THE UNIVERSITY OF
WARWICK

University of Warwick institutional repository: http://go.warwick.ac.uk/wrap

A Thesis Submitted for the Degree of PhD at the University of Warwick

http://go.warwick.ac.uk/wrap/76203

This thesis is made available online and is protected by original copyright.

Please scroll down to view the document itself.

Please refer to the repository record for this item for information to help you to cite it. Our policy information is available from the repository home page.

Modeling of Probe-Surface Interactions in Nanotopographic Measurements

By

Khalid Thamer Althagafy

University of Warwick

School of Engineering

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Engineering

September 2015

Table of Contents

Chapter 1 .... **1**

1. Introduction **1**

1.1. Overview 1

1.2. Research Questions and Aims and Objectives 4

1.3. Contribution 4

1.4. Thesis Organisation 7

1.5. Publications and Awards 8

Chapter 2 .... **9**

2. Literature review **9**

2.1. Introduction 9

2.2. Surface Metrology 11

2.3. Instrumentation Development 13

2.3.1 Surface Characterization 16

2.3.2 Statistical Analysis 19

2.3.3 Random Process Methods 21

2.4. Modern Instrumentation Developments 21

2.4.1 Contact Method 21

2.4.2 Non-contact Methods 27

2.5. Comparison between Instruments 36

2.6. Summary 36

Chapter 3 .... **38**

3. History of Stylus simulation **38**

3.1. Introduction 38

3.2. Stylus Tip Convolution 38

3.2.1. "Noise" in the Measurement: 41

3.3. Filtering 44

3.3.1. Development history 46

3.3.2 Electrical Filter systems 46
3.3.3 Envelope Filters system 46
3.3.4 Morphological Filters systems 48
3.4. Errors in using contact-probe/stylus surface measurements 50
3.4.1 Briefing on stylus 50
3.4.2 Noise and Vibration 51
3.4.3 Stylus errors and distortion 52
Chapter 4 55
4. Baseline software tool for stylus contact simulation 55
4.1. Introduction 55
4.2. Kinematic and Threshold Models 57
4.3. Atomic Force Microscopy (AFM) 61
4.3.1. AFM Scan Parameters 65
4.3.2. Approaching the Sample 66
4.3.3. Fourier Transforming Images 67
4.3.4. Removing Tilt and Low Frequency Interference 67
4.4. Numerical Simulation Process 70
4.4.1. Simulation software technique 74
4.5. Summary 82
Chapter 5 84
5. Sub-Micrometer contact of topography Styli 84
5.1. Introduction 84
5.2. Verification 89
5.3. Kinematic and Threshold Models 90
Chapter 6 99
6. Stylus tip size and shape effects in surface contact profilometry 99
6.1. Stylus Simulation 99
6.2. Styli and fine surface structure 100

6.3. Results 102

6.3.1. Verification 102

6.3.2. Simulation on real surfaces 105

6.3.3. Results discussion 106

6.4. Summary 120

Chapter 7 121

7. Implications for stylus profilometer uncertainty 121

7.1. Introduction 121

7.2. Testing the threshold models 123

7.3. Surface interpolation using 'fractal infill' methods 140

Chapter 8 154

8. Conclusion and further work 154

9. References 158

10. Appendix 164

10.1. Appendix 1: Simulation software technique (Extra explanation) 164

10.2. Appendix 2: Extra Results for Chapter 7 173

10.3. Appendix 3: Matlab Codes 179

A3.1 The main Matlab Code: 179

A3.2 The second Matlab Code: 197

A3.3 The third Matlab Code: 200

A3.4 The fourth Matlab Code: 201

10.4. Appendix 4: SPIP Software information 205

List of Figures

Figure 2.1 Surface measurement stages (Upper: Instrumentation, middle: decomposition & filtration and Lower: parameterisation & characterisation 13

Figure 2.2: simple RC-filter (low pass filter)..... 15

Figure 2.3: Some numerical parameters to define surface profile 18

Figure 2.4: Principle subsystems of a stylus instrument..... 24

Figure 2.5: Sinusoidal surface with a stylus [Elson and Bennett (1979)]..... 25

Figure 2.6: Principle of interaction between tip and sample..... 26

Figure 2.7; Laser triangulation method [adapted from (Groover, 2007)]..... 33

Figure 2.8: Principle of light scattering method 34

Figure 2.9; Schematic of confocal microscopy..... 35

Figure 3.1: The influence of vibration in physical versus mathematical tip convolution. 41

Figure 3.2 The presence of a "lateral" peak in trace 3 relative to trace 2. 43

Figure 3.3: Long and Short wavelength components diagram 45

Figure 3.4: Envelope filters method, locus of the ball's centre when traverses the surface (Adapted from: Muralikrishnan and Raja, 2009)..... 47

Figure 3.5: Closing and opening envelope of an open profile by a disk (Scott, 2000)..... 49

Figure 3.6: Stylus shapes: cone (left), pyramid (right) 50

Figure 4.1: Contact points of the stylus with rough surface 56

Figure 4.2: Scheme for the software suite..... 58

Figure 4.3: AFM (right), Probe holder (left) (Quesant, Q-Scope™ 788)..... 63

Figure 4.4: Probe Cantilever (right), Probe Tip enlarged in (left) 64

Figure 4.5: Samples used in this project..... 68

Figure 4.6: 3D lapped surface profile by AFM (Sample 1) 68

Figure 4.7: 3D Ground surface profile by AFM (sample 2A) 69

Figure 4.8: 3D Ground surface profile by AFM (sample 2B)..... 69

Figure 4.9: The effect of stylus tip radius on the measured surface profile..... 70

Figure 4.10 The principle of the simulation process of measuring a surface with a stylus 73

Figure 4.11:Summary of the numerical simulation algorithm..... 76

Figure 5.1: Scheme for the software suite. 88

Figure 5.2 Scanning across a noise-free sinusoidal surface..... 89

Figure 5.3: Modification of a surface detail by a 2 mm flat: (a) AFM image; (b) simulation..... 90

Figure 5.4: Positions of kinematic and 50 nm threshold contacts on a flat probe for three selected places from Figure 5.3b........ 92
Figure 5.5: Conical shape 60x60µm 93
Figure 5.6: Map of contact points with 149 nm........ 94
Figure 5.7: (a) Conical shape Stylus 10x10 µm (b) Map of contact points with 23nm threshold........ 95
Figure 5.8: (a) 3D 20x20 µm small map (AFM) image........ 96
Figure 5.9: 3D ground surface 10x10 µm (b) Map of contact with 149 nm threshold (c) with 40 nm threshold 97
Figure 5.103D Ground surface 8X8µm with tip at the centre. (b) Map of contact with 149 nm threshold........ 98
Figure 6.1: (5µm) Conical Stylus tip shape 101
Figure 6.2: (5µm) Spherical Stylus tip shape........ 101
Figure 6.3: (5µm) Perfect Pyramid Stylus tip shape........ 102
Figure 6.4(a) Locus & original surface (b) 2D cross section profile (c) Contacts distribution: of a 10µm conical tip on a random surface 103
Figure 6.5:(a) Locus and original surface (b) 2D cross section profile (c) Contacts distribution: of a 10µm conical tip on a random surface 104
Figure 6.6: AFM 3D image of a ground steel surface........ 105
Figure 6.7: AFM 2D profile of a ground steel surface........ 106
Figure 6.8: Original ground surface (Bottom) and perfect pyramid stylus tip locus (top)........ 107
Figure 6.9: 2D cross section of original ground surface profile and the traced pyramid stylus locus........ 108
Figure 6.10: Contact distribution on the perfect pyramid tip........ 108
Figure 6.11: Original ground surface (Bottom) and perfect Hemisphere stylus tip locus (top)........ 109
Figure 6.12: 2D cross section of original ground surface profile and the traced Hemisphere stylus locus 110
Figure 6.13: Contact distribution on the perfect hemisphere tip........ 110
Figure 6.14: Original ground surface (Bottom) and perfect conical stylus tip locus (top) 111
Figure 6.15: 2D cross section of original ground surface profile and the traced conical stylus locus........ 112
Figure 6.16: Contact distribution on the perfect conical tip........ 112

Figure 6.17: Original ground surface (Bottom) and truncated Hemisphere stylus tip locus (top) 113
Figure 6.18: Contact distribution on the truncated Hemisphere tip 113
Figure 6.19: Original ground surface (Bottom) and truncated conical stylus tip locus (top) 114
Figure 6.20: Original ground surface (Bottom) and truncated pyramid stylus tip locus (top) 115
Figure 6.21: Contact distribution on the truncated conical tip 115
Figure 6.22: Contact distribution on the truncated pyramid tip 116
Figure 7.1: Ground surface with Cone stylus (Sample 1) 124
Figure 7.2: Map of contact with 5nm threshold (Sample 1) 125
Figure 7.3: Map of contact with 20 nm threshold (Sample 1) 125
Figure 7.4; Map of contact with 5nm threshold (Sample 1) 126
Figure 7.5: Ground surface with Cone stylus (Sample 2) 126
Figure 7.6: Map of contact with 5nm threshold (Sample 3) 127
Figure 7.7: Map of contact with 20nm threshold (Sample 3) 127
Figure 7.8: Map of contact with 50nm threshold (Sample 3) 128
Figure 7.9: Three-D Ground surface with Cone stylus (Sample 4 - with Noise) 128
Figure 7.10: Map of contact with 5nm threshold (Sample 4) 129
Figure 7.11: Map of contact with 20nm threshold (Sample 4) 129
Figure 7.12: Map of contact with 50nm threshold (Sample 4) 130
Figure 7.13: Threshold Vs. Number of contacts for Conical stylus 132
Figure 7.14: Threshold Vs. Number of contacts for spherical stylus 133
Figure 7.15: Threshold Vs. Number of contacts for Flat stylus 134
Figure 7.16: Threshold Vs. Number of contacts for surface (sample 1) 135
Figure 7.17: Threshold Vs. Number of contacts for surface (sample 2) 135
Figure 7.18: Threshold Vs. Number of contacts for surface (sample 3) 136
Figure 7.19: Threshold vs. Number of contacts for surface (sample 4 -Noise) 136
Figure 7.20: Hemisphere stylus interact with rough surface 137
Figure 7.21: Map of contact points under kinematic condition (20 nm threshold) 138
Figure 7.22: Ground original surface (Zorig) using Cone stylus 143
Figure 7.23: Map of contact with 5nm threshold (Zorig) 143
Figure 7.24: Map of contact with 20 nm threshold (Zorig) 144
Figure 7.25: Map of contact with 50nm threshold (Zorig) 144
Figure 7.26: Sub-sampled Ground surface (Z) using sphere stylus 145

Figure 7.27: Map of contact with 5nm threshold (Z).. 145
Figure 7.28: Map of contact with 20 nm threshold (Z).. 146
Figure 7.29: Map of contact with 50 nm threshold (Zg).. 146
Figure 7.30: expanded "Patch" surface (Zexpp) using sphere stylus............................. 147
Figure 7.31: Map of contact with 5nm threshold (Zexpp)... 147
Figure 7.32: Map of contact with 20nm threshold (Zexpp).. 148
Figure 7.33: Map of contact with 50nm threshold (Zexpp).. 148
Figure 7.34: Expanded "Patch+filler" surface (Zexpf) using sphere stylus.................... 149
Figure 7.35: Map of contact with 5nm threshold (Zexpf)... 149
Figure 7.36: Map of contact with 20nm threshold (Zexpf)... 150
Figure 7.37: Map of contact with 50nm threshold (Zexpf)... 150
Figure 7.38: Threshold Vs. Number of contacts using spherical stylus.......................... 151

List of Tables

Table 4.1: AFM Scan parameters.......... 65
Table 6.1: Real surface roughness parameter 106
Table 6.2: Roughness parameters Errors 117
Table 7.1: Example of illustrative results for Steel, Brass and Aluminium.......... 121
Table 7.2: Kinematic/threshold contact points statistics for cone stylus 131
Table 7.3: Kinematic/threshold contact points statistics for sphere stylus 133
Table 7.4: Kinematic/threshold contact points statistics for flat stylus 134
Table 7.5: Kinematic/threshold contact points statistics by using sphere stylus 151
Table 7.6: Surface roughness parameters.......... 152

Abbreviations

AFM	Atomic force microscope
CCD	Charge-coupled device
CSI	Coherence scanning interferometry
E-system	Envelope-filter
F-operator	Form operator
ISO	International organisation for standardisation
L-filter	Large scale filter
LVDT	Linear variable differential transformer
M-system	Mean-line filter
NPL	National physics laboratory
OPD	Optical path difference
PC	Personal computers
PSI	Phase shifting interferometry
PZT	Piezoelectric translator
Ra	Roughness average
CLA	Roughness average = Ra = AA
Rq	Root mean square value corresponding to Ra
Sm	The mean spacing between profile peaks at the mean line
S-filter	Small scale filter
SPM	Scanning probe microscope
STM	Scanning tunnelling microscopes
WSI	Wavelength scanning interferometry
SPIP	Scanning probe image processor
CL	Actual height of the stylus at any point
Rku	Surface Kurtosis “peakedness” of the surface topography
Ry	Peak-Peak Height

Rsk	Surface Skewness, irregularity of the height distribution histogram
Zorig	A Square master array of original data
Zexp	an expanded array of same size of Zorig
Zexpp	Expanded array patch
Zexpf	Expanded array patch+ filler.

Acknowledgements

All thanks and praise are due to God the Almighty for his blessing that made this work possible and completed.

I would like to express my sincere thanks to my supervisor Professor Derek Chetwynd for firstly providing the opportunity and facilities to undertake this research, secondly for his continued guidance and encouragement throughout this project thesis. His valuable comments, suggestions and feedback during all the phases of this work were indispensable.

Special thanks must also go to Mr David Robinson and Mr Martin Davis whose enthusiasm and technical guidance have been greatly appreciated, and to Dr. Mujthaba Ahtamad for proofreading my thesis.

I would also like to thank my friend Ahmad Almakky for all his help and support. My thanks also extended to Dr Helmy Dowidar and Andy Huang for their help.

I would also like to express my gratitude to Umm AL-Qura University (UQU), for awarding me a scholarship to continue my PhD.

Last, by no means least, my special thanks my wife (Rahmah Alghamdi) and my children (Thamer, Smoo and Sama). Without your endless support and unlimited patience, I could never have completed this work.

This work is dedicated to my parents for enduring my absence while I was working on this project.

Declaration

I herewith declare that this draft thesis contains my own research performed under the supervision of Professor Derek Chetwynd, without assistance of third parties, unless stated otherwise. No part of this thesis results were previously submitted for a degree at any other university unless as publication under my name.

Abstract

Contact stylus methods remain important tools in surface roughness measurement, but as metrological capability increases there is growing need for better understanding of the complex interactions between a stylus tip and a surface. For example, questions arise about the smallest scales of topographic features that can be described with acceptable uncertainty, or about how to compare results taken with different types of probe. This thesis uses simulation methods to address some aspects of this challenge. A new modelling and simulation program has been developed and used to examine the measuring of the fine structure of the real and simulated surfaces by the stylus method. Although able to scan any arbitrary surface with any arbitrary stylus shape, the majority of the results given here uses idealized stylus shapes and 'real' ground steel surfaces. The simulation is not only used to measure the roughness of the surface but also to show the contacts distribution on the tip when scanning a surface. Surface maps of the fine structure of ground steel surfaces were measured by Atomic Force Microscopy (AFM) to ensure high lateral resolution compared to the capability of the target profilometry instruments. The data collected by the AFM were checked for missing data and interpolated by the scanning probe image processor (SPIP) software. Three basic computer generated stylus tips with different shapes have been used: conical, pyramid and spherical shapes. This work proposes and explores in detail the novel concept of "thresholding" as an adjunct to kinematic contact modelling; the tip is incremented downwards 'into' the surface and resulting contact regions (or islands) compared to the position of the initial kinematic contact. Essentially the research questions have been inquiring into the effectiveness of so-called kinematic contact models by modifying them in various ways and judging whether significantly different results arise. Initial evidence shows that examination of the contact patterns as the threshold increases can identify the intensity with which different asperity regions interact with the stylus. In the context of sections of the ground surface with a total height variation in the order of 500 nm to 1 μm, for example, a 5 nm threshold caused little change in contact sizes from the kinematic point, but 50 nm caused them to grow asymmetrically, eventually picking out the major structures of the surface. The simulations have naturally confirmed that the stylus geometry and size can have a significant effect on most roughness parameters of the measured surface in 3D. Therefore the major contribution is an investigation of the inherent (finite probe) distortions during topographic analysis using a stylus-based instrument. The surprising finding which is worthy of greater investigation, is how insensitive to major changes in stylus condition some of the popular parameters are, even when dealing with very fine structure within localized areas of a ground surface. For these reasons, it is concluded that thresholding is not likely to become a major tool in analysis, although it can certainly be argued that it retains some practical value as a diagnostic of the measurement process. This research will ultimately allow better inter-comparison between measurements from different instruments by allowing a 'software translator' between them. Short of fully realizing this ambitious aim, the study also contributes to improving uncertainty models for stylus instruments.

Chapter 1
1. Introduction

1.1. Overview

Despite the strong reasons for continuing to use contact stylus measurements of micro-topography, we must recognise that they will impose highly non-linear and practically undecipherable distortions into the measured geometry whenever the effective radius of the tip becomes significant compared to the size of the smallest features of interest on the surface being measured. While interest in smaller and smaller features continues to grow, it is not possible to similarly reduce the stylus tip size and retain a robust instrument. Thus, several concerns need to be addressed with some urgency. At its most basic, we wish to have good, reliable guidelines for when it is safe, in the sense that the distortions introduced are tolerable, to use a particular type of tip for a particular type and scale of topography. It must be recognised that there are increasing risks that an inadequate tip on an existing instrument will be used unthinkingly as the technical demands increase, leading to dangerously inaccurate measurements. Thus, the idea of discovering practical ways by which instruments might be able to warn of (if not strictly self-diagnose) potentially misleading results is very attractive. Extending this idea, we might even consider diagnosing the increased 'radius' of a worn tip. There is also increasing concern within International Standards over the lateral uncertainty in topographic measurements, expression of which depends directly on knowledge of how stylus tips interact with real topographies. Associated with, but distinct from, these matters of measurement uncertainty, there is the question of predicting method divergence between different types of topographic instrument. An understanding of how a probe interacts with a surface is clearly crucial to any attempt to model

how the results from one type of instrument relate to those from another type. This thesis aims to contribute to the longer term solution of these challenges.

As will become obvious from the reviews included here, it is almost impossible to gain precise experimental information about the contact between a stylus and a rough surface at the micrometre and smaller scales. There is no way to incorporate sensors into the contact region that is actually practicable in terms of precision, robustness or cost. The only plausible approach, certainly in the early stages of the research, for ranking rival methods and for setting the scope of what will ultimately be needed is to use computer models validated to some extent by their producing similar gross effects to those seen on real instruments. Hence, this work considers only numerical experiments based around new extensions of modelling ideas already demonstrated to have some physical relevance. Its starting point is the so-called 'kinematic model', a purely geometrical consideration of the locus of a ball rolling on a surface, or of how a suitably constrained stylus comes into initial contact with a rough surface. However, it is now essential to collect data on how the physical contact points relate to the positions around the stylus, not just to provide a map of how a modified surface might look, and to provide a software implementation that enables systematic study of it. From this might arise various inferences about lateral (and vertical) uncertainty and the sensitivity to changes in the nominal size and shape of a stylus. It is perhaps intuitively obvious, but will anyway be explored in some detail later, that the idealised 'first contact' might be misleading. For example, it could be vulnerable to quite small amounts of instrument noise. Even if it were not, a very narrow high-point could be selected that would be of virtually no functional significance. At this point it is necessary

to recall that there is longer-term interest in diagnostic processes that could run in close to real time on the limited computational power available within a commercial instrument. This provides a strong motivation for exploring algorithmically straightforward methods for responding to the potential oversensitivity of the kinematic model. It appears that even such basic approaches as manipulating the vertical quantisation, or applying threshold values, has not been studied in this context. We might, for example, ask questions such as if the stylus is pushed some distance into the surface, what size and position of contact would arise, in one or more 'islands' of continuous lateral extent? This raises yet more questions about whether such methods might have value as post-process diagnoses of the fitness of the probe for the measurement made. In effect, this would require running some type of stylus-surface simulation based on the actual instrument output and attempting to infer whether a reasonable model of the actual stylus would affect the data. Almost inevitably, this seems to demand data at a higher lateral resolution than is actually available and so a fast, not too approximate, interpolation process would be needed. It might be possible to exploit the common idea that surfaces show fractal-like behaviour over a reasonable bandwidth.

1.2. Research Questions and Aims and Objectives

Given the complexity of real surface data, and even more so if we allow for non-ideal tip geometries, there is no way to know how effective any such approaches might be other than to try them. This is the spirit in which this thesis approaches the challenges just discussed. So, its major objectives are to:

- Generate sophisticated and validated models for the interaction between engineering rough surfaces and contact probe in order to guide the art of the topographic instrument or high-tech surface.
- Review the basic kinematic model and suggest how statistically richer data can be obtained. In addition to a review of early models reporting 3D stylus simulations and their effect in a series of basic numerical experiments. Also to ensure that the correct and anticipated metrology needs are adopted from the available academic literature.
- Present a computational study of kinematic models in precision of real measurement noise.
- Investigate "threshold" methods that might lead to relaxation approaches and account for material properties (Non-Kinematic models).
- Provide a comprehensive study of measurement uncertainty in Nano-topograhic metrology.

1.3. Contribution

The summary of my major contribution is an investigation of the inherent (finite probe) distortions during topographic analysis using stylus-based instruments. While continually aware of physical systems, the main study is purely in modelling and simulation. Essentially, this work investigate about the effectiveness of so-

called kinematic contact models by modifying them in various ways and judging whether significantly different results arise. The motivation is ultimately to allow better inter-comparison between measurements from different instruments by allowing a ‘software translator’ between them. Short of fully realizing this ambitious aim, the study also contributes to improving uncertainty models for stylus instruments.

A new modelling and simulation program has been developed and used to examine the measuring of fine structure of the real and simulated surfaces by the stylus method. Although able to scan any arbitrary surface with any arbitrary stylus shape, the majority of the results given here uses idealized stylus shapes and ‘real’ ground steel surfaces. The simulation is not only used to measure the roughness of the surface but also to show the contacts distribution on the tip when scanning a surface. Surface maps of the fine structure of ground steel surfaces were measured by Atomic Force Microscopy (AFM) to ensure high lateral resolution compared to the capability of the target profilometry instruments. The data collected by the AFM were checked for missing data and interpolated by the scanning probe image processor (SPIP) software. Three basic computer generated stylus tips with different shapes have been used: conical, pyramid and spherical shapes. Various dimensions have been compared, ‘imperfection’ introduced and so on.

This work proposes and explores in detail the novel concept of “thresholding” as an adjunct to kinematic contact modelling; the tip is incremented downwards 'into' the surface and resulting contact regions (or islands) compared to the position of the initial kinematic contact. This work is the first to implement a version of kinematic stylus contact simulation that includes full geometric recording of

behaviour at the stylus tip; this will interest other specialists in the field because it is a step towards more complete uncertainty models. The ‘threshold’ concept was also new in this context and was an idea worth practical investigation to see if it has the sensitivity to be useful as a self-diagnostic tool within instruments.

Initial evidence shows that examination of the contact patterns as the threshold increases can identify the intensity with which different asperity regions interact with the stylus. In the context of sections of the ground surface with total height variation in the order of 500 nm to 1 μm, for example, a 5 nm threshold caused little change in contact size from the kinematic point, but 50 nm caused them to grow asymmetrically, eventually picking out the major structures of the surface.

The simulations have naturally confirmed that stylus geometry and size can have a significant effect on most roughness parameters of the measured surface in 3D. The surprising feature of these, worthy of greater investigation, is how insensitive to major changes in stylus condition some of the popular parameters are, even when dealing with very fine structure within localized areas of a ground surface. For these reasons, it is concluded that thresholding is not likely to become a major tool in analysis, although can be argued that it retains some practical value as a diagnostic of the measurement process.

Finally, the method works without gross errors, but is clearly approximate and probably generates distortions rather larger than would be wise for a diagnostic - given expected further increase in cost-speed effectiveness of the computers that will go on future instruments, it may be best, on the evidence of results here, not to

further pursue this level of approach but to revisit the general idea in a few years using more sophisticated algorithms.

1.4. Thesis Organisation

The discussions of the ideas, implementations, simulation experiments and the subsequent evaluation of methods and interpretation of results are set out as follows. Following this introduction, Chapter 2 provides essential background and critique about the past development and modern demands in surface metrology, including a little on the methods that will be used later to obtain data for the simulations. Chapter 3 then introduces more specifically the background relevant to stylus simulation upon which the new work will be built. Chapter 4 discusses more technical details of the simulation method and its practical implementation and also a baseline software tool for stylus contact Simulation. Chapter 5 presents refinements to 3D simulations of stylus effects in microtopography measurements. It briefly reviews how statistically richer data can be obtained by extending basic kinematic models, perhaps providing steps towards more sophisticated modelling of the contact process. The results concentrate on idealized and non-idealized styli operating beyond the limit of their expected resolving power.

Chapter 6 introduced the new and novel implementation of a suite of tools for studying stylus-surface interaction in simulation. A consistent pattern of data flows and archiving around a highly modular organization is advocated, allowing new features to be introduced in small, independently testable, modules. In this respect, the first features included and not previously reported concern the mapping of contact extent and intensity across the probe surface and the use of thresholding to investigate sensitivities to noise and instrument resolution.

Chapter 7 introduced more implications for stylus profilometer uncertainty and a further set of carefully integrated numerical experiments by using the threshold idea. Additionally, it give more details of how can we interpolate surface data in order to work with higher resolution stylus images, and demonstrate the general validity of an interpolation method that makes some physical sense in the context of surface metrology and tribology. Finally, chapter 8 gives the final conclusion and proposal for recommendation and further work.

1.5. Publications and Awards

Khalid T. Althagafy, Chetwynd D. G. Simulation OF Stylus Contact Patterns in Profilometry, Styli. Proc. 26th ASPE Annual Meeting, Denver, US, November 2011.

Khalid T. Althagafy, Chetwynd D. G. Investigations of Probe-Surface Interactions in Nanotopographic Measurements. 6th SIC international conference, London, UK, 2012.

Khalid T. Althagafy, Chetwynd D. G. Simulation Studies of Sub- Micrometre Contact of Topography, Styli. Proc. 27th ASPE Annual Meeting, San Diageo, US, 2012

Khalid T. Althagafy, Chetwynd D. G. Investigation of stylus tip-size effects in surface contact profilometry. Proceedings of the 13th Euspen International Conference – Berlin– May 2013

Awards:

1. Heidenhain_GMBH scholar award, European Society for Precision Engineering & Eanotechnology (EUSPEN), "In recognition of the strong academic accomplishment in a field related to Nano and Micro Metrology". Germany, 2013
2. Heidenhain_GMBH scholar award, European Society for Precision Engineering & Eanotechnology (EUSPEN), "In recognition of the strong academic accomplishment in a field related to Precision and Micro Engineering", Germany, 2013

Chapter 2
2. Literature review

2.1. Introduction

Industrial, commercial and academic interest in measuring surface roughness continues to increase for several reasons. Efficiency, lubrication, corrosion stability and probability of crack formation are examples of where the roughness of produced surfaces influence their performance, although often there is only an empirical understanding of the relationship. Roughness is also important to optics and to the perception of consumer products. A process control measurement is often used in manufacturing, even when its functional significance is less critical to the final product.

The contact of a small probe-tip against a surface is crucial to stylus-based surface roughness metrology, which remains a popular and important practical technique for both industry use and basic research. A finite stylus reduces measurement fidelity through direct geometrical effects, elastic distortion and often by interaction with dynamic effects that vary the contact force, sometimes to lift-off.

It is difficult to determine precisely the complex interactions that occur between a profilometer stylus and a surface. As will be discussed it is virtually impossible to directly instrument sub-micrometre scale contacts and analytic models are little help for pheromones that are not merely non-linear but slope-discontinuous. This thesis therefore focuses on simulation studies intended to improve metrological insights into how small contacts develop on rough surfaces (although it might well have other micro-tribology applications).

This study reports recent progress on extended studies into the contact regimes of surface profilometer styli, with a view especially to better understanding the uncertainty budgets. It simulates the 3D contact of an approaching (ideal or real) stylus. Interest here is specifically with behaviour at the limits of instrument capability and is studied by using AFM-derived data to represent small areas of engineering surfaces at resolutions to 1 nm vertically and 200 nm laterally. Typical conventional profilometers have stylus tips with characteristic dimensions of the order of micrometres and noise levels of tens of nanometres. Extending upon conventional kinematic simulations, its novelties include (i) detailed analysis of how contact points or patches scatter with respect to the normal axis and (ii) a "threshold" function for exploring numerical and physical sensitivity to very small surface features (potentially noise).

Discussion will summarize the main features of the software implementation (using the commercial package MATLAB) and the rationale for statistical parameters to summarize contact spatial distribution. The simplest form of thresholding is roughly equivalent to ideally pressing the probe into the surface. For the illustrations used here the modified surfaces are largely unchanged other than in terms of a slight vertical offset. Interest lies in how the shift interacts with additional surface features to alter the probe contact patterns.

This presentation reports recent progress on extended studies into the contact regimes of surface profilometer styli, with a view especially to better understanding the uncertainty budgets. It simulates the 3D contact of an approaching (ideal or real) stylus. Behaviour at the limits of instrument capability is studied by using AFM-

derived data to represent small areas of engineering surfaces at resolutions to 1 nm vertically and 200 nm laterally. Extending upon conventional kinematic simulations, its novelties include (i) detailed analysis of how contact points or patches scatter with respect to the normal axis and (ii) a “threshold” function for exploring numerical and physical sensitivity to very small surface features (potentially noise).

The discussion will summarise the main features of the software implementation (using commercial packages MATLAB and SPIP), the rationale for statistical parameters to summarize contact spatial distribution, and the use of these parameters to demonstrate the in-plane uncertainty to be expected when scanning fine ground surfaces with a typical (2 µm) stylus.

2.2. Surface Metrology

The science of measuring both the deviation of a workpiece from its ideal shape and the machining marks left by a machining tool on the workpiece surface is a good definition for surface metrology. Jiang et al. (2007a, p. 2049) defined surface metrology as “the science of measuring small-scale geometrical features on surfaces: the topography of the surface”. The geometrical features of a surface can be classified as Form, Waviness and Roughness (Whitehouse, 2011). Form is very long waves caused by error in slide ways. Waviness is irregularities of a longer wavelength caused by errors in the path of the tool due to, for example, vibration between the workpiece and a grinding wheel on the tool path. Roughness is irregularities often caused by the manufacturing process including the impression left by grinding or polishing.

The waviness and roughness features are known as surface texture, and the measurement of surface texture can be used to control the manufacture and predict the functional performance of the workpiece (Jiang and Whitehouse, 2012). For example, the surface roughness can monitor tool wear condition and predict the failure time of contact surfaces.

The surface roughness (for example Ra or Rq) is sensitive to the cutting time under different cooling / lubrication conditions due to the running-in process of cutting tools (Jiang et al., 2010) where Ra is the modulus of arithmetic mean of the magnitude of the deviation of the profile from the mean.

Measuring the surface texture of thin films is also vital in roll-to-roll manufacturing processes. It can characterise the defects in the films, which are directly related to the film performance and lifespan.

There are three stages for surface measurement (see Figure 2.1). First is acquisition of information about surface topography by instrumentation. Secondly, decomposition of this information into primary features (waviness, roughness and form) by filtration. Finally, there is characterisation of the surface by parameterisation.

There have been several fairly recent publications that give good coverage of the early and mainstream development of ideas and techniques for surface feature measurement.

This thesis therefore restricts discussion of such issues to a brief summary of issues directly influencing the research approach used; it generally, therefore, does not cite original sources but recommends the more accessible recent reviews.

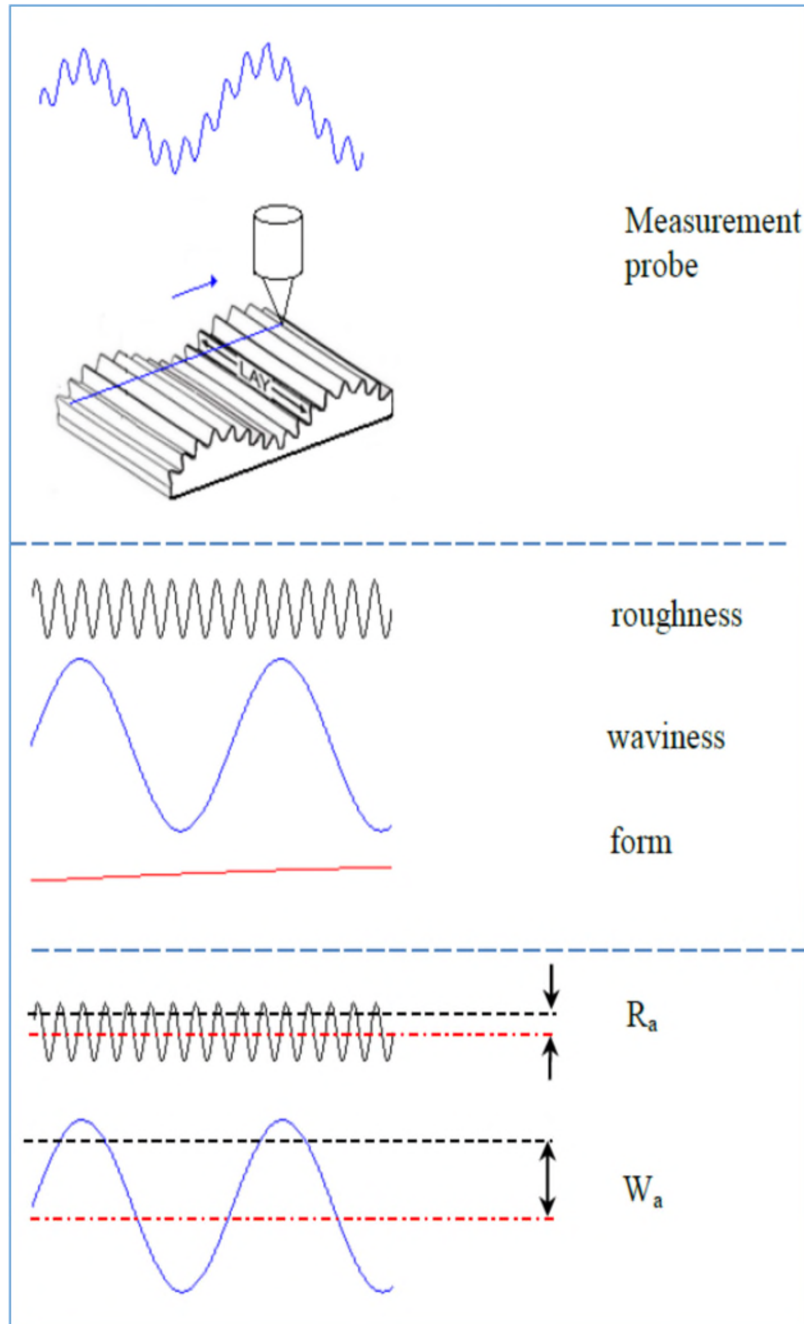


Figure 2.1 Surface measurement stages (Upper: Instrumentation, middle: decomposition & filtration and Lower: parameterisation & characterisation

2.3. Instrumentation Development

Traditionally, in the past, surfaces were assessed by either eye or thumbnail. Tomlinson at the National Physics Laboratory (NPL) implemented the first contact instrument as long ago as 1919 to measure the height of the roughness (Buice, 2007). However, the first practical instruments to measure axial profile were not

designed until the 1930s using a stylus-based method. Carl Zeiss designed the first non-contact optical instrument in Germany in1934 (Whitehouse, 2011).

In order to extract from the measured data the geometrical features of the surface that are of primary (functional) concern in a particular situation, filters with different wavelength cut-off boundaries have been developed. The first attempts at separating waviness from roughness introduced two filter types; M-system (mean-line filter) and E-system (envelope filter). Both were conceived of as generating a "reference line" representing the local trend of the workpiece surface, with respect to which other roughness features could then be described .The M-system and E-system were an electrical RC filter (low frequency pass) and a mechanical rolling circle filter respectively.

In 1963, the phase distortion produced by the single RC filter was corrected by adding a second RC circuit. The buffered 2CR (high pass to give roughness signal relative to the reference line) was introduced earlier, giving a sharper cut-off, but retaining phase distortion.

This filter was widely used in the 1960s and 70s especially after the filter was digitised at the turn of 1970. DJW 1986 (IMechE) is the real start for a 'phase corrected' version, which cannot be implemented in real time or by simple analogue circuitry.

They can, though, readily be emulated in digital computation. They only therefore become economically practicable on standard instruments from the mid-1980s.

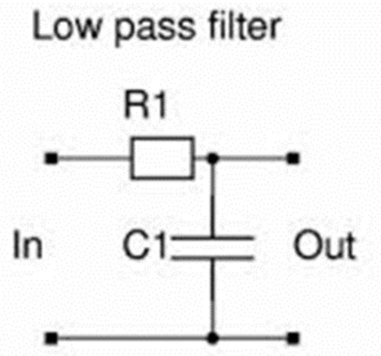


Figure 2.2: simple RC-filter (low pass filter)

Even a fully "zero-phase" digital filter with a 2RC characteristic showed some undesirable characteristics in terms of discriminating features of functional significance (and because different ways of implementing it could lead to marked inconsistencies). This was overcome in 1986 by introducing a Gaussian filter as a standardised phase-corrected profile filter. Further development led to an ISO profile filtration toolbox which consists of five classes namely; linear filters, morphological filters, robust filters and segmentation filters (Jiang et al.,2007a). For area filtration, the scale-limited concept is now considered. This concept is based on dividing the surface texture wavelengths into three scales namely: S-scale (contains small-scale lateral component), L-Scale (contains large-scale component), and form (contains the nominal form). Each scale has its own filter: the S-filter, L-filter and F-operator respectively. The area filtration is achieved by some combination of these filters. For example, to filter out the unwanted small-scale measurement noise and the nominal form, an (SF) filter can be used. Filters such as those utilising wavelets are still under development as modern manufacturing starts turning from traditional simple geometries toward free-form geometries.

2.3.1 Surface Characterization

The need to characterise surface topography was realised from the early pioneering days of the development of measuring instruments. The birth of statistical methods could be attributed to Abbott and Firestone when they proposed the use of the bearing area curve (Abbott and Firestone, 1933). The bearing area curve is based upon considering the contact properties of the surface; however it can be shown that it is equivalent to the cumulative probability distribution of surface heights. The most common parameter is the centre line average (Ra) value, and is easily obtained by direct reading of a meter on most stylus based measuring instruments. Thomas (1981) suggests that the Ra value has no physical significance and is used because it is convenient for the instrument maker. It is true that most of the statistical treatments of surface roughness are based upon the rms (Rq) value of the surface heights; however the Ra value has found its way into widespread use and it would be difficult to change this situation. For many surfaces, namely those having a Gaussian distribution of surface heights, there exists a simple relationship between the Rq and Ra values:

$$\frac{K}{M} > \frac{z''(x)}{\left(\delta + z(x)\right)} \tag{2.1}$$

$$Ra = \left(\frac{2}{\pi}\right)^{\frac{1}{2}} Rq \tag{2.2}$$

Ra and Rq are examples of height-dependent parameters of which there are many more; continental Europe favours a system based upon extreme value such as the highest peaks and the lowest valleys. These methods have been well documented elsewhere (Thomas, 1999) and are not discussed here. An important stage in the

development of surface characterization was the realization that surface topography is not merely a height-dependent phenomenon but its variation in the horizontal plane is also relevant. The variation of surface height with horizontal distance is often referred to as the texture of the surface, an open-textured surface, being one with a larger distance between peaks than a closely-textured surface.

The inadequacy of Ra and Rq values to define the texture of a surface can easily be demonstrated by considering two sine waves of equal amplitude but different wavelengths. In both cases Ra and Rq values would be the same, however the waveforms have a visible difference and this observation is also true for real surfaces. The advent of random process methods allowed the height and spatial variation of surface topography to be described. Random process methods have been the bases of many of the published mathematical models of surface topography, a notable example being Nayak (1971); following on from the earlier work of Longuet-Higgins (1957) on the properties of ocean waves. Early attempts to quantify the spatial properties of surfaces confined themselves to analogue techniques.

The use of digital computers in surface topography characterization has led to almost limitless possibilities of defining new parameters. Indeed the extent of the growth of new parameters has prompted one notable author to refer to the present state of affairs as a "parameter rash" (Whitehouse, 1982). With the aid of computers it is possible to calculate, within the confines of the discretization process, any of the random process functions or height-dependent parameters that have gained popular acceptance amongst research workers.

Examples of numerical parameters used for surface profiles are amplitude parameters such as, in the modern ISO notation Ra (also called AA or CLA in the early years and Rq and spatial parameters e.g. Sm (see Figure 2.3). The Rq is the root mean square value where, Ra is the mean modulus and they tend to closely correspond on most realistic surface types. The Sm is the mean spacing between profile peaks at the mean line (the mean of S values shown in Figure 2.3) (Smith and Chetwynd, 1992).

The manufacturing processes developed in the 1990s shifted the surface metrology from profile to areal characterisation. Areal parameters are known as field parameters and are widely used for surface characterisation. Field parameters consist of both S-parameters and V-parameters. The S-parameters describe both amplitude and spatial information. The V-parameters give volumetric information based on an areal material ratio curve (Jiang et al., 2007b).

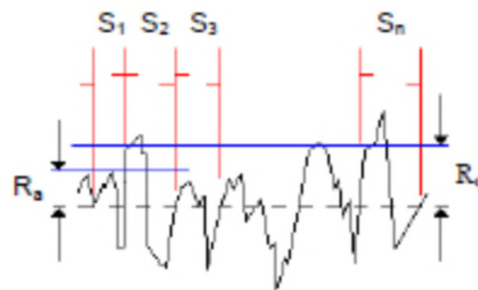


Figure 2.3: Some numerical parameters to define surface profile

2.3.2 Statistical Analysis

The majority of parameters are general and valid for any $M \times N$ rectangular image. But, for some parameters related to the Fourier transform, it can be assumed that the image is quadrangular (M=N). Before the calculation of the roughness parameters, it is recommended to carry out a slope correction by a 2nd or 3rd order polynomial plane fit. Moreover, roughness values depend heavily on measurement conditions especially scan range and sample density. Therefore, it is very important to include the measurement conditions when reporting roughness data. Furthermore, it is essential to note that these parameters based on local minimums and/or local maximums may be more sensitive to noise than other parameters. The parameters are divided into four groups as described in:

The Roughness Average (*Ra*) value can be written as:

$$Ra = \frac{1}{MN}\sum_{k=0}^{M-1}\sum_{l=0}^{N-1}\left|z(x_{k,}y_{l})\right| \tag{2.3}$$

The Root Mean Square (RMS) parameter *Rq*, is defined as:

$$Rq = \sqrt{\frac{1}{MN}\sum_{k=0}^{M-1}\sum_{l=0}^{N-1}\left[z(x_{k,}y_{l})\right]^{2}} \tag{2.4}$$

Both *Rq* and *Ra* require that the surface heights are measured with respect to a mean line. In analogue systems a capacitor-resistor (C-R) filter removes the trend and D.C. level from the profile signal before the parameters are calculated. With digital methods it is necessary to fit mean line to the raw data.

To describe the properties of a surface we may use standard statistical techniques. These techniques can be applied to the raw data or to describing specific features present in the profile record, e.g. heights, slopes, curvatures of the peaks etc. Such results can then be plotted as a statistical distribution which characterizes the surface.

The Surface Skewness (R_{sk}) describes the irregularity of the height distribution histogram, and is defined as:

$$R_{sk} = \frac{1}{MNR_q^3} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} \left[\left(z(x_{k,}y_l)\right)\right]^3 \qquad (2.5)$$

If R_{sk} = 0, a symmetric height distribution is indicated, for instance, a Gaussian like. On the other hand, if R_{sk}<0, it can be a bearing surface with holes or steps and, if R_{sk} >0 it can be a flat surface with peaks or plateau. However, values greater than 1.0 may indicate extreme holes or peaks on the surface.

The Surface Kurtosis (R_{ku}), describes the “peakedness” of the surface topography, and is defined as:

$$R_{ku} = \frac{1}{MNR_q^4} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} \left[\left(z(x_{k,}y_l)\right)\right]^4 \qquad (2.6)$$

As regards Gaussian, height distribution R_{ku} approaches three when increasing the number of pixels. Moreover, smaller values indicate broader height distributions and visa versa for values greater than three (SPIP Manual).

The Peak-Peak Height (R_y) is defined as the height difference between the image's highest and lowest pixel.

$$R_y = z_{max} - z_{min} \quad (2.7)$$

2.3.3 Random Process Methods

The use of purely height dependent parameters fails to characterise variations along the plane of a surface. Random process methods allow us to investigate the spatial properties of random signals. These methods owe much of their development to the communications industry which required techniques to describe complex signals. For example, correlation techniques can be used to recover weak signals from a source with a high level of background noise. Moreover, random process methods are equally applicable to the description of the properties of surfaces.

2.4. Modern Instrumentation Developments

Surface topography information collection using an instrument is carried out by either contact or non-contact methods. This section describes the basic principle of stylus and scanning probe microscopy (SPM) contact methods as well as optical non-contact methods (interferometry, laser triangulation, light scattering and confocal microscopy).

2.4.1 Contact Method

Physically, the principle of the contact can be described as such: given two solid bodies with known geometry pressed against each other with a force (or a pressure distribution) which caused deformations of the bodies. A theory of contact is required to predict the shape of the contact area and how it grows in size with increasing load, as well as the distribution of tractions, etc. Finally, it should enable

the components of deformation and stress in bodies to be calculated in the vicinity of the contact region. Before a problem in elasticity can be formulated, a description of the geometry of the contact surfaces is necessary. With assumptions of parabolic surface profiles and linear elastic, frictionless materials, (Hertz, 1896) proposed the first analytical solution for dry contact problems which laid the foundation for the subject of contact mechanics. Since then, the developments of contact theories mainly focus on eliminating assumptions made in the Hertz contact theory. Before giving the general dry contact model, Hertz analytical results are provided for a first insight of contact mechanisms.

2.4.1.1. Hertz Contact Theory

Hertz introduced the simplification that, for the purpose of calculating local deformations, each body can be regarded as an elastic half-space loaded over a small region of its plane surface (for details, see any elasticity textbook or, for a summary in our context Smith and Chetwynd, 1992). By this simplification, generally followed in contact stress theory, the highly concentrated contact stresses are treated separately from the general distribution of stress in the two bodies, which arises from their shape and the way in which they are supported. In order for this simplification to be reasonable, two conditions should be satisfied: the dimensions of the contact area must be small compared (a) with the dimensions of each body and (b) with the relative radii of curvature of surfaces. The first condition is necessary to ensure that the stress field calculated on the basis of a solid which is infinite in extent is not unduly influenced by the proximity of its boundaries to the highly stressed region. The second condition is to ensure firstly that the surfaces just outside the contact region approximate roughly to the plane surface of a half-

space, and second that the strains in the contact region are sufficiently small to lie within the scope of the linear theory of elasticity. Hertz also assumed the surfaces to be frictionless so that only the normal pressure is transmitted between two bodies. With those preparations, Hertz was able to give the analytical solution to contact problems. The contact stress (pressure) and elastic approach (indent) can be expressed as

$$a = \left(\frac{3PR}{4E^*}\right)^{\frac{1}{3}} \tag{2.8}$$

$$\frac{1}{E^*} = \frac{(1 - v_1{}^2)}{E_1} + \frac{(1 - v_2{}^2)}{E_2} \tag{2.9}$$

where v_1, v_2, E_1 and E_2 are the Poisson's ratios and elastic modulii of the two surfaces and R is the summit radius. For a small stylus tip, even with modest contact forces the pressure can appear to exceed the nominal yield point of the bulk material.

Surface layers are often not typical of the bulk and observations regularly show a little damage in practice (Thomas, 1982), but this issue remains as a regular point of controversy in the field.

2.4.1.2. Types of contact methods

Two types of contact methods are considered in this survey, the stylus method and scanning probe microscopy.

Stylus Based Methods

The stylus instrument is widely used for surface profile measurement. It can cover a large measurement range up to several millimetres with nanometre vertical resolution. The stylus instrument measures the geometrical features of the surface directly by a tactile method. The stylus tip is (usually) mounted on a cantilever which is coupled to a sensor such as Linear Variable Differential Transformer (LVDT). The LVDT detects the vertical movement as the stylus traverses the sample surface at a constant speed. As such, the generated voltage represents the convolution of the stylus tip with the surface profile. The output voltage from the translator is amplified and digitized in order to be manipulated by a computer (see Figure 2.4).

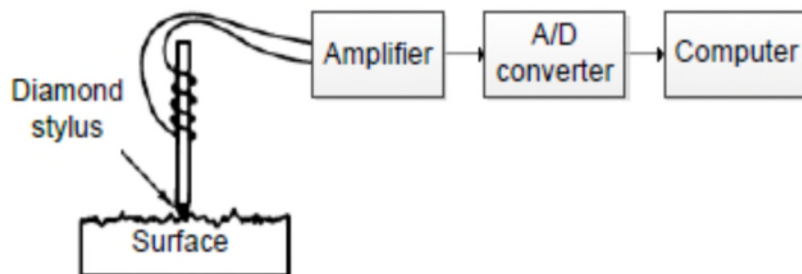


Figure 2.4: Principle subsystems of a stylus instrument

The lateral resolution depends on the radius of curvature of the tip and slope of the surface irregularities (Bennett and Dancy, 1981, Elson and Bennett, 1979). For instance, Figure 2.5 shows that the stylus tip contacts the sides of a valley in such a way as to prevent it penetrating to the full depth of the valley if the radius of the stylus, r or the amplitude of the sample, z increases.

The resolution issue just discussed is one of the main disadvantages. Another is sometimes argued to be that the stylus tip can damage the surface to be measured if

the contact stress exceeds the hardness of the surface. Therefore, the load of the stylus is controlled to be as low as possible such as 0.05mg (Bhushan, 1999). Furthermore, the measurement speed is relatively slow if a real surface measurement is to be undertaken. The measurement speed and the contact approach make stylus based methods inadequate for online/in-process measurement (Young et al., 1980, Bhushan, 2001).

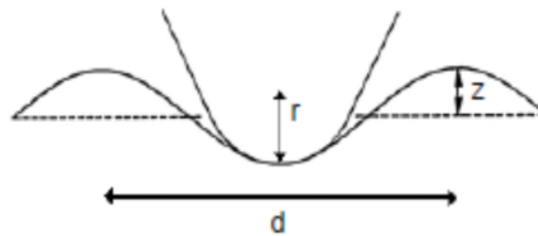


Figure 2.5: Sinusoidal surface with a stylus [Elson and Bennett (1979)]

Scanning Probe Microscopes (SPMs):

SPMs are commonly used for ultra-precision applications, right down to atomic scale resolution in surface measurement. The SPM has much in common with stylus instruments in the sense that a fine tip scans the sample surface to obtain real information. However, the main difference between the two methods lies in the detection of parameters (see Figure 2.6). In contrast to the stylus, where the geometry features are directly detected, the SPM detects, for example, charge density or atomic forces (Alvarez and Siqueiros, 2010). The following briefly describes two types of SPM probes, the scanning tunnelling microscopes (STM) and atomic force microscope (AFM). Binning and Rohrer (1982) achieved atomic-scale measurement resolution when they scanned a probe over a sample in a vacuum within a stabilised environment, which considered the first paper on STM.

The probe tip and the measured sample are brought to within angstroms of separation using a piezoelectric stage. Thus, the wave functions in the tip (i.e. electron clouds) overlap the wave functions in the sample surface. By applying voltage between the probe tip and the sample surface, a tunnelling effect is produced and a small current jumps the potential barrier across the tip and sample.

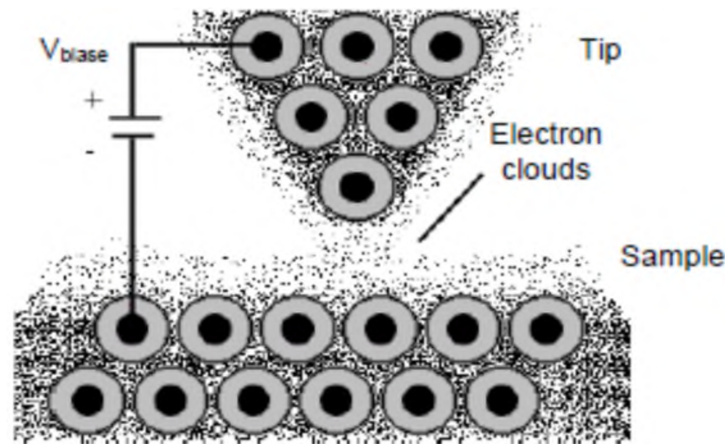


Figure 2.6: Principle of interaction between tip and sample

(Binnig and Rohrer, 1983, Chen, 2008).

Sensing this tunnelling current yields a very fine, but highly nonlinear measurement resolution for the gap between the tip and surface. Scanning the tip across the surface while servo-controlling its height to maintain constant current provides from the servo signal a relatively linear measurement of the surface heights. Because the effect is so localized to atoms at the very end of the stylus tip, STM is capable of sub-atomic resolution in all three axes. Their work led to a Nobel Prize in 1986.

The STM invention was soon extended into AFM instrumentation (which is now the most widely used SPM technique). The AFM basically consists of a tip of several nanometres radius of curvature attached to a micron scale cantilever. In

contrast to the STM, which uses the electron clouds as a measuring parameter, the ATM uses the atomic forces that are generated when the tip contacts or is in close proximity to a surface. The generated forces result in the bending of a cantilever to which the tip is connected. This deformation in the cantilever is usually detected by interferometry or optical deflection methods such as laser triangulation (Thirunavukkarasu, 2007). The measuring technique is similar to that of the STM, with constant bend of the cantilever (representing a constant interaction force) used as the null-servo condition. Although special AFM can deliver atomic resolution, most practical systems offer slightly less, but they have a larger range than STMs, are considerably more robust and do not require electrically conducting surface. There are several well-established variants of AFM, which need not be considered here. Because of the slowness and sensitivity of the measurement procedure these types of instruments do not fit usually with the industry environment for in-process measurement applications (Bhushan, 2001).

2.4.2 Non-contact Methods

Optical methods are almost always preferred over non-optical (e.g. capacitive methods) in non-contact surface texture measurement because of their fine vertical and lateral resolution. Also, data acquisition in optical techniques is fast which makes them suitable for online/in-process measurement applications. Furthermore, the non-contact concept gives them the lead for non-destructive measurement applications. The optical methods can be classified into interferometric and non-interferometric techniques. The interferometric methods are based on the detection of the interference of two beams of light. Non-interferometric methods are based

instead on detecting non-interfered light (scattered, incoherent) reflecting from a sample surface. The following sections describe some the main optical techniques.

2.4.3.1. Interferometry for Surface Metrology

Interferometry has enormous potential in surface metrology for in-process measurement. Interference fringes were discovered in 1802 by Thomas Young. However, Albert A. Michelson implemented the first interferometry instrument in 1882 (Michelson, 1982). From that time up until the 1970s these instruments were used for dimensional metrology rather than surface metrology. With the development of the laser, charge-coupled device (CCD) and personal computers (PC) interferometry found a place in surface metrology. Many algorithms have been developed to analyse interference fringes and hence reconstruct surface texture.

