Feature-based characterisation of surface topography

By
Nicola Senin

Thesis
Submitted to the University of Nottigham
for the degree of
Doctor of Philosophy

Faculty of Engineering
2018



University of Nottingham
UK | CHINA | MALAYSIA

Abstract

In recent years surface metrology has undergone a revolution: non-contact technologies have driven the transition from profile to areal measurement, and topography information can now be obtained in the form of three-dimensional geometric models. However, the conceptual approach underlying topography characterisation has not changed as much, and surfaces are still predominantly quantified in terms of "roughness".

This thesis explores *feature-based characterisation*, an approach that merges technologies from computer vision, image processing, geometric modelling, and statistical modelling, to forge a new set of tools for the analysis of three-dimensional surface topography. Feature-based characterisation provides the end-user with the capability of identifying, isolating and characterising any topographic formation of interest which may be found on a measured surface, addressing characterisation needs that may go well beyond the mere assessment of surface roughness.

Feature-based characterisation of surface topography offers new ways to approach currently challenging metrology problems, and offers new opportunities to explore original pathways in the development of advanced manufacturing processes, materials and products.

This thesis illustrates original methods developed by the candidate for feature-based characterisation, and presents a first attempt at unifying such methods into a comprehensive framework where feature-based characterisation is seen as an alternative to conventional characterisation based on quantifying roughness. Throughout the thesis, the foundational elements of feature-based characterisation framework will be illustrated and discussed with the help of examples from real-life applications.

Declarations

This thesis documents the research activities that I have carried on since May 2015 (the start of my doctoral studies). Despite having begun to work in surface metrology quite some time earlier (1999), it has been during the years of my doctorate that I was able to be fully productive in the generation of new methods for the analysis of surface data, and in particular in the development of a new theoretical foundation supporting an original, hopefully cohesive framework gathering all the methods which I have developed over the years, including the most recent ones.

When writing this thesis I found myself facing the problem of how to illustrate what is essentially a new theoretical framework, incorporating novel methods that I have devised during the doctorate, but also methods that date earlier than my doctoral studies. I have opted for a comprehensive illustration, to provide a more complete picture of the various aspects of feature-based characterisation, and to highlight the additional work of re-thinking and re-designing that I had to carry on during my doctoral studies, in order to fit the older methods into the new framework.

To address any doubts about how doctoral and pre-doctoral work has been handled in this thesis, I am reporting the following chapter-based breakdown.

- In *Chapter 1* the scenarios where conventional surface metrology may prove inadequate are introduced and described. Despite a few scenarios having been sporadically addressed in pre-doctoral work (e.g. dimensional inspection of microparts, characterisation of signature topographies from mechanical testing), the complete taxonomy is novel, and product of the doctoral studies.

- The review of the state-of-the-art presented in *Chapter 2* is original, and based on the novel perspective introduced by the proposed framework.

- The contents of *Chapter 3* (data structures to encode surface topography information) are original to the doctorate.

– The foundational elements of the framework illustrated in *Chapter 4* are original to the doctorate.

– *Chapter 5* covers feature characterisation, one of the fundamental components of the framework. Characterisation methods produced during my doctoral studies are: slicing and 2D datums (Section 5.6) principal component analysis on full-3D geometries (Section 5.3.2), the adaptation of the Angular Radial Transform (Section 5.4.3) and of the Scale-invariant Feature Transform (Section 5.4.5) for surface topography data. All the other methods have been produced at earlier dates, but have been re-designed for the framework, and investigated in terms of how they fit with the taxonomy of surfaces.

– Feature identification and extraction are covered in *Chapter 6*. Methods developed during the doctorate include: scan-free template matching (Section 6.3.2); the partitioning methods described in Section 6.4.3, except simple thresholding and k-means clustering; the contour stability analysis method (Section 6.4.7) and the solutions for transitioning from edge maps to feature maps (Section 6.4.7). The other methods discussed in the chapter have been re-designed for the framework and investigated on applicability to the previously introduced taxonomy of surfaces.

– In *Chapter 7* error and uncertainty in feature-based characterisation is discussed. The contents of the chapter have been entirely produced during the doctorate.

– *Chapter 8* illustrates methods for the generation of simulated surfaces. The contents have been entirely produced during the doctorate.

– *Chapter 9* illustrates the application of feature-based characterisation to a selected set of case studies. All the applications were investigated during the doctorate, except for the micro-target test case (Section 9.2.1).

– *Chapter 10* reports conclusions and outlook, reflecting the perspective that I have developed during the doctoral studies.

As of May 2018, I have authored or co-authored a total of 26 journal papers, 24 conference papers and 3 book chapters in subjects involving surface metrology. In addition, I have co-edited a book on geometric tolerances, covering also surface metrology. Of this research output, 10 journal papers and 11 conference papers have been conceived and produced during my doctoral studies. The list of publications that I have authored or co-authored in surface metrology is provided in a separate section at the end of this thesis. Several of such publications are also cited throughout the thesis when relevant

to the subjects described in the chapters. Some figure panels presented in this thesis have been taken from such previous works. Wherever this has happened, I have added proper references to the original works. Publications are a collective effort. However, in choosing the methods and results presented in this thesis, I have made sure to select only those that represent research work pursued by me. In any case, I have also reported my specific role in the publications that were authored or co-authored by me during the doctorate in the section where they are listed.

Finally, all the figures showing renders of datasets or analysis results, have been generated by software that I have personally developed to demonstrate the framework and the embedded methods, unless otherwise indicated in the related text or figure captions.

Nottingham, ...
Sept 2018 *Nicola Senin*

To my family

Acknowledgments

I would like to thank my supervisors Richard Leach and Simon Lawes for the guidance, and for the numerous and helpful discussions during my doctorate. Thank you Richard again for the incredible drive, and for the outstanding influence you have had on my professional, and personal, growth, ever since I met you at NPL in 2012.

I would also like to thank my first-year internal assessor, Joel Segal, for the helpful suggestions on how to improve this research work.

Thanks to the numerous academics, post-docs and students at the University of Nottingham, which I have interacted with over the years, and who have helped me by providing information, helpful hints and, most importantly, a large amount of test cases to challenge my ideas and methods.

Thanks also to all the past and present members of the Manufacturing Metrology Team, my second family outside family. Again, source of countless helpful discussions and interactions on the many issues that surround manufacturing metrology.

