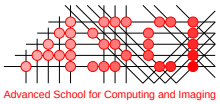
DETECTION AND MONITORING OF WEAR USING IMAGING METHODS

Jianbo Zhang

Composition of the Graduation Committee:

Prof. dr. ir. P. P. L. Regtien, University of Twente (promotor)
Dr. M. J. Korsten, University of Twente (assistant-promotor)
Prof. dr. W. L. Dalmijn, Delft University of Technology
Prof. dr. ir. J. Meijer, University of Twente
Prof. dr. ir. D. J. Schipper, University of Twente
Dr. ir. F. van der Heijden, University of Twente
Prof. dr. ir. A. J. Mouthaan, University of Twente (chairman & secretary)



This work was carried out in the graduate school ASCI. ASCI dissertation series number 128.

This research was conducted within the WEAR project (**TET. 5120**) supported by the Dutch Technology Foundation STW.

Keywords: online, wear detection, wear monitoring, image processing, Gabor filters, optimal filters, fractal analysis, reflectance model.

ISBN 90-365-2389-3

Email: Jianbo.Zhang@gmail.com

Printed by Wöhrmann Print Service, Zutphen, The Netherlands.

DETECTION AND MONITORING OF WEAR USING IMAGING METHODS

DISSERTATION

to obtain
the doctor's degree at the University of Twente,
on the authority of the rector magnificus,
prof. dr. W.H.M. Zijm,
on account of the decision of the graduation committee,
to be publicly defended
on Friday, September 8th, 2006, at 13.15

by

Jianbo Zhang

born on November 18, 1970
in Shouguang Shangdong, China

This dissertation is approved by:

Prof. dr. ir. P. P. L. Regtien	(promotor)
Dr. M. J. Korsten	(assistant promotor)

To my wife Yuanqing and my son Andy

Table of contents

List of Abbreviations and Symbols

Abbreviations

1D	one dimensional
2D	two dimensional
3D	three dimensional
ACF	autocorrelation function
AFM	atomic force microscopy
AR	autoregressive
BRDF	bi-directional reflectance distribution function
CC	cross correlation
CCD	charge coupled device
DOF	depth of focus
DSP	digital signal processor
FD	fractal dimension
FFT	fast Fourier transform
FIR	finite impulse response
FOV	field of view
IMAQ	image acquisition
IIR	infinite impulse response
MEMS	micro-electro-mechanical systems
MRF	Markov random field
PSD	power spectrum density
RMS	root mean square
SEM	scanning electron microscopy
STM	scanning tunneling microscopy
VI	virtual instrument
VIC	vertical intercept
WSS	wide sense stationary

Symbols

A_i, A_s, A_{im}	surface (image) area	45, 49, 58
c	power spectrum constant	113
d	diameter of an entrance aperture	59
D	fractal dimension	113
D_T	topological dimension	113
E	expectation	47, 86
E^*, E_1, E_2	Young's moduli of elasticity	141
$\mathcal{E}(m)$	complete elliptic integral of the second kind	51
f, u, v	spatial frequencies	76, 113
f_r	reflectance distribution	45
F	normal force or load	11
$F(\theta_i', \eta')$	Fresnel reflection coefficient	55
$g(x, y)$	Gabor function	76
$g_{mn}(x, y)$	Gabor filter bank	76
$G(u, v)$	Fourier transform of Gabor function	76
$G(\theta_i, \phi_i, \theta_r, \phi_r)$	geometrical attenuation factor	55
$\mathcal{G}(x, y)$	lowpass filter	86
h	height of ring light	49
$h(x, y)$	optimal filter	86
$\mathbf{h}$	vector form of optimal filter $h(x, y)$	87
H	hardness of material	11
$\mathbf{i}(x, y)$	vector form of a window of image $I(x, y)$	87
I_s, I_{im}	surface (image) irradiance	45, 59
I_n	image sequence	104
I_r	reference image	104
$I^r_{mn}(x, y)$	reference feature image	79
$I^d_{mn}(x, y)$	feature image	79
$J_U(\mathbf{h})$, $J_F(\mathbf{h})$	object functions	89, 90
k_{abr}, k_{adh}	wear coefficients	11, 12
L_i, L_r	radiance	45, 49
$P(f)$, $P(u, v)$	Fourier power spectra	113
r, R	radius	49, **??**, 141
$\mathbf{R}_{ii}$, $\mathbf{R}_{ii_f}$	autocorrelation matrices	88, 89
$\mathbf{R}_{ii_w}$, $\mathbf{R}_{zz}$	autocorrelation matrices	88
R_q	RMS surface roughness	46
s	sliding distance	11
T, T_k	thresholds	82, 83, 104
$\mathcal{T}$	correlation distance	47

v_1, v_2	Poisson's ratios	141
V, ν	volume	11, **??**
w	width	49
α	oblique angle of planar facet	47
β	spectral exponent	113
γ	off-axis angle	58
ε	albedo	54
η'	complex index of refraction	55
θ_i, ϕ_i	angle of incidence	45
θ_r, ϕ_r	angle of refelection	45
θ_n	angle of rotation	76
κ	scaling component	112
λ	eigenvalue	89
μ_w, μ^r_{mn}, $\mu_{w,k}$	mean values	81, 83, 87
$\mu_{b,k}$, μ_{ww}, μ_{wf}	mean values	83, 89
ξ, ξ_i	scaling factors	76, 111, 112
$\sigma_\hbar$, σ_α, σ^r_{mn}, σ_x	standard devitions	46, 47, 81, 76
σ_y, σ_u, σ_v, σ_g	standard devitions	76, 91
Φ_i, Φ_r	incident (reflected) radiant flux	45
τ	distance between two points	46
ω, ω_i, ω_r, ω_s	solid angles	45,49, 55, 58
$\hbar$	surface height	46

Chapter 1

Introduction

In this introductory chapter the challenges and significance of detection and monitoring of wear online in tribological research as well as industrial applications are presented, followed by the motivation, scope and objectives of the current research. In the end, the outline of the thesis is given.

1.1 Introduction

Wear is a type of surface damage that arises from the relative motion between interacting solid surfaces. It is a dynamic and complex process which incorporates surface and material properties, operating conditions, stresses, lubricants and geometry [6]. Wear plays an important role in determining life span of products or machine elements, and is one of the main causes of service, maintenance or finally replacement in their every aspect of life.

With the existing knowledge of materials science and mechanical tools (finite elements techniques) engineers are able to predict the functional behavior of a component of a product or process with respect to strength, fatigue etc. with reasonable accuracy. However, the lifetime of a component depends on wear. Wear is difficult to measure for its dynamic and complex nature.

Detection and monitoring of wear are rather important in tribological[1] research as well as in industrial applications. Some typical examples are: measurement of dynamics of wear processes, engineering surface inspection, coating failure detection, tool wear monitoring and so on. With wide and extensive use of wear-resistant materials, wear itself is becoming smaller and smaller, in some precision application, for instance, down to nanometer scale. Accordingly new techniques, methodologies, and instruments for detecting or monitoring of micro-wear are increasingly demanded. Due to dynamics and complex nature of a wear process, measurement of wear is usually conducted *offline* [1, 23, 57, 58, 79, 3, 82, 12, 35], i.e. during measurement the wear process needs to be interrupted and the specimen to be removed from the tester periodically to measure the evolution of wear as a function of time, number of cycles, or sliding distance. As an example, Fig. 1.1 shows some experimental results obtained from offline measurements, where wear volume of removed materials has a linear relationship with sliding distance. Tasan [83] developed a *semi-online* system to generate and measure dry sliding or rolling micro-wear in which wear is produced and measured without removing the specimen from the test rig. However, the wear process has to be stopped when measuring the profile of the specimen using interference microscope. Reliable *online* detection and monitoring of wear, in which the wear process is not interrupted and the wear environment (temperature, humidity, lubrication etc.) is not changed, remains a challenge to tribological research as well as to the industry [43, 100, 102].

This thesis will mainly focus on the investigation of new methodology, techniques, and instrument suitable for detection and monitoring of wear, particularly online and non-contact, using imaging methods.

In the next section, the motivation, research scope and objectives of the thesis

[1] Tribology is the science and technology of friction, wear and lubrication.

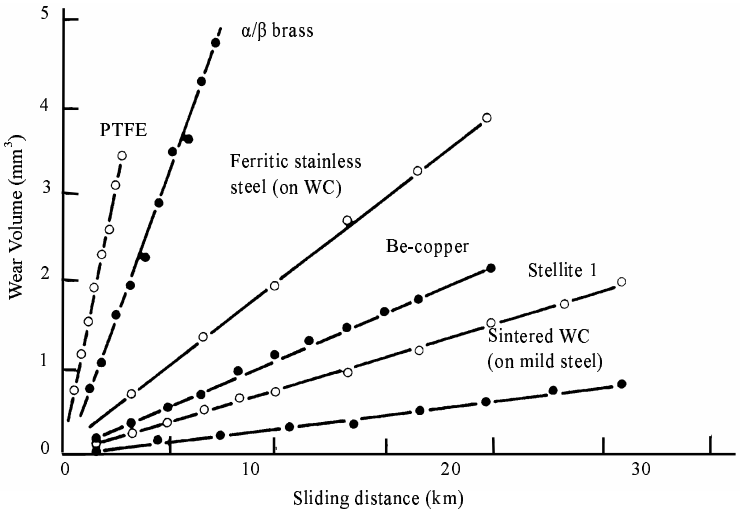


Figure 1.1: Wear volume removed from the specimen pin sliding against a tool steel ring as a function of total sliding distance from unlubricated pin-on-ring tests on the materials indicated (Adapted from [6]).

are described. Section 1.3 describes the thesis outline.

1.2 Context of the Project

1.2.1 Motivation

In tribological research, knowledge about the nature and extent of wear is of growing importance, because of the ever increasing demand for better performance of tools and components/products. On the other hand, in manufacturing or engineering industry, inspection and control of the cutting, machining or coating process to guarantee surface quality, require suitable damage/wear detection techniques/tools. Just a few applications for which no adequate measurement tools are available are:

- inspection of bore-polishing in combustion engines
- wear of electric shavers' blades
- damage of sheet products in metal forming processes

- coating failure in certain sheet production

Traditional methods to detect or quantify wear loose applicability because the amount of wear to be detected becomes smaller and smaller with the increasing wear resistance of many materials and reducing size of the components (micro-mechanisms). Besides, measuring (micro)wear during an experiment is not possible with available techniques such as interference microscopy, atomic force microscopy (AFM), scanning tunneling microscopy (STM), scanning electron microscopy (SEM) etc.

The goal of this research is the development of measurement tools (hardware and software) for the detection of micro- and macro-wear of materials online, based on imaging methods. These techniques allow faster and hence dynamic measurements of wear, and are essentially non-contact. This research is part of the project **TET. 5120** sponsored by STW (Dutch Technology Foundation). As such, other goals of the project are [83, 15, 38]:

- Further development of wear measurement method based on roughness data obtained with interference microscopy, and conducting the fast and accurate processing and analysis of large amounts of height data. Using this technique, wear measurement will take place offline, but with a high height resolution in the order of 1 nm.

- Development of techniques and instruments for studying friction in Micro-Electro-Mechanical Systems (MEMS) and for a reproducible interaction of sliding surfaces with well-defined surface properties (material and surface topology) to gain a better understanding of tribological effects in MEMS.

Eventually different paths in the big project converge and result into either a high height resolution or online wear measurement methodology.

Why Online

As mentioned in Section 1.1, wear is a dynamic and complex process which involves not only surface and material properties but operating conditions as well. Most of the existing techniques (which will be reviewed in Chapter 2) for wear measurement are offline and they have intrinsic restrictions:

- During measurement the specimen needs to be taken away from the wear tester and be measured under some instrument periodically, which is cumbersome and time-consuming.

- During mounting and dismounting steps extra damage to the specimen's surface may be introduced.

- It is hard to measure the same location of the surface. Repositioning is a big challenge.
- During measurement the wear process is interrupted and the wearing environment is hardly kept the same.
- It usually needs a well-trained operator to do the measurement (especially for interference microscopy, AFM, STM, SEM etc.). Otherwise operating errors may be introduced.

Because of the above restrictions of the existing techniques new online wear monitoring methodology, which offers higher accuracy in the sense of neither interrupting the wear process nor changing the wearing environment, is demanded in tribological research.

Why Imaging Methods

Some fundamentals of an imaging system will be addressed in Chapter 3. Before answering the question why using imaging methods, we first give the definitions of some terms associated with an imaging system and frequently used in this thesis, for the sake of avoiding ambiguity between optical imaging systems and optical profiling instruments.

An *image* is a two-dimensional (2D) (x, y) array of numerical values. An *intensity image* refers to an image where each (x, y) location represents an intensity, and is normally obtained by a video (microscope) camera or SEM. A *topographical image* or *range image* refers to an image where each (x, y) location represents a z height. This image is generally acquired by a profiling instrument, for instance by interference microcopy, STM or AFM. A *binary image* is derived from either an intensity or a topographic image. Each (x, y) location has a value of either "0" or "1", indicating which locations in the original image have some property, such as identified wear patterns in an intensity image or z height below a threshold value. In the scope of the thesis, we use *imaging* to represent a process by which intensity images are formed, unless otherwise stated explicitly.

Various research efforts have revealed that optical techniques best satisfy the requirements of online, non-contact measurement; other methods are usually deficient with regard to speed, flexibility and ability to perform non-contact measurement [91, 41]. With an optical imaging system many of the above-mentioned problems associated with existing offline techniques can be avoided. The non-contact nature of an optical sensor means that non-destructive monitoring of moving surfaces can be performed. This is important for measuring dynamics of a wear process. Optical imaging systems usually scan a finite area rather than along a line, and hence provide a more representative description of the surface.

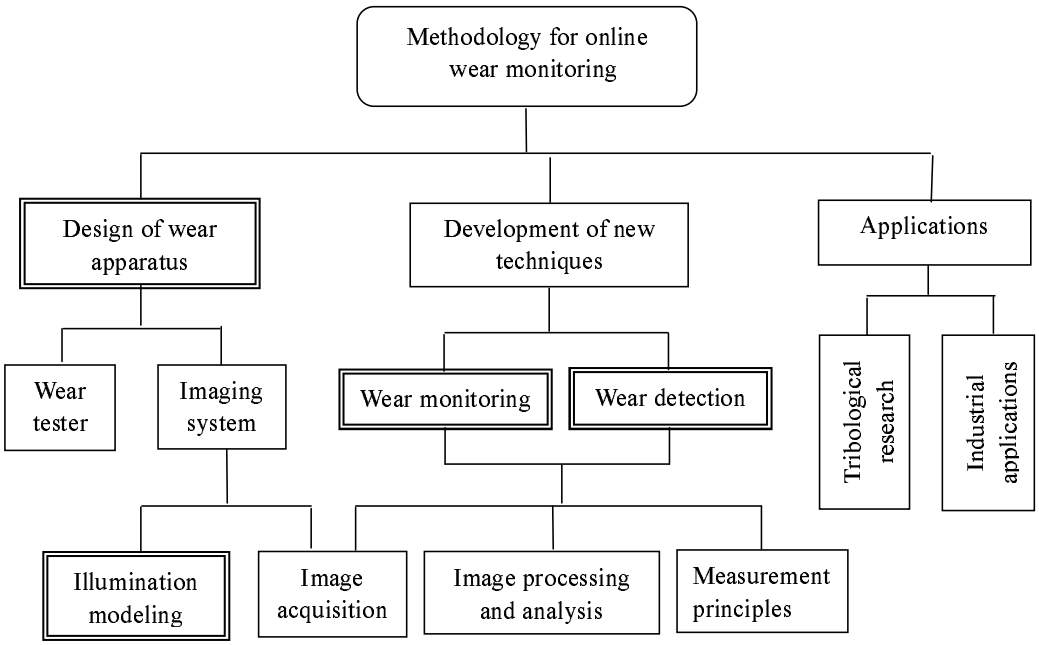


Figure 1.2: Overview of the research scope.

They also allow one to visualize and gain further understanding of the wear process by viewing and processing the temporal series of photomicrographs. More specific advantages of imaging methods will be presented in Chapter 3.

In summary, a computer-based imaging system is a good strategy for non-contact, online, and high speed inspection. It should be useful for online wear monitoring in tribological research and also suitable for industrial applications in an automated environment, where high speed, high accuracy and low cost are required.

1.2.2 Scope

The main goal of this research is to develop new methodology for online wear monitoring. The new methodology consists of design of an online wear testing apparatus, techniques for online wear detection and monitoring, and their applications. The overview of the research scope is shown in Fig. 1.2.

A wear apparatus which combines a wear tester and an imaging system will be designed. The wear tester should be capable of simulating different contact modes (point contact, line contact, conforming contact, for instance) and various operating conditions (varying load and velocity for instance). The imaging system

with proper illumination will be used to acquire real time images of the dynamic wear process generated by the wear tester. As illumination plays a crucial role in the imaging system, it will be modeled and evaluated. As such, non-contact, online, low cost and high speed will be necessary merits of the wear apparatus to be designed.

Development of new techniques for online wear detection and monitoring[2] will be the main theme of this research. These imaging-based techniques involve image acquisition, image processing, and associated measurement principles, such as filtering scheme, fractal analysis, etc.

With the developed techniques, identification of wear regions, observing some wear behavior, study of dynamics of particular wear processes, or estimation of degree of wear will become possible. This will benefit tribological research as well as wear inspection oriented applications: such as coating failure detection, engineering surface inspection, etc.

The highlighted items with double frames in Fig. 1.2 constitute the main part of the thesis.

1.2.3 Objectives

From the above discussion, the objectives of the thesis can be formulated below:

- Design of a simple, laboratory-use, cost-effective wear apparatus for wear testing and monitoring.
- Developing computer vision methods for the determination of wear in micro- and macro-structures, based on optical images. The method comprises both selecting a suitable measurement principle, data acquisition and data processing.
- These measurements take place online and are suitable to monitor the changes in the condition of a surface because of wear.

1.3 Thesis Outline

To describe the development of the new methodology for online monitoring, and to apply this methodology for tribological research or wear inspection-oriented applications, the thesis is structured as follows:

[2]The particular meaning of detection and monitoring in this thesis will be given in Chapter 3.

Chapter 2 provides an overview of the existing approaches to wear monitoring. The advantages and disadvantages of each approach are summarized.

Chapter 3 describes the design of a laboratory-use wear apparatus, including hardware and software, which aims to prevent the disadvantages of the offline methods discussed in Chapter 2. An overall perspective of the new methodology is also provided in this chapter.

Chapter 4 discusses choice of illumination strategy for the current imaging system. According to the strategy, ring light illumination is employed. An illumination model and a reflectance model based on Torrance-Sparrow model are presented and evaluated for ring light illumination.

Chapter 5 presents techniques for online wear detection, where two filtering approaches, namely unsupervised detection scheme using multichannel Gabor filters and supervised detection scheme using optimized filters are proposed, evaluated and compared.

Chapter 6 investigates techniques for monitoring of a dynamic wear process, dry sliding wear in particular. Two indicators of state of wear, derived from fractal analysis, are proposed and evaluated.

Finally conclusions and recommendations are provided in **Chapter 7**.

Chapter 2

Overview of Wear Monitoring

This chapter presents an overview of the existing approaches to wear measurement. For each approach, measurement principles, measurement range and resolution are addressed. The strengths and weaknesses of each approach are also summarized.

2.1 Introduction

Wear is the surface damage or removal of material from one or both solid surfaces as a result of relative motion to one another. Wear can be mainly classified into three groups [70, 3]:

- One term is the appearance of scar: e.g. pitted, spilled, scratched, polished, crazed, gouged and scuffed;
- The second way is in terms of physical mechanisms which remove the material or cause the damage, e.g. adhesion, abrasion, fatigue and oxidation;
- And the third is the conditions surrounding the wear situations, e.g. lubricated wear, unlubricated wear, metal-to-metal sliding wear, rolling wear, high stress sliding wear and high temperature metallic wear.

Wear is a dynamic and complex process which incorporates surface and material properties, operating conditions, stresses, lubricant oil film and geometry. Wear, as friction, is not a material property, but a system response. Operating conditions affect interface wear.

Due to complexity of a wear process, monitoring of wear is not an easy task. Different approaches and different measures have been used to determine amounts of wear, both qualitatively and quantitatively. A recent review of wear measurement techniques can be found in [78].

This chapter presents an overview of wear monitoring. General concepts regarding contact area and wear mechanisms are described in Section 2.2. Various approaches to wear measurement, measurement principle, range and resolution of each approach, as well as its associated advantages and disadvantages are presented in Section 2.3. Discussions about the existing approaches are given in Section 2.4, followed by conclusions in Section 2.5.

2.2 Contact Area and Wear Mechanisms

Before addressing approaches to wear monitoring, it is necessary to introduce some general concepts regarding the nature of the contact between two surfaces and wear mechanisms. The first aspect that will be considered is the area of contact.

Surfaces consist of hills (*asperities*) of varying heights and spacing and valleys of varying depths and spacing. When two nominally flat surfaces are placed in contact, due to surface roughness physical contact takes place at localized spots within the area that is defined by the macro-geometry. These points at which the actual contact occurs are referred to as *junctions*. The sum of the areas of all

the junctions constitutes the *real area of contact*. For most materials with applied load, this will be only a small fraction of the *apparent area of contact* that is determined through macro-geometry. Fig. 2.1 illustrates the situation.

The real area of contact is a function of the surface texture, material properties and interfacial loading conditions [6]. The ratio of real area of contact to apparent area of contact is as small as 10^{-4} in practical situations and the diameter of typical junctions is in the range of 1 to 100 microns [71, 45, 3]. The larger value mostly occurs for a very rough surface and high loads. Diameters of the order of 10 μm are more typical for normal contact situations.

The Hertz contact theory [26] is frequently used to determine stress levels as well as the size of contact regions. Appendix B presents how to use this theory to determine contact radius, i.e. *Hertzian radius*.

Repeated surface interactions and surface and subsurface stresses, developed at the interface, result in wear.

Due to complicity of a wear process, no general equations exist for characterizing all types of wear. For adhesive and abrasive wear, however, Archard's [1] equation proved to characterize these two main wear mechanisms quite well.

Adhesive wear occurs when two nominally flat solid bodies are in sliding contact, whether lubricated or not. It is formed due to localized bonding between contacting solid surfaces leading to material transfer between the two surfaces or the loss from either surface (see Fig. 2.2).

Based on experimental data of various unlubricated material pairs, the vast majority being metallic, Archard [1] first developed a model to derive the law of adhesive wear, frequently referred to as *Archard's equation*:

$$V = k_{adh}\frac{Fs}{H} \tag{2.1}$$

This equation describes that the amount of wear (volume V) is generally proportional to the applied load F and sliding distance s and generally inversely proportional to the hardness H of the surface being worn away, where k_{adh} is nondimensional wear coefficient dependent on the materials in contact. Fig. 1.1

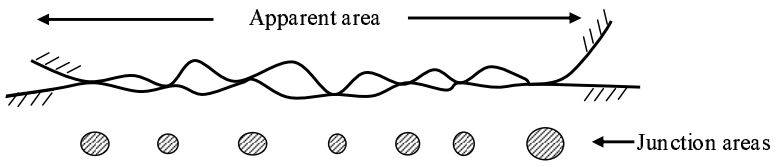


Figure 2.1: Schematic representation of the contact area between two rough surfaces.

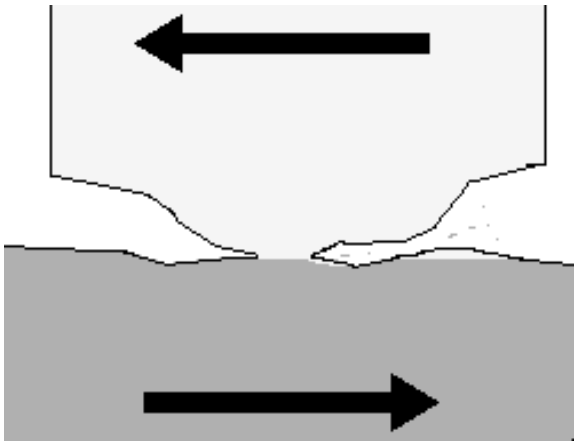


Figure 2.2: Adhesive wear is produced by the formation and subsequent shearing of welded junctions between two sliding surfaces.

shows the wear data obtained from pin-on-ring tests for a wide range of material combinations under unlubricated conditions in air.

Abrasive wear, on the other hand, is formed due to hard particles or hard protuberances forced against and moving along a solid surface. Material removal from a surface via plastic deformation can occur by several deformation modes which include plowing, wedge formation and cutting (see Fig. 2.3). An equation of the form similar to Archard's equation for adhesive wear is found to cover a wide range of abrasive situations, and is [3, 6]:

$$V = k_{abr} \frac{Fs}{H} \tag{2.2}$$

where k_{abr} is a nondimensional wear coefficient. From Eqs. (2.1) and (2.2), it can be deduced that the *wear rate* of sliding contact, namely wear volume over sliding distance V/s, is linearly dependent on either F or H, or some combination.

2.3 Approaches to Wear Measurement

In general the approaches to wear measurement can be classified into 6 types in terms of wear measures: mass difference, linear wear dimension, wear area, wear volume, topographical difference and other indirect measures. In this section, we will present each type of approaches in more detail.

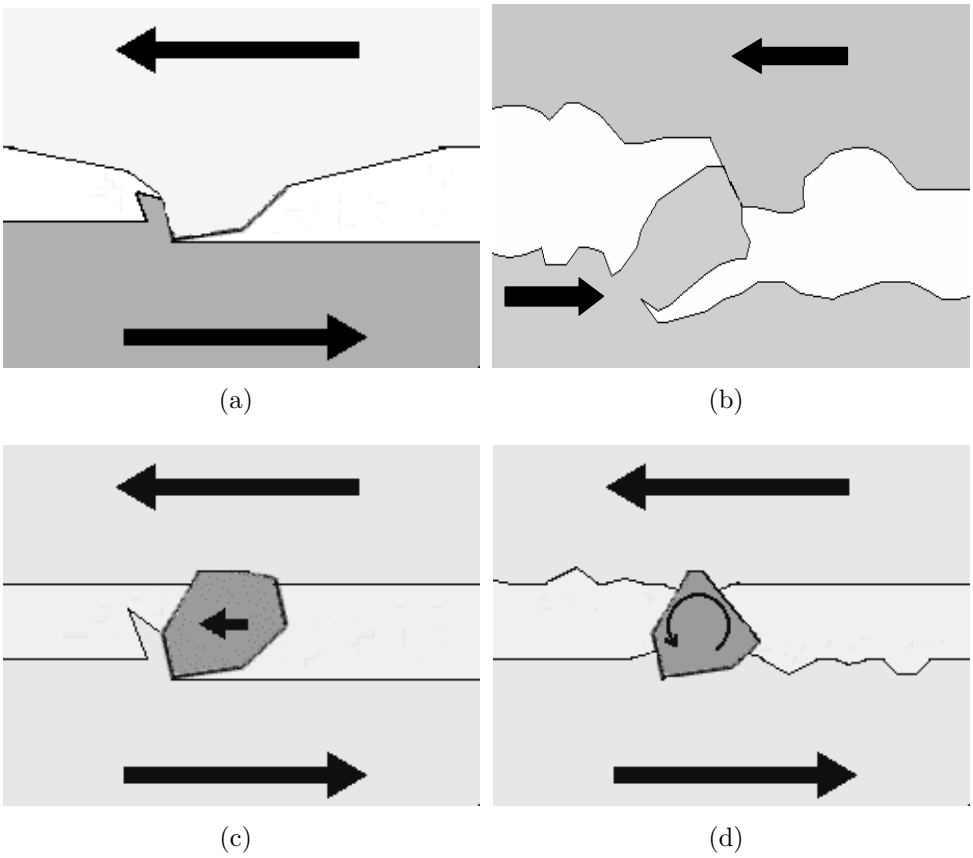


Figure 2.3: Two body abrasive wear occurs when one surface (usually harder than the second) cuts material away from the second (a) (b). This mechanism very often changes to three body abrasion as the wear debris then acts as an abrasive between the two surfaces (c) (d).

2.3.1 Mass Difference

Wear, whether large or small, usually leads to surface damage. The most common form of that damage is material loss which causes either mass change or dimensional change, which is discussed in the next subsection. Common examples include abrasive wear, filing, sanding and grinding for instance. Accordingly the most straightforward approach to quantify wear is weighing the sample before and after the wear test.

The amount of wear can be described by the absolute mass loss (in grams), or the rate of mass loss per unit of usage (grams per day), or by a fractional change in the mass of the part involved (for instance, 1% change per 200 hours of operation). Most ASTM[1] wear standards, however, report unit of wear in cubic millimeters of volume, rather than mass, so that materials with different densities can be better compared.

This approach is frequently used because of its relative simplicity and ease of performing a mass difference measurement [79, 3]. This approach is suitable for many laboratory tests as well as in real applications if the worn component is detachable and the mass difference is not too small relative to the total mass of the component.

However, with the wide use of wear resistant material, wear is becoming smaller and smaller. It is necessary to use mass measuring instruments with increased sensitivity. At some point, mass loss becomes too small for this approach to be feasible.

Another problem with this approach is that direct comparison of mass difference of materials can only be done if their densities are same. For bulk materials this is not a major obstacle. In the case of coatings however, this can be a major problem since their densities may not be known or easily determined.

The third problem is that the mass of the components may vary due to changing degrees of water absorption (notably for polymeric materials), plus absorption of water and other contaminants.

Further, a mass measurement does not measure plastically displaced material. And it is also sensitive to wear debris and transferred material that becomes attached to the surface and can not be removed.

Measuring mass difference gives no information about the distribution of wear over the component. This is a serious disadvantage in most applications since the wear life normally is limited by the maximum wear at some critical location rather than on the total wear.

[1] American Society for Testing and Materials.

2.3.2 Wear Dimension

An alternative approach to the mass loss measure of wear is based on linear dimensional measurements, such as wear scar depth or width. In many engineering applications, the concern of the component which is subject to wear is generally with the loss of a dimension, or a change in contour, not a mass loss per se. In such cases, monitoring of the dimensions of a part is a natural approach to assessing the amount of wear encountered. Frequently, such measurements lead to a replacement or service of a component.

The dimensional measurements may be conducted either continuously (online) during wear or by measuring the dimension change after wear. Matsunaga et al. [57] measured diameter changes of a steel pin wear scar. A novel use of wear dimension measure was developed by Glaeser [23] who placed microhardness indentations on the surface of a bronze bushing and measured their size changes during the running time of the bushing as an indicator of wear. One example of continuous measurements is measuring the pin length reduction in a pin-on-disk test using a high resolution displacement transducer [5]. This is a simple and elegant method, but is only applicable at relatively high wear rates and under laboratory conditions.

Although this wear dimension approach is frequently used, very small amounts of wear are difficult to measure, much as in the case of wear measurements using the mass loss method.

2.3.3 Wear Area

For some contact geometries, wear produces material loss over a restricted area. In some cases, those areas of wear can be measured and are proportional to the amount of wear. Examples would include worn areas on gear teeth, on bearings retainers and on sliding pads with contoured surfaces. If the curvature of the surface is known, then the amount of wear can be quantified on the basis of the area worn. Because many tribological components involve area contacts, in contrast with point and line contacts, area measure of wear is intuitive, easy and important. Mehrotra [58] did a laboratory study of wear using a block and ring test method involved ceramic materials where wear volume was calculated based on actual measurements of the scar width and length, and on the assumption that the scar surface curvature matched that of an unworn ring. In general, it is necessary to determine the surface profile associated with the worn region, which will be discussed in Subsection 2.3.5. Based on profile measurements, approximate area measures for calculating wear may or may not justified.

2.3.4 Wear Volume

Since most wear processes involve loss or displacement of material, volume can be used as an important measure of wear, i.e. volume of material removed or volume of material displaced. Besides, in some cases, it is necessary to directly measure wear volume. This generally occurs when the worn region is very irregular or asymmetric in shape, or when high accuracy in the result is needed. For scientific purposes this is the frequently used measure to quantify wear. Most wear standards in ASTM report wear in volume and many models of wear mechanisms are formulated in terms of volume, Archard's equation for adhesive wear [1], for instance.

Wear volume is the fundamental measure for wear when wear is equated with loss or displacement of material. This is the case most frequently encountered in engineering applications [3].

The volume measure of wear better enables a comparison among materials having different densities and also permits easy calculation of linear wear amounts or wear allowances [3]. The major disadvantages of this approach are: (1) the measurements are usually time-consuming; (2) it does not measure wear distribution.

2.3.5 Topographical Difference

Using the approaches based on the measurement of mass differences, volume differences, or displacements generated by wear, only global information about wear can be obtained. Local information, however, is more helpful and valuable since wear takes place at a microscopic scale between contacting asperities of the surfaces in contact. The surface profile has a big influence on the tribological properties and service life of machine elements. Therefore, measurement of topographical profile changes of finished or worn surfaces is frequently carried out in engineering and research [98, 22, 83, 38].

Since wear generally modifies surface geometry, the most widely used and sophisticated approach to monitoring wear is to directly determine the elevation of surface points by which the amount of wear is related to the topographical difference before and after wear. There are a lot of methods to measure a surface profile. These methods are generally named as range measurement methods as all of them produce elevation or range maps, typically in the form of a pixel array whose values represent surface elevation. In the ideal case, the pixels are square, indicating that the measurements have the same horizontal resolution in both x and y directions. Range measurement may be done by physical contact, for instance, using a contact profilometer (a stylus, for instance) to scan the elevation profile along a linear traverse, or using non-contact methods, for instance, using optical methods. In the following we present these methods in detail.

Contact Profilometry

The principle of contact profilometry is measuring surface profile along a line using a contact stylus. The strengths of this approach are: (1) simple; (2) high vertical resolution (0.1 nm) and wide range (vertical: as much as 100 μm; horizontal: 100 mm). Its weaknesses are: (1) the shape and size of the stylus determines its lateral resolution and limits the steepest slope that can be followed; (2) the stylus may add extra damage to the surface of a soft sample; (3) must be used in a fairly quiet and clean environment.

Noncontact Profilometry

STM and AFM belong to noncontact profilometry, which is in some respects an extension to the contact profilometry to smaller dimensions. The probe tip of AFM/STM is smaller and can therefore record smaller details, in the range of nanometer. Unlike the profilometer, in which the stylus tip is held in contact with the surface by mechanical forces, the AFM and STM rely on the interaction of the electron clouds around atoms in the surface and the tip. For STM, the scanning tip is moved to maintain a constant electron tunneling current between the tip and specimen. Accordingly it requires that the specimen be electrically conducting. The AFM/STM is used to scan a complete raster over a surface to produce a topographical image instead of a single elevation profile. AFM/STM is widely used to measure microscopic wear in tribological research [80, 82, 15].

The strengths of AFM/STM are: (1) capable of atomic resolution; (2) noncontact and nondestructive. The weaknesses are: (1) sensitive to vibration, must work under vibration isolation condition; (2) operation is complicated, experience and skills are needed to correctly focus and align these instruments on the sample; (3) narrow measurement range (vertical: up to 7 μm; lateral: up to 200 μm).

Interferometry

Interferometry derives surface profile data by interferometric comparison to a reference surface. This technique is widely used in tribological research to measure wear at asperity level [97, 12, 83, 38].

The strengths of this technique are: (1) noncontact and nondestructive; (2) comparatively faster (than AFM/STM) since an entire surface can be measured at once, with a lateral precision corresponding to the resolution of a conventional light microscope, or better than 1 μm; (3) high vertical resolution (down to 0.1 nm). The weaknesses are: (1) sensitive to ambient vibrations; (2) not suitable for highly curved surfaces; (3) changes in surface reflectivity due to changes in composition or local contamination present problems.

Stereoscopic Vision

Stereoscopic vision is a technique that reconstructs the 3D scene from 2D image series in which two cameras, or one camera from two positions can give relative depth or absolute 3D location. Once a topographical image is obtained, surface height and roughness can be computed to estimate wear.

Stereoscopic vision requires substantial computation to match each small image patch in one image with all the possible patches in the other image. Also, even after this matching is done, the best match at each point is sometimes incorrect due to image noise, occlusion, or perspective changes between the camera viewpoints. In addition, the resulting depth resolution is still much poorer than the lateral resolution. For wear measurement, an even more complicated problem arises because the subsequent images have been taken in different situations (at least one mounting and dismounting step is in between two successive images).

There are also some other range measurement methods, such as photometric stereo, structured light, confocal scanning light microscopy, acoustic microscopy and laser triangulation. Because of their respective inherent drawbacks, they are rarely used for wear measurement.

2.3.6 Indirect Approaches

In the above, we present some direct approaches to wear measurement in terms of direct measures of wear. These measures, mass, dimension, area, volume may convert from one to another in some cases if contact geometrical shape or density of the specimen is known.

There exist also some other indirect measures, such as time required to wear through a coating or load required to cause severe wear or a change in surface finish. In addition, some other indirect approaches to wear monitoring include vibration analysis [56], lubricant analysis [49], motor-current signature analysis [18], etc.

2.4 Discussions

The state of the art approaches to wear measurement are topographic methods using STM/AFM and interference microscopy. Such methods can be used to conduct high resolution measurement. They are mainly limited to fundamental research, i.e. measuring wear of some materials on a nanometer scale. On the other hand, the modern instruments used in such methods are sensitive to vibration, complicated to operate, comparatively slow and high cost, and not suitable for industrial applications.

Besides, nearly all the approaches to wear measurement discussed above are offline, i.e. with these methods, the wear process has to be interrupted and wear conditions are changed during the mounting and dismounting steps in the measurement. In this sense, offline methods are inaccurate, not suitable for measuring a wearing surface in motion.

Non-contact, low cost, high speed and reasonably accurate wear monitoring methods, particularly online, are highly demanded in tribological research as well as in industrial applications. To overcome the inherent drawback of offline methods, in this thesis a new methodology for online wear detection and monitoring, based on imaging methods is developed. These techniques allow faster and hence dynamic measurements of wear, and are essentially noncontact and low cost.

2.5 Conclusion

In this chapter an overview of existing methods for wear monitoring is presented. These methods are categorized into 6 types in terms of different wear measures. For each approach, its measurement principle, measurement range, resolution as well as its advantages and disadvantages are presented. From this overview, we may conclude that nearly all the existing approaches are offline and can not be used to measure wear of a surface in motion, ie. online. Non-contact, fast, accurate and cost-effective wear monitoring and detection methods, especially online, are demanded both in tribological research and in industrial applications.

Chapter 3

Design of an Online Wear Monitoring System

This chapter describes the online methodology for detection and monitoring of wear which aims to prevent the disadvantages of offline methods. Some basics of an imaging system are first presented. Based on the project targets formulated in Chapter 1, the design of a wear testing apparatus, including its hardware and software, is described. Some observations about wear and image sequences are given. In addition, an overall perspective of the new methodology can also be obtained in this chapter.

3.1 Introduction

Surface geometry and topography are important factors, among others, that affect the wear rate of a machine element; various wear mechanisms, whether adhesive or abrasive, involve net material loss or material redistribution, thus causing changes of surface topography and geometry the other way around. Therefore, by measuring and comparing worn and unworn values, either deterministically or statistically, of surface geometry, surface roughness or surface local height, for instance, one can monitor, detect, or quantify wear.

A rough surface can be considered as a two dimensional random signal. Many statistical parameters have been developed to characterize random nature of a surface, for example, roughness, height distribution function, autocorrelation function, power spectrum, skewness, kurtosis, etc [84, 12]. All these parameters can be measured by a lot of techniques, among which optical imaging techniques are widely adopted for their advantages, such as non-contact, fast, accurate, etc. Luk, Huynh and North [50] used a machine vision system to measure surface roughness of tool-steel in a production environment. Kiran, Ramamoorthy and Radhakrishnan [44] gave a review of evaluating surfaces of medium finished parts. Younis [99] proposed an online surface roughness measurement technique using image processing methods. Recently Wang and Wong developed a technique for measuring roughness on moving surfaces using the combined effects of interference and light scattering [92, 94].

In an optical imaging system, images are the physical carriers of surface information. Besides, images can be easily transformed into a form suitable for further (fast) processing. Hence, applying imaging methods for online, automated detection and monitoring of wear on a surface becomes possible. This is the theme of this thesis.