Nevertheless, it is important to mention here that optical interferometry techniques are extremely sensitive to environmental noise such as mechanical vibration, air turbulence and temperature drift (Adhikari, 2004). These noise sources cause errors in surface measurement and can produce invalid results.

Surface topography assessment at the nanometre and micrometre scales needs measurement instruments that can provide high vertical and lateral resolution. This is possible if microscope objective lenses are combined with interferometers so that the lateral resolution depends on the objective lens while the vertical resolution depends on the interferometer. When choosing an objective lens, trade-offs must be made between: the required field of view, the working distance and depth of focus of the objective lens, the amount of reflected light back through the objective lens, the surface gradient of the measured sample, and the required lateral resolution. The

lateral resolution depends, among other factors, on the magnification factor of the lens. High magnification lenses can provide micrometre and even sub-micrometre resolutions. For example, a 10X type objective lens gives approximately 1.4 µm of lateral resolution for an illumination source having wavelength 683 nm and an objective lens with numerical aperture 0.28.

This lateral resolution is based on the Rayleigh criterion which identifies the lateral resolution as:

$$r = 0.61\frac{\lambda}{NA} \qquad (2.10)$$

Where *r* is the lateral resolution, λ is the wavelength of the illumination source and *NA* is the numerical aperture. Vertical resolution depends mainly on the interferometer, generally ranging from sub-nanometre to a few microns based on the measurement technique that is used within the interferometer (e.g. phase shifting).

The four common interferometric objective configurations are Fizeau, Mirau, Michelson and Linnik. All of these interferometers are based on the interfering of two beams to obtain an interference fringe pattern, as represented by Equation 2.9. Each type of objective configuration has its own applications.

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2}|\gamma| \cos(\Delta\varphi) \qquad (2.11)$$

This equation clarifies the fact that the interference pattern has a sinusoidal pattern that represents the constructive and destructive interference fringes. Where *I*, is the intensity of the interference pattern, I_1 and I_2 are the intensities of the first and

second beams respectively. The coherence function $|\gamma|$ can be obtained practically by altering the OPD and registering the corresponding drop in intensity.

$\Delta\varphi$ is the phase difference between them (i.e. $\varphi_1 - \varphi_2$) where the phase φ can be defined as

$$\varphi=\frac{2\pi}{\lambda}nz \qquad (2.12)$$

where, λ is the wavelength of the light. The fringe visibility is defined as

$$v = \frac{(I_{max}-I_{min})}{(I_{max}+I_{min})} \qquad (2.13)$$

where, ν is the fringe visibility, I_{max} is the maximum fringe contrast and I_{min} is the minimum fringe contrast.

The fringe visibility is constant at any point along the optical path when the light beams are perfectly coherent (i.e. the emitted waves having the same frequency and phase in a similar manner to the waves emitted from lasers). White light, having a mix of wavelengths, is inherently incoherent and, in principle, does not generate interference fringes.

Phase Shifting Interferometry (PSI) is one of the most common and widely used techniques to measure super polished and smooth surfaces. The measurement resolution can be up to 1/1000 of a fringe under optimum conditions (Kafri, 1989).

The fundamental principle of this interferometry is to modulate (i.e. shift) the phase by physically moving one of the interferometer arms. This is often done by pushing

the reference mirror using a Piezoelectric Translator (*PZT*). After each phase shift the light intensity is registered. These registered intensities are then processed by a PSI algorithm in order to determine the original phase. It is necessary for the phase shift to be known for each step and computationally much simpler if the phase is shifted in equal steps to be shifted equally *n* times that is the mirror should be moved equal distances. This is because the interference pattern contains three unknown variables: the DC intensity bias, *Io*; the visibility, *v;* and the original phase, *φ*. As such, the minimum requirement is three sets of intensity readings (i.e. *n*=3). The following mathematical description is used for the case of *n* phase shifts (Schreiber and Bruning, 2007).

$$\begin{bmatrix} I_1 \\ I_2 \\ . \\ . \\ I_n \end{bmatrix} = \begin{bmatrix} I_o[1+\nu\cos(\varphi+\alpha_1)) \\ I_o[1+\nu\cos(\varphi+\alpha_2)) \\ . \\ . \\ I_o[1+\nu\cos(\varphi+\alpha_n)) \end{bmatrix} \tag{2.14}$$

where α is the induced phase shift. The most common algorithms to retrieve the original phase value are the Carré algorithm, three shift algorithm, four shift algorithm and five shift algorithm (Wyant et al., 1984, Carré, 1966, Schwider et al., 1983, Hariharan et al., 1987). They all require a known amount of phase shift to be applied except for the Carré algorithm. For instance, the three shift algorithm requires a shift of 90°. In 1966, Carré introduced a technique for PSI that is independent of the magnitude of the phase shift, thus it may be an arbitrary amount.

In practice, white light sources are widely used with special optical interferometers. The spacing of the interference fringes depends on the wavelength, except when the Optical path difference (OPD) is zero (Wyant, 2002). Hence, maximum visibility occurs when there is no OPD between the measured and reference paths, whereas the visibility is reduced when the OPD increases (Balsubramanian, 1980). Generally, the fringe visibility in white light interferometry is limited to a few micrometres, and this can be explicated as a measurement tool (Muhamedsalih, 2013).

Coherence Scanning Interferometer (CSI) is a combination of the White Light Interferometry with vertical scanning techniques. The CSI is widely used to assess surface topography by analysing the coherence function obtained from the interfering light beams (Petzing et al., 2010). Subsequently, the white light has a short coherence length, the maximum variation in the fringe visibility can be easily monitored by capturing a sequence of surface images when the reference mirror is scanned mechanically. An important point about all of these is that heights are captured in parallel for every pixel in the image (Muhamedsalih, 2013).

Wavelength Scanning Interferometry (WSI) is an interferometric method usually used to measure large absolute distances without any 2π phase ambiguity problems.

In summary**,** Interferometry is based on analysing the fringes produced by superposition of the beams which share a source, propagation axis and plane.

PSI is suitable for high precision measurement for super polished surfaces. When the optical path difference exceeds $\lambda/2$, CSI and WSI can be used for rough surfaces. Coherence Scanning Interferometer (CSI) is widely used for surfaces measurement,

having a large measurement range with sub-nanometre resolution. Wavelength Scanning Interferometry (WSI) is an alternative method to measure large discontinuous surfaces, but requires no mechanically moving parts to perform phase shifting operations (Muhamedsalih, 2013).

2.4.3.2. Laser Triangulation

Laser triangulation is widely used in dynamic measurement to determine the distance of the surface from the photo-detector, R, then to extract the surface texture. This method is based on trigonometric relationships between the incident angle of a laser beam, θ and two known distances. These are the distances of the sample surface and the worktable plane from the photo-detector, R and H respectively (Cui *et al.,* 2010).

Figure 2.7 shows the principle of operation. The surface height, *D* from the reference plane i.e. worktable, can be simply obtained by the equation

$$D = H - L\,cot\theta \tag{2.15}$$

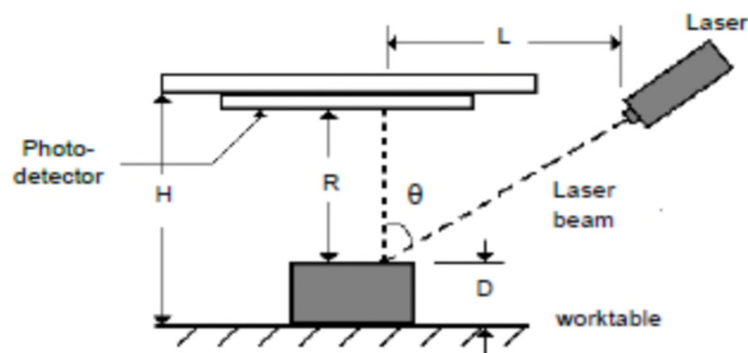


Figure 2.7; Laser triangulation method [adapted from (Groover, 2007)]

2.4.3.3. Light Scattering

Light scattering techniques are widely used for applications that require rapid and non-destructive surface roughness measurement (Vorburger *et al.,* 1993). These techniques are based on evaluating the scattering pattern obtained from a surface illuminated by a laser light (see Figure 2.7).

The scattering pattern reflected from a surface illuminated by laser light depends among other factors on the surface roughness. If the surface is perfectly smooth, then θi is equal to θr according to the law of reflection. However, increasing the surface roughness will diffuse the light away from θr. By analysing the diffusion angle, the surface roughness can be obtained.

This method has been the subject of much research (Church and Zavada, 1975, Bennett and Mattsson, 1989). Persson (1998) presents light scattering instruments for in-process measurement of surface roughness ranges of $0.09 \leq Ra \leq 0.16$ μm. The instrument used by Persson (1998) to illuminate a sample on a grinding machining while in operation was based on a He-Ne laser. The light intensity of the specular reflection was captured and analysed to determine the surface roughness. It was found that the result was of the same order as when a stylus instrument is used.

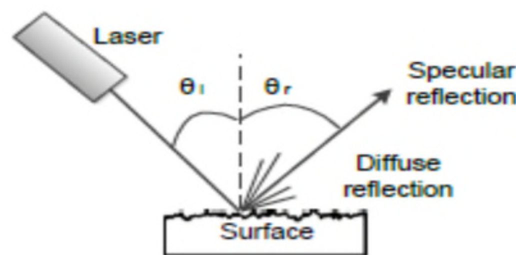


Figure 2.8: Principle of light scattering method

2.4.3.4. Confocal Microscopy

The enhanced image obtained using a confocal microscope when compared with a conventional microscope is produced by a pinhole aperture placed in front of the detector to reject out-of-focus light (see Figure 2.9). This concept was developed by Marvin Minsky in 1955 into a working instrument (Semwogerere and Weeks, 2005).

Since only the light that focused on the sample will pass through the pinhole, slices along the axial direction need to be viewed to reconstruct areal information about the surface. This can be achieved by capturing a set of images while the sample is scanned vertically (Conchello et al., 1994). Tony Wilson (1990) was a pioneer in developing the confocal microscope (Wilson T, 1990).

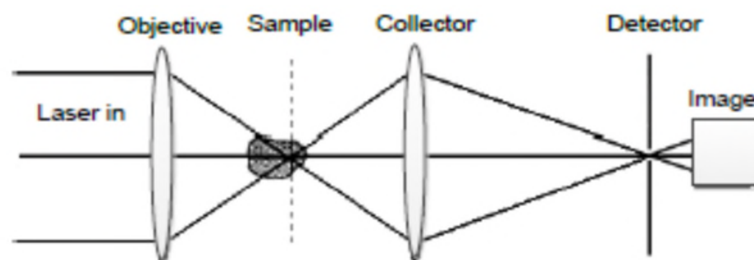


Figure 2.9: Schematic of confocal microscopy

Chromatic confocal microscopy introduced in 1994 by Tiziani and Uhde which has the potential to be used for (3-D) three-dimensional embedded metrology measurement. Colour impression was used to evaluate the intensity distribution and to discriminate the surface height. Only three images, it has been found, are required for areal topography which significantly improved the measurement throughput. Crucially this advantage makes the instrument suitable for online measurement applications (Tiziani and Uhde, 1994).

2.5. Comparison between Instruments

Contact and non-contact instruments work on different physical principles. The stylus measures the surface geometry, the optical interferometric techniques generally measure the optical path length and SPM measures either charge density or atomic forces. Besides surface measurement, the stylus can be used to measure physical properties such as workpiece hardness in addition to surface topography; film thickness can be obtained using some optical instruments (Stanford and Hagan, 2009). Therefore, even though all the instruments can retrieve surface topography information, the measurement principle may decide the application.

Generally, the measurement speed for non-contact methods is faster than for contact methods. The instruments have different measurement ranges and resolution capabilities. In contrast to the optical methods, the stylus could damage the surface and the measurement probe needs continuous wear condition monitoring. Additionally, higher vertical resolution can be achieved with optical techniques but the lateral resolution is not comparable. The highest vertical and lateral resolution can be obtained by using AFM, but it has the smallest vertical range.

2.6. Summary

Surface metrology consist of three stages: instrumentation; filtration and decomposition; characterisation and parameterisation. The geometrical features of the surface can be categorised into three features: form, waviness and roughness. Measurement of these features can be used to control the manufacture process and predict the functional performance of the workpiece. The Instrumentation stage acquires the surface topography information. Filtration decomposes the surface

texture into its primary features. As a final stage the surface is characterised by numerical parameters and functional curves.

Instrumentation can be divided into those using contact and non-contact methods. The non-contact method, measures the optical path length. It is a non-destructive method, has a relatively large measurement range up to 1 millimetre and sub-nanometre vertical resolution and has much potential for embedded measurement. The non-contact method can usually acquire the measurement data at high speed. The contact method, on the other hand, measures the geometric features of the surface physically. It can be destructive and is considered inadequate for embedded measurement. The measurement speed is relatively slow for areal surface measurement.

In summary, optical methods, particularly interferometers, are the techniques that have the most potential to be used for micro/Nano scale embedded surface metrology. It is also clear that all the techniques offer challenges when used near the limits of the ranges. Trends in industrial production demand much more measurement of small amplitude and small wavelength roughness. However, whether a physical probe tip or a spot of light, all probes are in some sense finite and of (slightly) indeterminate characteristics, they will all introduce particular non-linear distortion to fine surface detail.

Chapter 3
3. History of Stylus simulation

3.1. Introduction

This chapter introduces more specifically the background relevant to stylus simulation upon which the new work will be built.

Despite a growing divergence in surface metrology applications, there seems to be one element that is still common across the vast majority of instrumentation - a stylus tip (typically spherical, at least nominally) which mechanically makes contact with the surface. It is this "contact-based" approach which is used in detecting any engineering surface features and remains so far the most popular method despite the numerous alternative sensing approaches, including ultrasonic, capacitance, pneumatic and optical (Whitehouse, 1994).

In many practical regards, the underlying convolution of a stylus tip geometry over a surface is perhaps the most common or "unified" element among the vast majority of today's instruments. It may be argued, however, that the term "convolution" does not apply here when taken in the strictest mathematical sense, although it has been used in practice (Li, 1991).

3.2. Stylus Tip Convolution

In the stylus context, "convolution" is used to represent the generation of a surface profile as determined by the geometric or physical relationships between the stylus and surface. This "stylus convolution" differs from a pure mathematical convolution in that the latter typically incorporates all data points and the former is primarily based on peak interactions. Stylus or radius based convolution is a process

that causes a limitation on the surface wavelengths which are transmitted to the instrument's transducer.

However, this wavelength limitation, or more correctly- wavelength modification, cannot be generalized. Models have been developed to predict the transmission behavior of tip geometries; however, these only apply to very limited classes of surfaces.

The fundamental problem with predicting the transmission characteristics associated with stylus tip convolution is the fact that the tip convolution process is based primarily on peaks and thus varies from surface to surface. Mathematical models of the effects of this convolution process are typically based on idealized surfaces (such as sinusoidal or Gaussian) and generally these models incorporate all surface features, including valleys (Whitehouse, 2011, Raja 2009).

Although stylus tip convolution effects are not easily characterized or predicted mathematically, the stylus tip radius can be a functional and reproducible means of limiting wavelengths in surface metrology assessments. One common example is in the assessment of cam lobe geometry whereby, in application, a roller will follow the surface. In this application, it is very important to control the dynamics of the roller which ultimately generates the kinematics along the mechanism. To understand the functionally important cam surface attributes, it is essential that the surface be measured and analyzed from the roller's perspective - exploiting the features which the roller would "see" and ignoring those that it would not.

In so doing, the stylus acts as an artificial follower in the measurement process and

thus generates a data set directly related to the motion that the follower will produce in application. This "radius based" transmission is what is desired in order to functionally assess the surface geometry. Historically, this approach has been referred to as the envelope or e system, whereby the envelope of the stylus convolution is deemed to be the surface of interest (Weingraber, 1956).

Despite the fact that radius based wavelength limitation is often difficult to predict and can vary dramatically between profiles, it is a functionally important and metrologically reproducible means of assessing surface features and has therefore been included in the scheme for unification. In addition to the obvious importance relating to "radius based" wavelength limitation, tip convolution effects must also be understood in the context of wavelength limitations per the other approaches in the unification scheme.

Given that most of today's surface metrology instrumentation incorporates a nominally spherical stylus, the understanding of these transmission effects is necessary in selecting an appropriate tip radius when using Gaussian or ideal filters. For example, when assessing a surface with a Gaussian band-pass, it is necessary to utilize a stylus tip that does not significantly affect the surface wavelengths in the desired range.

The mathematical approach involves collecting a data set with a relatively small tip radius and subsequently performing a numerical convolution or simulation of a larger tip radius over the data set (Scott, 1992a). This mathematical approach is considerably more desirable in terms of factors such as the cost of purchasing and maintaining styli of varying sizes and the flexibility of being able to alter the tip

radius and not having to re-measure the component. Although similar results might be expected between the methods for most surfaces, they may not be identical. These divergences (although they are typically small) can be the result of many factors. Two of the primary factors will be discussed in the following sections.

3.2.1. "Noise" in the Measurement:

Disturbances (whether mechanical or electronic) which are present in the measurement process, translate directly into the measured profile. In any physical measurement, these effects will be present in the profile to some extent. Appling a mathematical convolution to the measured data set will tend to smooth out random noise. The geometrical transformation associates with stylus tip 'convolution' similarly smooth out many of these effects. If a physical measurement were made with the same tip radius that was mathematically applied, we may find high frequency attributes which are the result of the measuring process. The influence of this measurement "noise" is depicted graphically in Figure 3. 1.

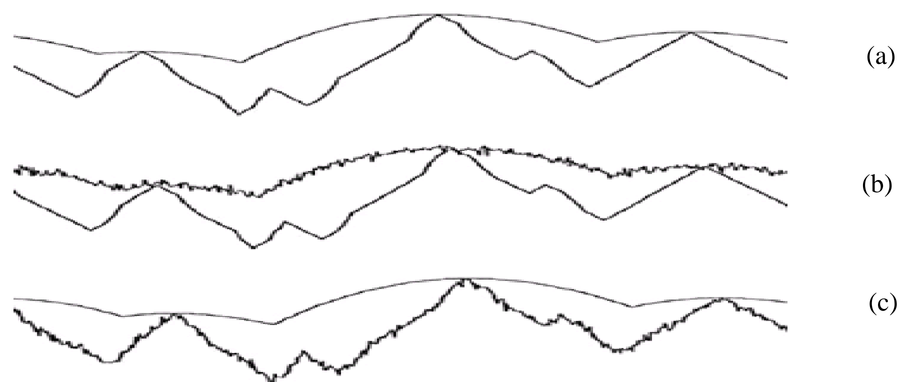


Figure 3.1: The influence of vibration in physical versus mathematical tip convolution

Figure 3.1a demonstrates the ideal situation in that there is no "noise". In this case, we would find that (apart from other influences) the mathematical convolution duplicates the physical convolution. However, in real applications, high frequency

noise will be present. If the surface were measured (physically) using the desired (large) tip radius, we may see a data set as is shown in Figure 3.1b, whereby the stylus path incorporates relatively high frequency changes as it moves over the profile peaks. In the mathematical convolution, the surface is first measured with a relatively small stylus (once again incorporating the noise) and then a circle is mathematically convolved over the profile. Figure 3.1c demonstrates how the convolution process tends to smooth out the high frequencies which were present in the underlying profile measurement. It should be noted that the radius convolved profile of Figure 3.1c is very similar to that of 3.1a, however the 3.1c profile may well be based in part on the peaks of the noise rather than purely from the base surface.

In a similar sense, irregularities in the stylus tip geometry (for example, from a systematic deviation from ideal shape or from chips or other wear damage) can introduce errors in the physically convolved data set. In a physical measurement using a relatively large probe (as in functional relationships to a cam-follower) it is probably right to ignore imperfections because it measured gross shape characteristics of the workpiece, attempting to remove/ignore the small-scale roughness. Larger-radius tips/followers can be made to close shape tolerance the small ones and slight wear damage is also likely to be small compared to the overall tip shape. In the case of a computational fitting process (be it stylus convolution, other motif methods, or conventional fillers), it can be implicitly assumed that the tip dimension is small compared to the shortest functionally significant wavelength in the assessment (this is something made clear only in more recent standards – early ones tended to, by default, allow the tip itself to define the short wavelength

limit). If this is so, error in the tip hardly matters. Hence this statement holds up or a typical E-system uses similar processes where it cannot get a tip significantly smaller than the functional limit. The errors may then be of quite significant magnitudes. This is what this thesis is considering: what errors/uncertainties should be specified because of tip effects, etc.

Another factor that can lead to a difference in assessments made via mathematical convolution of a circle versus physical convolution of (nominally spherical) stylus is the presence of peaks to the sides of the profile trace. When measuring a surface with a stylus tip of finite dimension, the potential for stylus-to-surface contact occurs over an area. Larger stylus tips have a greater potential for interacting with peaks that are to the sides of the trace. This is shown in Figure 3.2 where the high peak in the center of trace #3 could influence measurements along trace #2 depending on the tip radius being used.

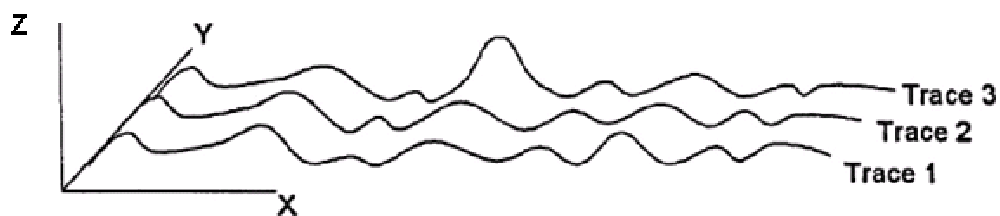


Figure 3.2 The presence of a "lateral" peak in trace 3 relative to trace 2.

The true convolution of a stylus tip in the presence of lateral peaks is best predicted through the three dimensional or areal measurement of the surface and the subsequent convolution of a sphere (Shunmugam and Rahadkrishnan, 1975). However, in practice this is very time consuming and more mathematically intensive than the simple measurement and analysis of single profiles.

3.3. Filtering

In commercial systems profile characterization is carried out on filtered data. In most cases a relatively simple capacitor-resistor network (2C-R) is used. This technique has several disadvantages, the roll-off characteristics are poor and some features longer than the nominal cut-off length are retained in the filtered profile. The most serious limitation is that there are only a few values of cut-off that may be selected (Stout and Blunt, 2000).

Roughness exists over a large range of spatial sizes. It is therefore important to limit the analysis to the range of sizes that are pertinent to the problem. To some extent this is achieved by a suitable choice of effective sampling interval. Systems that use computers to data log surface profiles have the advantage that sampling lengths and intervals can be easily altered. A second advantage is that digital filters can be used to carefully control the size of features that are included in the characterization process. Digital filters have different, usually sharper, roll-off characteristics than the standard C-R filters, and can be designed to be free from phase distortion. Unfortunately, it is not possible to compare results from digital filters with those obtained from C-R filters, due to their different characteristics. The greater flexibility of digital filters allows the choice of filter length to be tailored to the problem being studied. The limited bandwidth of wear processes has already been discussed; the same also applies for many other processes such as friction, contact mechanics, EHL etc. Knowledge of the mechanics of the process involved can lead to a choice of filter length that will highlight the topography of interest. Even the simple parameters are sensitive to filter length, for example Ra is particularly sensitive to the long wavelength components of topography. If these long

wavelengths are included in an analysis any changes in the shorter wavelengths may go unnoticed.

The small scale geometrical features on engineering surfaces are usually of complex shape and surface profiles and are therefore composed of a wide range of frequency components. As explained in previous sections on roughness and waviness, roughness corresponds to high frequency or short wave components, while, on the other hand, waviness has low frequency (long wave) components which tend to present low changing rates in surface profiles (Chetwynd DG., 2001), see Figure 3.3.

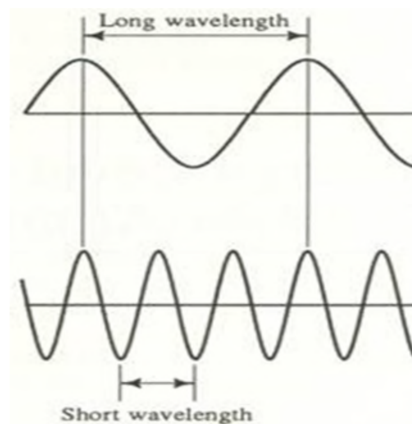


Figure 3.3: Long and Short wavelength components diagram

Filtering is applied as a procedure to separate out the surface-profile's frequency range of importance to the analysis. The filtering operation is able to carry out the following action: 1, in order to extract the roughness profile, allowing high frequency component through. 2, letting low frequency component pass through, thus the waviness profile is able to be extracted.

In sum, the filtering techniques act as a support or assistant tool which can help further analysis and can also reduce errors during the probe measurement.

3.3.1. Development history

The filtering operation has always been a pivotal factor in surface metrology analysis. The filtering operation as a modern technique has over 55 years of history (Muralikrishnan and Raja, 2009). Electrical Filters, Envelope Filters and Morphological Filters are the major members of the filtering operation family. The introduction and discussion of these three filters will be given in subsequent sub-sections, although the main details will concentrate on Envelope filter and Morphological filter systems.

3.3.2 Electrical Filter systems

This filter system treats the profile as a voltage source and applied a two-resistor-capacitor (2RC) network. The merit of this network is that it has memory; therefore the output value can be a function of the input at any instant and a function of prior values. Digital evaluation provides a whole range of filter characteristics, including those with no phase distortion. From (Muralikrishnan and Raja, 2009) explanations, 2RC can provide reduced weights to the voltages by running previous voltage and average current value.

The undesired phase would thrust into the output, when the network's memory assists averaging.

3.3.3 Envelope Filters system

The E-system which is short for Envelope Filters system was invented in the 1956's by Von Weingraber (Jiang et al., 1984). The basic clarification of this method is to simulate the features important in contact between mating surfaces and surface texture by determining the envelope traced by a ball rolling on this

surface. According to the track when the ball is rolling through the object's surface, the deviation from the envelope can be defined by the rolling track to see the several trends of the roughness along the surface. The envelope line is generated by the simulating process of the rolling ball. As shown in Figure 3.4:

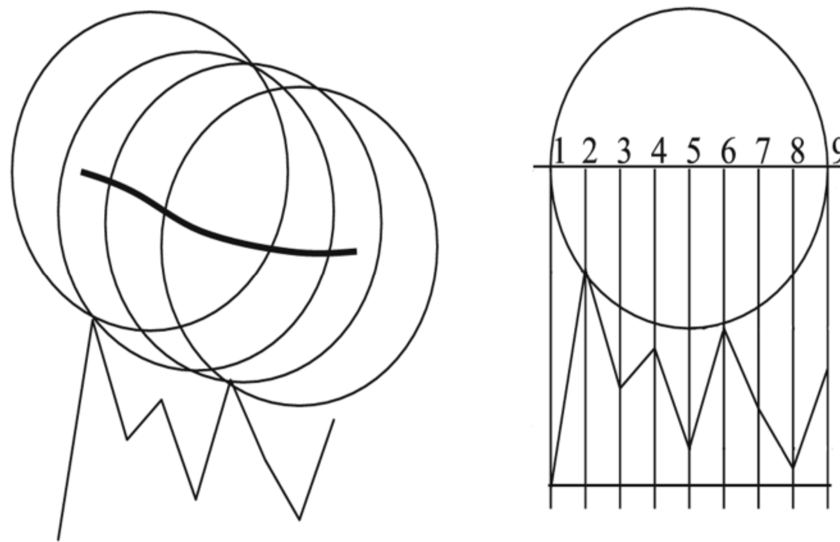


Figure 3.4: Envelope filters method, locus of the ball's centre when traverses the surface (Adapted from: Muralikrishnan and Raja, 2009)

It presented as an example the ball rolling through a surface. The black line shown on the track of the rolling ball is the envelope line. This ball is only doing vertical translations by assumption and making the peaks contact with the rough surface. The reason that some peaks did not contact with the rolling ball will be classified later in this sub-section.

Choosing a larger radius ball causes the envelope central locus to overlook the narrower peaks and valleys of the profile, so the central locus would be shown to be smoother. Otherwise, the centre line enters the narrow peaks and valleys when a small radius ball has been used.

A physical implementation of the smoothing would require thought about materials especially as regards stiffness and strength. But the underlying concept of the E-system is acting as an ideal rigid ball (i.e. a piece of geometry). Thus the software implementations in this thesis can be treated as an ideal material case.

3.3.4 Morphological Filters systems

Various concepts, from the E-system and its minor variants were linked formally to the mathematics of morphology and image processing to introduce the morphological filters in the late 1990's (Srinivasan, 1998) (Shan Lou, 2012). Morphological filter systems are the upgrade and superset technology of early envelope filters. More tools and capabilities were needed for morphological filters operation which provided better performance than envelope filters.

According to Scott (2000), morphological operations are able to be combined together to achieve a superimposed effect. And also a technique named scale-spaces can be a support tool by providing multi-resolution analysis to surface texture.

The techniques of envelope filters discussed in the previous section would be used in morphological filter systems as the superset of E-systems. There are two filters named as closing filters and opening filters.

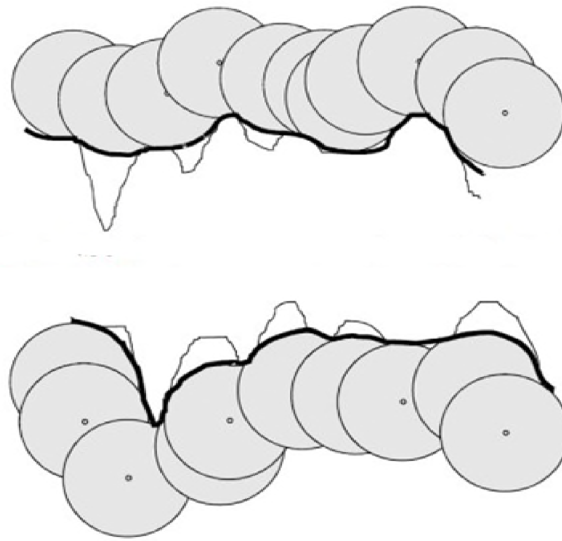


Figure 3.5: Closing and opening envelope of an open profile by a disk (Scott, 2000)

The filter acts on the profile to produce different mean lines. Figure 3.5 presented two examples of applying a closing operation and the opening operation on an open profile with the disk.

Achieving opening filter needs to place an infinite number of identical sizes and material disks and make contact with the profile from below. However; the closing filter can be gained by placing an infinite number of identical sizes and material disks in contact with the profile on its upper surface and also taking the disks' lower boundary. These two types of operation are acting in opposition by making comparison with each other which makes the efficiency in terms of surface measurement better than the traditional envelope filter systems. Both the opening and closing filters can be operated at the same time during the profile measuring period. These opening and closing operations are the secondary morphological operations produced by dilation and erosion which are the primary morphological operations (Muralikrishnan and Raja, 2009). Erosion reference lines are generated by a dilation followed by an erosion operation and referred to as a closing mean

line; on the other hand there would be an opening mean line. Then the combination of opening and closing filters can be made with various structuring element sizes which are called alternating sequential filters (ISO16610-41, 2010).

The extraction of different scales of information of the testing surface can be achieved by using scale space techniques to increase the structuring elements' size.

3.4. Errors in using contact-probe/stylus surface measurements

3.4.1 Briefing on stylus

The contact-probe/stylus instruments have been widely used in the surface metrology field. The first diamond stylus of 90° pyramid shape was developed in the UK which obtained two dimensions 2µm and 7µm at the tips (Whitehouse, 2002).

This type of stylus worked well on surfaces with an appreciable lay. The US subsequently invented a diamond cone which had an angle of 60° (Whitehouse, 1997).

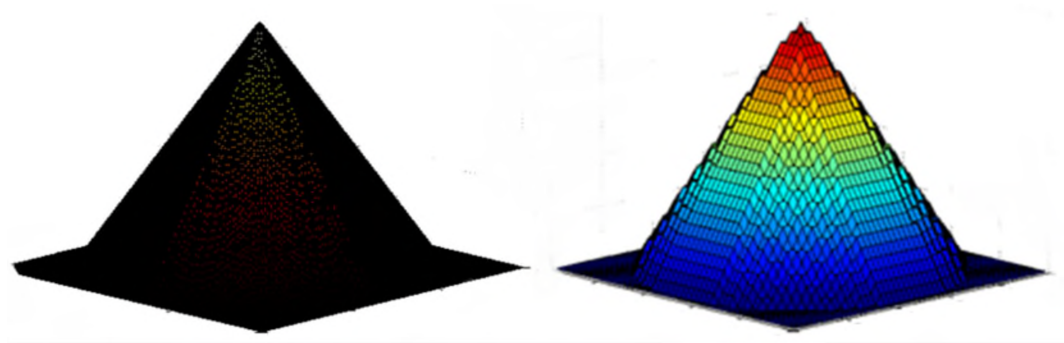


Figure 3.6: Stylus shapes: cone (left), pyramid (right)

As Figure 3.6 presents: the two shapes of stylus are those of cone and pyramid. Both of these stylus tip shapes have been suitable for a great many jobs. They are able to report accurately on the structure of fairly coarse structures with 'open'

texture, however, they are not suited to the smaller scale texture more much more commonly found in industry. So the stylus slope has to be considered, as regards which the tip can be infinitely sharp so that the stylus can go into the valleys to get more precise measurement results.

3.4.2 Noise and Vibration

Noise and vibration can be one of the major sources of interference for doing the contact-probe/stylus measurement on rough surfaces. This can be caused by inherent instrument noise, environment or by sound. For instance, when the contact-probe through the roughness surface caused some vibration and electrical noise in the pickup line. The most typical noise source is sinusoidal which has a range from 50Hz to 60Hz inclusive, short noise peaks (Mainsha, 2001). Noise contribution can be formally written as:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \mathbf{C}_I \mathbf{C}_S \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} + \begin{bmatrix} n_x(t) \\ n_y(t) \\ n_z(t) \end{bmatrix} \tag{3.8}$$

n_x, n_y n_z are the noise components, x, y and z presented the directions and t indicates time. The first two noise components would influence length and height measurement on x and y directions. And n_z affects the height measurement on z direction and also provides an indirect influence on determining the length. Value n_z will be considered and added into this project topic as affecting the height measurements. In sum, noise is clearly a critical major effect especially when measuring small dimensions, because the dimension of the object being measured is closer to the noise level.

3.4.3 Stylus errors and distortion

Despite the points discussed previously, stylus methods still retain a most significant role in surface metrology because of the more precise measuring and the fact that they are relatively robust against dirt compared with optical methods (non-contact ways).The probe stylus is one of the most critical and pivotal parts in surface metrology (Mainsha, 2001). The probe has to be sharper than the perturbation being measured. If a suitable probe is not selected for the rough surface measuring, damage on the surface and on the probe itself might occur. According to the work done by Church and Takacs (1990) because of effects on tip size, smooth surfaces are likely to be reported as rough, while rough surfaces are smoothed.

The stylus method is sometimes criticized because of claims that it can damage the surface (Whitehouse, 2002). Whitehouse mentioned the drawbacks to the theories' algorithms method on stylus techniques, for example possible distortion and stylus damage. He mentioned that there is a simple way to obtain a criterion for a stylus to achieve no surface damage. Whitehouse's method was focused essentially on kinematic theory not on the material of the stylus. Damage index ψ was the crucial factor in this theory in which it was defined that its value must be in the acceptable range of no damage occurring. Whitehouse pointed out, drawing on (Schmalz, 1936), which itself used the well-known Hertz contact theory, that a damage index ψ can be used:

$$\psi = \frac{1}{\pi}\left(\frac{\mathrm{W}}{\mathrm{H}^3}\right) \times \left(\frac{\mathrm{E}}{\mathrm{R}}\right)^{\frac{2}{3}} \times \frac{1}{\mathrm{H}} \quad (3.9)$$

Where H is the material's hardness being measured, E is the elastic modulus, and R is the radius of the stylus' tip, W is the normal force applied to the surface by the stylus, which is often taken as the basic (static) load from the instrument but should really include dynamic (inertial) effects as well. No damage should occur if $\psi < 1$.

During the period when Whitehouse's research was mainly focused on the kinematics theory the researches on different materials of stylus were also in progress. To follow on from the previous work, because of internal springs and probe acceleration driven by the surface, forces tend to be larger near to the peaks. So far there is no useful guideline beyond the simple damage index.

Drawing on analogy between a ball contact (as in the E-system) and a stylus tip, it is obvious that any surface features characterised by a dimension smaller or similar to that of the tip will be distorted. Small valleys will be bridged, sharp peaks made to appear non-rounded and so on. The stylus acts as a highly non-linear filter of short-wavelength features.

(McCool, 1984) was one of the first to study theoretically the influence of stylus tip radius on surface roughness measurements. One simulation model had been developed by him that magnitude of distortion can be assessed as well as the effect of record sampling frequency and length with tracing surface profiles. However, these are realizations of random processes having an indicated spectrum and do not apply generally.

These have been plenty of other attempts at modelling stylus contact since McCool, but we still have little practical understanding of how to predict the level of distortion in general.

Several publications were proposed that the measured surface could be “corrected” for distortion based on knowledge of the tip size. Just as in work done by Schwarz (Schwarz, 1994), the suggestion on different techniques for in-situ characterization of probe was presented which can minimise the probe distortions (Whitehouse, 2000). While valid for a few special measurement conditions, it is at best dubious for general surface arrangement because (1) in the highly non-linear distortion measuring there will not be a single inverse operation and (2) in practice the tip size and shape are not known to good precision.

Chapter 4
4. Baseline software tool for stylus contact simulation

4.1. Introduction

Interest in measuring surface roughness continues to increase for several reasons. Efficiency, lubrication, corrosion stability, and probability of crack formation are examples of where the roughness of produced surfaces has significant influence on their performance.

Surface micro-topography is measured most commonly by either a fine stylus contact or by one of several optical probing methods. Debate continues on their relative merits and demerits, but it is clear that both will be used for a long while to come. There is, therefore, a need for reliable ways of comparing the output of these methods. There is also a need for better guidelines or standards for applying styli to the fine or delicate surface structures that are increasingly demanded by applications across mechanical, electronic, optical and bio-medical engineering. As part of continuing work towards such ends, this project is mainly to study the geometrical interaction of styli with engineering surfaces.

The contact of a small probe-tip against a surface is crucial to stylus-based surface roughness metrology, which remains a popular and important practical technique for both industry use and basic research. A finite stylus reduces measurement fidelity through direct geometrical effects, elastic distortion and dynamic effects that vary the contact force, sometimes to lift-off. It is difficult to determine directly the complex interactions that occur between a profilometer stylus and a surface.

Obviously, surface features with dimensions, curvatures, *etc*. smaller than or even somewhat larger than those of the tip will be distorted during measurement. Therefore, simulations have concentrated on how the stylus contact modifies the truly existing topography. Except for some statistics reported in Dowdier and Chetwynd (2002), published work tends to ignore the effects on the stylus tip itself. Appropriate 3D simulations have awaited adequate low-cost computing power. It is now practicable to run numerical experiments that may: cast light on questions of in-plane uncertainty of stylus measurement; suggest improved ways to compare different instrument probing technologies; or lead to methods for self-diagnosis of stylus wear (Radhakrishnan, 1971, Althagafy and Chetwynd, 2011). Figure 4.1 shows a schematic of a stylus with radius R making a contact with a rough surface.

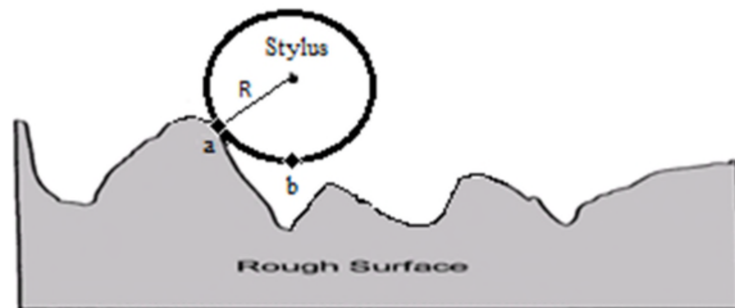


Figure 4.1: Contact points of the stylus with rough surface

All scanning surface instruments record the topography based on the assumption that contact is always at the centre line of the stylus tip at point b. They also in effect report the local height as that of the lowest point, b. This is almost always the case with very sharp tips, as in Atomic force Microscopy (AFM). In reality, this doesn't always happen with more general, robust instruments as the stylus could be touching the surface elsewhere, such as at point a, and sometimes might touch the surface simultaneously at more than one point. The likelihood of this happening

will depend on the stylus geometry (shape and size, including wear damage) and the roughness topography of the measured surface.

4.2. Kinematic and Threshold Models

Previous stylus-surface interaction simulations have concentrated on how the apparent (actually measured) topography is modified by the stylus contact from what truly exists.

There appears to have been very little research using such simulation methods to examine the effects on the stylus tip itself. This has been viewed as a lower priority question, but also it only makes sense to study it as a 3D simulation and so it has awaited the arrival of sufficient low-cost computing power.

It is timely, therefore, to give it more attention. For instance, it offers potential to help build simple empirical models that might cast light on questions of lateral uncertainty in stylus measurement, suggest ways to improve comparisons between different instruments, or lead to methods for self-diagnosing stylus wear.

The vertical distance between points and a and b on figure 4.1 contributes directly to the uncertainty of the height measurements and is often discussed in terms of a nonlinear filtering characteristic.

Less commonly considered is that the expected value of the horizontal distance between a and b is a major factor in the lateral uncertainties in the measured height maps.

The current work builds upon the pioneering work of a previously reported 3D stylus simulation (Althagafy and Chetwynd, 2011) that introduced a statistical summary of the points upon the stylus surface at which actual contact occurs.

It allowed use of both simulated non-ideally-shaped tips and, particularly, the use of high-resolution data from direct measurements of real styli (Radhakrishnan, 1971). Such extra data becomes important because (even with idealised stylus) the contact is rarely directly along the (metrologically assumed) central axis of the stylus system. The new simulation system introduces several refinements to this modelling approach. Firstly, it is clear that many simulations with subtle differences will be necessary to generate a statistically valid summary that could be compared to real instrument behaviour in order to infer the accuracy of the models. A highly modularized software suite is used, with bespoke parts implemented in MATLAB®.

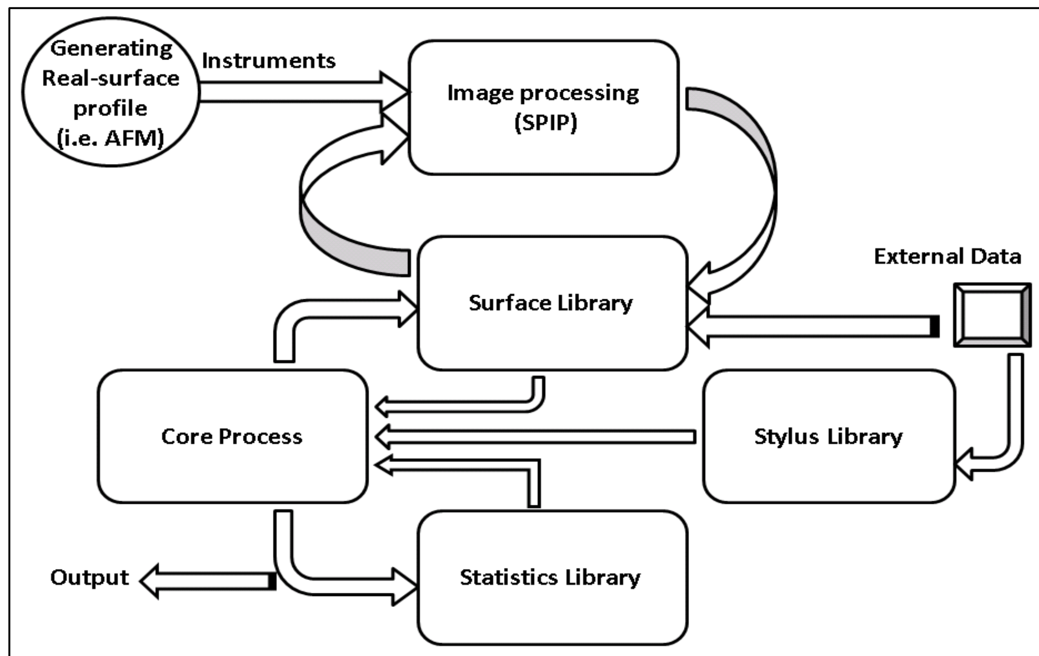


Figure 4.2: Scheme for the software suite.

Extensive use is made of libraries of surface maps: in addition to simulated and measured 'original surfaces' every 'modified surface' processed here is also entered, so giving a complete record and also enabling study of multiple stylus actions. There is a similar library of 'styli'. Every simulation runs by loading an appropriate pair from the libraries, processing them by a core operation and returning modified versions and statistical information. Higher level processing of surface information (for parameterizations, comparisons, etc.) is performed by a commercial package (in our case, SPIP®), which acts as an independent standard. This approach both maximizes capabilities for inter-comparisons from different sources and minimizes the size and complexity of the bespoke core routines, thereby easing the task of verifying them. For the present, while developing and evaluating the practicality of the new ideas, we are content to work with small maps, up to around 128 points square, to limit computational costs.

The underlying starting point for modelling probe contact remains essentially the kinematic model used previously: an algorithm is partially published in (Muralikrishnan, Raja, 2009). It is often called the Kinematic model because it is purely geometrical, assuming that: both stylus and surface are perfectly rigid; stylus and surface maps have parallel z axes; the stylus motion is constrained to be translation along the z axis; and a single point of contact occurs on the stylus moves towards the surface. The physical concept is that the central z-axis of a stylus map is positioned in turn above each (x, y) position on a surface map and then lowered along the z-axis until first contact is detected. The 'height' of this contact condition is then taken as the position of a fixed reference point on the z-axis of the stylus; e.g., point b on figure 4.1. This contact may not be on the stylus z-axis but its height

is recorded as the modified surface point at (x, y) on the surface map: this corresponds to where a real instrument would report such a point. Computationally, this is a dilatation operation that reduces to an array addition and search for a maximum.

The first enhancement is to record also the position(s) (note that there might be more than one point of contact, e.g., where a sphere enters a triangular groove) of the contact within the stylus array. This can be done compactly by simply building a 2D histogram of contact numbers at each point (Dowidar and Chetwynd, 2002), but it provides only a general statistical summary. More valuable information is gained by maintaining a history of contacts that correlates directly with the points on the surface map. It is then possible to revisit the same data sets using a separate bespoke module to explore in detail, e.g., the lateral error distribution between actual and reported height positions on the modified surface.

The kinematic model is unrepresentative and over-prescriptive because, among others, (i) real surface data is subject to instrument noise and so fine differences associated with the very peaks will not repeat; and (ii) neither surfaces nor styli are perfectly rigid and there will be elastic (possibly plastic) deformation at points of contact. Furthermore, it might be over-sensitive to the resolution of the surface data supplied by a specific instrument; a concern for measurement standards. A new threshold feature has been added to the kinematic model to address such points. On determining the kinematic condition, a search is made for other points which come within a specified distance of the stylus. This has immediate, obvious application in exploring how potential contacts might be lost or gained if resolution is reduced. It can also identify how many ‘contacts’ might be totally within the noise-floor of

an instrument (physically the routine is one-sided, but the dilation has already effectively treated the noise as one-sided).

Thresholding can also provide a first step towards more sophisticated models for approximating stiff but non-rigid contact. By a type of relaxation method, iterative modifications to the threshold might illustrate how the stylus settles elastically onto the surface. Each change of threshold would now be governed, e.g., by a force-balance using localized Hertzian contacts. Further details of these ideas and their practical implementation will be discussed later in chapters 5, 6 and 7.

4.3. Atomic Force Microscopy (AFM)

AFM is one type of a relatively new family of instruments called "scanning probe microscopes" (SPM). In the most general sense, these microscopes reveal information about the surface properties of materials by scanning the surface with a small probe. Information about the interaction of the probe with the surface is transmitted, via an Electronic Interface Unit, to a computer where an image of the surface is generated.

The first invention of this kind of instrument was the scanning tunneling microscope (STM), invented in the early 1980s. In the tunnelling microscope, a sharp metal tip is brought to within a few angstroms of a conducting or semi conducting surface. If a voltage is applied to the sample, an electron tunnelling current flows between the tip and surface, which varies exponentially with the distance between them.

The surface topography is revealed by scanning the tip over the surface in parallel lines while keeping the current constant (i.e., keep the tip-surface gap constant).

The STM is most commonly used on conducting surfaces. The Atomic Force Microscope or AFM was developed a few years after the STM. Unlike the STM, the surface studied by the AFM need not be conducting. The surface is scanned by a probe which may touch the surface or be a few angstroms away. The force between the probe and sample is used to generate an image of the surface topography.

The forces generated in an AFM are several orders of magnitude lower than in conventional stylus surface topography instrument and, also, the contact area between tip and surface is much smaller in an AFM. The net result is that whereas "profilometry" often risks damage to the surface because of the large forces and contact surface area, AFM reveals the surface topography and at the same time leaves the surface undamaged.

The techniques being developed and investigated in this work focus strongly on the application to conventional stylus instruments. They will have at most a slight relevance to high lateral resolution techniques such as AFM.

However, a specific challenge for simulation studies is to obtain truly representative data at high lateral resolution from surfaces with quite high roughness amplitudes.

AFM can provide this data but is not ideally studied to the target surfaces; there is a significantly increased risk of tip damage. On the other hand, the required resolution does not approach the limits of AFM capability.

Consequently, it was convenient to use an elderly instrument of modest performance (Quesant, Q-Scope™ 788). This section briefly reviews the use of AFM in the context of the needs for this thesis.

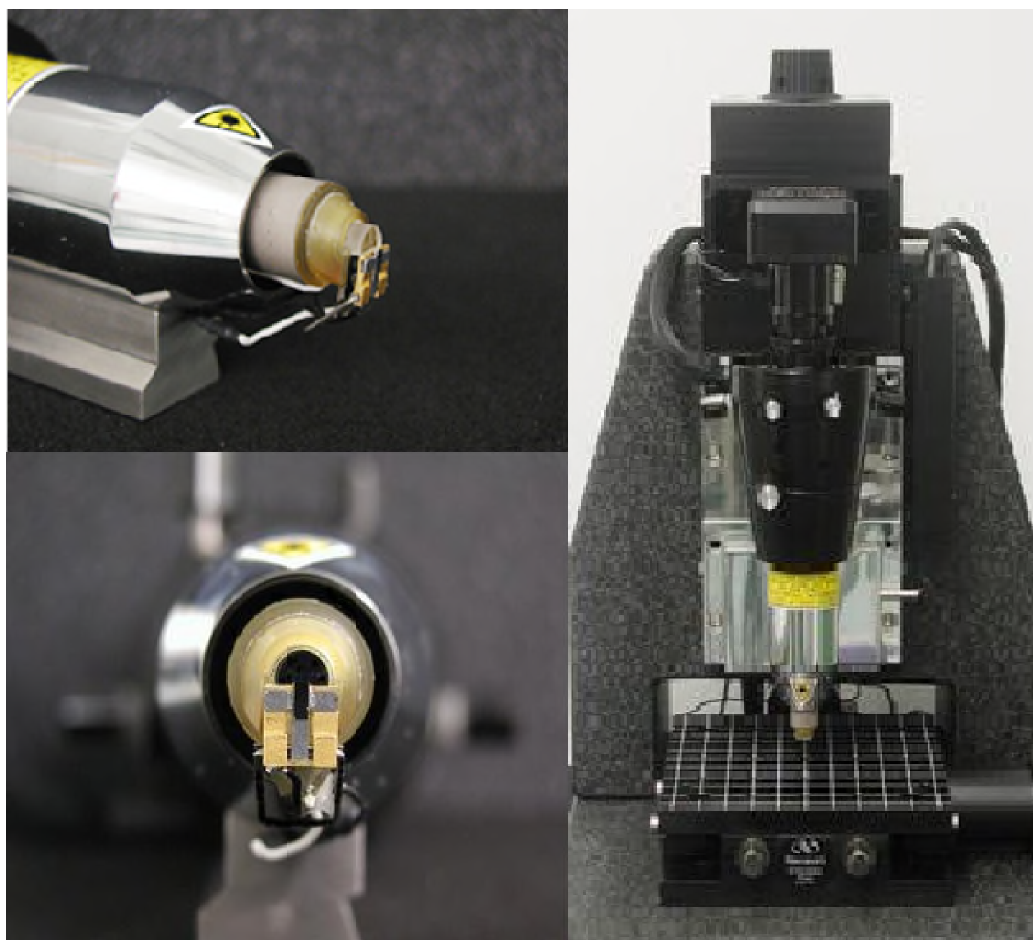


Figure 4.3: AFM (right), Probe holder (left) (Quesant, Q-Scope™ 788)

In essence, the AFM (see figure 4.3) consists of a small tip or probe mounted in a scanner. A stage supporting the scanner, an Electronic Interface Unit (EIU) and a computer complete the system. As the probe is "rastered" over the surface, the EIU maintains a condition of constant force between surface and probe by adjusting the height (*z* dimension) of the probe. The movement of the tip over the surface is controlled by a piezoelectric ceramic (or "tube scanner"), which can move in the *x, y,* and *z* directions in response to applied voltages. Movement of the piezo in the *x* and *y* directions scan the sample. A feedback circuit controls voltage applied to the *z* piezo so that the signal is held constant as the tip is scanned across the surface.

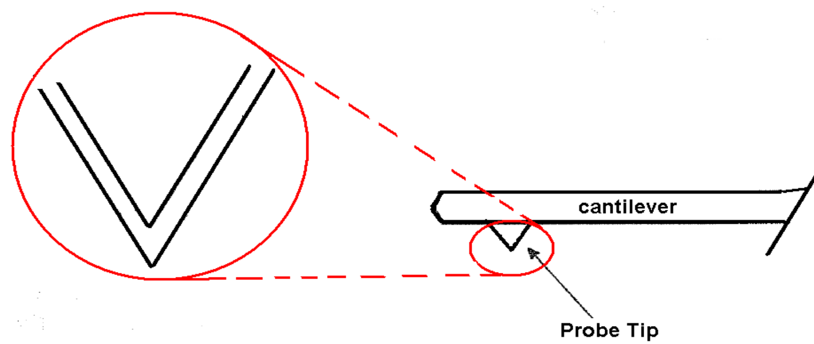


Figure 4.4: Probe Cantilever (right), Probe Tip enlarged in (left)

In the AFM, the tip is attached to a spring in the form of a cantilever. As the tip moves over the surface, the cantilever bends back and forth in the z direction. A laser beam is directed onto the cantilever and as the cantilever bends, the movement of the reflected beam is detected by a photo diode. In the usual mode of operation, constant force is related to constant cantilever deflection and so the piezo serves to keep a constant output from the photodetector.

The image of the surface is built up as a series of scan lines, each displaced in the y direction from the previous one. Each individual line is a plot of the voltage applied to the z piezo as a function of the voltage applied to the *x* piezo (Quesant, Q-Scope™ 788).

4.3.1. AFM Scan Parameters

It is important to choose the most appropriate instrumental variables for the scanned sample. The scan parameters are listed in table 4.1:

Table 4.1: AFM Scan parameters

Parameter	**Range**	**Used**
Scan Size	Max:20 µm	20 µm
Scan Rate	0.5 to 20 Hz	2 Hz
Z Dynamic Range	Max: 4.18 µm	4.18 µm
Scan Resolution	100 to 600	600
Scan Direction	0 to 359	90°
Image Mode	Z-height or Lateral Only	Z-height
Scan Profile	(Hills & Valleys)	Hills & Valleys

Scan Size indicates the square root of the area scanned. Thus, a value of 20 µm means a 20 µm x 20 µm or 400 μm^2 area is scanned. The smaller the value, the higher the magnification.

