on plane R when observed through camera lens. Due to the characteristics of the CCD array, the output of charge site IJ, which is equal to the illumination at point B, remain unchanged as if there was no height change at that area. In other words, the FTP cannot detect the phase change caused by the height of the object. Therefore, the height resolution, $\delta S$, of FTP for a particular optical geometry using CCD camera can be expressed as

$$\begin{aligned} \delta S &= \pm \bar{BC} \\ &= \pm \bar{AB} \end{aligned}$$

$$= \pm\frac{2s - w}{2M_c} \tag{9.17}$$

From the Eq. 9.17, it is clear that the measurement sensitivity of FTP is restricted by the physical structure of the CCD array in the camera.

# Bibliography

[1] M. Takeda and K. Motoh, "Fourier Transform Profilometry for the Automatic Measurement of 3-D Object Shapes", Applied Optics, Vol. 22 No. 24 (1983), 3977-3982.

[2] R. Jones and C. Wykes, "Holographic and Speckle Interferometry" Camberidge University Press, 1989.

[3] M. J. Howes and D. V. Morgan, "Charge-coupled Devices and Systems", John Wiley, 1979.

# Chapter 10

# Optical Measurement of 3-D Object Profile by FTP

## 10.1 Introduction

Apart from the inherent limitation of Fourier Transform Profilometry (FTP), the performance of FTP is also dependent on the experimental set-up, such as optical geometry and scene illumination, and the method to separate the phase information from the unwanted nonuniform reflectivity on the object surface. Chapter 8 has discussed the effect of optical geometry. In this chapter, the illumination technique will be discussed. Then, supported by experiment results, the comparison between the phase subtraction method and the spectrum shift method will be made.

## 10.2 Scene Illumination Technique

Since the grating is projected onto the subject surface, the illumination for the FTP has to be sufficiently homogeneous (even) to avoid any loss of grating definition. The following sections will address the methods of applying various types of illuminators and illumination techniques [1, 2] to enhance features and to provide homogeneous illumination so that the complexity of signal processing is reduced.

### 10.2.1 Basic Types of Illuminators

The types of illuminators covered in this section are grouped into the following categories. Each of these categories will be defined to clarify their use. Figure 10.1 illustrates their definitions.

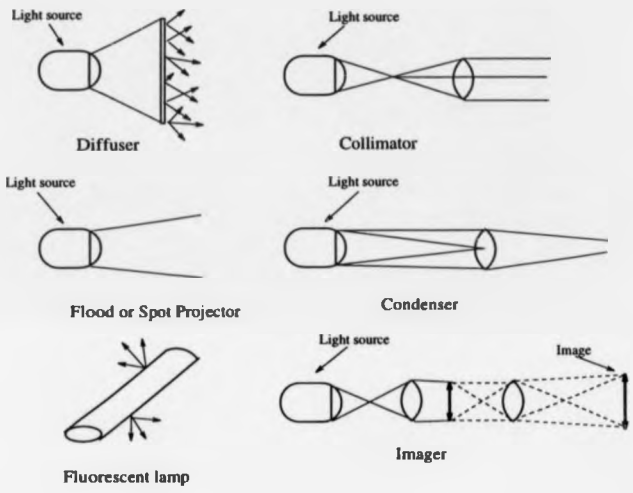



Figure 10.1: Basic types of illuminator

*Diffusers*

This term refers to illuminators that have the radiant energy emitting from a translucent material or reflecting from a diffuse surface. Examples of this are fluorescent lamps, light-tables and diffuse reflectors.

*Condenser*

A light source of this type changes the emitted direction of the light from an expanding cone into a condensing direction. This principle is used primarily in conjunction with imaging optics.

*Flood or spot projectors*

This type of illuminators also redirect the natural expanding cone, but are used to illuminate surface areas. Examples of this source are outdoor flood-

lights or headlamps.

*Collimators*

The rays of energy emitted from a pin-hole are very carefully redirected to form a beam of parallel light. If the collimator is perfect, the beam size will not change at all. Examples of this source are lasers and optical bench collimator.

*Imagers*

This term refers to illuminators that form an image of either the lamp filament or target at the object plane. Example of this type of illuminator is slide projector.

### 10.2.2 Illumination Techniques for FTP

This section deals with the arrangement of the illuminator related to the optical sensor. Schroeder H. E. [2] has given a precise and concise review about the illumination techniques for wide range of application. Only the illumination techniques concerned with FTP will be considered here.

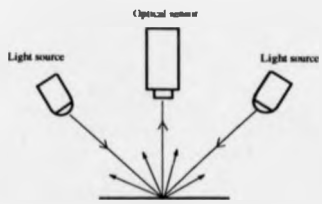



Figure 10.2: Front illumination

*Front illumination*

The purpose of this illumination, as shown as Figure 10.2, is to flood the area of interest with light so that the surface characteristics will act as the defining features in the image. It is possible that this technique can provide a homogeneous and shadow free illumination.

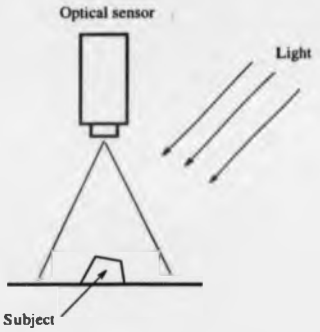



Figure 10.3: Offset illumination

*Offset illumination*

Figure 10.3 shows the block diagram of offset illumination. When the light source is offset, it produces a shadow of the object. The slide projector used in crossed-optical-axes geometry provides offset illumination.

*Porlarised lighting illumination*

A linear prolarised light source is used to illuminate the object. The optical sensor captures the image through a linear polariser. The polariser transmission axis of light source is normal to that of the optical sensor, as shown in Figure 10.4. Glare can be reduced by this technique. The technique is useful when FTP is used to measure an object with shiny surface.