Thanks to all the academics, post-docs, students and industrial researchers from other universities and companies in the UK, Europe, USA, and rest of the world, which I have met at conferences, seminars, or which I have interacted with more frequently through research meetings and projects. All our conversations left me with something, or maybe made me tweak something, or made me radically rethink some parts of my research, or scrap some others. As many, I have the perception that my ideas, are ideas I have come up with by myself. As many, I am most definitely wrong.

Finally, thanks to all my co-authors. Especially to my most frequent co-authors of recent years, for the incredible drive and unparalleled enthusiasm in embracing new research challenges.

Contents

Abstract iii

Declarations v

Acknowledgments ix

Glossary xvii

1 Introduction 1

1.1 Background 1

1.1.1 Basic assumptions of conventional surface metrology 2

1.1.2 Application scenarios where conventional surface metrology may be inappropriate 5

1.2 Main aims and objectives 10

1.3 Outline of this work 11

2 Review of current literature 15

2.1 Surface components and processing operations 15

2.2 Texture parameters 18

2.2.1 Field parameters 18

2.2.2 Feature parameters 28

2.3 Conventional surface metrology and non-conventional scenarios 30

2.4 Relationships with the dimensional/geometric inspection and verification of standard-sized parts 32

2.5 Previous work on feature-based surface characterisation 33

2.5.1 Commercial software 34
2.6 Summary and outlook 36

3 Surface topography information **39**
3.1 Two-dimensional imaging 39
3.1.1 Encoding of surface data from two-dimensional imaging 42
3.2 Line-scan measurement 43
3.2.1 Encoding of surface data from line-scan measurement 44
3.3 Vertical-scan imaging 45
3.3.1 Encoding of surface data from vertical-scan imaging 46
3.4 Volumetric imaging 47
3.5 Other topography measurement solutions that result in triangle meshes 50
3.6 Converting triangle meshes into height maps 51
3.7 An original approach to map-based encoding of topographic features . . 53
3.7.1 Implications related to map-based encoding of topographic features 53
3.8 Mathematical notation 56
3.9 Summary and outlook 57

4 Basic concepts underlying the feature-based characterisation framework **59**
4.1 Introduction 59
4.2 Determinism and randomness in topography 60
4.2.1 Random topographies as random functions 62
4.2.2 Randomness and F-operators 63
4.2.3 Randomness and S,L-filters 65
4.2.4 Randomness, determinism and characterisation pipelines 67
4.2.5 A quick recap of randomness vs determinism in topography . . . 67
4.3 Random fields to model topography 68
4.3.1 Topography as a random field 69
4.3.2 Introducing dependent random variables in the field 70
4.3.3 Covariance in random fields 70

4.3.4 Topography profiles as Gaussian random fields 71
4.3.5 Further assumptions to simplify the random field 71
4.3.6 The role of stationarity in surface metrology 73
4.4 Characterisation of non-stationary fields 75
4.4.1 Linking random fields to randomness intended as uncertainty in prediction 78
4.4.2 Random fields versus moving statistics 79
4.5 Feature technology 81
4.5.1 Shape complexity, spatial differentiation and feature technologies 81
4.5.2 Definitions 82
4.6 Summary and outlook 84

5 Feature characterisation **87**
5.1 Approaches to the characterisation of topography 87
5.1.1 Data preparation for feature characterisation 89
5.2 CAD-compare methods 91
5.3 Datum-based methods 93
5.3.1 Slicing and 2D datums 96
5.3.2 Datum axes 97
5.3.3 Similarity metrics for datum-based descriptions 103
5.4 Shape resampling and integration methods 103
5.4.1 Spin images 104
5.4.2 The ring projection transform - RPT 106
5.4.3 The angular radial transform - ART 107
5.4.4 Multi-scale descriptors 108
5.4.5 The Scale Invariant Feature Transform - SIFT 111
5.5 Spatial occupancy enumeration and cell decomposition methods 117
5.6 Characterisation of feature patterns 119
5.6.1 Reconstruction of feature patterns 119
5.6.2 Characterisation of feature patterns 121

5.7 Summary and outlook 123

6 Feature identification and extraction **125**

6.1 Introduction 125

6.2 An overview of feature identification methods 126

6.3 Template matching 126

6.3.1 Template matching by scanning 127

6.3.2 Scan-free template matching 128

6.3.3 Template matching, randomness and determinism 130

6.4 Template-less, segmentation-based identification 130

6.4.1 Introduction to segmentation 131

6.4.2 Phase I of segmentation - obtaining response maps 132

6.4.3 Phase II of segmentation: obtaining the partitions 136

6.4.4 Phase III of segmentation: post-processing of the partitions . . . 140

6.4.5 From segment maps to feature identification 142

6.4.6 Revisiting ISO 25178-2 morphological segmentation 142

6.4.7 Edge detection methods 143

6.5 Defect identification in regular patterns 150

6.6 Feature extraction 152

6.6.1 Feature boundary determination 152

6.6.2 Feature separation 154

6.7 Summary and outlook 155

7 Error and uncertainty in feature-based characterisation **157**

7.1 From uncertainty in measurement to uncertainty in the characterisation of surfaces 157

7.1.1 Uncertainty in surface metrology 157

7.2 Estimating uncertainty through replicate experiments 160

7.3 Estimating uncertainty in surface topography measurement via experimental evaluation of repeatability error and local bias 161

7.3.1 Estimating local bias 162

7.4 Experimental evaluation of repeatability and reproducibility error in feature-based dimensional assessment 166
7.5 Summary and outlook 169

8 Generation of simulated topographies **171**
8.1 The role and importance of generating topographies 171
8.2 Fundamentals of map generation 172
8.2.1 Layer-based generation 172
8.3 Deterministic topographies 173
8.3.1 Deterministic topographies from continuous and discrete functions 173
8.3.2 Deterministic topographies from triangle meshes 174
8.4 Random topographies 175
8.4.1 Pure random topographies 176
8.4.2 Interpolated random topographies 176
8.4.3 Multi-scale interpolated random topographies 177
8.4.4 Parametric topographies with random parameters 178
8.5 Matching topography specifications 179
8.5.1 Monte Carlo rejection sampling 179
8.5.2 Genetic optimisation 181
8.6 Generating partitions 182
8.7 Examples 183
8.7.1 Simulating the effects of operational life of a surface 183
8.7.2 Simulating measurement 184
8.8 Summary and outlook 185