As stated in the previous chapter, offline methods for wear monitoring have their inherent disadvantages. To avoid interruption of wear processes, an online methodology is presented for detection and monitoring of wear by an imaging system. The basics of an imaging system is stated in in Section 3.2. Next, according to the project targets and requirements, the design of an online wear monitoring system, including both hardware and software, are described in detail in Section 3.3. Section 3.4 introduces image registration which is a preliminary operation in image processing and will be used in following chapters. Finally, some observations about wear and image sequences are given in section 3.5 followed by a summary in Section 3.6.

3.2 Monitoring of Wear by an Imaging System

An imaging system usually consists of three main parts: scene, light source and image formation device [89]. *Scene* is referred to as an object or a collection of objects existing in the real 3-dimensional (3D) world and its/their surroundings. The scene of our interest in this thesis is the surface of a specimen which is suffering from wear. And an image is often formed by means of electromagnetic radiation: for instance, visible light, microwave (radar), infrared (thermography), X-rays and γ rays. The radiation from the light source interacts with the surfaces of the objects. Then part of the radiation is captured by an image formation device, which finally yields a 2D distribution of the radiation defined on the image plane. This distribution of radiation is called the *image*[1] of the scene. A sampled and quantized version of the 2D distribution is called a *digital image*. It is an array of numbers that represents a digital version of the image introduced above. The advantage of having such a representation is that it can be manipulated by a computer equipment, for instance, transforming one digital image into another (referred to as *image processing*).

An image formation process is a physical process that maps a 3D scene to a 2D image. A thorough understanding of it will help us design a suitable specific imaging system as well as interpret images meaningfully. Generally speaking, an image formation process involves interaction between light and a surface (illumination and reflection) and projection (camera model).

When electromagnetic radiation is incident on a rough surface, a portion of its energy, depending on the local angle of incidence, the polarization of the incident wave, and the electrical properties (permittivity, permeability, and conductivity) of the surface medium, will be reflected. The reflected beam will in return carry information about the surface. The radiation may be reflected either specularly or diffusely or both. Clearly, the relationship between the wavelength of radiation and the texture of the surface will affect the physics of reflection. Through a lens system, the reflected radiant energy can be projected to an image plane, where a film or a charge coupled device (CCD) is exposed. As such, 2D images of a 3D scene are formed. The interaction between light and a surface, i.e. illumination and reflection, will be dealt with in detail in Chapter 4 for its importance to the performance of a whole imaging system.

As stated in chapter 1, the main goal of this project is to investigate techniques of detection and monitoring of wear, particulary online, using imaging methods. First of all, it is necessary to make clear what detection and monitoring exactly

[1] It is referred to as intensity image in Chapter 1.

mean in this thesis.

Wear is a kind of surface damage mostly caused by relative motion. During a wear process, a wear scar or wear track with a particular pattern or shape will be present on the wearing surface. The evolution of the wear scar or wear track, either its size, depth, or shape, will reflect the wear status of the surface. Therefore, surveilling the surface by an imaging system is an intuitive solution for online detection and monitoring of wear.

Detection is meant to recognize, identify or separate the worn region, a wear scar or wear track, from the wearing physical surface. Most time we mean by detection in this thesis to identify the worn region in the image of the wearing surface. So our concerns about wear detection are:

- Whether or not wear takes place on the surface;
- If the answer is yes, what is its location, size or shape?

Monitoring on the other hand means to continuously inspect a specific wear process so that:

- The wear curve of particular material or contact mode can be drawn;
- Some wear phenomena can be observed;
- The relative wear degree/state of the surface can be known. Based on the wear status, the corresponding machine elements/parts can be warned to be maintained or serviced.

The main difference between detection and monitoring is that monitoring is more concerned with the dynamics of a specific wear process.

From the overview in chapter 2, one knows that wear monitoring can be done in many ways. So the question can be raised why using imaging and images for wear monitoring in this project. In Chapter 1 some answers have already been given. To be complete, the answers are further summarized below:

- By imaging means an automated monitoring/measurement system can be realized. It prevents any mechanical contact between the surface being monitored and the sensors of the monitoring system. It is, therefore, non-destructive so that no extra damage to the surface will be introduced.
- By imaging means (a video zoom microscope, for instance) perceiving details of a surface with small amount of wear (of the order of magnitude of a few microns[2]) which is too small for human vision.

[2]It depends on the lateral resolution of the imaging device.

- By analysis of temporal sequences of images, taken from the surface under test, the dynamics of the specific wear process can be studied through observing the gradual changes in the image set.
- An imaging system allows one to visualize and gain further understanding of the wear process by viewing and processing the temporal series of photomicrographs.
- Its online nature makes it not to disrupt the ongoing wear process. In this sense the measurement results are more correct than those achieved through offline methods.
- Digitized images can be fast processed by a computer using fast algorithms, therefore realtime monitoring of wear becomes possible.

Aside from the above-mentioned pros of the imaging means for wear monitoring, it, inevitably, also has some cons, for instance:

- An imaging system is sensitive to variance of illumination, hence a stable, uniform, and controlled illumination environment is needed.
- Wear area in an image is an indirect measure of which on a physical surface. To determine the correlation between them is not easy.
- Presence of wear particles in the image affects the result of wear detection. This problem can be solved by adding a mechanical component, say a blowing nozzle, to clean the surface.
- Surface contamination may mislead to the detection of wear in images.

Due to the limitation of imaging methods, we restrict our research on dry sliding wear processes only in this project.

According to the objectives formulated in Chapter 1, the design of the online wear monitoring system will be addressed in the next section.

3.3 System Design

As stated in chapter 2, wear measurement is traditionally done offline, i.e. during the measurement the wear process is interrupted. The operating conditions and environments of the wear process are changed as well. In this sense, the offline measurement is not accurate. As mentioned earlier, one of the objectives of this project is to design a simple, laboratory-use, cost-effective apparatus for online wear testing and monitoring, which should provide possibilities for studying and

monitoring of dynamic wear processes. The requirements to this apparatus are summarized bellow.

- Contact geometry is a critical factor in wear testing. The apparatus must be capable of simulating three types of sliding contacts: point contact (ball-on-disk), line contact (cylinder-on-disk) and conforming contact (flat-on-disk).
- Other factors besides contact type that significantly influence a sliding wear process include load, velocity, lubrication condition, and temperature. The apparatus must be able to vary load and velocity, but without lubrication and under room temperature.
- The surface of the wearing specimen must be monitored online, i.e. the wear process should not be interrupted and operating conditions should not be changed during wear monitoring.

3.3.1 Hardware Design

Following the above requirements, a simple, laboratory-use apparatus is designed, which combines a simplified Pin-on-Disk wear tester and an imaging system. Fig. 3.1 shows the schematic diagram of this apparatus, while the real picture of which is shown in Fig. 3.2.

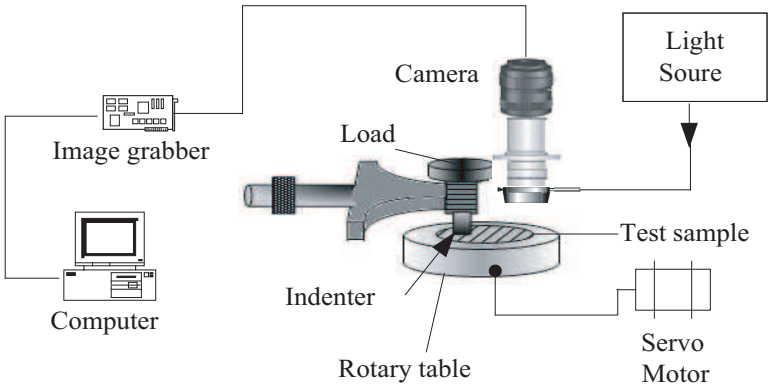


Figure 3.1: Schematic diagram of the wear testing and monitoring apparatus.

Simplified Pin-on-Disk Wear Tester

In the Pin-on-Disk wear tester, the pin is held stationary and the disk rotates. A wear test is conducted by loading a flat, cylinder or sphere shaped indenter (pin),

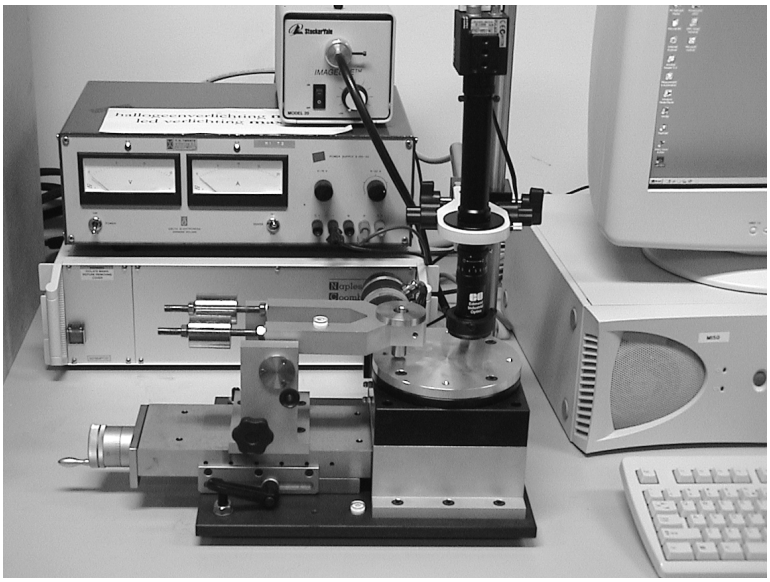


Figure 3.2: Picture of the wear testing and monitoring apparatus.

which is mounted on a stiff lever, onto a test sample that is fixed on a motor-driven rotary table. A servo motor controller is used to vary the rotational speed of the table. The load applied on the test sample is varied by removing or adding weights on the lever. The surface of the test sample is ground or machined. The pin made of harder material is used to evoke wear on the sample.

The rotary table is also a positioning stage which allows precise rotary positioning of objects by servo motor control. A precision worm and wheel with a 90:1 gear ratio provides repeatable, zero-backlash positioning with high load capacity. Servo motor control with a rear-mounted optical encoder ensures accurate indexing.

As shown in Fig. 3.3, the motion controller is part of the motion control system which includes amplifiers, motors and encoders. The motor converts current into torque which produces motion. The power amplifier converts a ± 10 Volt signal from the controller into current to drive the motor. The encoder translates motion into electrical pulses which are fed back into the controller.

Image Acquisition

The wearing surface is monitored by an imaging system, which is a combination of a video zoom microscope and a monochrome CCD sensor. Some specifications of the imaging system is listed in Table 3.1. The magnification of the imaging

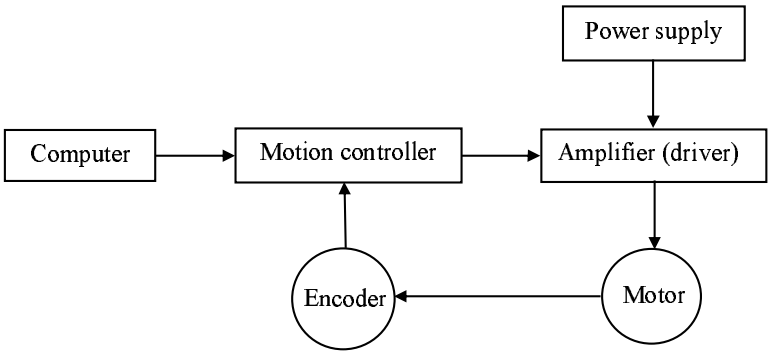


Figure 3.3: Elements of servo systems.

system is 2.5X to 10X, corresponding system resolution (lateral) $9.8\mu m$ to $4.4\mu m$. The optical axis of the system is perpendicular to the nominal plane of the test surface. Image sequences of the surface are captured through a frame grabber into the computer in real time.

Table 3.1: Specifications of the imaging system

Imager	2/3" progressive scan CCD
Active area	8.7mm × 6.9mm
Cell size	$6.7\mu m \times 6.7\mu m$
Active pixels	1300 (H) × 1030 (V)
Frame rate	15Hz
Primary magnification	2.5X − 10X
Object resolution	$9.8\mu m$ (2.5X) − $4.4\mu m$ (10X)
Field of View	3.5mm (H) × 2.8mm (2.5X) 0.87mm(H) × 0.69mm (V) (10X)
Depth of Focus	$30\mu m$ (2.5X); $6\mu m$ (10X)

The wear status of the surface of the test sample can be obtained by further processing of the images.

The image acquisition board employed is PCI-1428 from National Instruments Corporation. The PCI-1428 acquires digital images in real time and stores the images in onboard frame memory or transfers them directly to system memory. It has a high-speed data flow path, ideal for both industrial and scientific environments.

A 150-watt fiber optic illuminator, through a ring guide that is mounted at the end of the microscope, provides glare free and uniform illumination for the system. The intensity can be adjusted manually. Since illumination plays an important role on image quality, modeling of this illumination is addressed in detail in chapter 4.

Synchronization

The present wear apparatus combines motion control of a rotary table with image acquisition of a wearing surface. One of the challenges in integrating different processes is getting them to synchronize and work together. For instance, to image the same area of a moving surface after certain number of cycles (rotations). This is crucial for the study of the dynamics of the wear process by comparing the gradual changes of images.

In the rotary table, an optical encoder, which translates motion into electrical pulses is used. It provides feedback to the motion controller. The encoder density is 360,000 counts/revolution and the positioning accuracy is 2 arc minutes. For synchronization, a trigger line for synchronization is added in between the rotary table and the image grabber. Once the table moves to the same position in terms of the encoder count, a trigger signal is generated to trigger image acquisition. This can be implemented by programming with motion control software, which will be addressed next.

On the other hand, the trigger signal should also asynchronously reset the camera's scanning and shutter control, thus the acquisition starts immediately when the external trigger line (of the image grabber) is enabled.

In terms of the encoder density, positioning accuracy, position of the wear track, and the parameters of the camera, the ideal pixel shift can be calculated as follows:

> Positioning accuracy: $ac = 2\ arc\ minutes = 2/60/180 = 1/5400\ rad.$
> Pixel size of the CCD sensor: $6.7 * 6.7\ \mu m^2$.
> If the radius of the wear track is $r = 48\ mm$, then linear positioning accuracy: $lpa = ac * r = 8.9\ \mu m$.
> Hence, at magnification of 2.5X, the positioning accuracy in pixel level is about $8.9/(6.7/2.5) \approx 3.3\ pixels$.

This is the system positioning error in pixels. In practice, even larger pixel shifts among initial and successive images are found. This is mainly because of vibration and thermal expansion of the aluminum mounts that hold the specimen and the imaging system. The remaining pixel shift after synchronization can be

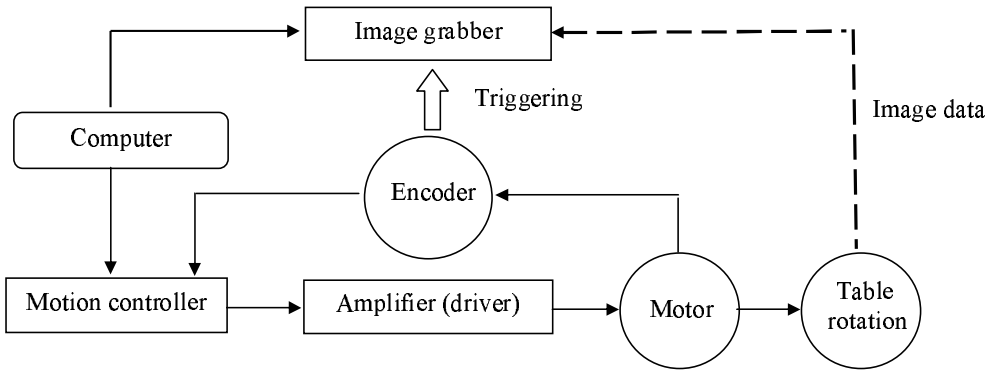


Figure 3.4: Synchronization of table rotation and image acquisition.

removed by *image registration*, an operation of aligning two or more images from the same surface, which will be addressed shortly in Section 3.4.

Fig. 3.4 illustrates the synchronization process. It is actually an extension of Fig. 3.3. The computer coordinates the table rotation via motion controller and the image acquisition via image grabber, both of which sit in the PCI slots of the computer. When the table moves to the same location (according to the encoder counts), a trigger signal will be generated to enable the external trigger line of the image grabber which starts immediately acquiring the image of the surface of the specimen that is mounted on the table. After one cycle of rotation, one surface image is acquired into the grabber's buffer.

After synchronization and image registration, images from the same surface area are obtained. Comparison of gradual changes of these images may lead to understanding of the behavior of a specific wear process.

3.3.2 Software

This wear testing and monitoring apparatus includes several software packages for motion control, image acquisition (IMAQ) and image processing and analysis. A block diagram of the overview of the software for the whole system (Fig. 3.5) illustrates the software constitution.

Motion Control Software

Commands[3] for specifying motion and machine parameters can be programmed and sent to the controller via the application programming interface (API) soft-

[3]The motion controller card accepts ASCII or binary commands.

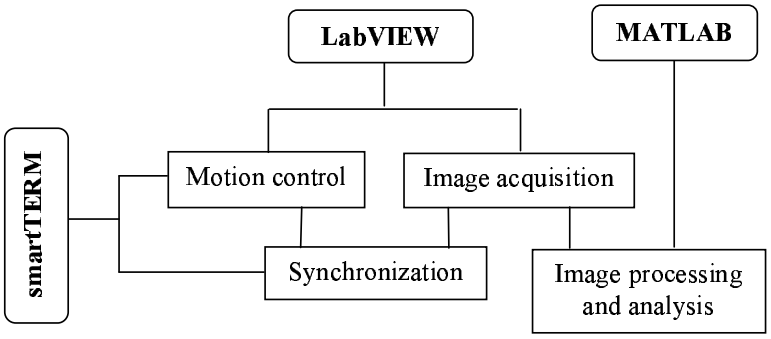


Figure 3.5: Overview of the software.

ware "smartTERM" from Galil[4]. In practice, motion control is implemented in LabVIEW[5] virtual instruments (VIs).

Software for Image Acquisition and Processing

The frame grabber PCI-1428 works with NI-IMAQ, the National Instruments complete image acquisition driver software, which performs all functions required for acquiring and saving images, shown in Fig. 3.6. But NI-IMAQ does not perform any image processing and analysis.

National Instruments IMAQ Vision is used for image processing and analysis, which, like NI-IMAQ, works with labVIEW as well. It adds powerful image processing and analysis to LabVIEW programming environment. Sometimes MATLAB is also used to perform some image processing and analysis considering its powerful functions for matrix operation and easy prototype programming.

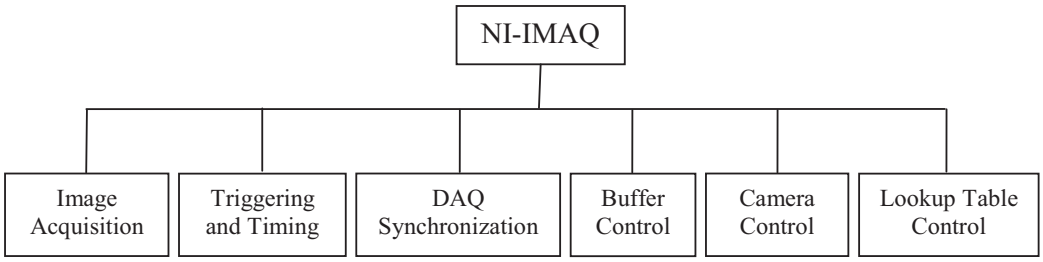


Figure 3.6: NI-IMAQ functions.

[4]Galil Motion Control, Inc. (http://www.galilmc.com/).

[5]LabVIEW is a graphical programming language from National Instruments, used primarily in data acquisition and instrumentation control.

3.4 Image Registration

Image registration is inherently basic, often used as a preliminary step in image processing applications. In the current project, to detect and compare changes of a series of images from the same wearing surface area are desired. So it is necessary for meaningful results to have the images registered. In this section a brief introduction to the method is given for image registration applied in the current project.

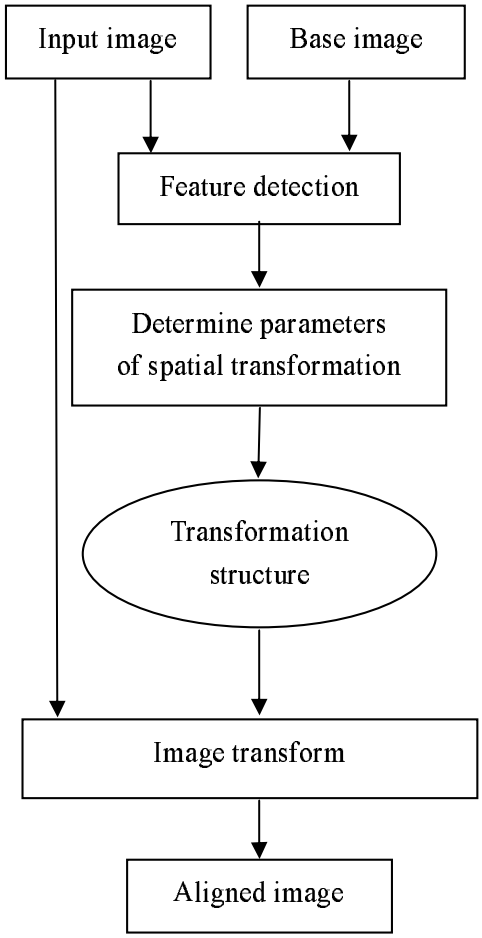


Figure 3.7: Overview of image registration process.

Image registration is the process of aligning two or more images of the same scene (taken at different times in the current case). Typically, one image, called the base image, is considered the reference to which the other images, called input

images, are compared. The objective of image registration is to bring the input image into alignment with the base image by applying a spatial transformation to the input image.

Due to the diversity of images to be registered as well as various types of degradations it is impossible to design a universal method applicable to all registration tasks. Every method should take into account not only the assumed type of geometric deformation between the images but also radiometric deformations, required registration accuracy and application-dependent data characteristics. Nevertheless, the majority of the registration methods consist of the following steps (see Fig. 3.7): (1) Feature detection; (2) Feature matching; (3) Determine parameters of spatial transformation; (4) Apply image transformation. A recent review on image registration can be found in [104].

In the current project, the pixel shift between the images, taken at different time with certain magnification, is mainly caused by mutually translation and slight rotation. This type of pixel shift between images can be easily solved by the simple normalized *cross correlation* (CC) method [2, 104].

Let S the *search area* of size $L \times L$ from the base image and W the *window* of size $M \times M$ from the input image be defined as shown in Fig. 3.8. Assume that S and W have K grey levels; i.e.,

$$\begin{aligned} &0 \leq S(i,j) \leq K-1 \\ &0 \leq W(m,n) \leq K-1 \\ &1 \leq i,j \leq L \\ &1 \leq m,n \leq M. \end{aligned}$$

Let $S_M^{i,j}(m,n)$ denote $M \times M$ the wholly contained subimage of S, i.e.

$$\begin{aligned} S_M^{i,j}(m,n) &\equiv S(i+m-1, j+n-1), \\ &\begin{cases} 1 \leq i,j \leq L \\ 1 \leq m,n \leq M. \end{cases} \end{aligned} \tag{3.1}$$

Each $M \times M$ subimage of S can be uniquely referenced by the specification of its upper left corner's coordinates (i,j), which will be used to define *reference points*. Assume that enough *a priori* information is known about the dislocation between the window and search area so that the parameters L and M may be selected with the virtual guarantee that, at registration, a complete subimage is contained in the search area as shown in Fig. 3.8. This assumption in practice is satisfied since the maximum pixel shift can be estimated as stated above. Image registration, therefore, is a search over the subset of the allowed range of reference points to find a point (i^*, j^*) which indicates a subimage that is most similar to the given window.

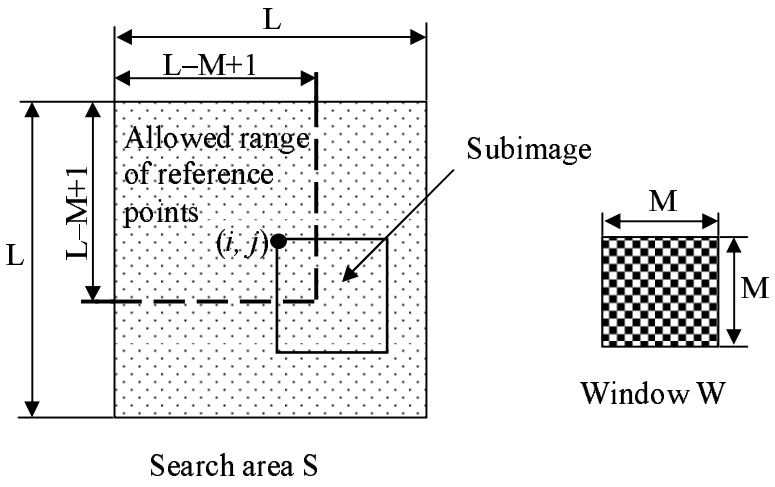


Figure 3.8: Search space.

The equation of the normalized CC is defined as:

$$R(i,j) = \frac{\sum_{m=1}^{M}\sum_{n=1}^{M}(W(m,n)-\mu_W)(S_M^{i,j}(m,n)-\mu_M(i,j))}{\sqrt{\sum_{m=1}^{M}\sum_{n=1}^{M}(W(m,n)-\mu_W)^2}\sqrt{\sum_{m=1}^{M}\sum_{n=1}^{M}(S_M^{i,j}(m,n)-\mu_M(i,j))^2}} \tag{3.2}$$

where μ_W and $\mu_M(i,j)$ denote the mean of the window W and the mean of subimage $S_M^{i,j}$ respectively.

In the correlation scheme the correlation surface $R(i,j)$ is searched for a maximum $(\tilde{i},\tilde{j})$. The procedure is successful if $(\tilde{i},\tilde{j})$ and (i^*,j^*) are equivalent.

Two main drawbacks of the normalized CC method are the flatness of the similarity measure maxima (due to the self-similarity of the images) and high computational complexity. Nevertheless the maximum can be sharpened by preprocessing or adding marks in the images. To reduce the computational load, a small window of size 81×81 which only covers the marks are used and the normalized correlation surface is calculated by fast Fourier transform (FFT) in the experiment.

After the above mentioned preliminary registration operations of the sequence of images, wear behavior, dry sliding wear in particular, are investigated, which will be presented in Chapter 6.

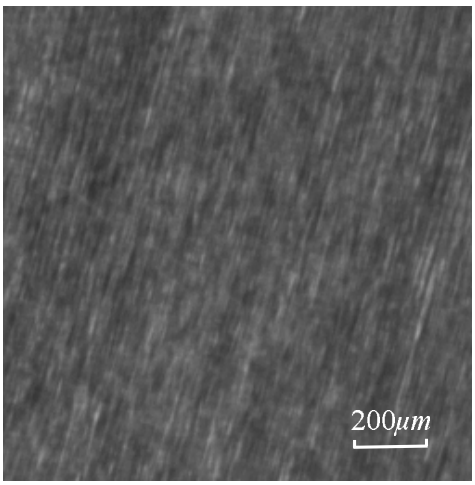
3.5 Observations

The online wear monitoring method is based on the analysis of sequences of photomicrographs, taken from the object under test. To study the dynamics of the wear process, a temporal series of images can be taken. Fig. 3.9 shows an image sequence, acquired by the imaging system of the wear apparatus, of the surface of a machined aluminium disk during a Pin-on-Disk dry sliding wear process. The pin has a spherical shape of diameter 3 mm and the applied normal load is 5 N.

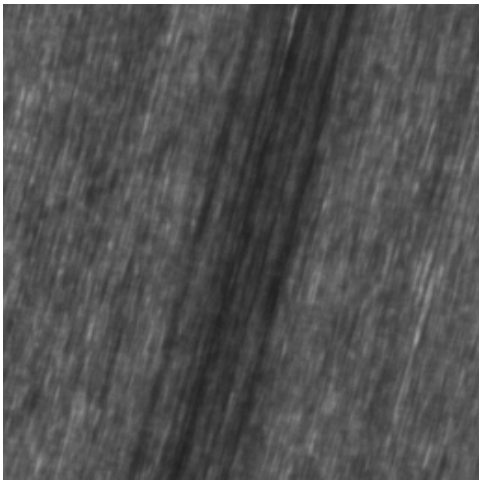
From this figure one can see that the gradual changes in the image set apparently reflect the wear process occurred on the aluminium disk surface. The wear track becomes wider with the increasing number of rotation cycles. A big wear particle appeared in Fig. 3.9 (d) and it was removed by the pin in subsequent rotation cycles. Chapter 5 and 6 will present how to identify the wear region from images, how to estimate the wear degree of the surface and wear behavior with respect to load, velocity etc.

3.6 Summary

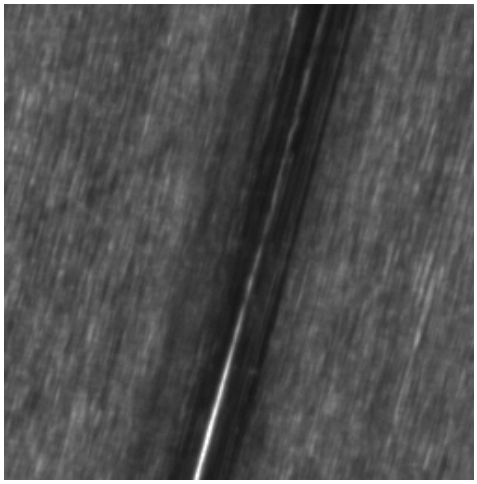
An online wear monitoring system using imaging methods is presented in this chapter. The basics and configurations of this system and its potential advantages are stated. Based on the project targets, the design of the wear testing and monitoring apparatus, including hardware and software, is described in detail. This chapter provides the basis for the following chapters.



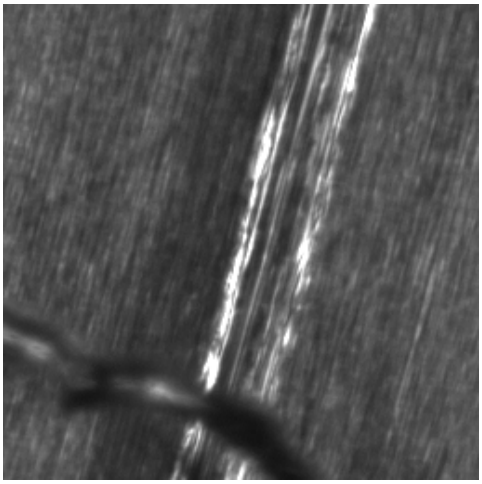
(a) Initial image without wear



(b) Image after 10 cycles of rotation

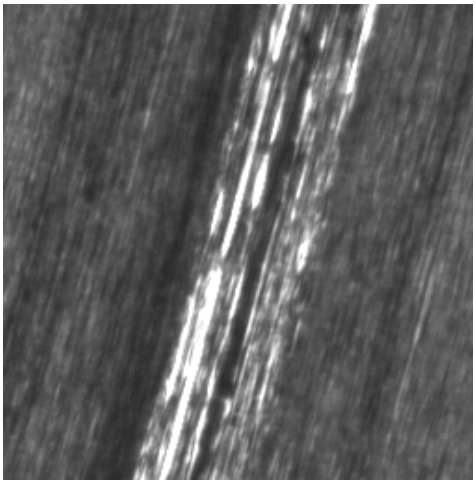


(c) Image after 20 cycles of rotation

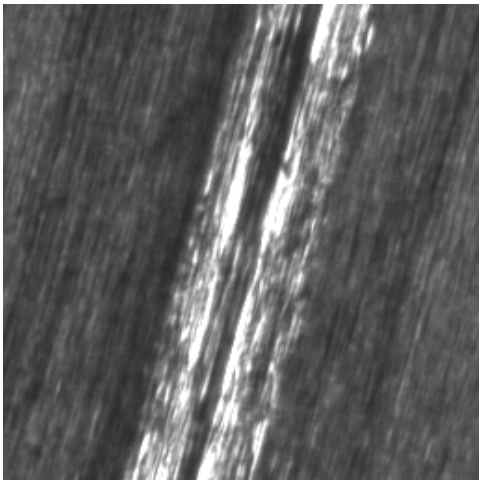


(d) Image after 30 cycles of rotation

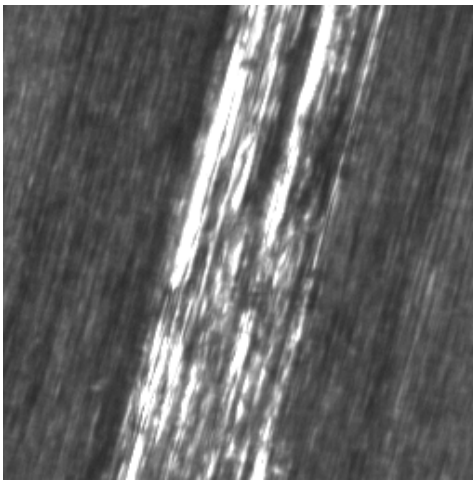
Figure 3.9: Series of images of an aluminium disk in a dry sliding wear process. The scale is illustrated in image (a), same scale for images (b) to (h).



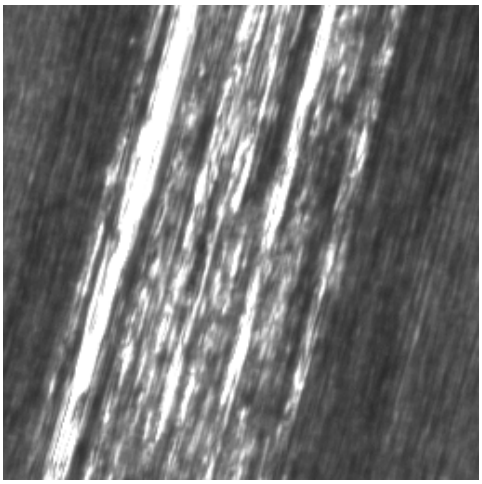
(e) Image after 40 cycles of rotation



(f) Image after 50 cycles of rotation



(g) Image after 60 cycles of rotation



(h) Image after 70 cycles of rotation

Figure 3.9: Series of images of an aluminium disk in a dry sliding wear process (continued).

Chapter 4

Ring Light Illumination Modeling

Illumination plays a crucial role in an imaging system. A good illumination scheme should effectively reveal features of interest of the specimen under observation and guarantee acquisition of adequate images. This chapter first discusses the choice of illumination strategy, according to which a fiber optic ring light illuminator is adopted to the current imaging system for wear monitoring. Emphasis is then put on modeling of ring light illumination. An illumination model and a reflectance model based on the Torrance-Sparrow model are presented and evaluated for ring light illumination.

4.1 Introduction

The main objective of the current project is to do online wear detection and monitoring via an imaging system, the first element of which is image acquisition. An aspect that is often overlooked in computer vision is the acquisition of *adequate* images: a proper illumination should be effective in revealing the features of interest of the specimen and guarantee sufficient contrast, whilst focussing must be correct to obtain sharp images. Any deficiencies of the initial images can cause big problems with image analysis and interpretation. Aside from camera choice, the type of illumination is perhaps the most critical element, which applies to all forms of imaging systems, for achieving adequate images.

Generally specimens are often observed by imaging systems under either reflected or transmitted illumination schemes, employing a variety of light sources and configurations, which are strategically positioned in appropriate locations. In many circumstances, reflected and transmitted light sources are combined to take advantage of particular specimen characteristics in a manner that most effectively reveals the features of interest. As stated in the previous chapter, a video zoom microscope is used to monitor micro-wear occurring on a relatively smooth surface (only proper illumination for microscopes is considered within the scope of this thesis). The features of interest in the current imaging system for wear monitoring are surface micro-geometry, morphology, and topography of the specimen, which influence and carry the information about wear on the surface [103, 102]. Accordingly, a good illumination scheme for the imaging system presented in Chapter 3 should effectively reveal these features and in the meantime guarantee sufficient contrast.

There exists no single optimum illumination strategy that is the correct choice for the wide variety of specimens that the imaging system is designed to accommodate [33]. Each specimen under observation can be illuminated by a variety of different mechanisms, and employing a nearly infinite number of variations or combinations of techniques. For a given specimen or object, although there may be several possible illumination schemes that produce acceptable results, a single approach may be discovered that, after careful refinement, produces exceptional results.

This chapter is organized as follows: Section 4.2 describes the choice of an illumination strategy, according to which a fiber optic ring light is adopted and introduced in Section 4.3; Section 4.4 presents an illumination model as well as a reflectance model for the ring light; and both are evaluated and validated; discussions are given in Section 4.5; followed by conclusions in Section 4.6.

4.2 Choice of Illumination Strategy

Specimen characteristics should be carefully considered in selecting an illumination strategy to suit the needs of visual observation, photomicrography, or digital imaging. The opacity of the specimen is, in general, the most important characteristic, and will determine the basic type of illuminator that should be employed, reflected, transmitted, or in some cases, a combination of both. Opaque specimens are typically illuminated from above (with reflected light), using orientations ranging from on-axis (parallel to the microscope optics) to highly oblique (up to 90° incident angle from the optical axis), as required to reveal the features or characteristics of interest.

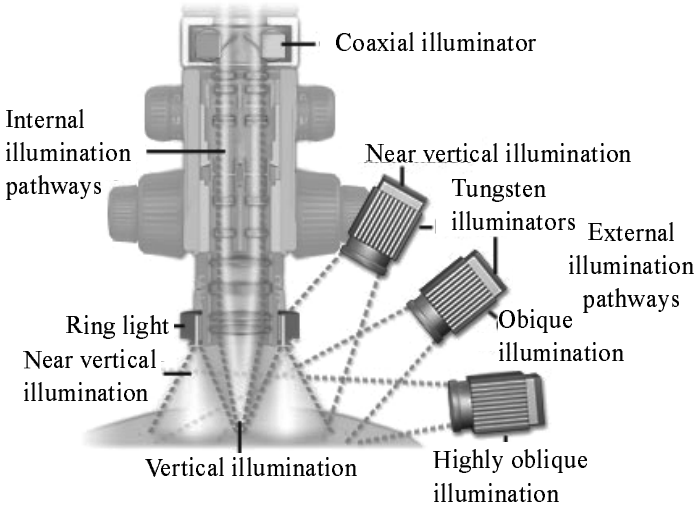


Figure 4.1: Reflected light illumination (adapted from [33]).

Once it has been determined that the specimen's opacity suggests the use of a particular general category of illuminator, then a number of other factors should be considered to further refine particular variations on the basic illumination scenario that will likely produce the desired results. Fig. 4.1 illustrates a variety of potential pathways for specimen illumination using reflected light. A simple tungsten (or tungsten-halogen) illuminator, shown oriented at different angles to the specimen surface, and a ring light mounted on the objective lens body, provide lighting that is independent of the microscope optical path. The illumination path for a coaxial illuminator, which functions within the microscope optical train, is illustrated in the cutaway section of the instrument. Opaque specimens most

commonly benefit from reflected illumination, while translucent and transparent objects usually produce the best results with some variation of transmitted illumination, such as brightfield, polarized, oblique, or darkfield [33]. Since mostly wear phenomena of opaque materials is studied in the current project, only reflected illumination is considered.

The geometrical profile, topography, and morphology of the specimen are all important factors in choosing and configuring illumination in a manner that will reveal the desired information. Specimens that are highly three-dimensional (have high relief) should be illuminated differently than those that are flat, smooth, or even highly polished. For example, highly oblique lighting can produce shadows on rough surface areas, obscuring surface detail that may be important (say a wear track). In the current project, the surfaces to examine are relatively smooth, diffuse light originating from directly above a specimen may uniformly illuminate both wear and wear-free regions with a certain degree of contrast. Nonuniform illumination in this case may make image segmentation difficult.

Laser light may be used in an optical measurement system; however, the coherence of laser light produces surface noise (intensity variations due to surface machining marks) that is approximately twice as high as that for conventional white light sources [19]. For this reason conventional white light is preferred.

Aside from opacity and geometrical profile, a number of other factors should be considered in planning a lighting strategy. These include the basic physical characteristics of the specimen, the type of information that is required from the examination, digital imaging requirements, and how the information will be utilized.

One of the crucial specimen characteristics that influence the choice of an appropriate lighting scheme, is the composition which directly affects both the surface and internal reflectivity. Metals, plastics, ceramics and glasses behave differently with regard to their appearance under different lighting conditions. In this project the research is restricted only on metallic surfaces. Such surfaces may produce artifacts by reflecting images of the light source(s) into the microscope objective. These reflections usually produce glare and obscure important detail, or distract attention from the important elements that are being observed and imaged. Diffuse light that limits such glare effect in this case is more preferred.