## 10.3 Comparision of Different Phase Extraction Methods

In order to compare two phase extraction methods, an experiment is conducted. Figure 10.5 shows the schematic digram of the experimental set-up. A slide projector was used to project a Ronchi grating[1] of 10 lines per millimeter onto an object surface. The object is an eating bowl stuck on a plane which serves as a reference plane. The object is illuminated by front

[1] Ronchi grating consists of opaque and transparent stripes of equal width

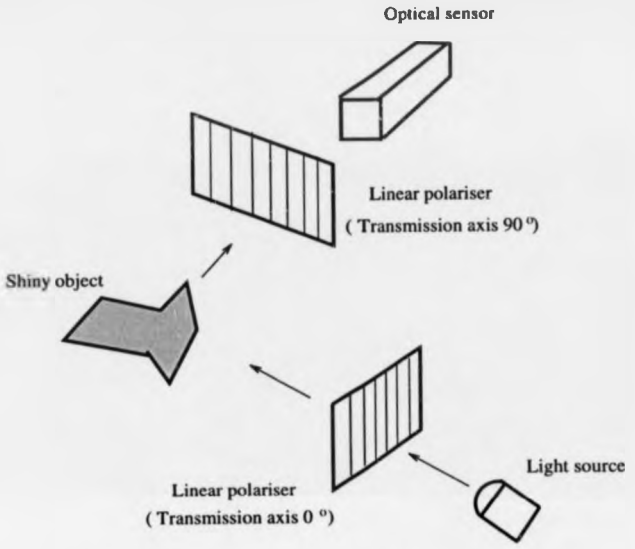



Figure 10.4: Porlarised lighting illumination

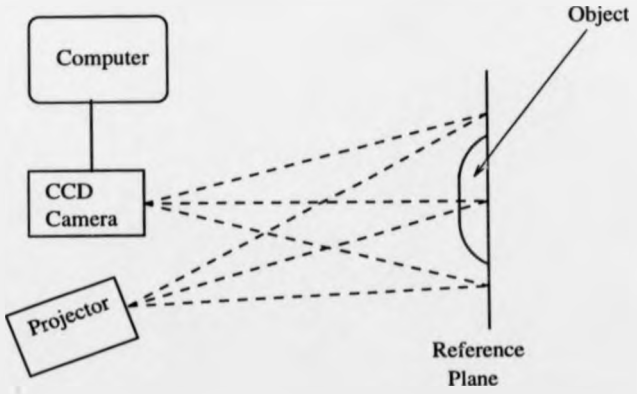



Figure 10.5: The schematic diagram of experimental set-up

illumination techniques using two diffuse reflector. The perturbed grating pattern was captured by a Cohu solid state camera with a 50 mm focal length lens. Figure 10.6 and 10.7 show the image of perturbed grating and the image of grating projected onto the reference plane respectively. Two methods are used to extract the phase information. The wrapped phase map is unwrapped by the Minimum Spanning Trees method.

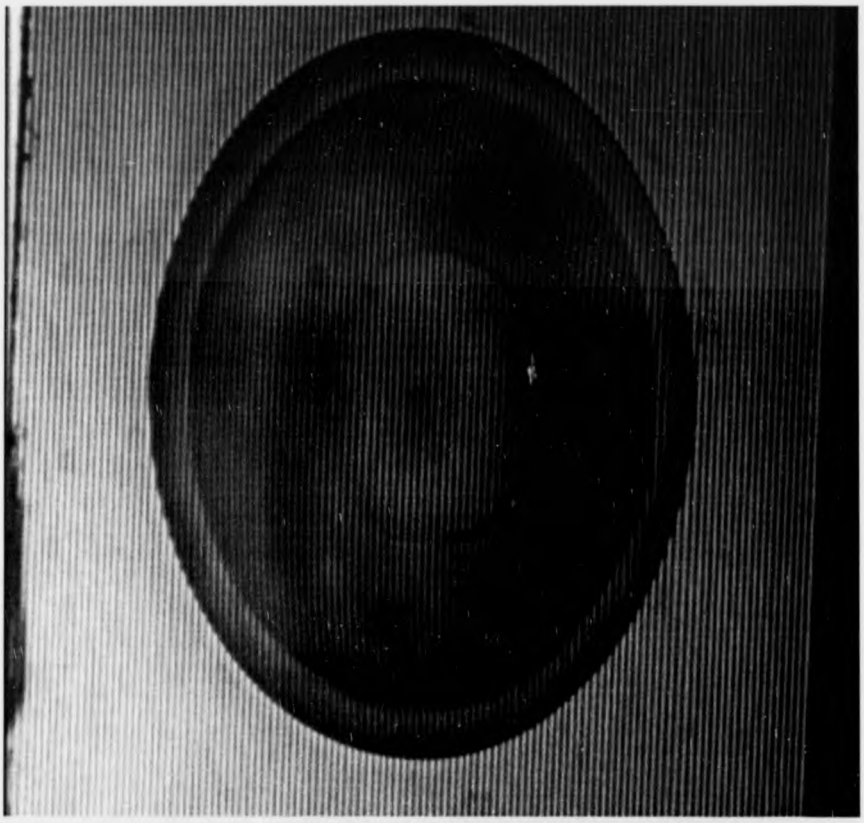


Figure 10.6: The image of perturbed grating

### 10.3.1 Results of the Phase Subtraction Method

The phase subtraction method [3] was implemented in Matlab. Appendix C.1 shows the M-file of the phase subtraction method. The fundamental spec-