9 Applications **187**
9.1 Research in advanced manufacturing processes 187
9.1.1 Signature topography of laser powder bed fusion of metals . . . 191
9.1.2 Spatter as function of tilt in electron powder bed fusion 195
9.1.3 Effects of post-processing on electron powder bed fusion surfaces 195
9.1.4 Material removal by electrolyte jet machining 196

9.2 Engineered topographies 198
9.2.1 Micro-targets 199
9.2.2 Micro-fluidics channels for lab-on-a-chip devices 200
9.2.3 Arrays of micro-pockets for friction reduction 202
9.2.4 High aspect-ratio micro-holes 205
9.3 Tribological testing and functional life of surfaces 207
9.3.1 Fretting wear (testing) 208
9.3.2 Fretting wear (functional life) 208
9.4 Summary and outlook 209

10 Conclusions **211**
10.1 Achievement of research objectives 211
10.2 Discussion 214
10.2.1 Usefulness (and relations to surface function) 214
10.2.2 Usability 216
10.2.3 The set-up challenge 217
10.2.4 Procedural definitions 218
10.2.5 Uncertainty and validation 218
10.2.6 Shape and randomness 219
10.2.7 Moving to fully three-dimensional geometric models 219
10.2.8 The role of machine learning 220
10.2.9 In-process metrology 221
10.3 Final remarks 222

Bibliography **224**

Author's publications **235**

Contributions to journal papers published during the doctorate **243**

Glossary

Feature A geometric element of interest (point, curve or region) within an *areal topography dataset* or a *profile topography dataset*. Without specifiers, typically used as a short form for *areal feature*, i.e. a region of interest within an *areal topography dataset*.

Feature characterisation The act of describing and encoding information pertaining shape, size, position, orientation of a *feature* of interest.

Feature extraction Within an existing *areal topography dataset* or *profile topography dataset*, the act of separating the data subset corresponding to a *feature* from the original topography, and encoding it as a separate, independent dataset.

Feature identification The act of recognising that a subset of a *topography dataset* represents an instance of a feature type (class) being searched for.

Segmentation The act of partitioning an *areal topography dataset* or a *profile topography dataset* into regions consisting of subsets of topologically connected locations. Each region is often referred to as *segment* and may or may not correspond to a *feature*.

Topography dataset A digital dataset containing information about the topography of a surface, i.e. the set of geometric formations of multiple sizes and shapes that form the surface. A topography dataset can be obtained by measurement, or generated algorithmically. An *areal topography dataset* encodes information pertaining a surface region. A *profile topography dataset* encodes information pertaining a cross-sectional profile obtained from the surface.

Chapter 1

Introduction

1.1 Background

In engineering we are interested in surfaces as they form the boundaries of our products. Surfaces define the accuracy and precision of the parts we manufacture, function is exerted through surfaces, components in assemblies interact through their surfaces, surfaces are the result of manufacturing processes applied to materials, and so provie information about them [1]. From a geometrical standpoint, surfaces are complex entities, comprised of intricate networks of formations spanning a wide range of scales of interest. In engineering measurement, the discipline of *dimensional metrology* is typically concerned with the largest scales of observation: dimensional metrology considers the shape of a surface as a whole, since the goal is the inspection and verification of a part geometry and dimensions [2]. Instead, *surface metrology* is concerned with the measurement of the smaller-scale formations existing on a surface [3], something which is often colloquially referred to as *roughness*. In this work, the term *conventional surface metrology* will be used to refer to surface metrology as it is commonly intended today; i.e. as a discipline concerned with the measurement and characterisation of surfaces in terms of their roughness. It will be shown that there are many situations where conventional surface metrology may not be adequate; this work covers such situations. This chapter deliberately introduces non-rigorous, but hopefully more intuitive, terminology, as the primary goal is here to illustrate some fundamental concepts and challenges. A more rigorous approach will be adopted from Chapter 2 onwards.

1.1.1 Basic assumptions of conventional surface metrology

Because of the nature of most engineering applications it has been historically involved with, the discipline of surface metrology has flourished in the measurement of surface formations covering μm and sub-μm scales. While in principle, many of the mathematical methods developed could be considered scale-independent, in reality, the majority of measurement instruments developed for surface metrology in the last decades are dedicated to μm and sub-μm scales of investigation. Even data processing has more or less implicitly evolved by considering the typical errors and artefacts generated at such scales by the available surface measurement technologies.

Surface metrology is governed by a few implicit but fundamental assumptions, which can be summarised as follows:

- *Nominal surfaces are regular and smooth*. That is, the surfaces imagined by designers/engineers when developing products, are (at the standard scale of observation) regular and smooth[1], and usually chosen from a small set of geometrical archetypes, such as plane, cylinder, etc. as they are easier to manufacture[2].
- *Manufactured surfaces are not as regular and smooth*. Fabricated surfaces differ from their nominal counterparts because manufacturing processes are not ideal, and introduce error at different scales.
- *Surfaces have a "lateral" and "vertical" direction*. Surfaces are assumed as geometric entities that spread over large areas, but have comparatively smaller extensions orthogonally to their definition areas. It is therefore convenient to imagine a local coordinate system that defines a "vertical" direction as orthogonal to the surface, and a "lateral" direction, directed anywhere as long as locally tangent to the surface.
- *Manufacturing error has "lateral" and "vertical" components*. When addressing the effects of manufacturing error on surfaces, it is convenient to imagine each departure from nominal (i.e. error) as characterised by a *lateral extension* (i.e. oriented along the surface) and a *vertical extension* i.e. oriented across (outwards or inwards) the surface. Surface error can then be classified based on extents (i.e. size/scale) in the lateral and vertical direction.

[1] Advanced applications may involve textured surfaces, i.e. surfaces where irregularity has been engineered for a specific purpose. However, these should be considered as special cases.

[2] with the notable exception of free-forms.

- *Dimensional metrology addresses large scale error* (either lateral, vertical or both), because large-scale error affects the shape and size of a part or product[3].
- *Surface metrology addresses small scale error* or better put, it addresses any surface error which is too small to have a significant effect on shape and size of the part or product[4]. Because of historical considerations on conventional manufacturing processes, small-scale surface error is inspected at μm and sub-μm scales, over surfaces of widths in the *mm* or larger scale.
- *Large-scale surface error must be taken out of the equation for surface metrology to operate*. In order to provide the measurement accuracies and resolutions needed to operate at small-scales, measurement technologies are often limited to small operational ranges. For this reason, surface metrology is often constrained in applicability to surface regions de-voided of large-scale error, and any residual source of large-scale error in measurement data should be removed before characterisation (see Chapter 2).