Scan Rate is the number of horizontal scan lines per second. In samples with sharp features a slow scan rate, e.g. 0.5 Hz, may be advantageous because the probe tip has more time to react to sharp image features. On the other hand, the slower the scan rate the longer the time required to acquire the image and greater vulnerability to thermal drift.

Scan Resolution is the number of data points per horizontal line, which is also equal to the number of lines within the scan. Higher scan resolutions reveal greater image detail but also take longer to acquire and process.

Z Dynamic Range affects the maximum Z height and the resolution achievable in Z. Its value depends on the type of scanning head used. The AFM used has a 40 μm head and the maximum Z height 4.18 μm.

Scan Direction is in degrees and is the direction of scan relative to the stage. The 90^0 setting scans from west to east and 0^0 scans from north to south.

Image Mode refers to the way Z height features are measured.

Scan Profile affects the relative partitioning above and below the sample plane of the available Z Dynamic Range which has been selected. The approach used was Hills and Valleys, which partitions Z equally above and below the sample plane. For the AFM used the Z Dynamic Range was 4.18 μm, and so Z can extend maximally 2.09 μm in each direction. Selecting Hills allows more available Z range above the sample plane, whereas Valleys do the opposite.

4.3.2. Approaching the Sample

The following procedure has been used to bring the probe tip to within half a millimeter or less from the surface of the sample.

The sample has been placed on the stage so that it is roughly under the probe tip. This will be to the left of the laser spot on the stage. Then, the probe tip is brought down to within about 0.5 mm or less from the sample, controlled manually and by eye.

From this point, using the soft touch feature in this AFM allows a fully automatic approach to the surface to prevent damaging of the sample surface or the probe.

4.3.3. Fourier Transforming Images

The objective of Fourier transformation is (in the present context) to remove unwanted interference and noise from the image. Interference is usually seen as periodic lines, whereas noise is often specks or dots scattered randomly throughout the image. Two dimensional Fourier transformation performs a mathematical operation on the image, allowing one to remove unwanted interference and noise by separating them in terms of spatial frequencies from the information which constitutes the image.

To process an image, a Fourier transform of the image should be applied. After the image has been transformed there are several ways to remove interfering frequencies.

4.3.4. Removing Tilt and Low Frequency Interference

As the image shows broad bands suggesting low frequency interference, AFM tilt removal feature has been used. Line by Line/H and Line by Line/V remove, respectively, low frequency horizontal and vertical interference.

Line by Line/H calculates the average height of all of the points on each scan line of the image and sets this value to zero. Line by Line/V calculates the average height of image lines (starting from the left) formed by pixels in the vertical direction.

Tilt in the image is caused by a deviation of the sample plane from the scanning plane. In general, it is desirable to remove tilt from an image, a process which is also referred to as levelling. The instrument software provides several automatic or semi-automatic levelling procedures; commonly, a least squares best-fit plane

would be used. Generally the simulations in this thesis take data from either a ground or lapped steel surface, see figures 4.5 to 4.8.

Figure 4.5: Samples used in this project

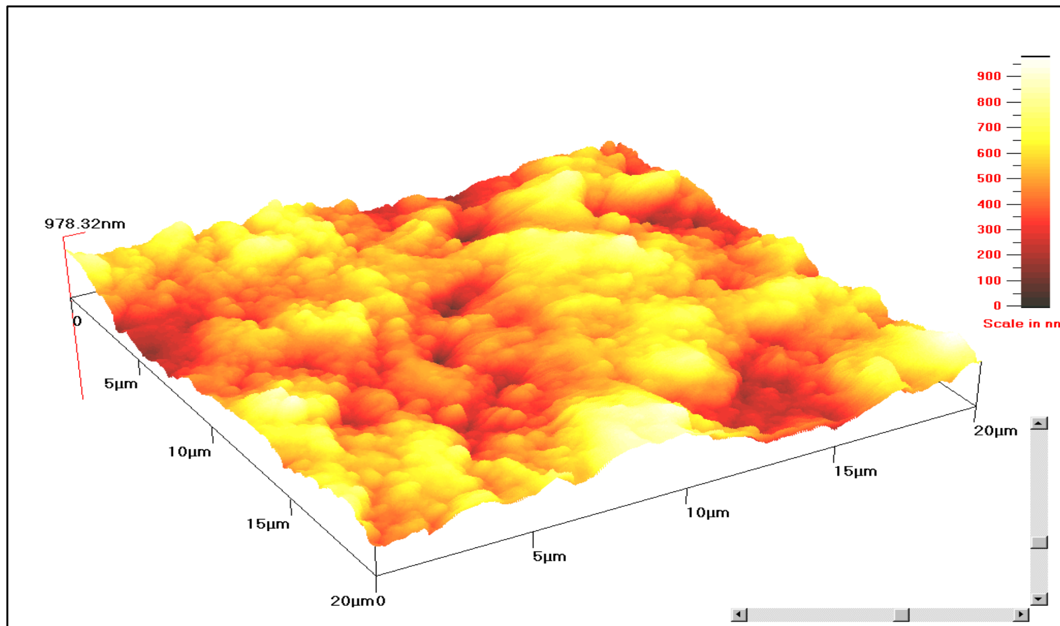


Figure 4.6: 3D lapped surface profile by AFM (Sample 1)

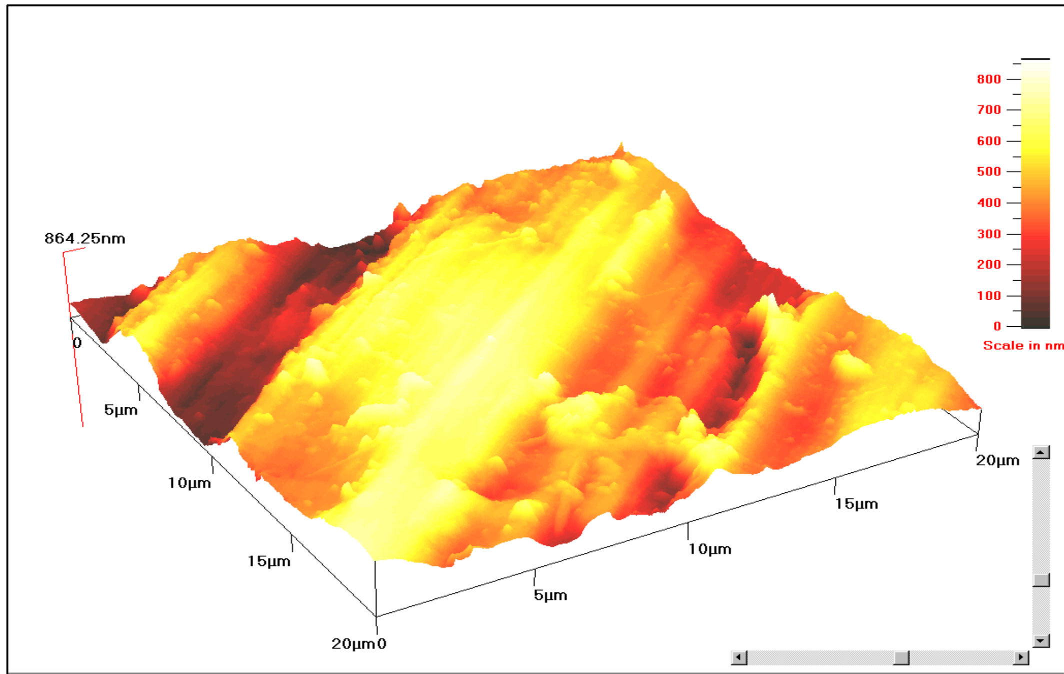


Figure 4.7: 3D Ground surface profile by AFM (sample 2A)

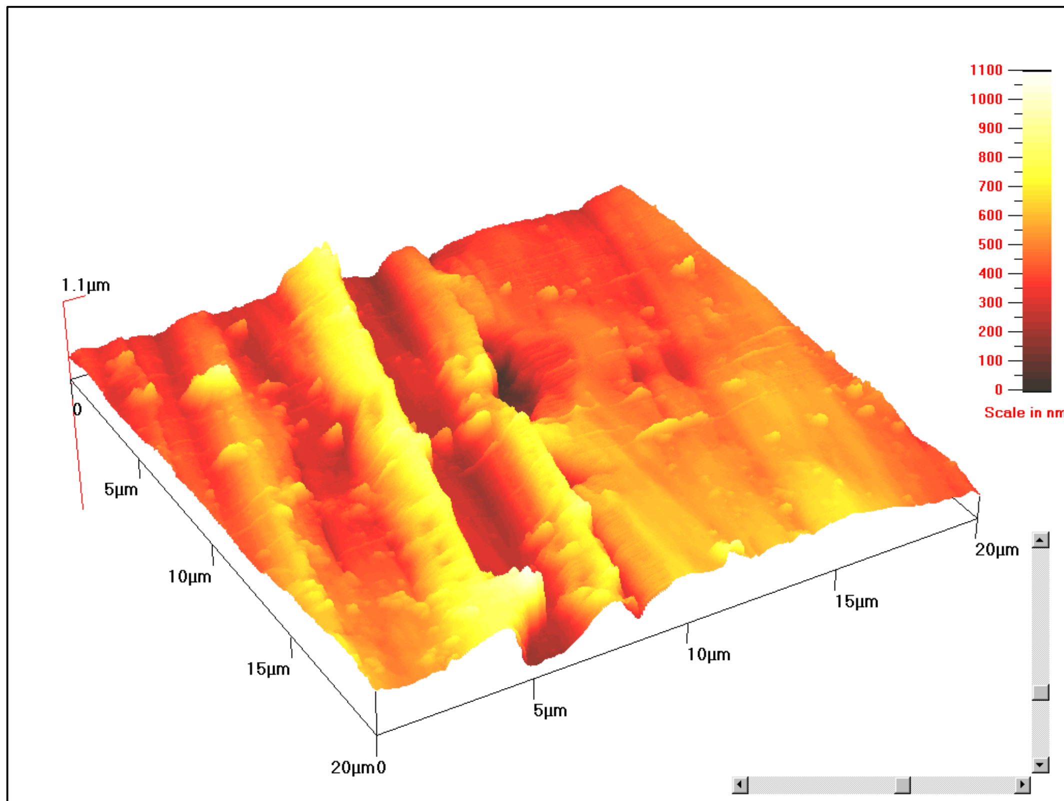


Figure 4.8: 3D Ground surface profile by AFM (sample 2B)

4.4. Numerical Simulation Process

In this chapter a numerical simulation process is used to gain data on the filtering effects of the stylus and on the distribution of contact points. The aim is to inform choices of stylus geometry (i.e. tip radius and cone angles) for the benefits of instrument design and uses and especially, for the information of standards organizations. Longer term, the simulation may help the development of instrument diagnostics for stylus wear and damage.

Figure 4.9 shows in 2D the effect of the stylus tip size on the measured profile of a surface. When measuring the real profile with a spherical tip, the measuring system will report the locus of a certain point of the tip normally the centre point. However, if the surface profile is scanned with the small tip, the locus of the centre point of the tip will represent the measured profile (locus 1).

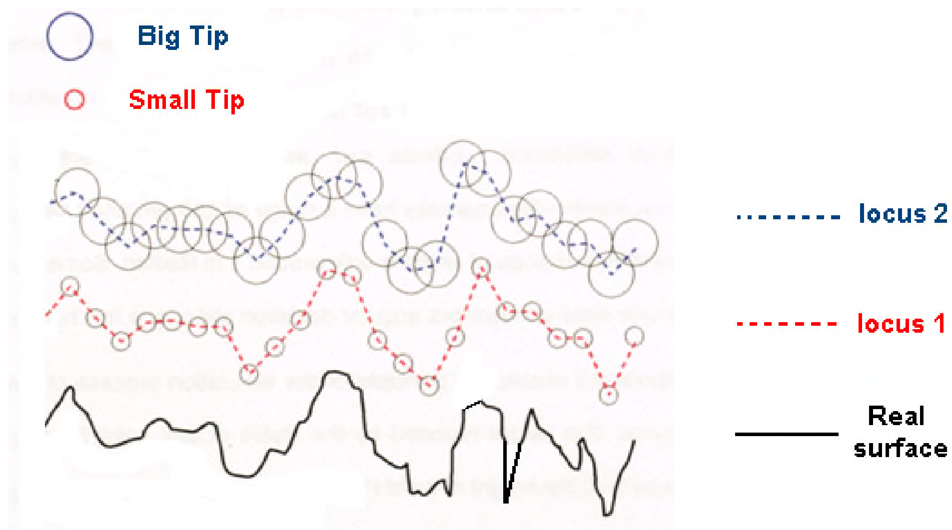


Figure 4.9: The effect of stylus tip radius on the measured surface profile

Obviously, the measured profile will be different from the real profile as the valleys tend to be sharpened while the peaks tend to be flattened. Rescanning the same real

profile with a bigger stylus tip produces another profile (locus 2) which is more different from the real profile. It is clear that locus 2 is more different from locus 1, which is of course logical as the smaller tip will be able to go deeper into the valleys than the bigger tip.

Before discussing the details of the simulation method and its practical implementation, it is useful to set the context by means of a qualitative physical summary of the process:

1. Assume an array representing a uniformly distributed set of spot height measurements that form a map of the surface topography and a much smaller map representing the shape of the stylus by a set of heights;

2. Choose a point on the surface map at which the stylus will be applied. In practical terms this can be done by extracting a sub-map (partition of the array) of the same size as the stylus array.

3. Set the 'stylus' above the 'surface' and determine the gap between them. In practical terms this amounts to taking the difference, point for point between the stylus and sub-surface arrays;

4. The position at which this gap is smallest indicates where the stylus would first make (kinematic) contact with the surface were it to be lowered vertically towards it;

5. The height of the centre-point of the stylus tip will then be the height of the surface contact point reduced by the difference between the heights of the stylus contact point and the stylus central point;

6. Record this contact data for later use and write the stylus height into the relevant position of an array for the modified surface;

7. Move to the next point of the surface by looping to 2 until the whole array is processed;

8. Note also that various further analyses can be built into this scheme, such as the introduction of a thresholding method proposed here for the first time to explore the sensitivity of the method to alternative surface features."

The basic core of the numerical simulation method is a simple extension of that used for many years to show how a circular (disc) 'stylus' filters (smooths) a roughness profile (Radhakrishnan, 1971). The extension from these 2-D methods to 3-D (real) surface modification is obvious, but appears to be hardly reported in the literature. Perhaps computing costs for the perceived benefit have been discouraging. It seems likely that a few investigations into statistical methods of predicting stylus filtering effects may have used it for testing without fully describing the techniques used (Wu, 1999).

The simulation is a kinematic process, based on the geometry of ideally rigid materials, etc. Consider first the 2D vision of a disc scanning surface profile, which is easy to visualize. A tip is lowered vertically towards a nominally horizontal test surface until a first contact point is encountered. The measured (reported) profile point is then recorded as the height of a reference point on the stylus (e. g. centre of spherical tip or lowest point) positioned at the lateral location of the centre of the tip. We assume that real styli are adequately constrained by their guiding mechanisms to be modelled as single degree of freedom systems. The slightly arcuate motion of many real profilometers is neglected. This is acceptable on spherical tips and, although less rigorous, for other shapes over small deflections. Effects such as contact friction are also ignored. Conventional instruments might

typically have a range of ±50 µm and use a 50 mm stylus arm, so the maximum change of angle is only around (1 m radian). Some modern, long range profilers cause more concern but angular deviation is still only a few hundredths of a radian (Dowidar, 2003).

Figure 4.10 shows the principle of the simulation process of measuring a surface with a stylus. The height reported by the stylus at any point P of the surface will not always represent the height of point (CL) unless the contact occurs exactly on the top of point (CL) with the centre of the stylus tip (point 5). The contact could happen at any point of the surface within the size of the stylus tip. To find out the actual height of the stylus at any point (CL) of the surface, the height of each point of the stylus is added to the height of each opposite point of the surface. The maximum value of the determined heights will represent actual height of the stylus when measuring point (CL). It will also give which point of the stylus it is that makes actual contact with the surface.

The stylus is then moved to the next point of the surface and its actual height is determined as well as the points of the stylus which make contact with the surface. This process is repeated on all surface points.

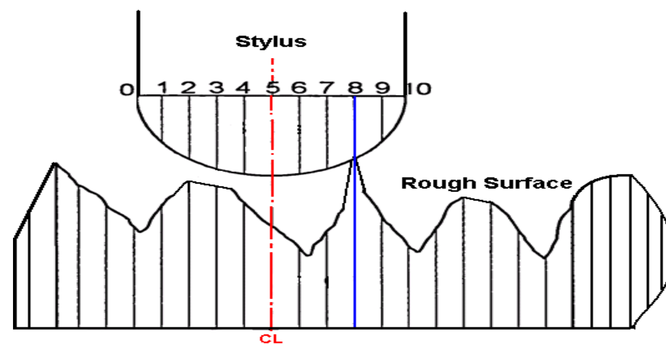


Figure 4.10: The principle of the simulation process of measuring a surface with a stylus

4.4.1. Simulation software technique

In this simulation, the surface could be any set of data representing a real or an arbitrary simulated surface. Also, the stylus could be any set of data representing a real or an arbitrary stylus shape. Both data sets of the surface and the stylus are dealt with as arrays.

Figure 4.11 shows the main features of the algorithm used in the new implementation of the simulation process. The surface and the stylus are represented by two sampling grids in arrays Z(N, M) and S(d,d) respectively, where N is the number of traces of the surface, M is the number of points in each trace and d is the tip diameter (for the ball, d=2r). In principle, Z is just an array of regularly spaced points representing x and y coordinates, with no preference of one over the others. However, it remains the case that most stylus instruments do indeed scan the surface in a series of parallel traces ('fast' and 'slow' axes!). Two new arrays L(N, M) and C(d,d) are created. The array L(N, M) has the same size as the surface array Z(N, M) and is used to report the locus of the stylus on each point of the surface. The other array C(d,d) has the same size as the stylus array S(d,d) and is used for counting the number of times that each point of the stylus has made contact with the surface. The central point of the stylus array S is set, in turn, above each point of the surface array Z(i, j). The actual height of the stylus is determined on each point of the surface Z(i, j) and is stored into the array L(i, j). Stylus points making contacts with the surface are determined. The counter C(a, b) of each point making contact with the surface will be increased by one while other points' counters will remain the same. The output of the simulation process will be the array L(N, M) representing the locus of the centre of the stylus S(d,d) when

scanning the surface Z(N, M) and the array C(d,d) representing the number of times that each point of the stylus has contacted the surface. The locus array will be compared to the actual surface array, either graphically or quantitatively to show the effect of the stylus geometry on the roughness parameters of the scanned surface. The contact array will show which part of the stylus is most often making contact with the surface and that helps in defining the appropriate stylus shapes for scanning such surfaces. The algorithm is implemented in MATLAB for ease of array handling, etc., at the expense of relatively poor computational speed. Appendix A shows the Matlab program used for the simulation process.

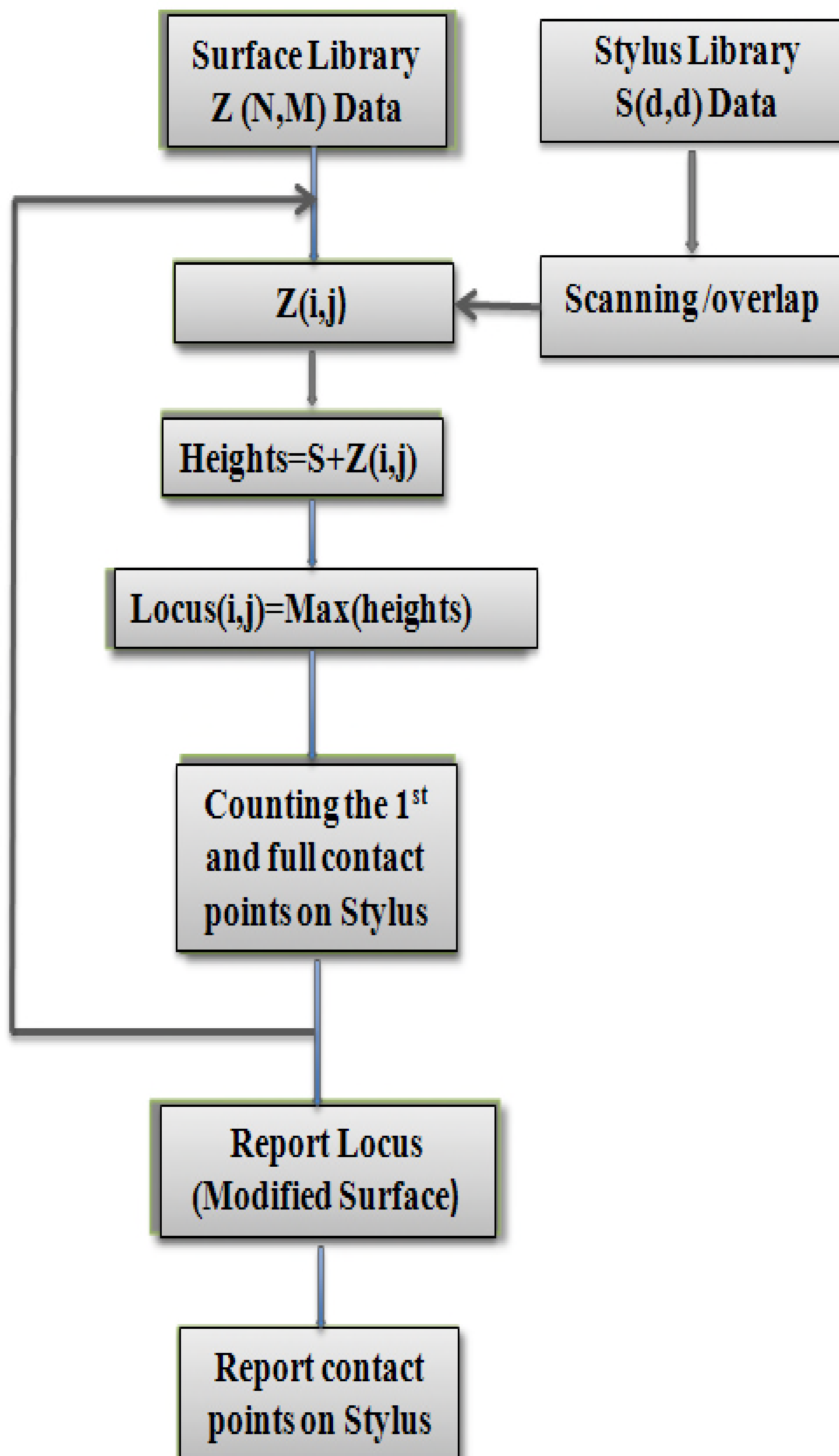


Figure 4.11: Summary of the numerical simulation algorithm

The way any particular software system (in the present case MATLAB) handles array indexing can offer advantages to the implementation of the stylus simulation, but can also place minor constraints on it. In the latter for example, it would be physically intuitive to consider the stylus shape map moving out from a central (z-axis) point identified as (0,0), but MATLAB indices two-dimensional arrays from (1,1) at one corner (which its internal plotting routines treat as "bottom, left"). On the other hand, its feature of enabling two-dimensional arrays to be handled consistently through a single index allows a more compact way to store the extra evaluation data that the new implementation provides. It is, therefore, worth reviewing the indexing ideas that will be used in some detail.

Generically, the description of the surface (or stylus) topography comprises a collection of spot heights (h or z) at a specific position (x, y) in a nominal reference plane; full description requires real-valued triples (x, y, z). A major simplification is provided by forcing the height measurements to be regularly spaced on a rectangular grid along the x and y-axes. Slightly greater simplification occurs if this spacing is forced to be the same along both x and y. The map can then be treated as a simple (real-valued) matrix, or 2D arrays, with integer indexing and with true position recoverable from the specific index and sampling interval. Thus on explicit triple for the spot height at a specific location (p_x p_y h_2), which could also be expressed h_2 (p_x , p_y) can be expressed, subject to nothing worse than an implied and consistent shift of origin, in array terms as Z(i,j) where p_x= i dx and p_y=j dy. Dx and dy are the real-valued sampling intervals along the axes, while i and j are integers.

The sampling intervals dx and dy must be the same in both the stylus array (s) and the surface array (z). They must represent values in the same units

The stylus shape is held in a square array S(d,d) for the common case where dx=dy, which is used throughout the present work. For computational efficiency, d should be kept small but it must nevertheless extend far enough to encompass all possible points of contact between the stylus and surface; for example, for a ball-tip of radius r, we might take d.dx=2r.

The basic kinematic stylus simulation (dilation) proceeds by a process equivalent to physically placing the centre-line of the stylus above each point of the surface in turn and each time recording the relative height of the stylus where it makes contact with the surface. The array of these heights forms the map of a modified surface (or stylus locus). The actual realization of this process used here is to extend from the surface map a suitably positioned sub-array having the same dimensions as the stylus array, compute the contact height for this sub-set, and record the height in the equivalent location of the modified surface array. This particular approach has two advantages. First, it allows effective exploitation of some in-built MATLAB functionality. The second is more subtle: while recording a fully-modified surface requires extracting a sub-array for each point of the original map, many statistical studies can operate with evaluations taken at a much smaller number of positions, which is made easy by this method.

The core algorithm has an outer structure that selects for processing appropriate points within the surface map, Z. In the case of a full scan to generate a modified surface, this will comprise nested loops incrementing through all index positions.

Assume, then, that a sub-array ZZ, of the same size as the stylus array S, has been extracted from Z at a position of interest Z(I,J). Since the kinematic model allows stylus motion only along the z-axis, the physical intuition of the next step is to place the stylus at an arbitrary distance directly above the extracted surface and then evaluate the size of the gap between surface and stylus at each point; the first contact will necessarily coincide with the position at which this gap is smallest.

For convenient visualisation (while noting that practical coding may benefit from other conventions) assume that S and ZZ are indexed with (0,0) being their central point. Since arbitrary shift on the z-axis is of no mathematical concern, we can then compute simply

$$Gap = S - ZZ$$

Note that this might be a sum or difference depending upon choice of plausible sign conventions in describing stylus shape. The minimum value of gap is found to be at position (i,j) and the reported stylus height will then be

$$h_{tip} = ZZ(i,j) + (S(i,j) - S(0,0))$$

This value is recorded as the surface output in the locus array at position L(I,J). Meanwhile, element C(i,j) in a counter array having the same dimension as S is incremented, therefore, building up a 3D histogram of how often each point of S is contacted.

Finally, a relative contact location is recorded in CL, having the same dimension as L, at the relevant point, for example:

$$CL(I,J) = i + Nj$$

where S is of size NxN. As long as S and Z are retained, this suffices to reconstruct any desired part of the contact history.

For visualization and for most parts of the coding, it is best to treat surface (and stylus) arrays in terms of (i,j) indices, that is in orthogonal x and y coordinates. However, it is sometimes convenient to think of the whole map as a linear array, row following row, with a single index P. Thus, we might use:

$$P= i+M(j-1)$$

If the I and j indices start at 1, MATLAB adopts this structure internally and either indexing system can be used.

The Kinematic simulation of stylus contact used to date, essentially and once the first contact (minimum gap) position is determined for each selected point on the surface map and an equivalent map of the locus of the reference point on the stylus logged into the same index position in the output (modified surface) array. While being moderately informative of how a stylus might affect a specific measurement, this kinematic model risks placing too much emphasis on an overly particular set of conditions. For example, because of noise and other measurement uncertainties any attempt to repeat the exercise on nominally the same surface might yield rather a different result even if the model stylus is exactly the same. Also, of course, real surfaces are not perfectly rigid and a real stylus will always settle slightly below its first contact through electric (and perhaps even plastic) deformation.

Consequently, a significant innovation within the present work is to introduce an additional thresholding phase after the kinematic process is complete. While simple to implement in software, it opens up a whole range of possibilities for exploring

stability and sensitivity of the modelling and adds further depth to the investigation of any proposed new stylus scheme.

The idealised physical intuition of the method is that once first contact is established, the stylus is forced to move downwards by a small distance, the threshold. The surface is treated as totally yielding to this operation, which might (or might not), therefore load to larges areas of the stylus intersecting with the original surface. By exploring the effects of the different threshold, we can evaluate the degree to which the kinematic contact is representative, or even possibly dominant.

Mathematically, the process involves only a slight extension. With the matrix gap computed for any specific stylus and extracted surface sub-matrix, we can write:

$$\text{thr - gap} = \text{gap} - (M_{ref} + Ths)$$

where, *Ths* is a thresholding constant. All points in (thr–gap) which have negative or zero values are then regarded as a parts of the thresholded contact. …

New challenges arise in providing ergonomically convenient ways to visualize the growth of contact regions as the threshold distance is increased, and describing the contact regions parametrically. One simple but effective way for handing the contact over a single stylus-sized patch is simply to define an array which has (one) wherever thr-gap ≤ 0 and (zero) elsewhere. (Actually, this is very easily achieved in in MATLAB because of the way it handles conditioned and logic operations).

The inbuilt plotting routines can then display an effective black and white plan map of which points are in contact. Of course, it would equally be easy just to set to zero

all positive gaps, leaving the negative values to reflect the relative weight of how influential each point in the contact region is likely to be. Practical experience so far suggests that this extra sophistication adds insufficient insight to be worth including.

4.5. Summary

This study has introduced a new implementation of a suite of tools for studying stylus-surface interaction in simulation. A consistent pattern of data flows and archiving around a highly modular organization is advocated, allowing new features to be introduced in small, independently testable, modules. In this respect, the first features included and not previously reported concern the mapping of contact extent and intensity across the probe surface and the use of thresholding to investigate sensitivities to noise and instrument resolution. The concepts are illustrated for clarity by a highly artificial flat probe and large parameter values for which many effects can be understood intuitively, as well as being readily visible on small-scale plots. However, even with this unrepresentative approach, the thresholds reveal information about the surface geometry that may be of functional significance.

The system is now being refined to add further features while simultaneously being used with more representative probe models to investigate, for example, the statistical significance of contact distributions.

The simulation program has been developed and used to simulate the real and non-real surfaces roughness measuring process by the stylus method. It can be used to scan any arbitrary surface with any arbitrary stylus tips shape. The simulation is not only used to measure the roughness of the surface but also to show the contacts

distribution on the tip when scanning a surface, information not previously considered. The theoretical results of the simulation in the following chapters have confirmed that the stylus geometry can have a significant effect on most roughness parameters of the measured surface in 3D. The contacts distribution on the tip has shown that most contacts do not occur at the central point of the stylus in most cases, even with idealized shapes

.

Chapter 5
5. Sub-Micrometer contact of topography Styli

5.1. Introduction

Stylus-based surface roughness metrology remains a popular and important practical technique for both industrial use and basic research, being complementary to optical methods. The nature of the contact at a small probe-tip is of clear concern to the fidelity of measurement, both directly in terms of potential distortion of functionally important data and for the inter-comparison of results from different instruments. Broadly similar concerns arise in some types of micro-tribology, but this study concentrates on the former, while anticipating other more general applications. It is intuitively obvious that any surface features having dimensions, curvatures and so on smaller than or even somewhat larger than those of the stylus tip will be distorted during measurement. Often regarded simply as a highly non-linear low-pass filter having (at best) poorly determined parameters, the probe interaction also has poorly understood effects on the measurement uncertainties, both normal to the surface (often called 'vertically') and in the surface plane ('horizontally' or 'laterally').

There are obvious, extreme challenges in attempting direct determinations of the complex interactions that occur between a profilometer stylus and a surface during a measurement of surface micro-topography. There is more hope that comparisons between real results and suitable numerical simulations might reveal indirect evidence about behaviour in the contact region. There is a long history (dating back to the obsolete E-system for generating measurement reference lines) of simulating the loci of the centres of circles, and latterly spheres, rolling on rough topography

(see, for example, (Muralikrishnan , Raja, 2009) for a brief review). It has been used primarily as an evaluation tool, but the concept also appears in attempts to compensate for stylus distortion by 'stylus deconvolution'. However, current models lack the sophistication to properly address comparisons or uncertainty analyses. They are, for example, purely kinematic, assuming perfect guide axes, rigid probes and rigid surfaces. They are actually a class of morphological filter: specifically, the centre locus is dilatation operation (Muralikrishnan , Raja, 2009).

Previous stylus-surface interaction simulations have concentrated on how the apparent (actually measured) topography is modified by the stylus contact from what truly exists. There appears to have been very little research using such simulation methods to examine the effects on the stylus tip itself. This has been viewed as a lower priority question, but also it only makes sense to study it as a 3D simulation and so it has awaited the arrival of sufficient low-cost computing power. It is timely, therefore, to give it more attention. For instance, it offers potential to help build simple empirical models that might cast light on questions of lateral uncertainty in stylus measurement, suggest ways to improve comparisons between different instruments, or lead to methods for self-diagnosing stylus wear.

The current work builds upon a previously reported 3D stylus simulation (Dowidar and Chetwynd, 2002) that introduced a statistical summary of the points upon the stylus surface at which actual contact occurs. It allowed use of both simulated non-ideally-shaped tips and, particularly, the use of high-resolution data from direct measurements of real styli (Dowidar and Chetwynd, 2004). Such extra data

becomes important because the contact is rarely directly along the (metrologically assumed) central axis of the stylus system.

The new simulation system introduces several refinements to this modelling approach. Firstly, it is clear that many simulations with subtle differences will be necessary to generate a statistically valid summary that could be compared to real instrument behaviour in order to infer the accuracy of the models. A highly modularized software suite is used, with bespoke parts implemented in MATLAB®. Figure 5.1 illustrates the scheme. Extensive use is made of libraries of surface maps: in addition to simulated and measured 'original surfaces' every 'modified surface' processed here is also entered, so giving a complete record and also enabling study of multiple stylus actions. There is a similar library of 'styli'. Every simulation runs by loading an appropriate pair from the libraries, processing them by a core operation and returning modified versions and statistical information. Higher processing of surface information (for parameterizations, comparisons, *etc*.) is performed by a commercial package (in our case, SPIP®), which acts as an independent standard. This approach both maximizes capabilities for inter-comparisons from different sources and minimizes the size and complexity of the bespoke core routines, thereby easing the task of verifying them. For the present, we are content to work with small maps, up to around 128 points square, to limit computational costs.

The underlying starting point for modelling probe contact remains essentially the kinematic model used previously: an algorithm published in (Dowidar and Chetwynd, 2002). The physical concept is that the central z-axis of a stylus map is

positioned in turn above each (x, y) position on a surface map and then lowered along the z-axis until first contact is detected. This contact may not be on the z-axis but its height is recorded as the modified surface point at (x, y): this corresponds to where a real instrument would report such a point. Computationally, this dilatation operation reduces to an array addition and search for a maximum.

The first enhancement is to record also the position(s) (note that there might be more than one contact, e.g., where a sphere enters a triangular groove) of the contact within the stylus array. This can be done compactly by simply building a 2D histogram of contact numbers at each point (Althagafy and Chetwynd, 2011), but more valuable information is gained by maintaining a history of contacts that correlates with the points on the surface map. It is then possible to revisit the same data sets using a separate bespoke module to explore in detail, e.g., the lateral error distribution between actual and reported height positions on the modified surface.

The kinematic model is unrepresentative and over-prescriptive because, among others, (i) real surface data is subject to instrument noise and so fine differences associated with the very peaks will not repeat; and (ii) neither surfaces nor styli are perfectly rigid and there will be elastic (possibly plastic) deformation at points of contact. Furthermore, it might be over-sensitive to the resolution of the surface data supplied by a specific instrument; a concern for measurement standards. A new threshold feature has been added to the kinematic model to address such points. On determining the kinematic condition, a search is made for other points which come within a specified distance of the stylus. This has immediate, obvious application in exploring how potential contacts might be lost or gained if resolution is reduced.

It can also identify how many 'contacts' might be totally within the noise-floor of an instrument (physically the routine is one-sided, but the dilation has already effectively treated the noise as one-sided). Thresholding can also provide a first step towards more sophisticated models for approximating stiff but non-rigid contact. By a type of relaxation method, iterative modifications to the threshold might illustrate how the stylus settles elastically onto the surface. Each change of threshold would now be governed, e.g., by a force-balance using localized Hertzian contacts.

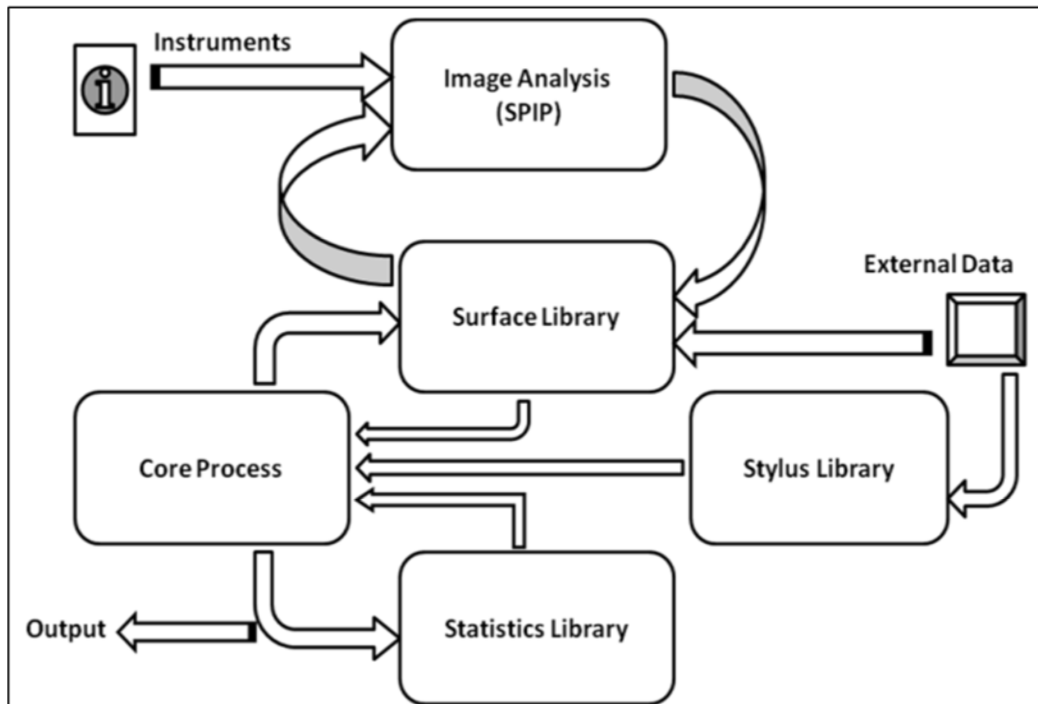


Figure 5.1: Scheme for the software suite.

The basic simulation suite (figure 5.1) is now established for use in extensive numerical experiments. Here, some early results are provided as a first demonstration of its capacity and potential for further analyses. In attempting to make the concepts more immediately obvious, this study shows only a physically

unrepresentative simulated stylus tip shape, involving an ideal truncation (flat), and a relatively large threshold.

5.2. Verification

The challenge of adequately verifying proper operation is greatly eased by the use of small modules that are used sequentially.

Running many trials of basic stylus shapes over simple surfaces, including single or clustered delta functions, rapidly identifies major bugs and increases confidence that no subtle errors remain. Just one example is given here, Figure 5.2. The pattern of a flat scanning across a noise-free sinusoidal prism is readily recognizable.

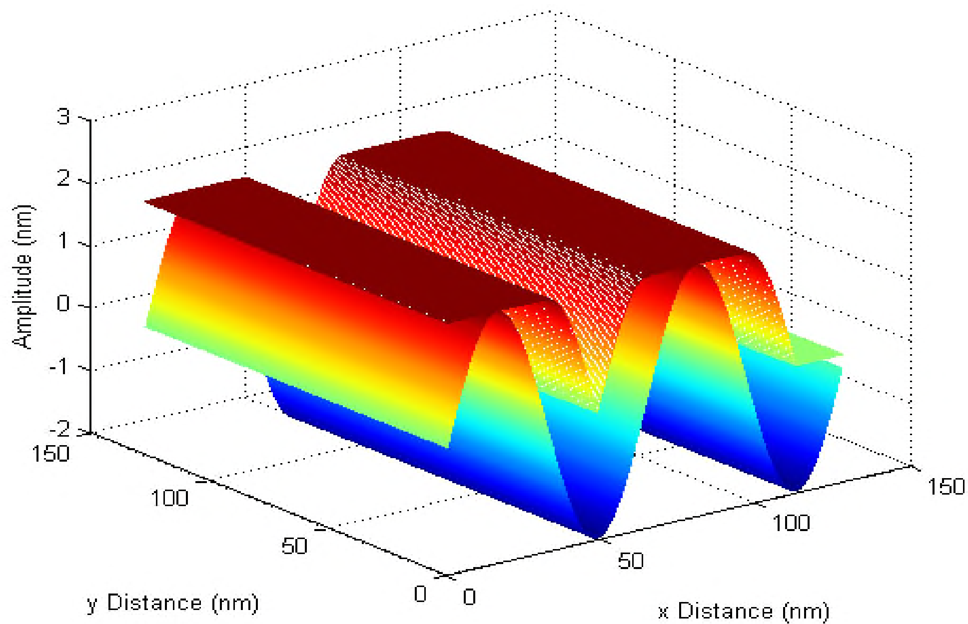


Figure 5.2 Scanning across a noise-free sinusoidal surface

It is also intuitively obvious that the contact pattern on the stylus has a low, uniform count (matching the flat regions on the modified surface) but there are higher values at the leading and trailing corner points (flank contact) and occasional multi-point contacts.

5.3. Kinematic and Threshold Models

Figure 5.3a shows a small map (128 by 128 points on a 0.2 µm sample grid) taken by atomic force microscopy from a typical ground steel surface. The fine detail has both sharply varying and plateau-like regions. For clarity of illustration, Figure 3b shows the severe effect of applying a ~2 µm square flat stylus (11 by 11 points) under kinematic conditions. The serious distortion is no surprise, with high-points being converted to plateaus, but, even so, the general structure is preserved: e.g., peak-to-valley amplitude has reduced merely from 758 nm to 746 nm.

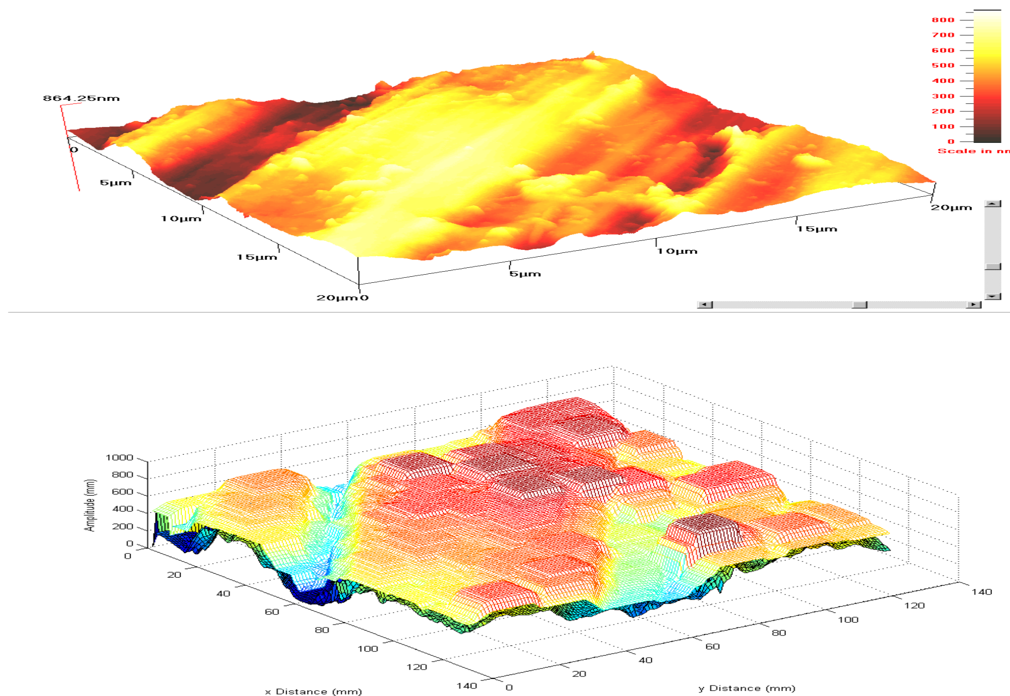


Figure 5.3: Modification of a surface detail by a 2 mm flat: (a) AFM image; (b) simulation

A typical profilometer set up to measure finish ground surfaces might have an effective noise floor of the order of 10 nm to 20 nm peak. Its digitized resolution might typically be 1 nm to 10 nm.

Alternatively, noise-related issues could be handled in terms of a fraction, say 0.5% to 0.1%, of full range (maybe a few tens of micrometres, typically). Hence, thresholds from 5 nm to 50 nm (for deliberate over-emphasis) seem relevant for study here. The simplest form of thresholding is roughly equivalent to ideally pressing the probe into the surface. For the illustrations used here the modified surfaces are largely unchanged other than a slight vertical offset. Interest lies in how the shift interacts with additional surface features to alter the probe contact patterns.

As an elementary illustration, Figure 5.4 shows the effects on the position of contacts on the probe when adding a 50 nm threshold to the case shown in Figure 5.3b.

For clarity, the maps show simply the presence or absence of contact at each point. Within the bounds of this flat probe, the actual area detected as 'in contact' is then highly analogous to the operation of the well-known bearing area. Note, though, that the pattern varies as the probe scans over the whole surface and more complex patterns will arise with more realistic probe shapes.

Centring the probe at (61, 66) – using ordinates, not dimensions, for simulations – puts it on a steep flank and kinematic contact is at one corner. Even 50 nm threshold causes only slight growth in the primary contact and the first signs of a second one nearby.

By contrast, (61, 66) and (109, 120) are in plateau regions, they show large contact growth, dominated by a single region and occasional secondaries.

In both cases, the kinematic contact is far from the centre of the region. Even at this small scale, growing the region develops a shape that correlates with the surface lay.

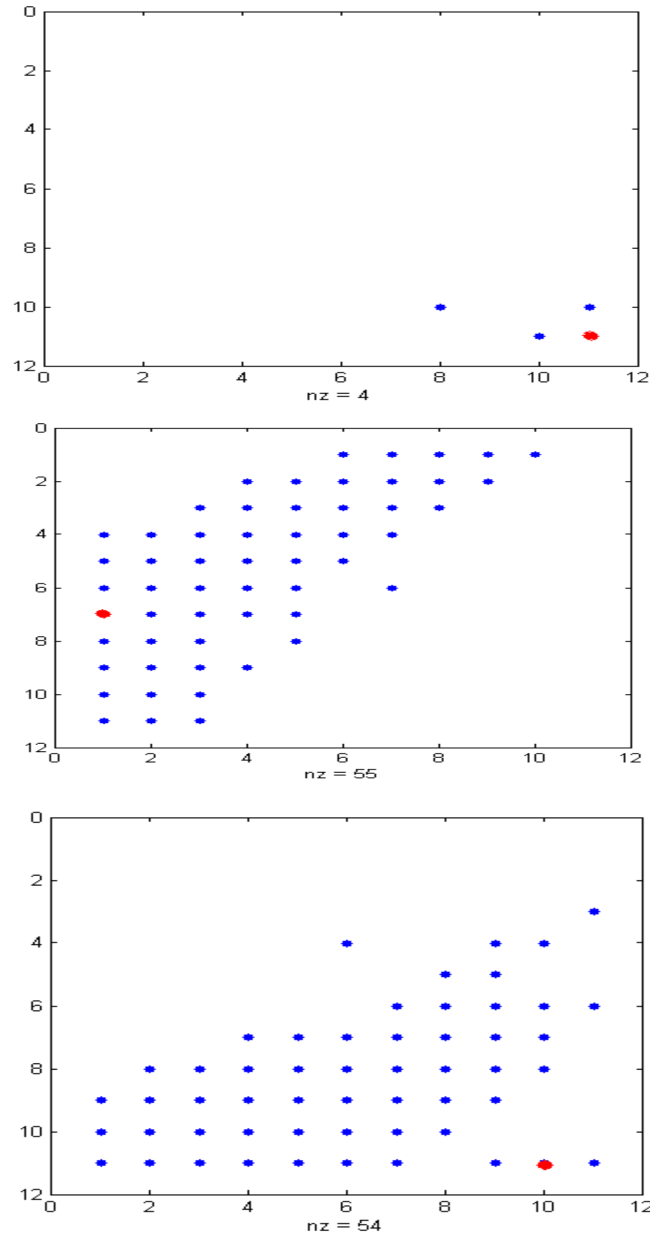


Figure 5.4: Positions of kinematic and 50 nm threshold contacts on a flat probe for three selected places from Figure 5.3b

In all three cases, the 'true' contact is significantly displaced laterally from the nominal line of action of the probe, a result also seen elsewhere (Dowidar and

Chetwynd, 2002) that has implications for measurement uncertainty of all topographic parameters except the simplest amplitude measures.

Although Figure 5.4 shows only the regions of contact, the intensity of contact at each point is of physical interest. One simple indication (not shown here) of this is to plot the probe maps in terms of how far each point on the probe would have encroached into the original surface to settle at the specified height relative to the kinematic condition. This gives, for example, a crude indication of the stress intensities associated with each contact, and so of its likely practical significance in defining the reported surface height of a real instrument.

As a further experiment, figure 5.5 shows an ideal conical stylus and a map indicating its contact against an ideal flat. The initial geometric contact is a single point.

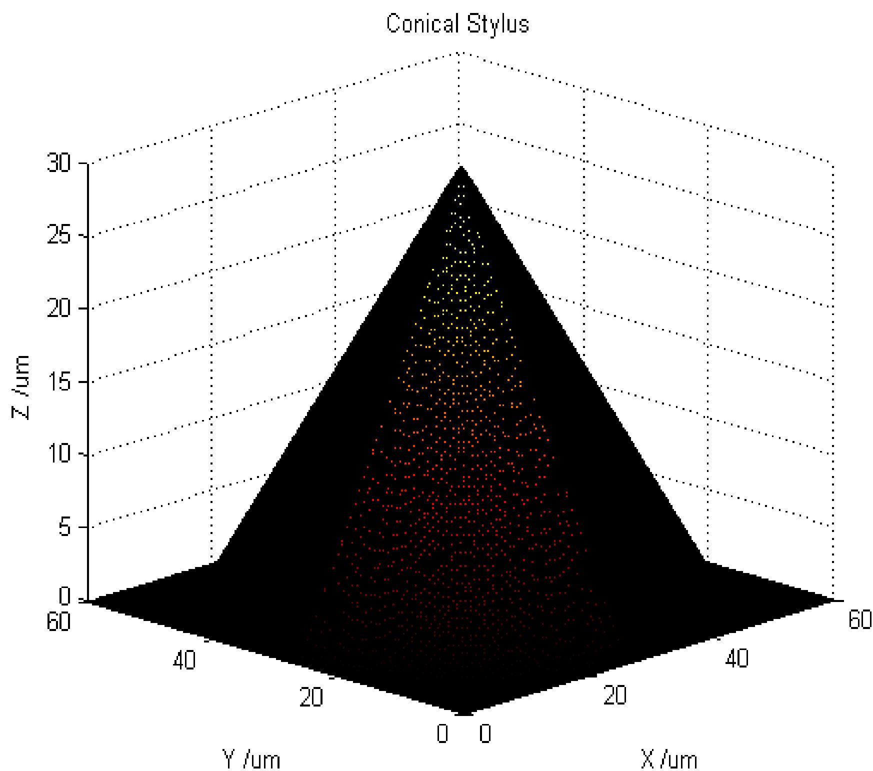


Figure 5.5: Conical shape 60x60µm

With a threshold, all surface points sufficiently close to the tip are considered to be in contact: a large value generates a symmetrical patch of contact. Here, we get effectively the same image as from an ideal indenter into a plastic material.

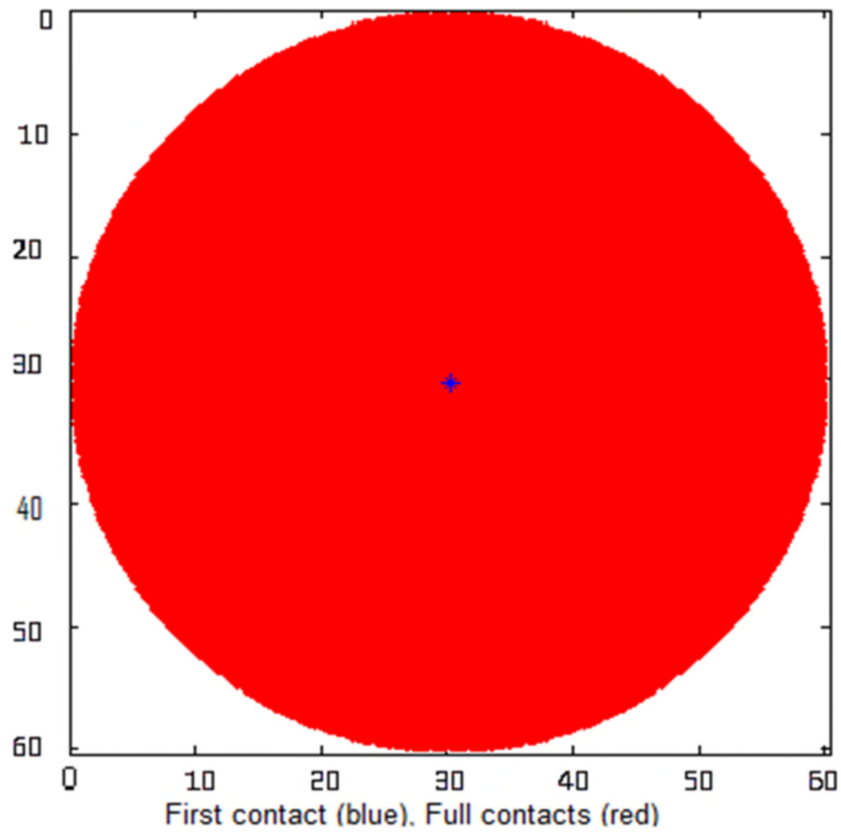


Figure 5.6: Map of contact points with 149 nm

Figure 5.7 shows just the tip of the 10X10 μm conical stylus and contact with a flat using a much smaller threshold. The basic behavior is the same, but resolution limits within the geometry become apparent.