Digital imaging requirements are another important factor that must be taken into consideration in choosing specimen illumination. The intensity of the lighting must also be adequate to ensure exposures that are of reasonable duration for the camera employed.

4.3 Fiber Optic Ring Light

Of all the illumination sources available for a microscope, fiber optic illuminators are probably the most versatile and popular. Many different light source designs, fiber types and configurations, and accessory attachments are available (see Fig. 4.2). A fiber optic light system can be configured to meet the stringent requirements of almost any application. Generally powered by high-intensity tungsten-halogen lamps, fiber optic illuminators are relatively bright sources, suitable for investigations of many specimens than are basic incandescent illuminators.

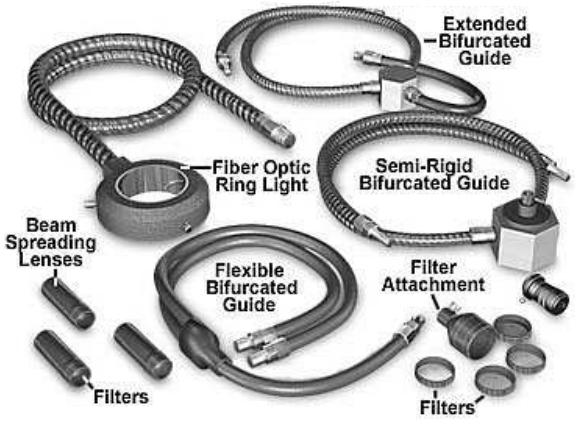


Figure 4.2: Fiber optic light guides and accessories (adapted from [33]).

The fiber optic ring light is one of the most widely used configurations among the fiber optic-based illuminators. A fixed means of attachment, surrounding the microscope objective, eliminates any varieties in adjustment, and ensures that the illumination is of consistent quality, and highly reproducible, from specimen to specimen. Because the illumination path is nearly coincident with the optical axis of the microscope, the viewing area is evenly illuminated, which will be addressed in the following section. In addition, the diffuse illumination provided by ring lights eliminates the shadows, while still providing adequate contrast for visual inspection. Because of the advantages of the fiber optic ring light, it is employed as the illuminator for the current imaging system according to the choice of illumination strategy presented in Section 4.2. The ring light will be modeled in the next section.

4.4 Modeling of Ring Light

Detection and monitoring of wear in the current project involve the analysis of intensity images resulting from the reflection of light. The apparent brightness of a point depends on illumination and its ability to reflect incident light in the direction of the sensor, i.e. image intensities are closely related to illumination and surface reflectance properties. Detection and monitoring of wear can greatly benefit from proper illumination and reflectance models.

Various reflectance models have been proposed and used in machine vision and computer graphics. Lambert [48] was the first to investigate the mechanisms underlying diffuse reflection. Phone [64] proposed a parameterized continuous function to represent specular reflectance, and used the model to produce computer-synthesized images of objects. Horn [28] developed shape-from-shading algorithms for machine vision using the Lambertian diffuse reflectance model and the double-delta specular reflectance model. Woodham [96] used the Lambertian model to estimate object shape by means of photometric stereo. Sanderson, Weiss and Nayar [81] used the double-delta specular model to determine the shape of specular surfaces by means of the structured highlight technique. He and Torrance et al. [25] developed a comprehensive physical model that described specular, directional diffuse, and uniform diffuse reflection by a surface. Recently Oren and Nayar [62] generalized the Lambertian model and demonstrated its usefulness in machine vision.

The above models and their applications have proven that the Lambertian model does reasonably well in describing diffuse reflections. On the other hand, specular models perform well only when the object surface is smooth. In general, approaches to the study of reflection can be classified into two categories: physical and geometrical optics. While geometrical models may be considered as mere approximations to physical models, they possess simpler mathematical forms that often render them more usable than physical models. Two specific models that have been reported to fit experimental data very well are the Beckmann-Spizzichino (physical optics) model [4] and the Torrance-Sparrow (geometrical optics) model [85], which have received considerable attentions from the vision research community. Both of them assume that the light source is at a great distance from the surface, so that all light rays that are incident upon the surface area are nearly parallel to one another. For ring light, however, these two models are not directly applicable and need to be further formulated.

4.4.1 Radiometric Definitions

To analyze illumination and surface reflection, it is useful to introduce some radiometric terms first. As shown in Fig. 4.3, all directions are represented by the polar angle θ and the azimuth angle ϕ. The light source is determined by its polar angle θ_i and azimuth angle ϕ_i. The monochromatic flux $d\Phi_i$ is incident on the surface area dA_s from the direction θ_i, and a fraction of it, $d\Phi_r$, is reflected in the direction (θ_r, ϕ_r). The *irradiance* I_s of the surface is defined as the incident flux density:

$$I_s = \frac{d\Phi_i}{dA_s} \tag{4.1}$$

The *radiance* L_r of the surface is defined as the flux emitted per unit fore-shortened area per unit solid angle. The surface radiance in the direction (θ_r, ϕ_r) is defined as:

$$L_r = \frac{d\Phi_r}{dA_s \cos\theta_r d\omega_r} \tag{4.2}$$

The *bi-directional reflectance distribution function* ($BRDF$) of a surface f_r is a measure of how bright the surface appears when viewed from a given direction, when it is illuminated from another given direction. The BRDF is defined as:

$$f_r = \frac{L_r}{I_s}. \tag{4.3}$$

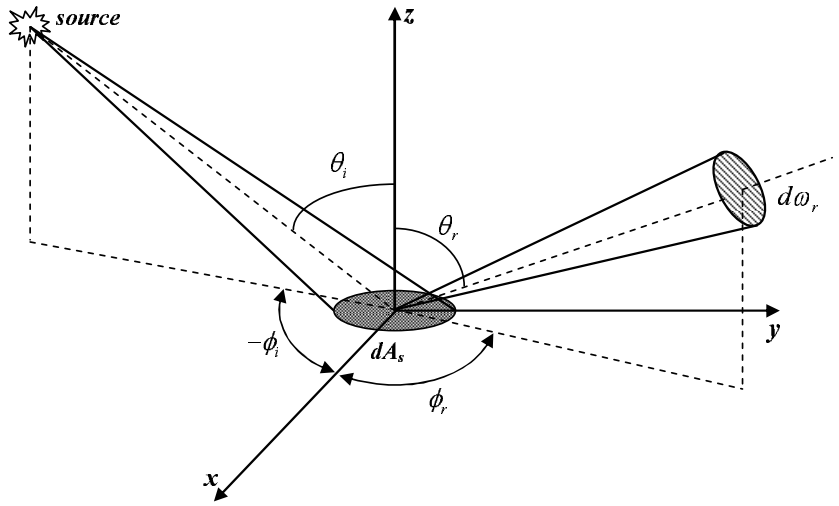


Figure 4.3: Basic geometry needed to define radiometric terms.

4.4.2 Surface Models

The manner in which light is reflected by a surface is dependent on, among other factors, the microscopic shape characteristics of the surface. A smooth surface, for instance, may reflect incident light in a single direction, while a rough surface will tend to scatter light in various directions, maybe more in some directions than others. To be able to accurately predict the reflection of incident light, one must have prior knowledge of the microscopic surface irregularities, in other words, a model of the surface is needed. An exact profile of a surface can be obtained by measuring the height at each point on the surface by means of profilometer sensors (refer to Subsection 2.3.5 of Chapter 2 for details). This method, however, is quite cumbersome (due to complicated operation, slow mechanical scanning or focusing, etc.) and also inapplicable in many practical situations, for instance the current online wear monitoring project in which the surface is in motion. Hence it is convenient to model a surface as a random process, where it is described by a statistical distribution of either surface height above a mean level, or its slope with respect to mean (macroscopic) surface slope.

Height Distribution Model

The surface height $\hbar$ may be expressed as a random function of the surface coordinates x and y. The shape of the surface is then determined by the probability density function of $\hbar$. Let $\hbar$ be normally distributed, with zero mean value and variance $\sigma_\hbar^2$. Then the distribution of $\hbar$ is given by:

$$p_\hbar(\hbar) = \frac{1}{\sqrt{2\pi}\sigma_\hbar} e^{-\hbar^2/2\sigma_\hbar^2} \tag{4.4}$$

The standard deviation $\sigma_\hbar$ is also the root mean square (RMS) of $\hbar$ and represents the *roughness* of the surface. $\sigma_\hbar$ will also be referred as to *RMS roughness*[1] hereafter. The surface is not uniquely described by the distribution of $\hbar$, however, as it does not give spatial information about how the height varies in the lateral direction. In order to strengthen the surface model, an *autocorrelation function* (ACF) that determines the correlation (or lack of independence) between the random values assumed by the height $\hbar$ at two points (x_1, y_1) and (x_2, y_2), separated by a distance τ. For the sake of simplicity, a one-dimensional (1D) ACF is presented here. The ACF $A(\tau)$ is defined as the expected product of two heights $\hbar(s)$

[1]Historically, RMS roughness is defined also as: $R_q = \sqrt{\frac{1}{L}\sum_{s=0}^{L}[z(s) - \overline{z(s)}]^2}$, where L is the evaluation length, $z(s)$ is the surface height at a point s in the surface profile and, $\overline{z(s)}$ is the average height of the surface profile.

and $\hbar(s+\tau)$ at points s and $(s+\tau)$:

$$A(\tau) = E[\hbar(s)\hbar(s+\tau)] = \frac{1}{L}\sum_{s=0}^{L}\hbar(s)\hbar(s+\tau) \tag{4.5}$$

where E denotes expectation and L the sample length. When $\tau = 0$, $A(\tau)$ is equal to the variance $\sigma_\hbar^2$ of the surface heights.

The ACF can be described by the fairly general function:

$$A(\tau) = e^{-\tau^2/\mathcal{T}^2} \tag{4.6}$$

where $\mathcal{T}$ is the *correlation distance*, for which $A(\tau)$ drops to the value e^{-1}. By varying the parameters $\sigma_\hbar$ and $\mathcal{T}$ of the surface model, one can generate surfaces that match in appearance the surfaces met in practice. Moreover, one also can use another height distribution function and/or another autocorrelation coefficient than the ones given above if the performance of the model is dissatisfied.

Slope Distribution Model

It is sometimes convenient to think of a surface as a collection of planar micro-facets as the scattering of light rays has been found to be dependent on the local slope of the surface and not the local height of the surface. A large set of micro-facets constitutes an infinitesimal surface patch that has a mean surface orientation **n**. Each micro-facet, however, has its own orientation which may deviate from the mean surface orientation by an angle α (see Fig. 4.4). We will also use α to represent the slope of each micro-facet. Surfaces can be modeled by a statistical distribution of the micro-facet slopes. If the surface is isotropic, the probability distribution of the micro-facet slopes may be assumed to be rotationally symmetric. Let the micro-facet slopes to be normally distributed, with zero mean value and standard deviation σ_α:

$$p_\alpha(\alpha) = \frac{1}{\sqrt{2\pi}\sigma_\alpha}e^{-\alpha^2/2\sigma_\alpha^2} \tag{4.7}$$

Surface roughness in this case is determined by a single parameter, namely σ_α. Larger values of σ_α may be used to model rougher surfaces. We will also refer σ_α as to *RMS slope*[2] hereafter. The advantages of using a single parameter come with the cost of a weaker model when compared to the height model. Given a probability distribution function of α, it is difficult to visualize the shape of the

[2]RMS slope is historically defined as $m_{rms} = \sqrt{\frac{1}{L}\sum_{s=0}^{L}[z'(s) - \overline{z'(s)}]^2}$, where $z'(s)$ is the derivative of the function $z(s)$ at point s.

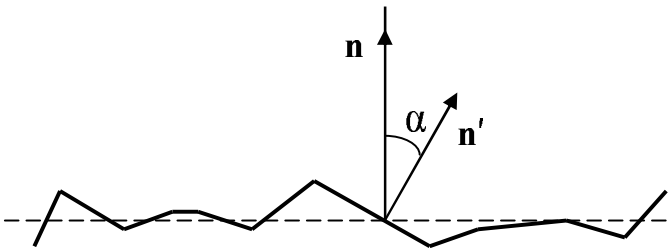


Figure 4.4: Surface model as a collection of planar microfacets.

surface and to estimate the RMS roughness. The slope model, though relatively ambiguous, is more directly applicable to the problem of surface reflection.

The slope model may also be derived from the height model. For example, if the surface height is assumed to be normally distributed with RMS roughness $\sigma_\hbar$ and correlation distance $\mathcal{T}$, the slope distribution can be determined from the height distribution as [4]:

$$p_\alpha(\alpha) = \frac{\mathcal{T}}{2\sigma_\hbar\sqrt{\pi}\cos^2\alpha}\exp\left[-\frac{\tan^2\alpha}{(2\sigma_\hbar/\mathcal{T})^2}\right] \tag{4.8}$$

And when rough surfaces are gently varying, the slopes (α) of most facets are small. The RMS slope σ_α maybe related to RMS roughness $\sigma_\hbar$ as [61]:

$$\sigma_\alpha = \frac{1}{\sqrt{2}}\tan^{-1}\frac{2\sigma_\hbar}{\mathcal{T}} \tag{4.9}$$

Note that in practice the values of both the RMS roughness and RMS slope are measuring instrument dependent. The length of the sample, the area resolution and the sampling rate all affect the roughness parameters − this being especially true for the RMS slope [93].

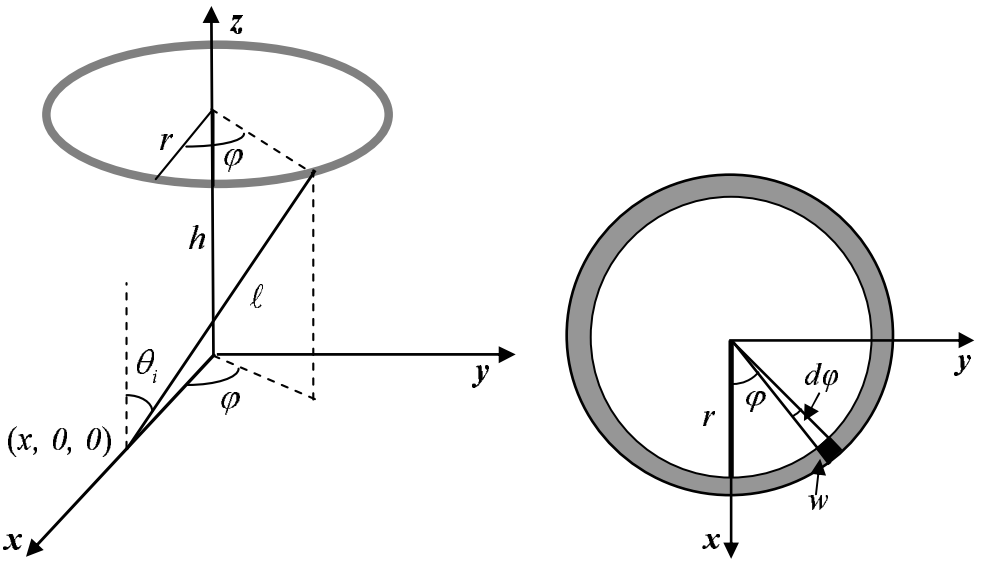
Both height and slope models have been used to develop surface reflectance models by the vision research community, for instance, Beckmann-Spizzichino [4] used a surface height distribution model to derive the reflectance model while Torrance-Sparrow [85] used a surface slope distribution model. In this thesis the slope distribution model is used to model the surface reflection under ring light illumination for its simplicity.

4.4.3 Ring Light Illumination Model

Before deriving the reflectance model, ring light illumination is modeled. The radiometric characteristics of the ring light are formulated first. As Fig. 4.5 shows,

suppose the *nominal* plane of the surface is located at $z = 0$ and the ring light is parallel to the nominal plane at $z = h$, its center lying at (0, 0, h) as illustrated in Fig. 4.5 (a). Furthermore, suppose the ring light is a circular strip with radius r and width w (which is assumed small) (see Fig. 4.5 (b)). Then the ring area element dA_i can be expressed as

$$dA_i = wrd\varphi \tag{4.10}$$



(a) Coordinate system used to derive the illumination model

(b) Top view of the ring light

Figure 4.5: Schematic view of the ring light.

Consider an infinitesimal surface area dA_s illuminated by a ring element dA_i at an angle of incidence θ_i, as shown in Fig. 4.6. Let $d\omega_i$ and $d\omega_s$ represent the solid angles subtended by the ring element dA_i and surface patch dA_s respectively. Then the flux incident on dA_s may be determined from the source radiance L_i as:

$$d\Phi_i = L_i d\omega_s dA_i \tag{4.11}$$

From the solid angles subtended by the surface and source areas, we obtain:

$$dA_i = d\omega_i \ell^2 \tag{4.12}$$

$$d\omega_s = \frac{dA_s \cos\theta_i}{\ell^2} \tag{4.13}$$

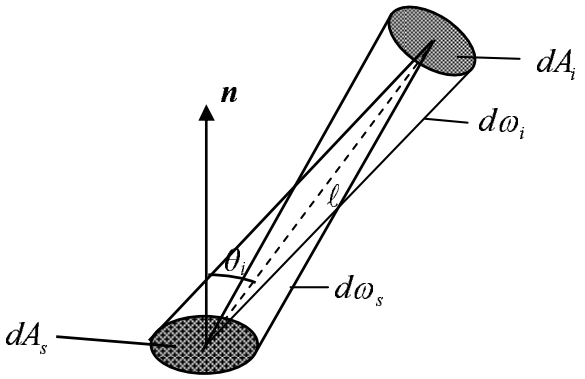


Figure 4.6: Dependence of the incident light energy on the source direction.

Substituting Eqs. (4.12) and (4.13) into Eq. (4.11), we obtain:

$$d\Phi_i = L_i d\omega_i dA_s \cos\theta_i \tag{4.14}$$

Thus the surface irradiance is determined from the above equation as:

$$dI_s = \frac{d\Phi_i}{dA_s} = L_i d\omega_i \cos\theta_i = \frac{L_i dA_i}{\ell^2}\cos\theta_i = \frac{L_i wr\cos\theta_i}{\ell^2}d\varphi \tag{4.15}$$

As the illumination scheme is cylindrically symmetrical, the surface irradiance needs to be calculated only along the x-axis. Thus the obtained irradiance along the x-axis will correspond to the irradiance along any line through the origin in the $z = 0$ plane. Consider the point[3] at the x-axis be $(x,\ 0,\ 0)$ as shown in Fig. 4.5. Because of cylindrical symmetry the irradiance from only a half-ring can be calculated, and the complete irradiance is then obtained by doubling the result. From Fig. 4.5 we obtain:

$$\ell = \sqrt{x^2 + h^2 + r^2 - 2xr\cos\varphi} \tag{4.16}$$

$$\cos\theta_i = \frac{h}{\sqrt{x^2 + h^2 + r^2 - 2xr\cos\varphi}} \tag{4.17}$$

By combining Eqs. (4.15), (4.16), (4.17) and integrating over an angle φ from 0 to π, the irradiance at the point $(x,\ 0,\ 0)$ is obtained:

$$I_s(x) = 2\int_0^{\pi} \frac{hwrL_i}{\sqrt{(x^2 + h^2 + r^2 - 2xr\cos\varphi)^3}}d\varphi, \tag{4.18}$$

which is equal to the irradiance at every point along the circle centered at the

[3]Here *point* means an infinitesimal surface patch, same hereafter in this chapter.

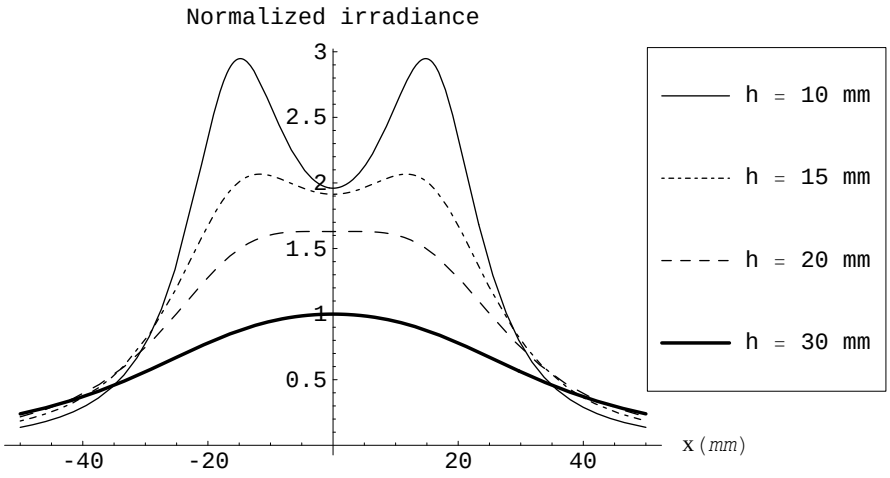


Figure 4.7: Irradiance along x-axis for different ring heights, computed according to Eq. (4.19).

origin with radius $|x|$. The solution to Eq. (4.18) is

$$I_s(x) = \frac{4hwrL_i}{(h^2+(x+r)^2)\sqrt{h^2+(x-r)^2}}\mathcal{E}(-\frac{4xr}{h^2+(x-r)^2}), \qquad (4.19)$$

where $\mathcal{E}(m)$ is the complete elliptic integral of the second kind.

According to Eq. (4.19), Fig. 4.7 shows the calculated irradiance distribution along the x-axis for different ring height h, where the ring radius $r = 16.65\ mm$ is fixed. The irradiance in this figure is normalized by the maximum for $h = 30\ mm$ (same hereafter unless otherwise explicitly stated). It can be seen that decreasing the ring height will introduce peaks in the irradiance distribution at approximately $|x| = r$ while increasing the ring height decreases the irradiance. When the ring is located at a proper height range, uniform illumination on the nominal plane of the surface is achieved when $|x| \leq r$ (for instance, when $h = 15\ mm$). In the current imaging system, the field of view (FOV) of the microscope (about $3.5\ mm \times 2.8\ mm$ at magnification 2.5X) is far smaller than the area that the ring covers, hence uniform illumination on the surface nominal plane can be easily achieved.

In order to achieve this relatively simple solution, in this model the output beam angular restriction of the ring light is not taken into account. In practice, however, a ring light guide does have a limit to the output beam angle to prevent the light rays from reaching the direction that is not wanted. This will be discussed again in the validation of the model in the following subsection.

Another note is that, in the above derivation, uniform illumination is achieved

on the surface nominal plane which is in parallel to the ring and perpendicular to the optical axis of the imaging system. The normal of the nominal plane is the mean orientation of the whole surface of interest, not necessarily the same as the mean orientation **n** of a particular infinitesimal surface patch discussed in Subsection 4.4.2. But one may consider the orientation of the nominal plane as the mean of all the orientations of the infinitesimal surface patches of interest. Bearing this in mind is very helpful to well understand the reflectance model discussed in Subsection 4.4.5.

In addition, since all the points on a real surface with certain roughness do not necessarily locate on the nominal plane, under ring light illumination the irradiance on a real surface still has certain contrast even if the ring light locates within the height range where uniform illumination is obtained on the nominal plane.

4.4.4 Validation of the Illumination Model

Some irradiance measurement from a real ring light was conducted to validate the illumination model. The ring light illumination in the experiment is provided by a StockerYale ImageLite^{TM} Model 20 fiber optic illuminator [31] which is coupled to a ring light guide [32] (see also Fig. 4.8).

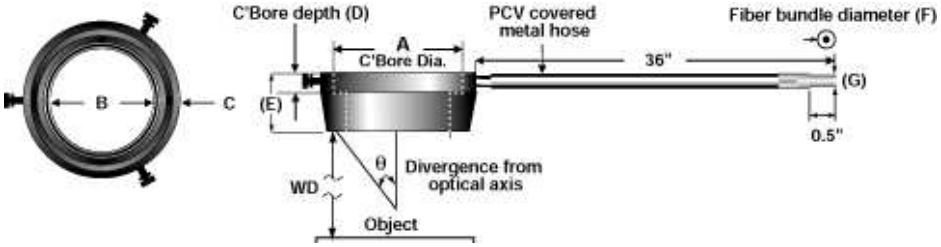


Figure 4.8: Technical image of a ring light guide (adapted from [30]).

The radius of the ring light is 16.65 mm. Due to cylindrical symmetry of the ring light illumination, we measured the irradiance distribution along the x-axis (as illustrated in Fig. 4.5) at four height positions, namely $h = 10$ mm, $h = 15$ mm, $h = 20$ mm and $h = 30$ mm. The irradiance was measured by a universal photometer (Hagner, model S2, the detector size of which is 10 mm in diameter). The measured results are shown in Fig. 4.9. For the same ring height $h = 30$ mm, the predicted and measured irradiance distributions along the x-axis are shown together in Fig. 4.10.

From Fig.4.9 we can see that the irradiance curves measured from the real ring light have a similar shape as those predicted by the illumination model. And when the ring height drops, the measured irradiance curves become wider like

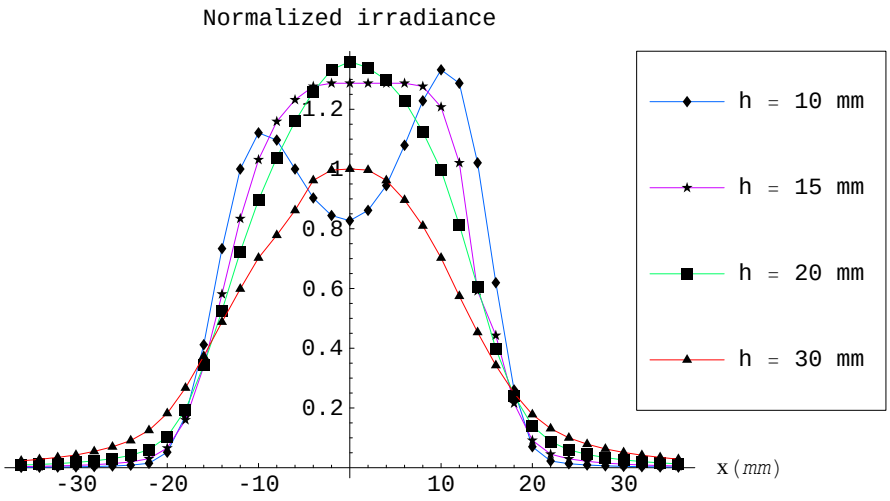


Figure 4.9: Irradiance distribution along x-axis measured at different ring height.

those predicted. In contrast to the predicted irradiance becoming larger as the ring height drops (see Fig. 4.7), the measured one first increases and then starts to decrease from a certain height on. This may be due to the fact that the real ring light has an output beam angle restriction (the output beam angle in this case is 20°, see Fig. 4.8) while the model does not have such angular restriction. For the same reason, the measured irradiance drops much quickly than that predicted by the model (see Fig. 4.10). For $h = 10\ mm$, the two peaks of the measured irradiance curve are not equal or symmetric, which is not expected as that in Fig. 4.7. This is mainly caused by the physical mounting error of the ring light in practice where the ring may tilt a little. The little height difference at both sides of the ring results in big irradiance difference especially when the ring height is small. On the other hand, during the measurement, the orientation of the sensor surface is not always the same as that of the nominal plane. Nevertheless, in practice, the ring light does not work at that low height position.

To exactly work out the illumination model for this particular ring light is beyond the scope of this thesis. Nevertheless, the irradiance distribution from the measurement more or less validates the illumination model for an ordinary ring light. On the other hand, as stated in the previous subsection, the FOV of the imaging system in practice only covers a small area in the center illuminated by the ring light, where uniform illumination is easily achieved.

As the objective of this research is to detect wear from image sequences, we

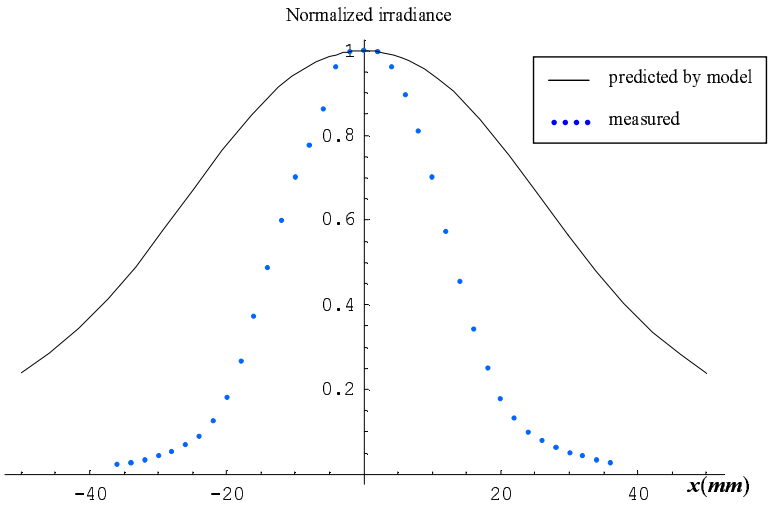


Figure 4.10: Comparison of modeled and measured irradiance distributions.

are interested in a reflectance model by which some insight into the characteristics of the images may be obtained. In the next subsection, we are going to derive a surface reflectance model under ring light illumination.

4.4.5 Reflectance Model Under Ring Light Illumination

Generally, in terms of reflectance characteristics surfaces can be classified as two categories: *Lambertian* and *non-Lambertian*. For a Lambertian surface, its radiance is independent of the viewing direction of the sensor; it appears equally bright from all directions. Its $BRDF$ is $f_r = \varepsilon/\pi$ where the *albedo* ε represents the fraction of incident energy that is reflected by the surface. Since most surfaces studied in the current project are non-Lambertian, in the following only non-Lambertian surfaces are considered.

For non-Lambertian surfaces, the radiance from any point on the surface is dependent on the angle of incidence as well as the viewing direction. Next a reflectance model will be derived for non-Lambertian surfaces illuminated by ring light, based on the Torrance-Sparrow model.

As mentioned previously, in the current project only wear occurring on metallic surfaces is of interest. For such surfaces, specular reflection is dominant. The Torrance-Sparrow model was developed with the aim of describing the mechanism for specular reflection by rough surfaces. In this model the surface is modeled as

a collection of planar micro-facets. As explained in Subsection 4.4.2, the surface has a mean surface orientation **n**, the slope α of each planar facet with respect to the mean orientation is described by a probability distribution. Each facet reflects incident light in the specular direction determined by its slope. Since the facet slopes are randomly distributed, light rays are scattered in various directions. Torrance and Sparrow have assumed the facet slopes to be normally distributed. Furthermore, they have assumed the distribution to be rotationally symmetric about the mean surface normal **n**. Accordingly, the facet slopes may be represented by a 1D normal distribution (later on in Section 4.5 we will further explain that the model is not sensitive to the slope distribution):

$$p_\alpha(\alpha) = ce^{-\alpha^2/2\sigma_\alpha^2} \tag{4.20}$$

Consider the geometry shown in Fig. 4.11, the surface area dA_s is located at the origin of the coordinated frame, and its surface normal points in the direction of the z-axis. The surface is illuminated by a beam of light that is incident on the surface at the angle $(\theta_i,\ \phi_i)$. We are interested in determining the radiance of the surface in the direction $(\theta_r,\ \phi_r)$. Only those planar micro-facets whose normal vectors lie within the solid angle $d\omega'$ are capable of specularly reflecting light flux that is incident at the angle $(\theta_i,\ \phi_i)$ into the infinitesimal solid angle $d\omega_r$. From the angles θ_i, ϕ_i, θ_r, and ϕ_r, we can determine the local angle of incidence θ_i' and slope α of the reflecting facets:

$$\theta_i' = \frac{1}{2}\cos^{-1}(\cos\theta_r\cos\theta_i + \sin\theta_r\sin\theta_i\cos(\phi_r - \phi_i)) \tag{4.21}$$

$$\alpha = \cos^{-1}(\frac{\cos\theta_r + \cos\theta_i}{2\cos\theta_i'}) \tag{4.22}$$

For an angle of incidence $(\theta_i,\ \phi_i)$, the radiance in the direction $(\theta_r,\ \phi_r)$ of a rough surface patch whose facet slopes are normally distributed with standard deviation σ_α may be expressed as (for detailed derivation refer to [85, 61]):

$$L_r(\theta_i,\phi_i,\theta_r,\phi_r) = k_{spec}G(\theta_i,\phi_i,\theta_r,\phi_r)F(\theta_i',\eta')\frac{L_i d\omega_i}{\cos\theta_r}e^{-\alpha^2/2\sigma_\alpha^2} \tag{4.23}$$

where $G(\theta_i,\phi_i,\theta_r,\phi_r)$ is a *geometrical attenuation factor* when shadowing and masking effects[4] are considered, $F(\theta_i',\eta')$ the Fresnel reflection coefficient, η' represents the complex index of refraction and k_{spec} a constant that determines the fraction of incident energy that is specularly reflected. As a matter of fact, the radiance in the viewing direction is a random variable as facet slopes are randomly distributed. L_r in Eq. (4.23) represents the expectation of the radiance in the direction $(\theta_r,\ \phi_r)$.

[4]Shadowing and masking effects arise from the obstruction of incident or reflected light rays, which depend on the angle of incidence and the angles of reflection.

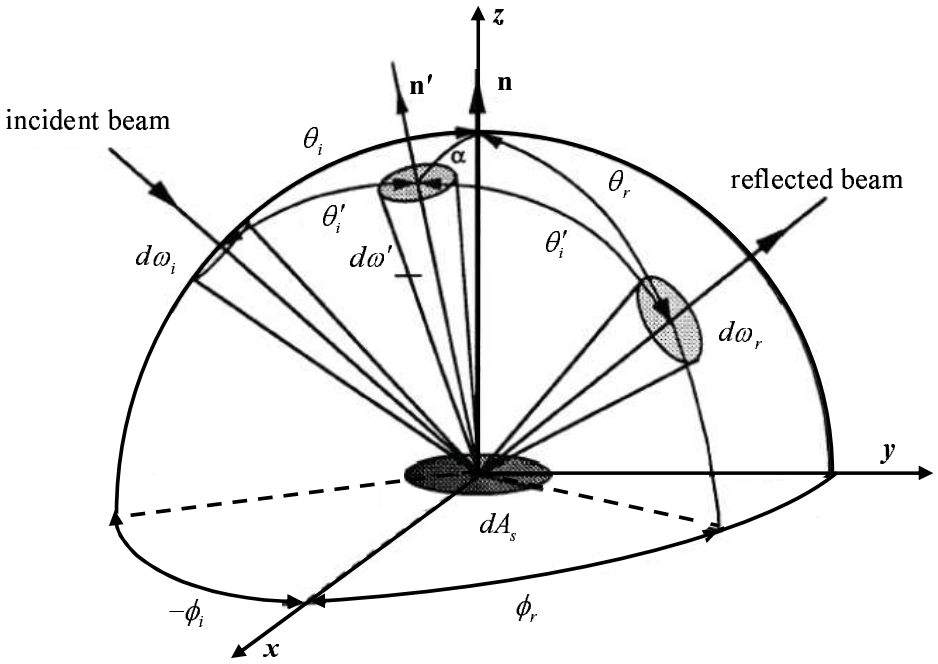


Figure 4.11: Coordinate system used to derive the Torrance-Sparrow model (adapted from [61]).

Next the radiance from any point(x, y) on the surface is derived when viewing from the direction (θ_r, ϕ_r) under ring light illumination. For simplicity, first we assume that within the surface area of interest all surface patches (which consist of planar micro-facets) have the same mean orientation **n** so that the orientation of the nominal plane is also **n**. Consider an arbitrary point (x, y) on the surface (see Fig. 4.12), for a ring element dA_i with radiance L_i illuminating at the angle (θ_i, ϕ_i), then the reflected radiance at the direction (θ_r, ϕ_r), according to Eq. (4.23), may be written as:

$$dL_r(x, y, \theta_r, \phi_r) = k_{spec} G(\theta_i, \phi_i, \theta_r, \phi_r) F(\theta_i', \eta') \frac{L_i d\omega_i}{\cos\theta_r} e^{-\alpha^2/2\sigma_\alpha^2} \tag{4.24}$$

From Fig. 4.12 we have:

$$\theta_i = \cos^{-1}(\frac{h}{\ell}) \tag{4.25}$$

$$\ell = \sqrt{h^2 + (x - r\cos\varphi)^2 + (y - r\sin\varphi)^2} \tag{4.26}$$

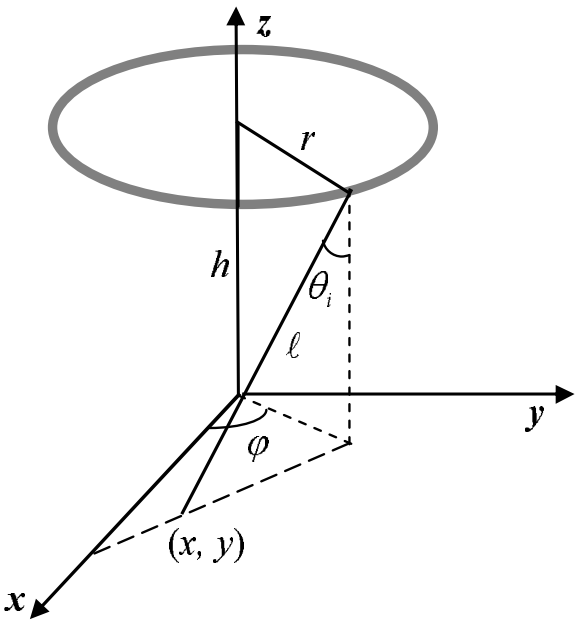


Figure 4.12: Coordinate system used to derive reflectance model under ring light.

$$\phi_i = \begin{cases} \tan^{-1}(\frac{y - r\sin\varphi}{x - r\cos\varphi}) & \text{if } r\cos\varphi \geq x \text{ and } r\sin\varphi \geq y; \\ \pi + \tan^{-1}(\frac{y - r\sin\varphi}{x - r\cos\varphi}) & \text{if } r\cos\varphi \leq x; \\ 2\pi + \tan^{-1}(\frac{y - r\sin\varphi}{x - r\cos\varphi}) & \text{if } r\cos\varphi \geq x \text{ and } r\sin\varphi \leq y. \end{cases} \tag{4.27}$$

Substituting Eqs. (4.10), (4.12) and (4.26) into Eq. (4.24) and integrating dL_r over φ from 0 to 2π, we obtain the reflected radiance at the direction $(\theta_r,\, \phi_r)$ from any point $(x,\, y)$ on the surface:

$$L_r(x, y, \theta_r, \phi_r) = \int_0^{2\pi} \frac{k_{spec} w r L_i G(\theta_i, \phi_i, \theta_r, \phi_r) F(\theta_i', \eta')}{(h^2 + (x - r\cos\varphi)^2 + (y - r\sin\varphi)^2)\cos\theta_r} e^{-\alpha^2/2\sigma_\alpha^2} d\varphi. \tag{4.28}$$

Eq. (4.28) represents the radiance of any point $(x,\, y)$ on the surface reflected at the direction $(\theta_r,\, \phi_r)$ under ring light illumination. In the next Subsection the relationship between image irradiance and surface radiance will be addressed.