With large-scale error out of the way, surface metrology operates under two main assumptions:

- *Small-scale error is uniformly spread over the surface*.This is again an assumption existing for historical reasons, based on considering how most conventional manufacturing processes operate. When upon observation, this may not be the case, surface metrology must provide further methods to remove any source of disuniformity, in order to re-establish a "uniform" distribution of the irregularities.
- *Small-scale error is inherently random*. Meaning that it is difficult to predict its value at a location, even when knowing it at other locations on the same surface.

For simplicity of illustration, in this chapter the condition where the listed assumptions are verified, will be referred to as conforming to the *uniform random roughness* paradigm. More accurate terminology and a mathematical explanation of this surface state will be provided in Chapter 4.

[3]Note that a common occurrence for large scale error is to be large "laterally" (i.e. extended over the surface), but very small "vertically" thus larger-scale error does not necessarily imply that less accurate instruments are required.

[4]the distinction can fail to properly represent a number of circumstances, such as in the measurement of very rough parts. However, it has a historical basis in the way many conventional manufacturing processes behave (i.e. they can produce very large and accurate surfaces where irregularities are at least one order of magnitude smaller, hence not affecting shape and size of the part)

The surface shown in Figure 1.1 represents a typical example of product whose surfaces, if taken individually, largely conform to the "uniform random roughness" paradigm. The part has been fabricated by sand casting. The nominal shapes of its surfaces are all easily recognisable as instances of planes, cylinders, etc. Because of manufacturing, there are shape/size errors; these errors are of concern of dimensional metrology. Surface metrology on the contrary is concerned with looking at the smaller scale formations on the surfaces. Within a smaller observational window, which can be imagined as covering a small portion of one of the surfaces of the part, a coarsely uniform pattern of formations at much smaller scales is visible. In this case, the pattern also appears quite random, since it comes from the surfaces of the expendable mould used to fabricate this part.



Figure 1.1: Part manufactured by sand casting. Its surfaces conform to the uniform random roughness paradigm.

Many of the choices that have been made historically in regards to how surfaces are measured, and how data are analysed, are strongly rooted in the conceptual paradigm of uniform random roughness. For example:

- *Surface measurement is typically performed by point-based sampling at constant spatial density* [4]. Which is something one would do either when there is not enough knowledge of the surface to inspect, or when uniformly spread, random irregularities are expected on the surface.
- *Data pre-processing aims to enforce the uniform random roughness paradigm.* Many of the procedures of form-removal and data filtering (see chapter 2) are designed to remove components incompatible with the uniform random roughness paradigm.
- *Surface characterisation by texture parameters relies on uniform random roughness to*

produce meaningful results. If the uniform random roughness paradigm can be enforced, then it becomes possible to characterise small-scale surface irregularities by using summary indicators that capture statistical properties of such irregularities. Most surface texture parameters defined by the ISO standards and commonly used in surface metrology are designed to this purpose. Amongst these, probably the most famous one is the mean arithmetic roughness *Ra*. Texture parameters will be discussed in detail in Chapter 2.

The assumption of there existing a scale of observation where surface irregularities can be seen as a uniform, evenly distributed noise, does indeed hold in many engineering applications where surface metrology has been historically needed the most. Many engineering surfaces exert their primary function through their overall shape, acting as the functional boundary of the part, for example acting as the interface between a part and its surroundings or other parts. Smaller scale formations on the surface do also matter, but not necessarily as much as a part geometry and dimensions. So the reasoning that smaller scale irregularities usually only matter in a cumulative sense (i.e. not individually) if often legit. So for example, friction, or any other tribological property, will be mostly determined by a multitude of individual, participating topography features, whilst single ones would not bear any particular relevance when considered alone. This reasoning is also consistent with the way parts are manufactured: typically the machining, forming or casting tools will focus on the realisation of a part shape and dimensions, and surface roughness is almost seen as an unavoidable side-effect[5].

1.1.2 Application scenarios where conventional surface metrology may be inappropriate

The conventional way of operation of surface metrology (i.e. a) processing measured surface data in order to isolate those components consistent with the uniform random roughness paradigm, and then b) summarise said data through summary indicators designed to be meaningful for surfaces conforming with the uniform random roughness paradigm), fails to be applicable in a series of circumstances, typically associated to specific types of application scenarios. Such scenarios will be colloquially referred to as "non-conventional" in the following, and they are the main subject of interest for this research work. Such non-conventional scenarios typically imply the presence of one of the following surface types:

[5] "almost" since there are also cases where irregularities are manufactured on the surfaces on purpose, to achieve some functional effect.

- multi-modal surfaces with textural transitions,
- surfaces where localised topographic formations (surface features) are of interest,
- multi-faceted surfaces resulting from the aggregated measurement of multiple smaller faces belonging to micro-parts and micro-assemblies; and
- surfaces featuring structured patterns.

Multi-modal surfaces with textural transitions

Some surfaces may feature multiple regions, each exhibiting a distinct organisation of surface irregularities. Such regions may be discriminable, for example because of visual appearance (texture, colour). Sharp or gradual transitions may exist between regions. Such surfaces are here referred to as having bi-modal, or multi-modal textures (e.g. see Figure 1.2); where each texture corresponds to a different state where the uniform random roughness paradigm is respected.

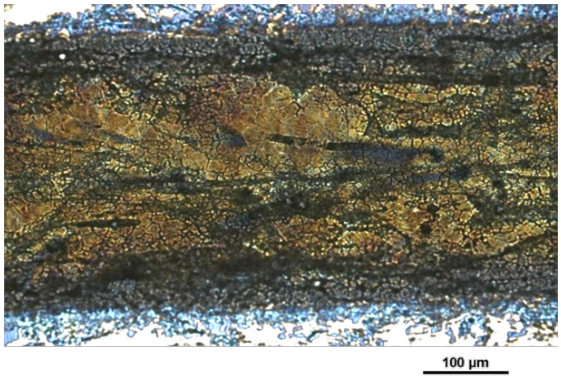


Figure 1.2: Bottom and side surfaces of an electrolyte jet machining (EJM) groove as seen from an optical microscope. The surface features multi-modal texture (presence of regions with distinct appearance). Sample courtesy of: J.Mitchell-Smith and A. Clare, University of Nottingham.

When multi-modal surfaces are encountered, it is often the case that all or at least a few of the "modes" are relevant to the characterisation. For some surfaces, even the transitions between the textural "modes" may have a significance and thus may need quantification (e.g. see Figure 1.3).