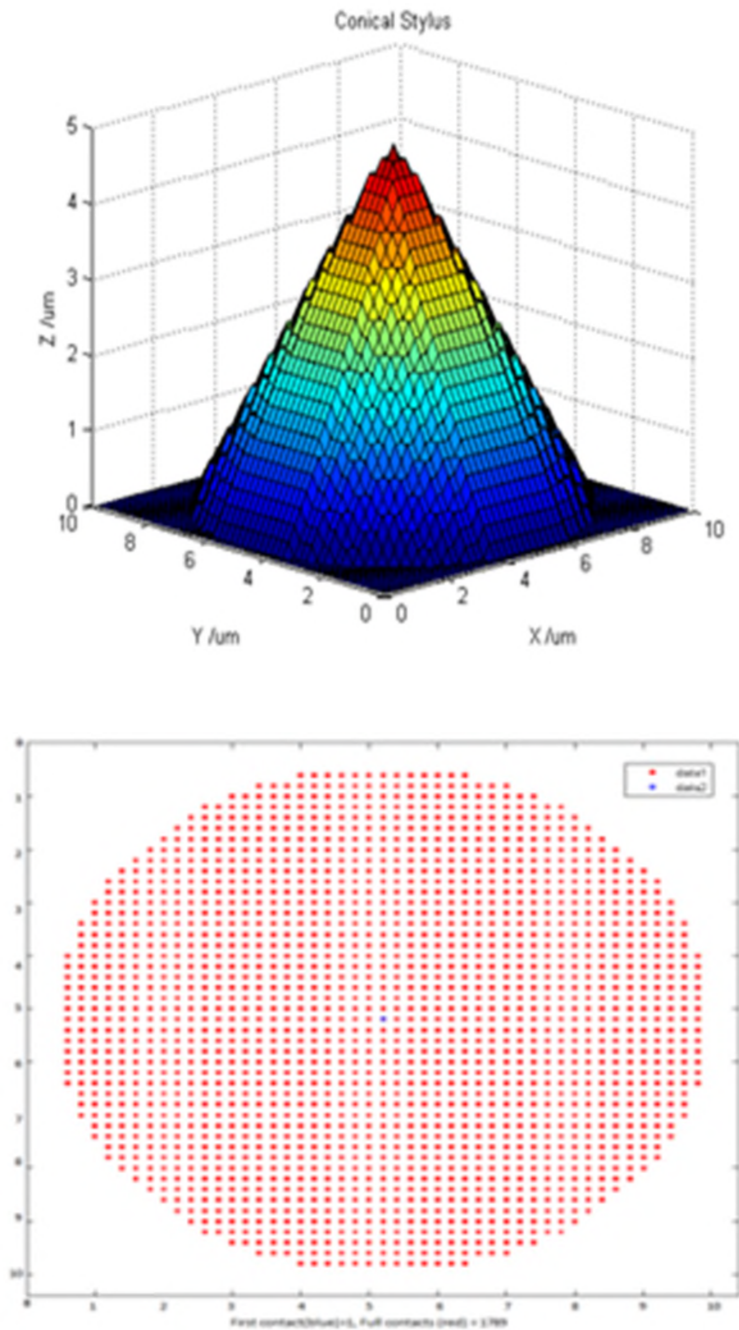


Figure 5.7: (a) Conical shape Stylus 10x10 μm (b) Map of contact points with 23nm threshold

Figure 5.8 shows a 20x20 µm small map (64 by 64 points on a 0.2 µm sample grid) taken by atomic force microscopy (AFM) from a typical ground steel surface. The fine detail has both sharply varying and plateau-like regions.

Figure 5.9 shows the stylus in (figure 5.7a) contacting local features of ground surface figure 5.8. Again using heavy threshold for clarity, key observations (fig 5.9b) are that the initial contact is not central and the patch grows asymmetrically, eventually picking out the major structures of the surface.

The total local height variation over this 10 µm by 10 µm is about 1000 nm, so thresholding to a few percent of height suggests real working simulations should be perhaps 10-20 nm.

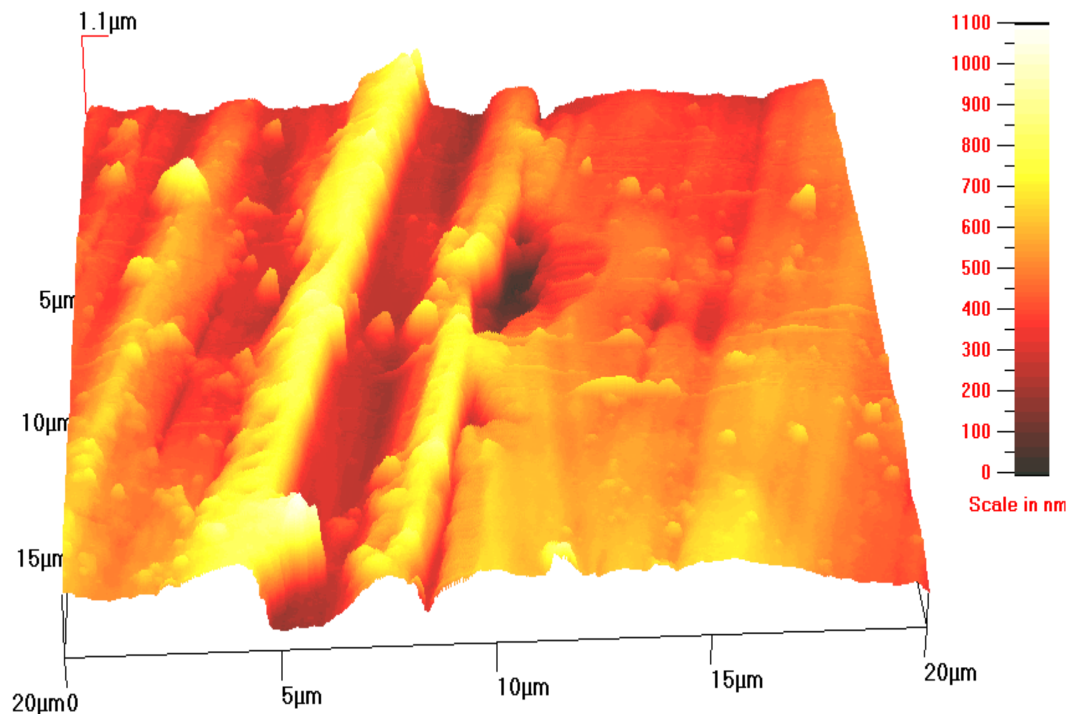


Figure 5.8: (a) 3D 20x20 µm small map (AFM) image

To measure such a surface we might expect to use an instrument range of perhaps 5000 nm, with inherent noise of up to 1% full scale. Thus thresholding to a few tens of nanometers is also appropriate for modeling noise sensitivity.

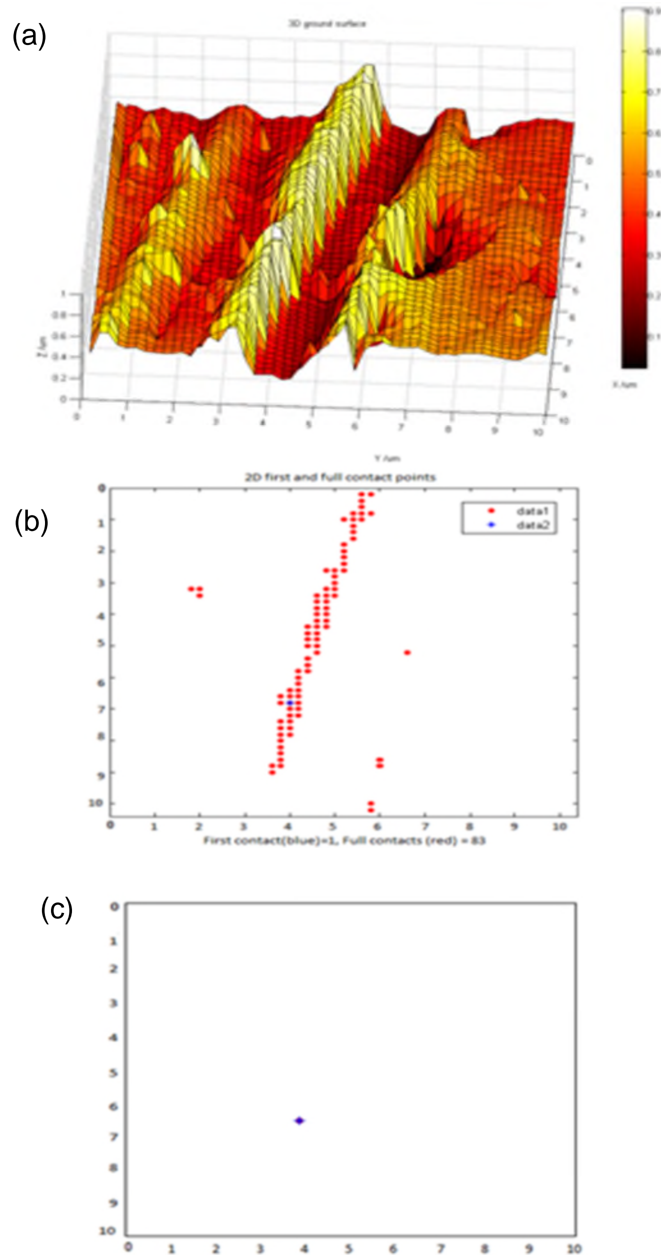


Figure 5.9: 3D ground surface 10x10 µm (b) Map of contact with 149 nm threshold (c) with 40 nm threshold

For the particular condition illustrated, the map remains essentially the single point (in blue at Figure 5.9b) for thresholds as high as 40 nm; there is a single, real peak

that is likely to dominate local contact behavior at the scale of this tip. Figure 5.9c shows the threshold up to 40 nm leaves single point, which is the first contact.

Comparing figure 5.9, and figure 5.10 (after moving the surface across and putting the tip in the middle), by applying heavy threshold, the pattern is very heavy threshold, and the pattern is very similar. It can be concluded that here is only dominant peak and it is not noise sensitive. In figure 5.9, the lateral error is about 1.5 um. It also showed that here (tip radius fairly large) little change in reported contact under local movements. (i.e. not much different to flat stylus in actual behavior).

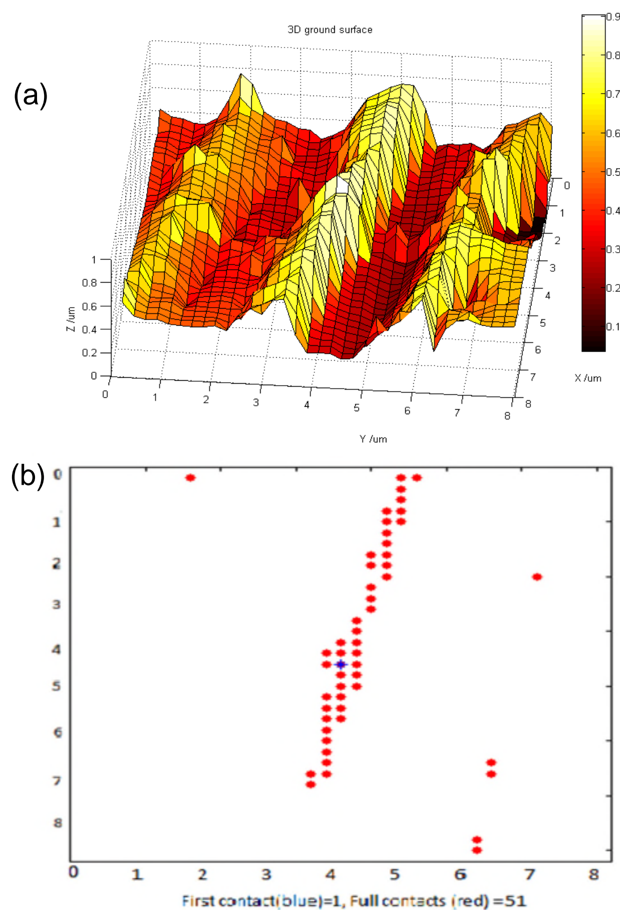


Figure 5.10: 3D Ground surface 8X8μm with tip at the centre. (b) Map of contact with149 nm threshold.

Chapter 6

6. Stylus tip size and shape effects in surface contact profilometry

6.1. Stylus Simulation

This chapter presents refinements to 3D simulations of stylus effects in microtopography measurements. It briefly reviews how statistically richer data can be obtained by extending basic kinematic models, perhaps providing steps towards more sophisticated modelling of the contact process. After a review of the new simulation scheme, results and discussion concentrate on idealized styli operating beyond the limit of their expected resolving power.

It is almost impossible to determine directly the complex interactions that occur between a profilometer stylus and a surface during a measurement of surface micro-topography. There is, however, a long history (dating back to the obsolete E-system of reference lines) of simulating the loci of circles, and latterly spheres, rolling on rough topography both as an evaluation tool and in attempts at 'stylus deconvolution'. Comparison between real results and such simulations can reveal indirect evidence concerning behaviour in the contact region, although current models are hardly sophisticated enough to make much impact. They are purely kinematic models, assuming perfect guides and rigid objects and are actually a class of morphological filter (specifically, the centre locus is dilation operation) (Muralikrishnan , Raja, 2009). Published work concentrates heavily on the effect of the 'measured surface', extended in one case to report some statistics on the contact to the stylus (Dowidar and Chetwynd, 2002). The present research takes this further, motivated by its potential to cast light on the in-plane uncertainty of

stylus measurements, suggest ways to improve comparisons between instruments or lead eventually to methods for self-diagnosis of stylus wear.

A new simulation gathers full data about stylus contact related to its location upon the surface. It introduces a threshold process (Althagafy and Chetwynd, 2011) by which the kinematic condition is violated in small increments and the growth of resulting 'contact areas' recorded. This is intended to give insight into sensitivity to instrument noise, repeatability, etc. It also has potential for modelling the contact process, by approximating stiff but non-rigid contact using relaxation techniques. The contact modelling is implemented in MATLAB® which is interfaced with the commercial topographic analysis software SPIP in order to provide a standard for parameter evaluation and comparison, and to translate between different instrument data formats and MATLAB arrays (Althagafy and Chetwynd, 2012). The surface could be any set of data representing a real or an arbitrary surface. Also, the stylus could be any set of data representing a real or an arbitrary stylus shape. Both data sets are dealt with as arrays.

6.2. Styli and fine surface structure

This chapter concentrates on study of the sensitivity to stylus shape and condition when detecting features of real surfaces at the very limits of conventional profilometer capabilities. It therefore uses relatively small arrays with a grid sampling interval of 0.1 μm. Three ideal computer generated stylus tips with different shapes have been used: conical, pyramid and spherical (Figures 6.1, 6.2 and 6.3). The tip radius and heights (for the conical and pyramid shape) are 3 μm and 5 μm. The tip angles of the conical and pyramid shapes are 90^0. Each tip has

been used in its perfect shape and with a quite severe truncation at 2μm below the original tip. The spherical tip offers a full hemisphere, not the more usual blend into a cone, allowing estimation of flank contact for different cone angles.

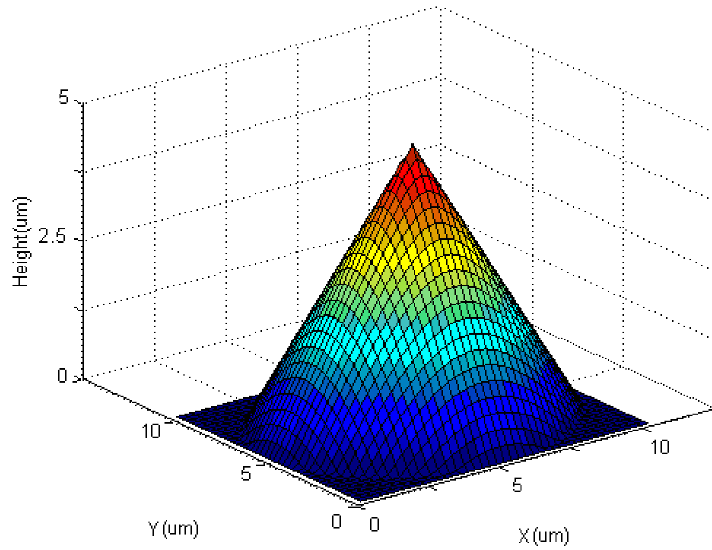


Figure 6.1: (5μm) Conical Stylus tip shape

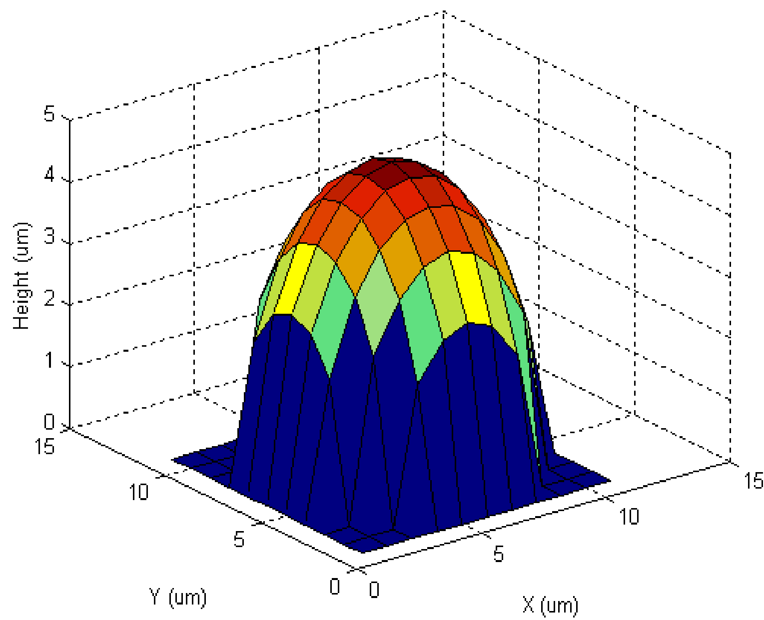


Figure 6.2: (5μm) Spherical Stylus tip shape

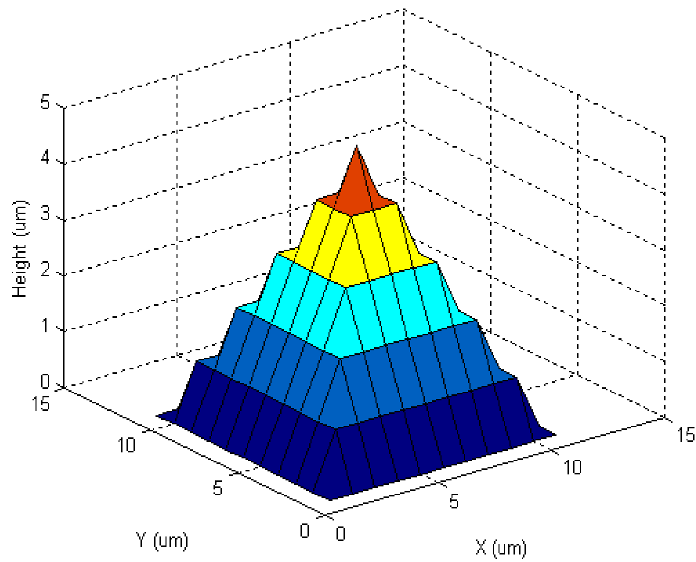


Figure 6.3: (5µm) Perfect Pyramid Stylus tip shape

6.3. Results

Many trials were run of the basic stylus shapes over simple computer generated surfaces, such as single or clustered delta functions. These can rapidly identify major bugs and increase confidence that no subtle errors remain in the simulation routines.

6.3.1. Verification

The challenge of adequately verifying proper operation is greatly eased by implementing the model and simulation as a set of small modules that are used sequentially. Running many trials of basic stylus shapes over simple surfaces, including single or clustered delta functions, rapidly identifies major bugs and increases confidence that no subtle errors remain.

The output of scanning the random surface with a 10 µm perfect conical tip is shown in figure 6.4. There is also a large visible difference between the original surface and the locus of the stylus. There is a visible difference from that generated by the spherical tip.

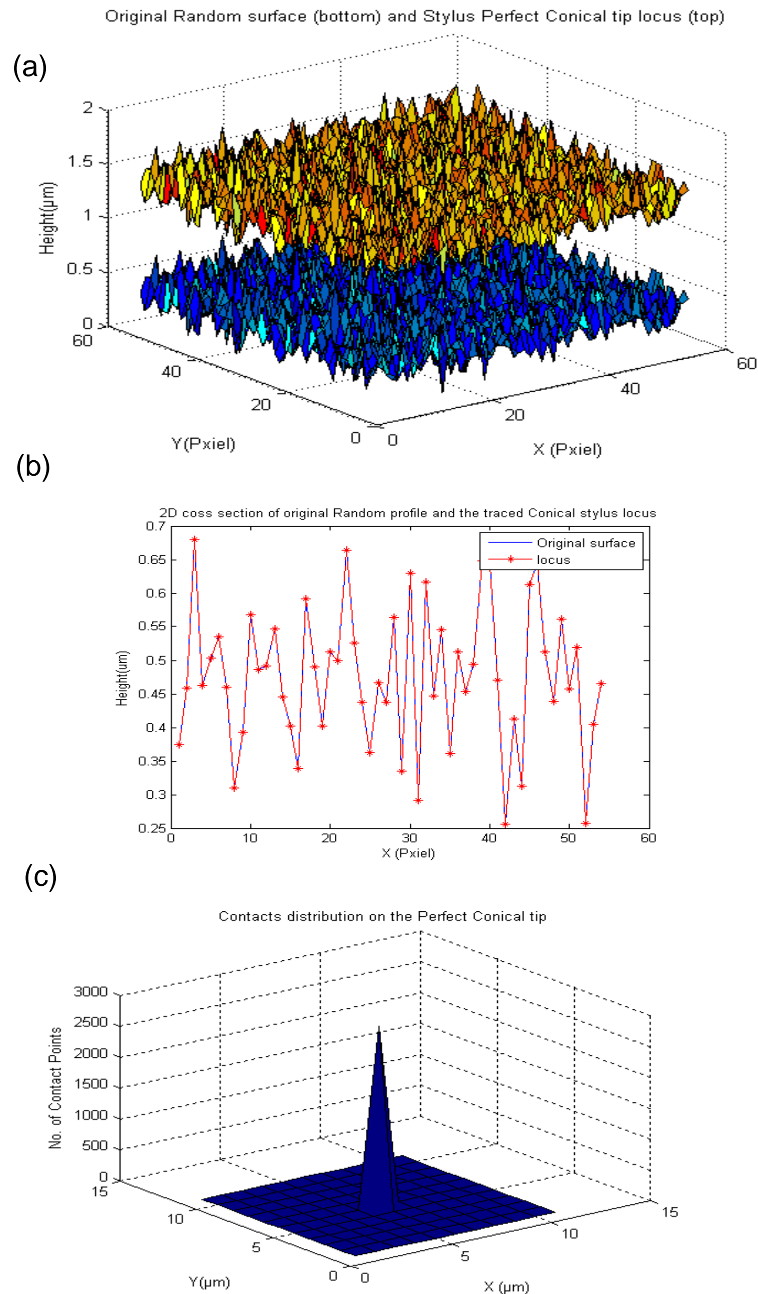


Figure 6.4(a) Locus & original surface (b) 2D cross section profile (c) Contacts distribution: of a 10µm conical tip on a random surface

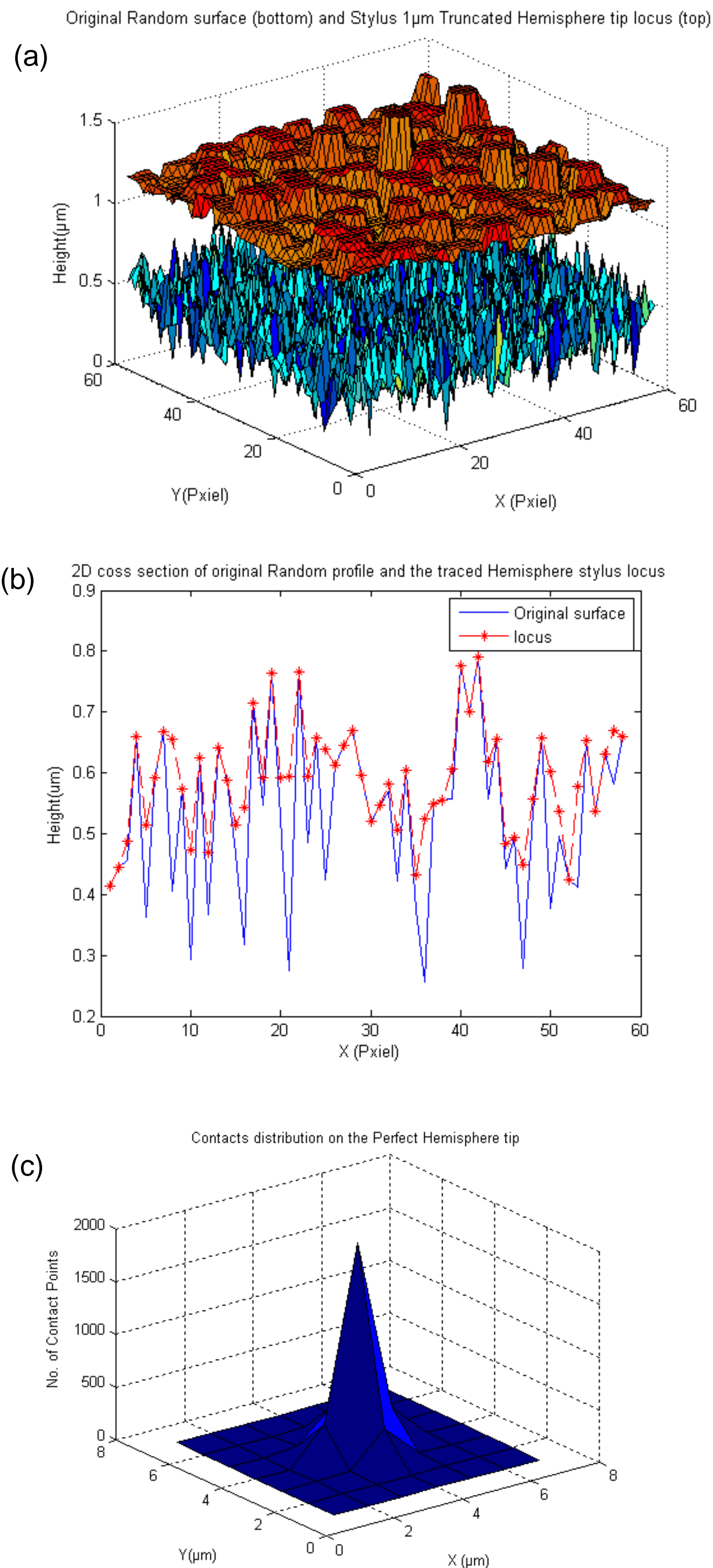


Figure 6.5:(a) Locus and original surface (b) 2D cross section profile (c) Contacts distribution: of a 10µm conical tip on a random surface

6.3.2. Simulation on real surfaces

The simulation verification has shown that the technique is working properly and giving significant results. The next stage is to apply the technique to a set of data representing real surfaces. The surface roughness of the real samples have been measured by Atomic Force Microscopy (AFM) and the data representing these surfaces have been converted to readable format in Matlab by using the Scanning Probe Image Processor (SPIP). The surface data have been scanned theoretically in 3D by the computer using the same tip shapes (pyramid, sphere and cone) with the two radius size 3μm and 5μm. The sample is ground steel. The data collected by AFM were checked for no missing data losses interpolation by SPIP software. The value of the roughness parameters of the samples are shown in table 6.1.

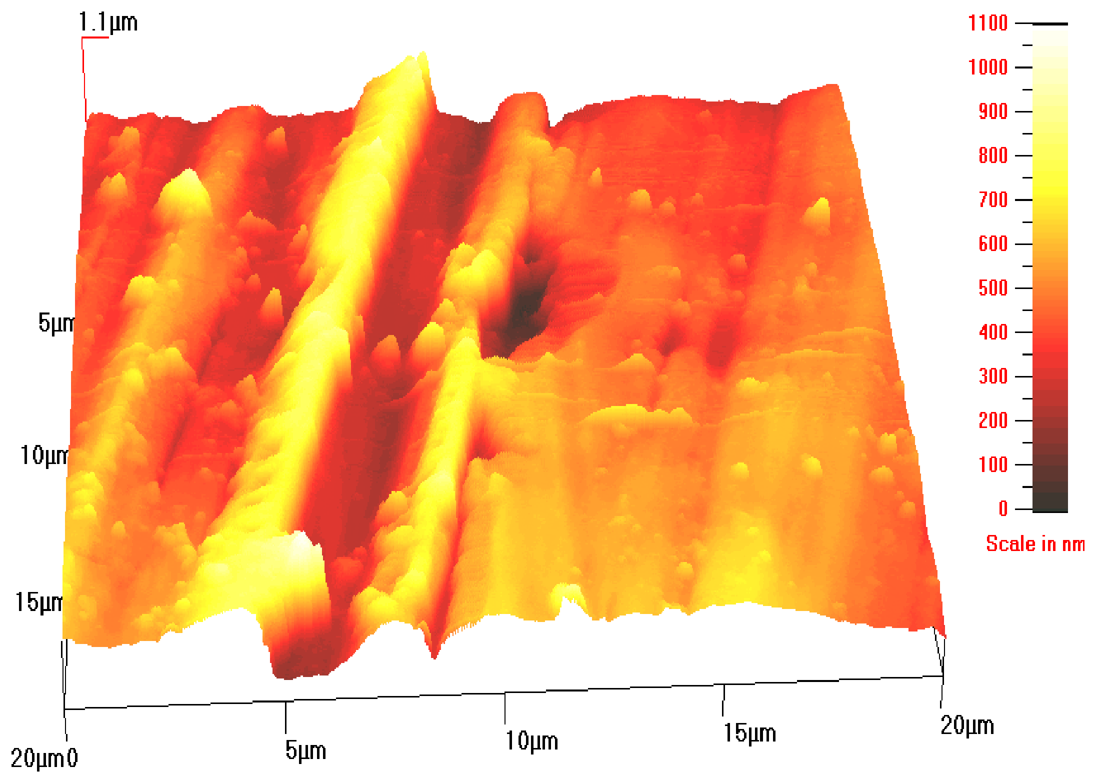


Figure 6.6: AFM 3D image of a ground steel surface

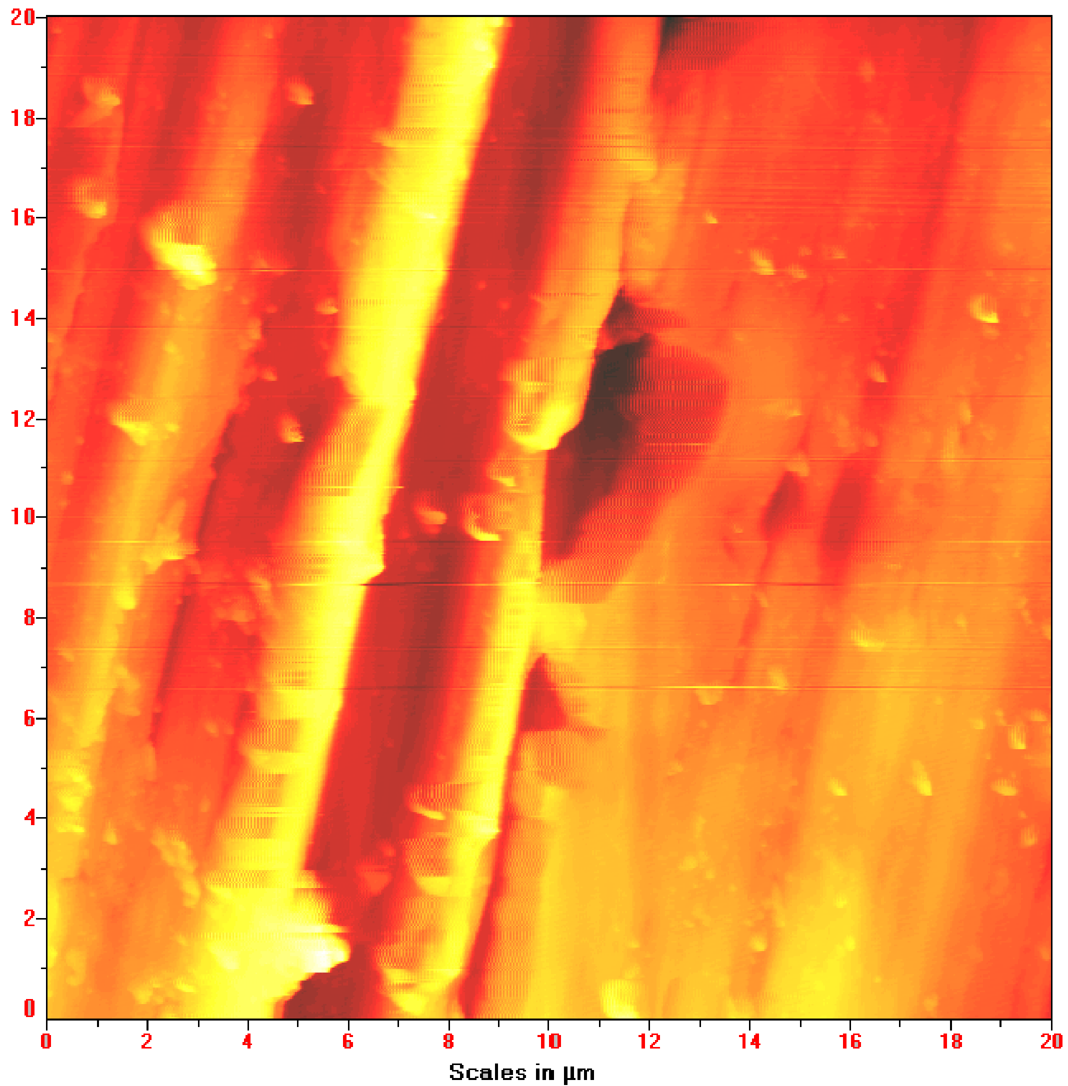


Figure 6.7: AFM 2D profile of a ground steel surface

Table 6.1: Real surface roughness parameter

Roughness Parameter	R_a	R_q	R_y	R_{sk}	R_{ku}
nm	458	481.5	920	1.135	1.37

6.3.3. Results discussion

The output of scanning the ground surface with the 5µm radius perfect pyramid tip is shown in figure 6.8. There is a little visible difference between the original

surface and the locus of the stylus. But, the roughness parameters Ra and Rq have been reduced from the original values of the surface. The contacts distribution on the tip shows all contacts. Moreover, the maximum number of contacts occurs on the central point of the tip.

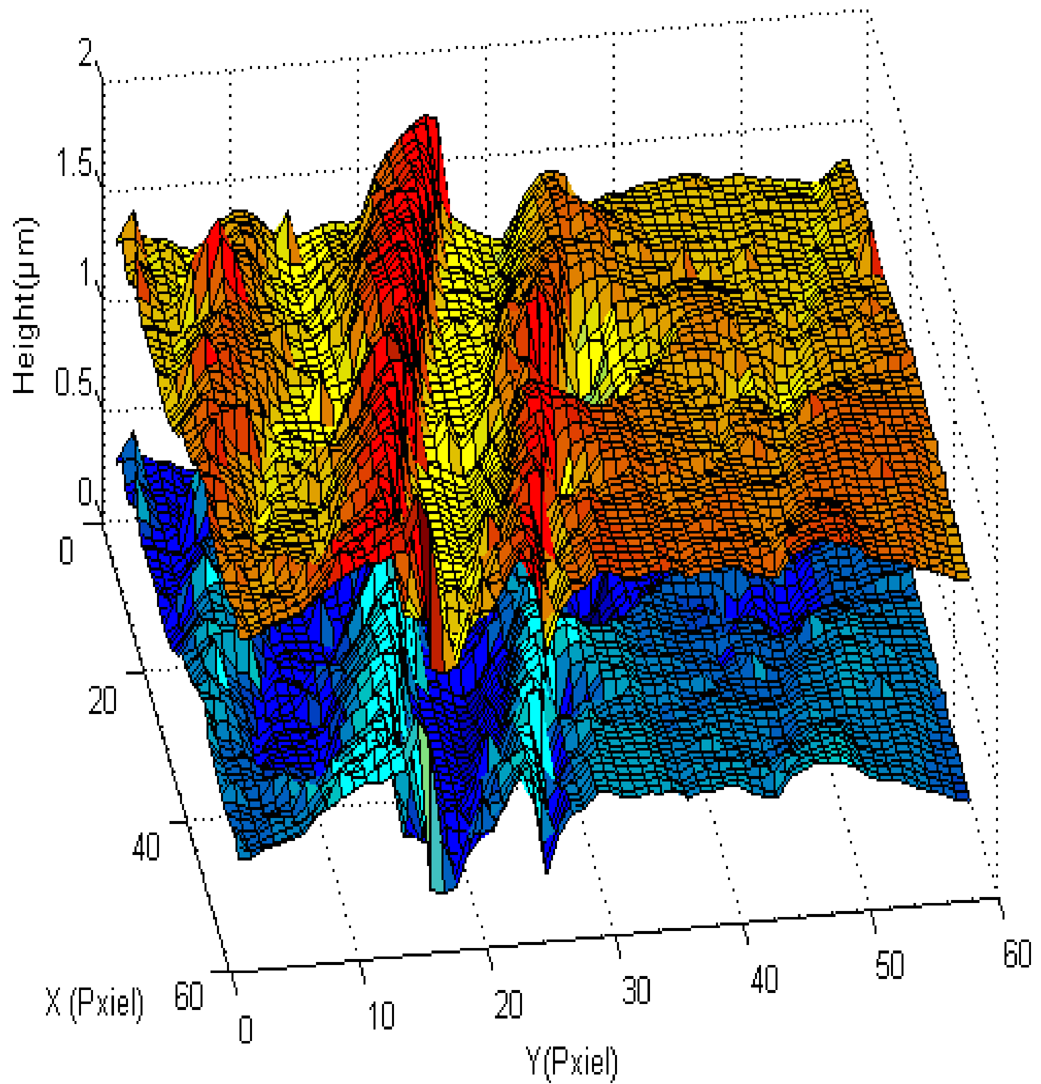


Figure 6.8: Original ground surface (Bottom) and perfect pyramid stylus tip locus (top)

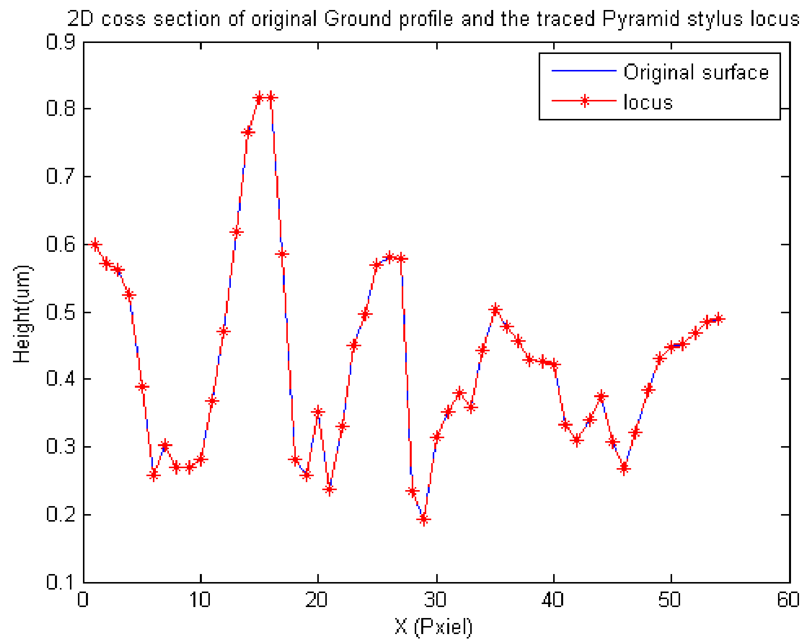


Figure 6.9: 2D cross section of original ground surface profile and the traced pyramid stylus locus

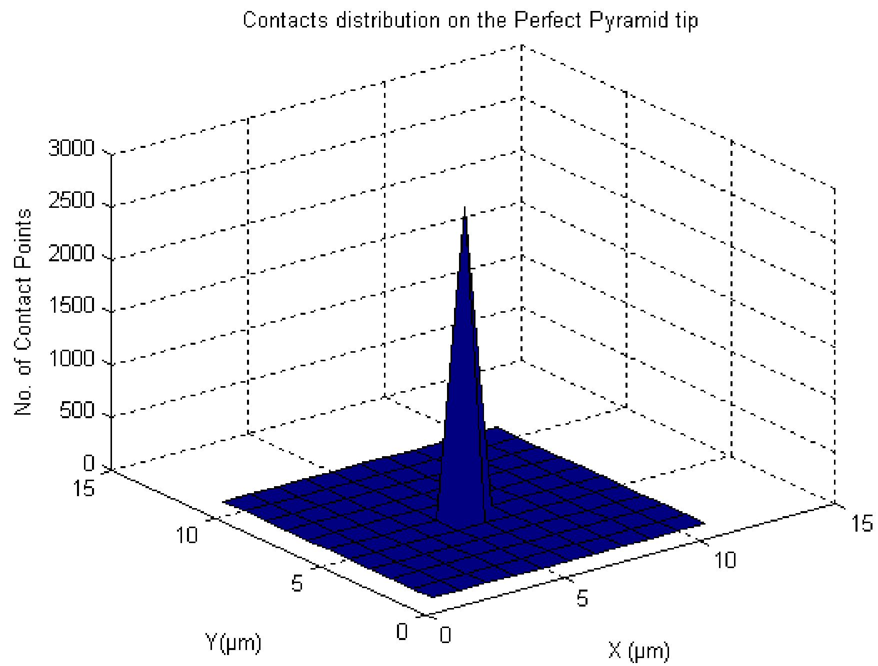


Figure 6.10: Contact distribution on the perfect pyramid tip

The output of scanning the ground surface with the 10 μm radius perfect spherical tip is shown in figure 6.11. There is visible difference between the original surface and the locus of the stylus. Moreover, the roughness parameters Ra and Rq have been reduced from the original values of the surface.

The contacts distribution on the tip shows that all contacts occur on the central line of the tip. Also the maximum number of contacts occurs on the central point of the tip.

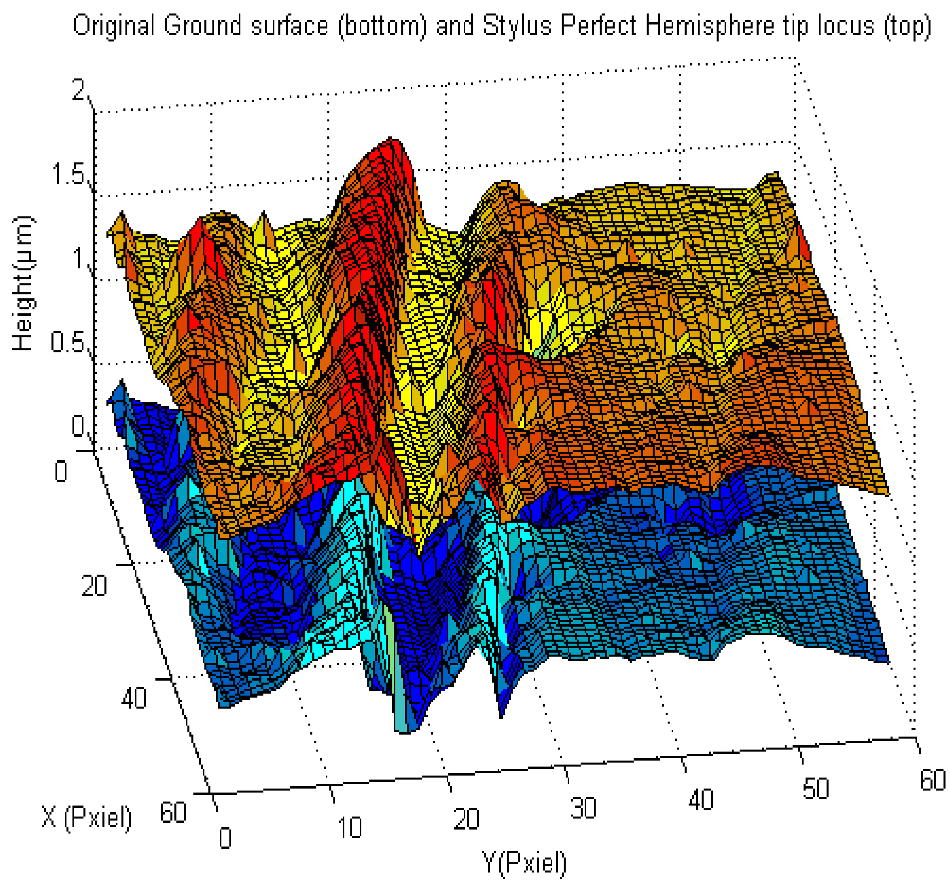


Figure 6.11: Original ground surface (Bottom) and perfect Hemisphere stylus tip locus (top)

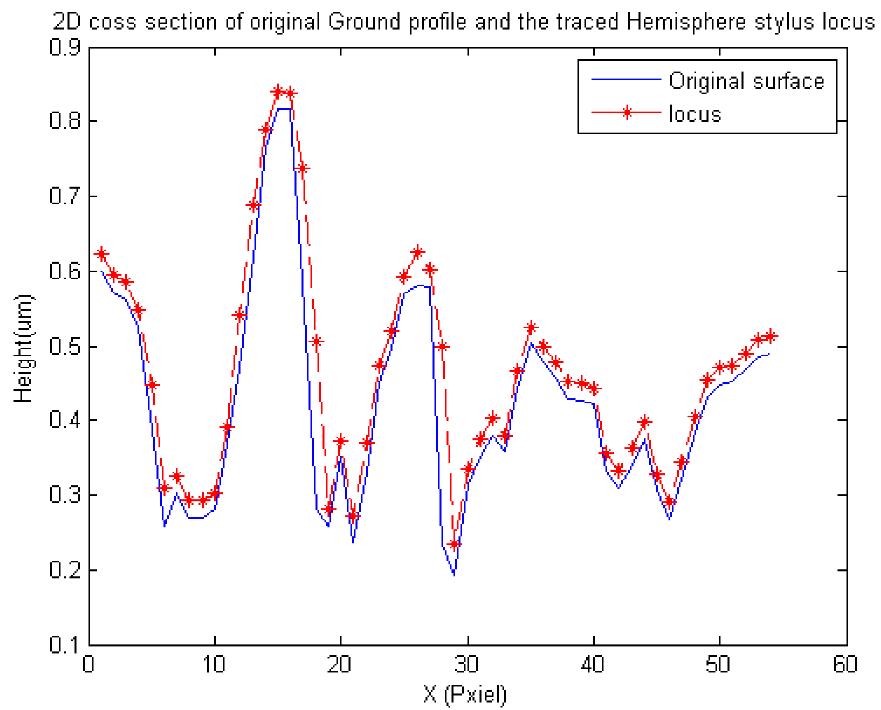


Figure 6.12: 2D cross section of original ground surface profile and the traced Hemisphere stylus locus

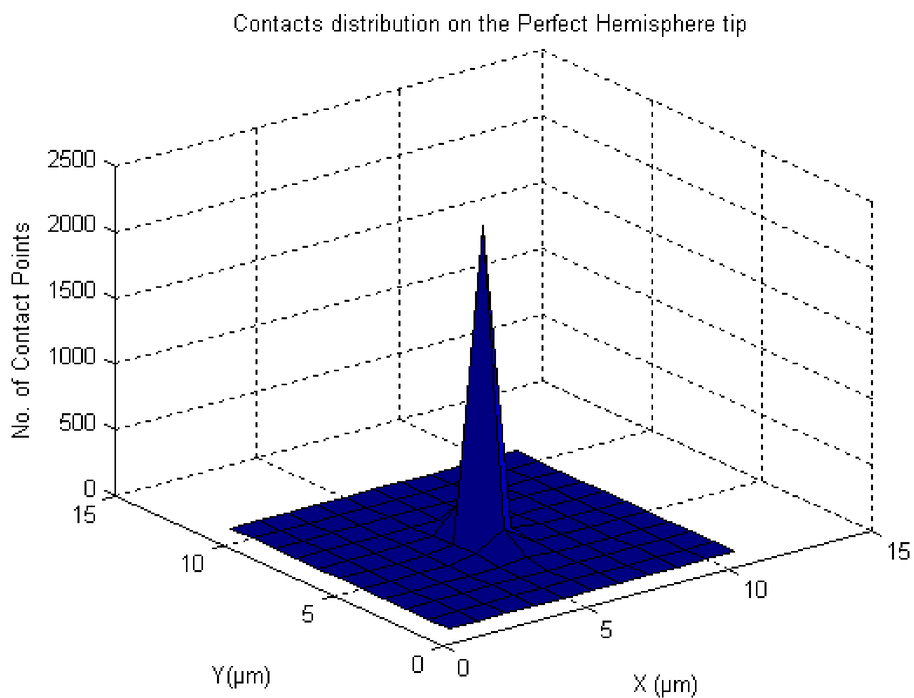


Figure 6.13: Contact distribution on the perfect hemisphere tip

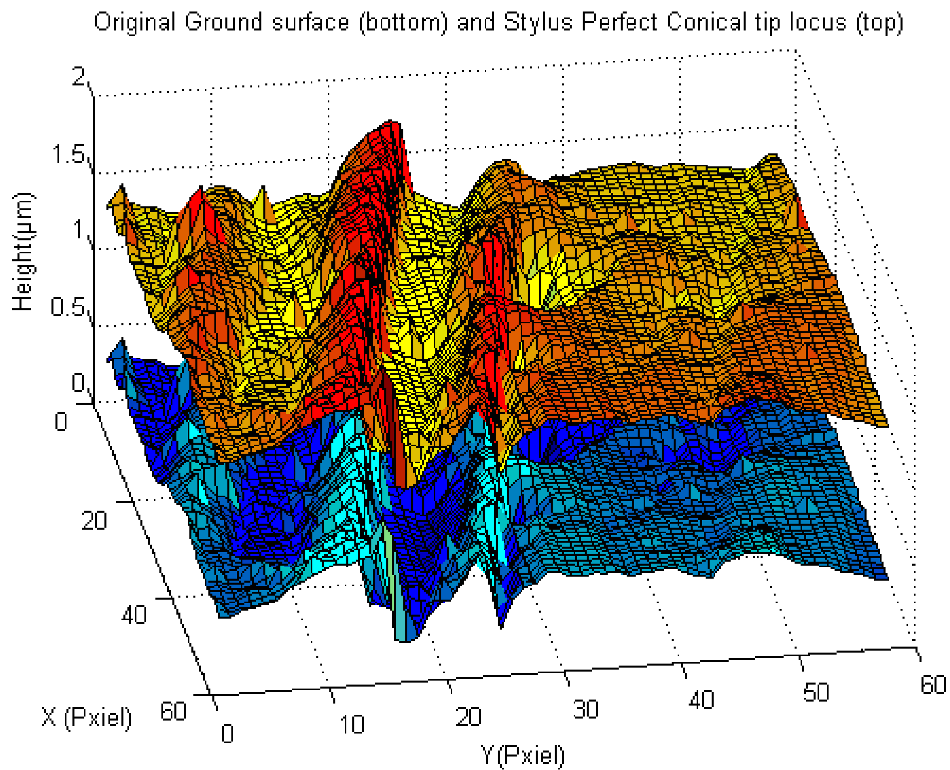


Figure 6.14: Original ground surface (Bottom) and perfect conical stylus tip locus (top)

The output of scanning the ground surface with a 10 µm perfect conical tip is shown in figure 6.14. There is also little visible difference between the original surface and the locus of the stylus.

The roughness parameters Ra and Rq are nearly the same as the original values of the surface. The contacts distribution on the tip shows that all contacts occur on the central point of the tip.

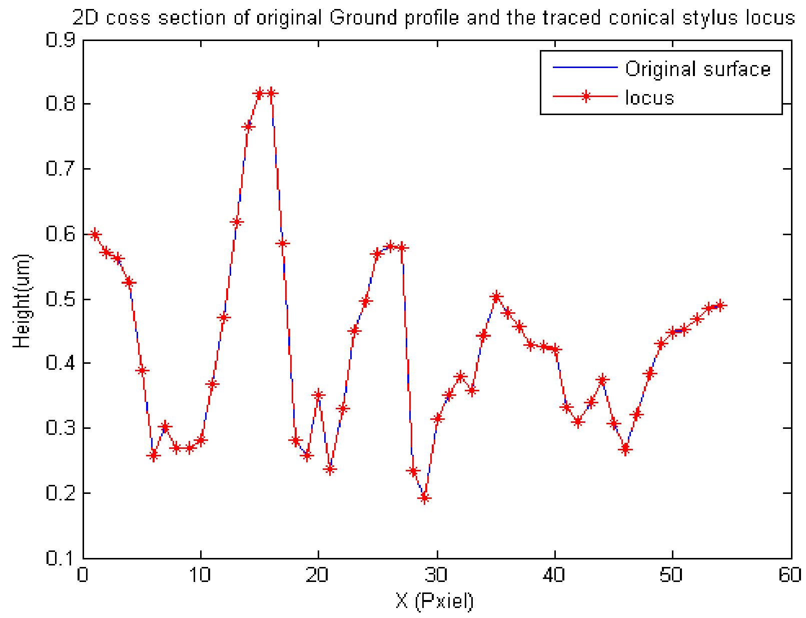


Figure 6.15: 2D cross section of original ground surface profile and the traced conical stylus locus

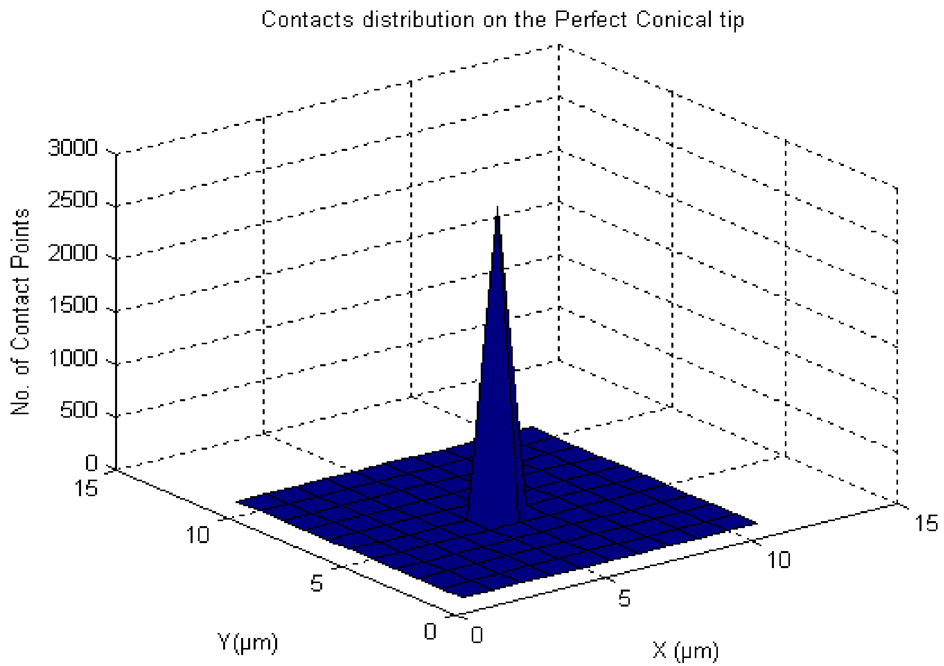


Figure 6.16: Contact distribution on the perfect conical tip

The contacts distributions of the 1 and 2 μm truncated tips when scanning the ground surface are shown in figure 6.18. It is clear that the ground surface gives very similar shapes of the contacts distribution on the spherical tips to the ones of the Sinusoidal surface as most contacts occur along the central line of the spherical tip but with more contacts spread around the edges of the tip end.