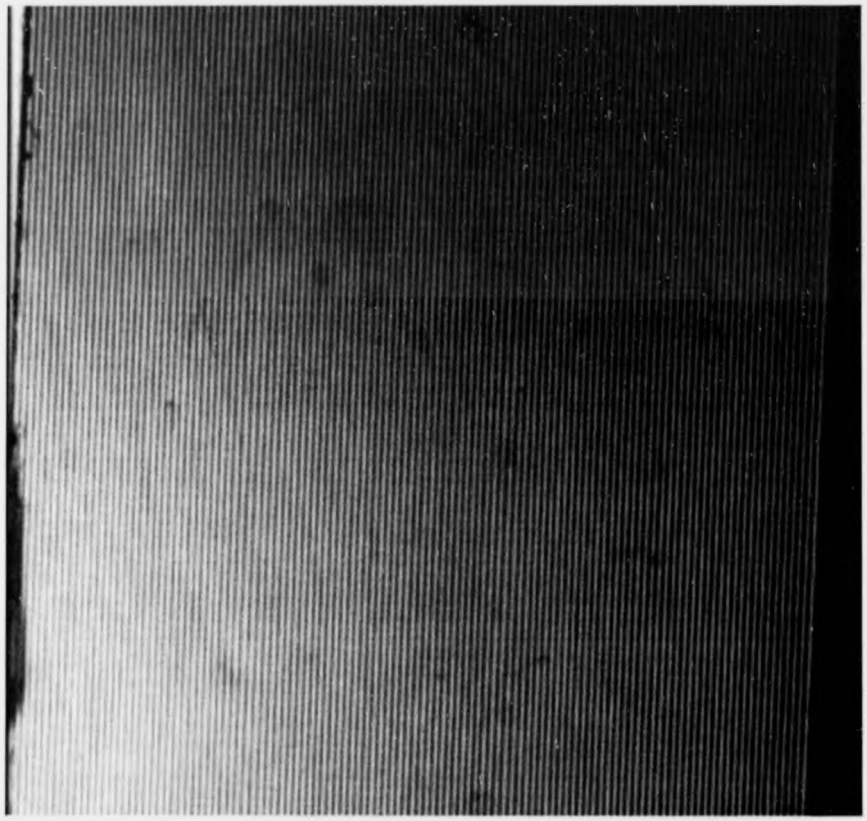


Figure 10.7: The image of Ronchi grating projected onto reference plane

trum was filtered by an ideal filter with band width of carrier frequency. A program was written to transfer Matlab data file format to FRANSYS data format. Figure 10.8 shows the wrapped phase map produced by the phase subtraction method.

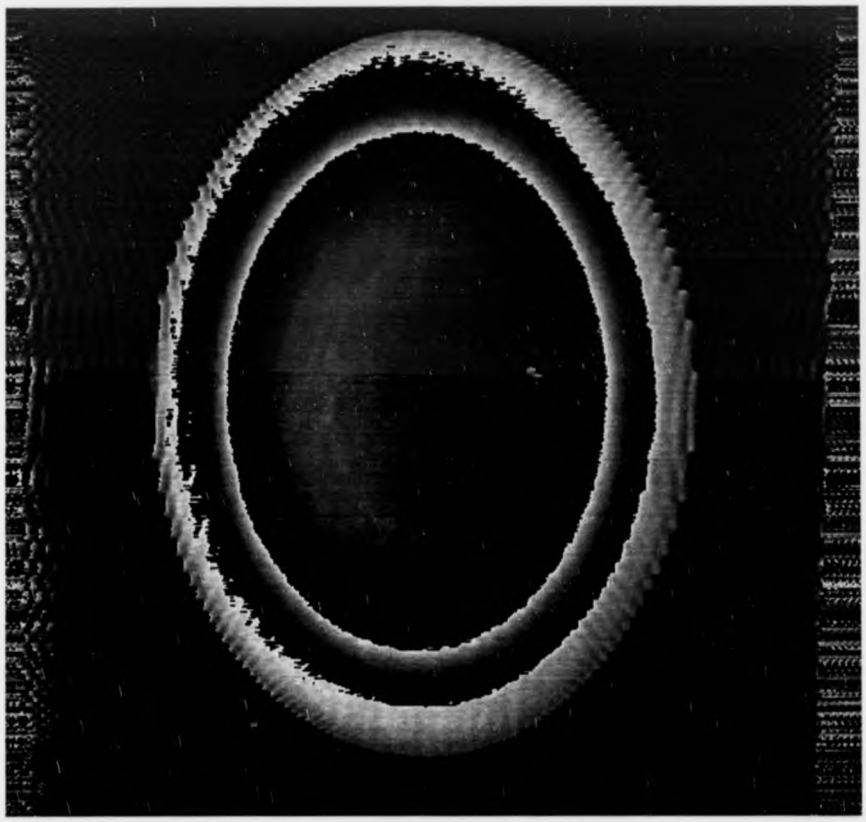


Figure 10.8: The wrapped phase map produced by phase subtraction method

The wrapped phase map was unwrapped by FRANSYS using the Minimum Spanning Trees method. Figure 10.9 and 10.10 show the 3-D plot of unwrapped data and its side view of unwrapped data respectively.

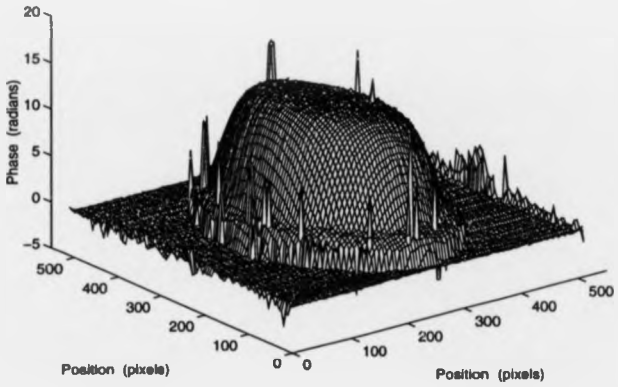



Figure 10.9: The 3-D plot of the unwrapped data produced by phase subtraction method

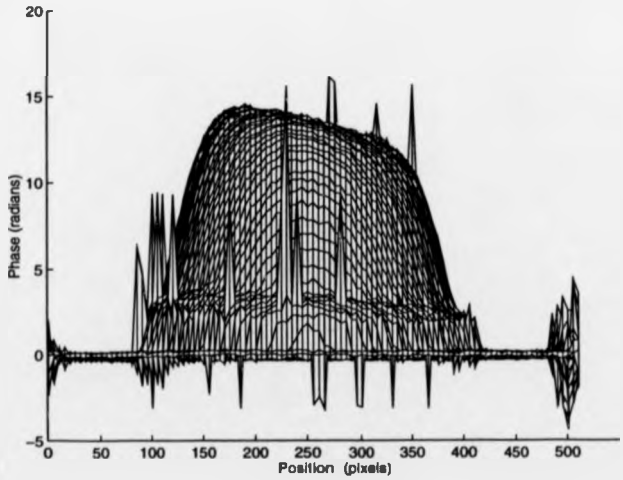



Figure 10.10: The side view of the unwrapped data produced by phase subtraction method

### 10.3.2 Results of the Spectrum Shift Method

Although the spectrum shift method [4] has been implemented in FRANSYS, the spectrum shift method implemented in Matlab was used in this experiment so that an apple-to-apple comparison can be made. Appendix C.2 shows the M-file of the spectrum shift method. Figure 10.11 shows the wrapped phase map produced by the spectrum shift method.