The most widely accepted method to characterise multi-modal topography is to handle each "mode" separately. In other words, regions must be carefully separated from each other so that, within each, the assumptions of uniform random roughness are verified enough to support the application of conventional investigation means (i.e. the computation of texture parameters).

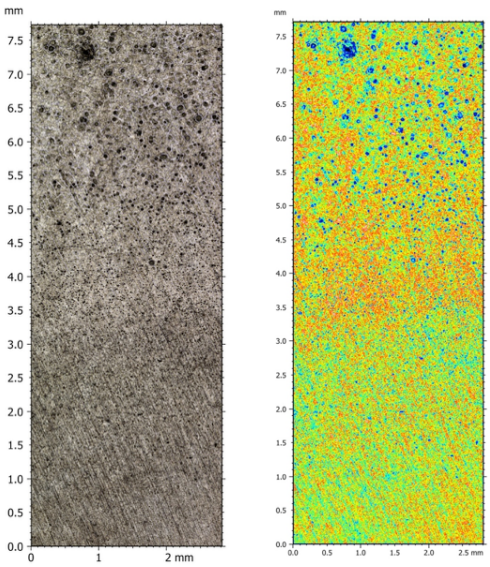


Figure 1.3: Transition between electron-beam melted region (above) and unaffected surface region (below). Where is the boundary between the two regions? Sample courtesy of J. Murray and A. Clare, University of Nottingham. Image obtained from the RGB overlay of a focus variation areal topography measurement; left) original grey scale image; right) image in false colours.

If the interest is specifically in the transition (e.g. see Figure 1.3), then there is currently no established solution for handling the characterisation task, except for qualitative inspection, or non-standard methods (e.g. texture parameters computed within moving windows - see Chapter 6).

Surfaces where localised topographic formations (surface features) are of interest

The second relevant class of applications, probably the one addressed the most in this thesis, is related to the investigation of localised surface features. Whilst a proper definition of surface feature will be given in Chapter 4, for the time being, it is sufficient to consider a localised surface feature as a region encompassing a topographic formation of interest located within the measured topography. Example applications belonging to this scenario are:

- inspection of damages or defects (e.g. see Figure 1.4),
- investigation of manufacturing process signature features (e.g. see Figure 1.5), and
- inspection of marks left by mechanical testing or functional life (e.g. see Figure 1.6).

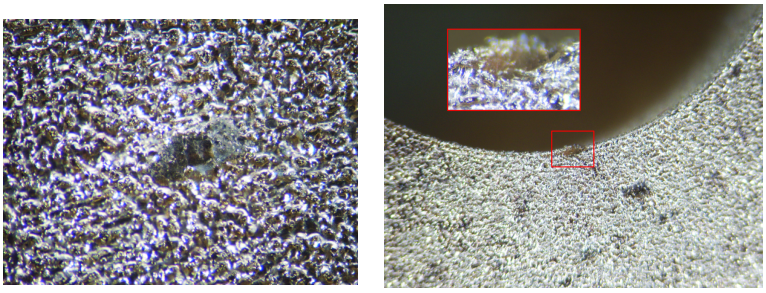


Figure 1.4: Localised defects on laser powder bed fusion surface (laser metal fusion by Sisma SpA); left) inclusion; right) detached edge. Images from SLR camera, courtesy of M. Moretti, University of Perugia, Italy.

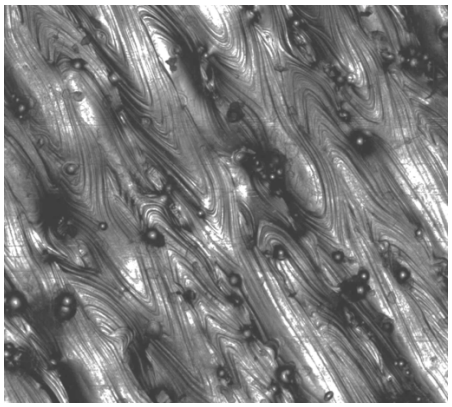


Figure 1.5: Top surface of a layer in laser powder bed fusion. Spatter particles and weld tracks (signature features) are visible. Image from the RGB layer of a confocal laser scanning microscope. Sample manufactured with a Renishaw selective laser melting machine (AM250).

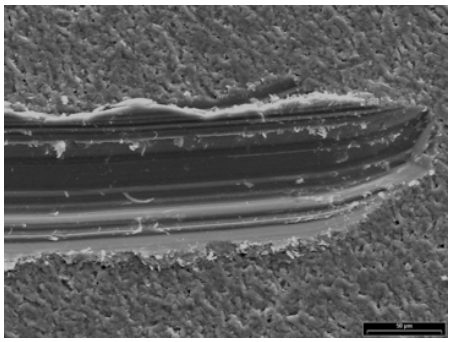


Figure 1.6: Groove obtained by scratch testing. Image from scanning electron microscopy. The groove is approximately $50\mu m$ wide.

Clearly the application scenarios introduced in this section and involving interesting localised topographic formations (surface features) are intimately related to the previously described application scenarios, involving multi-modal textures (as described in the previous section), since both scenarios in the end are about the identification and characterisation of regions of interest. However, depending on the contents and prop-

erties of the regions, different routes may be pursued for characterisation, as will be discussed in Chapters 4, 5 and 6.

Surfaces resulting from the aggregated measurement of multiple smaller faces belonging to micro-parts and micro-assemblies

As discussed later in Chapter 2 (Section: 2.4), it has recently become convenient to inspect micro-parts and micro-assemblies using the same measurement instruments (profilometers or microscopes) that would normally be applied for surface texture characterisation. Example applications involve the inspection of MEMS (e.g. see Figure 1.7.a), semiconductor components, microfluidics components (e.g. lab-on-a-chip, see Figure 1.7.b), and many other micro-parts. Characterisation of the individual faces in terms of their roughness (assuming the uniform random roughness paradigm can be enforced) may have limited uses, whilst the most requested application is the inspection and verification of dimensional and geometric properties.