4.4.6 Surface Radiance and Image Irradiance

It is necessary to find out the relationship between surface radiance and image irradiance as we are working on intensity images captured by the imaging system. Each pixel in a digital intensity image is represented in gray levels, which

are quantized measurements of image irradiance. Consider the image formation geometry shown in Fig. 4.13. For convenience, we will use the areas and solid angles shown in the figure to determine the surface radiance. The surface element dA_s, is projected by the lens onto an area dA_{im} on the image plane. Since the solid angles subtended from the center O of the lens by both areas dA_s and dA_{im} are equal, we can relate the two areas as:

$$dA_s = \frac{dA_{im} \cos\gamma}{\cos\theta_r} \left(\frac{z}{f}\right)^2, \tag{4.29}$$

where γ is the off-axis angle[5]. As the viewing direction θ_r varies, we see that

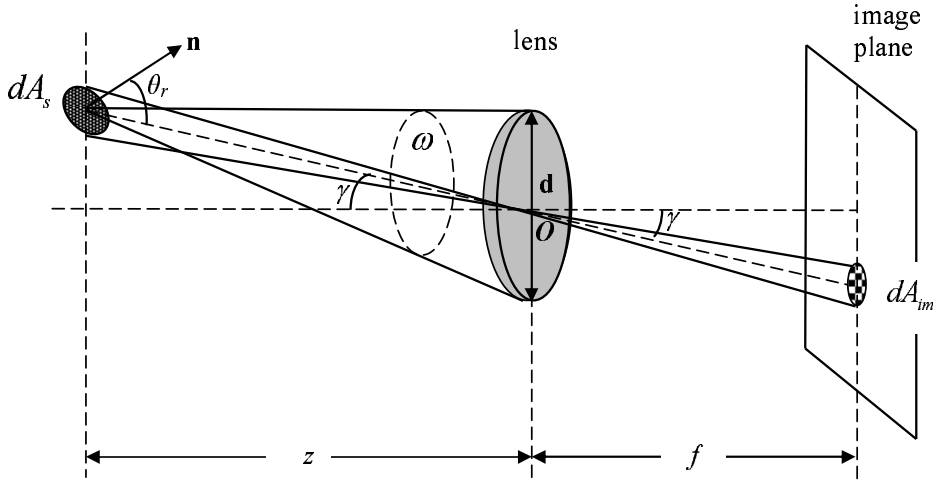


Figure 4.13: Image formation: light waves radiated by the surface area dA_s and gathered by the lens are projected onto an area dA_{im} on the image plane (adapted from [27, 61]).

the surface area dA_s that is projected onto the same image element (pixel) area changes as a function of θ_r. Since the image element area dA_{im} is constant for a given sensor, the surface area dA_s, must be determined from dA_{im}. All light rays radiated from dA_s that are incident on the lens area A_l are projected onto the image area dA_{im}. The solid angle ω subtended by the lens when viewed from the area dA_s is determined as

$$\omega = \frac{A_l \cos\gamma}{(z/\cos\gamma)^2} = \frac{\pi}{4}\left(\frac{d}{z}\right)^2 \cos^3\gamma \tag{4.30}$$

[5]The angle between the optical axis and the line from the center of the surface element to the entrance aperture nodal point.

where d is the diameter of the entrance aperture. The flux $d\Phi_r$ in Eq. (4.2) is the radiant flux received by the lens area A_l, which is projected to the image element area dA_{im}, i.e. the flux received by the image element is:

$$d\Phi_{im} = L_r \omega dA_s \cos\theta_r \tag{4.31}$$

By substituting Eq. (4.29) and Eq. (4.30) into Eq. (4.31), we get

$$d\Phi_{im} = L_r dA_{im} \frac{\pi}{4} \left(\frac{d}{f}\right)^2 \cos^4\gamma \tag{4.32}$$

Thus the image irradiance is:

$$I_{im} = \frac{d\Phi_{im}}{dA_{im}} = L_r \frac{\pi}{4} \left(\frac{d}{f}\right)^2 \cos^4\gamma. \tag{4.33}$$

That is, image irradiance is proportional to surface radiance. And the sensitivity of such an imaging system is not uniform over an image, but is constant for a particular point in the image. Ideally, an imaging device is calibrated so that this variation[6] in sensitivity as a function of γ can be removed. Other kinds of imaging systems, such as a microscope in the current case, lead to somewhat different expressions. Generally, however, image irradiance is proportional to surface radiance in such systems too [28]. Thus, the image irradiance is determined by the corresponding surface radiance and intrinsic camera parameters. For this reason, in the following only surface radiance L_r is discussed.

4.4.7 Evaluations

Because of the complexity of Eq. (4.28), it is very difficult, if not impossible, to find an analytical closed-form solution without simplification. For this reason numerical evaluations of this integral, instead, will be conducted in the following discussion.

We ignore the shadowing and masking effects in illumination and reflection, i.e. let $G(\theta_i, \phi_i, \theta_r, \phi_r) = 1$, and take $F(\theta_i', \eta')$ as a constant for simplicity. Note that for smooth surfaces, if $\theta_i' < 70°$, $F(\theta_i', \eta')$ is nearly constant [61] (see Fig. 4.14). Thus Eq. (4.28) may be simplified as:

$$L_r(x, y, \theta_r, \phi_r) = \int_0^{2\pi} \frac{k e^{-\alpha^2/2\sigma_\alpha^2}}{(h^2 + (x - r\cos\varphi)^2 + (y - r\sin\varphi)^2)\cos\theta_r} d\varphi. \tag{4.34}$$

where $k = k_{spec} w r L_i F(\theta_i', \eta')$ is considered as a constant. From Eqs. (4.25), (4.26), (4.21) and (4.22) we can see that α is a function of φ.

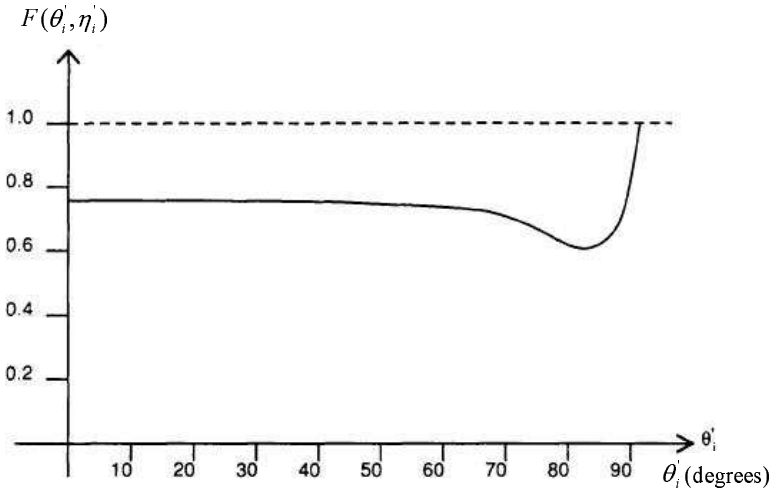


Figure 4.14: Typical plot of the Fresnel reflection coefficient as a function of the local incidence angle (adapted from [61]).

We first consider $\theta_r = 0$ in the evaluation. Fig. 4.15 shows the normalized radiance diagram reflected at the direction $\theta_r = 0$, where $h = 20\ mm$, $r = 16.65\ mm$, $\sigma_\alpha = 0.45$. Actually the radiance reflected at the direction $\theta_r = 0$ is rotationally symmetric along the z-axis under the assumption that all surface patches have the same mean orientation $\mathbf{n}$ as mentioned in Subsection 4.4.5. So in the following we only consider the radiance reflected from the points along the x-axis. Radiance curves along the x-axis for different surface roughness σ_α are given in Fig. 4.16, where $h = 20\ mm$ and $r = 16.65\ mm$. From this figure one can find that σ_α has a significant effect on the radiance, a larger roughness corresponding to a larger radiance in the direction $\theta_r = 0$. This is due to the fact that specular reflection is assumed to occur on each facet. For rougher surfaces (larger σ_α), light rays are scattered in more directions and more light rays reflected at the direction $\theta_r = 0$, whereas for smoother surfaces (small σ_α), light rays are more reflected to a small range of specular directions, less at the direction $\theta_r = 0$. In addition, one also can find that at a specific range of surface roughness (for instance, when $0.4 < \sigma_\alpha < \pi/3$ in the current case), uniform reflection is achieved (when $|x| < r$). When the viewing angle $\theta_r \neq 0$, the radiance will be not rotationally symmetric (see Fig. 4.17), where $h = 10\ mm$, $r = 16.65\ mm$ and $\sigma_\alpha = 1.2$), but symmetric with respect to the line $y = x\tan(\phi_r)$. Fig. 4.18 shows the radiance curves along the line $y = x\tan(\phi_r + \pi/2)$ when viewed from different directions, where $h = 10\ mm$, $r = 16.65\ mm$ and $\sigma_\alpha = 1$. Note that the term $\cos\theta_r$ appears in

[6]It is referred to as *vignetting* in photography.

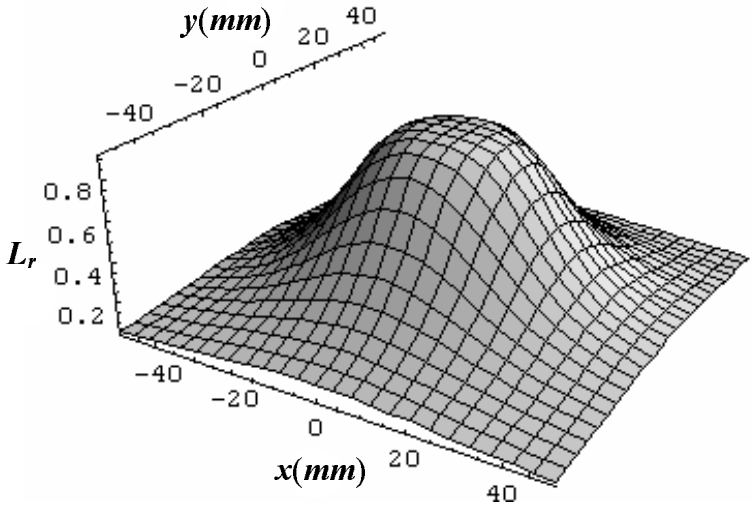


Figure 4.15: Normalized radiance viewed at the direction $\theta_r = 0$.

the
denominator of the model (see Eqs. (4.28) and (4.34)), therefore radiance increases with the increase of refection polar angle θ_r (see Fig. 4.17 and Fig. 4.18), with $\theta_r = \pi/2$ it approaching infinity. In reality this may not happen. As a matter of fact, with the increase of the angle of reflection θ_r, masking and shadowing will become dominant, i.e. the geometrical attenuation factor $G(\theta_i, \phi_i, \theta_r, \phi_r)$ decreases faster than $1/\cos\theta_r$. Therefore when θ_r is large, we must take masking and shadowing effects into consideration.

When considering masking and shadowing effect, the final expression of $G(\theta_i, \phi_i, \theta_r, \phi_r)$ is found to be [85, 61]:

$$G(\theta_i, \phi_i, \theta_r, \phi_r) = \min(1, \frac{2\cos\alpha\cos\theta_r}{\cos\theta_i'}, \frac{2\cos\alpha\cos\theta_i}{\cos\theta_i'}). \tag{4.35}$$

Fig. 4.19 shows the radiance curves when masking and shadowing effect is considered, where $h = 10\ mm$, $r = 16.65\ mm$ and $\sigma_\alpha = 1.2$. As mentioned earlier, the FOV of the camera is much smaller than the ring light covers. From the magnified view we can see that for a particular viewing angle θ_r the radiance fluctuation within the FOV is small.

From the above evaluations and analysis, we can see that, under ring light illumination, some parameters of the reflectance model, for instance, surface roughness σ_α and viewing direction θ_r have significant influence on the radiance. Surface radiance increases when surface roughness σ_α or viewing angle θ_r grows (Figs 4. 15 to 4.18). However, surface radiance within the FOV of the camera is not very

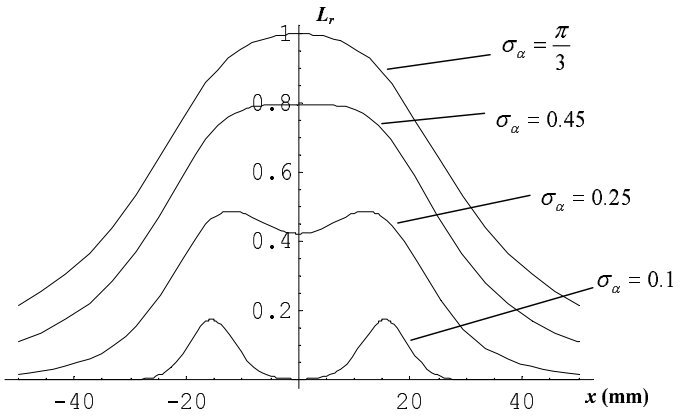


Figure 4.16: Radiance (normalized by the maximum for $\sigma_\alpha = \frac{\pi}{3}$) along x-axis viewed at the direction $\theta_r = 0$ for different surface roughness.

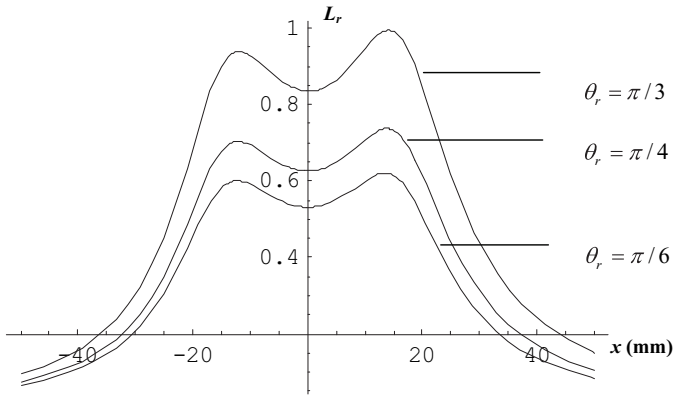


Figure 4.17: Radiance (normalized by the maximum for $\theta_r = \frac{\pi}{3}$) along x-axis reflected at directions of same azimuth angle $\phi_r = \pi$ but different polar angles.

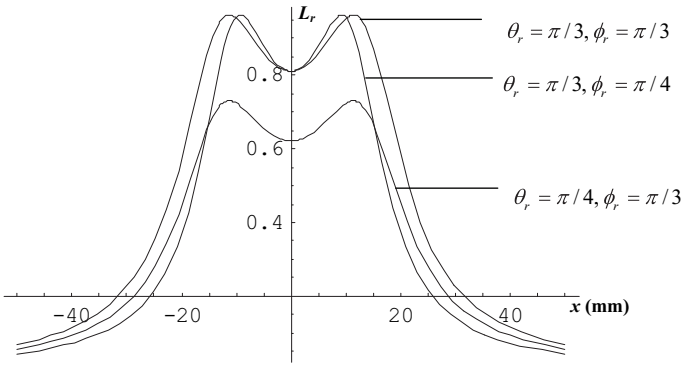


Figure 4.18: Radiance (normalized by the maximum for ($\theta_r = \frac{\pi}{3}$, $\phi_r = \frac{\pi}{3}$)) along the line $y = x\tan(\phi_r + \pi/2)$ viewed from different directions.

sensitive to the position, which is related to incidence direction of light rays. This is due to uniform illumination from the ring light.

Note that the formula derivation in Subsection 4.4.5 and the above evaluations are all under the assumption that the surface patches of interest have the same mean orientation **n**. In reality one surface patch has a mean orientation usually different from another, in other words, each surface patch has different local viewing direction θ_r (see Fig. 4.13) within the FOV of the camera. As it is very complicated to derive an analytical expression for this general case, here we only give some simple qualitative analysis. From above analysis and Figs. 4.17, 4.18 and 4.19, one can anticipate that surface radiance from different points varies due to the difference of local viewing direction θ_r from different surface patches. Hence, even under uniform illumination surface radiance varies with different local viewing directions (at fixed optical axis of the imaging system).

To qualitatively assess the reflectance model under ring light illumination described above, some experiments were conducted using the current online wear monitoring system (refer to Chapter 3). Fig. 4.20 shows an image of a machined aluminum surface with a wear track generated by the wear tester (with normal load 2N and 5 cycles of rotation). The image was captured at 2.5X magnification. According to the resolution of the imaging system, we assume the size of a surface patch in this case is about $100\mu m^2$. Surface roughness measurements[7] were carried out using a Mirau type white-light interference microscope (Micromap), which has a lateral resolution about $1\mu m$. Thus we may assume the facet size

[7]Such roughness measurements were conducted by Tasan [83] at Surface Technology and Tribology group, University of Twente.

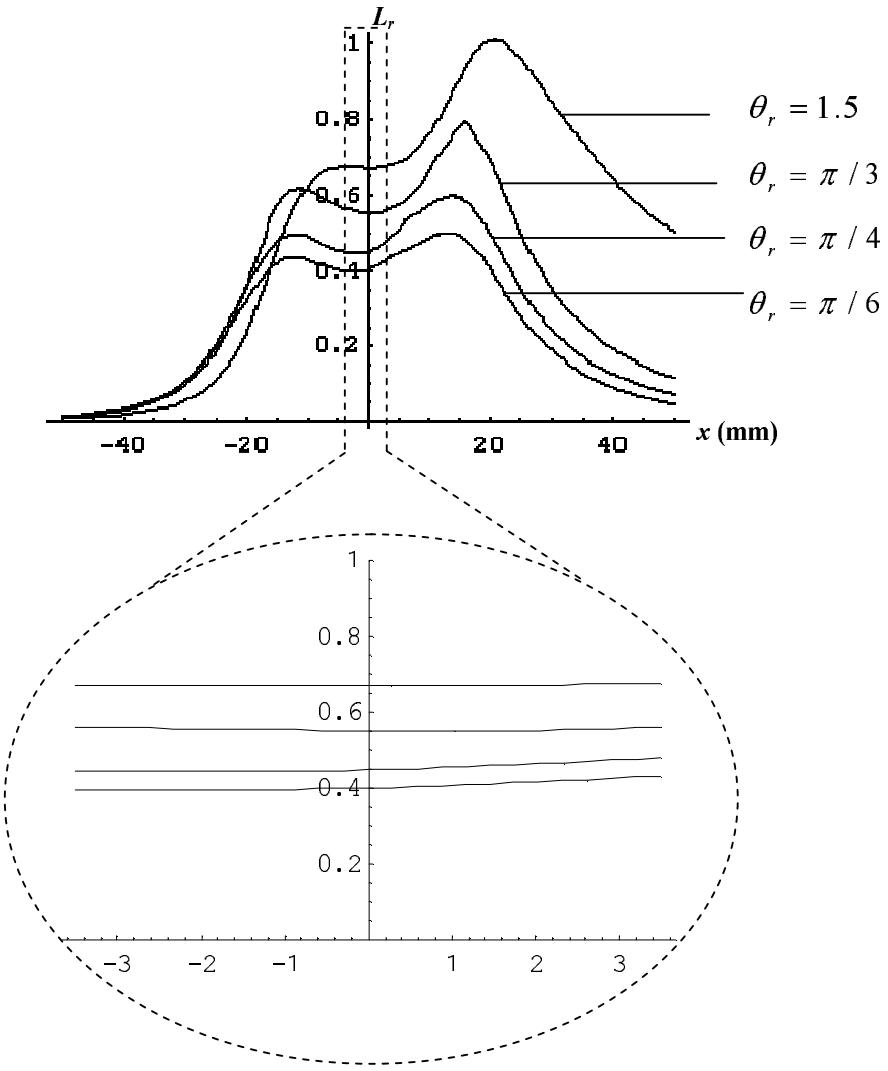
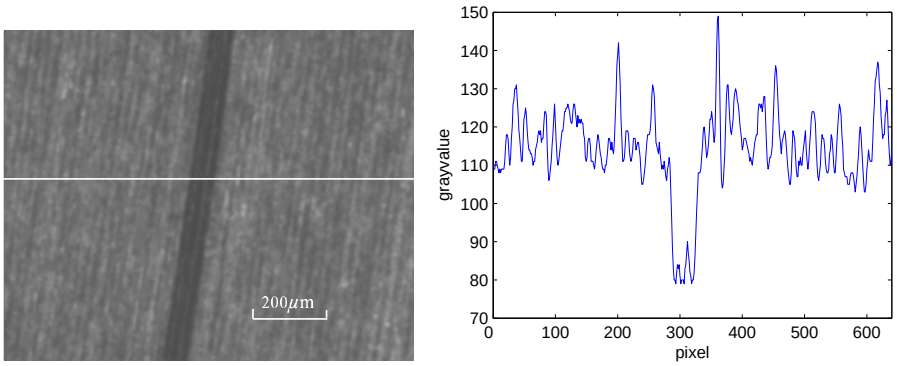


Figure 4.19: Radiance (normalized by the maximum for $\theta_r = 1.5$) along x-axis reflected at directions of same azimuth angle $\phi_r = \pi$ but different polar angles (masking and shadowing effect is considered).

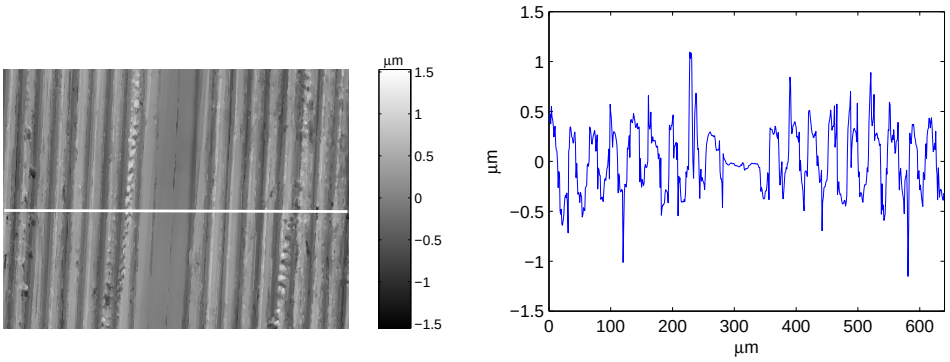


(a) Intensity image (b) Scan line indicated in (a)

Figure 4.20: An example image acquired by the current imaging system under ring light illumination.

is several μm^2. A graylevel coded topographical image of the same surface (not necessarily from the same area of the surface) is shown in Fig. 4.21. The measured RMS roughness R_q of the nonwear part of the surface is $0.34\mu m$ and the roughness of the wear track is $0.14\mu m$. From the measurements we see that the wear track becomes smoother after 5 cycles of rotation, i.e. smaller RMS roughness (see Fig. 4.21), which also means smaller RMS slope (Eq. (4.9)). According to Fig. 4.16, small RMS slope leads to small radiance, accordingly small image irradiance (intensity) (Eq. (4.33)). On the other hand, when the wear track becomes smoother, generally local viewing angles θ_r of surface patches become smaller when the optical axis of the imaging system is perpendicular to the nominal plane of the surface. According to Figs. 4.17, 4.18, 4.19, a small local viewing angle θ_r also leads to low radiance, accordingly low image intensity. Besides, according to the illumination model (see Fig. 4.7), when the local surface height becomes smaller, the equivalent height of the ring h becomes larger. This results in a lower surface irradiance and accordingly a lower radiance (image irradiance). All these together are consistent with the observation in Fig. 4.20 where the wear track becomes darker than the nonwear area of the surface in the image.

Hence, the variance of the image intensity (irradiance) as Fig. 4.20 illustrates reflects the differences of local viewing direction θ_r, local RMS roughness and local height of physical surface patches. In other words, when the ring light is in its uniform illumination range, radiance variance due to differences of local viewing angle and local roughness can still be clearly observed from the image.



(a) Graylevel coded topographical image of the same surface as shown in Fig. 4.20

(b) Profile of the scan line indicated in (a)

Figure 4.21: Surface roughness measurement.

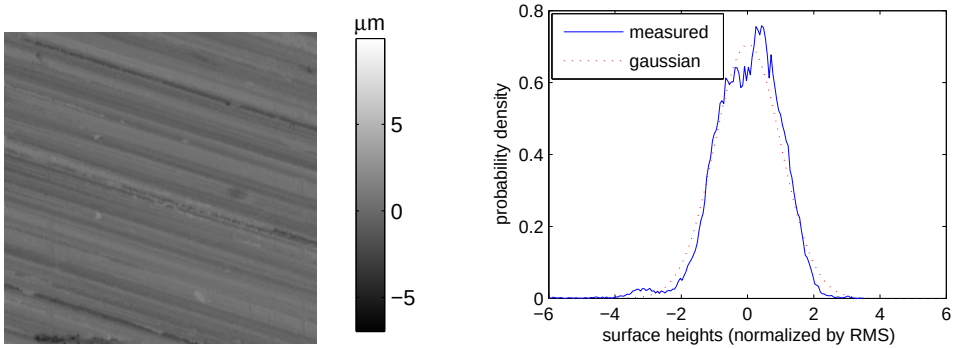
4.5 Discussions

The reflectance model presented above is independent of scale. Naturally, the size of a surface patch considered in the model depends on one's point of view. The smallest surface area resolved in a particular imaging situation is taken here to be a surface patch.

For the current imaging system in particular, the scale and resolution need to be considered as they determine how large or small wear can be observed. As mentioned earlier in Chapter 3 (see Table 3.1), the FOV of the microscope is about $3.5 \times 2.8mm^2$ at 2.5X. Its depth of focus (DOF) is about $30\mu m$ at 2.5X and $6\mu m$ at 10X. Its lateral resolution in object space is $9.8\mu m$ at 2.5X and $4.4\mu m$ at 10X. Based on the scale and resolution of the current imaging system, we can observe micro-wear occurring within the FOV on a relatively smooth surface with RMS roughness and peak to valley within the DOF. Thus in this case, we may assume the size of a surface patch is about $20\mu m^2$ at 10X and $100\mu m^2$ at 2.5X.

This reflectance model is derived from Torrance-Sparrow model by extending from a general point light source to a specific ring light. To make their model more generic, Torrance and Sparrow appended the Lambertian model [48] into the specular model to account for diffuse reflection that may result from multiple reflections or internal scattering. In the present model, however, we did not take diffuse reflection into consideration as under uniform illumination the diffuse reflection part becomes approximately constant from any point on a smooth surface.

In the derivation of the reflectance model we assumed the surface has a



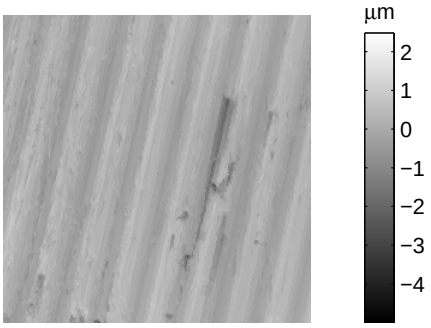
(a) Gray level coded topographical image

(b) Height distribution of the surface shown in (a)

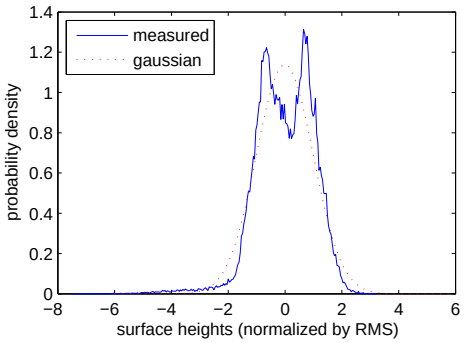
Figure 4.22: Surface roughness and its height distribution of a machined ball-bearing steel surface.

Gaussian slope distribution. Fig. 4.22 shows the height image of a machined ball-bearing steel surface and its height distribution. We can see that the Gaussian height distribution model fits the measured data quite well. And from Eq. (4.8), for gently varying rough surfaces with small slopes (α) of most facets, the Gaussian height distributed surfaces are also approximately Gaussian slope distributed and their roughness parameters are related in Eq. (4.9). Fig. 4.23 shows height distributions of the surface shown in Fig. 4.21 (a). From Figs. 4.23 (b) and (c) we can see that the height distribution of either the nonwear part of the surface or the whole surface with a wear track deviates from the Gaussian height distribution much more than the example shown in Fig. 4.22. However, the reflectance model for ring light illumination can be generalized to those surfaces whose slope distribution are not Gaussian shaped, which is often the case in practice. The reason is that, while the Gaussian roughness model appears explicitly in the Torrance-Sparrow model, it is integrated over all azimuth angles under the ring light illumination. As a result, the radiance is less sensitive to the actual surface roughness distribution although the surface radiance value is affected by the statistical surface roughness.

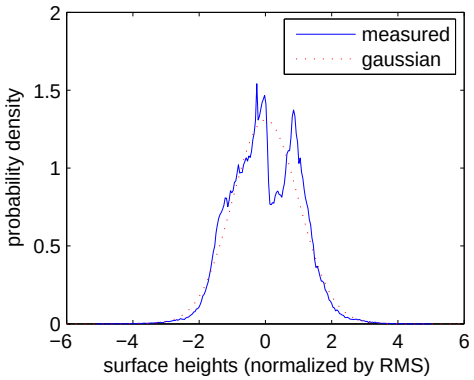
As we know that wear is the surface damage or removal or displacement of material from solid surfaces as a result of relative motion. A wear process either adhesive or abrasive (refer to Chapter 2) will change local surface heights, the mean orientation $\mathbf{n}$ of a local surface patch (local viewing angle θ_r meanwhile), local and overall statical surface roughness ($\sigma_\hbar$ or σ_α), hence, according to the above analysis the surface radiance and image irradiance thereafter will be changed as well. Therefore, by analyzing the image sequences of a wearing surface, captured



(a) Topographical image from the nonwear part of the surface shown in Fig. 4.21 (a)



(b) Height distribution of the surface shown in (a)



(c) Height distribution of the surface shown in Fig. 4.21 (a)

Figure 4.23: Height distributions of the surface shown in Fig. 4.21 (a).

by the current imaging system with ring light illumination, wear can be detected and monitored. Techniques for detection and monitoring of wear will be presented in Chapter 5 and Chapter 6 respectively.

4.6 Conclusions

The following conclusions may be drawn from this chapter:

1. According to the illumination strategy and the features of interest to reveal in the current project, fiber optic ring light may be a proper illumination choice for the imaging system for online detection and monitoring of wear.

2. Under ring light illumination, uniform illumination on the nominal plane of the surface can be achieved when the ring is located in a proper height range. Irradiance diagrams predicted by the illumination model are in good agreement with those measured by a photometer.

3. A specular reflectance model, based on the Torrance-Sparrow model, is formulated for ring light illumination. The parameters that affect surface radiance are discussed and evaluated. The radiance variation predicted by the model with respect to the parameters is in good agreement with the measurements from the current imaging system.

4. A wear process that alters surface parameters, such as surface local height, RMS roughness, RMS slope and mean orientations of local surface patches etc., results in changes of surface radiance, image irradiance and image texture thereafter.

Chapter 5

Wear Detection

Detection of wear can be regarded as a texture segmentation problem in most cases. This chapter presents two filtering approaches to wear detection where two types of filters, Gabor filters and optimized filters, are investigated. The experimental results conducted on real wear samples with various surface textures confirmed the usefulness of the filtering approaches to wear detection.

5.1 Introduction

Wear can take place at every phase in the life of parts or machine elements, including the manufacturing process. It plays an important role in determining life span of machine elements. Timely detection of wear, therefore, is highly demanded in many applications in order to predict remaining life of elements, avoid further/bigger damage to a whole system, safeguard product reliability and reduce potential cost.

Modern fabrication of parts typically involves machining, grinding, and polishing to remove material and to create a surface with specific macroscopic dimensions and also microscopic roughness. One of the purposes of this process is to reduce the risk of suffering wear during the surface use. Most surfaces appear as regular textures to a certain extent due to the quasi-periodical process of surface fabrication. Fig. 5.1 shows an example of a surface of stainless steel manufactured by turning operation. From this image one can see approximately parallel strips, corresponding to grooves or lays on a real surface.

As stated in the previous chapter, when a beam of light is reflected by a rough surface, the intensity and pattern of the scattered radiation depend on the roughness heights, the spatial wavelengths of the surface and the wavelength of the light. When one surface slides upon another, the one with softer material wears and its local worn area becomes smoother as the asperities are removed. This will result in apparent texture changes in the images captured by a properly designed imaging system [100], which is described in detail in chapter 3. As a consequence, the worn regions of the surface, referred to as wear patterns in images, can be characterized as texture changes with respect to the wear-free background. The area, shape and orientation of wear patterns carry the information about the wear state of the surface. Thus, wear detection can be considered as a texture analysis problem [101].

In this chapter filtering techniques for online detection of wear are investigated. Two types of filters, namely Gabor filters and optimized filters are applied. The photomicrographs of the surface of the specimen under test are first filtered by these filters so that wear patterns on the surface are strengthened while wear-free regions attenuated. Then wear can be detected by a subsequent classification step.

This chapter is organized as follows: filtering techniques are briefly reviewed in Section 5.2; two proposed approaches to wear detection, using Gabor filters and optimized filters, are described in detail in Section 5.3 and Section 5.4 respectively, where mathematical foundations are provided in Subsections 5.3.1 and 5.4.1; unsupervised approach is described in Subsection 5.3.2 and supervised approach in Subsection 5.4.2; experiments and results of these approaches are given in Subsections 5.3.3 and 5.4.3 followed by discussions in Section 5.5; in the end, conclusions

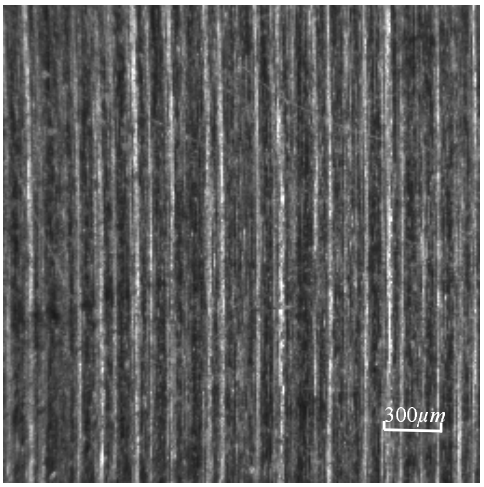


Figure 5.1: A surface image of stainless steel fabricated by turning.

are provided in Section 5.6.

5.2 Filtering Approaches

As stated in the previous section, most surfaces are present with a certain texture to some extent due to manufacture and machining processes, and wear patterns that occur on them mostly bring in texture changes, therefore, the wear detection problem encountered in textured surfaces can be regarded as a texture analysis problem at microscopic levels. As such, most methods for texture analysis can be used to detect wear on textured surfaces.

During the past decades, numerous approaches were presented for texture analysis [75, 86]. Tuceryan and Jain [86] classified the approaches to texture segmentation into five categories: statistical, geometrical, structural, model-based, and signal processing. The most commonly used techniques are the statistical, model-based and signal processing [72]. For most signal processing approaches, the textured image is passed through a linear transform, a filter or filter bank, followed by some energy measure. Recently Randen and Husøy [72] presented a review of these approaches and referred them as to *filtering approaches* due to the inherent similarities between these approaches. For statistical approaches, the textures are described by statistical measures. One typical and commonly applied and referenced method is the co-occurrence method introduced by Haralick [24]. For model-based approaches, some image model is assumed, for instance, the multi-resolution autoregressive (AR) model [55], and the Markov ran-

dom field (MRF) model [11, 10]. In this thesis we consider only filtering approaches on the wear detection problem for the following reasons: (1) The co-occurrence and AR schemes have a significant computational complexity [10, 72] which makes them not suitable for real-time wear detection; (2) Filtering approaches, on the other hand, can be implemented using fast algorithms, for instance FFT; and (3) they are usually insensitive to noise.

In filtering approaches, Fourier based techniques are good at obtaining global information. Wood [95] and Hosseini-Ravandi [29] used Fourier techniques to characterize textile web images and detect fabric defects. In the next chapter Fourier analytical techniques will be used to estimate the overall wear state of a surface. Fourier based techniques, however, do not provide, in general, enough information on local features. When wear only alters a small area of the image of the surface under detection, it is called *local* damage. Fourier analysis is usually not suitable for detection of local wear. To detect local wear patterns on the images of textured surfaces, methods that can localize features in the spatial as well as in the frequency domain are preferred in the current project. Digital wavelet transforms (DWT), used as multi-resolution spectral filters, provide both frequency and spatial local information about an image. They have been used to detect local defect in woven fabrics [40, 39]. However, it is difficult to characterize a texture pattern from the wavelet coefficients since the wavelet descriptors depend on pattern location [52].

The human eye is a highly efficient visual system and a robust pattern and texture analyzer. Gabor filters used in visual modeling have been successfully applied to a large variety of early vision tasks [14, 60]. An important property of Gabor filters is that they yield optimal joint localization, or resolution, in both spatial and spatial-frequency domains [13]. Because of such property, their mathematical tractability, as well as their ease of implementation for multi-channel filtering, Gabor filters are extensively used for texture analysis [7, 36, 17, 20, 46, 10], document analysis [54] and object detection [37, 8]. Studies of their ability to detect local microscopic wear with different scales and orientations in photomicrographs are unknown to the author.

If sufficient prior knowledge about certain texture, for instance, scales and orientations of some wear patterns in the current project is available, optimized filters can be designed to achieve optimal texture segmentation. Optimized Gabor filters [16] or infinite impulse response (IIR) filters with respect to one texture representation are filters with only a few free parameters, thus search space is very restricted [73]. A general form of finite impulse response (FIR) generally has many more free parameters and is often used in texture segmentation [73, 47].

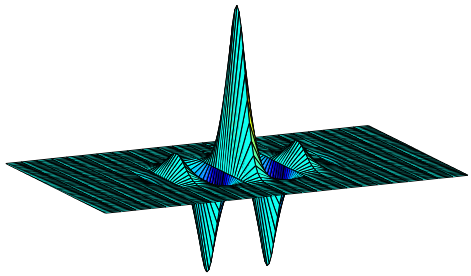
Multichannel Gabor filters and Optimized filters with optimal energy separation will be investigated in this chapter for wear detection for their advantages

mentioned above. In the following sections, we will first present the mathematical background of these two filters, then discuss how to apply them to wear detection in detail, respectively.

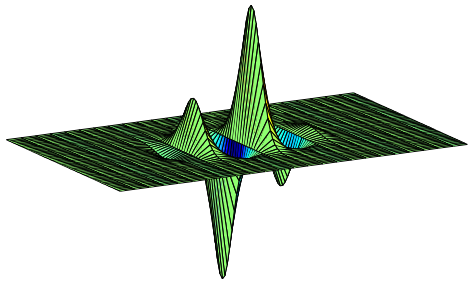
5.3 Wear Detection by Gabor Filters

5.3.1 Introduction to Gabor filters

In the spatial domain, the Gabor function is a complex exponential modulated by a Gaussian function. The Gabor function forms a complete but nonorthogonal basis set and its impulse response in the 2D plane has the following general form [46, 54]:



(a) Real component



(b) Imaginary component

Figure 5.2: Perspective view of a typical Gabor function in spatial domain.

$$g(x,y) = \frac{1}{2\pi\sigma_x\sigma_y}\exp[-\frac{1}{2}(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2})+2\pi j u_0 x] \tag{5.1}$$

where u_0 denotes the radial frequency of the Gabor function along $x-$axis (i.e. the 0° orientation), σ_x and σ_y the standard deviation of the Gaussian envelope along the x and y axes. Fig. 5.2 shows the perspective plot of a typical Gabor filter in the spatial domain. In the frequency domain, the Gabor function acts as a band-pass filter and the Fourier transform of $g(x,y)$ is:

$$G(u,v) = \exp[-\frac{1}{2}(\frac{(u-u_0)^2}{\sigma_u^2}+\frac{v^2}{\sigma_v^2})] \tag{5.2}$$

where

$$\sigma_u = \frac{1}{2\pi\sigma_x}, \tag{5.3}$$

$$\sigma_v = \frac{1}{2\pi\sigma_y}. \tag{5.4}$$

A self-similar[1] filter bank can be obtained by appropriate dilation and rotation of $g(x,y)$ through the generating function

$$g_{mn}(x,y) = \xi^{-m} g(x',y') \tag{5.5}$$

where $x' = \xi^{-m}(x\cos\theta_n + y\sin\theta_n)$, $y' = \xi^{-m}(-x\sin\theta_n + y\cos\theta_n)$, $\xi > 1$, $m = 1,2,\ldots,P$, $n = 1,2,\ldots,Q$. The integer subscripts m and n represent the index for scale (dilation) and orientation (rotation), respectively. P is the total number of scales and Q is the total number of orientations in the self-similar Gabor filter bank. For each orientation n, the angle θ_n is given by

$$\theta_n = (n-1)\pi/Q. \tag{5.6}$$

The scale factor ξ^{-m} in (5.5) ensures that the energy $E_{mn} = \int_{-\infty}^{\infty}|g_{mn}(x,y)|^2 dxdy$ is independent of m [54]. Thus, all the filters in the Gabor filter bank have the same energy, irrespective of their scale and orientation.

The nonorthogonality of the Gabor functions implies that there is redundant information in the filtered images. As illustrated in [54] the parameters of the Gabor filter bank, ξ, σ_x and σ_y are computed using the following formulas to ensure that the half-peak magnitude responses of adjacent filters touch each other as shown in Fig. 5.3 to reduce redundancy:

[1]By self-similar we mean Gabor functions with similar geometrical shape but different scale and orientation.