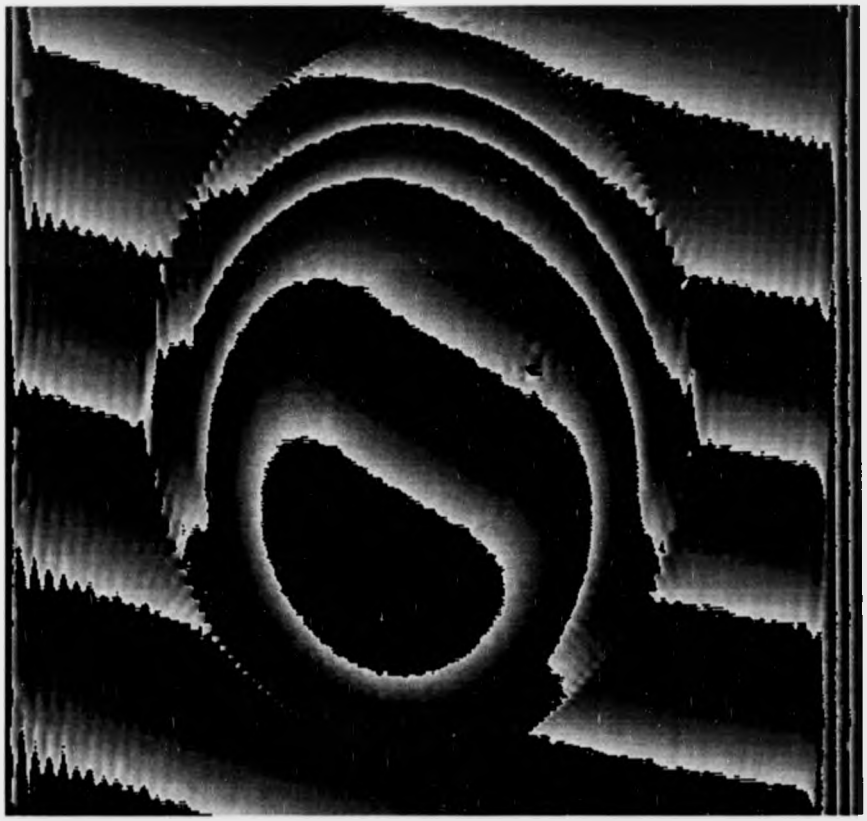


Figure 10.11: The wrapped phase map produced by spectrum shift method

The wrapped phase map was unwrapped by FRANSYS using Minimum Spanning Trees method. The same project file of FRANSYS used in the phase subtraction method is used. Figure 10.12 and 10.13 show the 3-D plot of unwrapped data and its side view of unwrapped data respectively.

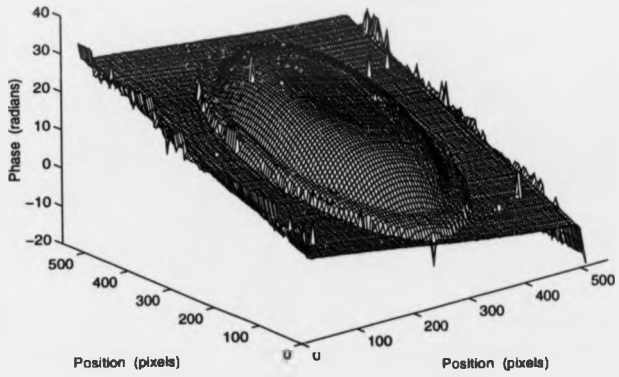



Figure 10.12: The 3-D plot of the unwrapped data produced by spectrum shift method

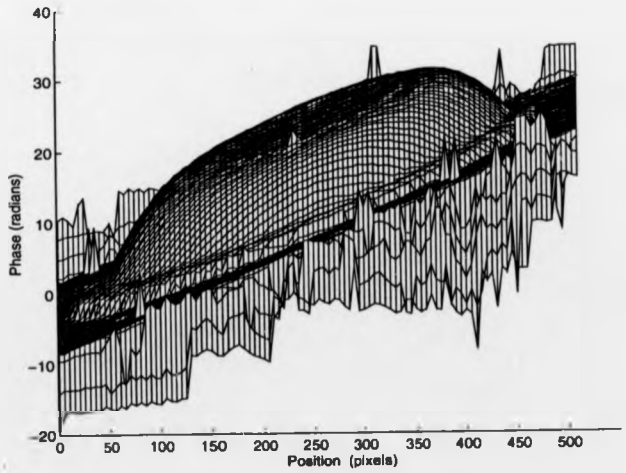



Figure 10.13: The side view of the unwrapped data produced by spectrum shift method

Referring to Figure 10.12, it is noted that two ramps running across the image have been imposed on the unwrapped data. One of the ramps is caused by the fact that the grating is not exactly aligned with the vertical axis of the frame. Another is caused by optical geometry. These superposed deformations have been eliminated. This has been accomplished by fitting a polynomial to data at the edge of the image which represent the flat surround and subtracting this fitted curve across the unwrapped phase map. The similar method has been used to eliminated the superimposed deformation due to the distortion of the camera lens. Figure 10.14 and 10.15 show the 3-D plot of corrected unwrapped data and its side view respectively.

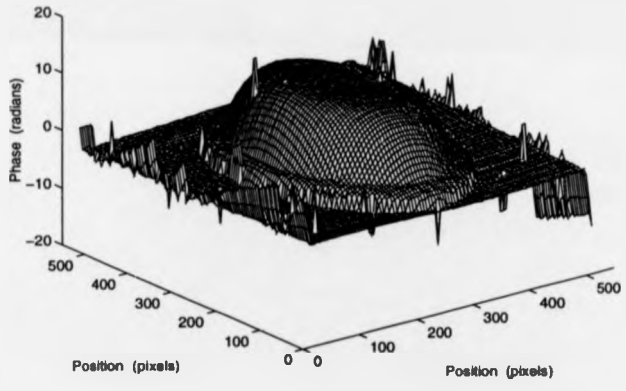



Figure 10.14: The 3-D plot of the corrected unwrapped data

### 10.3.3 Phase Subtraction Method Vs Spectrum Shift Method

In PTF, several sources of errors appear. These include:

- The errors caused by Fast Fourier Transform algorithm.
- nonlinearity of recording media.
- error introduced by optical geometry.