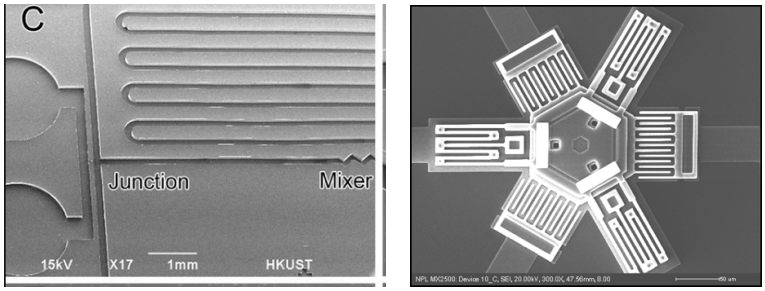


Figure 1.7: Deterministic micro-parts and micro-features; left) channels for microfluidics; right) micro electro-mechanical system (MEMS). Images from scanning electron microscopy. Courtesy of the National Physical Laboratory, Teddington, UK.

Again it is clear that even this latter scenario shares most of the fundamental issues identified for the previous two. Once again the challenge is that regions of the original surface must be isolated (in this case corresponding to the individual surfaces of the micro-part or assembly to be inspected), and each surface may need to undergo a different treatment.

Surfaces featuring structured patterns This latter scenarios encompasses those surfaces that are comprised of multiple, engineered features organised in regular layouts (structured patterns), usually designed with the explicit purpose of achieving some functional effect. For example in Figure 1.8 (left), a pattern of laser-machined micro-pockets has been fabricated to reduce the coefficient of friction in bearing applications. Figure 1.8 (right) shows an example of retro-reflector surface. Evidently, structured pat-

terns feature regularity and periodicity aspects that cannot be captured by texture parameters provided by conventional surface metrology, as the uniform random roughness paradigm does not hold. In particular, while the uniformity may be respected, the randomness typically is not. As for the previous scenarios, also this one implies the need for partitioning the topography, so that the pattern units can be isolated and investigated, and properties related to the regularity and layout of the entire pattern can be studied.

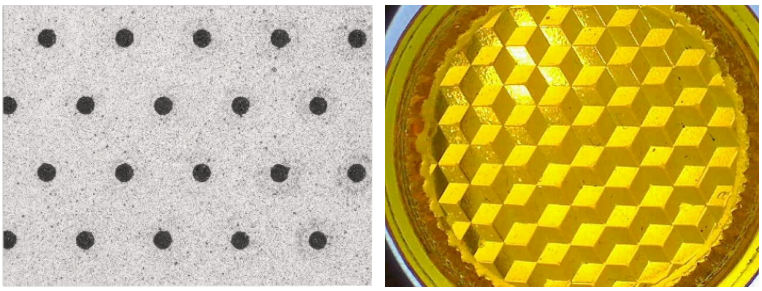


Figure 1.8: Example surfaces featuring structured patterns; left: laser-machined micro-pockets to reduce the friction coefficient in bearing applications (from [5]); right: retro-reflector (from:http://academic.greensboroday.org/ regesterj/potl/Waves/Reflection/ReflPyr10xB.jpg")

1.2 Main aims and objectives

The main aims of the research work reported in this thesis are: a) the development of a set of flexible methods to support feature-based characterisation of surfaces, and b) the establishment of a comprehensive framework, providing the theoretical foundation to support and categorise the methods of feature-based characterisation, suitable to cover all the application scenarios previously introduced as "non-conventional". The term "feature-based characterisation" is here used to refer to either the characterisation of a surface in terms of its topographic features, or the characterisation of one or multiple features themselves.

In detail, the research objectives addressed in the doctorate are the following:

- Identify a mathematical formalism and associated digital representation of surface topography which can handle both conventional and non-conventional (i.e. feature-based) characterisation approaches.
- Identify the conditions of applicability of conventional characterisation methods (i.e. based on texture parameters) in formal/mathematical terms, and describe conditions of non-applicability.

- Introduce a formal definition of *feature* and *feature-based characterisation*, and formalise the phases of the feature-based characterisation process, providing a consistent definitional and taxonomic framework that can accommodate existing and new methods dedicated to each phase.
- Develop new methods dedicated to the various phases of feature-based characterisation, compensating and complementing the existing ones, and define how they fit within the framework.
- Demonstrate the methods and the overall framework of feature-based characterisation through the application to a selected set of test cases.
- Provide a critical overview of accomplishments as well as pointers to currently unsolved challenges.

1.3 Outline of this work

In addition to this introductory chapter, the remainder of the thesis is organised as follows.

- *Chapter 2* is dedicated to a review of current state of the art on surface metrology (i.e. the "conventional" approach to surface topography characterisation). The underlying assumptions about the nature of the investigated topography are discussed (i.e. uniform random roughness) and their influence on the meaning and applicability of the conventional surface analysis pipelines is illustrated. In the last part of the chapter, current literature related to how some of the "non-conventional" application scenarios have been handled so far is also reported. In fact, although this work is the first example of an attempt to establish a general framework that transcends the conventional approach to surface metrology, and actually tries to encapsulate it within a broader picture, other researchers in various capabilities have address and solved case-specific problems. These efforts are also reported in Chapter 2.
- *Chapter 3* introduces and discusses an original mathematical formalism capable of handling topography for conventional characterisation, as well as topography of individual features resulting from feature-based characterisation. The chapter begins by discussing the types and nature of datasets that can be obtained by using the measurement instruments currently available for surface metrology. As this work is not about measurement, but rather it focuses on what do to with measured data once it has been acquired, the descriptions of the actual measurement

technologies are be limited to the minimum necessary to understand why topography data is available in the current formats. A discussion on the properties and peculiar aspects of the available datasets is important to set the boundaries of what can and cannot be done, and helps better understanding the challenges of topography data analysis and processing presented in the following chapters. The chapter then proceeds by introducing a formal representation of topography that can handle both entire surfaces, and individual features.

– In *Chapter 4* several foundational elements of the proposed, original framework for feature-based surface topography characterisation are presented. The meaning of *determinism and randomness* for surface topography is discussed. Elements pertaining the theory of *random fields* are introduced to discuss the fundamental problem of how to conjugate randomness and geometric information. *Topography decomposition by spatial partitioning* is also presented as the original means to handle the complexity which typically accompanies topography information, and as the basic idea underlying the feature-based approach to the investigation of surfaces. The concept of *topography feature* is given a formal definition.

– *Chapter 5* is dedicated to presenting and discussing the original contributions provided by the doctoral work for *feature characterisation*, one of the fundamental pillars of the feature-based framework. The goal of characterisation is to find suitable ways to describe and encode information pertaining shape, size, position, orientation of surface features of interest. Characterisation is also fundamental to establish a way to compare feature instances. Characterisation solutions may be more or less suitable to handle features associated with different amount of randomness, thus multiple solutions are introduced and discussed in the chapter.