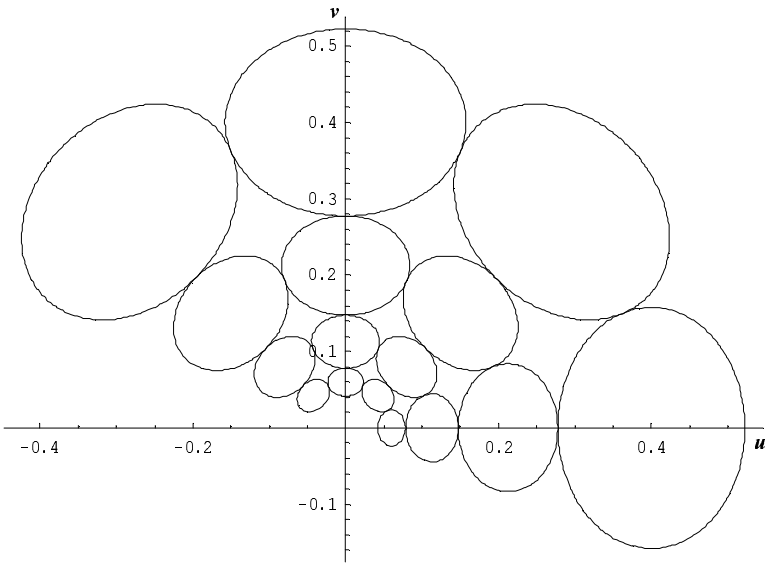


Figure 5.3: The contours indicate the half-peak magnitude responses of the 4×4 Gabor filter bank. The filter parameters used are $u_h = 0.4$, $u_l = 0.06$, $P = 4$ and $Q = 4$.

$$\begin{aligned} \xi &= \left(\frac{u_h}{u_l}\right)^{\frac{1}{P-1}} \\ \sigma_x &= \frac{\sqrt{2\ln 2}(\xi+1)}{2\pi u_h(\xi-1)} \\ \sigma_y &= \left(2\pi\tan(\frac{\pi}{2Q})\sqrt{\frac{u_h^2}{2\ln 2} - \frac{1}{(2\pi\sigma_x)^2}}\right)^{-1} \end{aligned} \tag{5.7}$$

where $u_0 = u_h$, and u_l and u_h are the lowest and highest frequencies of interest. The mathematical derivation of these formulas (5.7) can be found in Appendix A. A bank of self-similar Gabor filters formed by rotation (varying n) and dilation (varying m) of the basic Gabor filter (5.1) is used to perform power spectrum sampling of the image of the surface under detection.

Each of the complex Gabor filters has real (even) and imaginary (odd) parts that can be conveniently implemented as the spatial masks of size $N \times N$. In order to have a symmetrical region of support, N is preferred to be an odd number.

In [46] the real and imaginary part of each of the complex Gabor functions are used for filtering. In our work, only the real part of Gabor function is adopted,

i.e., Eq. (5.1) is replaced with Eq. (5.8)

$$g(x,y) = \frac{1}{2\pi\sigma_x\sigma_y}\exp[-\frac{1}{2}(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2})]\cos(2\pi u_0 x) \tag{5.8}$$

Since the real part of the Gabor function acts as a blob (or bar) detector [8] and the imaginary part acts as an edge detector [59, 9], using the combination of real and imaginary parts of the Gabor function might offer better localization of wear patterns than only using the real part. However, in surface inspection, the detection of wear is more important than the localization of wear or the edges. Therefore, the goal of precise wear localization or edge localization has not been considered in this work. In addition, our experiments also show that the imaginary part does not provide further improvement for the detection of wear patterns. On the other hand, speed is a big concern for real-time online wear detection and monitoring. Timely detection of wear from real time images requires high computation speed. Using only the real part greatly saves computational load but not much reduces the performance of detection.

5.3.2 Unsupervised Wear Detection

Segmentation of a similar class of local wear patterns with *a priori* knowledge about the orientation and size of a sample wear can be regarded as *supervised wear detection*. When the approximate orientation and size of wear patterns are known, the power spectrum sampling of the image plane is not necessary. In such cases, the segmentation can be achieved by only one Gabor filter, from the Gabor filter bank (or one optimized filter, discussed in Section 5.4) that can provide best discrimination of texture features against the wear pattern. In most practical cases the dimension and orientation of local wear occurred on surfaces vary randomly. Therefore, a complete automation of visual wear inspection process requires *unsupervised wear detection* that can be performed online. The term "unsupervised wear detection" refers to the detection of an unknown class of wear patterns for which there is no training. We will focus on unsupervised wear detection in this section.

Multichannel filtering theory for the processing of visual information in the biological model of the human visual system has inspired various texture segmentation algorithms [46, 54]. A variation of this algorithm, which can be used for online wear detection, was presented in [100, 101]. This modification of the multichannel filtering algorithm for online wear detection leads to a reduction in computational complexity and false alarm. In this subsection we will present the multichannel filtering scheme in detail.

Unsupervised wear detection requires simultaneous surface inspection at local and global scales. Multichannel filtering allows multiresolution analysis of surface texture. Fig. 5.4 illustrates the procedure of this approach.

Multichannel Gabor Filtering

Each image is filtered with a bank of Gabor filters detailed in Subsection 5.3.1. Each of these Gabor filters is selectively tuned to a narrow range of frequency and orientation. The octave (dyadic) band decomposition is commonly used for wavelet decomposition and was also used in this work for the selection of frequency bands for the power spectrum sampling of the image. Each of the Gabor functions Eq. (5.5) can be implemented as a spatial mask of size $N \times N$. The selection of mask size and number of channels has big influence on detection performance. As a compromise between computational complexity and performance, 16 Gabor filters distributed at four scales ($P = 4$) and four orientations ($Q = 4$) as shown in Fig. 5.5 are selected empirically in this work. Each of these Gabor filters was implemented as a spatial mask of size 7×7. Every image $I(x, y)$ (image under detection $I^d(x, y)$ or reference image $I^r(x, y)$) is filtered by each of 16 Gabor filters and the magnitude of the filtered image $I_{mn}(x, y)$ is computed using Eq. (5.9):

$$I_{mn}(x, y) = |I(x, y) * g_{mn}(x, y)| \tag{5.9}$$

where "*" denotes 2D convolution.

Nonlinearity

Next, a local nonlinear function is used to rectify the multichannel filter response. This nonlinear function transforms both negative and positive amplitudes to positive amplitudes. There are several nonlinear functions that are good for prior texture segmentation, such as the magnitude $|\cdot|$, the squaring $|\cdot|^2$ and rectified sigmoid $|\tanh(x)|$ nonlinearity [72]. The magnitude nonlinearity requires minimum computations and is, therefore, preferred in this work. This nonlinearity is inherent while computing the magnitude of Gabor filter images (5.9), therefore, no extra computational burden is added. The 16 images, $I^d_{mn}(x, y)$, represent the features of the image under detection.

Feature Difference

The filtered images from different channels are called *feature images*. The same operation is also applied to a wear-free reference image to get reference feature images $I^r_{mn}(x, y)$, which are used to compute feature difference images for recognizing wear features. The reference feature images are computed at the beginning

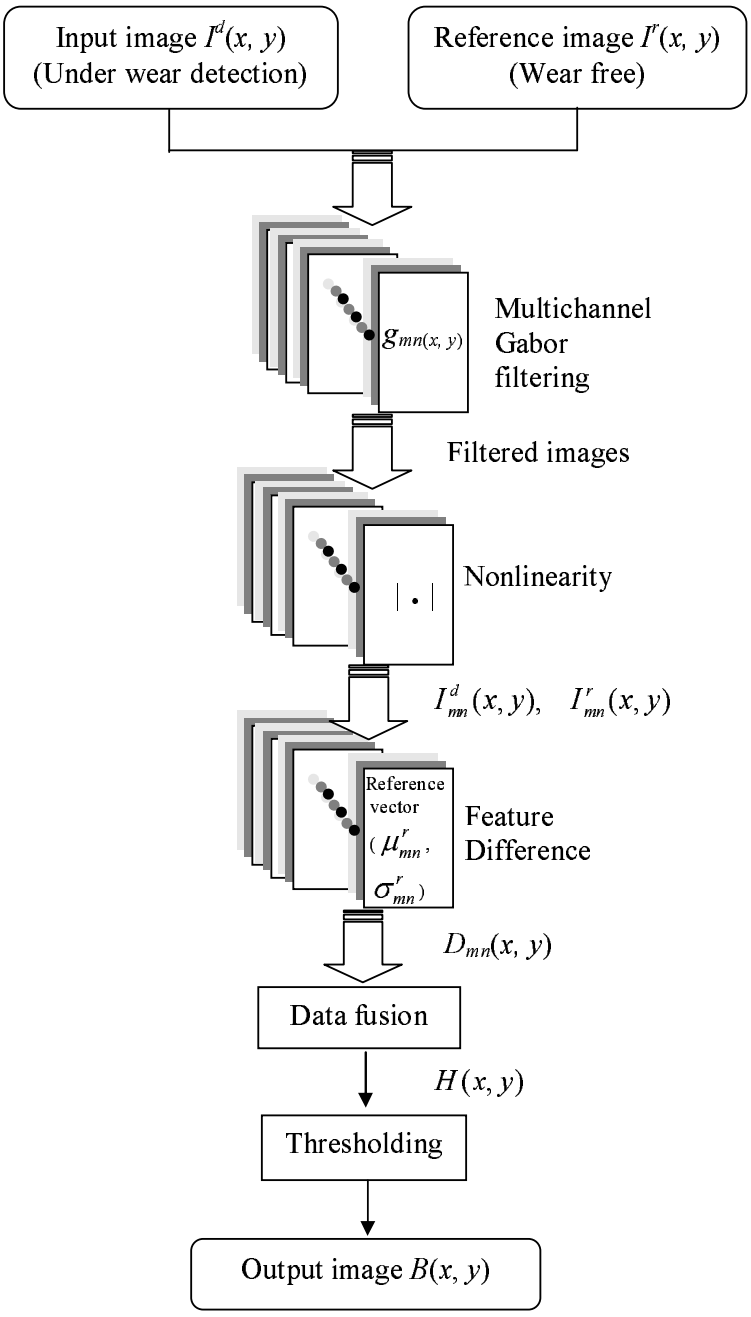


Figure 5.4: Unsupervised wear detection procedure.

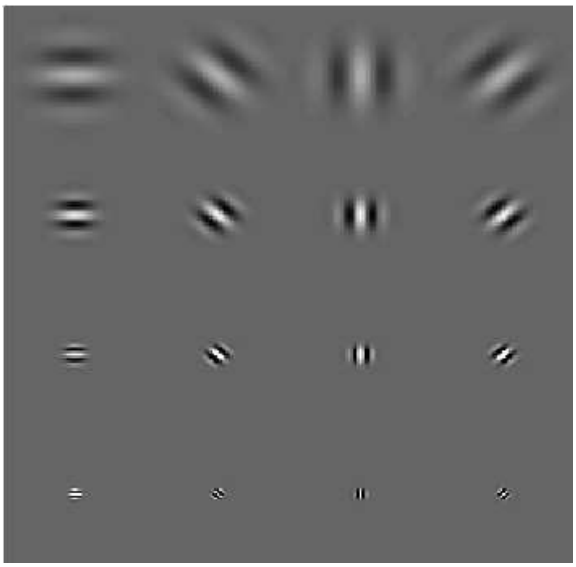


Figure 5.5: 4×4 real Gabor filters in spatial domain.

of the wear detection as a *calibration* operation, which does not increase computational burden of real-time online detection.

The first order statistics are fundamental for visual characterization of texture [42]. Therefore, the mean μ^r_{mn} and the standard deviation σ^r_{mn} from each of these 16 images $I^r_{mn}(x, y)$ are used to discriminate wear in the image under detection. A decision rule must be chosen for the characterization of pixels in $I^d_{mn}(x, y)$ based on reference features of $I^r_{mn}(x, y)$. The decision rule selected in this work is [46]:

$$D_{mn}(x,y) = \begin{cases} I^d_{mn}(x,y) & \text{if } |I^d_{mn}(x,y) - \mu^r_{mn}| \geq k.\sigma^r_{mn} \\ 0 & \text{otherwise} \end{cases} \tag{5.10}$$

The parameter k determines the sensitivity and is chosen to control the probability of false rejection. An empirically determined value of $k = 3$ was found suitable and used for all the experiments in this work. The next step is to combine pixels from the difference images $D_{mn}(x, y)$ to be a unique feature image so that only possible wear features remain.

Data Fusion

As the information gathered by different channels from the same image is often uncertain, fuzzy or incomplete, it is better to combine all the feature difference

images. The approach proposed in [20, 46] for data fusion is attractive for its computation simplicity, therefore is also adopted in this work. The fusion procedure is: (1) vector addition of all pixels of $D_{mn}(x, y)$ from different orientations n at the same scale m, followed by (2) geometrical mean of resultant pixels at adjacent scales. This can be formulated as:

$$C_m(x, y) = \sum_{n=1}^{Q} D_{mn}(x, y), \tag{5.11}$$

$$H(x, y) = \frac{1}{P-1} \sum_{m=1}^{P-1} \sqrt{C_m(x, y) C_{m+1}(x, y)}. \tag{5.12}$$

The first step of the fusion scheme in Eq. (5.11) generates a set of 4 images $C_m(x, y)$ from the addition of all the pixels from different orientations n at the same scale m with $m = 1, ..., 4$. Through this step, the information on the likely wear features obtained in the 4 orientations is concentrated into a single image $C_m(x, y)$ for each scale level m. Then by Eq. (5.12), the geometrical mean of resultant pixels of $C_m(x, y)$ at adjacent scales creates 3 images; pixels from all these 3 images are averaged to produce a unique fused image $H(x, y)$. Through the second step, only those wear features that appear in at least two adjacent resolution levels are considered. Thus, the fused image $H(x, y)$ contains the joint contribution from the 16 ($P \times Q$) channels.

Essentially this data fusion scheme preserves those wear features that occur at any orientations by vector addition, yet reduces false alarms by considering only those which appear in at least two adjacent resolution levels through the operation of geometrical mean.

Thresholding

The final stage of wear detection is the binarization of $H(x, y)$ to separate the wear patterns and suppress the pixels not belonging to them. In this stage, a binary image $B(x, y)$ is provided where local wear patterns appear segmented from the regular textured background. This is achieved by thresholding $H(x, y)$. Values below the threshold are considered as belonging to the background, and those above the threshold are considered as belonging to wear patterns.

From the histogram of one fused image (Fig. 5.6) we can find two peaks, each peak belonging to one class (the left peak to background noise, the right peak to wear patterns). For this histogram shape, the best threshold value T can be found through an automatic iterative isodata algorithm developed by Ridler and Calvard [77, 34]. The algorithm initially segments the histogram of $H(x, y)$ into two parts using a starting threshold value such as T_0 being equal to half of the

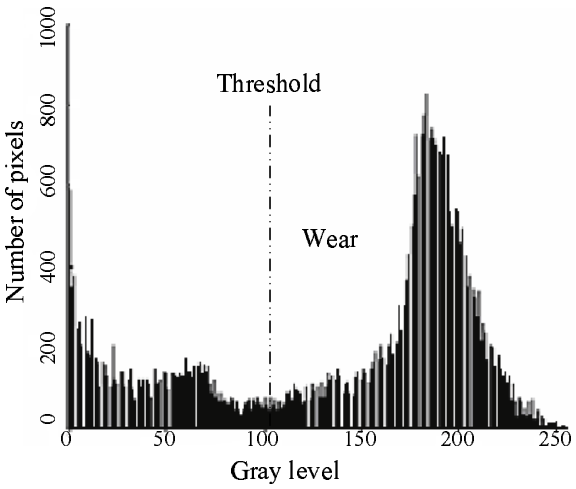


Figure 5.6: Histogram of a fused image $H(x, y)$.

maximum dynamic range of $H(x, y)$. The sample mean $\mu_{w,0}$ of the gray values associated with the foreground pixels (wear patterns) and the sample mean $\mu_{b,0}$ of the gray values associated with the background pixels are computed. A new threshold value T_1 is then computed as the average of these two sample means. The process is repeated, based upon the new threshold, until the threshold value does not change any more, i.e.

$$T_k = (\mu_{w,k-1} + \mu_{b,k-1})/2, \text{ until } T_k = T_{k-1}. \tag{5.13}$$

The thresholding operation further reduces the probability of false alarm.

In summary, through the above procedure (Fig. 5.4), wear patterns with different sizes and orientations can be detected.

5.3.3 Experiment and Results

The performance of the multichannel Gabor filtering scheme for online wear detection described above was evaluated through a series of experiments. The experimental setup introduced in Chapter 3 has been applied. Images of the wear samples under detection were acquired and digitized into 512×512 pixels with eight-bit pixel depth. Three wear samples with different surface textures were tested. One is a machined steel ring with a wear track in the middle (Fig. 5.7(d)), the second is a vapor-polished aluminium bar with wear scratches (Fig. 5.7(e)), the third one is a ground steel bar with wear scratches (Fig. 5.7(f)). In the experiment, a bank of 4×4 Gabor filters (Fig. 5.5) described in Subsection 5.3.2

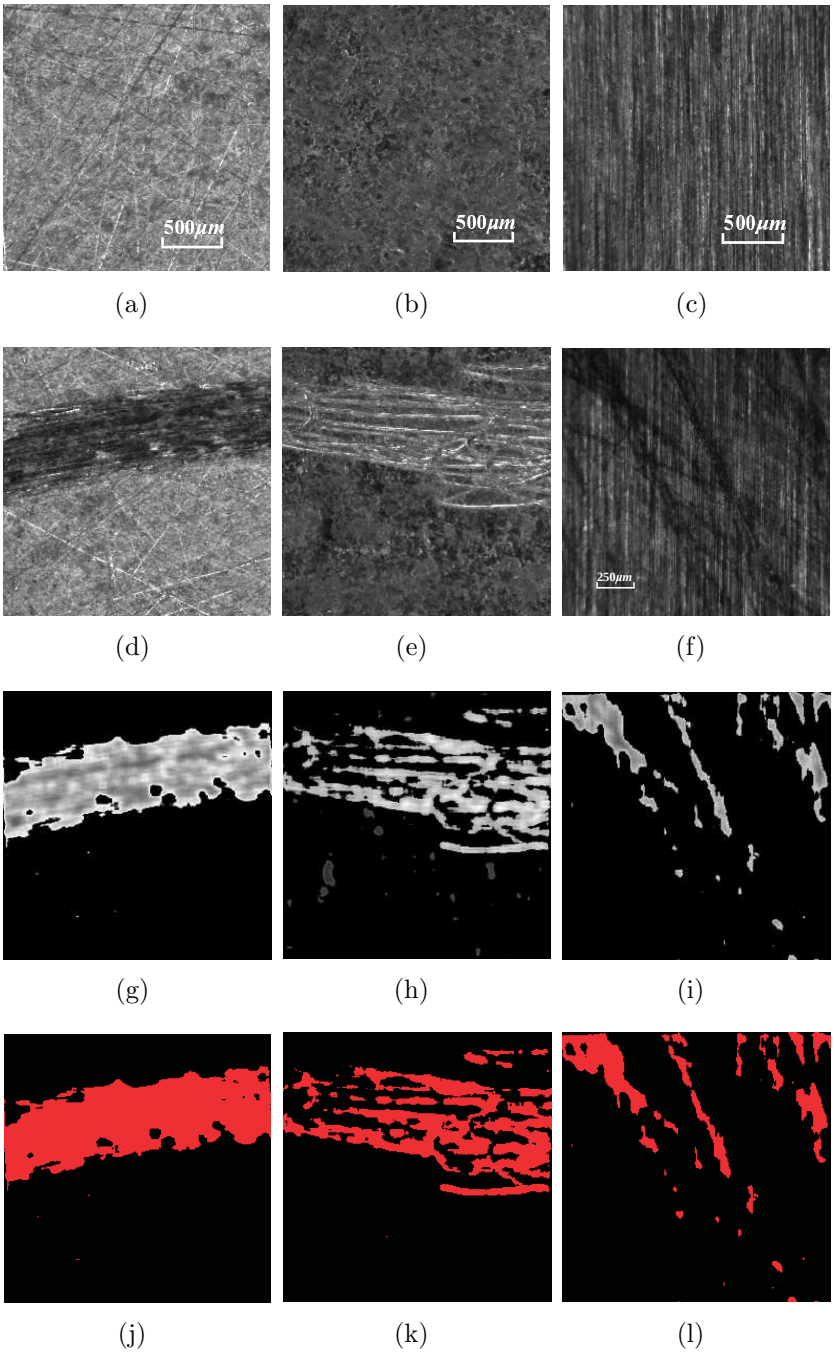


Figure 5.7: Unsupervised wear detection on test samples: reference images of the samples in (a), (b) , (c) ; images of the samples with wear in (d), (e), (f); corresponding filtered images in (g), (h), (f); segmented wear patterns in (k), (m), (n).

are adopted for detection of wear patterns. Each Gabor filter is implemented as a spatial 7×7 mask for a computational compromise. The segmentation of wear patterns in Fig. 5.7 becomes more exact as the mask size is increased from 7×7 to 9×9 or 11×11. However, a small mask size corresponds to small computation load, which is desirable for realtime wear detection. The frequency range of Gabor filters in the filter bank depends on the range of wear patterns to be detected. The experimental results are shown in Fig. 5.7, of which the top row (Figs. 5.7(a), (b) and (c)) are wear-free images captured during the calibration phase; the second row (Figs. 5.7(d), (e) and (f)) are images acquired during wear processes; the third row (Figs. 5.7(g), (h) and (i)) are images after multichannel filtering and fusion operation; the bottom row (Figs. 5.7(j), (k) and (l)) shows the segmented wear patterns. From these experiments one can find that, using the proposed multichannel filtering algorithm, the wear patterns (wear scratches or wear track) occurring on the surfaces of the three samples, irrespective of their sizes and orientations, are successfully detected.

After the segmentation of the wear region, important wear features such as wear area, centroid, area moments and products of inertia, wear orientation, and so on, can be calculated from the separated wear patterns to further characterize the wear patterns detected.

For an image of 512×512 pixels and a 4×4 Gabor filter bank implemented in masks of size 7×7, the wear detection procedure discussed above takes several seconds on a general-purpose Pentium IV computer. Hence, realtime implementation of this approach requires additional digital signal processor (DSP) hardware.

5.4 Wear Detection by Optimized Filters

The Gabor filters and the infinite impulse response (IIR) filters are the filters with only a few free parameters and therefore the search space for optimization is very restricted. Better optimization results might be obtained when the number of free available parameters of a filter is large. A general FIR filter has generally more free parameters than an IIR or a Gabor filter. The single biggest advantage of FIR filters is that they can implement any impulse response, provided it is of finite length.

Optimized filters have been used for texture segmentation [73] and defect detection [47]. In this section, a new approach to online wear detection using linear FIR filters with optimized energy separation is investigated. The surface textures are modeled by their autocorrelation functions. The linear FIR filters that guarantee optimal discrimination of energy in local regions rather than optimal representation are used in this work.

5.4.1 Mathematical Foundations

The feature extraction model used to design optimal filters is illustrated in Fig. 5.8. This model has been used in several references [73, 47]. Basically, the objective of the optimal filter $h(x, y)$ is to extract those frequencies where the wear-free textures have low signal energy and the texture with wear has high signal energy. If this is accomplished, the wear regions in the composite inspection image can be segmented by analysis of the local texture energy.

As shown in Fig. 5.8, filtering of the acquired image $I(x, y)$ with the filter $h(x, y)$ generates a new image $s(x, y)$:

$$s(x, y) = h(x, y) * I(x, y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} h(m, n) I(x - m, y - n) \tag{5.14}$$

where * denotes 2D convolution and $h(x, y)$ is an $M \times N$ optimal filter to be designed. The energy of pixels in a gray-level image $s(x, y)$ is defined as $E[s^2(x, y)]$, where E stands for the expectation. The wear patterns in the inspection images are assumed to be local rather than global. Therefore, a local operator is needed to compute the energy of the filtered image $s(x, y)$ in a local region. A local energy measure can be achieved by smoothing the image $s^2(x, y)$ with a smoothing filter $\mathcal{G}(x, y)$. That is, the feature image is computed by nonlinearity:

$$z(x, y) = |s(x, y)|^2 = s^2(x, y) \tag{5.15}$$

and smoothing:

$$w(x, y) = z(x, y) * \mathcal{G}(x, y) \tag{5.16}$$

Finally, a thresholding operation is applied to the feature image for associating a class label with wear regions.

In this subsection, the feature mean and variance are first expressed in a matrix/vector notation, which are used for designing optimized filters.

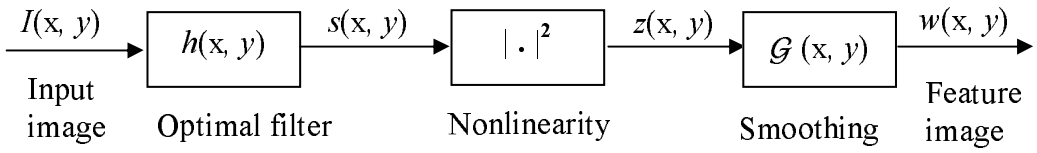


Figure 5.8: Block diagram of the feature extraction model.

Vector Formulation

By lexicographically ordering of the columns of the filter $h(x, y)$, it can be represented as a vector of length $Z = M \times N$:

$$\mathbf{h} = \begin{pmatrix} h(0,0) \\ \vdots \\ h(0,N-1) \\ \vdots \\ h(1,0) \\ \vdots \\ h(1,N-1) \\ \vdots \\ h(M-1,N-1) \end{pmatrix} \tag{5.17}$$

Likewise, the $M \times N$ window of $I(x, y)$ around pixel (x, y) can be written as

$$\mathbf{i}(\mathbf{x},\mathbf{y}) = \begin{pmatrix} I(x,y) \\ \vdots \\ I(x,y-N+1) \\ \vdots \\ I(x-1,y) \\ \vdots \\ I(x-1,y-N+1) \\ \vdots \\ I(x-M+1,y-N+1) \end{pmatrix} \tag{5.18}$$

Then, for every pixel in $s(x, y)$, the output in (5.14) can be rewritten as

$$s(x,y) = \mathbf{h}^T \mathbf{i}(x,y) \tag{5.19}$$

The Feature Mean

Assume that the inspection image is a random process, which is wide sense stationary over the region of interest. Using the vector formulation, the mean value of the feature image $w(x, y)$ can be derived as follows [73]:

$$\mu_w = E[w(x,y)] = E[z(x,y) * \mathcal{G}(x,y)] \tag{5.20}$$

The smoothing filter $\mathcal{G}(x, y)$ is a unit gain low pass filter. Therefore, the mean feature value at the output of this filter is equal to the mean feature at the input.

Assuming that the filter coefficients are such that $\sum_{x,y} \mathcal{G}(x,y) = 1$, (5.20) can be written as

$$\begin{aligned} \mu_w &= E[w(x,y)] = E[s^2(x,y)] \\ &= \mathbf{h}^T E[\mathbf{i}(x,y)\mathbf{i}(x,y)^T]\mathbf{h} \\ &= \mathbf{h}^T \mathbf{R}_{ii} \mathbf{h} \end{aligned} \tag{5.21}$$

where $\mathbf{R}_{ii} = E[\mathbf{i}(x,y)\mathbf{i}(x,y)^T]$ in (5.21) is the autocorrelation matrix of the vector $\mathbf{i}(x,y)$. The autocorrelation matrix of the vector $\mathbf{i}(x,y)$ can be constructed from the 2D autocorrelation function of the window image $\mathbf{i}(x,y)$ [73]. Since the $\mathbf{R}_{ii}$ is symmetric, the derivative of the mean feature value μ_w is given by

$$\frac{\partial \mu_w}{\partial \mathbf{h}} = \frac{\partial (\mathbf{h}^T \mathbf{R}_{ii} \mathbf{h})}{\partial \mathbf{h}} = 2\mathbf{R}_{ii}\mathbf{h} \tag{5.22}$$

The Feature Variance

The variance of the feature image $w(x,y)$ is given by

$$\begin{aligned} \sigma_w^2 &= E[(w(x,y) - \mu_w)^2] \\ &= E[w^2(x,y)] - \mu_w^2 \end{aligned} \tag{5.23}$$

By substituting the vector form of (5.16), i.e. $w(x,y) = \mathbf{w}^T \mathbf{z}(x,y)$, into (5.23) we get

$$\begin{aligned} \sigma_w^2 &= E\{[\mathbf{w}^T \mathbf{z}(x,y)][\mathbf{w}^T \mathbf{z}(x,y)]\} \\ &= \mathbf{w}^T \mathbf{R}_{zz} \mathbf{w}^T - \mu_w^2 \end{aligned} \tag{5.24}$$

where

$$\mathbf{R}_{zz} = E[\mathbf{z}(x,y)\mathbf{z}^T(x,y)] \tag{5.25}$$

The autocorrelation matrix $\mathbf{R}_{zz}$ can be constructed from the autocorrelation function of the image $z(x,y)$. By approximating the inspection image as a separable autoregressive process [63] of order one, the simplified expressions for the variance and its derivative can be developed [73]:

$$\sigma_w^2 \approx 2\psi_w (\mathbf{h}^T \mathbf{R}_{ii} \mathbf{h})^2 \tag{5.26}$$

$$\frac{\partial \sigma_w^2}{\partial \mathbf{h}} \approx 8\psi_w (\mathbf{h}^T \mathbf{R}_{ii} \mathbf{h}) \mathbf{R}_{ii} \mathbf{h} \tag{5.27}$$

where ψ_w is some scalar, such that $\partial \psi_w / \partial \mathbf{h} \approx 0$.

Object Functions for Optimization

The objective of designing optimal filters is to discriminate wear regions from wear-free regions from an inspection image. Thus the obvious aim of an optimal filter is to achieve the maximum separation between average local energies of wear regions μ_{ww} and wear-free regions μ_{wf} at the filter output.

The measure of relative *distance* between the average feature values can be used for the optimization of a single filter with respect to the discrimination between two textures. Unser [88] proposed the object function

$$J_U(\mathbf{h}) = \frac{(\mu_{ww} - \mu_{wf})^2}{\mu_{ww}\mu_{wf}} \tag{5.28}$$

for the design of optimal texture transforms. Randen and Husøy [73] have adapted this object function for the designing of optimal filters for texture segmentation. Optimization with respect to $J_U(\mathbf{h})$ is easily developed and entails the solution of

$$\frac{\partial J_U(\mathbf{h})}{\partial \mathbf{h}} = 0 \tag{5.29}$$

The chain rule for differentiation yields

$$\frac{\partial J_U(\mathbf{h})}{\partial \mu_{ww}} \frac{\partial \mu_{ww}}{\partial \mathbf{h}} + \frac{\partial J_U(\mathbf{h})}{\partial \mu_{wf}} \frac{\partial \mu_{wf}}{\partial \mathbf{h}} = 0 \tag{5.30}$$

Combining (5.21) (5.22) and (5.30) we get

$$\mathbf{R}_{ii_f}^{-1} \mathbf{R}_{ii_w} \mathbf{h} = \lambda \cdot \mathbf{h} \tag{5.31}$$

where

$$\lambda = \frac{\mu_{ww}}{\mu_{wf}} = \frac{\mathbf{h}^T \mathbf{R}_{ii_w} \mathbf{h}}{\mathbf{h}^T \mathbf{R}_{ii_f} \mathbf{h}} \tag{5.32}$$

The (5.31) is an eigenvalue equation where the filter $\mathbf{h}$ is the eigenvector and λ is the eigenvalue. The optimal filter coefficient vector is found as the eigenvector of $\mathbf{R}_{ii_f}^{-1} \mathbf{R}_{ii_w}$ with the maximum corresponding object function value $J_U(\mathbf{h})$. The proof can be found in [73].

The main disadvantage of the object function $J_U(\mathbf{h})$ is that the feature variances, σ_{ww} and σ_{wf}, are not taken into account. Consequently, the optimal filters designed with respect to this object function can only achieve maximally large separation of feature means μ_{ww} and μ_{wf}. If the variances of local energy estimate, σ_{ww} and σ_{wf}, are large, then the feature distribution may considerably overlap. Therefore, a good object function should not only yield a large distance between

the mean values of features, but also yield low feature variances. An object function well known in pattern recognition literature that takes the feature variance into account is the Fisher criterion [21]:

$$J_F(\mathbf{h}) = \frac{(\mu_{ww} - \mu_{wf})^2}{\sigma_{ww}^2 + \sigma_{wf}^2} \tag{5.33}$$

As in the previous case, the optimal filter that maximizes the object function $J_F(\mathbf{h})$ can be found by equating the partial derivative of $J_F(\mathbf{h})$ to zero, i.e.

$$\frac{\partial J_F(\mathbf{h})}{\partial \mathbf{h}} = 0 \tag{5.34}$$

Randen and Husøy [73] have developed a *closed form* solution for the above equation using the approximate expressions of feature variance and its derivative. By substituting the expressions for mean and variance and their corresponding derivatives into (5.34), we get

$$\begin{aligned}&(\psi_{wf}(\mathbf{h}^T\mathbf{R}_{ii_f}\mathbf{h})^2 + \psi_{wd}(\mathbf{h}^T\mathbf{R}_{ii_w}\mathbf{h}) \cdot (\mathbf{h}^T\mathbf{R}_{ii_f}\mathbf{h}))\mathbf{R}_{ii_w}\mathbf{h} \\ &= (\psi_{ww}(\mathbf{h}^T\mathbf{R}_{ii_w}\mathbf{h})^2 + \psi_{wf}(\mathbf{h}^T\mathbf{R}_{ii_w}\mathbf{h}) \cdot (\mathbf{h}^T\mathbf{R}_{ii_f}\mathbf{h}))\mathbf{R}_{ii_f}\mathbf{h}\end{aligned}$$

which can be further simplified as

$$\mathbf{R}_{ii_f}^{-1}\mathbf{R}_{ii_w}\mathbf{h} = \frac{\mathbf{h}^T\mathbf{R}_{ii_w}\mathbf{h}}{\mathbf{h}^T\mathbf{R}_{ii_f}\mathbf{h}} \cdot \mathbf{h}$$

or

$$\mathbf{R}_{ii_f}^{-1}\mathbf{R}_{ii_w}\mathbf{h} = \lambda \cdot \mathbf{h} \tag{5.35}$$

Equation (5.35) is identical to the eigenvalue problem discussed in the previous approach, i.e. (5.32). Thus the coefficients of the optimal filter can be obtained from the eigenvectors of $\mathbf{R}_{ii_f}^{-1}\mathbf{R}_{ii_w}$ which generates the maximum object function $J_F(\mathbf{h})$.

Interestingly optimizing with respect to two different object functions leads to the same eigenvalue problem, different eigenvectors, however, that maximize different object functions may be selected as the optimal filters. It is possible that one eigenvector maximizes both object functions so that only one optimal filter is obtained (this will be demonstrated by Figs. 5.14 (a), (b) and (c) later in experiments).

To sum up, the optimal filters corresponding to two object functions suggested by Unser [88] and Fisher [21] are computed as follows:

1. Computing the correlation matrices $\mathbf{R}_{ii_f}$ and $\mathbf{R}_{ii_w}$;

2. Computing the eigenvectors of $\mathbf{R}_{ii_f}^{-1}\mathbf{R}_{ii_w}$;

3. Selecting the eigenvector yielding maximum object function $J_U(\mathbf{h})$ or $J_F(\mathbf{h})$;

4. Inverse lexicographical reordering the elements of $\mathbf{h}$ to obtain the optimal filter $h(x, y)$.

The detailed procedure for designing optimal filters is also illustrated in Fig. 5.9.

5.4.2 Supervised Wear Detection

As stated in Section 5.3.2, if the priori knowledge of wear patterns to be detected is available, such detection of known wear patterns can be regarded as supervised wear detection. From some known category of wear samples, specific optimal filters can be designed to detect it. First of all, the selection of parameters for designing these optimal filters is discussed in this subsection.

Size of the Optimal Filter

The dimensions of an FIR filter are related to its bandwidth. Filters for large bandwidth require smaller dimension and vice versa [51]. Appropriate dimensions of optimal filter can be determined from the spectral characteristics of the surface image. The spectral characteristics of wear can be entirely different from its wear-free background. Therefore, as will be seen in Subsection 5.4.3, the size of an optimal filter depends on the spectral characteristics of wear patterns to be detected. A symmetric region of support is required for accurate edge localization and therefore, only odd size filter masks are designed.

Selection of a Smoothing Filter

An important element in the design of an optimal filter is the choice of a smoothing filter. Among several candidate filters, Gaussian low pass smoothing filter is commonly used [73, 47] since it is separable and yields optimal joint resolution in spatial and frequency domain:

$$\mathcal{G}(x, y) = \frac{1}{\sqrt{2\pi}\sigma_g} \exp(-\frac{x^2 + y^2}{2\sigma_g^2}) \tag{5.36}$$

The choice of bandwidth σ_g determines the frequencies to be included for local energy estimation at the output. Finite approximation of the above filter (5.36)

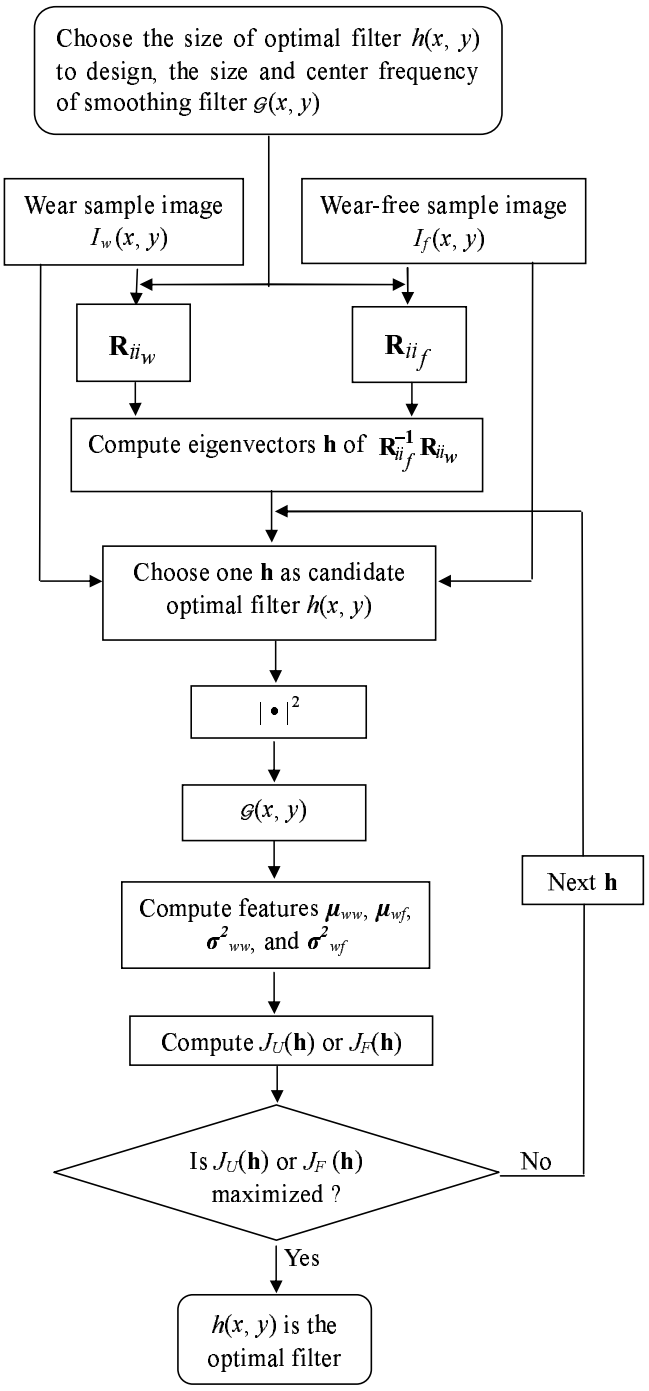


Figure 5.9: Procedure of designing optimal filters.

implemented as separable convolution masks are used as smoothing filter. Jain and Farrokhnia [36], and Kumar and Pang [47] suggested the following choice

$$\sigma_g = \frac{1}{2\sqrt{2}u_0} \tag{5.37}$$

where the center frequency u_0 is the estimation of the highest frequency of the background texture.