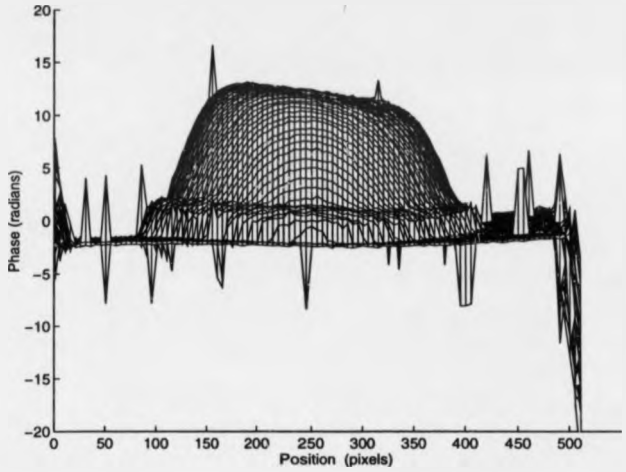



Figure 10.15: The side view of the corrected unwrapped data

- the influence of random noise.

The comparison will be made based on the effect on the errors listed above. However, the aliasing error will not be considered because it is associated with the resolution of CCD.

In the spectrum shift method, the fundamental frequency spectrum is filtered and then be translated along the frequency axis by carrier frequency. After the fundamental frequency spectrum is shifted, the carrier is eliminated from areas with grating. However, a carrier is introduced in the areas without grating. Figure 10.11 shows the effect on the wrapped phase map. The spurious phase data will increase the complexity of phase unwrapping processing. Furthermore, the carrier frequency has to be accurately located. If the error in the carrier frequency position is $\delta w$, using the shift theorem of Fourier Transform, the deduced phase distribution $\phi'(x, y)$ will be given by

$$\phi'(x, y) = \exp(-2\pi i x \delta w)\phi(x, y)$$

It means that the recovered phase object will be the true phase object modulated by a slow broad variation across the image. In some parts of the image this modulation may add a phase shifts that are larger than those

produced by a large fraction of the object. Thus, this effect represents a serious distortion of the data [5]. The grating must be exactly aligned with the vertical axis of the frame, otherwise, a ramp will be imposed on the unwrapped data along the y axis. Another ramp caused by optical geometry is also imposed along x axis. Although these superimposed ramps can be cancelled by the curve fitting method described in section 10.3.2, this method may also cancel the real deformation. The nonlinearity of lens has to be corrected by another curve fitting which suffers from same potential problem.

In the phase subtraction method, the errors caused by misalignment, distortion of lenses and optical geometry can be treated as a non-zero initial phase. Since the non-zero initial phase is cancelled by the phase subtraction method, the unwrapped data are free from these errors. The spurious phase data are eliminated because there is no translation of spectrum. Furthermore, only the position of fundamental frequency spectrum is required in this method. It is easier than locating the exact position of carrier frequency. As the carrier frequency is removed by subtraction, there will be no low modulation caused by the error in the carrier frequency position. Although two images are required, the phase subtraction method is still able to apply for measurement of instantaneous quality and dynamic system. It is because the background image is taken once before any force or disturbance is introduced. The main drawback is that this method is easily influenced by noise. The result of the phase subtraction method, as shown in Figure 10.9, is noisier than that of the spectrum shift method, as shown in Figure 10.12.

For measurement of out-of-plane deformation, the perturbed grating images are recorded at different load values. In the case of the spectrum shift method, the two perturbed grating images are computed separately, and then two computed phases are subtracted from each other in order to obtain the out-of-plane deformation. However, the phase subtraction method can combine two separated calculation into one, if the perturbed grating image recorded before load is applied is used as a reference perturbed grating image.

# Bibliography

[1] Graeme Awcock, "Scene Constraints" Lecture Note, University of Brighton.

[2] H. E. Schroeder "Practical Illumination Concept and Technique for Machine Vision Application" Robot Sensors Volume 1, IFS (Publication) Ltd., p.229-p.244.

[3] M. Takeda and K. Motoh, "Fourier Transform Profilometry for the Automatic Measurement of 3-D Object Shapes", Applied Optics, Vol. 22 No. 24 (1983), 3977-3982.

[4] M. Takeda, H. Ina and S. Kobayashi, "Fourier-transform Method of Fringe-pattern Analysis for Computer-based Topography and Interferometry",J. Optical Society of America, Vol. 72 No. 1 (1982), 156-160.

[5] K. A. Nugent, "Interferogram Analysis Using an Accurate Fully Automatic Algorithm" Applied Optics, Vol. 24 No. 18(1985), 3101-3105.

# Chapter 11

# Conclusion and Further Work

## 11.1 Conclusion

There are a variety of optical system to produce fringe patterns: such as Moire and interference of light. The required information is encoded in the phase of the fringe patterns. The phase of the fringe patterns can be related to a wide range of physical quantity. Some of these physical quantity are as follows: deformation of soil model, stress/stain and shape of a object. In the past, the fringe patterns were analysed manually by an expert. it is a time-consuming and tedious task. Furthermore, an expert is not alway available. Due to advances in photoelectric devices and the development of computer engineering, automatic analysis of fringe patterns is possible. The procedure of the automatic fringe analysis can be summarized as following:

1. Photoelectric detection of fringe pattern.
2. Calculation of wrapped phase map, i.e. the calculated phase is modulo $2\pi$
3. Unwrapped the phase to from a contiguous solution.