– In *Chapter 6* the original contributions produced in the doctorate for *feature identification and extraction* are presented. The characterisation methods discussed previously in Chapter 5 are revisited as fundamental means to implement *feature identification*. Once features have been identified, they may need to be extracted and handled as independent datasets; thus, *feature extraction* is also discussed in the chapter. The extraction moment can play also a key role in determining the accuracy of a feature boundaries, which in turn affect many of the characterisation processes aimed at determining size and shape properties of the feature. So, *feature boundary detection* and its influence on feature-based characterisation results is discussed as well.

– *Chapter 7* is dedicated to understanding how *measurement error* is associated to measured datasets and propagates through the feature-based characterisation

pipeline ultimately affecting the results of the analysis. In the chapter, original approaches to the formulation of *uncertainty budgets* starting from assessment of error components, developed during the doctoral studies, are introduced and discussed.

– *Chapter 8* is dedicated to presenting original methods for the generation of *artificial topographies*. The availability of a comprehensive and robust set of methods and technologies for the generation of artificial topographies with known properties is fundamental to validate any surface topography characterisation pipeline, including those belonging to feature-based characterisation. A set of flexible tools devised for this purpose during the doctorate is introduced and discussed.

– In *Chapter 9* the feature-based characterisation framework is demonstrated and validated through a selected set of *applications*. The applicability and advantages of feature-based characterisation are illustrated and discussed. The selected applications cover research and development of advanced manufacturing processes, high value-added micro-products and micro-parts, the mechanical and tribological characterisation of materials, and the study of surfaces after their functional life.

– *Chapter 10* is dedicated to the *conclusions*. The main results, and open issues, pertaining feature-based characterisation of surface topography are introduced and discussed. An outlook on future perspectives is also provided.

Chapter 2

Review of current literature

This chapter covers the methods and ideas that form the basis of what is currently referred to as the discipline of *surface metrology*. Most of the contents of this chapter come from the current ISO standard for surface topography characterisation: *ISO 25178-2 Geometrical Product Specifications (GPS) –Surface texture: Areal –Part 2: Terms, definitions and surface texture parameters* [6], though references are made to the previous and now superseded standard ISO 4287 *Geometrical Product Specifications (GPS) –Surface texture: Profile method –Terms, definitions and surface texture parameters* [7]. A comprehensive overview of both standards can be found in [3]. The role of this chapter is to provide just enough elements to illustrate the approach and the underlying assumptions that characterise surface metrology as it is currently known.

Elsewhere in this thesis, the ISO-specified approach to the characterisation of surfaces will be often referred to as *conventional surface metrology*, to distinguish it from the original methods discussed in the other chapters of this thesis.

2.1 Surface components and processing operations

The discipline of surface metrology is concerned with the measurement of the smaller-scale formations existing on a surface [3]. A typical surface measurement usually results in the acquisition of a relatively small portion of a larger surface (see Figure 2.1). The acquired region is here referred to as *measured surface*, or *measured profile* according to the old ISO 4287 standard. The measured surface/profile will naturally differ from the actual, underlying portion of the real surface it was acquired from, because of the non-ideal nature of any measurement process. The measured surface is therefore sometimes referred to as *mechanical surface* or *electromagnetic surface* depending onto whether it has been acquired through contact probing or non-contact (often, optical) probing [6,

8].

As stated earlier, the current standard ISO 25178-2 [6] is concerned with areal measurement (i.e. the result of measurement is a surface that can be imagined as defined by a function $z = f(x,y)$). Because of historical limitations of the early measurement technologies, the older standard ISO 4287 [7] addresses *profiles* instead (i.e. imagined as defined by a function $z = f(x)$). Despite the change in dimensionality (i.e. going from profiles to areal datasets) the main methods and ideas are shared between the old and new standard, therefore both approaches will be considered in the following.

The measured surface (or profile) can be seen as comprised of two main components (Figure 2.1). The first is *form*: a large-scale, underlying component that represents the nominal shape of the part in the region being measured; e.g. plane, cylinder, hemisphere and polynomial (for surfaces); or line, polynomial curves and splines (for profiles). The second component is *texture*: the collection of all the other smaller-scale formations that are visible on the surface/profile in addition to form. Form and texture together are often referred to as *topography* [3, 6].

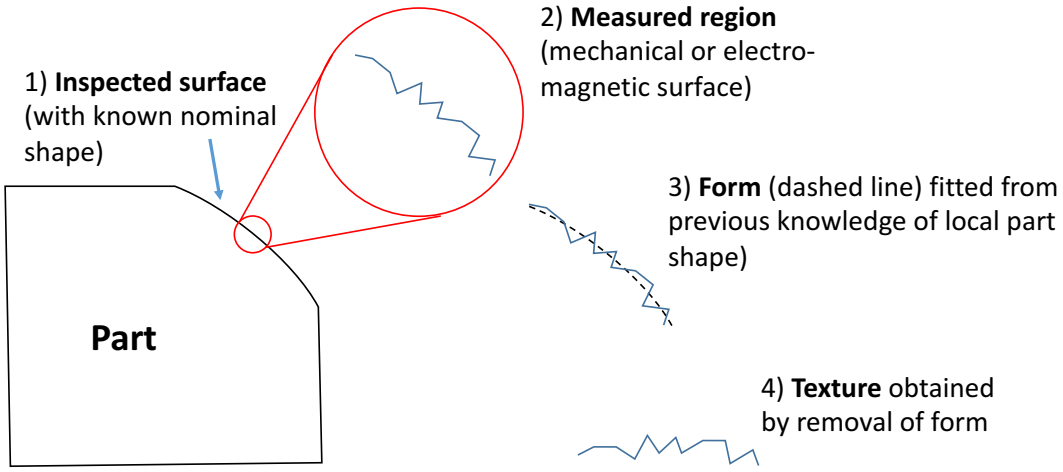


Figure 2.1: Surface components (shown as profiles for better clarity)

The old standard ISO 4287 [7] also further divides texture into *waviness* (i.e. larger-scale texture) and *roughness* (i.e. smaller-scale texture). This differentiation has now been abandoned because forcing a binary subdivision of scales may not always be appropriate and more freedom is left for users to decompose texture information into as many scales as they see fit. However, the term *roughness* in particular has remained as a widely adopted colloquial term, and will be often used throughout this text to refer to "small-scale" topography formations.