Calibration

Once appropriate parameters are selected, optimized filters with respect to particular object functions can be designed according to the procedure stated in Subsection 5.4.1 (see also Fig. 5.9). The designing of optimal filters is usually done at the very beginning, i.e. the calibration phase, hence it does not add any computational burden to the wear detection phase, suitable for online implementation.

For calibration, two patches of images from one particular known category of wear and wear-free samples, respectively, are applied. They are of zero mean[2] and same size. After the optimal filters are designed, they are applied to detecting wear with similar patterns as those appearing in the wear samples by filtering (see Fig. 5.8). Finally, a thresholding operation (Eq. (5.13)) as stated in Subsection 5.3.2 is applied to the filtered image (feature image) to classify wear regions. In the next subsection, experimental results of this supervised approach to wear detection will be presented.

5.4.3 Experiment and Results

The supervised wear detection method using optimized filters was validated by a series of experiments. The experimental setup in Chapter 3 has been applied. Same images of the wear samples with 512×512 pixels applied in the previous approach were also tested in this experiment. In the calibration phase, only two image patches of same size (32×32 or 64×64) selected from wear-free and known category of wear samples respectively were used to design the optimal filters with respect to two object functions $J_U(\mathbf{h})$ and $J_F(\mathbf{h})$. If the complete 512×512 pixels image is utilized for designing the optimal filter, the discriminating effect of wear from its large wear-free background diminishes due to the inherent averaging that takes place while computing its correlation matrix $\mathbf{R}_{ii_w}$. Furthermore, the computational time for computing the correlation matrices, $\mathbf{R}_{ii_w}$ and $\mathbf{R}_{ii_f}$, for the complete image is significantly high. Therefore, only a small image patch from the

[2]To ensure average gray level insensitivity.

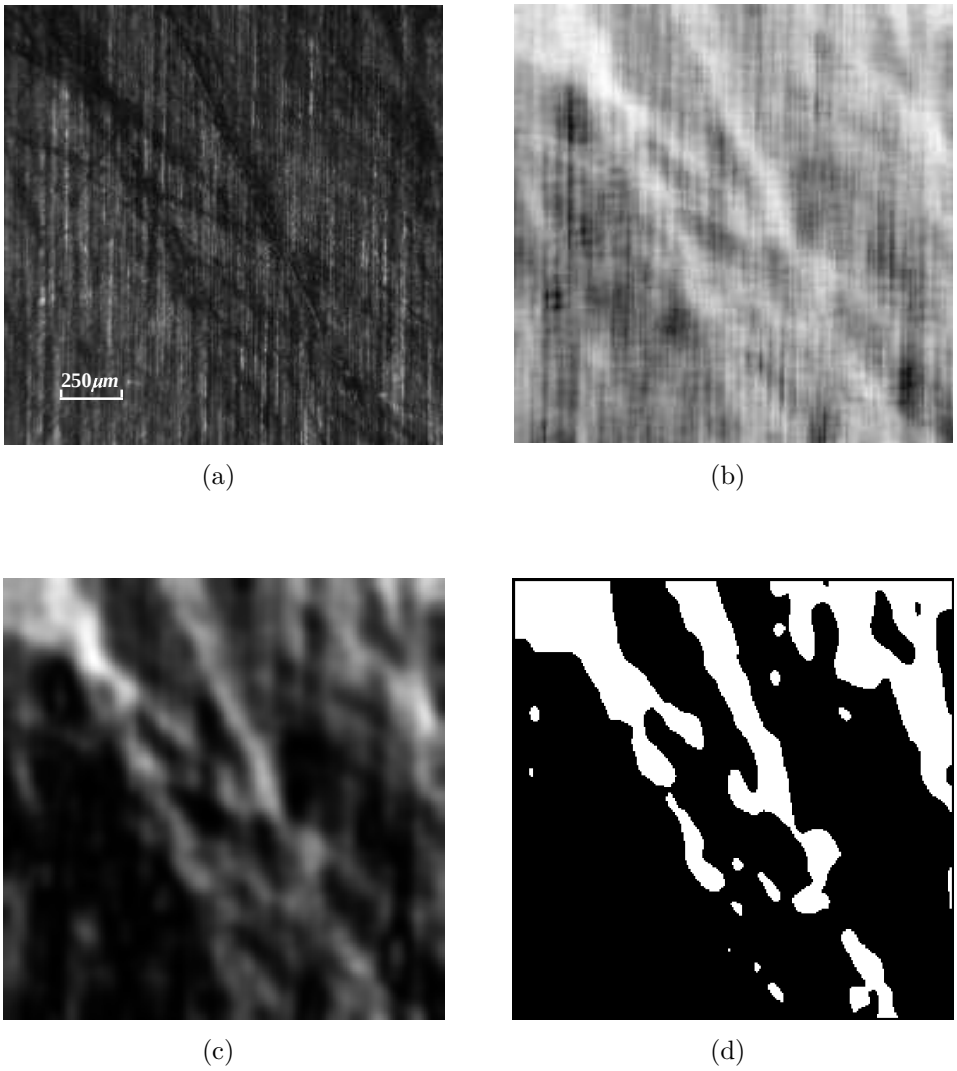
region of image having known category of wear (and equal sized image patch from a wear-free image) is utilized for designing optimal filters. The size of this image patch is empirically determined and it depends on the spatial extent of wear in an image.

In all the experiments, the analyzed images were made to have zero mean value. A Gaussian low-pass filter as discussed above was applied for smoothing. For all the experiments reported in this section, the size of the smoothing filter is empirically fixed as 9×9, unless otherwise explicitly stated.

Fig. 5.10 shows one of the wear detection results. In this experiment an optimal filter of size 9×9 with respect to object function $J_U(\mathbf{h})$ was designed to detect wear. The object function (5.28) for each of the 81 eigenvectors is shown in Fig. 5.11. The eigenvector that maximizes $J_U(\mathbf{h})$ is picked up as the optimal filter. Fig. 5.13 shows its magnitude frequency response. It can be seen that the magnitude frequency response exhibits passbands where local energy estimate is high (corresponding to wear) and stopbands elsewhere. Fig. 5.10(a) is the image of the wear sample, a ground steel with wear scratches, whose magnitude frequency response is shown in Fig. 5.13. Fig. 5.10(b) shows the image after filtering with the optimal filter. The local energy estimate is shown in Fig. 5.10(c). As seen from this image, the average local energy for the region corresponding to wear is much greater than that of wear-free region, so wear can be easily be segmented by simple thresholding introduced in Subsection 5.3.2 (see Fig. 5.10(d)).

More experiments for difference wear samples (also used in the previous method) were conducted to test optimal filters designed with these two object functions $J_U(\mathbf{h})$ and $J_F(\mathbf{h})$. The results are shown in Fig. 5.14. The size of the optimal filters is 7×7. Figs. 5.14(a), (d) and (g) are sample images with wear track or scratches. The corresponding local energy estimates with the optimal filters designed using object function $J_U(\mathbf{h})$ are shown in Figs. 5.14(b), (e) and (h); and object function $J_F(\mathbf{h})$ in Figs. 5.14(c), (f) and (i). It can be seen that the optimal filters designed in these experiments were robust. All wear patterns in these images were successfully detected no matter whether wear track or wear scratches. For the wear track in Fig. 5.14(a) and wear scratches in Fig. 5.14(g), the optimal filters designed are the same for $J_U(\mathbf{h})$ and $J_F(\mathbf{h})$. However this is not always the case. Different optimal filters can be obtained for different object functions. For instance, a slight difference can be seen in Fig. 5.14(e) and (f) due to different optimal filters applied. More examples will be shown in Section 5.5.

Fig. 5.15 shows the detection results for the image in Fig. 5.10(a), using the optimal filter designed with respect to object function $J_U(\mathbf{h})$. It is found that optimal filters of larger size deliver better results, i.e. less false alarms and better detection. However, larger size also means more computational load. Hence, in practice, a trade-off between better results and faster speed must be made. For



(a) (b)

(c) (d)

Figure 5.10: Results of supervised wear detection: (a) image under detection, (b) after filtering (a) with 9×9 optimal filter, (c) local energy estimate of image in (b), (d) segmented wear regions after thresholding energy image in (c).

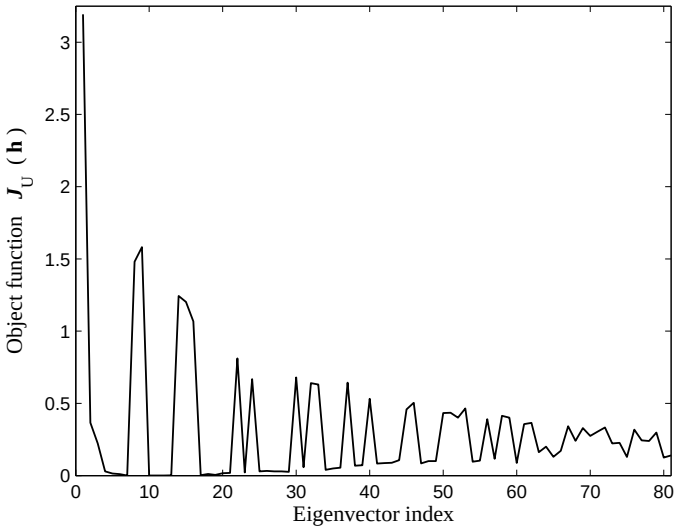


Figure 5.11: Object function $J_U(\mathbf{h})$ for each of the 81 eigenvectors $\mathbf{h}$.

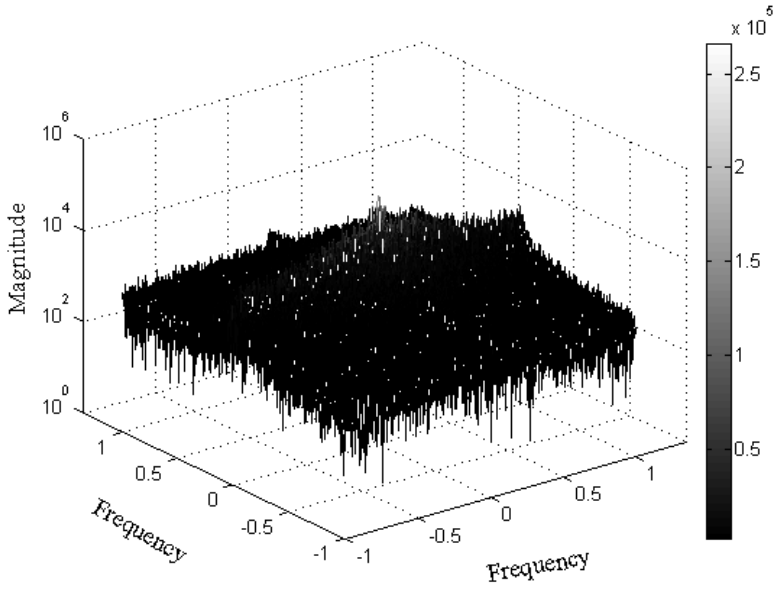


Figure 5.12: Amplitude frequency response of the image shown in Fig. 5.10(a).

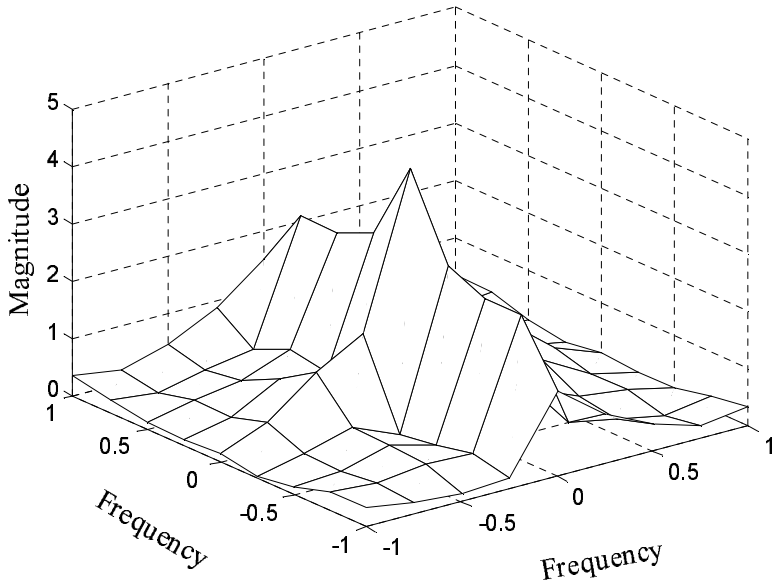


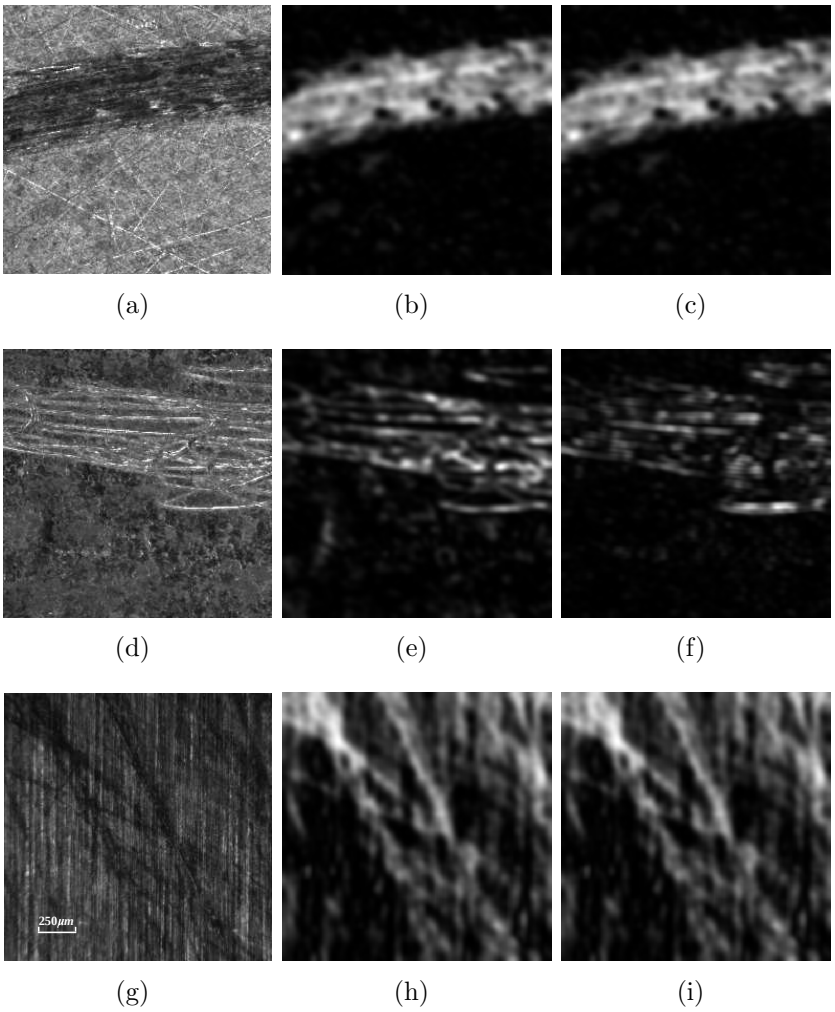
Figure 5.13: Amplitude frequency response of 9×9 optimal filter.

this example, a filter of size 7×7 (see Fig. 5.15(b)) can give good results with marginal compromise in performance.

5.5 Discussions

The lack of an appropriate criterion for wear detection makes it very difficult to compare the performance of a wear detection technique on various types of wear. A commonly used criterion for the quantification of image segmentation results is the percentage of misclassified pixels. However, the class of the pixel in an image does not necessarily reflect the class of the corresponding point on the real (physical) surface. For this reason, we only discuss the performance of wear detection techniques qualitatively, by looking at and comparing the detection results on images.

As discussed in Subsection 5.4.1, optimal filters designed with respect to the object function $J_F(\mathbf{h})$ should deliver better results than those designed with respect to $J_U(\mathbf{h})$ since the former takes variance into consideration. But this also makes the designing of optimal filters more complicated as it is very difficult to find a closed-form solution if no approximation is being made. In our experiments we found that for larger sized optimal filters designed with $J_F(\mathbf{h})$ may fail to detect any wear (see Fig. 5.16(a)). This implies that having more free parameters to

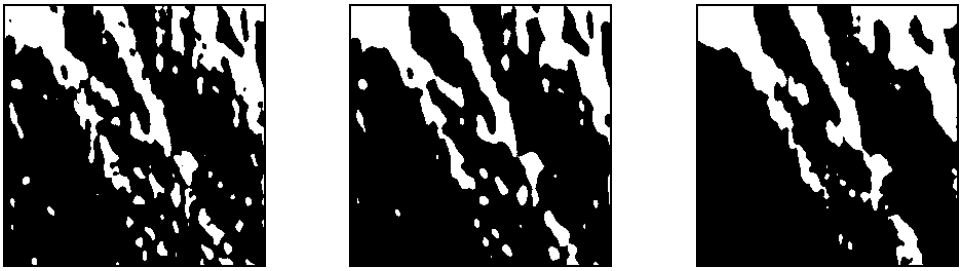


(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

Figure 5.14: Images of wear samples under detection in (a), (d) and (g) respectively; corresponding local energy estimates with the optimal filters of size 7×7, designed using object function $J_U(\mathbf{h})$ in (b), (e) and (h); object function $J_F(\mathbf{h})$ in (c), (f) and (i).



(a) (b) (c)

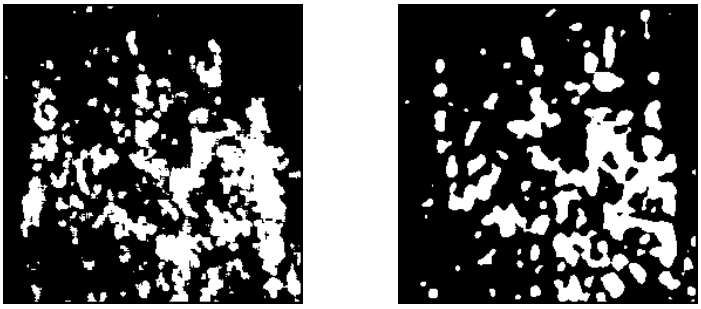
Figure 5.15: Detection results for the image in Fig. 5.10(a), using the optimal filter of size (a) 3×3, (b) 7×7 and (c) 13×13, designed with object function $J_U(\mathbf{h})$.

optimize for an FIR filter does not have to be always an advantage. Optimization might be cumbersome and errors may appear because of a non-optimal step or setup. For instance, when the size of the smoothing filter becomes larger, optimal filters designed with $J_F(\mathbf{h})$ also get easier to fail (see Fig. 5.16(b)).

By comparing the results (see Fig. 5.7, Fig. 5.10, Fig. 5.14 and Fig. 5.15) of these two different filtering approaches for wear detection, i.e. unsupervised (Gabor filters) and supervised (Optimized filters), we can see that the multichannel Gabor filtering approach provides better results, for instance, better detection and less false alarm for same filter sizes. This is because multichannel Gabor filtering covers all the scales and orientations of wear patterns whereas optimal filters rely much on the resembling degree of the wear under detection and the representative wear sample for calibration. But the latter approach is more computational efficient than the former one since only one optimal filter instead of a bank of filters is used.

5.6 Conclusions

Wear detection on machined surfaces can be regarded as a texture segmentation problem. This chapter develops two filtering approaches to online wear detection, the unsupervised detection scheme using multichannel Gabor filters and the supervised detection scheme using optimized filters. Both approaches can successfully detect wear patterns present on textured surfaces, machined surfaces, for instance. Experiments conducted on various real wear samples confirmed the usefulness of these two approaches.



(a) (b)

Figure 5.16: Detection results for the image in Fig. 5.10(a) using the optimal filter designed with respect to the object function $J_F(\mathbf{h})$, where optimal filter size: (a)13×13 , (b) 7×7; smoothing filter size: (a) 7×7, (b) 13×13.

For both approaches, filter size affects the performance of wear detection. Trade-offs between speed and performance must be made for realtime applications. The unsupervised scheme using multichannel Gabor filters needs little prior knowledge of wear and was found to offer better detection results. Compared to Gabor filters, the supervised scheme using optimized filters obtained successful results with less computationally complex filters. However it needs better insight, representative wear samples for instance, in the design (calibration) phase.

Chapter 6

Wear Monitoring

The dynamics of wear processes are studied in this chapter by analyzing image sequences of a wearing surface. Wear behavior of dry sliding wear in particular, under different operating conditions, such as varying normal load, different sliding speeds, is investigated. Fractal values, namely fractal dimension and intercept, computed from the power spectra of images of a wearing surface, are adopted as indicators of the dynamic wear process. Experimental results show that progressive changes of fractal values might reveal the wear status of the surface.

6.1 Introduction

Image processing techniques for online wear detection were presented in the previous chapter. These techniques can be used in inspection-oriented applications: such as engineering surface inspection, coating failure detection, tool wear monitoring, etc. Wear behavior and dynamics of wear processes, on the other hand, are of interest in tribological research as well as in engineering applications.

With the existing knowledge of materials science and mechanical tools (finite elements techniques) engineers are able to predict the functional behavior of a component of a product or process with respect to strength, fatigue etc. with reasonable accuracy. However, the lifetime of a component depends on wear. Due to the complexity of a wear process, measurement of wear is traditionally conducted offline [12, 83]. Measuring wear using existing techniques (refer to Chapter 2) has intrinsic restrictions: First, measurement takes place offline, i.e. during the experiment, the specimen must be removed from the tester periodically to measure the evolution of wear as a function of time or sliding distance. It is very difficult to re-measure at the same location on the specimen before and after the wear test. Second, the mounting and dismounting steps unavoidably bring in errors. In tribological research as well in engineering applications, reliable online measurement or monitoring of a dynamic wear process is demanded and remains a challenge [102].

Wear as a type of surface damage is a 3D phenomenon. In Chapter 2 a variety of wear measures have been reviewed, among which volume is the fundamental measure for wear when wear is equated with loss or displacement of material. In engineering applications, however, the concern is generally with the loss of a dimension, or the presence of a wear scar, not a volume per se. These changes and the volume loss are related to each other through the geometry of the wear scar and therefore can be correlated in a given situation. One important aspect to recognize is that the relationship between wear volume and wear dimensions, for instance, depth, width, or area, is not necessarily a linear one. Using the imaging system described in Chapter 3, some 1D and 2D wear measures such as wear width, wear area, and surface roughness can be obtained.

In this chapter we will study some wear behavior and dynamics of wear processes by analyzing these measures estimated from image sequences of a wearing surface. In particular, online monitoring of dry sliding wear under varying conditions, such as different normal forces, different velocities, is investigated. Wear curves, for instance, wear area versus sliding cycles, are drawn and checked with Archard's law. Besides, to estimate the degree of wear and further predict the remaining life of a component, some indicators of state of wear are proposed and validated.

This chapter is organized as follows: Section 6.2 presents necessary preliminary operations of sequence of images for subsequent image analysis; Section 6.3 gives some observations on wear behavior, in which Subsection 6.3.1 studies the evolution of wear area against rotation cycles, followed by experiments in 6.3.2, result analysis in 6.3.3 and conclusions in 6.3.4 respectively; estimation of the degree of wear is described in Section 6.4; fractal analysis of dry sliding wear is addressed in Section 6.5, in which experimental results and analysis are given in Subsection 6.5.1, followed by discussions in 6.5.2 and conclusions in 6.5.3; finally, summary to this chapter is provided in Section 6.6.

6.2 Image Sequence

To study the dynamics of a wear process, a temporal series of images is taken using the imaging system introduced in Chapter 3. Each image is characterized by a set of data (pixels in grey-value). The wear process is described by the gradual changes in this data set.

To monitor the dynamic wear process by tracking the gradual changes in the image data set, it is necessary to have all the images captured from the same surface area. This requires image acquisition and surface motion be synchronized. Besides, after synchronization, due to repetitive positioning errors of the rotary table, thermal expansion of the aluminum mounts that hold the camera and the specimen, as well as vibration, there still exists pixel shift i.e. misalignment, between the initial and successive images. Accordingly, these images need to be further registered. Synchronization and image registration have been described in Chapter 3.

After the above mentioned preliminary operations, wear behavior of dry sliding wear can be investigated.

6.3 Observation of Wear Behavior

In most cases, wear occurs through surface interactions at asperities. During relative motion, first, the softer surface may be plastically deformed, or material on the contacting surface may be displaced so that the properties of the solid body, at least at or near the surface, are altered, but little or no material is actually lost. Later, material may be removed from a surface and may result in the transfer to the mating surface or may break loose as a wear particle. Hence, a wear process can be divided into two phases: initial phase and material transfer or particle-forming phase. In the initial phase, there is no net volume or mass loss of the interface, although one of the surfaces is worn (with a net volume or mass loss).

The initial phase of a wear process is studied by observing and comparing gradual changes in the sequence of images of a wearing surface through the imaging system with well-controlled, uniformly distributed ring light illumination.

During a wear process, the micro-geometry of the asperities on a surface will change, which results in the change of the surface topography. The microscopic waveform of a surface profile modulates the incident light beams into scattered beams whose intensities and scattering angles can be described as functions of the incident angles, wavelengths and slope roughness of the surface topography (refer to Chapter 4). Accordingly, the dynamic wear process can be described by the image changes of a same surface area.

An intuitive way to compare and detect changes between images is simply subtracting one from another, if the images are in alignment and acquired under uniform illumination. From the difference images, the wear area can be identified simply by thresholding, i.e. if the pixels in the images satify

$$|I_n - I_r| > T, \tag{6.1}$$

where I_n denotes the sequence of images, I_r the reference image without wear, and T the empirically determined threshold, then these pixels make up the wear area. After thresholding, binary images are obtained from which the wear area can be easily computed by counting pixels. The procedure of identification of the wear area from image sequences is shown in Fig. 6.1.

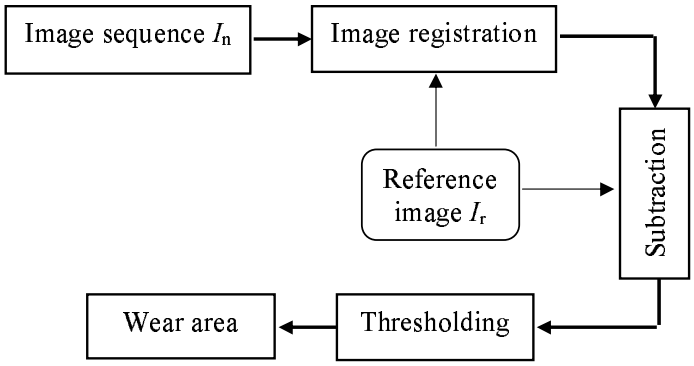


Figure 6.1: Identification of wear area by thresholding

Note that the wear area identified in a reflection intensity image is equal to the physical wear area in a real surface times the magnification factor of the imaging system.

After identifying wear areas in the sequence of images, we may study the behavior of the initial phase of a wear process. For instance, investigation of how the wear area changes against rotation cycles.

6.3.1 Wear Area Versus Rotation Cycles

To simplify the study, the Pin-on-Disk (refer to Chapter 3) dry sliding case is considered. If the tip of the pin has a spherical shape and it is made of much harder material than the disk, the wear of the tip is very small and can be neglected. With a certain load, the pin slides on the disk surface when the disk rotates, thus a circular wear track is generated on the disk. Next let us derive the relationship between the wear area A of the wear track and the number of the rotation cycles n.

Assume that the surface of the disk is ideally flat and the pin is loaded sufficiently to cause the disk (made of softer material) deform. Further assume that in each rotation the pin follows exactly the same track. As such, on the disk a circular wear track is generated as a result of the wear. The geometrical relationship is shown in Fig. 6.2, where the wear track is of width d and maximum depth y. If the radius of the spherically shaped pin is R, then for $y/R \ll 1$,

$$y \approx \frac{d^2}{8R}. \tag{6.2}$$

And the wear volume is:

$$V \propto \frac{1}{2} y d = \frac{d^3}{16R}. \tag{6.3}$$

The wear area A is proportional to width of the wear track d, i.e.

$$A \propto d. \tag{6.4}$$

According to Achard's equation, V is proportional to the sliding distance s, which is equivalent to the number of rotation cycles n, i.e.

$$V \propto n. \tag{6.5}$$

Hence, by comparing Eqs. (6.3), (6.4) and (6.5), we get

$$A \propto n^{1/3} \tag{6.6}$$

Likewise, according to Archard's equation, wear volume V is proportional to normal load F, therefore wear area should be proportional to $F^{1/3}$.

6.3.2 Experiment

To validate the above model, some experiments of dry sliding wear are conducted using the experimental setup stated in Chapter 3. In these experiments, a spherically shaped indenter (ball bearing steel 52100) of diameter 3 mm is sliding, under

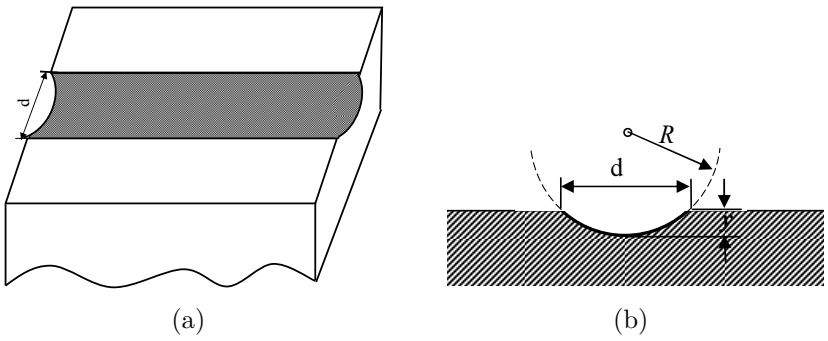


Figure 6.2: Geometrical relationship of the wear track.

different normal loads and varied rotation speeds, against a machined aluminum flat disk with average RMS roughness $R_q \approx 0.7\ \mu m$ and peak to valley 6 μm. The operating conditions are listed in table 6.1. The online wear behavior of the disk surface is of our interest.

Table 6.1: Operating conditions of wear experiment

Material	Indenter: ball bearing steel Disk: aluminium
Normal load	1 N, 5 N
Pin shape and size	Sphere with diameter 3 mm
Hertzian radii	27.1 μm, 46.3 μm
Rotation speed (linear velocity)	0.5 rpm, 1 rpm (2.5 mm/s, 5.0 mm/s)
Lubrication	None
Temperature	20°C
Image size	640 × 480 (pixels)

Some measures must be taken to guarantee the correctness of the experiment. First, the microscope is positioned through the linear translation stage so that the camera view covers the whole wear track. In practice the wear track is positioned in the middle of the image; second, in order to study the progressive changes of wear area for a same part of the surface, the rotation of the disk and image acquisition is synchronized. One reference (initial) image (without wear) is first acquired before loading the pin and starting the rotation and then the pin is loaded and one image is captured after each rotation. All the images in the sequence are registered with respect to the reference image. In this way the images from the

same area of the wearing surface are obtained. Then identification of the wear area is done as described earlier in this section (see also Fig. 6.1).

One experimental result of wear area identification is shown in Fig. 6.3 where the applied load is 5 N and the rotation speed is 0.5 rpm. The threshold T is automatically determined as the maximum value of a wear free window area in a difference image.

As stated in Chapter 2, most typical junctions for normal contact situations have diameters of the order of 10 microns. In the current case, when the load is 5 N, the calculated Hertzian radius is 46.3 μm. The wear track width after 1 cycle of rotation is about 60 μm, which is a few times larger than the normal typical junction size.

Microscopically scrutinizing the wear track, one can see that it is highly merged from different grooves generated at different rotations. From several independent Pin-on-Disk wear tests it is found that only one wear track is generated (each cycle of rotation follows the same track). There is, however, some randomness due to eccentricity of the rotary table as well as the difference of tensile stress of the pin holder and changing friction force between the pin and the table surface. In the example shown in Fig. 6.3, the average width of the wear track after 15 cycles of rotation is about 200 μm.

The experimental results of wear area versus rotation cycles for different loads and speeds are shown in Fig. 6.4. In these experiments, all the images are acquired under the same illumination (same intensity level) and the thresholds are chosen in the same way as mentioned earlier.

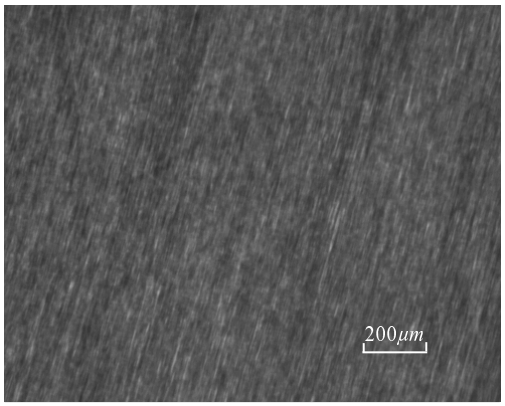
6.3.3 Result analysis

Fig. 6.4 (a) shows wear area against rotation cycles for the same rotation speed 0.5 rpm (also same linear speed) and different loads 1 N and 5 N. Fig . 6.4 (b) illustrates the result for the same load 5 N and different rotation speeds 0.5 rpm and 1 rpm. The results are least squares fitted to the curve

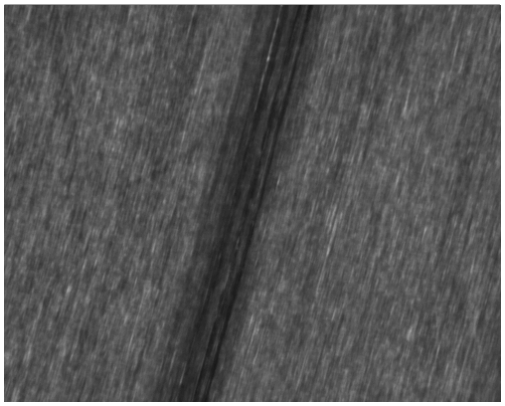
$$A = k \cdot n^{1/3} + b, \tag{6.7}$$

where k and b are two free parameters of the curve. The curve fitting is only done for the initial phase of a wear process. In the particle-forming phase, the presence or removal of large particles on or from the images leads to large deviations of the identified wear area.

From Fig. 6.4 (a) and (b) we can see that in the initial phase of wear processes, the wear area is approximately proportional to $n^{1/3}$, where n is the number of rotation cycles.



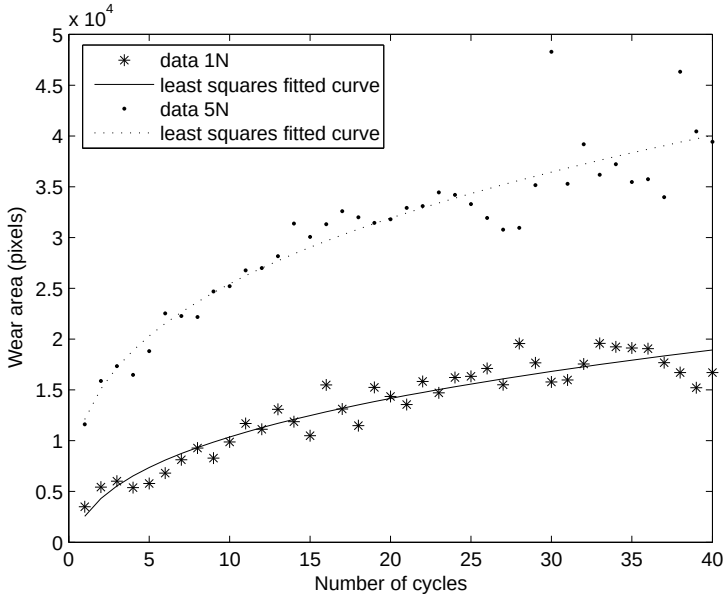
(a) Initial (reference) image



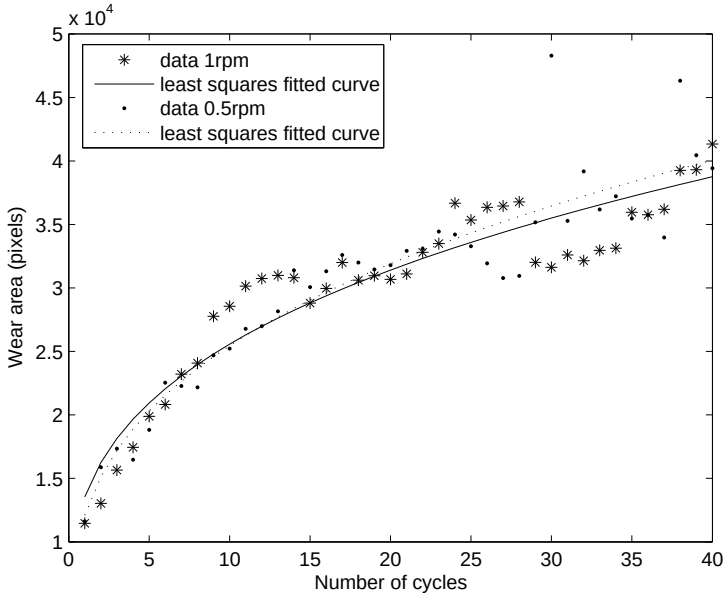
(b) Image after 15 cycles of rotation

(c) Binary image after thresholding the difference image between (a) and (b)

Figure 6.3: An experimental result of wear area identification.



(a) Same rotation speed (0.5 rpm), different loads



(b) Same load (5 N), different rotation speeds

Figure 6.4: Wear area versus rotation cycles.

From Fig. 6.4 (a) we do not observe that the wear area is proportional to $F^{1/3}$ for these two forces in the initial phase of a wear process. Due to the shortage of the test samples, we are not able to try more loads to achieve a statistical result. But, generally we can see that a larger load causes larger wear.

From Fig. 6.4 (b) we can see that the rotation speed (low) does not have influence on the wear area. This confirms Eqs. (2.1) and (2.2), i.e. wear rate is independent of the sliding velocity. However, from the experiments we can see that larger load makes a wear process finish the initial phase faster.

6.3.4 Conclusion

From the experimental results and above analysis we can conclude that the wear area identified by the simple thresholding method in an image sequence might be a proper measure for a dry sliding wear process. And for the initial phase of a wear process, wear area measured in this way may be proportional to (number of cycles)$^{1/3}$, which is consistent with Archard's equation.

6.4 Estimation of Degree of Wear

In the previous section, we discussed a quantitative measure, wear area for studying dry sliding wear behavior. Another interesting research point of online wear monitoring is to estimate the degree of wear which indicates remaining life of a product.

As discussed in the previous chapter, for some applications, coating failure detection for instance, it could be sufficient to obtain a qualitative answer like "a type of wear occurs on the bottom right of a surface". The "what" and "where" questions can thus cover the entire range from "there is something". Furthermore, except extracting early-warning information from the detection of wear, it may be more useful to monitor wear dynamically at multiple levels, i.e. determining wear state instead of a binary decision of "there is wear" versus "no wear" because it gives an indication of the remaining lifetime of a product or a machine element. Representation in state of wear allows for lower variance, less confusion, and might help characterize a specific wear process as well. Therefore, it is necessary to define some indicators of state of wear. Some quantitative wear measures, for example, wear scar width or area discussed earlier, loose applicability for estimation of the degree of wear as they are not suitable for characterizing either the randomness of surfaces or the dynamic nature of a wear process. As for surface roughness measures, whether RMS roughness or RMS slope, ignores any spatial correlation and structural information within the surface. This profoundly limits their capability

to characterize surfaces as well as dynamics of wear processes.

Generally speaking, in a dynamic case statistical measures work better. Fractal values as surface morphological measures might be good alternatives. In this section, indicators of state of wear will be investigated using fractal image analysis. Next, we give a brief introduction to fractals.

6.4.1 Fractal Geometry

Conventionally point, line, surface and volume have topological dimensions of 0, 1, 2 and 3 respectively. But many real-world lines and surfaces have a roughness that reaches out into a higher dimension, and expressing that as a fractional value may be understood as a qualitative measure of the roughness. It is interesting that much of the natural world seems to obey fractal geometry, rather than Euclidean geometry.

Fractals were first introduced by Mandelbrot [53] and have been being rapidly assimilated into many diverse fields of physics and mathematics. Fractal geometry is the geometry of the broken up, the pitted and pocked, the tangled and twisted, the turbulent and the chaotic [87]. The central theme of fractal geometry is that nature, although seemingly complex, exhibits a fundamental property generally known as *self-similarity*. In other words, however complex the shape and/or dynamic behavior of a system, its features at one scale resemble those at other scales. Consider a bounded set S in Euclidean n-space. The set S is said to be self-similar if S is in the union of M distinct (non-overlapping) copies of itself, each of which has been scaled down by a ratio $\xi < 1$ in all coordinates. The *fractal dimension* (FD) D, also known as *Hausdorff dimension*, is given by

$$M\xi^D = 1 \text{ where } D = -\frac{\ln M}{\ln \xi} \tag{6.8}$$

The ranges in the value of D characterize the type of fractal (see Table 6.2).