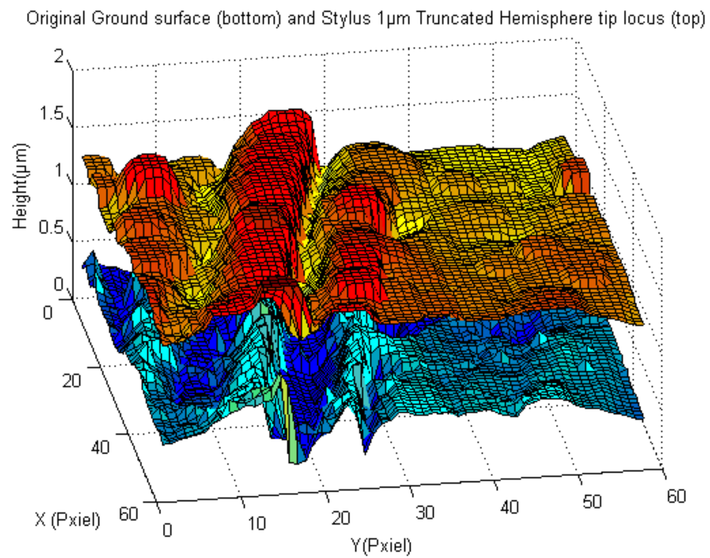


Figure 6.17: Original ground surface (Bottom) and truncated Hemisphere stylus tip locus (top)

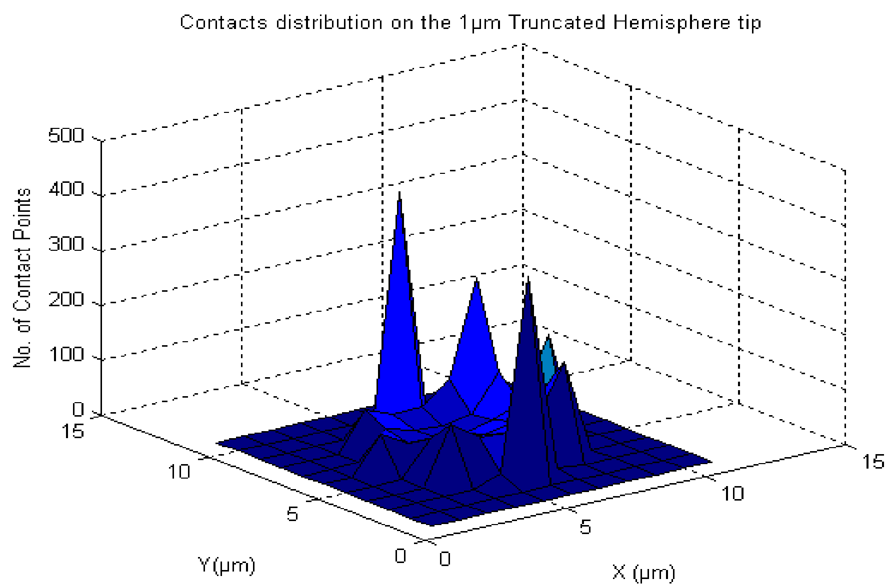


Figure 6.18: Contact distribution on the truncated Hemisphere tip

In the conical tip are also shown slightly different distributions as most contacts occur around the corners of the tip. This is due to the existing irregularities on the ground surface which don't exist on the perfect sinusoidal shape. It is also noticed that the centre of all tips have the minimum number of contacts with the surface.

Additionally, it is visible that scanning the ground surface with different dimensions 3µm has shown nearly the same locus and contact distribution shape as the 5µm tips.

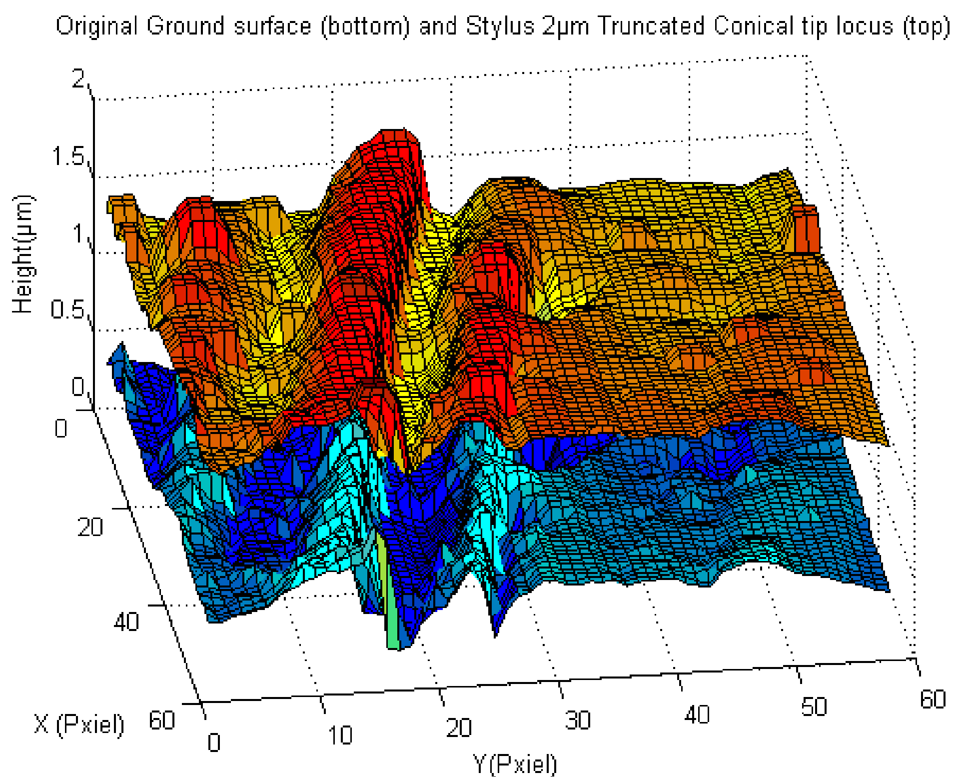


Figure 6.19: Original ground surface (Bottom) and truncated conical stylus tip locus (top)

The truncated pyramid tip shows a different result from the truncated conical tip which shows the growth of contacts at the central point of the tip.

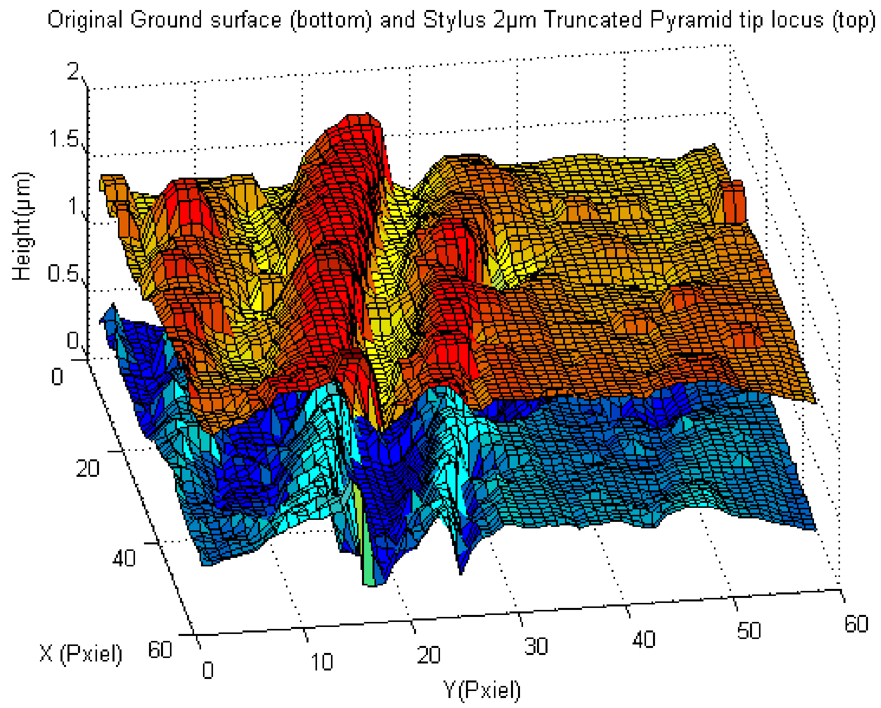


Figure 6.20: Original ground surface (Bottom) and truncated pyramid stylus tip locus (top)

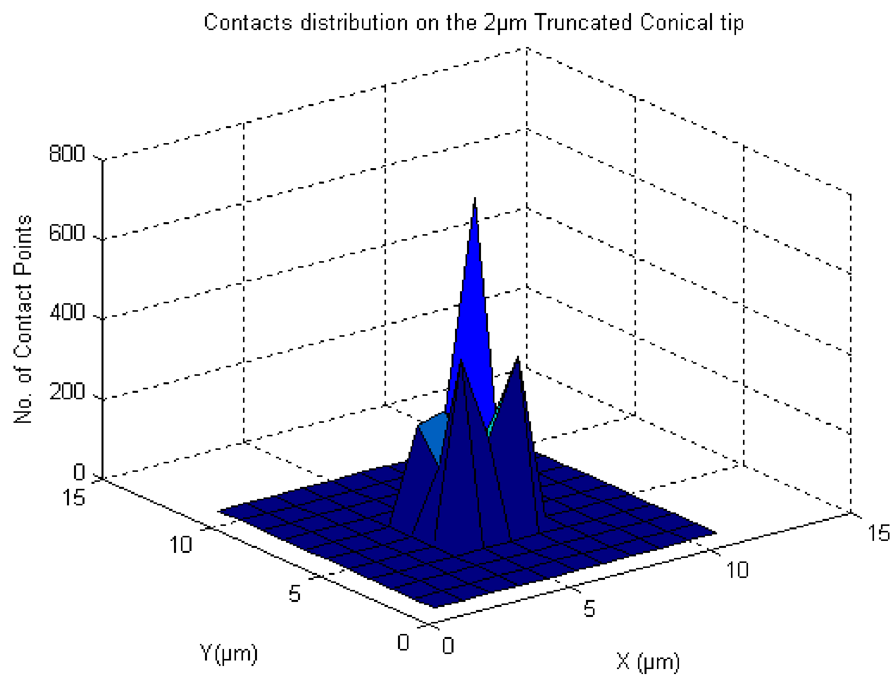


Figure 6.21: Contact distribution on the truncated conical tip

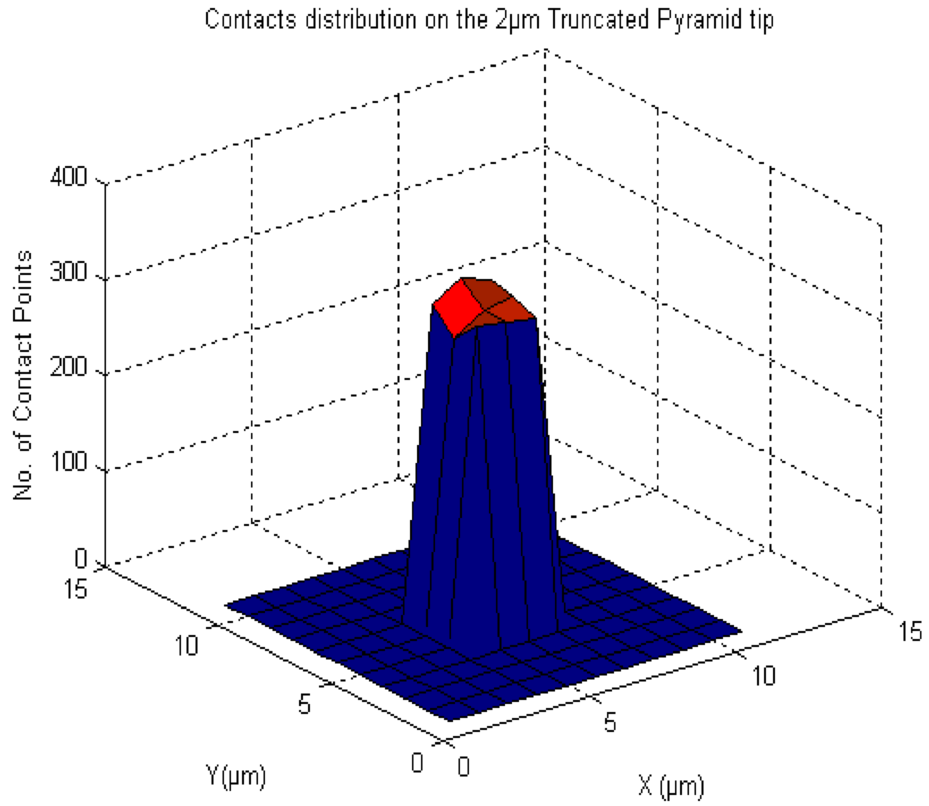


Figure 6.22: Contact distribution on the truncated pyramid tip

Table 6.2 shows the percentage error of the roughness parameters of different outputs when scanning the ground shape with different tips. The maximum deviation occurs when using the truncated 5 µm spherical tip.

Table 6.2: Roughness parameters Errors

		Roughness Parameter	Sa	Sq	Sy	Ssk	Sku
			458 nm	**481.5 nm**	**920 nm**	**1.135**	**1.37**
Stylus Tip	**Size**	**Shape**	%Error=100x(Measured Value - Actual Value) / Actual Value				
Pyramid	3µm	Perfect	0.00	0.00	0.00	0.00	0.00
		2µ Worn	-2.1	-1.5	-7	-1.8	-3.7
	5µm	Perfect	0.00	0.00	-0.02	0.00	0.00
		2µ Worn	-3.62	-3.44	-6.88	-0.877	-1.43
Sphere	3µm	Perfect	-1.09	-1.4	-17.2	-0.88	-0.72
		2µ Worn	-34	-29	-11	-7.96	-16.7
	5µm	Perfect	-6.76	-6.76	-1.6	-18.4	-0.877
		2µ Worn	-65.34	-56.34	-29.9	-10.57	-20
Cone	3µm	Perfect	0.00	0.00	0.00	0.00	0.00
		2µ Worn	-2	-1.4	-7.39	-1.7	-3.64
	5µm	Perfect	0.00	0.00	0.00	0.00	0.00
		2µ Worn	-3.38	-3.34	-7.45	-0.877	-1.4

The 5 µm spherical tip with 2 pm truncation has the worst effect on the roughness parameters as it has highly reduced Ra and Rq from their actual values.

The initial verification of the computer simulation in this study was based on computer-generated data, which are not real in the surface roughness measurements. This is to demonstrate the idea and show the extreme cases of the effect of the stylus shape on the roughness of the measured surface. Tips have been used in their perfect shapes; spherical, pyramid and conical. Even the worn tips have been truncated uniformly producing a perfect worn shape where the top and bottom of the tip are totally flat and parallel. In addition, the tips have been used in

their ideal orientations where the central line of the tip is vertical and the base of the tip is horizontal. All other sources of error including mechanical and electrical error have been neglected.

The simulation on the computer-generated surfaces has shown that the stylus geometry has an effect on all measured roughness parameters. Even with the perfect conical and pyramid tips where the contact with the scanned surface and the tip is always at one point; there is an effect. This effect may not be too great on some roughness parameters but it nevertheless exists. On the other side, demonstrating the contacts distribution on the tip has shown a very good prediction of what could be happening on that small area of the stylus contacting the surface. Scanning real surfaces by computer-generated tips has pushed the simulation one step toward the real metrology.

The simulation program has scanned surfaces produced by different machining processes (e.g. lapping and grinding). The actual values of the roughness parameters of these surfaces have been previously determined by different software (Scanning Probe Image Processor, SPIP™). The simulation has shown that the stylus shape and size have affected all the surface roughness parameters of all specimens. Generally, from the simulation on the real and non-real surfaces, it has been noticed that the measured values of the roughness parameters Sa, Sq and Sy are always less than the actual values of the surface. It has been also noticed that the conical and pyramid tip shapes give closely similar results of the roughness parameters and the contacts distribution. This is simply because the two types of tips are very similar in shape and size at their ends.

Additionally, the contacts distribution on the tip does not only depend on the tip shape but also on the topography of the measured surface. When the sinusoidal surface is scanned by the spherical tip, if the wavelength and the amplitude of the surface are slightly bigger than the tip, the tip will not be able to fully penetrate the valleys and touch the surface and most contacts will occur on the sides of the tip. The lowest point of the tip will only touch the surface at the highest points of the peaks where the tip has the maximum contacts on its sides and the minimum contacts on its central point. If the wavelength of the surface is much bigger than the stylus tip and the amplitude is smaller, the contacts on the central point of the tip will relatively increase and the contact area of the tip will decrease.

If the stylus tip is bigger than the wavelength of the surface, the contacts will mainly be concentrated on the central point of the tip and within a small area around it.

From the simulation of scanning real and non-real surface, it has also been noticed that if contacts distribution is typical of real instruments, there is considerable lateral uncertainty in where the repeated surface heights occur, perhaps approaching the nominal tip dimension. This reduces doubt on the benefit of closely spaced samples of surface.

As expected, the maximum deviation occurs when using the 5 µm spherical tip with 2 µm truncation, but many cases show only small errors.

The simulations show that in most cases contact does not occur at the central point of the stylus, even with idealized shapes other than perfectly shaped ones. With the spherical tip, the mean position of the contact is close to the centerline, while its

standard deviation is about 0.5 μm on 3 μm radiuses. It tends to the radius of the flat on truncated ones.

6.4. Summary

A new simulation program has been developed and used to examine the measuring fine structure of real surfaces by the stylus method. Although able to scan any arbitrary surface with any arbitrary stylus shape, the results given here use idealized styli and ‘real’ ground steel surfaces.

The simulations have naturally confirmed that stylus geometry and size can have a significant effect on most roughness parameters of the measured surface in 3D. The surprising feature of these, worthy of greater investigation, is how insensitive to major changes in stylus condition some of the popular parameters are, even when dealing with very fine structure within localized areas of a ground surface.

Chapter 7

7. Implications for stylus profilometer uncertainty

7.1. Introduction

Basically, to avoid positional uncertainty in the presence of small tilts, It is better to record what happens with various simulations using good input data (e.g., sampled from an original surface map that has been levelled by something like SPIP software). The reasoning is that at the scales we are studying, there is no clear, functionally relevant way of distinguishing an overall tilt from a local one, where we happened to have a sample of the 'flank' of some inherent waviness. If so, then we can presume that all our data is valid for looking at overall uncertainty ranges and so on.

Taking a typical Talysurf configuration, with 2 µm radius hard tip and around 1 mN contact force, basic Hertzian stresses (hard ball on softer flat, or possibly hard ball on softer similar size ball to recognise asperity contact) give a rough idea of what might happen with thresholds. Such calculations are not very sensitive to individual parameter values because of all the fractional power relationships - which is good for a rough guide! Reckoning the ball-on flat gives illustrative results for Steel, Brass and Aluminium, such as:

Table 7.1: Example of illustrative results for Steel, Brass and Aluminium

Material	E (GPa)	Contact radius (µm)	Approach distance (nm)	contact area (µm²)
Steel	210	0.2	19	0.12
Brass	120	0.23	27	0.17
Aluminium	70	0.27	37	0.28

This suggests that if we are thinking in terms of whether local compression (approach) might pick up extra contact points, thresholds up to only 20-30 nm look of any great physical significance - so possibly going occasionally as high as 50 nm in simulation makes experimental sense, but we should concentrate effort at smaller values. The used grids 0.2 um samples, in which case each point in area counts represents 0.04 um^2 (each spot represents an area above that value if it has only just come into the map, but has that area if it was already in the map at the previous threshold step!). There is an expectation to accumulate 5-10 spots with thresholds of 20-30 nm. Comparing such arguments to what happens in the simulation results may eventually tell us a bit more about local structure, noise etc., (Incidentally, the Hertzian calculations show peak contact stresses between 5 and 12 GPa for the above conditions, way above normal elastic limits and so stressing that, in the absence of significant observed damage, the local ability to support the stylus must be a great deal better than bulk-material predictions. Of course, this is relatively well known, but it adds another complexity to thinking about what is going on between real measurements and simulations.)

There is another possible tilt-like effect that could be built into simulations. Consider a typical (Talysurf) with an analogue pick-up. The pen-shift offset will usually 'compensate' for up to about 25 μm of positional variation with respect to the centre of the stylus range. The tip is mounted on an arm of at least 50 mm length to the pick-up pivot, so, depending on the pen-shift position, an ideal flat on a stylus might actually be 'tilted' by up to 0.5 mrad. (On very rough surfaces, this will be an error source varying with profile height, but it will look almost 'systematic' within a specific measurement of a good quality surface.) If, for example, a flat on a stylus

is itself 2 um diameter, we have only about 1 nm of height uncertainty according to where on the flat first contact happens. This seems trivial enough not to be further studied at present - but it should nevertheless be recorded that it has been considered and put to one side. For comparison, if we had a surface of about 5 mm across, we might well proceed with a tilt error equivalent to a couple of um height change form side-to-side: this is also a tilt of odd tenths of a (mradian). The more 'interesting' thing to simulate is the difference between ideal spherical tips and those with 'worn' flats, either normal to the tip axis or deliberately inclined. The latter is justified because measurement conditions can be asymmetric: the stylus lever arm action means that there may be heavier loads and more wear on the outer edge of the tip (i.e., the trailing side of the normal scan). However, if better instrumental technique means that the stylus is lifted on the 'return' strokes, the wear might become somewhat more pronounced on the leading side of the normal scan. Overall, all this goes into almost too much detail, because real styli have imperfect profiles and the orientation of those 'fault features' will almost certainly be dominated by how the tip happened to be bonded onto the stylus arm.

7.2. Testing the threshold models

A further set of carefully integrated numerical experiments was carried out using the threshold idea, to demonstrate why it was an idea worth studying; and show its real worth (or conversely show clearly why it should be ruled out).

While, in terms of its purpose, it is easiest to think of each ‘experiment’ as independent, it is possible to save some work by setting up the actual execution so that there is a useful level overlap (the same reference data sets, some conditions

being common to more than one way of exploring a range of possibilities and so on.).

The data presented can support decisions about which areas (if any) should have priority for the larger-scale studies that could provide reliable answers. Each experiment has been run on several different examples of real surfaces, each logged via AFM at 0.2 µm high(ish) lateral resolution.

For a 0.2 µm fine grid, the stylus characteristic dimension will need to be around 2 µm for a decent practical compromise between resolution in the simulation and overall array sizes.

Stylus: Cone

Threshold: 0.005, 0.05 µm, and 0.05 µm respectively.

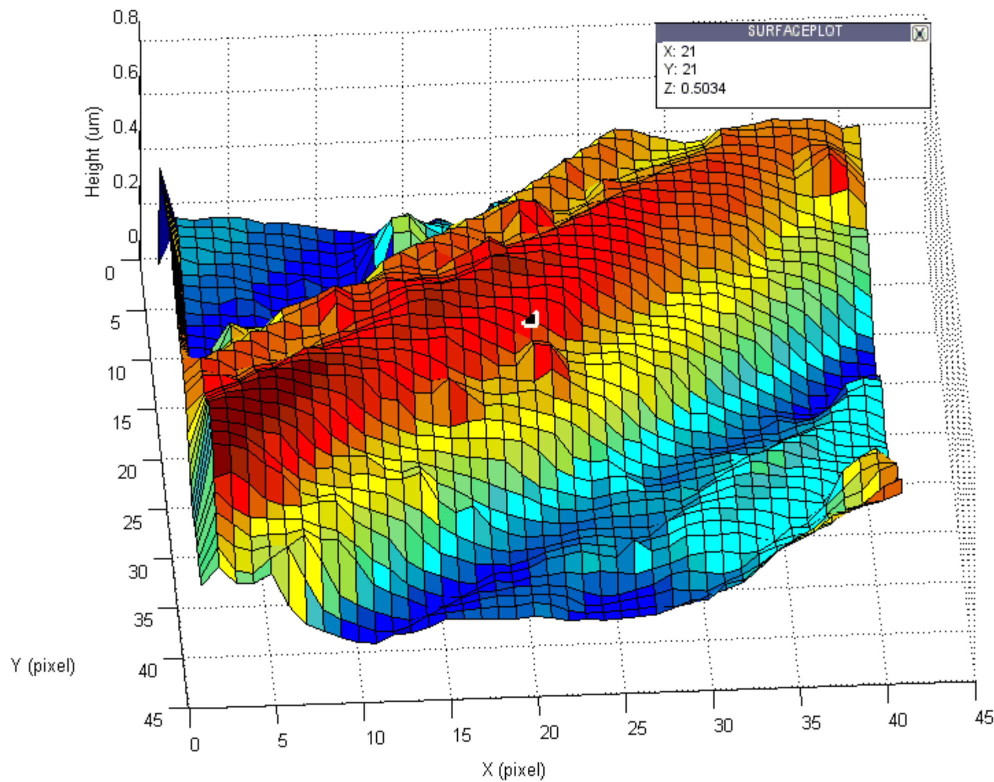


Figure 7.1: Ground surface with Cone stylus (Sample 1)

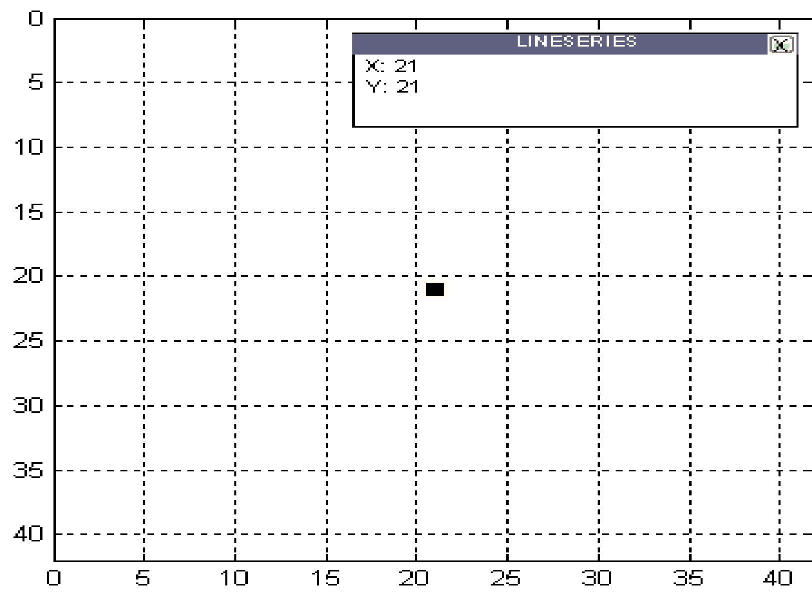


Figure 7.2: Map of contact with 5nm threshold (Sample 1)

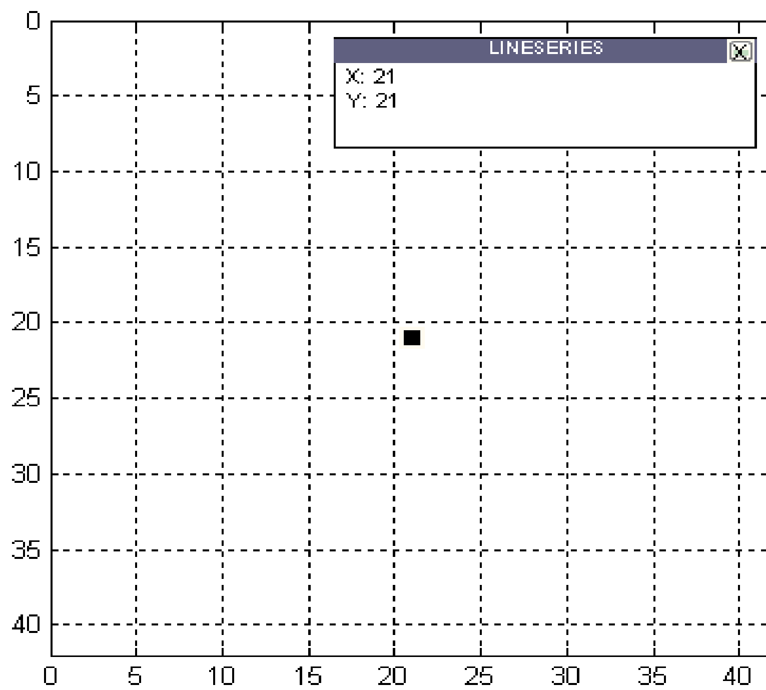


Figure 7.3: Map of contact with 20 nm threshold (Sample 1)

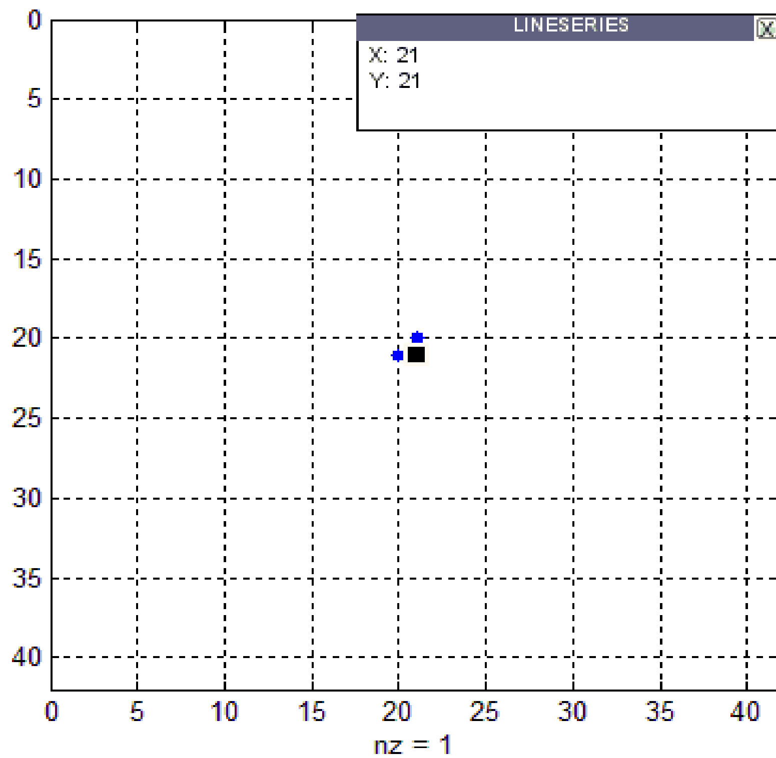


Figure 7.4; Map of contact with 5nm threshold (Sample 1)

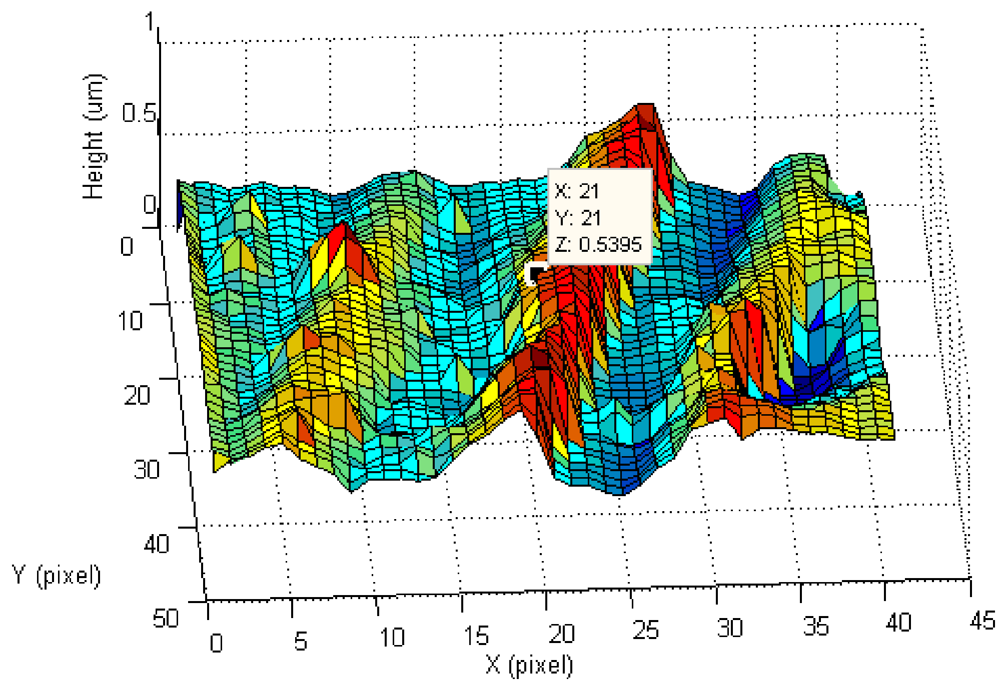


Figure 7.5: Ground surface with Cone stylus (Sample 2)

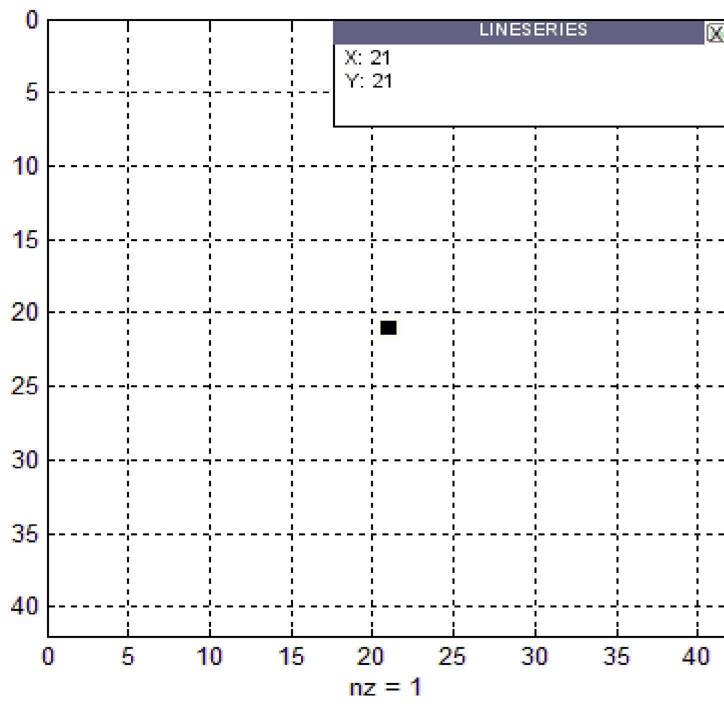


Figure 7.6: Map of contact with 5nm threshold (Sample 3)

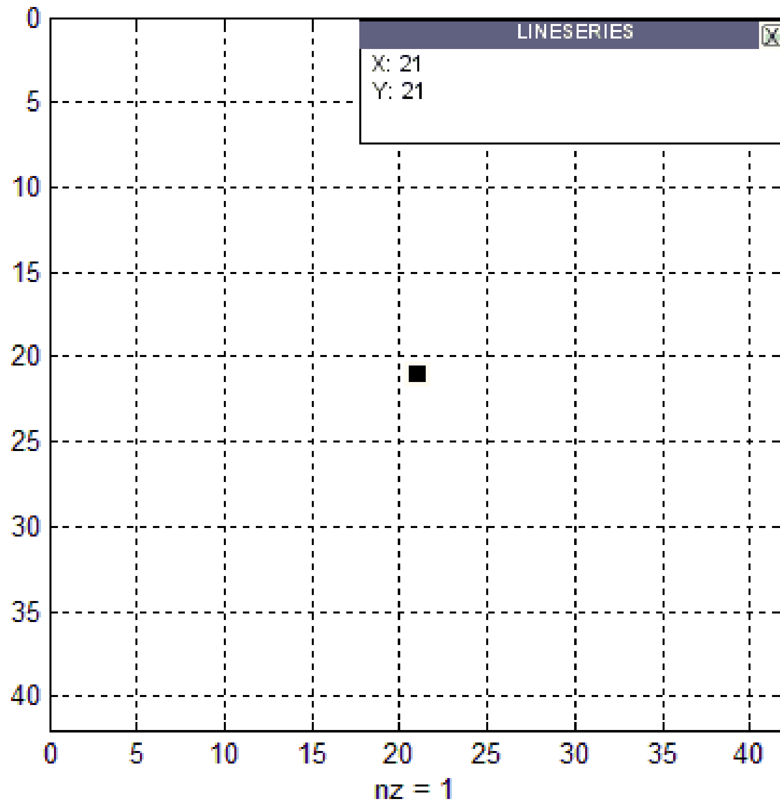


Figure 7.7: Map of contact with 20nm threshold (Sample 3)

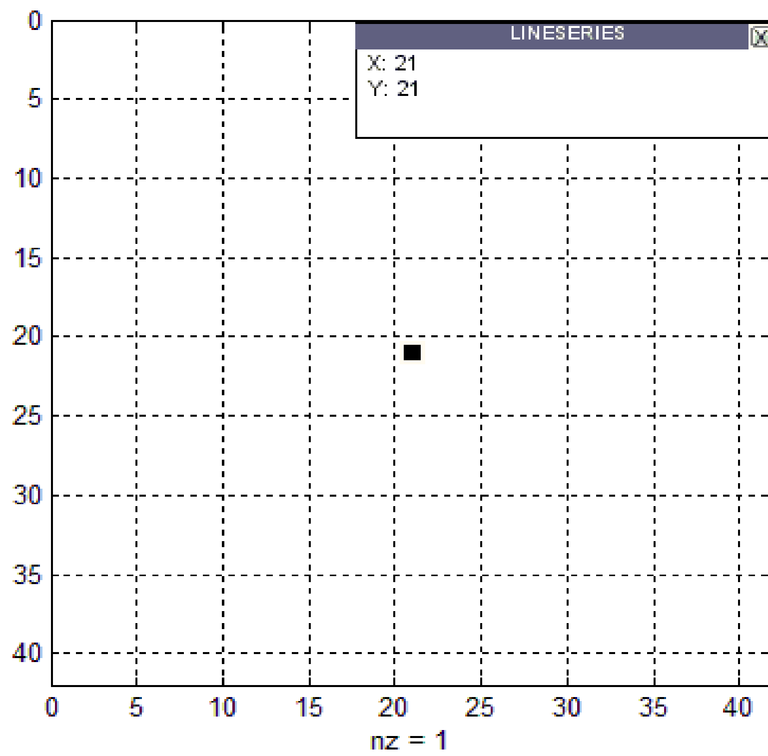


Figure 7.8: Map of contact with 50nm threshold (Sample 3)

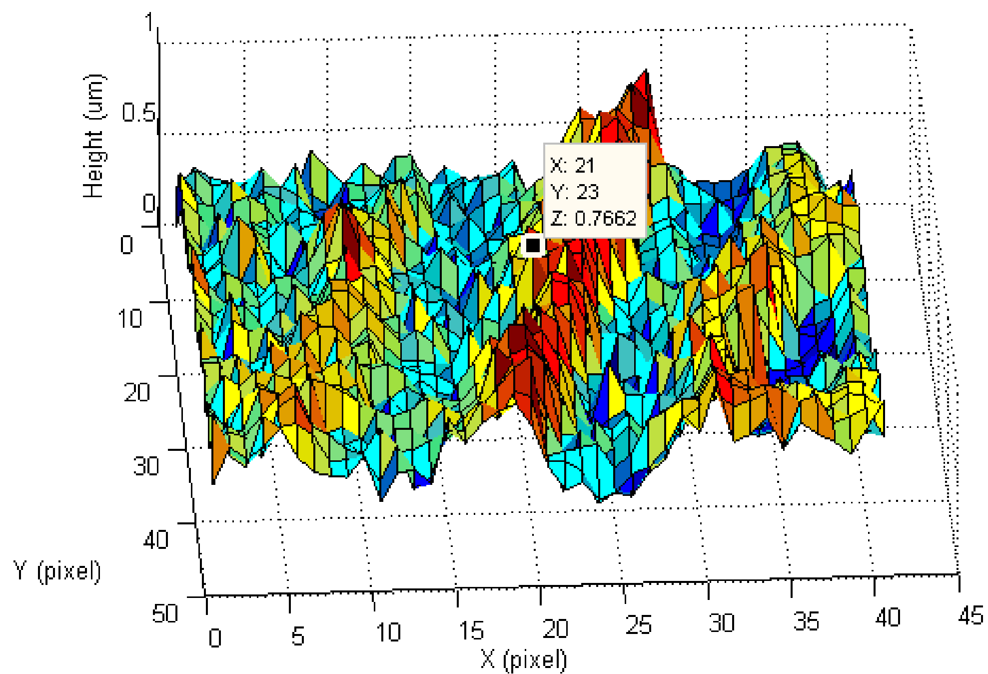


Figure 7.9: Three-D Ground surface with Cone stylus (Sample 4 - with Noise)

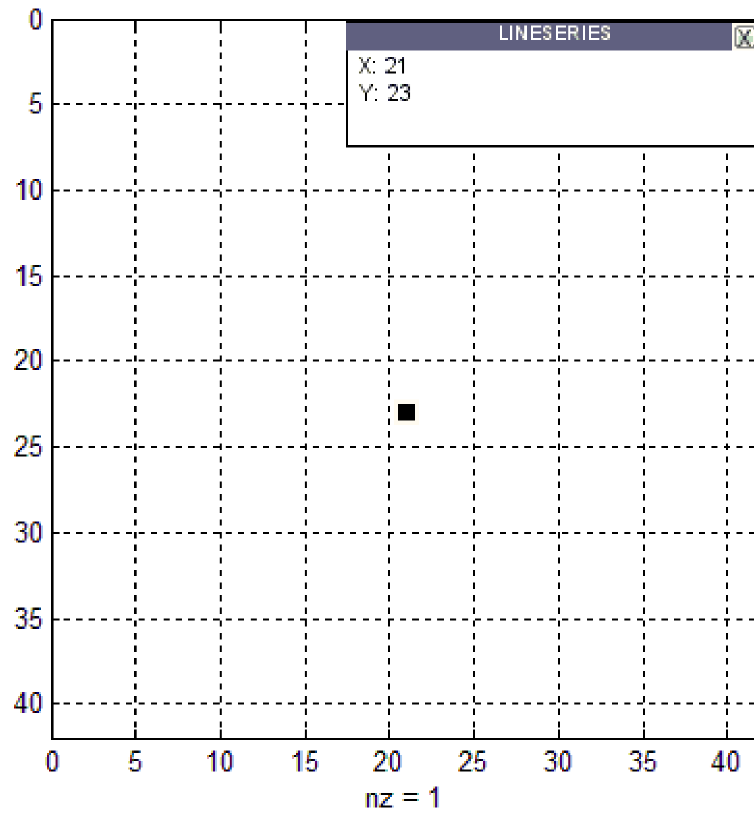


Figure 7.10: Map of contact with 5nm threshold (Sample 4)

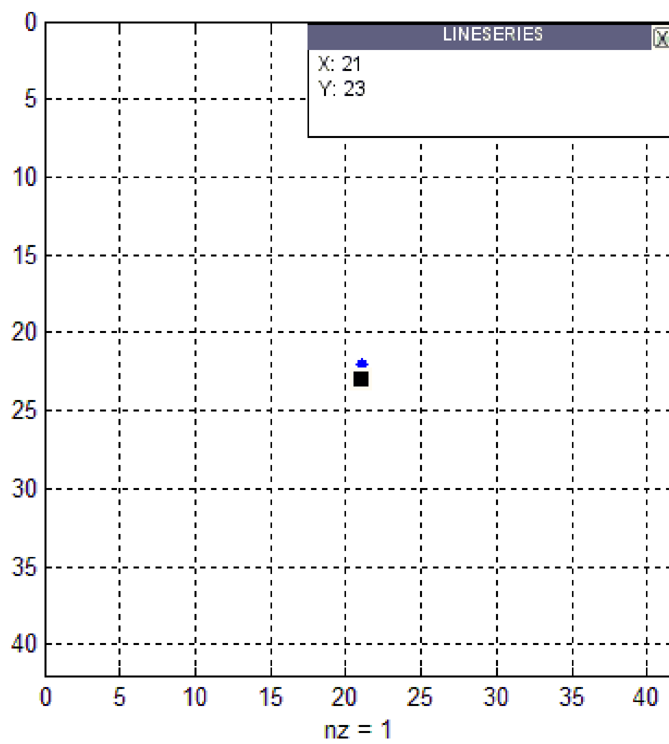


Figure 7.11: Map of contact with 20nm threshold (Sample 4)

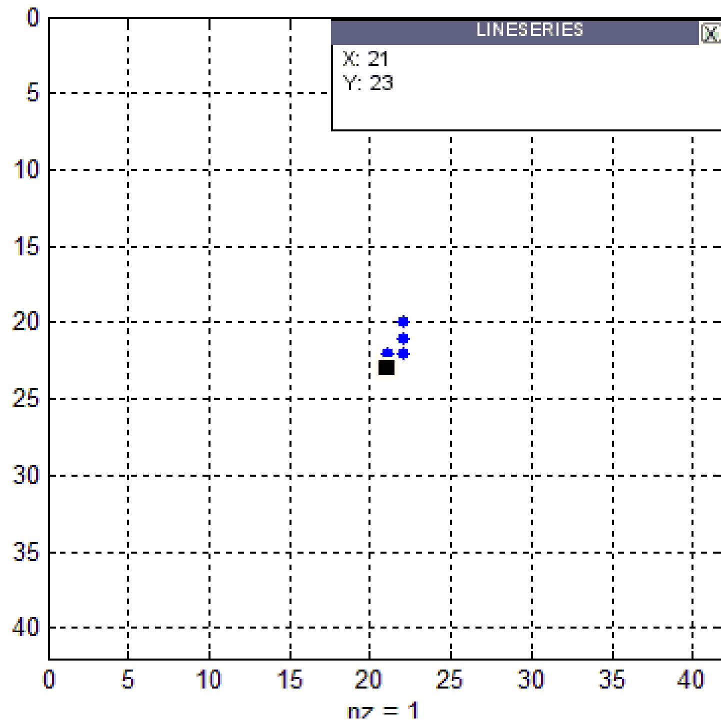


Figure 7.12: Map of contact with 50nm threshold (Sample 4)

Figures 7.1, 7.5 and 7.9 show a small map (40 by 40 points on a 0.2 µm sample grid) taken originally by atomic force microscopy from a typical ground steel surface and converted to Matlab©. The fine detail has both sharply varying and plateau-like regions. Surface maps were scanned in simulation by a 2 µm radius ideal conical stylus with 90º. A typical profilometer set up to measure finish ground surfaces might have an effective noise floor of the order of 10 nm to 20 nm peak. Its digitized resolution might typically be 1 nm to 10 nm. Alternatively, noise-related issues could be handled in terms of a fraction, say 0.5% to 0.1%, of full range (maybe a few tens of micrometres, typically). Hence, thresholds from 5 nm to 50 nm (for deliberate over-emphasis) seem relevant for study here. The simplest form of thresholding is roughly equivalent to ideally pressing the probe into the surface. For the illustrations used here the modified surfaces are largely unchanged

other than a slight vertical offset. Interest lies in how the shift interacts with additional surface features to alter the probe contact patterns.

Table 7.2: Kinematic/threshold contact points statistics for cone stylus

Surface\Cone	Sample 1	Sample 2	Sample 3	Sample 4 (+noise)
Kinematic contact – offset from centre (µm)	0.15	0	0.3	0.3
+ threshold @ 5 nm: Number of contact points	1	1	1	1
Centroid position (dominant)/offset (µm)	0.15	0	0.3	0.3
– Number of islands	1	1	1	1
+ threshold @ 20 nm: Number of contact points	1	2	1	2
Centroid position (dominant)/offset (µm)	0.15	0	0.3	0.3
– Number of islands	1	1	1	1
+ threshold @ 50 nm: Number of contact points	3	4	1	5
Centroid position (dominant)/offset (µm)	0.15	0	0.3	0.3
– Number of islands	1	1	1	1

The initial geometric kinematic contact is a single point which is offset from the centre of an ideal conical stylus tip as shown by the data in table 7.2. For all samples, the different thresholds cause the same offset from the stylus centre for each sample. With a threshold, contact points sufficiently close to the tip are considered to be in contact. Both the kinematic contact point and the centroid position of the dominant peak island (after threshold) have the same offset from the centre. Obviously because of the ideal conical shape, the contact points occur (With a very high probability) at the stylus centre, but some freak asperities could lead to an offset first contact.

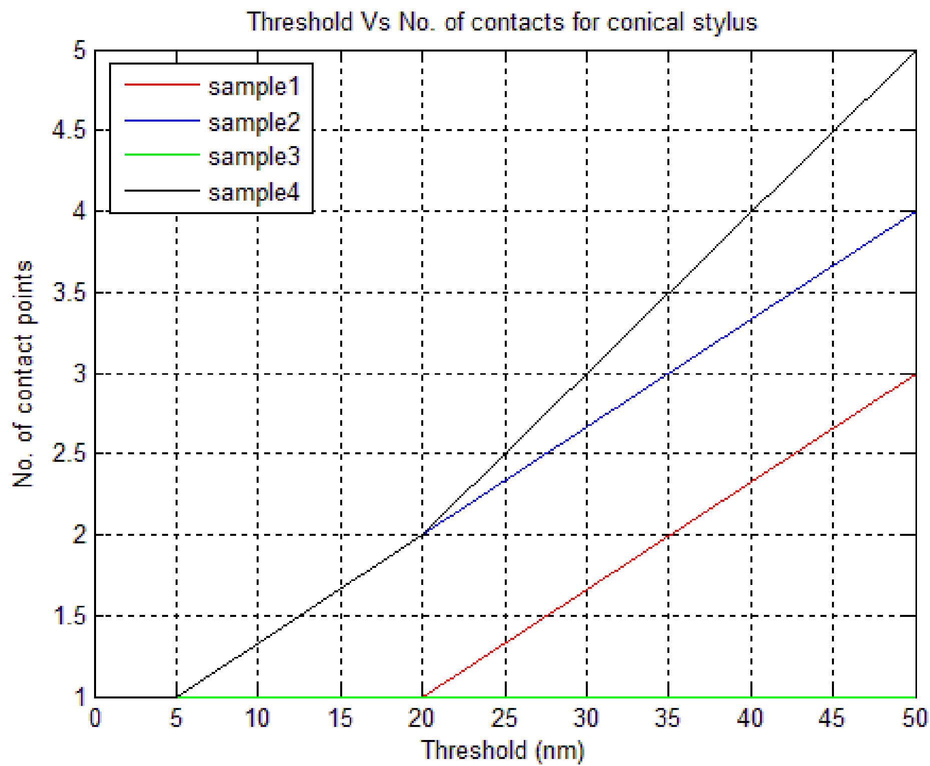


Figure 7.13: Threshold Vs. Number of contacts for Conical stylus

Figure 7.13 shows the relationship between threshold and number of contacts when scanning four different surfaces by using an ideal conical stylus. The graph shows that there has been a noticeable increase in the number of contact points as the threshold increases.

It shows that all four surfaces have the same growth of contact while applying up to 5 nm threshold. Then, the number of contact points in three surfaces gradually increased at 20 nm and 50 nm threshold except one surface (sample 3) which remains at one single contact point based on the rough surface profile features. Comparing sample 3 and sample 4 which is sample 3 plus noise, it can be noticed that there is a significant increase in the number of contacts after adding random noise to sample 3.

Table 7.3: Kinematic/threshold contact points statistics for sphere stylus

Surface\Sphere	Sample 1	Sample 2	Sample 3	Sample 4
Kinematic contact – offset from centre (μm)	0.3	0	0.46	0.46
+ threshold @ 5 nm: No. of contact points	4	2	2	1
Centroid position (dominant)/offset (μm)	0.22	0	0.46	0.46
– Number of islands	1	1	1	1
+ threshold @ 20 nm: No. of contact points	21	8	3	2
Centroid position (dominant)/offset (μm)	0.22	0	0.46	0.31
– Number of islands	1	2	1	1
+ threshold @ 50 nm: No. of contact points	37	19	7	3
Centroid position (dominant)/offset (μm)	0	0.08	0.54	0.31
– Number of islands	2	2	2	2

The initial geometric kinematic contact is a single point which is offset from the centre of the sperical stylus tip as shown by the data in table 7.3. For all samples, the different thresholds cause the same offset for each sample.

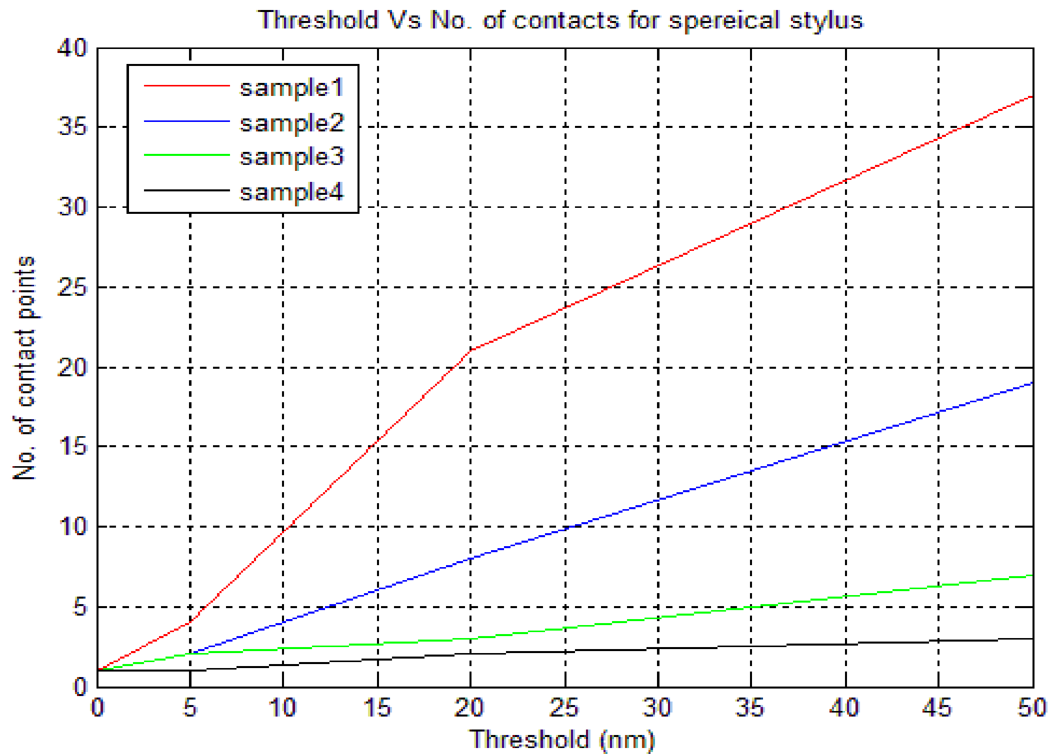


Figure 7.14: Threshold Vs. Number of contacts for spherical stylus

Figures 7.14 and 7.15 show the relationship between threshold and the number of contacts in-crease when scanning four different surfaces by using an ideal sphere and flat styli, respectively.