There are two prominent analytical techniques for the generation of wrapped phase map from fringe patterns: phase-stepping (PS) techniques and discrete Fourier transform (DFT) techniques. The PS techniques calculate the phase from a series of intensity measurement with a fixed mutual phase shift. Therefor, the PS methods are sensitive to non-linearity in the response of the

detector and background variations. A number of PS algorithms have been evolved to lessen the problem related to PS techniques. Several PS algorithms have been described. On the other hand, the DFT techniques rely on retrieving the phase information in the spatial frequency plane. The precision of the DFT methods depends upon the spatial measurement of fringe position; rather than intensity. The 1-D DFT technique demands vertical carrier fringes. Otherwise, a ramp, of which the gradient corresponds to the vertical component of the carrier fringe, is imposed on the computed phase. The 2-D DFT technique is free from this restriction. The multi-channel Fourier fringe analysis and the spatio-temporal frequency multiplex are two prominent methods to exploit the surplus space in the frequency domain so that more information can be stored in a single fringe image. These techniques have been discussed.

The spectral leakage is one of the inherent error sources of DFT techniques. A weighting function has to apply to the image data prior to transformation to alleviate the problem. Two investigations are conducted to find out the combined effect of weighting function and filtering window on accuracy of DFT techniques. A computer generated interferogram is used in both experiments so that the spectral leakage is the only error source. Firstly, the 1-D DFT method is examined. The statistical analysis of full field errors reveals that the Hamming function performs better than other function on average and produces the lowest maximum error while the Rectangle filter is used. However, the lowest mean and the lowest maximum full field error is produced by the jointed effect of the Hanning function and the Papoulis filter. The highest mean accuracy can be achieved by the Papoulis filter combined with the Hamming, the Rectangle or the Papoulis function, after one per cent of pixels are masked out on both sides of each raster line. It can be concluded that, when 1-D DFT technique is used, the Papoulis filter should be used for high accuracy. The Hanning function should be employed whenever the maximum error is important. However, the Papoulis function should be used, if the mean error is the critical parameter. Another investigation is concerned with 2-D DFT method. The 2-D window can be generated from 1-D window by either outer-product method or circularly rotated method. The 1-D Hamming, the 1-D Papoulis and the 1-D Rectangle window are

use as a prototype in this examination. When the full field and partial field error are taken into account, it is deduced from the statistial analysis that the Papoulis outer-product function and the Rectangle outer-product filter is the best combination to diminish the spectral leakage caused by 2-D DFT technique. Furthermore, the experiment indicated that the outer-product method is better then the circularly rotated method when the same prototype is employed to form a filter window.

Recent research work are concentrated on developing an analytic method that integrates the computing simplicity of PS techniques with the dynamic event analysis capability of DFT techniques. The spatial version of phase-stepping technique simultaneously acquires N phase shifted fringe patterns in N discrete spatial locations. The main problem of this approach is that it requires additional calibration procedures; for example, the alignment of the cameras. In the spatial-carrier phase shifting (SCPS) technique, a large amount of tilt is introduced so that the phase difference between successive pixels is equal to $\pi/2$. Then, three successive pixels are used to recover the phase using the standard phase shifting method. If the amount of tilt does not introduce accurately, a linear phase term will be added into the retrieved phase. A novel Spatial Phase-stepping (SPS) technique has been developed. The SPS technique can apply to fringe pattern with any spatial carrier frequency. Furthermore, this technique removes the introduced tilt automatically. The SPS technique utilizes a group of pixels that spread over the whole wave length of spatial carrier to compute the phase value of the centre pixel. Obviously, the resolution of computed phase in x direction is reduced by a factor that is equal to one wave length of spatial carrier fringe. An error analysis is conducted. The result indicates that the errors associated with the region in where there is no phase change are zero. It also shows that the maximum error occurs in the region where the phase distribution varies rapidly. Therefore, it can be conjectured that the absolute phase error is directly proportional to the gradient of the phase surface across the whole spatial carrier wave.

Two examples have been presented to illustrate the capability of the SPS technique. Both examples are related to the interferometric fluid flow measurement. From the fluid flow results, a comparison can be made between

the DFT and SPS techniques. SPS is nearly 10 times faster than the DFT approach in the first example. A comparing of Figure 6.17 with Figure 6.18 shows that the results produced by both techniques are qualitatively similar. However, some local details have been lost in the result produced by the SPS technique. This occurs because the resolution of the image is reduced after the original interferogram is spatially divided. In the second example, the result retrieved by SPS technique, which is shown in Figure 6.23, matches the result computed by the DFT technique from an image of resolution 600 DPI, shown in Figure 6.24. This implicates that the SPS method provides a solution to the fringe field with approximately 40 per cent of the resolution of the equivalent DFT.

Since the phase calculated by any one of above techniques give princial values ranging from $-\pi$ to $\pi$, the computed phase is wrapped into this range. A procedure has to be employed to correct the $2\pi$ phase jumps. A robust two-dimensional hierarchical phase unwrapping strategy which is based on the Minimum Spanning Tree (MST) approach has been reviewed. The MST approach to phase unwrapping is particularly aimed at addressing the problems posed by natural or aliasing induced discontinuities. The phase unwrapping strategy utilizes the concept that, unless an error is present, the same solution should be produced along the common boundary of the adjoining areas of the phase map. The phase unwrapping strategy is divided into two levels. The low level cover the procedures required to unwrap phase pixel to pixel, within tiles. The high level considers the assembly of tiles into a whole field solution. It has been noted that tile size is major parameter to achieve the best unwrapped solution. The strategy has to run a number of time in order to determine the optimal tile size. It is undesirable for real time analysis. A new adaptive tile sizing approach is proposed in this thesis.

Previous research work have shown that the DFT techniques are a suitable analysis methods for interfergram. This project is also directed towards applying a similar approach, which is called as Fourier Transform Profilometry (FTP), to analyse a image of an object onto which Ronchi grating is projected. A review of FTP has been made. Two different optical geometry, crossed-optical-axes and parallel-optical-axes, had been proposed. Previous work compared the two optical geometries based on their ease of construc-

tion. In this thesis, the comparison of two optical geometry is based on the maximum range of measurement. It is shown that the maximum range of FTP measurement is

$$\left|\frac{\partial h(x,y)}{\partial x}\right|_{max} < \left(\frac{1}{3} - \tan^2\theta\right)\frac{l_o}{d} \tag{11.1}$$

Where :

$\theta$ = angle between optical axes of a projector lens and that of a camera lens

$d$ = the distance between the center of exit pupil of projector lens and the center of entrance pupil of camera lens

$l_o$ = the distance between the center of entrance pupil of camera lens and reference plane

In the case of parallel-optical-axes geometry, $\theta$ is equal to zero. Equation 11.1 become the same equation shown in Takeda et al [1]. It means that Takeda et al [1] just showed a special case of Equation 11.1. Since $\tan^2\theta$ is alway positive, it is clear that parallel-optical-axes geometry can provide a wider range of measurement then the crossed-optical-axes geometry if the value of $l_o$ and $d$ is same.