Table 6.2: Fractal values of typical objects

Fractal Dimension	Fractal Type
$0 < D < 1$	Fractal Dust
$1 < D < 2$	Fractal Signal
$2 < D < 3$	Fractal Surface or Image
$3 < D < 4$	Fractal Volume

Natural fractals do not in general possess such deterministic self-similarity. Most of them exhibit statistical self-similarity. Such fractal sets are composed of

M distinct subsets each of which is scaled down by a ratio $\xi < 1$ from the original and is identical in all statistical respects to the scaled original. Furthermore, the scaling ratios need not to be the same for all the scaled down copies. Certain fractal sets are composed of the union of M distinct subsets each of which is scaled down by a ratio $\xi_i < 1$, $1 \leq i \leq M$ from the original in all coordinates. The FD is given by a generalization of Eq. (6.8), namely

$$\sum_{i=1}^{M} \xi_i^D = 1 \tag{6.9}$$

Finally, there are *self-affine* fractal sets which are scaled by different ratios in the different coordinates. For example, consider a curve $f(x)$ which satisfies $f(\xi x) = \xi^\kappa f(x)$, $\forall\ \xi > 0$, where ξ is a scaling factor and κ is the scaling exponent. In other words this implies that the curve is self-affine if a scaling of the x-coordinate by ξ gives a scaling of the f-coordinate by a factor ξ^κ. To summarize, the definition of a fractal is reprinted, according to Mandelbrot, who originally coined the term in 1982. *A fractal is by definition a set for which the Hausdorff dimension strictly exceeds the topological dimension, and a fractal is a shape made up of parts similar to the whole in some way.* In this study, we mainly focus on fractal surfaces whose fractal dimensions are larger than 2.

6.4.2 Fractal Image Analysis

The application of fractal geometry to imaging science and computer vision is now receiving increasing attention. Fractal analysis has been used to describe complex physical phenomena such as turbulence, brittle fracture of materials, machining and tool wear [43, 80]. Fractal characterizations have also been used to describe complex two or three-dimensional surfaces, such as deposited surfaces [80], wear-erosion surfaces [74, 76], wear particles [65, 66, 67, 68, 69]. To our knowledge, the evolvement of dry sliding wear on a machined surface has not been investigated by online analyzing its images using fractal analysis techniques. In the following, we will use fractal analysis to assign numerical values to indicate the degree of wear for a wearing surface under dry sliding situation.

Fractal dimension is the most important parameter in fractal analysis. Various methods have been proposed to estimate fractal values [80]: Fourier, Kolmogorov, Korcak, Minkowski, root mean square, slit island, etc. These methods differ in computational efficiency, numerical precision and estimation boundary. However, a strong correlation was reported between the relative ranking of fractal values obtained from different fractal measuring techniques [74, 80]. Among these methods the Fourier analytical technique is the most promising one in that it has several

advantages: first, this technique provides fractal information as a function of surface orientation so that it reveals and can characterize anisotropy (an example is shown in Fig. 6.6); second, the Fourier method is relatively insensitive to the presence of the noise in images; third, there exists a fast algorithm, the FFT, which provides efficient implementation particularly for realtime application; fourth, the intercept of the log (power spectrum) versus log (frequency) has also been interpreted and correlated with surface roughness, which will be addressed shortly; fifth, the computation of fractal values is based on explicit formula and potentially more accurate computationally. For these reasons Fourier analysis is adopted to estimate fractal values in this work.

Next this method is addressed in a little more detail. Consider the power spectrum of an ideal one-dimensional fractal signal with fractal dimension D. This has the formula,

$$P(f) = c|f|^{-\beta} \tag{6.10}$$

where c is a constant and β the spectral exponent is related to its fractal dimension D by [87]

$$D = \frac{3D_T + 2 - \beta}{2} \tag{6.11}$$

where D_T is the topological dimension (for a surface, $D_T = 2$). In the following we describe how to use this equation to compute the fractal dimension over a surface or 2D image. The procedure begins by computing the 2D Fourier transform $F(u, v)$ of the spatial surface or image $I(x, y)$ by means of an FFT and then estimating the power spectral density function (PSD)

$$P(u, v) = |F(u, v)|^2, \tag{6.12}$$

where u and v are the spatial frequencies (number of waves per unit wave length) in the x and y directions respectively. Fig. 6.5 shows an anisotropic fractal surface (gray level coded topographical image), whose PSD is illustrated in Fig. 6.6[1]. The radial magnitude lines originating from the center of the PSD combine the information for all profiles in the same corresponding direction. Sampling the PSD along each of these radial lines is equivalent to sampling profiles in different directions from the original surface. The FD is computed from each of the radial lines in a similar fashion, i.e. fitting a least squares error line to the data. In Fig. 6.6 a radial line is selected at an angle θ from the PSD and is redrawn in Fig. 6.7.

The PSD magnitudes at angle θ are selected by converting $P(u, v)$ to the polar coordinate system $P(f)$ such that $f = \sqrt{u^2 + v^2}$, $u = f\cos\theta$ and $v = f\sin\theta$ where f denotes radial frequency. The PSD in the direction θ is the collection of magnitudes at those coordinates. A least squares error line is fitted to the

[1]It is actually the logarithm plot of the PSD for better visualization.

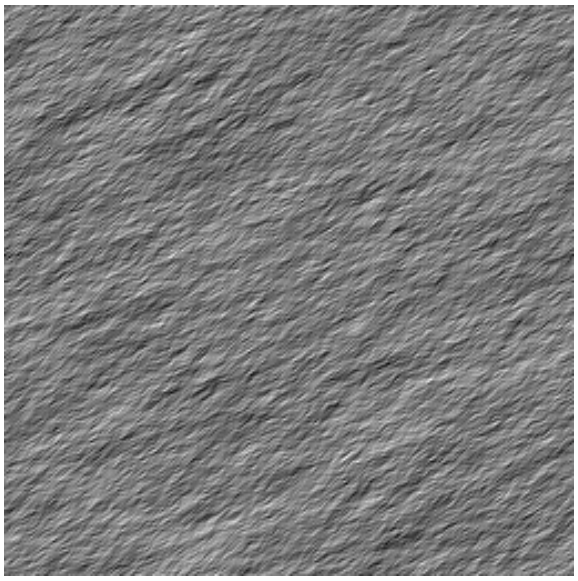


Figure 6.5: A simulated anisotropic fractal surface (gray level coded topographical image).

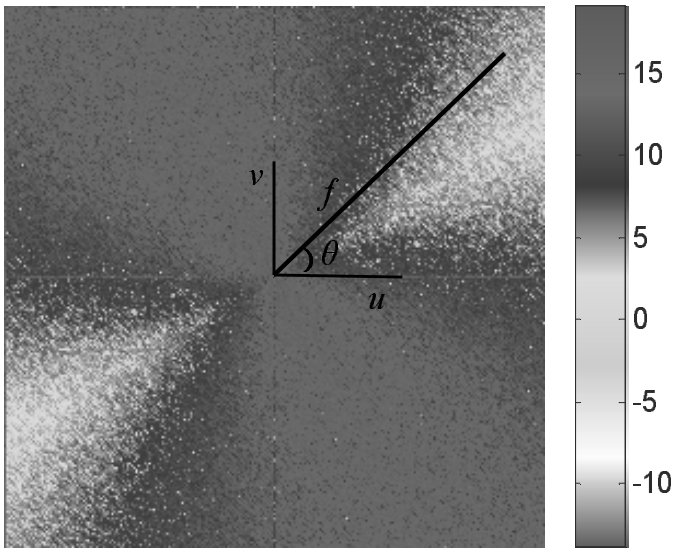


Figure 6.6: PSD of the fractal surface shown in Fig. 6.5.

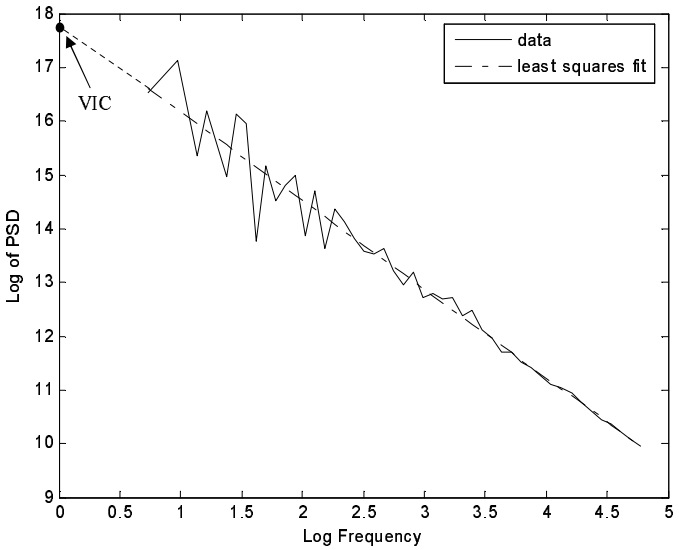


Figure 6.7: Power spectral density along the radial line shown in Fig. 6.6.

spectral power magnitudes over a log-log scale. By ignoring the first few Fourier coefficients near the center of the PSD that contains the low frequency information, any structural information for the entire surface data is skipped [80]. The slope of the fitted line is related to the FD from Eq. (6.11) by ($D_T = 2$):

$$D = \frac{8 - \beta}{2} \tag{6.13}$$

This is the FD for the angle θ. The angle θ is varied to build up a collection of FDs along each radial line. Because the PSD is symmetric the angle only needs to be varied from $0 \leq \theta \leq \pi$. By plotting each of the FDs on polar coordinates the directionality of the FD can be examined. This type of plot is called a *rose plot* [80]. Anisotropic rose plots of slope and FD of the surface shown in Fig. 6.5 are given in Figs. 6.9 (a) and (b) respectively.

In addition to computing the PSD, the phase distribution should also be checked to ensure randomness, a necessary property for a surface or image to be fractal [80]. This is an important criterion that is not always tested in measurement procedures. The same series of amplitudes would correspond to a decidedly nonfractal profile if the phase values were not randomized. Fig. 6.8 shows the plot of the histogram of the phase values for the terms in the Fourier series of the simulated surface shown in Fig. 6.5. They are indeed uniformly random.

Meanwhile, another interesting parameter that needs to be considered is the

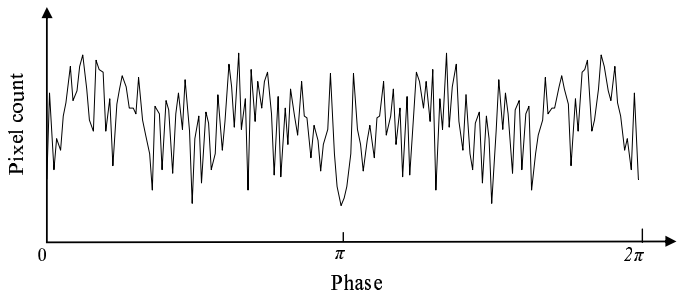


Figure 6.8: Histogram of the phase values for the terms in the Fourier series of the surface shown in Fig. 6.5.

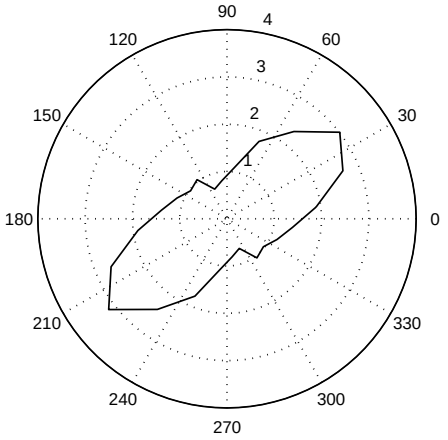
vertical intercept (VIC) of the regression line (see Fig. 6.7) with the vertical axis located at the origin. Russ [80], Rawers and Tylczak [74] argued that it has to do with surface roughness. Simply from the Eq. (6.10) we can see that the constant of proportionality c, which corresponds to the VIC of the fitted line, describes the overall magnitude of the roughness. Clearly this is an important parameter, since two surfaces which have the same fractal dimension (same β) but differ in this parameter will have different properties, and may be produced in different ways. Hence VIC is considered as another fractal value in this work. Fig. 6.9(c) shows the rose plot of VIC.

The simulated surface shown in Fig. 6.5 is anisotropic in structure, which can be easily recognized from its PSD shown in Fig. 6.6, but almost isotropic in RMS roughness. From the rose plots of slope, FD and VIC (see Fig. 6.9), we can see that the slope and corresponding FD characterize the anisotropy of the surface in structure, while the nearly isotropic VIC reflects the isotropy of the surface in RMS roughness.

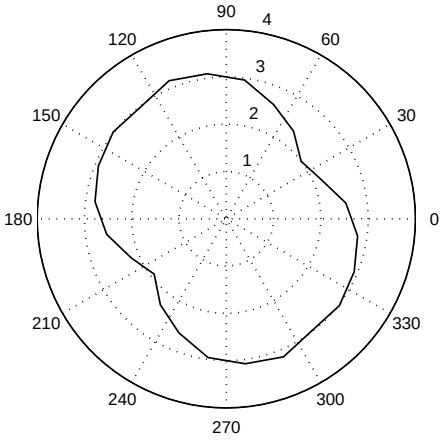
6.4.3 Indicators of State of Wear

As stated in the beginning of this section, one of the objectives of wear monitoring is to estimate the state or degree of wear of the surface so that the remaining lifetime of the concerning product or machine element can be predicted, given that the same wear conditions are kept. Hence, it would be necessary and helpful to define/find some indicators of state of wear, by which the wear status (or degree of wear) of the surface being monitored can be described quantitatively to a certain extent. These indicators may also further help understand or characterize a specific wear process.

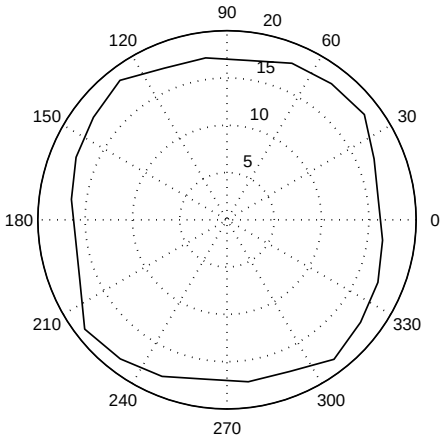
Generally a wear process, irrespective of wear type or mechanism, will somehow vary the surface features: surface roughness, elevation, and/or surface tilt,



(a) Slope



(b) FD



(c) VIC

Figure 6.9: Rose plots of slope, FD and VIC of the anisotropic surface shown in Fig. 6.5.

which result in morphological changes of the surface. Thus, some measures which characterize surface morphology and/or roughness may also be taken as indicators of state of wear.

Fractal values, FD and VIC, the important parameters in fractal analysis described in 6.4.2, are measures of the morphology, texture, and/or roughness in the surface (or images). The variation in surface structure caused by a wear process, which corresponds to morphological changes in grayvalue images captured by an imaging system, can be numerically characterized by fractal values. Thus fractal values may be potential indicators of wear state of a surface.

Next, online monitoring of dry sliding wear will be investigated using fractal analysis by assigning numerical values, namely fractal values, with the ultimate goal of relating them to the state of wear that might lead to predict the remaining life of a product or understand some specific wear phenomena.

6.5 Monitoring of Dry Sliding Wear

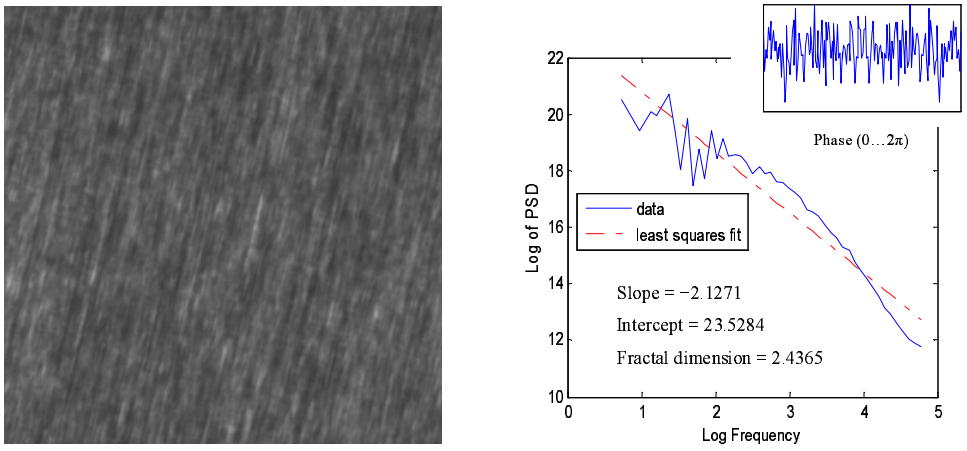
Many machined or man-made natural surfaces are of fractal nature, i.e. they have details at many scales and also part of a surface at a large scale is similar to a part of the surface at a small scale. The brightness images resulted from scattering or re-emission of light or other signals from surfaces are related in some unique way to surface roughness, elevation, and local surface slope. Consequently, the brightness patterns which are observed from a fractal surface will be mathematically a fractal as well [80]. Fig. 6.10 shows an image of a machined surface and associated fractal values. Hence it seems reasonable to expect the amount of wear to be revealed by progressive changes in fractal values, by comparing them from surface images of the same materials under similar lighting and viewing conditions. In the current study we restrict ourselves on machined surfaces only.

Considering the advantages of Fourier analytical technique on fractal analysis of images as introduced in Subsection 6.4.2, we will apply it to the online monitoring of dry sliding wear.

6.5.1 Experimental Results and Analysis

To validate the feasibility of fractal analysis techniques on online monitoring of wear, some wear experiments under dry sliding condition with different normal loads and rotary speeds were conducted. The experiments are similar to those described in Subsection 6.3.4. The operating conditions are listed in table 6.3. The online wear behavior of the disk surface is of our interest.

Likewise, in order to study the progressive changes of fractal values of a same



(a) (b)

Figure 6.10: (a) A surface image of a machined aluminium disk, and (b) power spectrum of its Fourier transform, plotted as directionally averaged of log (power) vs. log (frequency); The upper right part of (b) shows its phase distribution that is uniformly random.

surface area, the rotation of the disk and image acquisition must be synchronized as addressed in chapter 3. In this way the images from the same area (with negligible maximum 5 pixels error in each direction) of the wearing surface are acquired after each rotation and then fractal analysis of these images is done as described in Subsection 6.4.2. The fractal values are computed in 24 directions and the directionally averaged values are considered as the overall fractal values of the image. Images from the first 5 cycles and the subsequent every 5 cycles are used for analysis.

Some intermediate results of the experiment under the conditions that normal load is $5N$ and rotation speed 0.5rpm are shown in Fig. 6.11. From this figure we can see that the initial image of the machined surface (without wear) and those with dry sliding wear tracks are indeed fractals in that (1) their power spectra show linear variation between the logarithm of power spectrum and the logarithm of frequency; (2) their phase spectra are randomly distributed (a necessary property for fractals mentioned earlier).

Aside from checking randomness of phase spectrum, there may also be a problem in dealing with the presence (or absence) of a few frequencies in the spectrum. In terms of the physical surface (or image), the peak corresponds to a single sinusoid superimposed on the otherwise fractal surface. The sinusoid dominates the visual appearance of surface (or image). However, the process of fitting a linear

Table 6.3: Operating conditions of wear experiment

Material	Indenter: ball bearing steel Disk: aluminium
Pin shape and size	Sphere with diameter $3mm$
Normal load	1N, 2N, 5N
Hertzian radii	27.1μm, 34.1μm, 46.3μm
Rotation speed (linear velocity)	0.5rpm, 1rpm (3.1mm/s, 5.0mm/s)
Lubrication	None
Temperature	20°C
Image size	512 × 512 (pixels)

least squares line to hundreds of points in the power spectrum tolerates having a few points widely scattered from the main trend, i.e. the presence of a few widely scattered frequencies is only a minor component of the log (power spectrum) vs. log (frequency) plot. Another note is that the Fourier transform method assumes that the signal being processed is periodic. For an image, it means that the image is one tile in an endlessly repeating array. This in turn means that the upper and lower edges of the image, and the left and right edges of the image, should match. If they do not, it requires many high frequency terms to handle the abrupt step. It is necessary to ignore these terms as they do not contain information on the FD. These values are bypassed in all of the analyses hereafter.

Fig. 6.12 gives the plot of fractal values versus cycles of rotation of the experiment in Fig. 6.11. We can see that fractal values vary with the rotation cycles: the FD decreases and the VIC increases with the number of rotation cycles. This can be reasoned as follows: When the hard pin is sliding on the relatively softer disk with a certain load, generally the asperities of the disk are first pressed and plastically deformed; with the sliding continues, they are further removed from the surface as wear debris. This makes the surface locally smoother (local roughness decreases), but overall roughness of the viewed surface increases relatively. This is consistent with the changes of the VIC, so VIC may be regarded as a surface roughness indicator. On the other hand, with the press and removal of the asperities, the magnitudes of the low frequencies increase, which result in the drop of the FD consequently.

It can also be observed from Fig. 6.12 that the wear process reaches an almost stable stage after about 40 cycles. The fluctuations of the FD and VIC after 40 cycles are mainly caused by the formation of large wear debris, which are present in the images and may be removed by the pin in subsequent rotations. In the

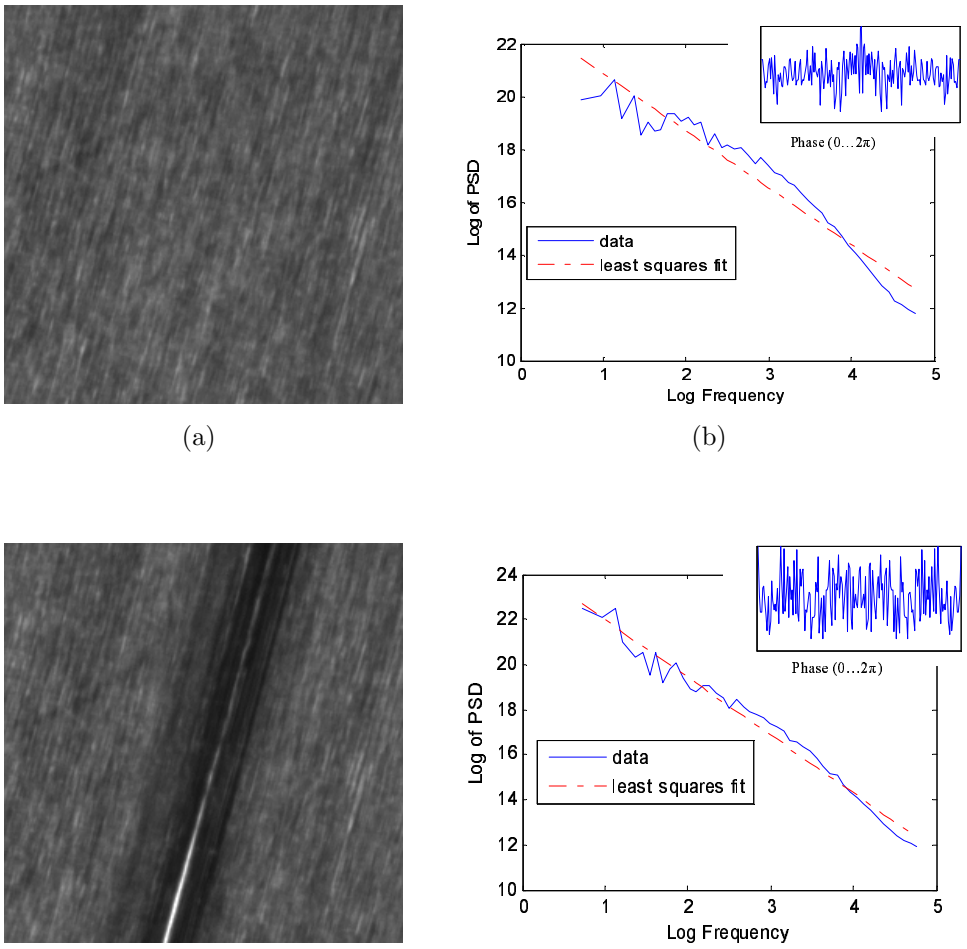
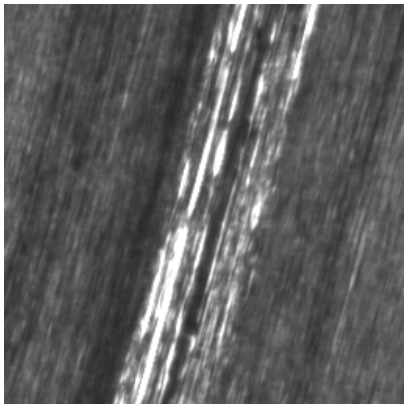
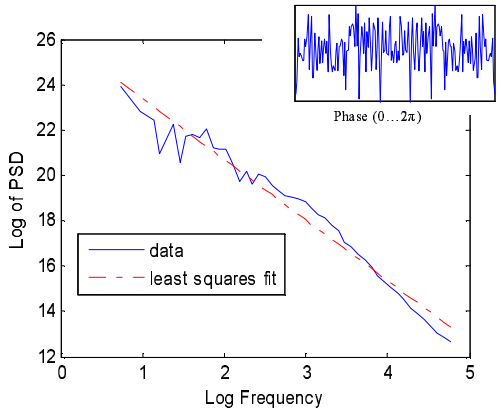


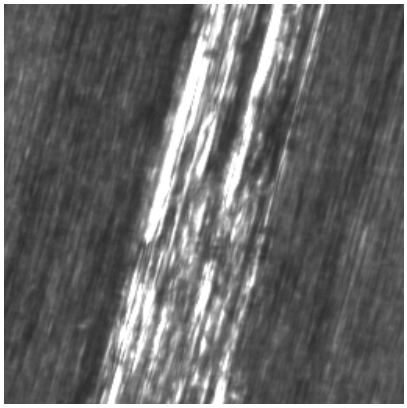
Figure 6.11: Experimental results of a dry sliding wear process: (a) a surface image of a machined aluminium disk before wear, and (b) its power spectrum, plotted as directionally averaged of log (PSD) vs. log (frequency), the upper right part of which shows its phase distribution; (c) is the image of the surface shown in (a) after 20 cycles of dry sliding rotation with normal force 5 N and rotary speed 0.5 rpm; and (d) is its corresponding power spectrum and phase distribution.



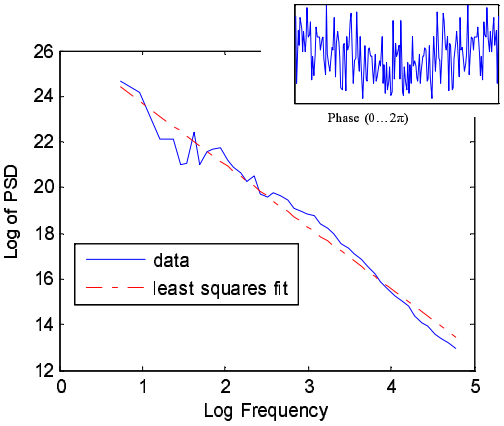
(e)



(f)



(g)



(h)

Figure 6.11: (continued) (e) (g) are the images of the surface in (a) after 40 and 60 cycles of dry sliding rotation, respectively, with normal force 5 N and rotary speed 0.5 rpm; and (f) (h) are their corresponding power spectra and phase distributions.

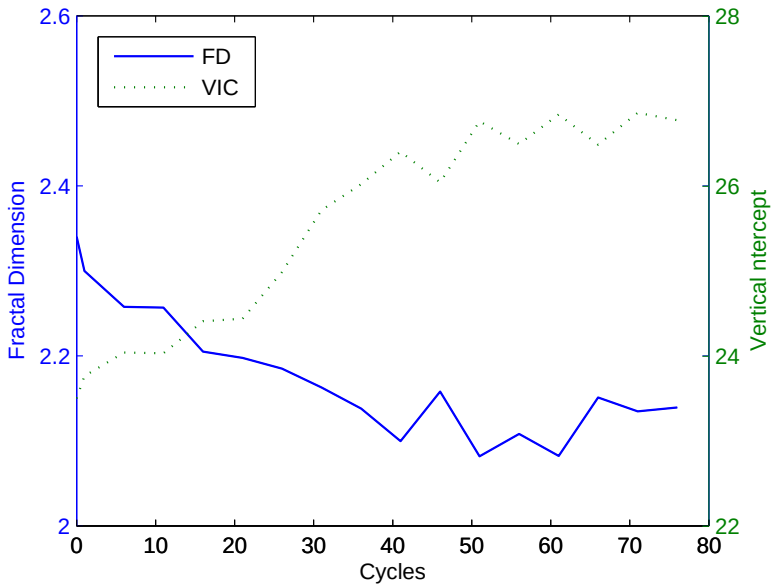


Figure 6.12: Fractal values vs. cycles of rotation under 5 N load and 1 rpm speed.

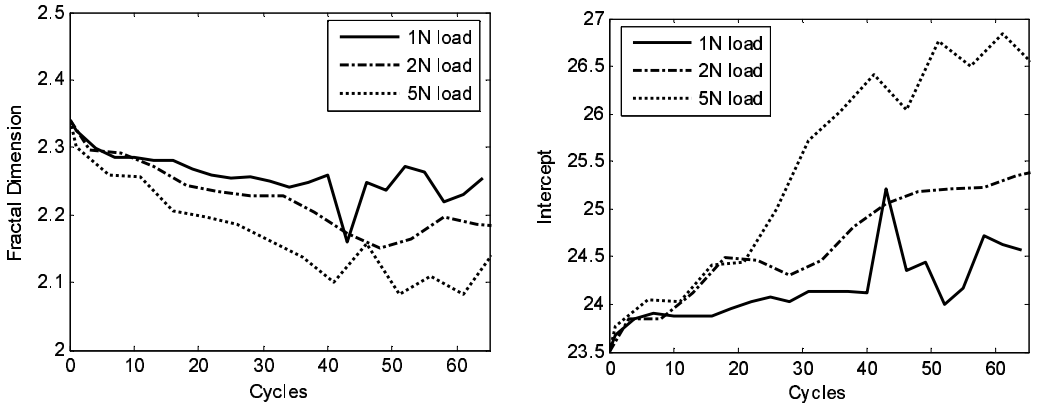
following we will study the effects of normal load and rotation speed on dry sliding wear using fractal analysis.

Effect of Load

Fig. 6.13 shows the results of variations of fractal values versus rotation cycles under different normal loads but same rotation speed. From this figure one can find that the larger the normal load, the faster the FD falls and the VIC rises. This is consistent with wear phenomena observed by offline methods, i.e. larger normal load evokes larger amount of wear on the surface.

Effect of Rotation Speed

Fig. 6.14 shows the results of variations of fractal values versus rotation cycles under different rotation speeds but same normal load. From this figure we can see that the speed (low) does not have big influence on the fractal values, but we still can observe some differences between these two curves. This can be reasoned as follows: for larger speed, the sliding wear reaches an intermediate stable state faster, but with subsequent sliding, wear debris loses from the surface, which results in a sharper change on fractal values, then the wear reaches another stable state



(a) (b)

Figure 6.13: FD (a) and VIC (b) vs. rotation cycles for same rotation speed (0.5 rpm), different loads.

more quickly.

Relation of Fractal Values and State of Wear

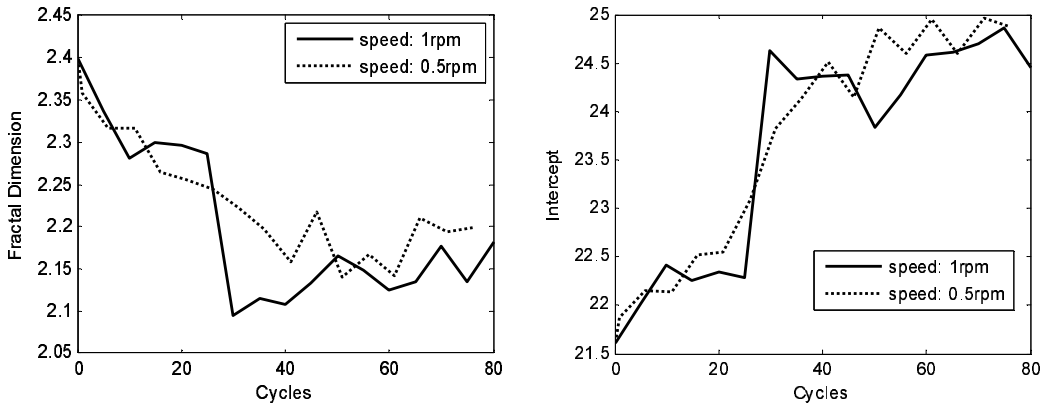
From the above experimental results and analysis we can see that fractal values numerically indicate the overall state of wear of the topography of a surface in a dynamic wear process. Hence we may take them as indicators of state of wear for online monitoring of dry sliding wear. Furthermore, since they can be easily computed from the FFT of surface images, they are suitable for realtime wear inspection oriented applications as well.

6.5.2 Discussions

To understand well the results of monitoring of a dry sliding wear process using fractal analysis, we must keep the following in mind:

1. The camera must always stay in focus during the observation of the wear process; otherwise unsharp images may cause incorrect fractal values. FD, for instance, may drop because of the absence of the terms of high frequencies in unsharp images.

2. The presence of wear particles causes the fluctuation of fractal values, therefore they become imprecise after a certain number of cycles of rotation.



(a) (b)

Figure 6.14: FD (a) and VIC (b) vs. rotation cycles for same load (5 N), different rotation speeds.

3. The surface images are self-affine not self-similar, i.e. the scale of the vertical direction (in gray values) is different from those of the spatial directions and the scales of the spatial directions may also differ. For this reason, the FD of a surface image is not necessarily in the range of 2 to 3, or the same as that obtained from other methods.

4. The overall magnification of the imaging system has a big influence on the calculation of the fractal values because of self-affinity. Therefore the magnification must be identical for comparison of the results from different wear experiments.

5. Fractal values give only a relative indication of the overall state of wear for a wear process. The real state of wear must be evaluated by correlation with other wear measurements, possibly offline.

6.5.3 Conclusions

The following conclusions may be drawn from this section:

1. Images of machined surfaces after dry sliding wear are still fractals.

2. Fractal analysis, Fourier analytical techniques in particular, can be used to the monitoring of dry sliding wear online.

3. Fractal values, namely FDs and VICs computed from PSDs of a sequence of images of a wearing surface relatively indicate the overall state of wear in a dry sliding case.

6.6 Summary

This chapter presents techniques for online wear monitoring. By analyzing image sequence of a wearing surface, dynamics of wear processes on roughness level are studied. Wear behavior of dry sliding wear in particular, for instance wear area against sliding cycles under different operating conditions, such as varying normal loads, different sliding velocities, is investigated. Wear area in the initial phase of a wear process identified by the simple thresholding method is found to be likely consistent with Archard's equation.

To estimate the degree of wear, fractal values, namely fractal dimension and vertical intercept, computed from the power spectra of images of a wearing surface, are adopted as indicators of the dynamic wear process. Experimental results show that progressive changes of fractal values might reveal the wear status of the surface.

Chapter 7

Conclusions and Recommendations

This chapter presents some final conclusions on this research as well as some recommendations for further improvement and future research.

7.1 Conclusions

The goals of this thesis, as formulated in Subsection 1.2.3, were to design a simple, low cost, laboratory-use apparatus for conducting wear test as well as online wear measurement, to use this apparatus to develop techniques for online wear detection and monitoring, and to study wear dynamics of ad hoc dry sliding wear. In this section the thesis is concluded by discussing to what extent these research goals have been accomplished.

7.1.1 Review of Wear Monitoring

In Chapter 2 the existing methods for wear monitoring were reviewed. These methods were generally categorized into 6 types in terms of different wear measures. Among them the state of the art methods are topographic methods using modern high resolution instruments such as STM, AFM and interference microscope. Such topographic methods are mainly limited to fundamental research, i.e. measuring wear of some materials on a nanometer scale. On the other hand, the modern instruments used in such methods are sensitive to vibration, complicated to operate, comparatively slow and high cost, and not suitable for industrial applications.

Nearly all the existing methods, irrespective of their measurement principles, are offline. Such inherent nature makes them not suitable for measuring a wearing surface in motion. As such using these methods we can not accurately characterize the behavior of a dynamic wear process.

Non-contact, low cost, high speed and reasonably accurate wear monitoring methods, particularly online, are highly demanded in tribological research as well as in industrial applications. This thesis was aimed to develop such low-cost online methodology.

7.1.2 Wear Apparatus Design

Chapter 3 developed a simple, cost-effective, laboratory-use apparatus for online wear testing and monitoring. The apparatus combines a simplified Pin-on-Disk wear tester and an imaging system. The wear tester part is capable of simulating three types of sliding contacts: point contact, line contact and conforming contact by employing different pins of spherically, cylindrically and flat shaped tips respectively. Other important factors for wear testing, load and velocity for instance, can also be varied. The imaging system part consists of a video zoom microscope and a CCD camera. It is capable of acquiring images of a surface in motion in real time.

Using the designed wear apparatus, dry sliding wear test for three contact modes can be conducted. Online wear detection and monitoring can be performed by analyzing the images captured by the imaging system. This wear apparatus offers more accurate results than traditional offline wear measurement instruments in the sense of neither interrupting the dynamic wear process nor changing wear environment or conditions. Furthermore, it is fast, noncontact, cost effective and easy to operate.

With the current wear apparatus we are able to estimate surface roughness statistically, but can not measure surface topography.

7.1.3 Ring Light Illumination Modeling

Considering the crucial role illumination plays in an imaging system, Chapter 4 discussed the choice of illumination strategy for a microscope. According to this strategy, a fiber optic ring light was found to be a proper illuminator for the imaging system designed in Chapter 3, under which the wear features of interest were revealed. A ring light illumination model was then developed. It showed that uniform illumination on the nominal plane of a surface can be achieved when the ring is located in a proper height range. Irradiance diagrams predicted by the model were in good agreement with those measured by a photometer. Uniform illumination is even more easily achieved for a microscope (the current case) due to its small FOV.

Modeling of surface reflection is a very complex problem. A specular reflectance model, based on the Torrance-Sparrow model, was formulated for ring light illumination. The parameters, such as local viewing angles, RMS roughness, RMS slope, etc. that affect surface radiance were also discussed and evaluated. The radiance variation predicted by the model with respect to the parameters was found in good agreement with the measurements from the current imaging system. In addition, the reflectance model for ring light illumination can be generalized to those surfaces whose slope distribution are not of the assumed Gaussian shape due to the inherent integration over all azimuth angles under the ring light.

A wear process that alters surface parameters, such as surface local height, RMS roughness, RMS slope and mean orientations of local surface patches etc., results in changes of surface radiance, image irradiance and image texture thereafter. The models developed in Chapter 4 helped interpret wear patterns present on the specimen's surface images and provided theoretical basis for detection and monitoring of wear by analyzing image sequences.

7.1.4 Online Wear Detection

Wear detection on machined surfaces can be regarded as a texture segmentation problem. Chapter 5 developed two filtering approaches to online wear detection: unsupervised detection scheme using multichannel Gabor filters and supervised detection scheme using optimized filters. Both approaches can successfully detect wear patterns present on textured surfaces, machined surfaces for instance. Experiments conducted on various real wear samples confirmed the usefulness of these two approaches.

Using only the real part of a Gabor filter, reasonable performance of wear detection was found. This is useful for realtime applications due to its computational efficiency.

For both approaches, filter size affects the performance of wear detection. Trade-offs between speed and performance must be made for realtime applications. The unsupervised scheme using multichannel Gabor filters needs little prior knowledge of wear and was found to offer better detection results. Compared to Gabor filters, the supervised scheme using optimized filters obtained successful results with less computationally complex filters. However it needs better insight, representative wear samples for instance, in the design (calibration) phase.

These two approaches were only qualitatively evaluated. Quantitatively evaluation needs to be done by correlation with measurements from offline methods.