Table 7.4: Kinematic/threshold contact points statistics for flat stylus

Surface\Flat	Sample 1	Sample 2	Sample 3	Sample 4
Kinematic contact – offset from centre (μm)	2.6	2.9	0	1.23
+ threshold @ 5 nm: No. of contact points	2	2	1	4
Centroid position (dominant)/offset (μm)	2.6	2.9	0	1.23
– Number of islands	1	1	1	2
+ threshold @ 20 nm: No. of contact points	15	2	1	5
Centroid position (dominant)/offset (μm)	2.5	2.9	0	1.23
– Number of islands	1	1	1	3
+ threshold @ 50 nm: No. of contact points	63	7	3	10
Centroid position (dominant)/offset (μm)	2.2	2.9	0	0.42
– Number of islands	3	2	3	4

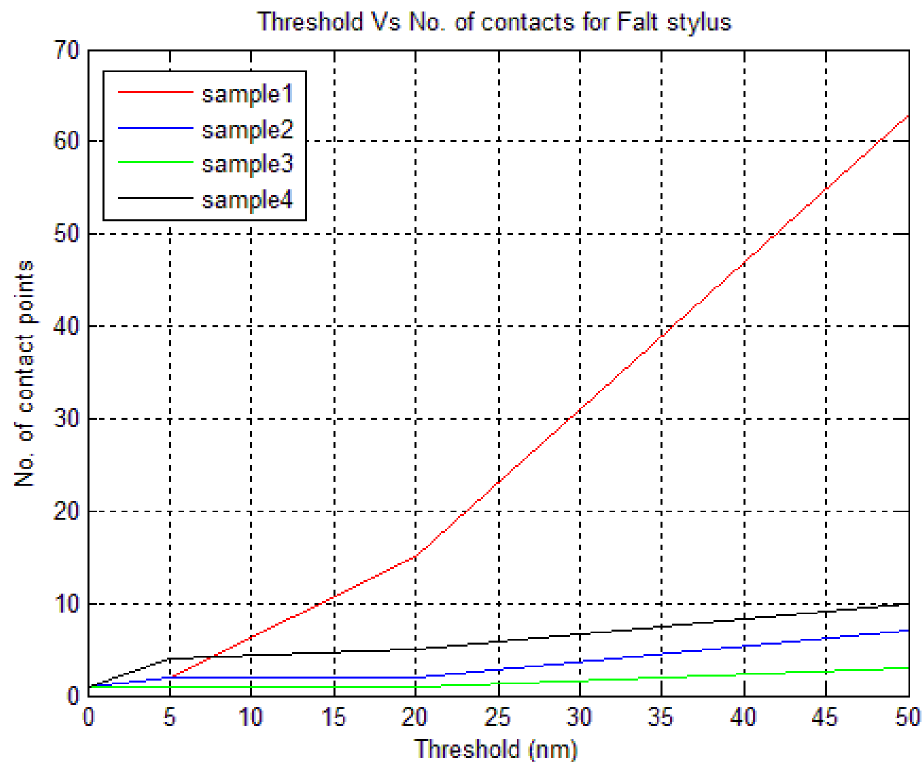


Figure 7.15: Threshold Vs. Number of contacts for Flat stylus

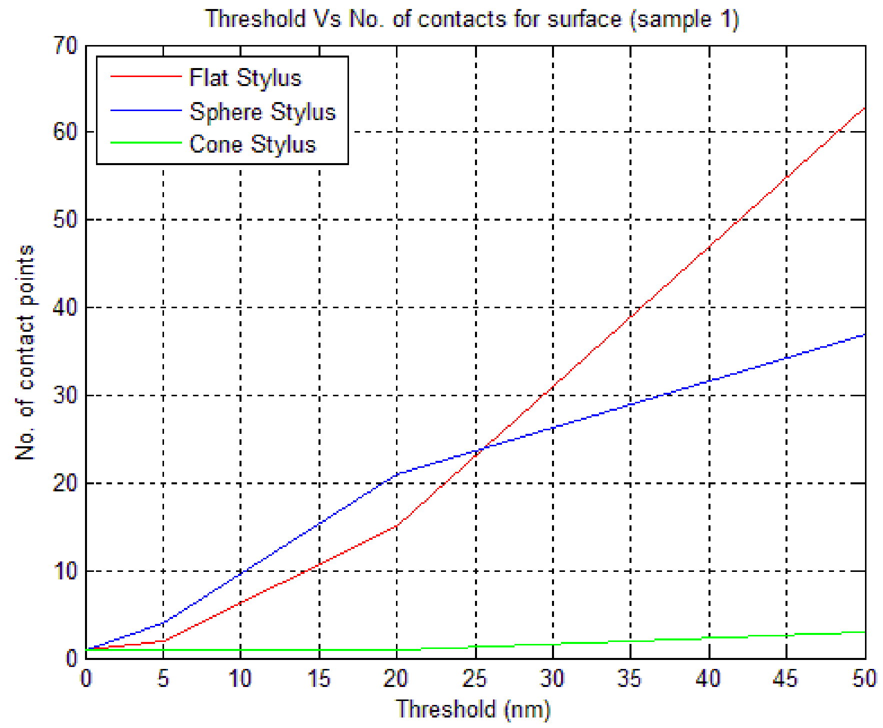


Figure 7.16: Threshold Vs. Number of contacts for surface (sample 1)

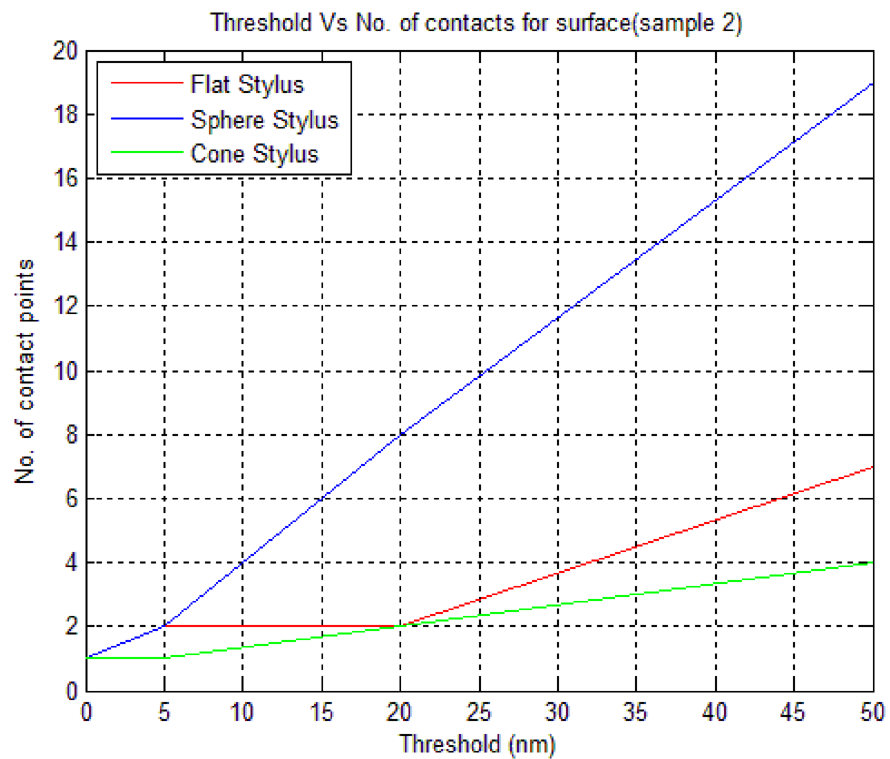


Figure 7.17: Threshold Vs. Number of contacts for surface (sample 2)

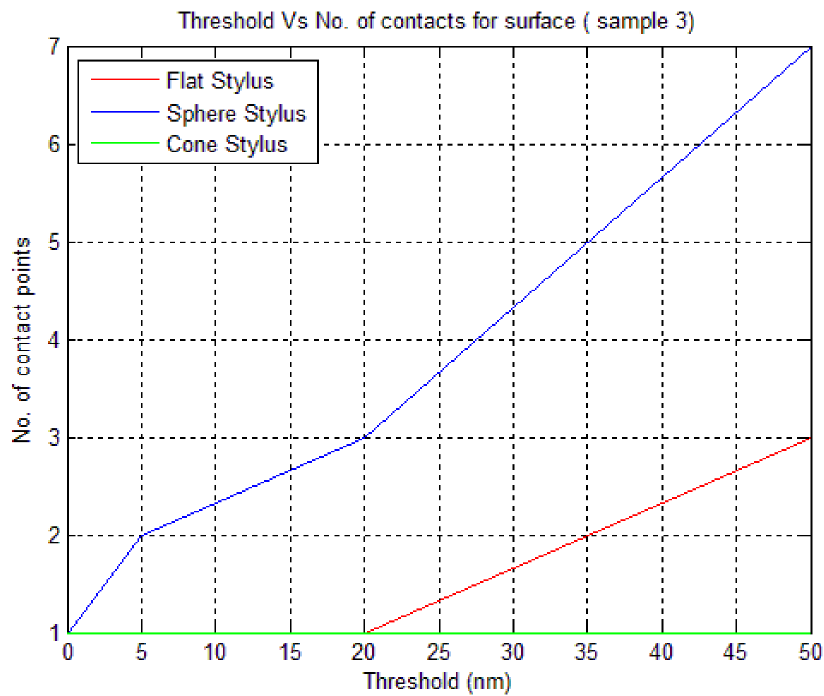


Figure 7.18: Threshold Vs. Number of contacts for surface (sample 3)

Figures 7.16 7.17, 7.18 and 7.19 show comparisons between threshold and number of contacts increase when scanning four different surfaces by using ideal flat, sphere and conical styli.

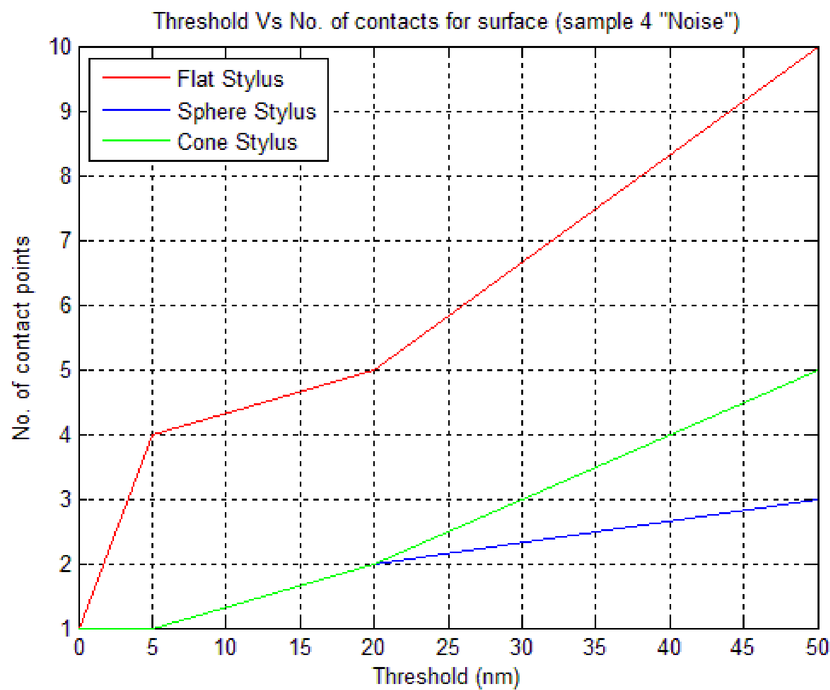


Figure 7.19: Threshold vs. Number of contacts for surface (sample 4 -Noise)

In figure 7.20 a contact between a rigid sphere and rough surface is shown schematically. For illustration, there are the following examples of calculating the contact offset and Z errors by using the Hertz contact theory (covered in chapter 2 and 3) to each contact spot;

For the Sphere:

Stylus centre has a coordinate point (20,20) on the stylus contact map.

While the Kinematic contact point at (18,19)

$\varepsilon_{lat} = 0.2\ \mu m$, which ε_{lat} is the lateral uncertainty (roughly standard deviation or could be an island centroid)

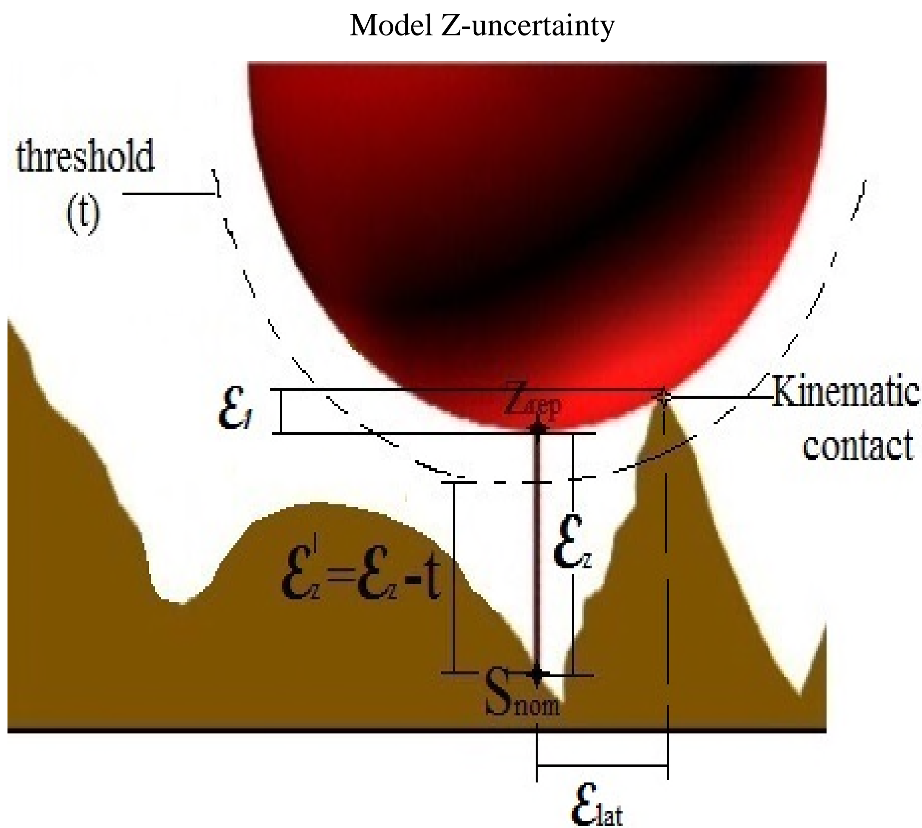


Figure 7.20: Hemisphere stylus interact with rough surface

Mathematically, calculated the *Z* error as (real *Z* contact height – reported height) or *Z* error as (real *Z* nominal (centre) – reported height).

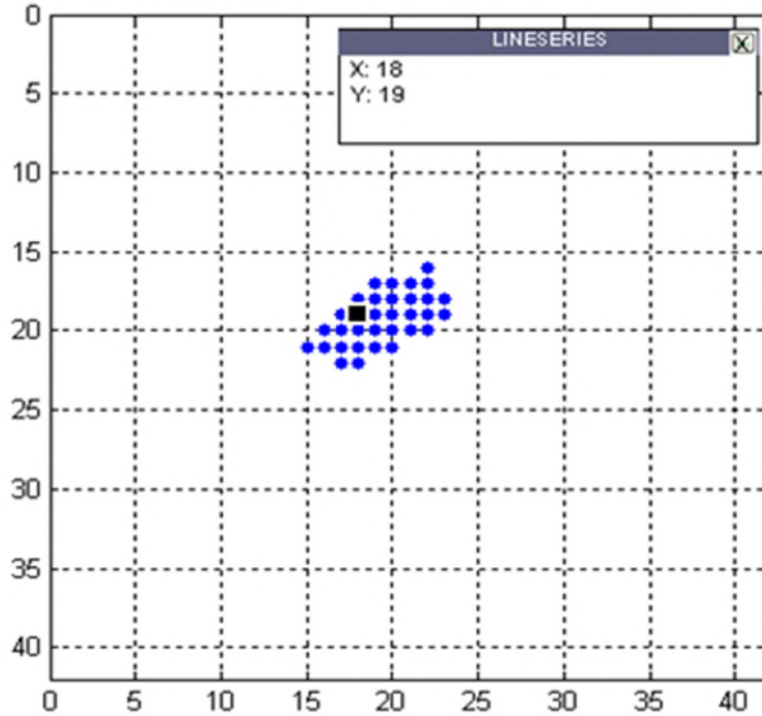


Figure 7.21: Map of contact points under kinematic condition (20 nm threshold)

$$z_{rep} \approx z_{cont} \approx 0.71$$

$$\varepsilon_z = 0.09\mu m$$

$$\varepsilon'_z = 0.09\mu m - thr$$

$$\varepsilon'_z = 0.09\mu m - 0.05 = 0.04\ \mu m$$

$$\varepsilon_1 \sim \frac{{\varepsilon_{lat}}^2}{2R} = \frac{(0.2)^2}{2x2} = \frac{0.05}{4} = 0.125\ \mu m = 12.5\ nm$$

Sphere with (0.05 µm) applied threshold appears to intercept the nominal contact place. Therefore, after applying the Hertz Contact theory ε_z no more than 30 nm.

Hertz Contact: Sphere vs flat

$$\varepsilon_z = \left[\frac{9F^2}{16\,R\,E^{*2}}\right]^{1/3}$$

where R= 2 µm = 2x10^{-6} m

E^* is the Young's modulus which is the ratio of stress (pressure 'pa') to strain (dimensionless), and so Young's modulus has units of pressure in SI unit is therefore the Pascal (Pa or N/m2).

$E^* \approx$ 100 GPa = 1x10^{11} pa, and then E^{*2} =1 x 10^{22} Pa (Assuming the material is brass, the stainless steel is ~ 200 GPa

F= 1 mN = 1 x 10^{-6} N

$$\varepsilon_z = \left[\frac{9F^2}{16\,R\,E^{*2}}\right]^{1/3}$$

$$\varepsilon_z^3 = \frac{9\;x\;10^{-6}}{(16)(2\;x10^{-6})(1x10^{22})} = \frac{9x10^{-6}}{3.2\;x\;10^{17}} = 2.8\;x10^{-23}m^3$$

$$\varepsilon_z = 30\;nm$$

For the conical stylus, as "practically" correct x= y=21 and z=0.62

Stylus centre has a coordinate of point (21,21) on the stylus contact map and the Kinematic contact occurs at the same point.

For the Flat stylus: $\varepsilon_{lat} = 2.2\ \mu m$

$$z_{rep} = z_{cont} \approx 0.76$$

$$\varepsilon_z = 0.14\mu m$$

$$\varepsilon'_z = 0.14\mu m - thr$$

$$\varepsilon'_z = 0.14\mu m - 0.05 = 0.09\ \mu m$$

7.3. Surface interpolation using 'fractal infill' methods

The following numerical experiment has been carried out to interpolate surface data in order to work with higher resolution stylus images. Basically, it aims to develop and demonstrate the general validity of an interpolation method that makes some physical sense in the context of surface metrology and tribology. Better might be to take a larger surface patch, sub-sample it by some reasonable factor (4 to 10), interpolate to a new variant by the same factor and compare this to the original, simply subtract them to show the differences. Some form of line interpolation is probably a good starting point (illustrative algorithm outlined in the appendix). Then, random or fractal-based small disturbances could be introduced on top of that line interpolation.

An interpolator for surfaces sampled on regular grids has been used to consider a representation of a surface as a set of point heights sampled on a regular rectangular grid; for convenience, a square grid, sampled at spacing *h* in each direction, is used here. The task is to superimpose these points within a finer grid, spacing *δh*, infilling the intermediate points of the fine grid with broadly representative values; the exact heights of the real points are to be preserved. As a first stage a 'smooth', interpolation can be used, comprising a set of straight lines that form a warped plane

to indicate the general trend between actual points (a true plane can fit only three points, not the four here required).

The basic procedure is repetitive for infilling between any adjacent points. Therefore we consider only four points at the corners of one square 'cell' of the original grid, z_{II}, z_{IJ}, z_{JI}, z_{JJ}. Four slopes are defined to represent the trend along the edges of this cell,

$$\Delta_{IJ}=\frac{z_{IJ}-z_{II}}{h}\ ;\quad \Delta_{JI}=\frac{z_{JI}-z_{II}}{h}\ ;\quad \Delta_{JJI}=\frac{z_{JJ}-z_{JI}}{h}\ ;\quad \Delta_{JJJ}=\frac{z_{JJ}-z_{IJ}}{h}$$

Hence, height can be predicted for any position along an edge by, e.g.,

$$z_{xI}=z_{II}+\Delta_{JI}x \quad \text{etc.}$$

It has been chosen (arbitrarily) to construct the interpolation as lines parallel to *x*. The interpolation is forced to be in terms of an integer factor *n*, so imposing the constraint on parameter choices that $h = n\delta h$. (*n*-1) new points will be inserted parallel to an axis between any two original ones. So, in this new grid, the original corner points of the cell above will be at relative indices (0, 0), (0, *n*), (*n*, 0) and (*n*, *n*) respectively. Using these relative coordinates for notational convenience, a set of spot heights are now determined along the lines parallel to the *y*-axis from z_{00} to z_{0n} and from z_{n0} to z_{nn}:

$$\left.\begin{aligned} z_{0k} &= z_{00}+\Delta_{IJ}(k\delta h)\\ z_{1k} &= z_{10}+\Delta_{JJI}(k\delta h)\end{aligned}\right\}\ k=0,1,\cdots,n$$

Pairs of points has been connected, so generated for a given *k* by an intermediate slope (i.e. for a straight line parallel to the *x*-axis)

$$\Delta_k=(z_{1k}-z_{0k})/h$$

These slopes can be used to fill in as many points as we wish along lines parallel to *x*, specifically those for the *n*-by-*n* interpolation as

$$z_{jk} = z_{0k} + \Delta_k(j\delta h) \qquad j = 0, 1, \cdots, n$$

This process preserves the original data at the cell corners; as presented above it effectively overwrites them with the same values. The process as written allows for the final interpolated array being a separately identified one, constructed only at a final step. A specific implementation could well write the cell corners and then the edge lines into the output array at earlier stages and then the in-fill step need only operate from index 1 to (n-1). As written, it overwrites the whole of the edges with their previous values.

The process has generated the set of z_{jk} heights as a quite 'smooth' patch of high resolution data that in some sense represents the likely local trend of the surface between points of the original data set. This trend (in the form of a 'twisted plane') could be perturbed in various ways. For example, a small random variation could be added point-by-point to represent measurement noise, or we could add in any other suitably scaled data set thought to have potential physical relevance. The latter might, for example, involve drawing on ideas from fractals and self-similarity to use a scaled-down version of some of the original data.

Finally the line-based trends match smoothly as we progress around points of the original grid, because the edge lines are common (indeed some time could be saved at the cost of a more complex implementation by not repeating their computation for each patch). However, where additive perturbations are included, later ones may well overwrite previous ones on the cell edges, which might give rise to larger steps than are desirable; it would, of course be possible to blend these by some type of average of all the perturbations relevant to application at any common point.

Fractal Infill: Sphere stylus on Zorig, Z, Zexpp and Zexpf respectively

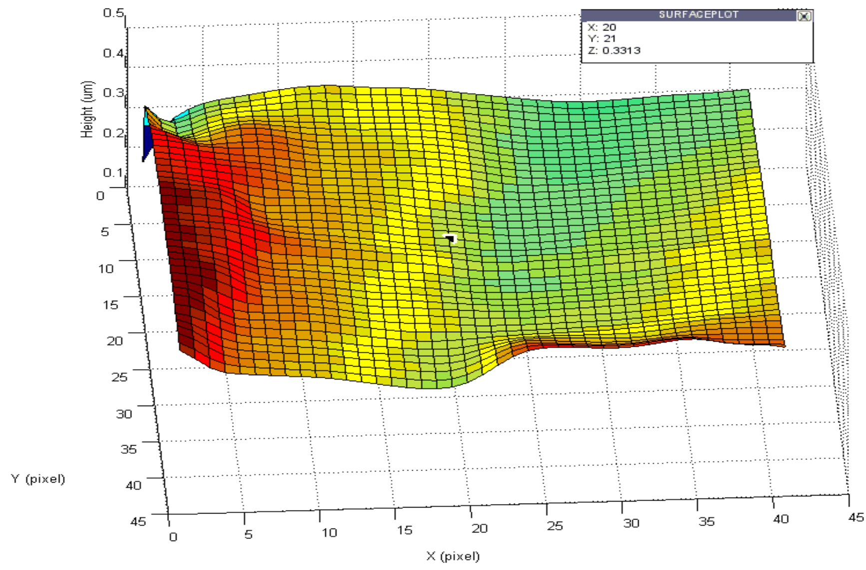


Figure 7.22: Ground original surface (Zorig) using Cone stylus

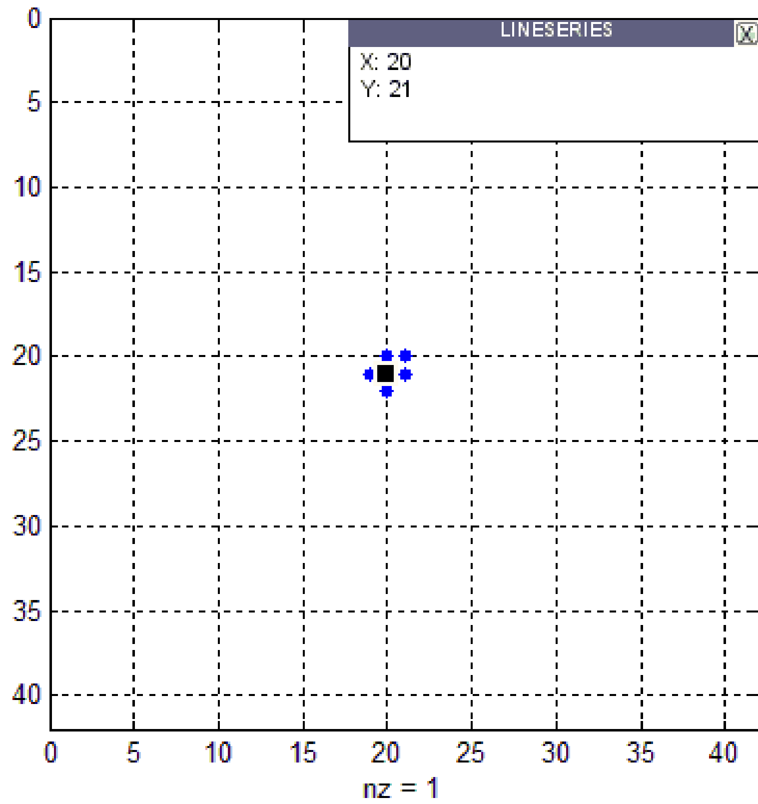


Figure 7.23: Map of contact with 5nm threshold (Zorig)

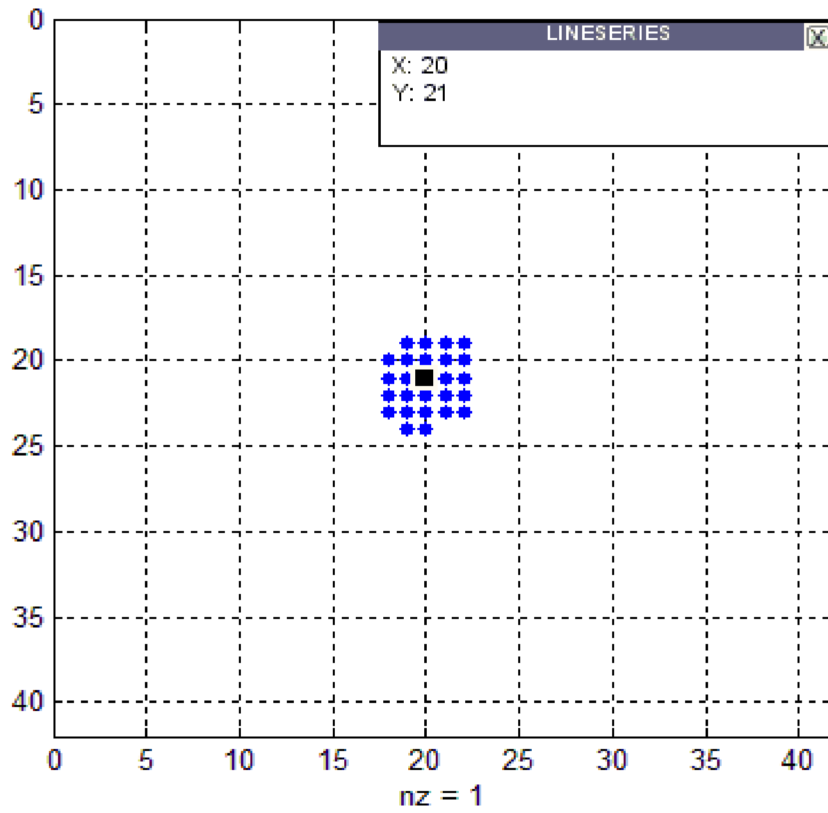


Figure 7.24: Map of contact with 20 nm threshold (Zorig)

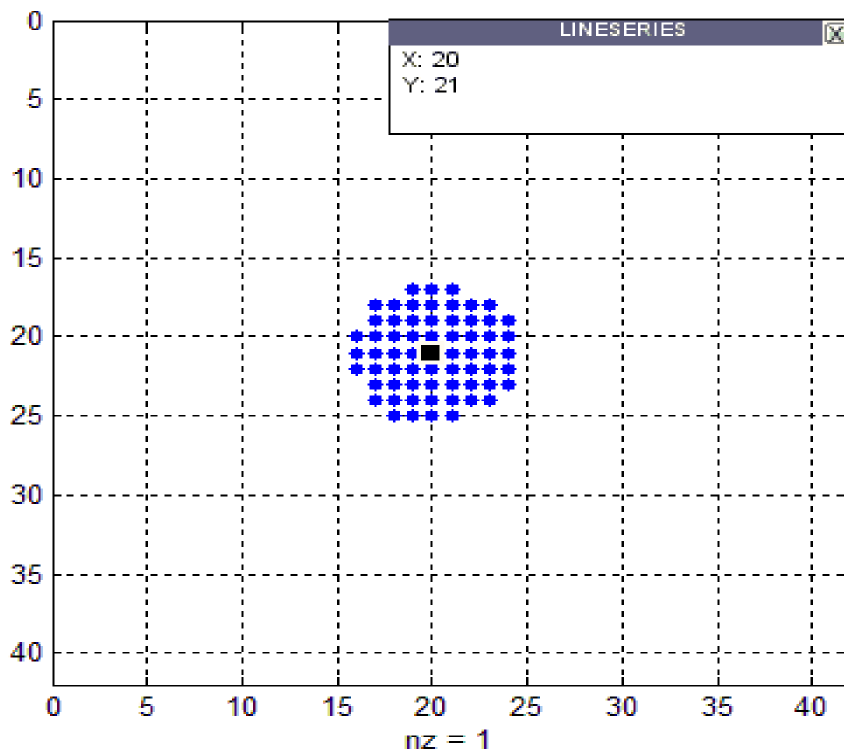


Figure 7.25: Map of contact with 50nm threshold (Zorig)

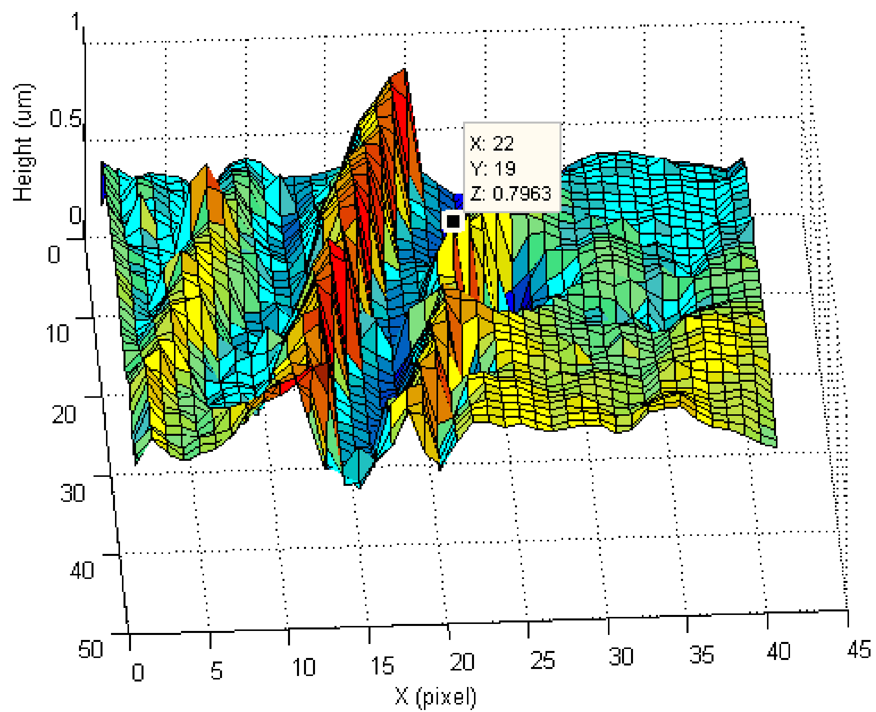


Figure 7.26: Sub-sampled Ground surface (Z) using sphere stylus

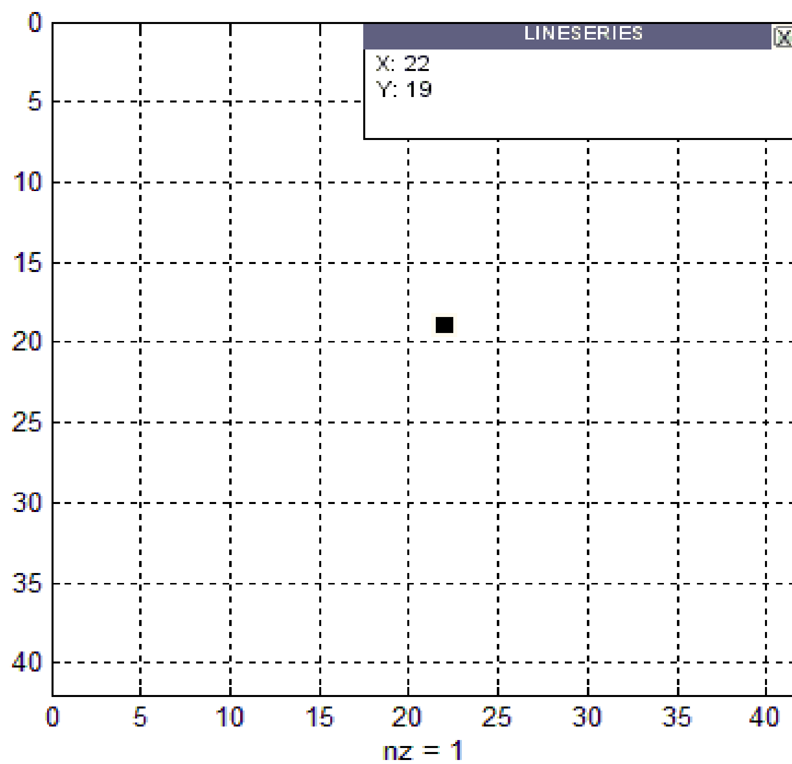


Figure 7.27: Map of contact with 5nm threshold (Z)

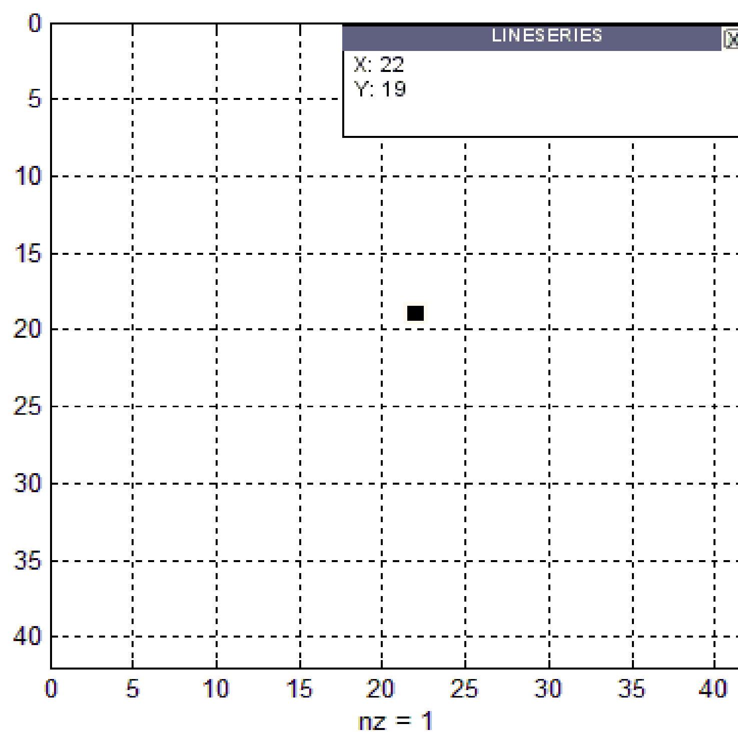


Figure 7.28: Map of contact with 20 nm threshold (Z)

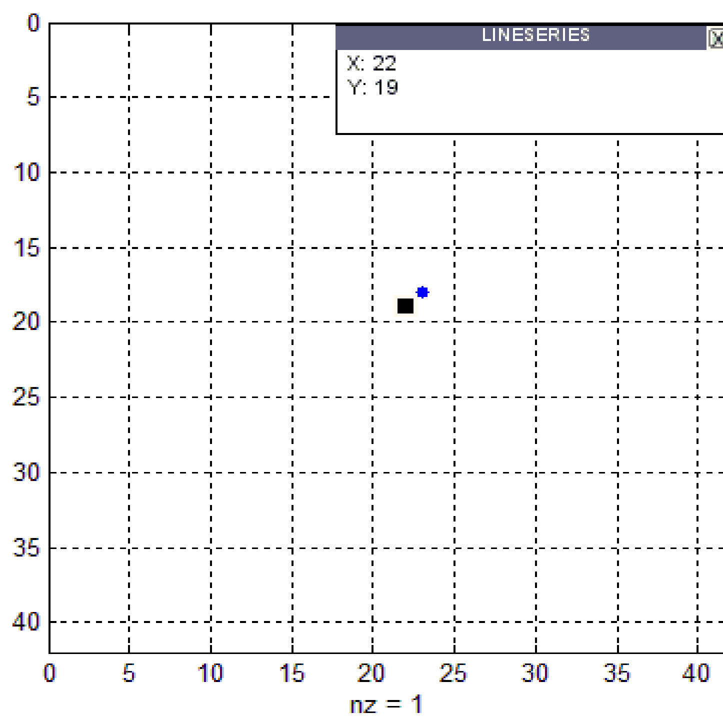


Figure 7.29: Map of contact with 50 nm threshold (Zg)

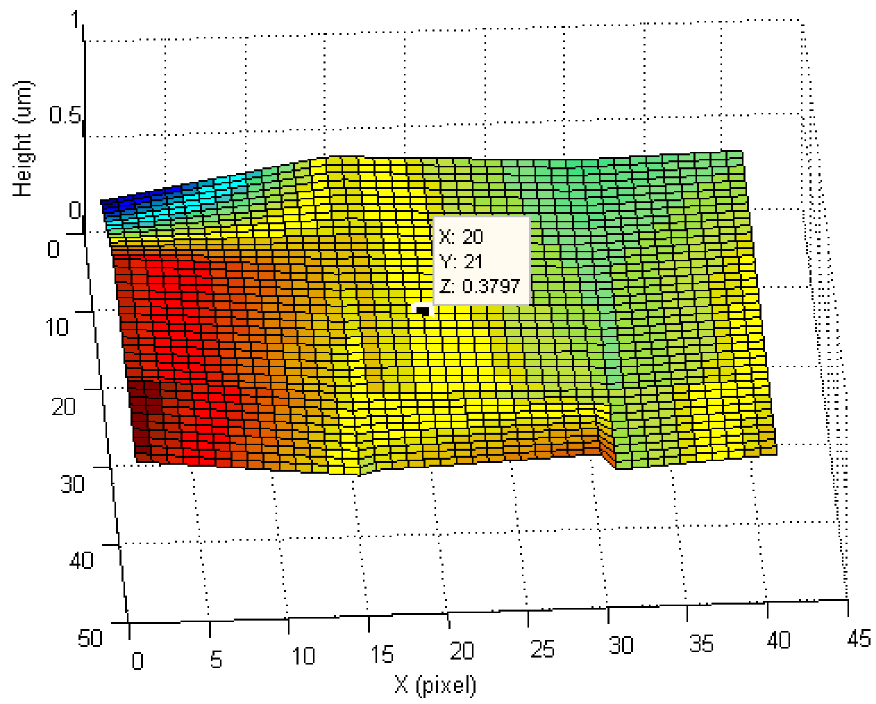


Figure 7.30: expanded "Patch" surface (Zexpp) using sphere stylus

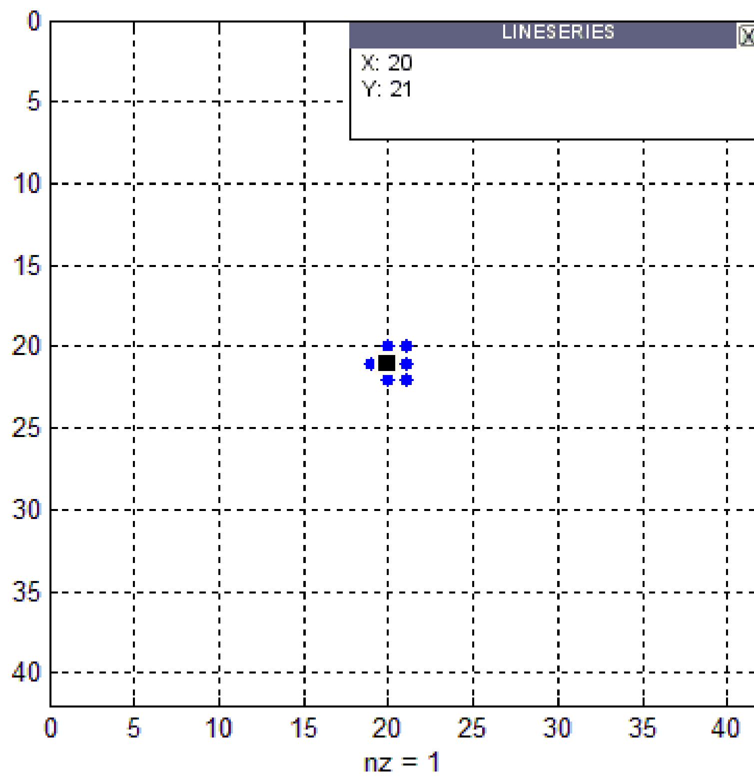


Figure 7.31: Map of contact with 5nm threshold (Zexpp)

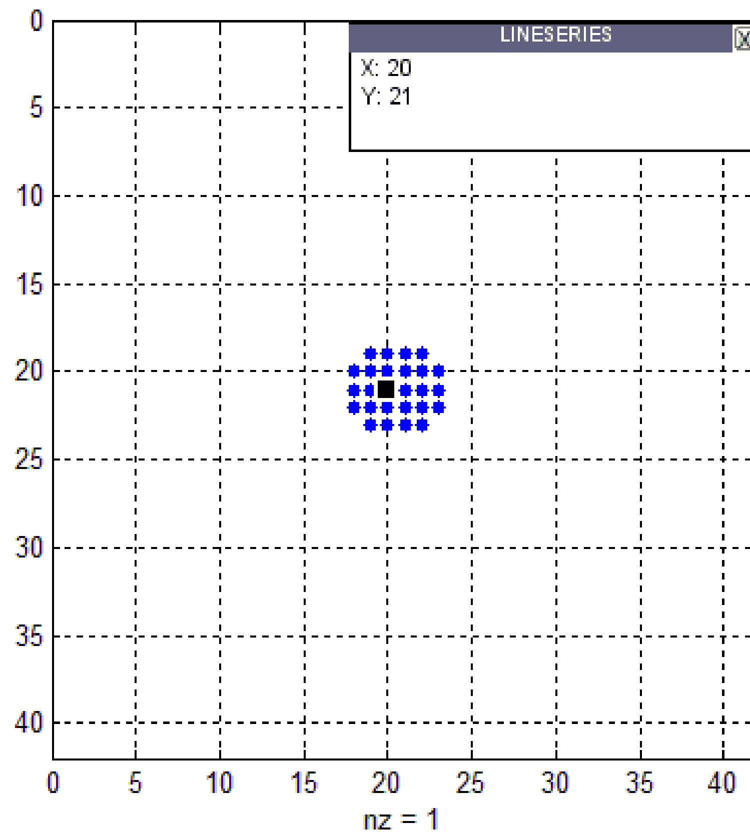


Figure 7.32: Map of contact with 20nm threshold (Zexpp)

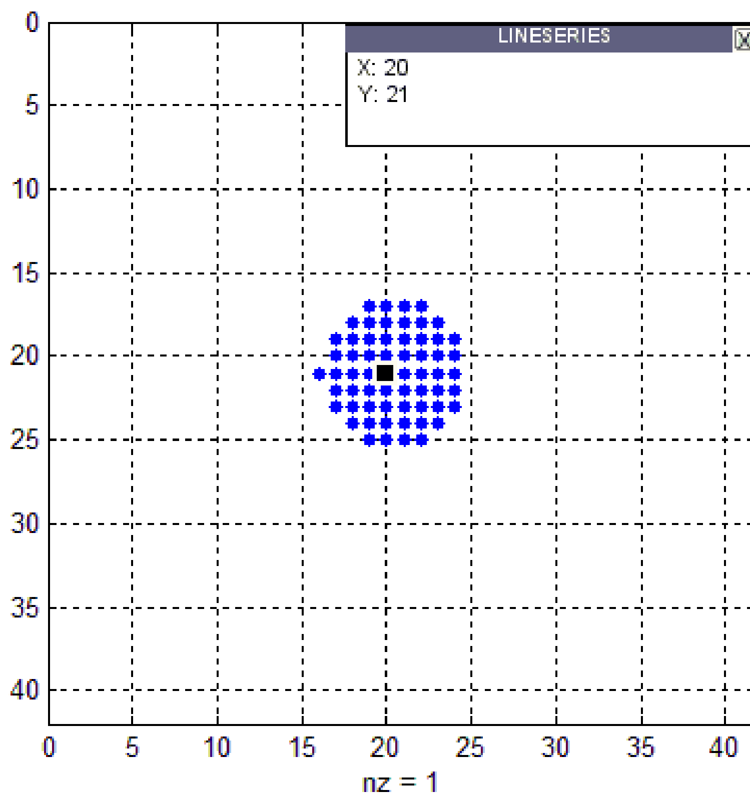


Figure 7.33: Map of contact with 50nm threshold (Zexpp)

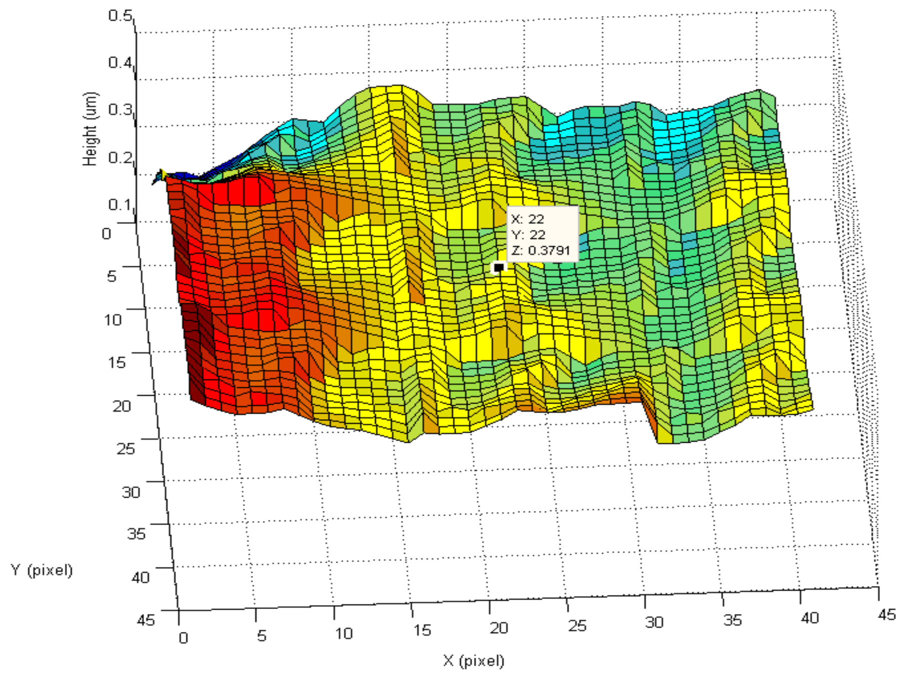


Figure 7.34: Expanded "Patch+filler" surface (Zexpf) using sphere stylus

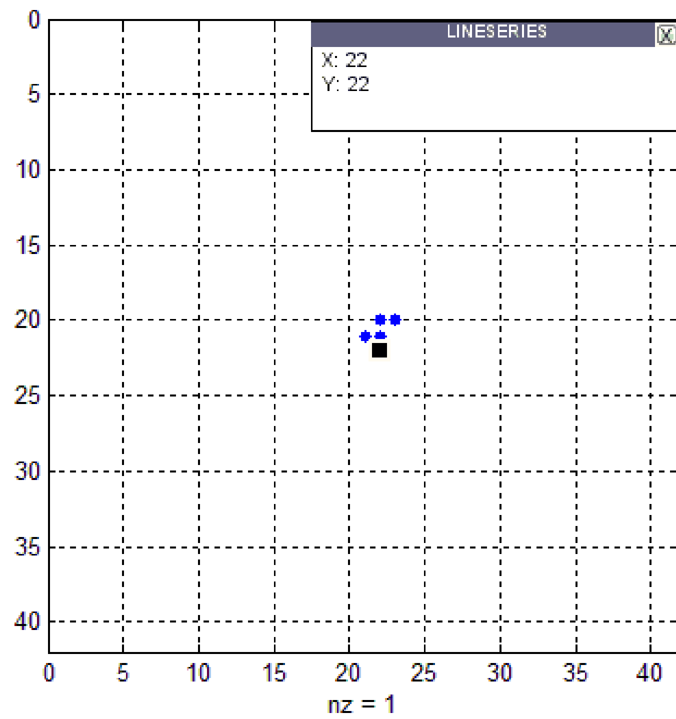


Figure 7.35: Map of contact with 5nm threshold (Zexpf)

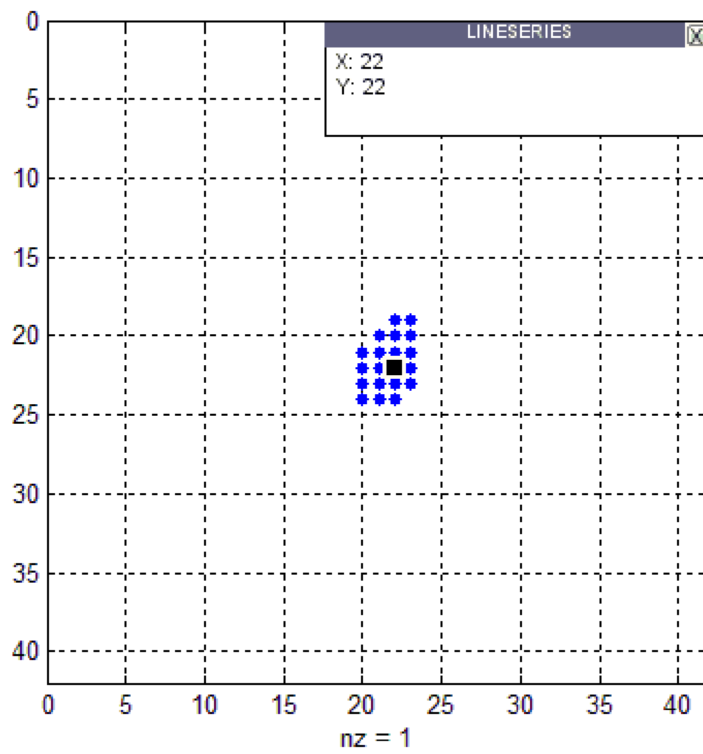


Figure 7.36: Map of contact with 20nm threshold (Zexpf)

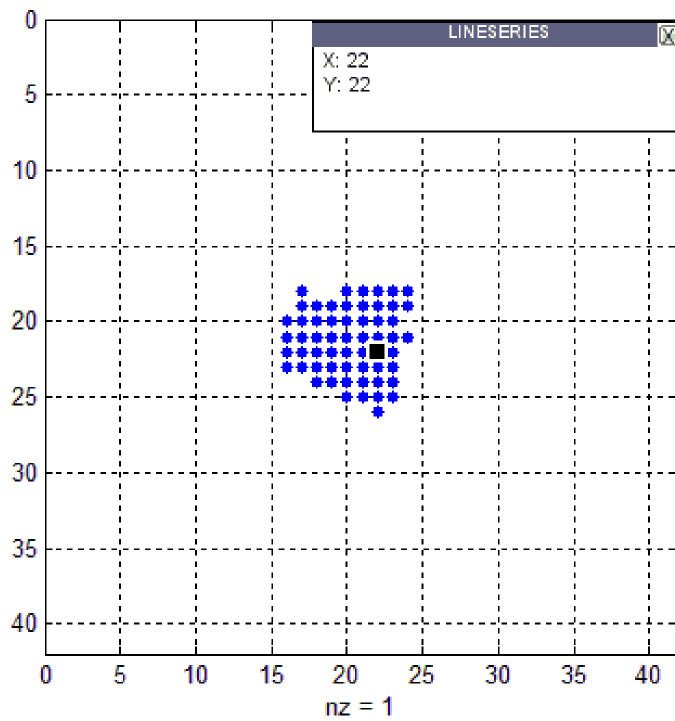


Figure 7.37: Map of contact with 50nm threshold (Zexpf)

Figures 7.22, 7.26, 7.30 and 7.34 show small maps of square master array Zorig, sub-sampled array Z, expanded interpolated 'patch' Zexpp and plane interpolated patch plus fractal 'noise' Zexpp, respectively.

Table 7.5: Kinematic/threshold contact points statistics by using sphere stylus

Surface\Sphere	Sample 1 (Zorig)	Sample 2 (Z)	Sample 3 (Zexpp)	Sample 4 (Zexpf)
Kinematic contact – offset from centre (µm)	0	0.31	0	0.31
+ threshold @ 5 nm: Number of contact points	6	1	7	5
Centroid position (dominant)/offset (µm)	0	0.31	0	0.31
– Number of islands	1	1	1	1
+ threshold @ 20 nm: Number of contact points	26	1	26	20
Centroid position (dominant)/offset (µm)	0	0.31	0.78	0.23
– Number of islands	1	1	1	1
+ threshold @ 50 nm: Number of contact points	65	2	61	58
Centroid position (dominant)/offset (µm)	0	0.39	0	0
– Number of islands	1	1	1	1

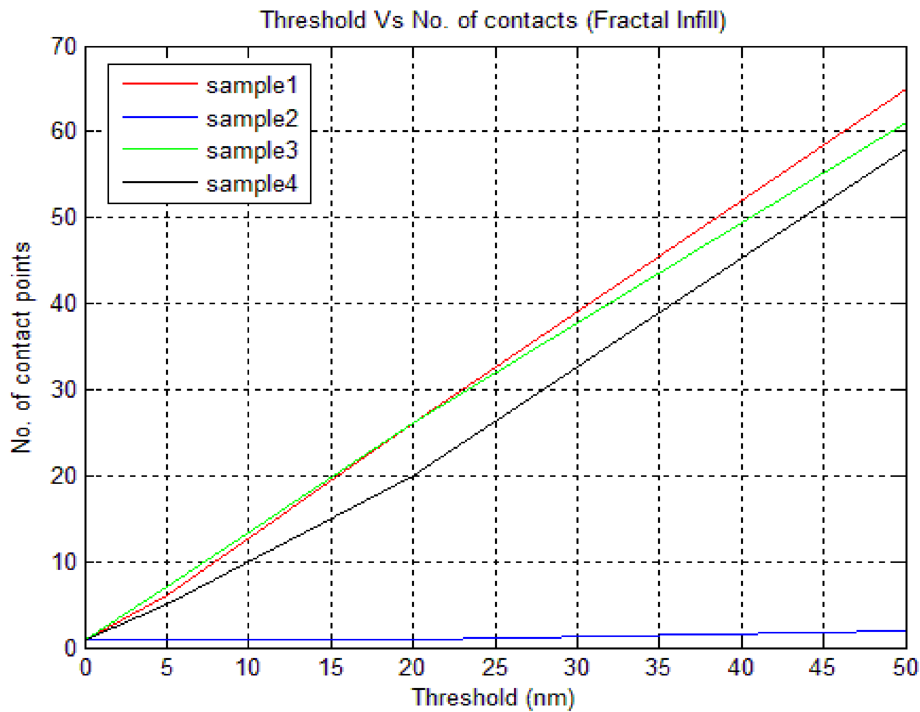


Figure 7.38: Threshold Vs. Number of contacts using spherical stylus

Hence, thresholds from 5 nm to 50 nm (for deliberate over-emphasis) seem relevant for study here. The simplest form of thresholding as explained in earlier chapters is roughly equivalent to ideally pressing the probe into the surface. For the illustrations used here the modified surfaces are largely unchanged other than a slight vertical offset. Interest lies in how the shift interacts with additional surface features to alter the probe contact patterns.

The initial geometric kinematic contact of the original master surface Zorig and the expanded interpolated 'patch' Zexpp are a single point at centre of the ideal spherical stylus tip and this is shown by the data in table 7.5. For all samples, the different thresholds cause the same offset from the stylus centre for each sample. With a threshold, contact points sufficiently close to the tip are considered to be in contact. Both kinematic contact point and the centroid position of the dominant peak island (after threshold) have the same offset from the centre. Obviously because of the ideal conical shape, the contact points occur at the stylus centre.