The FTP method decodes the 3-D shape information from the phase information stored in a 2-D image with spatial carrier fringe. The phase information can be separated from the image signal by two method; the phase subtraction method and the spectrum shift method. A experimental comparison has been made between these two methods. The results show that the phase subtraction method is less susceptible to nonlinearity of recording media and systematic optical geometry error. On the other hand, the spectrum shift method is better in term of computing time and noise immune.

There are some limitations of FTP. Firstly, the object size which can be inspected in one view by FTP is limited by the area of grating image which is formed within the field of view of the camera. Furthermore, the FTP's depth of field is limited by the highest F-number of the camera. It is also found that the spatial sensitivity of FTP using CCD camera is restricted by the physical structure of the CCD array in the camera.

## 11.2 Further Work

### 11.2.1 The Further Work of Spatial Phase-stepping Technique

In chapter 6, the error analysis associated with SPS technique has been evaluated for the application of a three phase-step algorithm. Further work could now be directed towards the investigation of the errors associated with different numbers of phase-steps in the algorithm.

The SPS analysis algorithm has been generically derived from the extraction of data from carrier fringe encoded interferograms. Its capability is demonstrated by solving the transonic flow field in this thesis. In addition to flow field measurement, the SPS technique can also be used in a broad range of application. Examples are the analysis of polarisation images, in-plane deformation measurement and profilometry. Future studies should explore further the potential shown by the SPS technique.

The resolution of the available CCD cameras has increased rapidly, to the point where resolutions of 2048x2048 pixels are now commonly available. At this point it would seem possible that the CCD camera will replace today's film based technology. With such high resolution CCD camera and SPS technique, a novel instrument for real time fringe analysis operating in the industrial environment could be constructed. With the current increases in computing power, it is now possible to compute the true three-dimensional spatial position of an object. If this is coupled with the SPS analysis approach and the use of structured light, it may soon be possible for machines to have close to human optical resolution with true three-dimensional perception abilities in real time.

### 11.2.2 Fringe Pattern Analysis by Walsh Transform

The DFT, which is based on the complete orthogonal system of sine and cosine function, has been demonstrated to be a suitable analysis method for fringe patterns. The question arises whether the fringe patterns can be analysed by other mathematical transforms based on another system of functions. Since sine and cosine form a complete orthogonal functions system,

it is reasonable to investigate other mathematical transforms which rely on another system of orthogonal function.

| | Sine-cosine Functions | Walsh Functions |
|---|---|---|
| Parameters | Amplitude<br>Frequency<br>Phase<br>— | Amplitude<br>Sequency<br>Delay<br>Time base |
| Mathematical transform | Fourier transform | Walsh transform |
| Power spectrum | Frequency spectrum | Sequency spectrum |
| Filters | Time-invariable, linear | Periodically time variable, linear |
| Characterization | Frequency response of attenuation and phase shift | Sequency response of attenuation and delay |
| Modulation | Amplitude, phase, frequency modulation | Amplitude, time position, time base, code modulation |

Table 11.1: List of features of sine-cosine functions and Walsh functions

The orthogonal transform most commonly used in engineering applications are as follows: Fourier, Walsh, Haar and Cosine. Among these transform, The Walsh transform, which is based on the Walsh functions [2], are of great interest since Walsh functions are the only known functions with desirable features comparable to sine-cosine functions for use in communications [3]. Furthermore, the fast Walsh transform [4] is up to an order of magnitude faster than the fast Fourier transform because no multiplications are needed.

Table 11.1 lists features of sine-cosine functions and Walsh functions. The mathematical theory of Walsh transform corresponds to Fourier transform used for sine-cosine functions. The DFT techniques are an analytical phase measurement methods. Referencing to Table 11.1, the phase used for sine-cosine functions corresponds to the delay used for Walsh functions. It is suggested that Walsh transform could probably be employed to analyse the fringe patterns using the concept of delay measurement.

# Bibliography

[1] M. Takeda and K. Motoh, "Fourier Transform Profilometry for the Automatic Measurement of 3-D Object Shapes", Applied Optics, Vol. 22 No. 24 (1983), 3977-3982.

[2] J. L. Walsh, "A Closed Set of Orthogonal Functions", Ann. J. Math., 55, 5, 1923.

[3] H. F. Harmuth, "Applications of Walsh Functions in Communications", IEEE Spectrum, p. 82-91, Nov. 1969.

[4] D. F. Ellioot and K. R. Rao, "Fast Transforms", Academic, 1982.

# Appendix A

# List of Publications Based on the Thesis

P. H. Chan, P. J. Bryanston-Cross and S. C. Parker, "Spatial phase stepping method of fringe-pattern analysis", Optics and Lasers in Engineering, Vol. 23, p. 343-354, (1995).

P. H. Chan, P. J. Bryanston-Cross and S. C. Parker, "Analysis of fringe-pattern using a spatial phase stepping method with automatic phase unwrapping", Measurement Science and Technology, Vol. 6, No. 9, p. 1250-1259, (1995)

P. J. Bryanston-Cross, P. H. Chan and T. R. Judge, "Analysis of lateral shearing interferogram using 2-D DFT techniques", Proc. of Engineering Applications of Optical Diagnostic Techniques, 7th December, 1994, Cranfield University.

P. H. Chan, P. J. Bryanston-Cross and T. R. Judge, "Studies of Fourier Transform Profilometry", Proc. SPIE, Vol. 2178 p. 165 - 176.