7.1.5 Online Wear Monitoring

Using the wear apparatus developed in Chapter 3, some behaviors and dynamics of ad hoc dry sliding wear processes were investigated in Chapter 6. The wear area identified by the simple thresholding method in an image sequence might be a proper measure for a dry sliding wear process. And for the initial phase of a wear process, wear area measured in this way was preliminarily found to be proportional to (number of cycles)$^{1/3}$, which is consistent with Archard's equation (obtained offline) [1].

Images of machined surfaces after dry sliding wear were found to be still fractals. Techniques of fractal analysis, Fourier analytical techniques in particular, can be used to monitor wear, by analyzing image sequences of a wearing surface. Fractal values, namely FDs and VICs computed from PSDs of a sequence of images of a wearing surface relatively indicate the overall state of wear in a dry sliding case.

Online monitoring of dry sliding wear was only investigated on relatively slow speeds. For a high moving speed, the motion blur of the images need to be considered.

7.1.6 Overall Conclusions

The research goals formulated in Subsection 1.2.3 for the study of methodology for online wear monitoring have been accomplished. With certain limitations the imaging methods presented in this thesis may provide a noncontact, economical, fast and accurate solution to the online wear monitoring problem.

The designed laboratory-use wear apparatus, together with the imaging techniques are of potential usefulness not only in tribological research but in some industrial applications as well: such as cutting tool wear monitoring, damage detection of metal sheets in production phase, coating failure inspection etc. Some preliminary experiments of primer damage detection conducted on metal plates produced by Corus [90] validated the approaches developed in Chapter 5.

The same as other methods, the methodology for online wear monitoring presented in this thesis inevitably have some limitations. In the following we list some general limitations:

- The inspection surfaces need to be relatively smooth and clean;
- The wear apparatus and the imaging methods were developed to study only dry wear (not lubricated);
- The proposed approaches to wear detection work only on textured surfaces;
- The fractal image analysis approach to wear monitoring is applicable only to fractal surfaces (machined surfaces for instance).

7.2 Recommended Improvements and Future Work

In the following some recommendations are provided for further improvement of this research and some problems that may be explored for future study.

7.2.1 Wear apparatus

- Presence of wear particles causes errors in image analysis. By designing a mechanical part to sweep the wear particles away from the FOV of microscope may improve the performance;
- Adding the adapter designed by Canner [83, 38] for free-rolling contact (see Fig. 7.1) to the wear tester may extend its functionality to rolling wear;

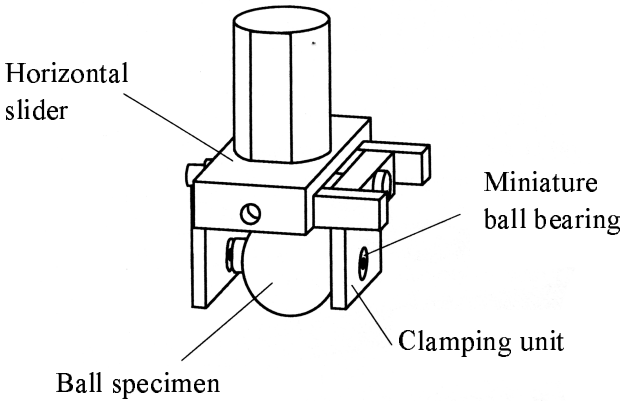


Figure 7.1: An adaptor for free-rolling contact experiment.

- Using a sophisticated stereo microscope instead of the current one, topographic images may be obtained.

7.2.2 Illumination

In the current imaging system, a fiber optic ring light is employed as illumination for its advantages (presented in Chapter 4). The ring light provides uniform and diffused illumination to the specimen. This might also be a disadvantage for observing wear on the specimen's surface. The reason is that, surface texture is an important feature for wear detection. We should keep as much as possible surface textural information in the image. Ring light illumination, however, reduces texture contrast for its diffused nature (360° coverage). For instance, the surface shown in Fig. 4.21 (a) has obvious patterns of lays and grooves, which are not clearly seen in Fig. 4.20 (a). To compensate this shortcoming of ring light illumination, it might be good to combine a directional oblique illumination so that surface textural information remains in the image to some extent, meanwhile the shadows are illuminated by the ring light to some degree to prevent the lose of the surface details. An optimal illumination might be achieved by adjusting the angle of the directional oblique light and intensity of the ring light.

7.2.3 Wear Detection and Monitoring

- Some specific wear behavior or phenomena can be further studied by characterizing the detected wear image textures and correlating with offline mea-

surements.

- Correlating fractal values measured by the current imaging methods with those obtained from offline topographical methods may lead to an algorithm for estimating the remaining life of a certain element of certain material.

Appendices

Appendix A

Derivation of the Parameters of the Gabor Filter Bank

The following derivation is mainly based on the explanatory notes from [54]. The Gabor filter bank introduced in Chapter 5 is

$$g_{mn}(x,y) = \xi^{-m} g(x',y'), \tag{A.1}$$

where

$$\begin{cases} x' = \xi^{-m}(x\cos\theta_n + y\sin\theta_n) \\ y' = \xi^{-m}(-x\sin\theta_n + y\cos\theta_n) \end{cases}.$$

$\xi > 1$, $m = 1, 2, \ldots, P$, $n = 1, 2, \ldots, Q$, and $\theta_n = (n-1)\pi/Q$.

Given u_l, u_h, P, and Q, we are going to find ξ, σ_x and σ_y so that the half-peak magnitude response of adjacent filters tough each other.

The Fourier transform of Eq. (A.1) is

$$G_{mn}(u,v) = \xi^{m-1} G(\xi^{m-1}(u\cos\theta + v\sin\theta), \xi^{m-1}(-u\sin\theta + v\cos\theta)), \tag{A.2}$$

where $G(u,v)$ is the Fourier transform of Gabor filter $g(x,y)$ (see Eq. (5.2)). Thus

$$G_{11}(u,v) = \exp(-\frac{1}{2}(\frac{((u-u_h)^2}{\sigma_u^2} + \frac{v^2}{\sigma_u^2})) \tag{A.3}$$

$$G_{21}(u,v) = \xi \exp(-\frac{1}{2}(\frac{((\xi u-u_h)^2}{\sigma_u^2} + \frac{(\xi v)^2}{\sigma_u^2})) \tag{A.4}$$

$$G_{31}(u,v) = \xi^2 \exp(-\frac{1}{2}(\frac{((\xi^2 u-u_h)^2}{\sigma_u^2} + \frac{(\xi^2 v)^2}{\sigma_u^2})) \tag{A.5}$$

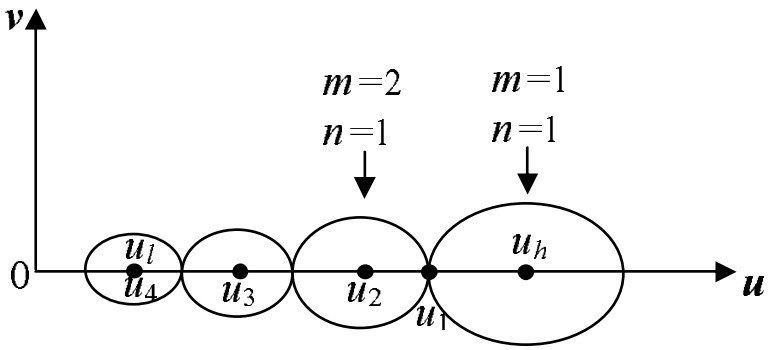


Figure A.1: The contours of half-peak magnitude responses of adjacent Gabor filters with the same orientation but different scales.

We want $G_{11}(u,v)$ and $G_{21}(u,v)$ to touch at half-peak magnitude as illustrated in Fig. A.1, i.e.

$$\exp(-\frac{1}{2}(\frac{((u_1-u_h)^2}{\sigma_u^2}+0) = \frac{1}{2} \tag{A.6}$$

$$\xi\exp(-\frac{1}{2}(\frac{((\xi u_1-u_h)^2}{\sigma_u^2}+0) = \frac{\xi}{2} \tag{A.7}$$

From Eq. (A.6) and Eq. (A.7) we get

$$\begin{cases} u_1 = -\sigma_u\sqrt{2ln2}+u_h \\ u_1 = \frac{u_h}{\xi}+\frac{\sigma_u}{\xi}\sqrt{2ln2} \end{cases} \tag{A.8}$$

From Eq. (A.8) we obtain

$$\sigma_u = \frac{(\xi-1)u_h}{(\xi+1)\sqrt{2\ln 2}} \tag{A.9}$$

Then according to Eq. (5.3),

$$\sigma_x = \frac{(\xi+1)\sqrt{2\ln 2}}{2\pi u_h(\xi-1)} \tag{A.10}$$

Eq. (A.7) gives ellipse equation with center $u_2 = u_h/\xi$. Similarly, for $m=3$, $n=1$, we get $u_3 = u_h/\xi^2$ according to Eq. (A.5). Hence, for $m=P$, we have

$$\begin{aligned} u_l &= \frac{u_h}{\xi^{P-1}}, \\ \xi &= \left(\frac{u_h}{u_l}\right)^{\frac{1}{P-1}}. \end{aligned} \tag{A.11}$$

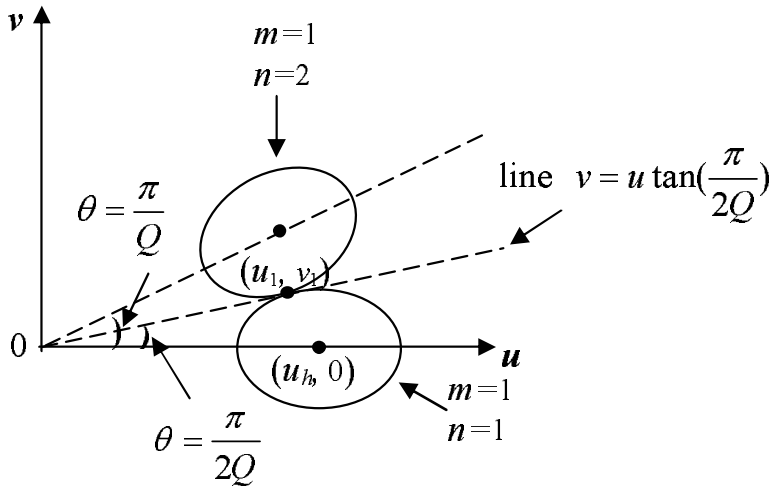


Figure A.2: The contours of half-peak magnitude responses of adjacent Gabor filters with the same scale but different orientations.

For the rotated Gabor functions at $\theta = (n-1)\pi/Q$, they have to touch also at half-peak magnitude (see Fig. A.2). We need the two ellipse to touch at half-peak, which is the same as the line $v = u\tan\frac{\pi}{2Q}$ being touched to the ellipse $G_{11}(u,v) = 1/2$, i.e. intersecting the line $v = ku$ with

$$\frac{u-u_h}{\sigma_u^2 2\ln 2} + \frac{v^2}{\sigma_v^2 2\ln 2} = 1 \tag{A.12}$$

for a single point (u_1, v_1) by eliminating v from Eq. A.12, and setting the discriminator to zero for a unique solution in the equation of u, one gets the following equation:

$$4u_h^2 b^4 = 4(b^2 + a^2k^2)(b^2u_h^2 - a^2b^2) \tag{A.13}$$

where $b^2 = \sigma_v^2(2\ln 2)$, $a^2 = \sigma_u^2(2\ln 2)$, $k = \tan\pi/2Q$. Hence

$$\sigma_v = \frac{\tan(\frac{\pi}{2Q})}{\sqrt{2\ln 2}}\sqrt{u_h^2 - \sigma_u^2(2\ln 2)} \tag{A.14}$$

Then according to Eq. (5.4), we get

$$\sigma_y = \left(2\pi\tan(\frac{\pi}{2Q})\sqrt{\frac{u_h^2}{2\ln 2} - \frac{1}{(2\pi\sigma_x)^2}}\right)^{-1} \tag{A.15}$$

Appendix B

Hertzian Radius

During the contact of two surfaces, contact will initially occur at only a few points (*asperities*) to support the normal load (force). In most contact situations, asperities may be deformed either elastically or plastically. In the experiments conducted using the designed wear apparatus, a spherically shaped pin is in contact with a smooth surface. In the following we consider the simpler idealized case of a single asperity loaded on a homogeneous and high elastic (rigid) flat body (see Fig. B.1). Such contact of two elastic solids are referred to as *Hertzian contact* [26].

During the compression by a normal force F, the spherical body deforms. If it does not deform, its profile would overlap with rigid body as shown by the dotted lines in Fig. B.1. The elastic deformation results in displacement of the surface outside the footprint such that the contact area is less than the overlap area resulting from intersection of the dotted surface with the flat plane.

We now consider the problem of elastic deformation of the sphere of radius R in solid contact with an applied normal load F. The contact area is circular, having a radius r. From the Hertz analysis [26], the contact radius, referred to as *Hertzian radius*, is:

$$r = \left(\frac{3FR}{4E^*}\right)^{1/3} \tag{B.1}$$

where E^* represents the effective modulus of elasticity, defined by:

$$\frac{1}{E^*} = \frac{1-v_1^2}{E_1} + \frac{1-v_2^2}{E_2} \tag{B.2}$$

The parameters E and v are Young's modulus of elasticity and Poisson's ratio, respectively; subscripts 1 and 2 refer to the two bodies. From Eq. (B.1), note that the area of contact increases as normal load $F^{2/3}$.

For example, in our experiment of ball bearing steel against Aluminum disk:

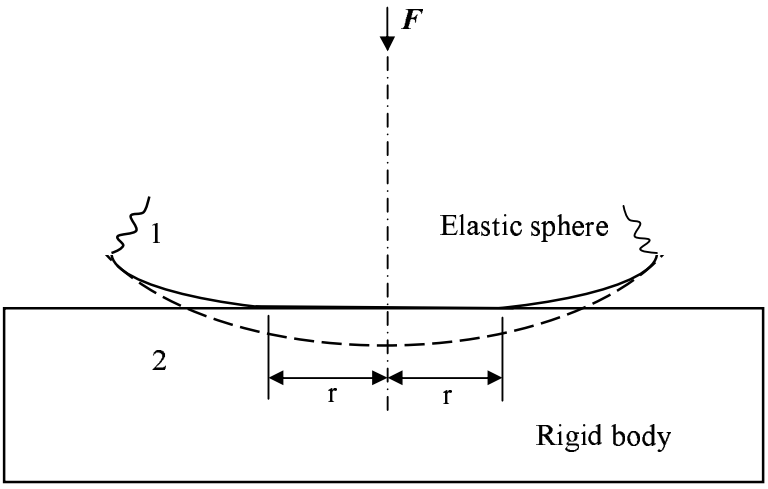


Figure B.1: Schematic of normal contact of an elastic sphere and a rigid body.

Ball bearing steel: $E_1 = 195\ GPa$, $v_1 = 0.28$
Aluminum: $E_2 = 69\ GPa$, $v_2 = 0.33$
$E^* = 56.6873\ GPa$
For $R = 1.5\ mm$,
the Hertzian contact radii for different normal loads are:
$r = 27.1\ \mu m$ ($F = 1$ N), $r = 34.1\ \mu m$ ($F = 2$ N),
$r = 46.3\ \mu m$ ($F = 5$ N).

Summary

Wear is traditionally measured offline. This thesis discusses a new methodology for online detection and monitoring of wear. The study comprises the design of an online wear testing apparatus and the development of techniques for online wear detection and monitoring using imaging methods.

A simple, cost-effective, laboratory-use apparatus for online wear testing and monitoring was designed, which combines a simplified Pin-on-Disk wear tester and an imaging system. Using the designed wear apparatus, dry sliding wear tests for three contact modes (point, line and conforming) can be conducted, meanwhile images of a surface in motion can be acquired in real time. Online wear detection and monitoring can be performed by analyzing a sequence of images of a wearing surface. This wear apparatus offers more accurate results than traditional offline wear measurement instruments in the sense of neither interrupting the dynamic wear process nor changing wear environment or conditions. Furthermore, it is fast, non-contact, cost effective and easy to operate.

Considering the crucial role illumination plays in an imaging system, a fiber optic ring light, according to the choice of illumination strategy, was chosen as a proper illuminator for the imaging system, under which the wear features of interest were revealed. An illumination model and a reflectance model were then developed for ring light illumination. Both models were numerically evaluated and validated by experiments. Good agreement between the models and the measurement results was found. A wear process alters surface parameters, and results in changes of surface radiance, image irradiance and image texture. The proposed models help to interpret wear patterns present on the specimen's surface images and provide a theoretical basis for detection and monitoring of wear by analyzing image sequences.

Wear detection on machined surfaces can be regarded as a texture segmentation problem. Two filtering approaches, namely an unsupervised detection scheme using multi-channel Gabor filters and a supervised detection scheme using optimized filters were developed for online wear detection. Both approaches can successfully detect wear patterns present on textured surfaces, machined surfaces for instance.

Experiments conducted on various real wear samples confirmed the usefulness of these two approaches.

Monitoring of a dynamic wear process can be conducted by analyzing a sequence of images from a wearing surface. The wear behavior of an ad hoc dry sliding wear process was investigated under different operating conditions. Fractal analysis techniques were also applied to estimate the state of wear during this process. Experimental results showed that fractal values, more precisely fractal dimensions and vertical intercepts computed from power spectra of a sequence of images of a wearing surface are potential good indicators of the state of wear.

Samenvatting

Slijtage wordt gewoonlijk offline gemeten. Dit proefschrift handelt over een nieuwe methode voor het online detecteren en monitoren van slijtage. Het onderzoek omvat het ontwerp van een testapparaat voor het online bestuderen van slijtage, alsmede de ontwikkeling van technieken voor het online detecteren en meten van slijtage met behulp van optische technieken.

Er is een eenvoudig en goedkoop testapparaat ontworpen voor online slijtagemetingen, bestaande uit een combinatie van een pin-on-disk slijtagemodule en een visionsysteem. Met dit instrument kunnen dry sliding slijtageproeven worden uitgevoerd met drie verschillende vormen van contactoppervlak (punt, lijn en conform), waarbij gelijktijdig beelden worden opgenomen van het oppervlak dat aan slijtage onderhevig is. Het online detecteren en monitoren van slijtage gebeurt door het opnemen van beeldreeksen van het slijtende oppervlak en deze te analyseren. Met dit instrument kan het slijtageproces nauwkeurig worden bestudeerd, beter dan met traditionele testapparaten, in die zin dat het slijtageproces niet hoeft te worden onderbroken en de omgevingscondities en slijtageparameters constant blijven. Bovendien is de meetmethode snel, contactloos, relatief goedkoop en gemakkelijk uit te voeren.

Belichting speelt een essentiele rol in beeldverwerkende systemen. Een ringvormige lichtbron bleek de beste belichting op te leveren; met deze belichtingsstrategie werden goede beelden verkregen van de relevante slijtagekenmeren. Er zijn modellen ontwikkeld voor zowel de belichting van getextureerde oppervlakken als de reflectie ervan. Beide modellen zijn numeriek ge?valueerd en met experimenten getoetst. De experimentele resultaten waren in goede overeenstemming met de modelvoorspellingen. Door slijtage veranderen oppervlakteparameters, die op hun beurt resulteren in een andere lichtemissie, beeldvorming en waargenomen textuur. De voorgestelde modellen bieden hulp bij het interpreteren van slijtagepatronen die in het beeld van het te bestuderen oppervlak verschijnen, en verschaffen een theoretische basis voor het detecteren en monitoren van slijtage bij de analyse van de beeldreeksen.

Het vaststellen van slijtage van bewerkte oppervlakken is terug te voeren tot

een segmentatieprobleem. Er zijn twee filtermethoden ontwikkeld voor het online vaststellen van slijtage: unsupervised detectie met behulp van meerkanaals Gabor-filters, en supervised detectie gebruik makend van optimale filters. Beide methoden maken het mogelijk slijtagepatronen op getextureerde oppervlakken (zoals bewerkte mataaloppervlakken) succesvol te detecteren. De bruikbaarheid van beide methodes is door middel van experimenten overtuigend aangetoond.

Door het analyseren van beeldreeksen van slijtende oppervlakken kan het dynamische proces van slijtage worden gevolgd. Het gedrag van een specifiek gekozen dry sliding slijtageproces is bestudeerd onder verschillende slijtage-condities. Voor het schatten van de slijtagetoestand is mede gebruik gemaakt van fractaal-analyse. Uit experimenten is gebleken dat fractale parameters - en in het bijzonder fractale dimensie en vertikale doorsnijding - berekend uit de vermogensspectra van een beeldreeks van een slijtend oppervlak, goede indicatoren zijn van de slijtagetoestand.

Bibliography

[1] J. F. Archard. Contact and rubbing of flat surfaces. *Journal of Applied Physics*, 24(8):981–988, 1953.

[2] D. I. Barnea and H. F. Silverman. A class of algorithms for fast digital image registration. *IEEE Transactions on Computers*, C-21(2):179–186, 1972.

[3] R. G. Bayer. *Mechanical Wear Prediction and Prevention*. Marcel Dekker, Inc., New York etc., 1994.

[4] P. Beckmann and A. Spizzichino. *The Scattering of Electromagnetic Waves from Rough Surfaces*. Pergamon Press, Oxford, etc., 1963.

[5] F. Bergman, P. Hedenqvist, and S. Hogmark. The influence of primary carbides and test parameters on abrasive and erosive wear of selected PM high speed steels. *Tribology International*, 30(3):183–191, 1997.

[6] B. Bhushan. *Principles and Applications of Tribology*. Wiley, New York etc., 1999.

[7] A. C. Bovik, M. Clark, and W. S. Geisler. Multichannel texture analysis using localized spatial filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(1):55–73, 1990.

[8] D. P. Casasent and J. S. Smokelin. Neural net design of macro gabor wavelet filters for distortion-invariant object detection in clutter. *Optical Engineering*, 33(7):2264–2271, 1994.

[9] C. M. Chen, H. H. Lu, and K. Han. A textural approach based on gabor functions for texture edge detection in ultrasound images. *Ultrasound in Medicine and Biology*, 27(4):515–534, 2001.

[10] D. A. Clausi. Comparison and fusion of co-occurrence, Gabor and MRF texture features for classification of SAR sea-ice imagery. *Atmosphere-Ocean*, 39(3):183–194, 2001.

[11] C. H. Cohen, Z. Fau, and S. Attali. Automated inspection of textile fabrics using textured models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13:803–808, 1991.

[12] M. B. d. Rooij. *Tribological Aspects of Unlubricated Deepdrawing Process*. PhD thesis, University of Twente, the Netherlands, 1998.

[13] J. G. Daugman. Uncertainty relation for resolution in space, spatial frequency and orientation optimised by two dimensional visual cortical filters. *Journal of the Optical Society of America*, 2(7):1160–1169, 1985.

[14] J. G. Daugman. High confidence visual recognition of persons by a test of statistical independence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(11):1148–1161, 1993.

[15] S. Deladi. *MEMS Generated and AFM-Based Surface Modification*. PhD thesis, University of Twente, the Netherlands, 2005.

[16] D. Dunn and W. E. Higgins. Optimal Gabor filters for texture segmentation. *IEEE Transactions on Image Processing*, 4(7):947–964, 1995.

[17] D. Dunn, W. E. Higgins, and J. Wakeley. Texture segmentation using 2-D Gabor elementary functions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(2):130–149, 1994.

[18] D. M. Eissenberg and H. D. Haynes, editors. *Motor-Current Signature Analysis*, volume 18 of *Friction, Lubrication and Wear Technology*. ASM Handbook, ASM International, 1992.

[19] H. A. ElMaraghy and D. J. Bullis. Expert inspector of surface defects. *Computers in Industry*, 11:191–205, 1989.

[20] J. Escofet, R. Navarro, and M. S. Millan. Detection of local defects in textile webs using Gabor filters. *Optical Engineering*, 37(8):2297–2307, 1998.

[21] K. Fukunaga. *Statistical Pattern Recognition*. Academic Press, San Diego, CA, 2nd edition, 1990.

[22] R. Gahlin and S. Jacobson. A novel method to map and quantify wear on a micro-scale. *Wear*, 222(2):93–102, 1998.

[23] W. A. Glaeser. Wear measurement techniques using surface replication. *Wear*, 40:135–137, 1976.

[24] R. Haralick, K. Shanmugam, and I. Dinstein. Textural features for image classification. *IEEE Transactions on Systems, Man, Cybernetics*, 3:610–621, 1973.

[25] X. D. He, K. E. Torrance, F. X. Sillion, and D. P. Greenberg. A comprehensive physical model for light reflection. *Computer Graphics*, 25(4):175–186, 1991.

[26] H. Hertz. Ueber die beruhrung fester elasticher korper. *Journal Fur Die Reine Und Angewandte Mathematik*, 92:156–171, 1881.

[27] B. K. P. Horn. *Robot Vision*. The MIT Press, 1986.

[28] B. K. P. Horn and R. W. Sjoberg. Calculating the reflectance map. *Applied Optics*, 18(11):1770–1779, 1979.

[29] S. A. Hosseini-Ravandi. Fourier transform analysis of plain weave fabric appearance. *Textile Research Journal*, 65(11):676–683, 1995.

[30] http://www.edmundoptics.com/.

[31] http://www.edmundoptics.com/IOD/DisplayProduct.cfm?productid=1384.

[32] http://www.edmundoptics.com/IOD/DisplayProduct.cfm?productid=1432.

[33] http://www.microscopyu.com/articles/stereomicroscopy/stereoreflected.html.

[34] http://www.ph.tn.tudelft.nl/courses/FIP/noframes/fip Segmenta.html.

[35] P. Huang and J. Wei. A new method to measure wear cracks by using reflected light intensity. *Wear*, 231:153–157, 1999.

[36] A. K. Jain and F. Farrokhnia. Unsupervised texture segmentation using Gabor filters. *Pattern Recognition*, 24(12):1167–1186, 1991.

[37] A. K. Jain, N. L. Ratha, and S. Lakshmanan. Object detection using Gabor filters. *Pattern Recognition*, 30:295–309, 1987.

[38] J. Jamari. *Running-in of Rolling Contacts*. PhD thesis, University of Twente, the Netherlands, 2006.

[39] W. J. Jasper, S. J. Garnier, and H. Potlapalli. Texture characterization and defect detection using adaptive wavelets. *Optical Engineering*, 35(11):3140–3149, 1996.

[40] W. J. Jasper and H. Potlapalli. Image analysis of mispicks in woven fabric. *Text. Res. J.*, 65(1):683–692, 1995.

[41] K. I. Jolic, C. R. Nagarajah, and W. Thompson. Non-contact, optically based measurement of surface roughness of ceramics. *Measurement Science and Technology*, 5:671–684, 1994.

[42] B. Julsez. Visual pattern discrimination. *IRE Transactions on Information Theory*, 8:84–92, 1962.

[43] A. A. Kassim, M. Zhu, and M. A. Mannan. Texture analysis using fractals for tool wear monitoring. In *Proceedings International Conference on Image Processing*, volume 3, pages 105–108, Rochester, NY, USA, 2002.

[44] M. B. Kiran, B. Ramamoorthy, and V. Radhakrishnan. Evaluation of surface roughness by vision system. *International Journal of Machine Tools and Manufacture*, 38(5-6):685–690, 1998.

[45] D. Kuhlmann-Wilsdorf. *Fundamentals of Friction and Wear of Materials*. Asm. 1981.

[46] A. Kumar and G. Pang. Defect detection in textured materials using Gabor filters. *IEEE Transactions on Industry Applications*, 38(2):425–440, 2002.

[47] A. Kumar and G. Pang. Defect detection in textured materials using optimized filters. *IEEE Transactions on Systems, Man and Cybernetics Part B-Cybernetics*, 32(5):553–570, 2002.

[48] J. H. Lambert. *Photometria Sive De Mensura De Gratibus Luminis, Colorum Et Umbrae*. Eberhard Klett, Augsburg, Germany, 1760.

[49] F. E. Lockwood and R. Dalley, editors. *Lubricant Analysis*, volume 18 of *Friction, Lubrication and Wear Technology*. ASM Handbook, ASM International, 1992.

[50] F. Luk, V. Huynh, and W. North. Measurement of surface roughness by a machine vision system. *Journal of Physics E-Scientific Instruments*, 22:977–980, 1989.

[51] A. Mahalanobis and H. Singh. Application of correlation filters for texture recognition. *Applied Optics*, 33:2173–2179, 1994.

[52] S. Mallat. Wavelets for a vision. In *Proceedings of the IEEE*, volume 84, pages 604–614, 1996.

[53] B. B. Mandelbrot. *The Fractal Geometry of Nature*. Freeman Press, Oxford, 1982.

[54] B. S. Manjunath and W. Y. Ma. Texture features for browsing and retrieval of image data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(8):837–842, 1996.

[55] J. Mao and A. K. Jain. Texture classification and segmentation using multiresolution simultaneous autoregressive models. *Pattern Recognition*, 25(2):173–188, 1992.

[56] H. R. Martin, editor. *Vibration Analysis*, volume 18 of *Friction, Lubrication and Wear Technology*. ASM Handbook, ASM International, 1992.

[57] M. Matsunaga, Y. Ito, and H. Kobayashi. Wear test of bucket teeth. *American Society of Mechanical Engineers*, pages 336–342, 1979.

[58] P. K. Mehrotra, editor. *Mechanisms of Wear in Ceramic Materials*. Wear of Materials. American Society of Mechanical Engineers, 1983.

[59] R. Mehrotra, K. R. Namuduri, and N. Ranganathan. Gabor filter-based edge detection. *Pattern Recognition*, 25(12):1479–1494, 1992.

[60] R. Navarro, A. Tabernero, and G. Cristobal. *Image Representation with Gabor Wavelets and Its Applications*, volume 97 of *Advances in Imaging and Electronic Physics*. Academic Press, 1996.

[61] S. K. Nayar, K. Ikeuchi, and T. Kanade. Surface reflection - physical and geometrical perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(7):611–634, 1991.

[62] M. Oren and S. K. Nayar. Generalization of the Lambertian model and implications for machine vision. *International Journal of Computer Vision*, 14:227–251, 1995.

[63] A. Papoulis. *Probability, Random Variables and Stochastic Processes*. McGraw-Hill, New York, 3rd edition, 1991.

[64] B. Phong. Illumination for computer generated pictures. *Communications of ACM*, 18:311–317, 1975.

[65] P. Podsiadlo and G. W. Stachowiak. The development of the modified hurst orientation transform for the characterization of surface topography of wear particles. *Tribology Letters*, 4(3-4):215, 1998.

[66] P. Podsiadlo and G. W. Stachowiak. Evaluation of boundary fractal methods for the characterization of wear particles. *Wear*, 217(1):24, 1998.

[67] P. Podsiadlo and G. W. Stachowiak. Scale-invariant analysis of wear particle surface morphology. I. theoretical background, computer implementation and technique testing. *Wear*, 242(1-2):160–179, 2000.

[68] P. Podsiadlo and G. W. Stachowiak. Scale-invariant analysis of wear particle surface morphology. II. fractal dimension. *Wear*, 242(1-2):180–188, 2000.

[69] P. Podsiadlo and G. W. Stachowiak. Scale-invariant analysis of wear particle surface morphology. III. pattern recognition. *Wear*, 242(1-2):189–201, 2000.

[70] B. Pugh. *Friction and Wear: A Tribology Textbook for Students.* Newnes-Butterworth, London, 1973.

[71] E. Rabinowicz. *Friction and Wear of Materials.* John Wiley and Sons, New York, 1965.

[72] T. Randen and J. H. Husøy. Filtering for texture classification: A comparative study. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(4):291–310, 1999.

[73] T. Randen and J. H. Husøy. Texture segmentation using filters with optimized energy separation. *IEEE Transactions on Image Processing*, 8(4):571–582, 1999.

[74] J. Rawers and J. Tylczak. Fractal characterization of wear-erosion surfaces. *Journal of Materials Engineering and Performance*, 8(6):669–676, 1999.

[75] T. R. Reed and J. M. H. du Buf. A review of recent texture segmentation and feature extraction techniques. *CVGIP: Image Understanding*, 57:359–372, 1993.

[76] L. M. F. Ribeiro, A. L. Horovistiz, G. A. Jesuino, L. R. de O Hein, N. P. Abbade, and S. J. Crnkovic. Fractal analysis of eroded surfaces by digital image processing. *Materials Letters*, 56(4):512–517, 2002.

[77] T. W. Ridler and S. Calvard. Picture thresholding using an iterative selection method. *IEEE Transactions on Systems, Man and Cybernetics*, (SMC-8):630–632, 1978.

[78] A. W. Ruff and P. Blau (Ed.), editors. *Wear Measurement*, volume 18 of *Friction, Lubrication and Wear Technology.* ASM Handbook, ASM International, 1992.

[79] M. J. H. Ruscoe, editor. *A Predictive Test for Coin Wear in Circulation*. Wear of Materials. American Society of Mechanical Engineers, 1987.

[80] J. C. Russ. *Fractal Surfaces*. Plenum Press, New York, 1994.

[81] A. C. Sanderson and L. E. Weiss. Structured highlight inspection of specular surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(1):44–55, 1988.

[82] J. Schofer and E. Santner. Quantitative wear analysis using atomic force microscopy. *Wear*, 222:74–83, 1998.

[83] Y. C. Tasan. *Measurement of Deformation in Rolling and Sliding Contacts*. PhD thesis, University of Twente, 2005.

[84] T. R. Thomas. *Rough Surfaces*. Imperial College Press, London, 2nd edition, 1999.

[85] K. E. Torrance and E. M. Sparrow. Theory for off-specular reflection from roughened surfaces. *Journal of the Optical Society of America*, 57(9):1105–1114, 1967.

[86] M. Tuceryan and A. K. Jain. *Texture Analysis*. Handbook Pattern Recognition and Computer Vision. World Scientific, Singapore, 1993.

[87] M. J. Turner, J. M. Blackledge, and P. R. Andrews. *Fractal Geometry in Digital Imaging*. Academic, San Diego; London, 1998.

[88] M. Unser. Local linear transforms for texture measurements. *Signal Process*, 11(1):61–79, 1986.

[89] F. van Heiden. *Image Based Measurement Systems*. John Wiley and Sons Ltd., 1994.

[90] R. Visser, A. van de Vrie, I. Wassink, and I. van Welbergen. Wear detection on metal plates. Technical report, University of Twente, the Netherlands, 2003.

[91] T. V. Vorburger and E. C. Teague. Optical techniques for online measurement of surface topography. *Precision Engineering*, 3(2):61–83, 1981.

[92] W. Wang, P. L. Wong, J. B. Luo, and Z. Zhang. A new optical technique for roughness measurement on moving surface. *Tribology International*, 31(5):281–287, 1998.

[93] D. Whitehouse. *Surfaces and Their Measurement*. Hermes Penton Science, London, 2002.

[94] P. L. Wong and K. Y. Li. In-process roughness measurement on moving surfaces. *Optics and Laser Technology*, 31:543–548, 1999.

[95] E. J. Wood. Applying fourier and associated transforms to pattern charactersization in textiles. *Textile Research Journal*, 60:212, 1990.

[96] R. J. Woodham. Photometric stereo: A reflectance map technique for determining surface orientation from image intensity. In *Proceedings of SPIE*, volume 155, pages 136–143, 1978.

[97] J. C. Wyant and K. Creath. Advances in interferometric optical profiling. *International Journal of Machine Tools and Manufacture*, 32(1/2):5–10, 1992.

[98] S. Yin, J. Li, and M. Song. Surface profile measurement using a unique microtube-based system. *Optics Communications*, 168(1-4):1–6, 1999.

[99] M. A. Younis. Online surface roughness measurements using image processing towards an adaptive control. *Computers Industrial Engineering*, 35(1-2):49–52, 1998.

[100] J. Zhang, M. Korsten, and P. P. L. Regtien. A vision system for online wear detection. In *XVII IMEKO World Congress*, pages 1960–1964, Dubrovnik, Croatia, 2003.

[101] J. Zhang, M. J. Korsten, and P. P. L. Regtien. Application of multichannel filtering for online wear detection. In *The Ninth Annual Conference of ASCI*, pages 144–150, Heijen, the Netherlands, 2003.

[102] J. Zhang, P. P. L. Regtien, and M. J. Korsten. Monitoring of dry sliding wear using fractal analysis. *Metrology and Measurement Systems*, XII(2):110–120, 2005.

[103] J. Zhang, P. P. L. Regtien, and M. J. Korsten. Monitoring of dry sliding wear using fractal analysis. In *10th IMEKO TC10 International Conference on Technical Diagnostics*, pages 133–138, Budapest, Hungary, 2005.

[104] B. Zitova and J. Flusser. Image registration methods: A survey. *Image and Vision Computing*, 21(11):977–1000, 2003.

Index

Acknowledgements

This thesis will have never been accomplished without the help from a lot of people. I would like to take this opportunity to thank all those who gave me support and encouragement, in one way or another, in the research as well as in my personal life during the past years at University of Twente, Enschede, the Netherlands.

First of all, I would like to give my best thanks to my promotor, prof. Paul Regtien. He made it possible for me to come to the Netherlands to do my PhD after an interesting interview held in January 2001. Doing a PhD in the Netherlands for me not only means a nice career opportunity abroad but also an easy way to get reunion with my wife, who was doing her TWAIO[1] at Technical University of Eindhoven, the Netherlands at that time. Thus Paul helped me solve the family reunion problem indirectly. And most importantly, he always gave me support with his expertise and patience in the research. Particularly, we had intensive discussions together on the draft thesis, and he also translated the summary into Dutch. Further, his consistent encouragement gave me much confidence to meet the challenges ahead.

Next, many thanks should be given to my supervisor, dr. Maarten Korsten. We had a lot of interesting and informative discussions together on the research project. He often pointed out the critical points of a problem very quickly, and his recommendations or remarks mostly helped me find better solutions to the problems. With his extensive knowledge of computer vision and image processing, Maarten helped me a lot in the research.

Besides Paul and Maarten, I would like to thank the other members of my graduation committee: prof. W. L. Dalmijn, prof. D. J. Schipper, prof. J. Meijer, and dr. F. van der Heijden. I really appreciate the time they have taken to read and comment on my thesis.

The research described in this thesis has been conducted as part of the project **TET. 5120**, which was sponsored by STW (the Dutch Technology Foundation). I appreciate very much all the members of the user committee of this project,

[1] TWAIO stands for *Tweejarige assistent in opleiding* in Dutch, a two-year postgraduate trainee program in the Netherlands.

especially dr. ir. M. B. de Rooij, dr. C. Tasan and dr. J. Jamari from Surface Technology and Tribology group for their collaboration, help, and constructive remarks.

Four students, Robert Visser, Anneke van de Vrie, Ingo Wassink and Herwin van Welbergen participated in my project. I am grateful to their participation and contribution.

I am also very grateful to the secretaries I have been with in the past years, Joan, Liesbeth and Carla. They helped me so much in supporting my work, stay, travel, as well as dealing with bureaucratic matters, for instance, visa, work permit, residence permit, etc. Without their help, surviving in the Netherlands would be very difficult for me.

The technical staffs, Alfred and Egbert were always there for help when I had any technical problems with my experimental setup, computers, whatsoever. I am grateful to their support.

I was so lucky to have two Romanian roommates, Nicolae and Valer (not at the same time but one after another). We share the same values on many issues and most time we have similar opinions on international politics, culture, bureaucracy, corruption, whatsoever. Valer even becomes my good friend and he brightened my life in Holland! Special thanks to them.

The other colleagues from the Measurement and Instrumentation group: Aditiya, Nanno, Zweitze, Jaap, Zoran, Ciska, Gloria and all the colleagues from the Control Engineering group made the wonderful time for me. Thanks a lot.

It would be never enough to express my gratitude to my wife, Yuanqing Guo. Without her endless love and consistent support many things in my life, including this dissertation, would have never happened.

I should not forget to mention and thank the *VIP*, Andy, my son. He is the best *result* I achieved in the past years in the Netherlands. He has been making me see many more beautiful colors of life as a parent and inspiring me a lot in his unique way since his birth.

All my family members gave me a lot of support in my life. Particularly, my mother-in-law came to the Netherlands from China and stayed with us for three months to help look after the newborn baby. Many thanks to all of them.

Last but not the least, I would like to acknowledge all of my Chinese and international friends. Their friendship makes my life in the Netherlands enjoyable and full of fun. I am not going to write down their names here as the list will be too long and it is very likely to miss someone. Thank all of you, my friends!

Jianbo Zhang

July 2006, Enschede