The initial geometric kinematic contact of the sub-sampled array Z and the fractal 'noise' Zexpp are a single point which is 0.31 μm offset from the centre of the sperical stylus tip as shown by the data in table 7.5.

Table 7.6: Surface roughness parameters

Roughness Parameter	Sa	Sq	Sy	Szk	Sku
Zorig/nm	92.549	126.35	1083.9	0.20741	4.0414
Z/nm	92.109	125.9	987.27	0.071706	4.2633
Zexpp/nm	78.523	105.86	964.7	0.51802	4.2823
Zexpf/nm	78.458	105.95	964.74	0.5003	4.2931
Error_p = Zexpp-Z	-0.1475	-0.16	-0.023	6.22	0.0044
Error_f = Zexpf-Zorig	-0.15	-0.16	-0.11	1.41	0.0623

where, Zorig is a square master array of original data, Z is sub-sampled from the original data Zorig. Zexp is an expanded array of same size of Zorig , Zexpp = patch and Zexpf = patch+ filler.

Table 7.6 shows the surface roughness parameters. It can be noticed that the Surface Skewness Ssk is always < 0 which means it can be a bearing surface with holes. If Ssk = 0, a symmetric height distribution is indicated; while, if Ssk > 0 it can be a flat surface with peaks. Values numerically greater than 1.0 may indicate extreme holes or peaks on the surface.

It always recommended to subtract the mean before calculating the roughness parameters. This is a gentle way to assure that a high mean level does not contribute to the roughness values. However in some cases it may need to preserve the height values such that the bearing (dominating) height value is kept at zero to obtain a better number for the skewness, Ssk parameter.

The Surface Kurtosis, Sku, describes the "peakedness" of the surface topography, the surface height distributions Sku approaches 4.0 when increasing the number of pixels. Smaller values indicate broader height distributions and vice versa for values greater than 4.0.

Many roughness parameters are extremely sensitive to plane distortions and it is therefore important to have the surface image corrected as long wave structures should be reduced by High Pass filtering before calculating the roughness parameters.

Chapter 8

8. Conclusion and further work

This study has introduced a new implementation of a suite of tools for studying stylus-surface interaction in simulation. A consistent pattern of data flows and archiving around a highly modular organization is advocated, allowing new features to be introduced in small, independently testable, modules. In this respect, the first features included and not previously reported concern the mapping of contact extent and intensity across the probe surface and the use of thresholding to investigate sensitivities to noise and instrument resolution. The concepts are illustrated for clarity by a highly artificial flat probe and large parameter values for which many effects can be understood intuitively, as well as being readily visible on small-scale plots. However, even with this unrepresentative approach, the thresholds reveal information about the surface geometry that may be of functional significance.

The system is now being refined to add further features while simultaneously being used with more representative probe models to investigate, for example, the statistical significance of contact distributions.

The simulation program has been developed and used to simulate the real and non-real surfaces' roughness measuring process by the stylus method. It can be used to scan any arbitrary surface with any arbitrary stylus tip shape. The simulation is not only used to measure the roughness of the surface but also to show the contacts distribution on the tip when scanning a surface, information not previously considered. The theoretical results of the simulation have confirmed that the stylus geometry can have a significant effect on most roughness parameters of the

measured surface in 3D. The contacts distribution on the tip has shown that most contacts do not occur at the central point of the stylus in most cases, even with idealized shapes.

This work proposes and explores in detail the novel concept of “thresholding” as an adjunct to kinematic contact modelling; the tip is incremented downwards 'into' the surface and resulting contact regions (or islands) compared to the position of the initial kinematic contact. This work is the first to implement a version of kinematic stylus contact simulation that includes full geometric recording of behaviour at the stylus tip; this will interest other specialists in the field because it is a step towards more complete uncertainty models. The ‘threshold’ concept was also new in this context and was an idea worth practical investigation to see if it has the sensitivity to be useful as a self-diagnostic tool within instruments.

Evidence shows that examination of the contact patterns as the threshold increases can identify the intensity with which different asperity regions interact with the stylus. In the context of sections of the ground surface with total height variation in the order of 500 nm to 1 µm, for example, a 5 nm threshold caused little change in contact sizes from the kinematic point, but 50 nm caused them to grow asymmetrically, eventually picking out the major structures of the surface.

The simulations have naturally confirmed that stylus geometry and size can have a significant effect on most roughness parameters of the measured surface in 3D. The surprising feature of these, worthy of greater investigation, is how insensitive to major changes in stylus condition, some of the popular parameters are, even when dealing with very fine structure within localized areas of a ground surface. For these reasons, it is concluded that thresholding is not likely to become a major tool in

analysis, although it is argued that it retains some practical value as a diagnostic of the measurement process.

Finally, the method works without gross errors, but is clearly approximate and probably generates distortions rather larger than would be wise for a diagnostic - given expected further increase in cost-speed effectiveness of the computers that will go on future instruments, it may be best, on results here, not to further pursue this level of approach but to revisit the general idea in a few years using more sophisticated algorithms.

Considerations for continuing this work could be as follows:

- A large program of practical measurements and complementary simulation should be undertaken in order to provide guidance on a useful value for lateral uncertainty.
- This project was mainly based on using computational software such as MATLAB package for initiating the simulation process because of the time limitation of this work. Most of the tests were aided by software systems. It is significant to test this simulation process by combining these with hardware systems. Therefore real laboratory sessions need to be applied for future works.
- Further study will need to use representative real surface data and study effects over a wider range of roughness analysis tools. This is both to reformat instrument data into a MATLAB compatible file and to provide an industry standard analysis suite that can be considered as a ‘standard’, the simulation should be interfaced to SPIP (Scanning Probe Image Processor)

or a similar commercial package. This possibility was explored within this project and deemed plausible but too time-consuming to adopt immediately.

- Most of the developments and tests on the simulation in this thesis did not consider the materials of stylus tips; MATLAB package is currently not able to process materials selection of stylus tips for simulation. Thus, extending the codes provided in this dissertation to various materials of stylus tips or combining with other programming language need to be considered in future works. This is in order to identify how the contact distribution is affected by choosing different materials of stylus tips for the simulation process.

9. References

Abbott, E. J. & Firestone, F. A. (1933). Specifying Surface Quality: A method based on Accurate measruement and comparsion, *Mechanical Engineering*, 55, 569-573.

Adhikari, R.(2004). *Sensitivity and Noise Analysis of 4 Km Laser Interferometric Gravitational Wave Antennae*. Ph.D, Massachuestts Institute of Technology.

Althagafy, K. T., Chetwynd, DG., (2011). Simulation of stylus contact patterns in profilometry. Styli. Proc. *26th ASPE Annual Meeting Conference*, Denver-Colorado, US, 464-467.

Althagafy, K. T., Chetwynd, DG., (2012). Simulation studies of Sub-micrometer contact of Topography, Styli. Proc.*27th ASPE Annual Meeting*, San Diageo, US, 2012.

Alvarez, L. & Siqueiros, A. J. M. (2010). Scanning Probe Microscopy *Formatex*, 1302-1309.

Balsubramanian, N. (1980). *Optical system for surface topography measurement*. U.S. patent 4,340,306.

Bennett, J. M. & Dancy, J. H. (1981). Stylus profiling instrument for measuring statistical properties of smooth optical surfaces. *Applied Optics*, 20, 1785-1802.

Bennett, J. M. & Mattsson, A. L. (1989). *Introduction to Surface Roughness and Scattering*, Washington, D.C.

Bhushan, B. (1999). *Principles and Applications of Tribology*, John Wiley & Sons, Inc.

Bhushan, B. (2001). *Modern Tribology Handbook*, CRC Press LLC.

Binnig, G. & Rohrer, A. H. (1983). Scanning Tunneling Microscopy. North-Holland, *Surface Science*, 126 236-244.

Buice, E. S. (2007). *Implementation of A Dynamuc Positioning Machine for Nano-scale Engineering*. PhD, North Carolina University.

CABRAL, A. & REBORDÃO, J. (2007). Accuracy of frequency-sweeping interferometry for absolute distance metrology, *Opt. Eng.,* 46, 073602, 1–10

Carre, P. (1966). Installation et utilisation du comparateur photoelectrique et interferentiel du Bureau International des Poids et Mesures. *Metrologia*, 2, 13-23.

Chen, C. J. (2008). *Introduction to Scanning Tunneling Microscopy*, OXFORD UNIVERSITY PRESS.

Chetwynd. D.G.(2001). *Surface measurement and characterization*. In D. E.Mainsah, Metrology and Properties of Engineering Surfaces (pp. 2-11). Dordrecht, Netherlands: Kluwer Academic Publishers.

Church, E. L. & Zavada, J. M. (1975). Residual surface roughness of diamond-turned optics. *Applied Optics*, 14, pp. 1788-1795.

Church, E., & Takacs, P. (1990). Effects of the non-vanishing tip size in mechanical profile measurements. Proc. *SPIE* The international society for optical engineering (pp. 504-514,515).

Conchello, J.-A., Kim, J. J. & Hansen, A. E. W. (1994). Enhanced three-dimensional reconstruction from confocal scanning microscope images. *Applied Optics*, 33, 3740-3750.

Cui, F. K., Song, Z. B., Wang, X. Q., Zhang, F. S. & Li, Y. (2010). Study on laser Triangulation Measurement Principle of Three Dimensional Surface Roughness, *Advanced Materials Research*, 136, 91-94.

Dowidar HAM, Chetwynd DG. (2004) A replication technique for measuring profilometer styli. Proc. ICS 2004, *The XI International Colloquium on Surfaces*, Chemnitz; II 84-92

Dowidar, H.A.M. (2003), *The geometrical interaction of the stylus and the measured surface in 3D roughness measurements*, University of Warwick.

Dowidar, HAM. and Chetwynd, DG. (2002) Distribution of surface contacts on a simulated probe tip. FLM Delbressine et al. (eds) Proc. *3rd International euspen Conference*, Eindhoven; 757-760.

E.Mainsha, D. (2001). *Metrology and Properties of Engineering Surfaces,* Kluwer Academix Publishers. Boston, USA, (pp. 211-213).

Elson, J. M. & Bennett, J. M. (1979). Relation between the angular dependence of scattering and the statistical properties of optical surfaces. *Optical Society of America*, 69, 31-47.

Groover, M. P. (2007). *Fundementals of Modern Manufactureing: Materials, Processes*, and Systems, John Wiley & Sons, Inc.

Hariharan, P., Oreb, B. F. & Eiju, T. (1987). Digital phase-shifting interferometry: a simple error-compensating phase calculation algorithm. *Applied Optics*, 26, 2504-2506.

Jiang, F., Li, J., Yan, L., Sun, J. & Zhang, S. (2010). Optimizing end-milling parameters for surface roughness under different cooling/lubricating conditions. *Inter. Journal of Advanced Manufacturing Tech*., 51, 841-851.

Jiang, X, Scott, P.J, Whitehouse, D.J, Blunt, L. (2007) Paradigm shifts in surface metrology. Part I. Historical philosophy. *Proceedings of the Royal Society A*, 463:2071-2099

Jiang, X., Scott, P., Whitehouse, D. & Blunt, L. (2007a). Paradigm shifts in surface metrology. Part I. Historical philosophy. *Proceedings of the Royal Society A*. 463, 2049-2070.

Jiang, X., Scott, P., Whitehouse, D. & Blunt, L. (2007b). Paradigm shifts in surface metrology. Part II. The current shift. *Proceedings of the Royal Society A*. 463, 2071-2099.

Kafri, O. (1989). Fundamental limit on accuracy in interferometry. *Optics Letters*, 14, 657-658.

Kemao, Q., Fangjun, S. & Xiaoping, W. (2000). Determination of the best phase step of the Carr´e algorithm in phase shifting interferometry. *Measurement Science and Technology*, 11, 1220-1223.

Kikuta, H., Iwata, K. & Nagata, R. (1986). Distance measurement by the wavelength shift of laser diode light. *Applied Optics*, 25, 2976-2980.

Li M., (1991), "*Interpretation and Evaluation of Stylus Profiling Techniques*", PhD Thesis, University of Warwick, Coventry

Longuet-Higgins, M.S. (1957), The statistical analysis of a random, moving surface. *Philosophical Transactions of the Royal Society of London*. Series A, Mathematical and Physical Sciences: p. 321-387.

McCOOL, J. (1984). Assessing the effect of stylus tip radius and flight surface topography measurements. *In Journal of Tribology* (pp. 202-210).

Michelson, A. A. (1882) Interference phenomena in a new form of refractometer, *American Journal of Science*, Series 3 Vol. 23:395-400

Muhamedsalih, Hussam (2013), *Investigation of Wavelength Scanning Interferometry for Embedded Metrology*. Doctoral thesis, University of Huddersfield.

Muralikrishnan B, Raja J. (2009), *Computational Surface and Roundness Metrology*. Springer-Verlag London; Ch 2, 8

Nayak, P.R. (1971), Random process model of rough surfaces. *Journal of Lubrication Technology*. 93: p. 398.

Persson, U. (1998). In-process measurement of surface roughness using light scattering. Elsevier Science, *Scientific American, Wear* 215, 54-58

Petzing, J., Coupland, J. & Leach, R. (2010). *Good Practice Guide No. 116: The Measurement of Rough Surface Topography using Coherence Scanning Interferometry*, NPL.

Q-scope™ (*AFM 788 Manual*)

Radhakrishnan, V. (1971), Analysis of surface profiles by computational methods. *Journal of the Institution of Engineers*. India. 51: p. 272.

Schmalz, G. (1936). Technische Oberflaechenkunde, *Springer*. Germany, (pp66 and 269)

Schreiber, H. & Bruning, J. H. (2007). *Optical Shop Testing*, John Wiley & Sons, Inc.

Schwarz, U. (1994). Tip Artefacts in scanning probe microscopy. *Journal of Microscopy* (pp. 183-197).

Schwider, J., Burow, R., Elssner, K., Grzanna, J., R., S. & Merkel, K. (1983). Digital wave-front measuring interferometry: some systematic error sources. *Applied Optics*, 22, 3421-3432.

Scott P. J. (1992a,) "A New Unified Approach to Surface Texture Instrumentation", *Proceedings of the ASPE Annual Meeting*,

Semwogerere, D. & Weeks, E. R. (2005). *Confocal Microscopy*, Taylor & Francis.

Shan Lou, X. J. (2012). An Introduction to Morphological Filters in Surface University of Huddersfield, EPSRC Centre for Innovative Manufacturing in Advanced Metrology. *Computing and Engineering Researchers' Conference*, University of Huddersfield.

Shunmugam, M. S, Rahadkrishnan V. (1975), "Two- and Three-Dimensional Analysis of Sufaces According to the E-System", *Proc. The Institution of Mechanical Engineers*, 188,691-697

Smith and D.G. Chetwynd (1992), *Foundation of ultraprecision Mechanisms Design*, Gorden and Breach.

Srinivasan, V. (1998). Discrete morphological filters for metrology. *Proceedings 6th ISMQC Symposium on Metrology for Quality Control in Production* (pp. 188,691-699).

Stanford, J. & Hagan, P. (2009). An Assessment of the Impact of Stylus Metallurgy on Cerchar Abrasiveness Index. *Coal Operators' Conference*, 348-355.

Stout, K.J. and Blunt, L. (2000) *Three Dimensional Surface Topography*, Penton Press, London.

Thiel, J., Pfeifer, T. & Hartmann, M. (1995). Interferometric measurement of absolute distances of up to 40 m. *Measurement*, 16, 1-6.

Thirunavukkarasu, (2007). Optical Sensing Limits in Contact and Bending Mode Atomic Force Microscopy. *Experimental Mechanics* 47, 841–844.

Thomas, T. and G. Charlton, (1981) Variation of roughness parameters on some typical manufactured surfaces. *Precision Engineering*. 3(2): p. 91-96.

Thomas, T.R. (1982), *Rough surfaces*, Longman, London, UK.

Thomas, T.R. (1999), *Rough surfaces*: World Scientific.

Tiziani, H. J. & Uhde, A. H.-M. (1994). Three-dimensional image sensing by chromatic confocal microscopy. *Applied Optics*, 33, 1838-1843.

Von Weingraber H., (1956), "Zur definition der Oberflachenrouheit Werk Strattstechnik", *Masch. Bau*, 46,

Vorburger, T.V, Marx, E., Lettieri, T.R. (1993). Regimes of surface roughness measurable with light scattering. *Applied Optics*, 32. pp. 3401-3408

Whitehouse D. J., (1994), *Handbook of Surface Metrology*, Institute of Physics Publishing, Bristol

Whitehouse, D. (1982), The parameter rash—is there a cure? *Wear*, 83(1): p. 75-78.

Whitehouse, D. (1997). *Surface metrology*. In Measurement Science & Technology (pp. 955-972).

Whitehouse, D. (2002) *Surfaces and their measurement*. Hermes Penton Ltd, London

Whitehouse, D. J. (2011). *Handbook of Surface and Nanometrology*, Taylor and Francis Group.

Wilson T, (ed.). (1990), *Confocal microscopy*. Academic Press: London, etc. 426 pp. #33.

Wu, J.J. (1999), Spectral analysis for the effect of stylus tip curvature on measuring rough profiles. *Wear*, 230(2): p. 194-200.

Wyant, J. C. (2002). White Light Interferometry. *Proceedings of SPIE*, 4737, 98-107.

Wyant, J., Koliopoulos, C., Bushan, B. & George, O. (1984). An Optical Profilometer for Surface Characterization of Magnetic Media. *ASLE Transactions*, 27, 101-113.

Yang, H.-J., Nyberg, S. & Riles, K. (2007). High-precision absolute distance measurement using dual-laser frequency scanned interferometry under realistic conditions. *Nuclear Instruments and Methods in Physics Research Section A*, 575, 395-401.

Young, R. D., Vorburger, T. V. & Teague, E. C. (1980). In-Process and On-Line Measurement of Surface Finish. *CIRP Annual,* 29, 435–440.

10. Appendix

10.1. Appendix 1: Simulation software technique (Extra explanation)

The surface map that is to be processed by the stylus is stored in a two dimensional array (or matrix). The studied samples in this thesis had a matrix dimension of either 64×64 or 128×128 depending on the selected resolution. The sample dimensions are represented by L_x and L_y ; for the cases covered in this thesis $L_x = L_y = 20\mu m$. Each cell in the matrix holds the value of a measured point elevation on the studied surface the cell coordinates representing an *x-y* position. These heights are stored in a matrix called A as shown in the following:

$$\mathbf{A} = \begin{bmatrix} a_{(1,1)} & a_{(1,2)} & a_{(1,3)} \\ a_{(2,1)} & a_{(2,2)} & a_{(2,3)} \\ a_{(3,1)} & a_{(3,2)} & a_{(3,3)} \\ \vdots & \vdots & \vdots \end{bmatrix} \qquad (A.1)$$

The number of measured points in the x and y directions are referred to as M and N. Since we use a fixed lateral range, the more points taken in the measurements the higher the resolution is. The discrete distances in the x and y directions will be:

$$dx = \frac{L_x}{M} \quad ; \qquad (A.2)$$

$$dy = \frac{L_y}{M} \qquad (A.3)$$

Knowing that each point has two coordinates, a point can be represented by:

$$P = i \times j \qquad (A.4)$$

Therefore for a sample having a resolution of 128×128 the discrete values of $dx = dy = 0.15625\mu m$, while for a sample having a resolution of 64×64 the discrete values are $dx = dy = 0.3125\mu m$. So as an example for a matrix cell having the coordinates of $i = 16$ and $j = 14$ its numbering notation can be found $P = 16 * 14 = 224$. By providing the i and j matrix cell coordinates for the studied matrix $\mathbf{A}_{i,j}$ the researcher can access the point of interest and based on the following two equations can find the coordinates for the point of interest:

$$x_P = i.dx \qquad (A.5)$$

While in y direction:

$$y_P = j.dy \qquad (A.6)$$

This means from equation (A.5) and (A.6) that

$$P_{ij}(x_P, y_P) = P_{ij}(i.dx, j.dy) \qquad (A.6')$$

Therefore for the previously studied point numbered 224 in a sample having a resolution 128×128 the coordinates of the point of interest on the x axis is:

$$x_P = 16 \times 0.15625 = 2.5\mu m$$

While on the y axis coordinate the point coordinate is:

$$y_P = 14 \times 0.15625 = 2.2\mu m$$

Meaning that point 224 has the coordinates of $P_{224}(2.5, 2.2)$. The next step is to know the height at that point or for any point of interest. This can be achieved by

the user when he enters the i and j cell values for matrix $\mathbf{A}_{i,j}$ this will result in retrieving the value of the measured surface height H_P:

$$a_p(i,j) = H_P \tag{A.7}$$

Meaning for the previously studied point 224 that

$$H_{224} = a_{224}(16,14) = 0.8\ \mu m$$

Based on the fact that the matrix size is big, an idea was adopted which proposed matrix $\mathbf{A}_{i,j}$ to be broken into mini matrices$\mathbf{AA}_{i,j,b,bb}$, these mini matrices represent an area of the studied surface. This method helps the researcher to conduct analysis on individual surface areas. The intended size of the matrix to be broken up to depends on the resolution of the sample and is specified by the researcher in the code input section. As an example the matrix cells in matrix $\mathbf{A}_{i,j}$ shown in red represent the studied region

$$\mathbf{A}_{i,j} = \begin{bmatrix} a_{(1,1)} & a_{(1,2)} & a_{(1,3)} & a_{(1,4)} \\ a_{(2,1)} & a_{(2,2)} & a_{(2,3)} & a_{(2,4)} \\ a_{(3,1)} & a_{(3,2)} & a_{(3,3)} & a_{(3,4)} \\ a_{(4,1)} & a_{(4,2)} & a_{(4,3)} & a_{(4,4)} \\ a_{(5,1)} & a_{(5,2)} & a_{(5,3)} & a_{(5,4)} \end{bmatrix} \tag{A.8}$$

Each broken matrix is represented in the following matrix called $\mathbf{AA}_{i,j,b,bb}$, for the purposes of the representation case here:

$$\mathbf{AA}_{\mathrm{i,j,b,bb}} = \begin{bmatrix} a_{(1,1)} & a_{(1,2)} & a_{(1,3)} \\ a_{(2,1)} & a_{(2,2)} & a_{(2,3)} \\ a_{(3,1)} & a_{(3,2)} & a_{(3,3)} \end{bmatrix} \quad \text{(A.9)}$$

The broken up matrix has two new additional cells representing the two dimensional coordinate system of each broken up area of the studied surface, while the first two cells represent the cell numbers in the broken up matrices.

So as an example if the user wants to study a certain area of the studied surface having $b = 14$ and $bb = 15$ the following matrix would be extracted from the main one called A:

$$\mathbf{AA}(:,:,14,15) = \begin{bmatrix} 0.832 & 0.354 & 0.654 \\ 0.338 & 0.563 & 0.238 \\ 0.344 & 0.298 & 0.452 \end{bmatrix}$$

If the researcher wants to go into more detail and extract the data at a specific point then by specifying $i = 2\ \ j = 3$ the obtained height at that point is

$$\mathbf{AA}(2,3,14,15) = 0.298\ \mu\text{m}$$

There are two mapping matrices in the code, these are represented by quad for loops the first two for loops are responsible for breaking up the main matrix into small area regions in the XY plane while the second two for loops are responsible for storing the broken matrices into the new matrix $\mathbf{X}_{\mathrm{i,j,b,bb}}$ (X is a general term given to any broken matrix).

The Stylus matrix is represented by $\mathbf{C}_{\mathrm{i,j}}$ this matrix is read into MATLAB for analysis:

$$\mathbf{C1}_{\mathrm{i,j}} = \begin{bmatrix} C1_{(1,1)} & C1_{(1,2)} & C1_{(1,3)} & C1_{(1,4)} \\ C1_{(2,1)} & C1_{(2,2)} & C1_{(2,3)} & C1_{(2,4)} \\ C1_{(3,1)} & C1_{(3,2)} & C1_{(3,3)} & C1_{(3,4)} \\ C1_{(4,1)} & C1_{(4,2)} & C1_{(4,3)} & C1_{(4,4)} \\ C1_{(5,1)} & C1_{(5,2)} & C1_{(5,3)} & C1_{(5,4)} \end{bmatrix} \quad \text{(A.10)}$$

The power of such a matrix is that it provides the researcher with modelling tools, meaning that by modelling different stylus configurations new types of stylus designs can be introduced into the field of research. Therefore when the required stylus accuracy is not achieved design modifications are undertaken. This modelling process is just the first step, once the required accuracy is obtained by the numerical tests then the researcher can proceed to the second step of producing the stylus in the workshops to conduct lab experiments at a later stage.

The second part of the code breaks matrix $\mathbf{C1}_{\mathrm{i,j}}$ into specific surface regions then assigns the area to a matrix called $\mathbf{A1}_{\mathrm{i,j,b,bb}}$ for analysis; this matrix represents the segmented area of the stylus surface.

$$\mathbf{A1}_{\mathrm{i,j,b,bb}} = \begin{bmatrix} a1_{(1,1)} & a1_{(1,2)} & a1_{(1,3)} \\ a1_{(2,1)} & a1_{(2,2)} & a1_{(2,3)} \\ a1_{(3,1)} & a1_{(3,2)} & a1_{(3,3)} \end{bmatrix} \quad \text{(A.11)}$$

An intermediate matrix tool named $\mathbf{pr}_{\mathrm{i,j,b,bb}}$ is used in the code to handle data processing for the measured surface segmented regions where $\mathbf{pr}_{\mathrm{i,j,b,bb}} = \mathbf{A1}_{\mathrm{i,j,b,bb}}$, therefore the intermediate matrix can be written as follows:

$$\mathbf{pr}_{\mathrm{i,j,b,bb}} = \begin{bmatrix} \mathrm{Pr}_{(1,1)} & \mathrm{Pr}_{(1,2)} & \mathrm{Pr}_{(1,3)} \\ \mathrm{Pr}_{(2,1)} & \mathrm{Pr}_{(2,2)} & \mathrm{Pr}_{(2,3)} \\ \mathrm{Pr}_{(3,1)} & \mathrm{Pr}_{(3,2)} & \mathrm{Pr}_{(3,3)} \end{bmatrix} \tag{A.12}$$

In previous steps it was stated that:

$$\mathbf{C1}_{\mathrm{i,j,b,bb}} = \mathbf{A1}_{\mathrm{i,j,b,bb}} = \begin{bmatrix} \mathrm{a1}_{(1,1)} & \mathrm{a1}_{(1,2)} & \mathrm{a1}_{(1,3)} \\ \mathrm{a1}_{(2,1)} & \mathrm{a1}_{(2,2)} & \mathrm{a1}_{(2,3)} \\ \mathrm{a1}_{(3,1)} & \mathrm{a1}_{(3,2)} & \mathrm{a1}_{(3,3)} \end{bmatrix} \tag{A.13}$$

Matrix $\mathbf{atopr}_{\mathrm{i,j,b,bb}}$ results from the algebraic addition process of the two broken up surface segments for the stylus and the measured surface.

$$\mathbf{atopr}_{\mathrm{i,j,b,bb}} = \mathbf{C1}_{\mathrm{i,j,b,bb}} + \mathbf{pr}_{\mathrm{i,j,b,bb}} \tag{A.14}$$

Hence, the resulting matrix in (1.14) has come from the addition of the stylus and measured surface heights at each point.

The next essential step is the creation of matrix $\mathbf{M}_{\mathrm{b,bb}}$ its role is to take reference point height, this point is located in the corner of the matrix area at $\mathbf{atopr}_{1,1,\mathrm{b,bb}}$ (as default).

$$\mathbf{M}_{\mathrm{b,bb}} = \mathbf{atopr}_{1,1,\mathrm{b,bb}} = \mathrm{H}_{\mathrm{ref}} \tag{A.15}$$

Relying on the conditional statement during the matrix mapping process for the segmented regions if this condition is satisfied then the point of the first contact is found:

$$\mathbf{atopr}_{\mathrm{i,j,b,bb}} > \mathbf{M}_{\mathrm{b,bb}} \tag{A.16}$$

Then at that point matrix $\mathbf{M}_{b,bb}$ can be equated with matrix $\mathbf{atopr}_{i,j,b,bb}$:

$$\mathbf{M}_{b,bb} = \mathbf{atopr}_{i,j,b,bb} \tag{A.17}$$

From this conditional statement the first contact point is found where its coordinates are located in the storage matrix for the x axis direction:

$$i = i_{max} \tag{A.18}$$

While for the y axis direction it is:

$$j = j_{max} \tag{A.19}$$

So for example if the conditional statement was satisfied during the two dimensional mapping process of the matrix for a point having $i_{max} = 45$ and $j_{max} = 67$ at area segment having the coordinates $b = 13$ and $bb = 19$ then it can be written that:

$$\mathbf{firstcontact}_{i_{max},j_{max},b,bb} = \mathbf{firstcontact}_{45,67,13,19} = 234\ \text{nm}$$

Visualization matrix, where the first contact point is assigned a value of 1 while the cells that do not satisfy the conditional statement are assigned a value of 0:

$$\mathbf{firstcontact}_{i,j,b,bb} = \begin{bmatrix} 1 & 0 & 0 & .. \\ 1 & 0 & 0 & .. \\ 0 & 0 & 0 & :: \\ .. & .. & .. & .. \\ .. & .. & .. & .. \end{bmatrix} \tag{A.20}$$

The sample mapping process requires the use of the full contact matrix repeatedly for each segmented area therefore resetting its cell values to zero is a requirement:

$$\mathbf{fullcontact}_{i,j,b,bb} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{4.21}$$

An example of a full contact matrix with assigned values showing points, the red part of the matrix shows the segmented region of study:

$$\mathbf{fullcontact}_{\mathrm{i,j,b,bb}} = \mathbf{atopr}_{\mathrm{i,j,b,bb}} = \begin{bmatrix} 112 & 453 & 209 & 211 \\ 231 & 213 & 214 & 209 \\ 345 & 347 & 345 & 315 \\ 234 & 312 & 301 & 313 \\ 231 & 345 & 305 & 301 \end{bmatrix}$$

Then through the following algebraic operation with the use the of the previously found matrices $\mathbf{atopr}_{\mathrm{i,j,b,bb}}$ and $\mathbf{M}_{\mathrm{i,j}}$**:**

$$\mathbf{fullcontact}_{\mathrm{i,j,b,bb}} = \mathbf{atopr}_{\mathrm{i,j,b,bb}} \geq \mathbf{M}_{\mathrm{i,j}} - \mathbf{atopr}_{\mathrm{i,j,b,bb}} \tag{A.22}$$

$$\mathbf{thr}_{\mathrm{i,j}} = \mathrm{constant} \tag{A.23}$$

In previous steps the matrices $\mathbf{atopr}_{\mathrm{i,j}}$, $\mathbf{M}_{\mathrm{i,j}}$ and $\mathbf{thr}_{\mathrm{i,j}}$ had been found**.** The gap is calculated through the following algebraic operation for all the surface segmented regions:

$$\mathbf{gap}_{\mathbf{i,j,b,bb}} = \mathbf{atopr}_{\mathrm{i,j}} - (\mathbf{M}_{\mathrm{i,j}} + \mathbf{thr}_{\mathrm{i,j}}) \tag{A.24}$$

Finally the gap matrix for each area segment is calculated:

$$\mathbf{gap}_{\mathbf{i,j,b,bb}} = \begin{bmatrix} \mathrm{gap}_{(1,1)} & \mathrm{gap}_{(1,2)} & \mathrm{gap}_{(1,3)} \\ \mathrm{gap}_{(2,1)} & \mathrm{gap}_{(2,2)} & \mathrm{gap}_{(2,3)} \\ \mathrm{gap}_{(3,1} & \mathrm{gap}_{(3,2)} & \mathrm{gap}_{(3,3)} \end{bmatrix} \tag{A.25}$$

Table A-1 Code variable description.

Variable name	Description
Number_of_matricies_x	Broken Matrix number on the x axis
Number_of_matricies_y	Broken Matrix number on the y axis
Number_of_cells_in_x	Range of cells in the x direction
Number_of_cells_in_y	Range of cells in the y direction
M	X loop number
N	Y loop number
i	X loop variable
j	Y loop variable

10.2. Appendix 2: Extra Results for Chapter 7

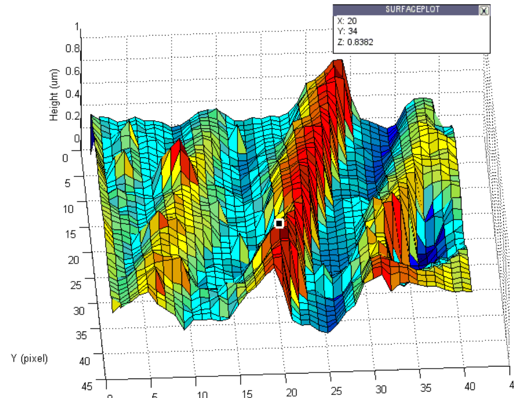


Figure A.1: 3D Ground surface with Flat stylus (Sample 1)

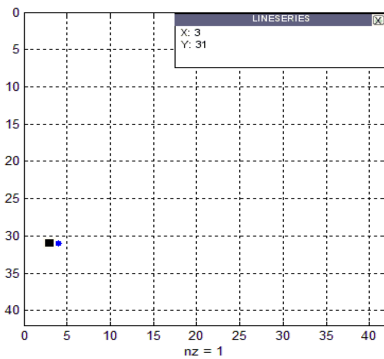


Figure A.2: Map of contact with 5 nm threshold

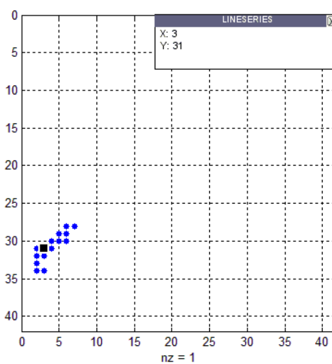


Figure A.3: Map of contact with 20 nm threshold

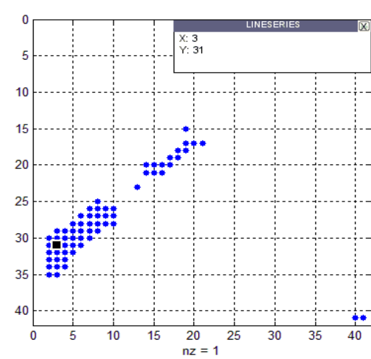


Figure A.4: Map of contact with 50 nm threshold

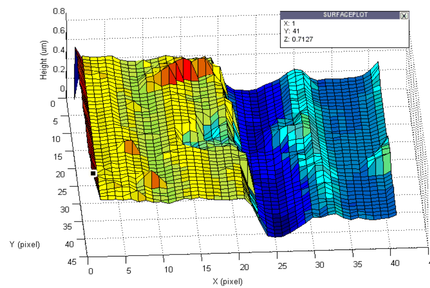


Figure A.5: 3D Ground surface with Flat stylus (Sample 2)

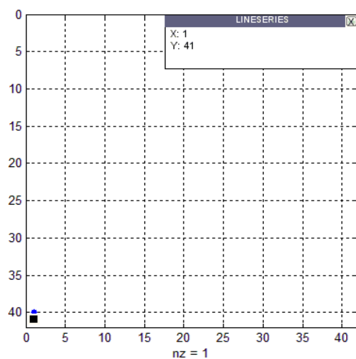


Figure A.6: Map of contact with 5 nm threshold

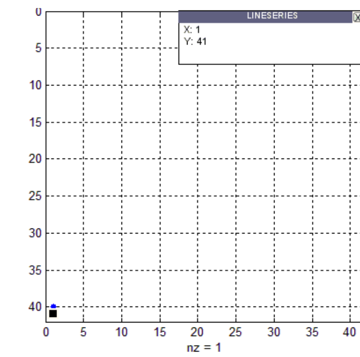


Figure A.7: Map of contact with 20 nm threshold

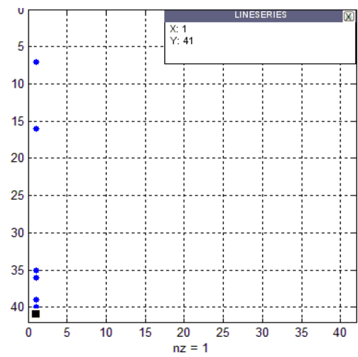


Figure A8: Map of contact with 50 nm threshold

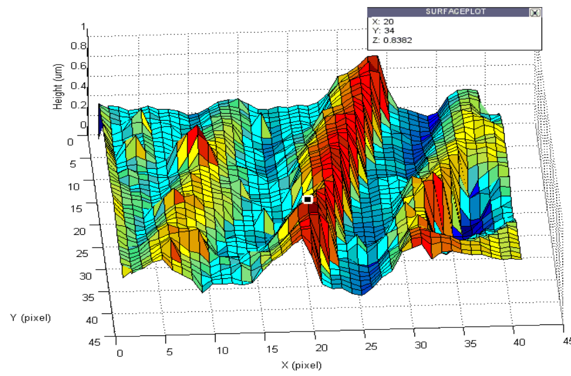


Figure A.9: 3D Ground surface with Flat stylus (Sample 3)

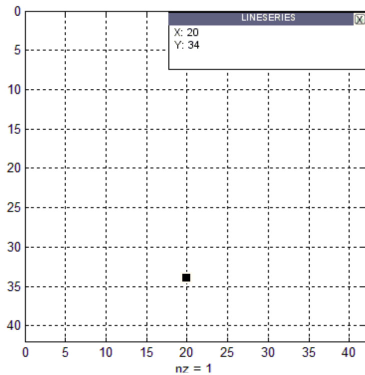


Figure A.10: Map of contact with 5 nm threshold (Sample 3)

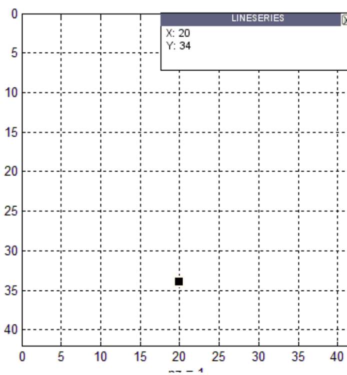


Figure A.11: Map of contact with 20 nm threshold (Sample

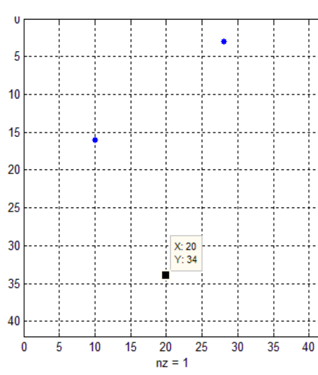


Figure A.12: Map of contact with 50 nm threshold (Sample 3)

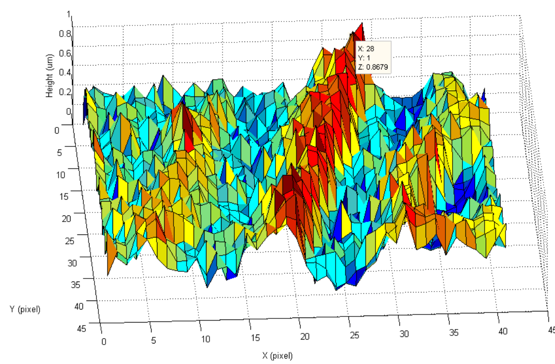


Figure A.13: 3D Ground surface with Flat stylus (Sample 4- Noise)

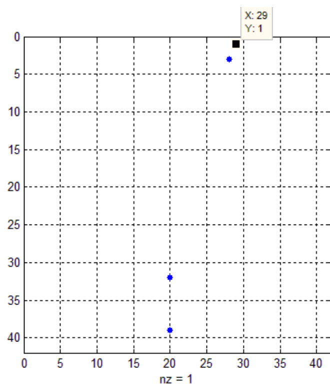


Figure A.14: Map of contact with 5nm threshold (Sample 4)

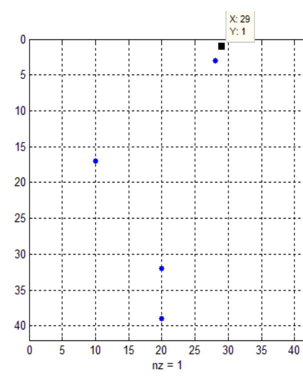


Figure A.15: Map of contact with 20 nm threshold (Sample 4)

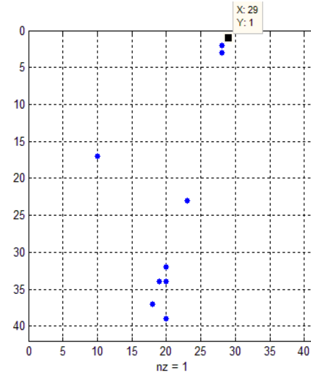


Figure A.16: Map of contact with 50 nm threshold (S 4)

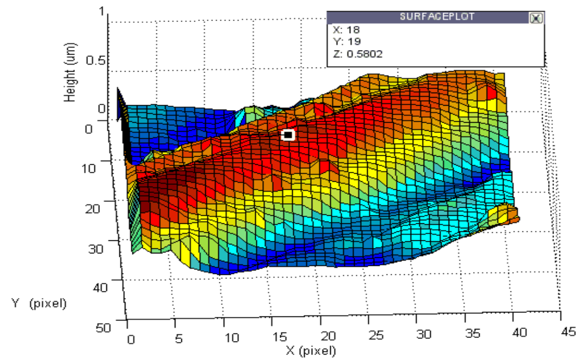


Figure A.17: 3D Ground surface with Sphere stylus (Sample 1)

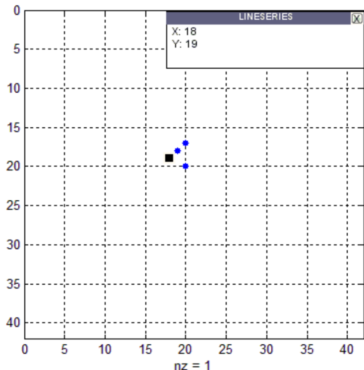


Figure A.18: Map of contact with 5nm threshold (Sample 1)

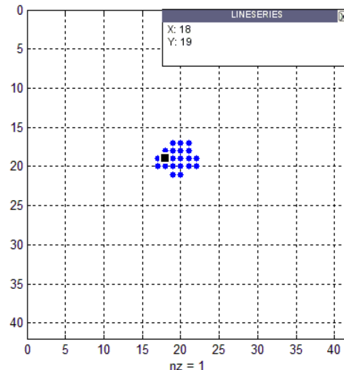


Figure A.19: Map of contact with 20nm threshold (Sample 1)

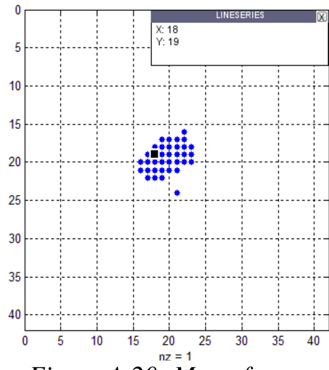


Figure A.20: Map of contact with 50nm threshold (Sample 1)

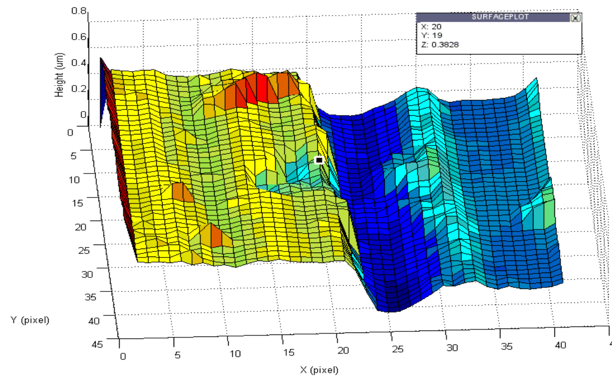


Figure A.21: 3D Ground surface with Sphere stylus (Sample 2)

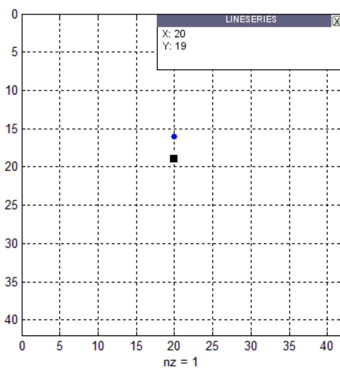


Figure A.22: Map of contact with 5nm threshold (Sample 2)

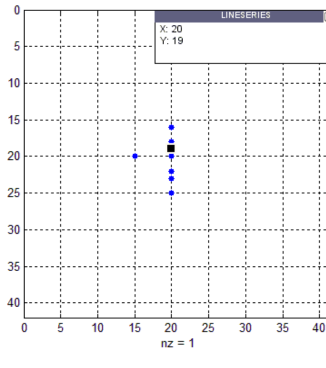


Figure A.23: Map of contact with 20nm threshold (Sample 2)

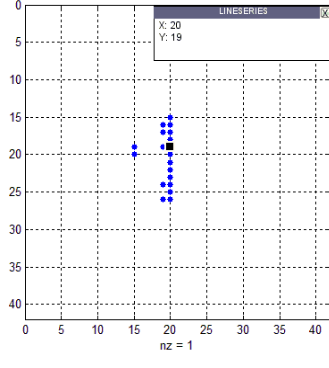


Figure A.24: Map of contact with 50nm threshold (Sample 2)

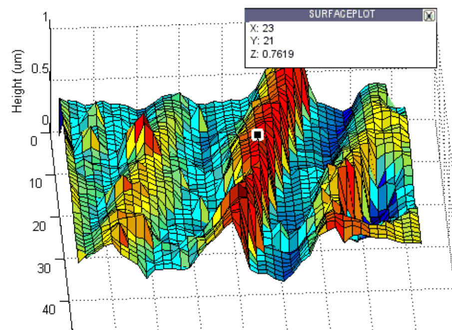


Figure A.25:3D Ground surface with Sphere stylus (Sample 3)

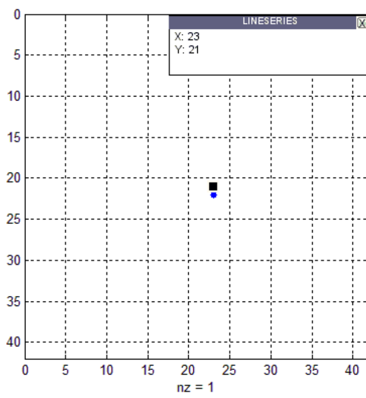


Figure A.26: Map of contact with 5nm threshold (Sample 3)

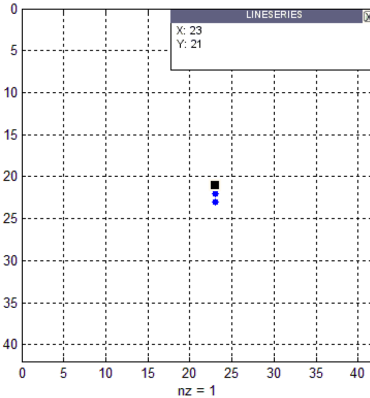


Figure A.27: Map of contact with 20nm threshold (Sample 3)

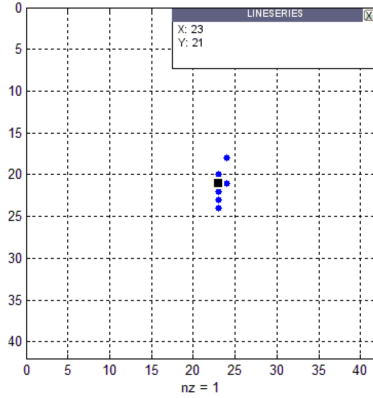


Figure A.28: Map of contact with 50nm threshold (Sample 3)

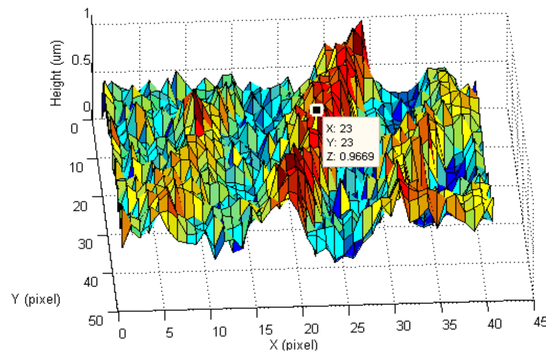


Figure A.29: 3D Ground surface with Sphere stylus (Sample 4 - Noise)

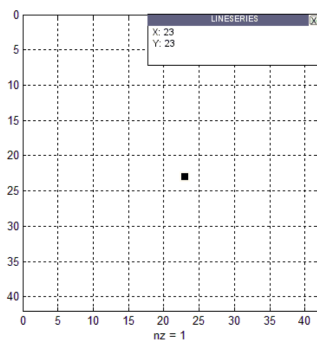


Figure A.30: Map of contact with 5nm threshold (Sample 4)

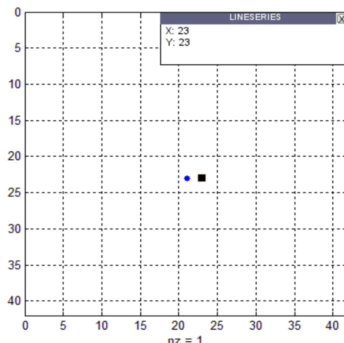


Figure A.31: Map of contact with 20nm threshold (Sample 4)

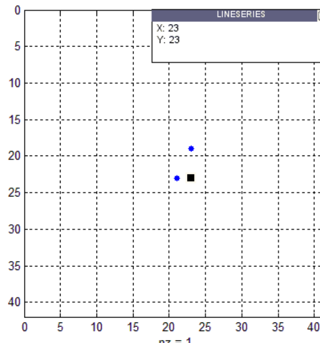


Figure A.32: Map of contact with 50nm threshold (Sample 4)

Stylus: Cone

Threshold: 0.005, 0.05 µm, and 0.075 µm respectively.