# Appendix B

# Theoretical Aspect of Fluid Flow Measurement

The refractive index, $n$, of a gas is related to its density, $\rho$, by the Gladstone-Dale relationship.

$$K = \frac{n-1}{\rho} \tag{B.1}$$

where K is the Gladstone-Dale constant. A linear relationship exists between the optical density of the fluid as defined by its refractive index and its physical density. In the interferometer, a light wave passing through the test flow interferes with the reference wave which passes the outside of the test section in an environment with constant refractive index, $n_0$. If both waves are propagating along optical paths of equal lengths, the difference in their optical paths is given by

$$\triangle l(x,y) = \int_{Z_1}^{Z_2} [n(x,y,z) - n_0] dz \tag{B.2}$$

Eq. B.2 shows that the fringe field observed in the interferometer may be related directly to the dynamic density variation and to the flow characteristics within the inspection volume. Among many possible arrangements of the interferometer, the Mach-Zehnder interferometer , as shown in Figure B.1, is a popular choice for fluid flow measurement because the test beam and reference are widely separated. The following discussion is derived from Merzkirch [1].

Due to the integration in Eq. B.2, the signal obtained by the interferome-

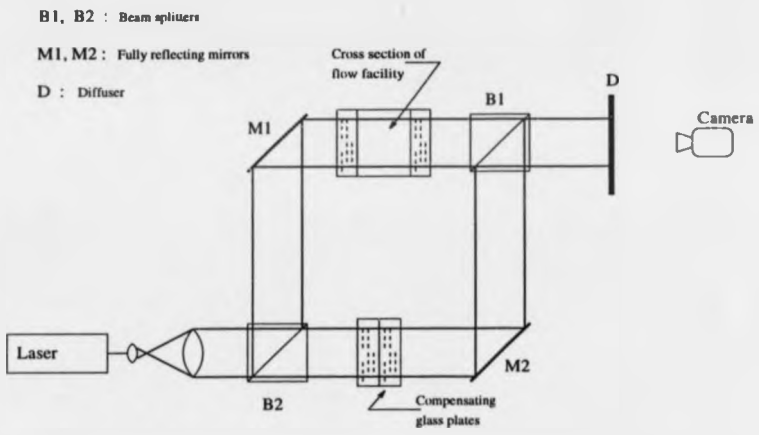



Figure B.1: Basic arrangement of a Mach-Zehnder interferometer

ter from a three-dimensional flow is two-dimensional. It is essential to direct the light beam in different directions through the fluid flow to resolve the dependence of $n$ on $z$. However, the number of necessary viewing direction is reduced to one, if the flow is axisymmetric. In the case of axisymmetric flow, the refractive index depends on the two coordinates $x$ and $r = \sqrt{y^2 + z^2}$, as shown in Figure B.2, if the light propagated again in z-direction.

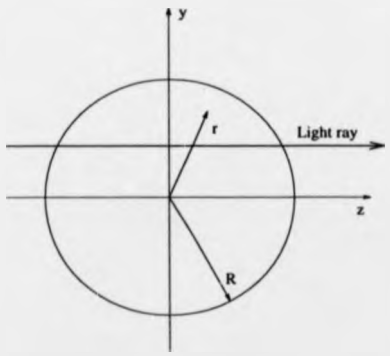



Figure B.2: Cross-section of axisymmetric flow field. The light propagates in the Z-direction

Therefore, the total phase difference, $\phi(x, y)$, for the axisymmetric case

is

$$\phi(x,y) = \frac{2\pi}{\lambda} \int_{r=y}^{R} \frac{n(x,r) - n_0}{\sqrt{r^2 - y^2}} d(r^2) \tag{B.3}$$

Where $R$ is the radius of the axisymmetric flow field and $x$ and $y$ are the coordinates in the recording plane. Eq. B.3 can be directly inverted since it is an integral equation of classical Abel type. This yields

$$n(x,r) - n_0 = -\frac{1}{2\pi} \frac{\lambda}{\pi} \frac{d}{d(r^2)} \int_{r}^{R} \frac{\phi(x,y)}{\sqrt{y^2 - r^2}} d(y^2) \tag{B.4}$$

A disadvantage of the Abel inversion is the needed differentiation, $d/d(r^2)$, of the experimental data, $\phi(x,y)$. Such a differentiation is associated with the generation of errors.

# Bibliography

[1] W. Merzkirch,"Techniques of Flow Visualization", AGAD-AG-302, ISBN 92-835-0438-0

# Appendix C

# Matlab M-files for FTP

## C.1 M-file of phase subtraction method

```
%
% M-file of phase subtraction method
%

load ref2.dat
load imag1.dat
%
for x=1:512
temp(x,:)=ref2(x,:).*papo;
end
for x=1:512
ref2f(x,:)=fft(temp(x,:));
end
clear temp
for x=1:512
temp(x,:)=imag1(x,:).*papo;
end
for x=1:512
imag1f(x,:)=fft(temp(x,:));
end
clear temp
```

```
%
clear temp
clear ref2
clear imag1
%
w=[zeros(1,62) ones(1,126) zeros(1,68)]
for x=1:512
temp(x,1:256)=ref2f(x,1:256).*w;
temp(x,257:512)=zeros(1,256);
end
for x=1:512
ref2a(x,:)=ifft(temp(x,:));
end
clear temp
%
for x=1:512
temp(x,1:256)=imag1f(x,1:256).*w;
temp(x,257:512)=zeros(1,256);
end
for x=1:512
imag1a(x,:)=ifft(temp(x,:));
end
clear temp
%
temp=imag1a.*(conj(ref2a));
resu1=imag(log(temp));
save resu1 resu1
```

## C.2 M-file of spectrum shift method

```
%
% M-file of spectrum shift method
%
load imag1.dat
```

```
for x=1:512
temp(x,:)=imag1(x,:).*papo;
end
for x=1:512
imag1f(x,:)=fft(temp(x,:));
end
clear temp
clear imag1
%
w=[zeros(1,62) ones(1,126)  zeros(1,68)];
for x=1:512
temp(x,1:256)=imag1f(x,1:256).*w;
temp(x,257:512)=zeros(1,256);
end
%
for x=1:512
temp1(x,:)=[temp(x,125:512) temp(x,1:124)];
end
clear temp
for x=1:512
imag1a(x,:)=ifft(temp1(x,:));
end
clear temp1
%
resu=imag(log(imag1a));
save resu resu
```

DX

216381