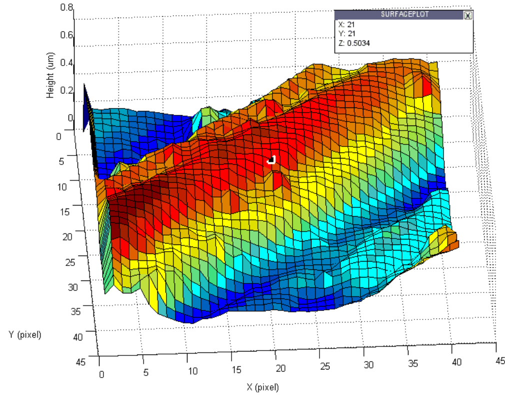


Figure A.33: Ground surface with Cone stylus (Sample 1)

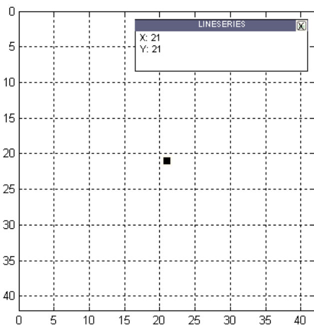


Figure A.34: Map of contact with 5nm threshold (Sample 1)

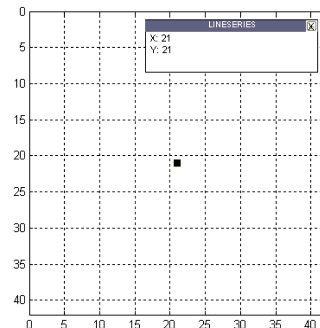


Figure A.35: Map of contact with 5nm threshold (Sample 1)

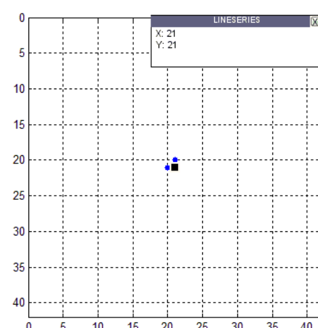


Figure A.36: Map of contact with 5nm threshold (Sample

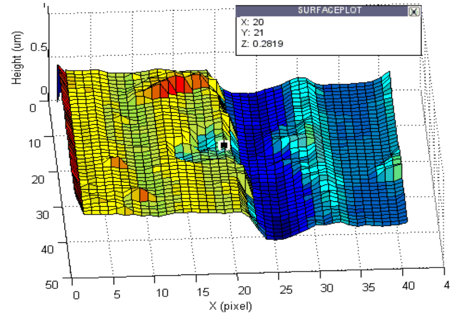


Figure A.37: Ground surface with Cone stylus (Sample 2)

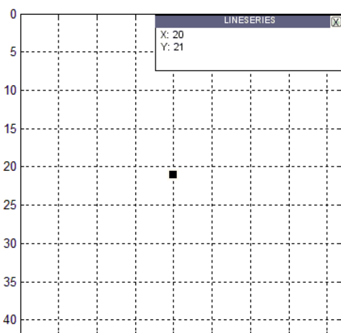


Figure A.38: Map of contact with 5nm threshold (Sample 2)

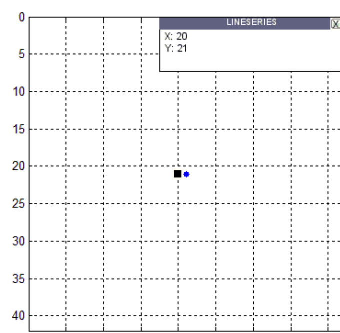


Figure A.39: Map of contact with 20nm threshold (Sample 2)

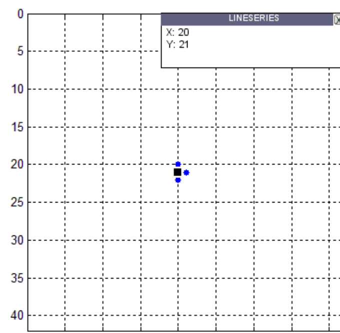


Figure A.40: Map of contact with 50nm threshold (Sample 2)

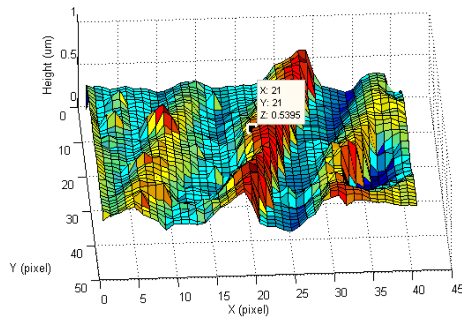


Figure A.41: Ground surface with Cone stylus (Sample

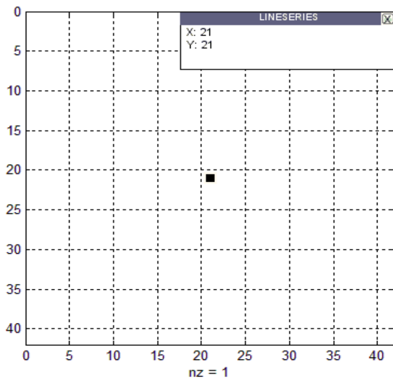


Figure A.42: Map of contact with 5nm threshold (Sample 3)

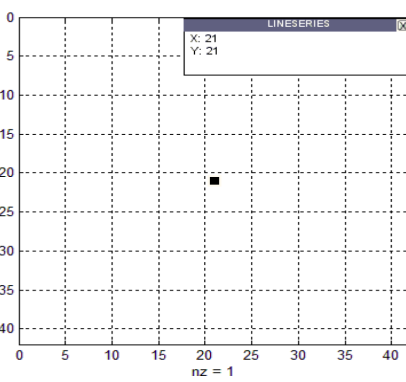


Figure A.43: Map of contact with 20nm threshold (Sample 3)

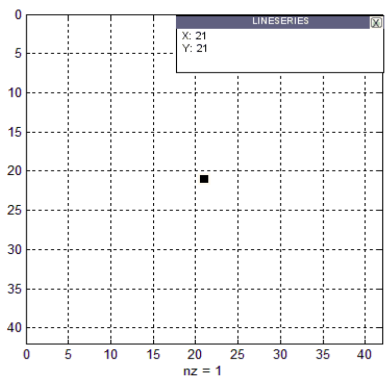


Figure A.44: Map of contact with 50nm threshold (Sample 3)

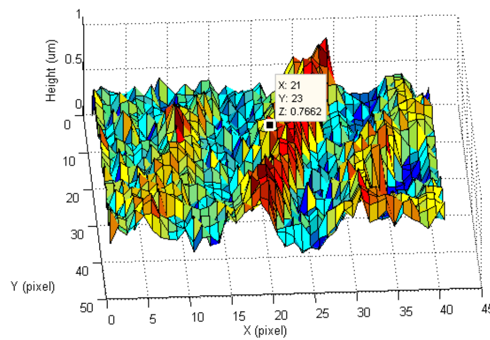


Figure A.45: 3D Ground surface with cone stylus (Sample 4 - Noise))

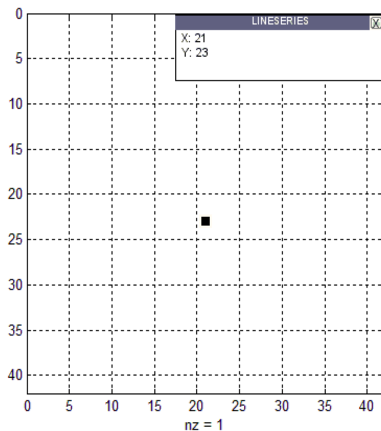


Figure A.46: Map of contact with 5nm threshold (Sample 4)

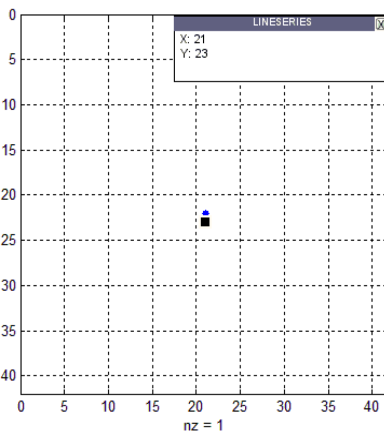


Figure A.47: Map of contact with 20nm threshold (Sample 4)

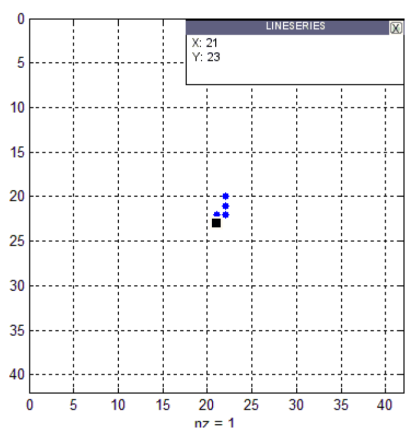


Figure A.48: Map of contact with 50nm threshold (Sample 4)

10.3. Appendix 3: Matlab Codes

A3.1 The main Matlab Code:

```
clc
clear

tic
%Constants Input Section
SampleDimensions
%Lab Data Input
DataLoadSection
%Lab Data Assigned a name by user
%Assigning_names_to_Matricies
%Breaking Main Matrix to Parts
%Main_Matrix_Break_Up
% Calculating Average of Each Single Matrix Created
%Minin_Matrix_Averaging
%DATA ANALYSIS
%Assinging_the_values_into_a_new_matrix_for_visulization
%Finding Surface Conatct Matrix
%Finding_Surface_Conatct_Matrix
%Viewing Surface Contact
%Viewing_Surface_Contact
%Indvidual Matrix Analysis for full and first contact
%shipwreck
%Whole Matrix first and full contacts
WholeMatrix
whos
%To Write Later
%Mesh_generation_for_Total_Matrix
%Mini Matrix Mesh
%Mesh_generation_for_Mini_Matrix
%Centroid_of_Contact
toc


% mid Matrix cell number in x direction'
Number_of_matricies_x =M/(alpha+1);
% mid Matrix cell number in y direction'
Number_of_matricies_y=N/(alpha+1);
% 'The Total Number of Matricies in x direction'
Number_of_matricies_x;
% 'The Total Number of Matricies in y direction'
Number_of_matricies_y;

% Matrix cell number in x direction'
Number_of_cells_in_x=M/Number_of_matricies_x;
% Matrix cell number in y direction'
Number_of_cells_in_y=N/Number_of_matricies_y;

%Matrix Break Up to Mini Matrcies
%b and bb refers to the coordinate of the small matrix
    b=0;
for ii=1:Number_of_matricies_x;
    b=b+1;
```

```
    bb=0;
    for jj=1:Number_of_matricies_y;
        bb=bb+1;

        xx=b;
        for iii=1:Number_of_cells_in_x;
             xx=xx+1;
             yy=bb;
        for jjj=1:Number_of_cells_in_y;
            yy=yy+1;
        aa(iii,jjj,b,bb )=a(xx,yy);
        end
        end



    end
end

 b=0;
for ii=1:Number_of_matricies_x;
    b=b+1;
    bb=0;
    for jj=1:Number_of_matricies_y;
        bb=bb+1;

        xx=b;
        for iii=1:Number_of_cells_in_x;
             xx=xx+1;
             yy=bb;
        for jjj=1:Number_of_cells_in_y;
            yy=yy+1;
        aa1(iii,jjj,b,bb )=a1(xx,yy);
        end
        end



    end
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%Minin_Matrix_Averaging
   b=0;
for ii=1:Number_of_matricies_x;
    b=b+1;
    bb=0;
    for jj=1:Number_of_matricies_y;
        bb=bb+1;
        c=size(aa(:,:,b,bb));
        cc=size(aa(:,:,b,bb));
        aaa(:,:,b,bb )=sum(sum(aa(:,:,b,bb )))/(c(1)*cc(2));
    end
end

%Minin Matrix Averaging 1
```

```
   b=0;
for ii=1:Number_of_matricies_x;
    b=b+1;
    bb=0;
    for jj=1:Number_of_matricies_y;
        bb=bb+1;
        c1=size(aa1(:,:,b,bb));
        cc1=size(aa1(:,:,b,bb));
        aaa1(:,:,b,bb )=sum(sum(aa1(:,:,b,bb )))/(c1(1)*cc1(2));
    end
end
```

%%%%%%%%%%%%%%%%%%%%

% finding contact points

```
clc
b=0;
for ii=1:Number_of_matricies_x;
    b=b+1;
    bb=0;
    for jj=1:Number_of_matricies_y;
         bb=bb+1;
         aa(:,:,b,bb);
         c1=ans(:,:);
         thr(:,:,b,bb)=threshold;


%pr=zeros(10,10);

pr=aa1(:,:,b,bb);


 pr(:,:,b,bb)=aa1(:,:,b,bb);
 atopr(1:10,1:10,b,bb)=c1(:,:)+pr(:,:,b,bb);
%interaction function
%%
M(b,bb)=atopr(1,1,b,bb);
for i=istart:iend;
    for j=jstart:jend;
        if atopr(i,j,b,bb)>M(b,bb)
            M(b,bb) =atopr(i,j,b,bb);           %Holding max vdm
            imax=i;           % index of the current max
            jmax=j;

        end
    end
end
firstcontact(1:10,1:10,b,bb)= zeros(size(atopr(i,j,b,bb)));
firstcontact(imax,jmax,b,bb)=1;          % Map of tip, all zero
except simple contact point
fullcontact(1:10,1:10,b,bb)=zeros(size(atopr(1:10,1:10,b,bb)));
fullcontact(1:10,1:10,b,bb)=atopr(1:10,1:10,b,bb)>=M(b,bb)-
thr(:,:,b,bb);
```

```
gap(:,:,b,bb)=atopr(1:10,1:10,b,bb)-(M(b,bb)-thr(:,:,b,bb));
    end
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%




clc
clear all
close all
format short;


traces=64;                                % Number of scanned traces
n=64;                                      %number of pints in each
trace

%Surface in three Dimensional

%Creating sinewave pattern surface

% angle=0:pi/30:10*pi;   % for 300X300  %angle=0:pi/6.3:10*pi; for
64X64
% z= sin(angle);
% for i=1 : traces-1
% z=[z; sin(angle)];
% end

%Creating randum surface
for surface=1:3;


%Surface in three Dimensional

%Creating sinewave pattern surface
if surface==1 ;
angle=0:pi/6.3:10*pi;
z= sin(angle);

 for i=1 : traces-1
z=[z; sin(angle)];
zsinoriginal=z;
 end
end
.txt;

   if surface==2 %sample2(l)
 z=randn(64);
 z=z/(max(max(z)));

 zrandoriginal=z;
   end
```

```
   if surface==3 %sample3(m)
      z=xlsread('KAL90ROV_AFM.xlsx');
      z=z';
      %%%z=m;
      p1=1.053;
      rt3=p1;
      zrealoriginal=z;

   end

   p1=1.099;                                       % Maximum peak to valley
of the surface (to make the height the same as the AFM)
   r1=5;                                      %Tip base 2*r1
   h=5;                                 %Complete tip height

    a=size(z);
    traces=a(1,1);
    n=a(1,2);
    z=z-min(min(z));
    z=z/(max(max(z)));
    z=(z*p1);

    v=z;
   % Inputs


   for tipshape=1:3;
      f=0;

      % Creating Pyramid Tip shape

      if tipshape==1
         b=r1;
         stylus=ones(2*r1+1);
         for x=0 : h;
            [stylus(r1-b+1:r1+b+1,r1-b+1:r1+b+1)]=x;
            b=b-1;
         end
      end
         % Creating Hemisphere Tip shape

      if tipshape==2


         stylus=zeros(2*r1+1);
         for y=-r1:r1,
            for x=-r1:r1,
               xbar=x+r1+1;
               ybar=y+r1+1;
               factor=r1^2-x^2-y^2;
               if factor<0
                  stylus(xbar,ybar)=stylus(xbar,ybar);
               else
                  stylus(xbar,ybar)=sqrt(factor);
               end
            end
         end
```

```
%end
%end

      % Creating Conical Tip shape
   else


      b=r1;
      stylus=zeros(2*r1+1);
      for x=0 : h;
         for y=-b:b,
            for k=-b:b,
               xbar=y+b;
               ybar=k+b;
               factor=b^2-y^2-k^2;
               if factor>0
                  stylus(xbar+x+1,ybar+x+1)=x;
               end
            end
         end
         b=b-1;
      end

   end


   g=size(stylus);

   for q=1:3,
      z=v;
      for u=1:g(1,1),
         for k=1:g(1,2),
            if stylus(u,k)>h-f;
               stylus(u,k)=h-f;
            end
         end
      end

      %Calculations the locus

      contact=zeros(2*r1+1);
      for i=r1+1:traces-r1,
         for j=r1+1:n-r1,
            xy=z(i-r1:i+r1, j-r1:j+r1);
            position=stylus+xy;
            maxi=max(max(position));
            loc(i-r1,j-r1)=maxi;
            for k=-r1:r1,
               for m=-r1:r1,
                  kbar=k+r1+1;
                  mbar=m+r1+1;
                  if position(kbar,mbar) < maxi
                     contact(kbar,mbar)=contact(kbar,mbar);
                  else
                     contact(kbar,mbar)=1+contact(kbar,mbar);
                  end
               end
```

```
      end
   end
end
%Calculation of roughness parameters


z=z(r1+1:traces-r1, r1+1:n-r1);
z=z-(min(min(z)));
raa=sum(sum(abs(z)))/((n-(2*r1+1))*(traces-(2*r1+1)));
loc=loc-(min(min(loc)));
ra=sum(sum(abs(loc)))/((n-(2*r1+1))*(traces-(2*r1+1)));

if surface==1,
   if tipshape==1,
      if f==0,
         locgp1=loc;
         contactgp1=contact;
         stylusgp1=stylus;
         raag=raa;
         ragp1=ra;
         zsingp1=z;
      end
      if f==1,
         locgp2=loc;
         contactgp2=contact;
         stylusgp2=stylus;
         ragp2=ra;
         zsingp2=z;
      end
      if f==2,
         locgp3=loc;
         contactgp3=contact;
         stylusgp3=stylus;
         ragp3=ra;
         zsingp3=z;
      end

   end



   if tipshape==2,
      if f==0,
         locgh1=loc;
         contactgh1=contact;
         stylusgh1=stylus;
         ragh1=ra;
         zsingh1=z;
      end
      if f==1,
         locgh2=loc;
         contactgh2=contact;
         stylusgh2=stylus;
         ragh2=ra;
         zsingh2=z;
      end
      if f==2,
         locgh3=loc;
         contactgh3=contact;
```

```
            stylusgh3=stylus;
            ragh3=ra;
            zsingh3=z;
         end

      end

      if tipshape==3,
         if f==0,
            locgc1=loc;
            contactgc1=contact;
            stylusgc1=stylus;
            ragc1=ra;
            zsingc1=z;
         end
         if f==1;
            locgc2=loc;
            contactgc2=contact;
            stylusgc2=stylus;
            ragc2=ra;
            zsingc2=z;
         end
         if f==2;
            locgc3=loc;
            contactgc3=contact;
            stylusgc3=stylus;
            ragc3=ra;
            zsingc3=z;
         end

      end
   end
   if surface==2;
      if tipshape==1;
         if f==0;
            loclp1=loc;
            contactlp1=contact;
            styluslp1=stylus;
            raal=raa;
            ralp1=ra;
            zrand=z;
            zrandlp1=z;
         end
         if f==1;
            loclp2=loc;
            contactlp2=contact;
            styluslp2=stylus;
            ralp2=ra;
            zrandlp2=z;
         end
         if f==2;
            loclp3=loc;
            contactlp3=contact;
            styluslp3=stylus;
            ralp3=ra;
            zrandlp3=z;
         end
```

```
    end


    if tipshape==2;
        if f==0;
            loclh1=loc;
            contactlh1=contact;
            styluslh1=stylus;
            ralh1=ra;
            zrandlh1=z;
        end
        if f==1;
            loclh2=loc;
            contactlh2=contact;
            styluslh2=stylus;
            ralh2=ra;
            zrandlh2=z;
        end
        if f==2;
            loclh3=loc;
            contactlh3=contact;
            styluslh3=stylus;
            ralh3=ra;
            zrandlh3=z;
        end

    end

    if tipshape==3;
        if f==0;
            loclc1=loc;
            contactlc1=contact;
            styluslc1=stylus;
            ralc1=ra;
            zrandlc1=z;
        end
        if f==1;
            loclc2=loc;
            contactlc2=contact;
            styluslc2=stylus;
            ralc2=ra;
            zrandlc2=z;
        end
        if f==2;
            loclc3=loc;
            contactlc3=contact;
            styluslc3=stylus;
            ralc3=ra;
            zrandlc3=z;
        end

    end
end
if surface==3;
    if tipshape==1;
        if f==0;
            locmp1=loc;
            contactmp1=contact;
```

```
        stylusmp1=stylus;
        raam=raa;
        ramp1=ra;
        zreal=z;
        zrealmp1=z;
    end
    if f==1;
        locmp2=loc;
        contactmp2=contact;
        stylusmp2=stylus;
        ramp2=ra;
        zrealmp2=z;
    end
    if f==2;
        locmp3=loc;
        contactmp3=contact;
        stylusmp3=stylus;
        ramp3=ra;
        zrealmp3=z;
    end

end


if tipshape==2;
    if f==0;
        locmh1=loc;
        contactmh1=contact;
        stylusmh1=stylus;
        ramh1=ra;
        zrealmh1=z;
    end
    if f==1;
        locmh2=loc;
        contactmh2=contact;
        stylusmh2=stylus;
        ramh2=ra;
        zrealmh2=z;
    end
    if f==2;
        locmh3=loc;
        contactmh3=contact;
        stylusmh3=stylus;
        ramh3=ra;
        zrealmh3=z;
    end

end

if tipshape==3;
    if f==0;
        locmc1=loc;
        contactmc1=contact;
        stylusmc1=stylus;
        ramc1=ra;
        zrealmc1=z;
    end
    if f==1,
```

```
                    locmc2=loc;
                    contactmc2=contact;
                    stylusmc2=stylus;
                    ramc2=ra;
                    zrealmc2=z;
                 end
                 if f==2,
                    locmc3=loc;
                    contactmc3=contact;
                    stylusmc3=stylus;
                    ramc3=ra;
                    zrealmc3=z;
                 end
              end
           end
           f=f+1;
        end
           end
     end

     save khalid

   %plotting data

  % surf(z)
  % hold on
  % surf(loc+p1)

  % shading interp
  % colormap copper
  %pppppppppppppppppppppppppppppppppppppp

   surf(zsingp1);hold on; surf(locgp1+1);
  xlabel('X (Pxiel)');
  ylabel('Y(Pxiel)');
  zlabel('Height(µm)');
  title('Original Sinosoidal surface (bottom) and Stylus Perfect
  Pyramid tip locus (top)');
  figure;
  plot(zsingp1(28,:),'b'); hold on; plot(locgp1(28,:),'r--*');
  legend ('Original surface','locus');
  xlabel('X (Pxiel)');
  ylabel('Height(um)');
  title('2D coss section of original Sinosoidal profile and the
  traced Pyramid stylus locus');
  figure(1111);
  surf(contactgp1);
  xlabel('X (µm)');
  ylabel('Y(µm)');
  zlabel('No. of Contact Points');
  title('Contacts distribution on the Perfect Pyramid tip');
  figure;
  surf(zsingp2);hold on; surf(locgp2+1);
  xlabel('X (Pxiel)');
  ylabel('Y(Pxiel)');
  zlabel('Height(µm)');
  title('Original Sinosoidal surface (bottom) and Stylus 1µm
  Truncated Pyramid tip locus (top)');
```

```
figure(2222);
surf(contactgp2);
xlabel('X (µm)');
ylabel('Y(µm)');
zlabel('No. of Contact Points');
title('Contacts distribution on the 1µm Truncated Pyramid tip');
figure;
surf(zsingp3);hold on; surf(locgp3+1);
xlabel('X (Pxiel)');
ylabel('Y(Pxiel)');
zlabel('Height(µm)');
title('Original Sinosoidal surface (bottom) and Stylus 2µm
Truncated Pyramid  tip locus (top)');
figure(3333);
surf(contactgp3);
xlabel('X (µm)');
ylabel('Y(µm)');
zlabel('No. of Contact Points');
title('Contacts distribution on the 2µm Truncated Pyramid tip');

 figure;
 surf(zrandlp1);hold on; surf(loclp1+1);
 xlabel('X (Pxiel)');
ylabel('Y(Pxiel)');
zlabel('Height(µm)');
title('Original Random surface (bottom) and Stylus Perfect Pyramid
tip locus (top)');
figure(666);
plot(zrandlp1(:,28),'b'); hold on; plot(loclp1(:,28),'r--*');
legend ('Original surface','locus');
xlabel('X (Pxiel)');
ylabel('Height(um)');
title('2D coss section of original Random profile and the traced
Pyramid stylus locus');
figure(4444);
surf(contactlp1);
xlabel('X (µm)');
ylabel('Y(µm)');
zlabel('No. of Contact Points');
title('Contacts distribution on the Perfect Pyramid tip');
 figure;
 surf(zrandlp2);hold on; surf(loclp2+1);
 xlabel('X (Pxiel)');
ylabel('Y(Pxiel)');
zlabel('Height(µm)');
title('Original Random surface (bottom) and Stylus 1µm Truncated
Pyramid tip locus (top)');
figure(5555);
surf(contactlp2);
xlabel('X (µm)');
ylabel('Y(µm)');
zlabel('No. of Contact Points');
title('Contacts distribution on the 1µm Truncated Pyramid tip');
 figure;
 surf(zrandlp3);hold on; surf(loclp3+1);
 xlabel('X (Pxiel)');
ylabel('Y(Pxiel)');
zlabel('Height(µm)');
```

```
title('Original Random surface (bottom) and Stylus 1µm Truncated
Pyramid tip locus (top)');
figure(6666);
surf(contactlp3);
xlabel('X (µm)');
ylabel('Y(µm)');
zlabel('No. of Contact Points');
title('Contacts distribution on the 2µm Truncated Pyramid tip');

figure;
 surf(zrealmp1);hold on; surf(locmp1+2);
 xlabel('X (Pxiel)');
ylabel('Y(Pxiel)');
zlabel('Height(µm)');
title('Original Ground surface (bottom) and Stylus Perfect Pyramid
tip locus (top)');
figure;
plot(zrealmp1(:,28),'b'); hold on; plot(locmp1(:,28),'r--*');
legend ('Original surface','locus');
xlabel('X (Pxiel)');
ylabel('Height(um)');
title('2D coss section of original Ground profile and the traced
Pyramid stylus locus');
figure(7777);
surf(contactmp1);
xlabel('X (µm)');
ylabel('Y(µm)');
zlabel('No. of Contact Points');
title('Contacts distribution on the Perfect Pyramid tip');
figure;
 surf(zrealmp2);hold on; surf(locmp2+2);
 xlabel('X (Pxiel)');
ylabel('Y(Pxiel)');
zlabel('Height(µm)');
title('Original Ground surface (bottom) and Stylus 1µm Truncated
Pyramid tip locus (top)');
figure(8888);
surf(contactmp2);
xlabel('X (µm)');
ylabel('Y(µm)');
zlabel('No. of Contact Points');
title('Contacts distribution on the 1µm Truncated Pyramid tip');
 figure;
 surf(zrealmp3);hold on; surf(locmp3+2);
 xlabel('X (Pxiel)');
ylabel('Y(Pxiel)');
zlabel('Height(µm)');
title('Original Ground surface (bottom) and Stylus 2µm Truncated
Pyramid tip locus (top)');
figure(9999);
surf(contactmp3);
xlabel('X (µm)');
ylabel('Y(µm)');
zlabel('No. of Contact Points');
title('Contacts distribution on the 2µm Truncated Pyramid tip');

%hhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhh
  surf(zsingh1);hold on; surf(locgh1+1);
```

```
xlabel('X (Pxiel)');
ylabel('Y(Pxiel)');
zlabel('Height(µm)');
title('Original Sinosoidal surface (bottom) and Stylus Perfect
Hemisphere tip locus (top)');
figure;
plot(zsingh1(28,:),'b'); hold on; plot(locgh1(28,:),'r--*');
legend ('Original surface','locus');
xlabel('X (Pxiel)');
ylabel('Height(um)');
title('2D coss section of original Sinosoidal profile and the
traced Hemisphere stylus locus');
figure(1000);
surf(contactgh1);
xlabel('X (µm)');
ylabel('Y(µm)');
zlabel('No. of Contact Points)');
title('Contacts distribution on the Perfect Hemisphere tip');
figure;
surf(zsingh2);hold on; surf(locgh2+1);
xlabel('X (Pxiel)');
ylabel('Y(Pxiel)');
zlabel('Height(µm)');
title('Original Sinosoidal surface (bottom) and Stylus 1µm
Truncated Hemisphere tip locus (top)');
figure(1100);
surf(contactgh2);
xlabel('X (µm)');
ylabel('Y(µm)');
zlabel('No. of Contact Points');
title('Contacts distribution on the 1µm Truncated Hemisphere
tip');
figure;
surf(zsingh3);hold on; surf(locgh3+1);
xlabel('X (Pxiel)');
ylabel('Y(Pxiel)');
zlabel('Height(µm)');
title('Original Sinosoidal surface (bottom) and Stylus 2µm
Truncated Hemisphere tip locus (top)');
figure(1200);
surf(contactgh3);
xlabel('X (µm)');
ylabel('Y(µm)');
zlabel('No. of Contact Points');
title('Contacts distribution on the 2µm Truncated Hemisphere
tip');

 figure;
 surf(zrandlh1);hold on; surf(loclh1+1);
 xlabel('X (Pxiel)');
ylabel('Y(Pxiel)');
zlabel('Height(µm)');
title('Original Random surface (bottom) and Stylus Perfect
Hemisphere tip locus (top)');
figure(777);
plot(zrandlh1(:,28),'b'); hold on; plot(loclh1(:,28)+0.35,'r--*');
% 0.22 to draw the locus above the original surface
legend ('Original surface','locus');
xlabel('X (Pxiel)');
```

```
ylabel('Height(um)');
title('2D coss section of original Random profile and the traced
Hemisphere stylus locus');
figure(1300);
surf(contactlh1);
xlabel('X (µm)');
ylabel('Y(µm)');
zlabel('No. of Contact Points');
title('Contacts distribution on the Perfect Hemisphere tip');
 figure;
 surf(zrandlh2);hold on; surf(loclh2+1);
 xlabel('X (Pxiel)');
ylabel('Y(Pxiel)');
zlabel('Height(µm)');
title('Original Random surface (bottom) and Stylus 1µm Truncated
Hemisphere tip locus (top)');
figure(1400);
surf(contactlh2);
xlabel('X (µm)');
ylabel('Y(µm)');
zlabel('No. of Contact Points)');
title('Contacts distribution on the 1µm Truncated Hemisphere
tip');
 figure;
 surf(zrandlh3);hold on; surf(loclh3+1);
 xlabel('X (Pxiel)');
ylabel('Y(Pxiel)');
zlabel('Height(µm)');
title('Original Random surface (bottom) and Stylus 2µm Truncated
Hemisphere tip locus (top)');
figure(1500);
surf(contactlh3);
xlabel('X (µm)');
ylabel('Y(µm)');
zlabel('No. of Contact Points');
title('Contacts distribution on the 2µm Truncated Hemisphere
tip');

figure;
 surf(zrealmh1);hold on; surf(locmh1+2);
 xlabel('X (Pxiel)');
ylabel('Y(Pxiel)');
zlabel('Height(µm)');
title('Original Ground surface (bottom) and Stylus Perfect
Hemisphere tip locus (top)');
figure;
plot(zrealmh1(:,28),'b'); hold on; plot(locmh1(:,28)+0.07,'r--*');
legend ('Original surface','locus');
xlabel('X (Pxiel)');
ylabel('Height(um)');
title('2D coss section of original Ground profile and the traced
Hemisphere stylus locus');
figure(1500);
surf(contactmh1);
xlabel('X (µm)');
ylabel('Y(µm)');
zlabel('No. of Contact Points');
title('Contacts distribution on the Perfect Hemisphere tip');
figure;
```

```
 surf(zrealmh2);hold on; surf(locmh2+2);
 xlabel('X (Pxiel)');
ylabel('Y(Pxiel)');
zlabel('Height(µm)');
title('Original Ground surface (bottom) and Stylus 1µm Truncated
Hemisphere tip locus (top)');
figure(1600);
surf(contactmh2);
xlabel('X (µm)');
ylabel('Y(µm)');
zlabel('No. of Contact Points');
title('Contacts distribution on the 1µm Truncated Hemisphere
tip');
 figure;
 surf(zrealmh3);hold on; surf(locmh3+2);
 xlabel('X (Pxiel)');
ylabel('Y(Pxiel)');
zlabel('Height(µm)');
title('Original Ground surface (bottom) and Stylus 2µm Truncated
Hemisphere tip locus (top)');
figure(1700);
surf(contactmh3);
xlabel('X (µm)');
ylabel('Y(µm)');
zlabel('No. of Contact Points');
title('Contacts distribution on the 2µm Truncated Hemisphere
tip');

%cccccccccccccccccccccccccccccccccccccccccccccccccc
 surf(zsingc1);hold on; surf(locgc1+1);
xlabel('X (Pxiel)');
ylabel('Y(Pxiel)');
zlabel('Height(µm)');
title('Original Sinosoidal surface (bottom) and Stylus Perfect
Conical tip locus (top)');
figure;
plot(zsingc1(28,:),'b'); hold on; plot(locgc1(28,:),'r--*');
legend ('Original surface','locus');
xlabel('X (Pxiel)');
ylabel('Height(um)');
title('2D coss section of original Sinosoidal profile and the
traced Conical stylus locus');
figure(1800);
surf(contactgc1);
xlabel('X (µm)');
ylabel('Y(µm)');
zlabel('No. of Contact Points');
title('Contacts distribution on the Perfect Conical tip');
figure;
surf(zsingc2);hold on; surf(locgc2+1);
xlabel('X (Pxiel)');
ylabel('Y(Pxiel)');
zlabel('Height(µm)');
title('Original Sinosoidal surface (bottom) and Stylus 1µm
Truncated Conical tip locus (top)');
figure(1900);
surf(contactgc2);
xlabel('X (µm)');
ylabel('Y(µm)');
```

```
zlabel('No. of Contact Points');
title('Contacts distribution on the 1µm Truncated Conical tip');
figure;
surf(zsingc3);hold on; surf(locgc3+1);
xlabel('X (Pxiel)');
ylabel('Y(Pxiel)');
zlabel('Height(µm)');
title('Original Sinosoidal surface (bottom) and Stylus 2µm
Truncated Conical tip locus (top)');
figure(1900);
surf(contactgc3);
xlabel('X (µm)');
ylabel('Y(µm)');
zlabel('No. of Contact Points');
title('Contacts distribution on the 2µm Truncated Conical tip');

 figure;
 surf(zrandlc1);hold on; surf(loclc1+1);
 xlabel('X (Pxiel)');
ylabel('Y(Pxiel)');
zlabel('Height(µm)');
title('Original Random surface (bottom) and Stylus Perfect Conical
tip locus (top)');
figure(888);
plot(zrandlc1(:,28),'b'); hold on; plot(loclc1(:,28),'r--*');
legend ('Original surface','locus');
xlabel('X (Pxiel)');
ylabel('Height(um)');
title('2D coss section of original Random profile and the traced
Conical stylus locus');
figure(2000);
surf(contactlc1);
xlabel('X (µm)');
ylabel('Y(µm)');
zlabel('No. of Contact Points');
title('Contacts distribution on the Perfect Conical tip');
 figure;
 surf(zrandlc2);hold on; surf(loclc2+1);
 xlabel('X (Pxiel)');
ylabel('Y(Pxiel)');
zlabel('Height(µm)');
title('Original Random surface (bottom) and Stylus 1µm Truncated
Conical tip locus (top)');
figure(2100);
surf(contactlc2);
xlabel('X (µm)');
ylabel('Y(µm)');
zlabel('No. of Contact Points');
title('Contacts distribution on the 1µm Truncated Conical tip');
 figure;
 surf(zrandlc3);hold on; surf(loclc3+1);
 xlabel('X (Pxiel)');
ylabel('Y(Pxiel)');
zlabel('Height(µm)');
title('Original Random surface (bottom) and Stylus 2µm Truncated
Conical tip locus (top)');
figure(2200);
surf(contactlc3);
xlabel('X (µm)');
```

```
ylabel('Y(µm)');
zlabel('No. of Contact Points');
title('Contacts distribution on the 2µm Truncated Conical tip');

figure;
 surf(zrealmc1);hold on; surf(locmc1+2);
 xlabel('X (Pxiel)');
ylabel('Y(Pxiel)');
zlabel('Height(µm)');
title('Original Ground surface (bottom) and Stylus Perfect Conical
tip locus (top)');
figure;
plot(zrealmc1(:,28),'b'); hold on; plot(locmc1(:,28),'r--*');
legend ('Original surface','locus');
xlabel('X (Pxiel)');
ylabel('Height(um)');
title('2D coss section of original Ground profile and the traced
conical stylus locus');
figure(2300);
surf(contactmc1);
xlabel('X (µm)');
ylabel('Y(µm)');
zlabel('No. of Contact Points');
title('Contacts distribution on the Perfect Conical tip');
figure;
 surf(zrealmc2);hold on; surf(locmc2+2);
 xlabel('X (Pxiel)');
ylabel('Y(Pxiel)');
zlabel('Height(µm)');
title('Original Ground surface (bottom) and Stylus 1µm Truncated
Conical tip locus (top)');
figure(2400);
surf(contactmc2);
xlabel('X (µm)');
ylabel('Y(µm)');
zlabel('No. of Contact Points');
title('Contacts distribution on the 1µm Truncated Conical tip');
 figure;
 surf(zrealmc3);hold on; surf(locmc3+2);
 xlabel('X (Pxiel)');
ylabel('Y(Pxiel)');
zlabel('Height(µm)');
title('Original Ground surface (bottom) and Stylus 2µm Truncated
Conical tip locus (top)');
figure(2500);
surf(contactmc3);
xlabel('X (µm)');
ylabel('Y(µm)');
zlabel('No. of Contact Points');
title('Contacts distribution on the 2µm Truncated Conical tip');
```

A3.2 The second Matlab Code:

```
clc
clear all
close all
% Demo code fragment for testing 'fractal infill'
% NOT EXECUTABLE - mainly psuedo-code made up of comments
% Assume square master array Zorig, that will be sub-sampled into
Z.
% For each version of Z, use ideas from FractalInfillDemo1.m
% to reconstruct expanded array Zexp of same size as Zorig.
% Then compare Zexp to Zorig to see effectiveness of the
'prediction'.

nzorig = 601;   %say, 'binary' width easy to divide up!
mag = 14;  % or 2 or 8: sub-sampler and later expansion factor
%Zorig=ones(nzorig);
%Zorig=xlsread('z.xlsx');
%Zorig=xlsread('KAL90ROV_Filter_Result_AFM_asc.xlsx');
%DDDDD=importdata('KAL90ROV_Filter_Result_AFM.asc');
Zorig = importdata('KAL90ROV_Filter_Result_AFM.asc');
%   Zorig=Zorig-min(min(Zorig));     % To scale the values to be
between [0-max-min)] (Making the min=0)
%   Zorig=Zorig/(max(max(Zorig)));   %Dimensionless: scaling the
values to be in the range of (0-1)
%   p1=0.925;
%   Zorig=(Zorig*p1);
Z = Zorig(1:mag:nzorig, 1:mag:nzorig); % sub-sampled every mag
points.
nz=fix(nzorig/mag)+1;  % points in reduced Z
nf=mag+1;   % size of interpolation patch needed for re-expanding
ns=nzorig;   %full size of expanded (stylus) map
%
% Build plane and fractal infill expansions of Z:
% this section is essentially same ideas as in FractalInfillDemo1.
%
% grab a bit of surface for infill, nf square, taken from middle
of Z
centrepoint=fix(nz/2)+1; halfsize=fix(nf/2); %to define region
top=centrepoint+halfsize; low=centrepoint-halfsize;
filler=Z(low:top,low:top);  %
frac_scale = 1/nf;    %to scale infill heights to width scales
%use TwistPlane in its 'leveling mode (zero the corners)
C=[filler(1,1), filler(1,nf); filler(nf,1), filler(nf,nf)];
fitplane=TwistPlane(C, nf); %get twisted plane matching four
corners
filler=filler-fitplane;   %leveling operation
filler=frac_scale*filler; %reduced height, leveled
% end of building infill (not most effciient - spelled out for
clarity)
%
% Now expand each foursome of Z into a Zexp* with twisted plane
% interpolation and superposed leveled fractal infill each time
% (using always same one here!!)
Zexpp=zeros(ns);    %placeholder full array to allow easy index
notations.
Zexpf=zeros(ns);
```

```
%%work through rows and columns from 1 to nz-1
%inefficient becuase of ovverlaps on edges in ZS, but easy to
follow!
ii=1;  %row index points in expanded array
for i=1:(nz-1);
    jj=1;  %col index points in expanded array
    for j=1:(nz-1);
        C=[Z(i,j), Z(i,j+1); Z(1+1,j), Z(i+1,j+1)]; %corners
        patch=TwistPlane(C, nf);
        iil=ii+(nf-1); jjl=jj+(nf-1);  %upper limts of patch
        Zexpp(ii:iil,jj:jjl)=patch;  %plane interpolated
        Zexpf(ii:iil,jj:jjl)=patch+filler; %and with fractal
'noise'
        jj=jj+nf;   %increment the outer index col count
    end
    ii=ii+nf;   %increment the outer row index
end
figure (1);surf (Zorig);
shading interp
colormap copper
figure (2);surf (Z);
shading interp
colormap copper
figure (3);surf (Zexpp);
shading interp
colormap copper
figure (4);surf (Zexpf);
shading interp
colormap copper
% Zexp* now preserves orignal points of Z in correct places as in
Zorig
%
%Get estimate of how much time this process takes (compared to
stylus sim)
%
% Next phase is to run various comparisons
% Save data files and pass t, say, SPIP to evalaute differences in
% parameters from Zorig, ZZexpp and Zexpf.
% Simply plot graphs to show variations?
%Error_p = Zexpp-Z; %and plot, express rms , etc.
%Error_f = Zexpf-Zorig; %ditto
% Also, apply stylus to each with small threshold steps to see
whether
% they show different patterns of kinematic contact, island
growth,etc.
% If not, e.g., extra effort is proven not worthwhile in our
context.

%%%Roughness for Zorig:
SaZorig=sum(sum(Zorig(:,:)))/(nzorig*nzorig);
% meanlevel=SaZorig;
% Zorig=Zorig-meanlevel;  %leaving tilt
ff=sum(sum(Zorig(:,:)));
fff=ff^2;
ffff=fff/(nzorig*nzorig);
SqZorig=(ffff)^0.5
SskZorig=(sum(sum(Zorig(:,:)))^3)/(nzorig*nzorig*(SqZorig^3))
```

```
SkuZorig=(sum(sum(Zorig(:,:)))^4)/(nzorig*nzorig*(SqZorig^4))
SzZorig=max(max(Zorig(:,:)))-min(min(Zorig(:,:)))
%%%Roughness for Z:
SaZ=sum(sum(Z(:,:)))/(nz*nz);
% meanlevel=SaZ;
% Z=Z-meanlevel;  %leaving tilt
ff=sum(sum(Z(:,:)));
fff=ff^2;
ffff=fff/(nz*nz);
SqZ=(ffff)^0.5
SskZ=(sum(sum(Z(:,:)))^3)/(nz*nz*(SqZ^3))
SkuZ=(sum(sum(Z(:,:)))^4)/(nz*nz*(SqZ^4))
SzZ=max(max(Z(:,:)))-min(min(Z(:,:)))
%%%Roughness for Zexpp:
SaZexpp=sum(sum(Zexpp(:,:)))/(ns*ns);
% meanlevel=SaZexpp;
% Zexpp=Zexpp-meanlevel;  %leaving tilt
ff=sum(sum(Zexpp(:,:)));
fff=ff^2;
ffff=fff/(ns*ns);
SqZexpp=(ffff)^0.5
SskZexpp=(sum(sum(Zexpp(:,:)))^3)/(ns*ns*(SqZexpp^3))
SkuZexpp=(sum(sum(Zexpp(:,:)))^4)/(ns*ns*(SqZexpp^4))
SzZexpp=max(max(Zexpp(:,:)))-min(min(Zexpp(:,:)))
%%%Roughness for Zexpf:
SaZexpf=sum(sum(Zexpf(:,:)))/(ns*ns);
% meanlevel=SaZexpf;
% Zexpf=Zexpf-meanlevel;  %leaving tilt
ff=sum(sum(Zexpf(:,:)));
fff=ff^2;
ffff=fff/(ns*ns);
SqZexpf=(ffff)^0.5
SskZexpf=(sum(sum(Zexpf(:,:)))^3)/(ns*ns*(SqZexpf^3))
SkuZexpf=(sum(sum(Zexpf(:,:)))^4)/(ns*ns*(SqZexpf^4))
SzZexpf=max(max(Zexpf(:,:)))-min(min(Zexpf(:,:)))
%
Error_pSa = SaZexpp-SaZ %and plot, express rms , etc.
Error_pSq = SqZexpp-SqZ %and plot, express rms , etc.
Error_pSsk = SskZexpp-SskZ %and plot, express rms , etc.
Error_pSsz = SzZexpp-SzZ %and plot, express rms , etc.

Error_fSa = SaZexpf-SaZorig; %ditto
Error_fSq = SqZexpf-SqZorig %and plot, express rms , etc.
Error_fSsk = SskZexpf-SskZorig %and plot, express rms , etc.
Error_fSsz = SzZexpf-SzZorig %and plot, express rms , etc.

save ('Zorig','-ascii')
```

A3.3 The third Matlab Code:

```
clc
clear all
close all
% Demo code fragment for crude 'fractal infill'
% Assume master array Z, sampled at hz=1 um, of nz=17 points
square.
% Build expanded surface ZS to match stylus grid at hs=0.25 um
% scaling factor is x4, but infill is then on a grid of 5x5 points
% conventionally always use odd numbers of points, to allow
centre-zero.
nz=17; Z=ones(nz); hz=1.0; hs=0.25;  %placeholder values!
nf=hz/hs+1;   %size of interpolation patch
ns=(nz-1)*(hz/hs)+1;   %full size of expanded (stylus) map
%
%If wanted, grab a bit of surface for infill
%needs to be nf square, taken, e.g., from middle of Z
filler=Z(7:11,7:11);  %set hardwired for nf=5 here!!!
frac_scale = 1/nf;    %to scale infill heights to width scales
%use TwistPlane in its 'leveling mode (zero the corners)
C=[filler(1,1), filler(1,5); filler(5,1), filler(5,5)];
fitplane=TwistPlane(C, nf); %get twisted plane matching four
corners
filler=filler-fitplane;   %leveling operation
filler=frac_scale*filler; %reduced height, leveled
% end of building infill (not most effciient - spelled out for
clarity)
%
%Now expand each foursome of Z into ZS with twisted plane
interpolation
%superpose leveled fractal infill each time (using always same one
here!!)
ZS=zeros(ns);    %placeholder full array to allow easy index
notations.

%%work through rows and columns from 1 to nz-1
%inefficient becuase of ovverlaps on edges in ZS, but easy to
follow!
ii=1;  %row index points in expanded array
for i=1:(nz-1)
    jj=1;  %col index points in expanded array
    for j=1:(nz-1)
        C=[Z(i,j), Z(i,j+1); Z(1+1,j), Z(i+1,j+1)]; %corners
        patch=TwistPlane(C, nf);
        iil=ii+(nf-1); jjl=jj+(nf-1);  %upper limts of patch
        ZS(ii:iil,jj:jjl)=patch;  %plane interpolated, Z heights
preserved
        ZS(ii:iil,jj:jjl)=ZS(ii:iil,jj:jjl)+filler %to add fractal
'noise'
        jj=jj+nf;   %increment the outer index col count
    end
    ii=ii+nf;   %increment the outer row index
end
%ZS now has full population, wiht all orignal points of Z
preserved
%because TwistPlane preseves them, so ensuring corner of filler
are zero.
```

A3.4 The fourth Matlab Code:

```
%%%Last work on 2 December 2013
%%%continue 29/03/2015(Cone)
clc
close all
clear all

%surface(1)
%a=xlsread('G2ss201.xls');
%a=ones(128)
% dx = 0.156; % spacing in x in mm
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%29/03/2015
% dy = 0.156; % spacing in y in mm
% nx = 128; % number of points along x
% ny = 128; % number of points along y

% surface (2)
% a=xlsread('KA90CO2_AFM.xlsx');%64
% dx = 0.313; % spacing in x in mm
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%29/03/2015
% dy = 0.313; % spacing in y in mm
% nx = 64; % number of points along x
% ny = 64; % number of points along y

%surface (3)
% a=xlsread('zoriginal_scaled01_64.xlsx');%64
% dx = 0.313; % spacing in x in mm
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%29/03/2015
% dy = 0.313; % spacing in y in mm
% nx = 64; % number of points along x
% ny = 64; % number of points along y

% surface (4)
a=xlsread('zoriginal_scaled01_64noise01.xlsx');%64
dx = 0.313; % spacing in x in mm
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%29/03/2015
dy = 0.313; % spacing in y in mm
nx = 64; % number of points along x
ny = 64; % number of points along y

x= (0:1:nx-1)*dx; % generate x array
y = (0:1:ny-1)*dy; % generate y array
for j = 1:ny
    for i = 1:nx
    z(j,i) = a(j,i);
end
end
z=z/(max(max(z)));
p1=0.8148130;
z=(z*p1);
figure (1111); surf(z);
xlabel('x Distance (pixel)');
ylabel('y Distance (pixel)');
zlabel('Amplitude (um)');
title ('z');
shading interp;
```

```
%Generating a sphere of radius 11 µm, assuming a sampling interval
of
%1µm in x and z.
%The Cone can be generated as shown below.
Radius =2; % radius in mm
 dx =0.1; % spacing in x in mm
 dy =0.1 ; % spacing in y in mm
% nx = 128; % number of points along x %%for 128
% ny = 128; % number of points along y

nx = 64; % number of points along x   %%for 64
ny = 64; % number of points along y
x= (0:1:nx-1)*dx; % generate x array
y = (0:1:ny-1)*dy;
mx = Radius/dx; % number of points in one radius of the ball
my = Radius/dy;
%end
j = 1;
for yr = -Radius:dy:Radius;
i = 1;
for xr = -Radius:dx:Radius;
if sqrt(xr^2+yr^2) <= Radius;
B(j,i) = -sqrt(xr^2+yr^2);
else
    B(j,i) =-2;
end
i = i + 1;
end
j = j + 1;
end
%B=B;
figure(555); surf(B);

%The locus of the center of the sphere can be obtained as shown
below.
z1(ny+2*my+1,nx+2*mx+1) =0;
z1(my+1:ny+my,mx+1:nx+mx) = z;
for j=my+1:ny+my;
for k = mx+1:nx+mx;

    C(j-my,k-mx) = max(max(z1(j-my:j+my,k-mx:k+mx)));
% compare sum of profile and ball heights

end
end
figure (2); surf(z);
hold on;
mesh(C);
xlabel('x Distance (pixel)');
ylabel('y Distance (pixel)');
zlabel('Amplitude (um)');
title ('z & loc');
shading interp;
hold off
% G = fspecial('gaussian',[1 1],2);
% Ig = imfilter(z,G,'same');
% figure,imshow(Ig);
```

```
% C = conv2(z,G);
%
% return
% generating a Gaussian Filter S
% a = sqrt(log(2)/pi);
% lamdacX = 0.08; % cutoff along x in mm
% lamdacY = 0.08; % cutoff along y in mm
% x = (-lamdacX:dx:lamdacX-dx)';
% y = (-lamdacY:dy:lamdacY-dy)';
% mx = size(x,1); % number of points along x
% my = size(y,1); % number of points along y
% for i = 1:mx
%     for j = 1:my
%         S(j,i)=(1/(a^2*lamdacX*lamdacY))*exp(-
pi*(x(i)/a/lamdacX)^2*(-pi*(y(j)/a/lamdacY)^2));
%     end
% end
% S = S/sum(sum(S));
% figure (3); surf(x,y,S)
% title ('Gaussian');
% shading interp
%
% dx = 0.002; % spacing in x in mm
% dy = 0.002; % spacing in y in mm
% nxc = 135; % number of points along x
% nyc = 135; % number of points along y
% x= (0:1:nxc-1)*dx; % generate x array
% y = (0:1:nyc-1)*dy; % generate y array
%
% for i = 1:nxc
%     for j = 1:nyc
%       C2 = conv2(z,S);
%     end
% end
%
%       C2 = conv2(z,S);
% figure (4),surf(C2);
% shading interp;
% title ('conv');
% shading interp
% return
% w = C2(my/2+1:ny+my/2,mx/2+1:nx+mx/2);
% figure (5),surf(w);
%
% shading interp
 %return
  %%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%% extra
 a=z(1:size(B),1:size(B));        % get archive data
%%%update14/12/2013
%C=C(1:size(B),1:size(B));
% a=C;
%%%%%%%%%%%%%%%
pr=B;        % get stylus module (pr=probe)
%%%%%%%%%[prx,pry]=size(pr)  ;             % basic index range
%%%%%%%%%%%[sx,sy]=size(a1);              % overall diameter

istart=1;
iend=size(B);
```

```
i=istart:iend;
jstart=1;
jend=size(B);
j=istart:jend;

thr=0.05; %0, 0.005, 0.01, 0.02, 0.05, 0.1
% 0.057, 0.07, 0.85 and 0.105
%figure (77); plot(z(:,26),'b'); hold on; plot(B(:,26)+thr,'r');
hold on; plot (C(:,26));%hold on;plot(ssss(1:7,4)+3,'k-*');

%%%%%%%%%%%%%%%%a=a(i:i+10,j:j+10);
%pr=zeros(21);                          %Probe shape (11x11)
%%%%%%%%%%%%%%%%M=a(1,1)+pr(1,1); % canceled,we used atopr by
differnt way
atopr=a+pr;                             %interaction function
M=atopr(1,1);
for i=istart:iend;
    for j=jstart:jend;
        if atopr(i,j)>M;
            M =atopr(i,j);           %Holding max vdm
            imax=i;              % index of the current max
            jmax=j;

        end
    end
end
firstcontact= zeros(size(atopr));            %11x11
firstcontact(imax,jmax)=1;          % Map of tip, all zero except
simple contact point
fullcontact=zeros(size(atopr));
fullcontact=atopr>=M-thr;

gap=atopr-(M-thr);

%a=a';
% dx = 0.1; % spacing in x in mm
% dy = 0.1; % spacing in y in mm
% nx = 41; % number of points along x
% ny = 41; % number of points along y
% x= (0:1:nx-1)*dx; % generate x array
% y = (0:1:ny-1)*dy;
% figure (121); surf (x,y,a);
figure (121); surf (a);
xlabel('X (pixel)');
ylabel('Y (pixel)');
zlabel('Height (um)');
figure (1211); surf (atopr);
figure (122) ;  waterfall(pr);
figure (123); spy(fullcontact,'b'); hold on ;
spy(firstcontact,'r')
grid on;
set(gca, 'GridLineStyle', ':');
figure (124); spy(fullcontact,'b');
grid on;
set(gca, 'GridLineStyle', ':');
```

10.4. Appendix 4: SPIP Software information

Measurement:

The following measurement tools allows us to make a wide range of measurements interactively by drawing the available shapes in an image. The common shapes, which can be measured are listed below:

1. Line measurement; primarily used for measuring length and height differences.

2. Polygon measurement; primarily used for measuring length, width, height, area and volume.

3. Ellipse and Circle and measurement; primarily used for measuring radius and length.

Measurands

The shapes can be quantified by comprehensive set of parameters called measurands, which are described below:

- ***Area:***

$$Area = \frac{\sum_i (x_i + x_{i-1}) \cdot (y_i - y_{i-1})}{2}$$

The Area is calculated from the shapes periphery, i.e. the closed polygon that surrounds the feature. The area is calculated using:

- ***Perimeter***:

The Perimeter is calculated from the shapes periphery as:

$$Perimeter = \sum_i \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2}$$

- ***Centre of Gravity:***

The centre of gravity can be calculated using three different methods: Periphery, Area and Volume.

$$CG_x = \frac{\sum_i (x_i + x_{i-1})^2 \cdot (y_i - y_{i-1})}{Area}$$

$$CG_y = \frac{\sum_i (y_i + y_{i-1})^2 \cdot (x_i - x_{i-1})}{Area}$$

Periphery or Outline

The centre of gravity based on the shapes periphery, i.e the closed polygon that surrounds the shape. Here, centre of gravity is calculated using:

Area

The centre of gravity based on a shapes interior scan-point.

In this case the centre is calculated as the average value of all horizontal respectively vertical coordinates within the shape.

$$CG_x = \frac{\sum_{x,y} x}{Area}$$

$$CG_y = \frac{\sum_{x,y} y}{Area}$$

Volume

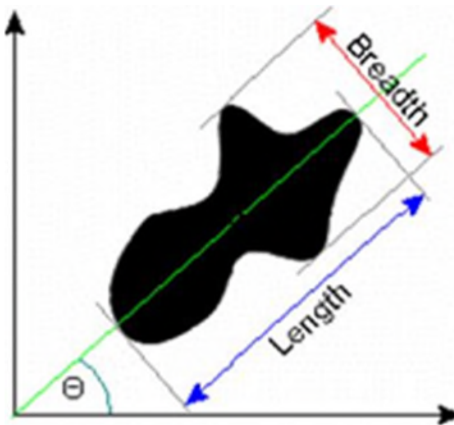
The centre of gravity, where each point at x,y position within the shapes is weighted by that scan points height.

$$CG_x = \frac{\sum_{x,y} x \cdot height}{Volume}$$

$$CG_y = \frac{\sum_{x,y} y \cdot height}{Volume}$$

Orientation:

Orientation gives the angle of the axis of momentum. To obtain the Orientation we find the line which best fits all the points in the object, actually only the points describing the periphery are used. This line is the "axis of momentum". Having the moment axis it’s simply a matter of calculating the angle to the x-axis.



$$M_x = \sum_{x,y} x^2 - \frac{\left(\sum_{x,y} x\right)^2}{Area}$$

$$M_y = \sum_{x,y} y^2 - \frac{\left(\sum_{x,y} y\right)^2}{Area}$$

$$M_{xy} = \sum_{x,y} (x \cdot y) - \frac{\sum_{x,y} x \cdot \sum_{x,y} y}{Area}$$

$$\Theta = \tan^{-1}\left(\frac{M_x - M_y + \sqrt{(M_x - M_y)^2 + 4 \cdot M_{xy}^2}}{2 \cdot M_{xy}} \right)$$

Length:

Length (also known as the maximum caliper) is defined as the longest cord along the angle given by the moment's axis to the x-axis. In other words, if we rotated the shape so that the moment's axis becomes parallel to the x-axis, then Length is the extension of the bounding rectangle in the horizontal direction.

- For lines this is the distance between the most distant points of the line.
- For circles this equals the diameter

- For ellipses it equals the length of the major axis

Breadth / Width:

Breadth (also known as the minimum caliper) is defined as the longest cord perpendicular to the angle ® given by the moments axis to the x-axis. In other words, if we rotated the feature so that the moments axis becomes parallel to the x-axis, then Breadth is the extension of the bounding rectangle in the vertical direction.

Maximum Feret's diameter:

Maximum Feret diameter is the longest distance between any two points on the periphery. Diameter (Heywood):

The diameter is expressed as the diameter of a circle having an area equivalent to the shape's area.

$$Diameter = \sqrt{\frac{4}{\pi} \cdot Area}$$

Radius is defined as half the size of that of the Heywood Diameter described above.

Elongation:

Elongation is a measure indicating how elongated a shape is. A square or circle will return the value zero. As these shapes changes towards a long rectangle or ellipse the returned value converges towards 1.0.

$$Elongation = \frac{|Length - Breadth|}{Length}$$

Aspect Ratio:

Aspect Ratio is the aspect ration defined as Length over Breadth. The aspect ration will from this definition always be greater than or equal to 1.0. The aspect ratio of both a circle and square is 1.0, whereas other shapes will have a value less than 1.0.

$$AspectRatio = \frac{Length}{Breadth}$$

Form Factor:

Form Factor provides a measure that describes the shape of a feature. The form factor is defined by the formula below.

$$FormFactor = \frac{4 \cdot \pi \cdot Area}{Perimeter^2}$$

Roundness:

Roundness describes the shapes resemblance to a circle. The roundness factor of a shape which will approach 1.0 the closer the shape resembles a circle.

$$Roundness = \frac{4 \cdot Area}{\pi \cdot Length^2}$$

Compactness:

Compactness is a measure expressing how compact a feature is. From the formula below, a circle will have a compactness of 1.0, a squares compactness is 1.1284, whereas elongated and irregular shapes results in values less than 1.0.

$$Compactness = \frac{\sqrt{\frac{4 \cdot Area}{\pi}}}{Length}$$

Convexity:

Convexity is the ratio between the convex polygons perimeter (shown in green) and the perimeter of the shape itself.

$$Convexity = \frac{ConvexPerimeter}{Perimeter}$$

Solidity:

Solidity is a measure describing the resemblance of the shapes area with it's convex area.

$$Solidity = \frac{Area}{ConvexArea}$$