According to both the old ISO 4287 and the current ISO 25178-2 standards, once acquired, the measured surface/profile must be subjected to a series of processing operations summarised in Figure 2.2:

- The application of an *F-operator* to remove form (form removal would result in the *total profile* in the old ISO 4287). The F-operator has two additional effects: it removes tilt (i.e. levelling error due to sample placement on the instrument) and it sets the mean z value to zero[1];
- the application of an *S-filter* to remove texture at very small scales, assuming that at such scales, texture is only the realisation of measurement instrument noise. This operation results in the *S-F surface* (*primary profile* according to the old ISO 4287);
- the application of an *L-filter* that splits texture components into two groups based on a reference scale threshold, essentially replicating the old ISO 4287 method of dividing waviness and roughness components. The small-scale part is referred to as the *S-L surface* (*roughness profile* in the old ISO 4287).

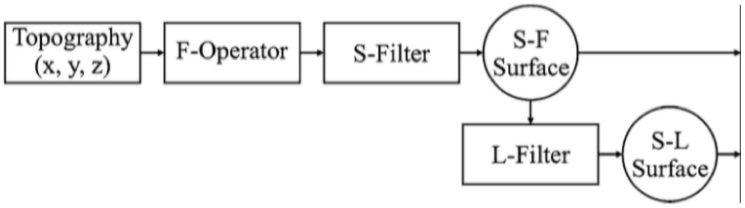


Figure 2.2: Data processing pipeline of conventional surface metrology (from [3])

The F-operator is often implemented by total least-squares fitting of the surface to a known analytical model (e.g. a plane, a polynomial surface, etc.). Alternatively, if no prior knowledge of the underlying form is available, a second-order regression filter according to the standard ISO 16610-71 [9] can be adopted. The fitted model is then subtracted from the original surface to obtain a residual, form-deprived, surface. The S and L filters are usually implemented starting from a convolution smoothing operator (e.g. using a Gaussian or similar kernel to define the weights for a moving average - see the standard ISO 16610-61 [10]), or equivalently via fast Fourier transform in the spatial frequency domain. Since convolution smoothing behaves as a low-pass, for the *L-filter*, which must be a high-pass, the residual after subtracting the filtered result from the *S-F surface* is kept. A common alternative to the conventional Gaussian filter defined in ISO 16610-61 [10] is the robust Gaussian filter defined in the standard ISO

[1] both effects being due to the application of a fitting operation, see the next point

16610-71 [9]. Robust methods are particularly useful when implementing the L-filter and challenging topography features, i.e. usually height outliers, are believed to be present.

Both the *S* and *L* filters necessitate setting values for the cut-off wavelengths (now referred to as *nesting indices* according to the standard ISO 25178-2 [6]). In case the conventional filters are applied, the cut-off correspond to 50% transmission, an important aspect to keep in mind because it implies that several undesired wavelengths are not entirely suppressed. Cut-off values depend on application. However, reference values can be found in the standard ISO 25178-3 [11].

2.2 Texture parameters

The usually untold reason for the data processing pipeline (F-operator, S-filter and L-filter) described above and in Figure 2.2 is that the measured surface/profile needs to be reduced to a form compliant with the "uniform, random roughness" paradigm introduced in Chapter 1. In Chapter 4 it will be demonstrated that pursuing compliance to such paradigm allows to lump all height observations (i.e. all the measured points of a profile or areal dataset) into a single random variable representing height. This random variable can then be described by computing sample statistics / estimating population parameters. Using such an approach, an arbitrarily complex topography can be reduced to a concise set of scalar, summary indicators.

2.2.1 Field parameters

The first class of texture parameters illustrated in this section is referred to as *field parameters* (as opposed to *feature parameters* illustrated later). For areal topographies (i.e. the new ISO 25178-2 standard) the complete denomination for this parameters is *areal field parameters* where the "field" is the entire area of the measured surface (usually referred to as *sampling area* or *definition area*). Equivalent parameters do exist for profiles, as defined in the older ISO 4287 [7]. Many of the texture parameters reported in the following sections have the phrase "*of the scale-limited surface*" added to their name, to explicitly indicate they must be computed on the dataset resulting from the application of the S and L filters. Areal field parameters are classified into categories based on what type of properties they capture from the surface. Some of the most used ones are discussed in the following. All the examples reported in the next sections have been artificially generated using Monte Carlo rejection sampling, the technique is illustrated in Chapter 8.

Height parameters

Height parameters capture properties of the probability distribution of heights, in most cases by replicating the definitions of basic sample statistics such as the 2^{nd}, 3^{rd} and 4^{th} order moments: *variance*, *skewness* and *kurtosis*[2]. Clearly, to be meaningful these parameters assume that the condition of "uniform random roughness" have been assured through the F-operator and the S and L filters, so that $z(x)$ or $z(x,y)$ has zero mean at any location where the surface has been measured. This means that height (z) can be seen as error with respect to an ideal $z = 0$ mean plane.

Root mean square height of the scale-limited surface - Sq The root mean square height is defined by the following equation:

$$Sq = \sqrt{\frac{1}{A}\iint_A z(x,y)\,dxdy}. \tag{2.2.1}$$

Sq quantifies the dispersion of heights about the mean ($z = 0$). As the equation shows, Sq is not really a variance, but rather a standard deviation (because of the square root). Such choice is justified by the advantage of quantifying dispersion on the same scale height values are defined at. In Figure 2.3 two surface topographies and related height probability distributions are shown corresponding to smaller and larger value of Sq respectively[3].

In reference to Figure 2.3 it should be noted that symmetric, Gaussian height distributions are shown for simplicity. However, they may not necessarily be, as standard deviation can be computed on any distribution.

Skewness of the scale-limited surface - Ssk . Skewness, the 3^{rd}-order moment, is defined as:

$$Ssk = \frac{1}{Sq^3}\frac{1}{A}\iint_A z^3(x,y)\,dxdy. \tag{2.2.2}$$

Skewness describes the symmetry of the distribution of heights about its mean. In surface metrology, skewness is normalised by dividing by Sq^3. $Ssk < 0$ indicates that the distribution is skewed towards height values above the mean, whilst $Ssk > 0$ indicates the opposite; $Ssk = 0$ indicates a symmetric distribution. In Figure 2.4 two topography

[2] the 1^{th}-order moment, the mean, is zero because of the F-operator (see earlier)

[3] Probability distribution as opposed to probability density function because the dataset is supposed discrete. Regardless of the Sq equation being provided in continuous form.

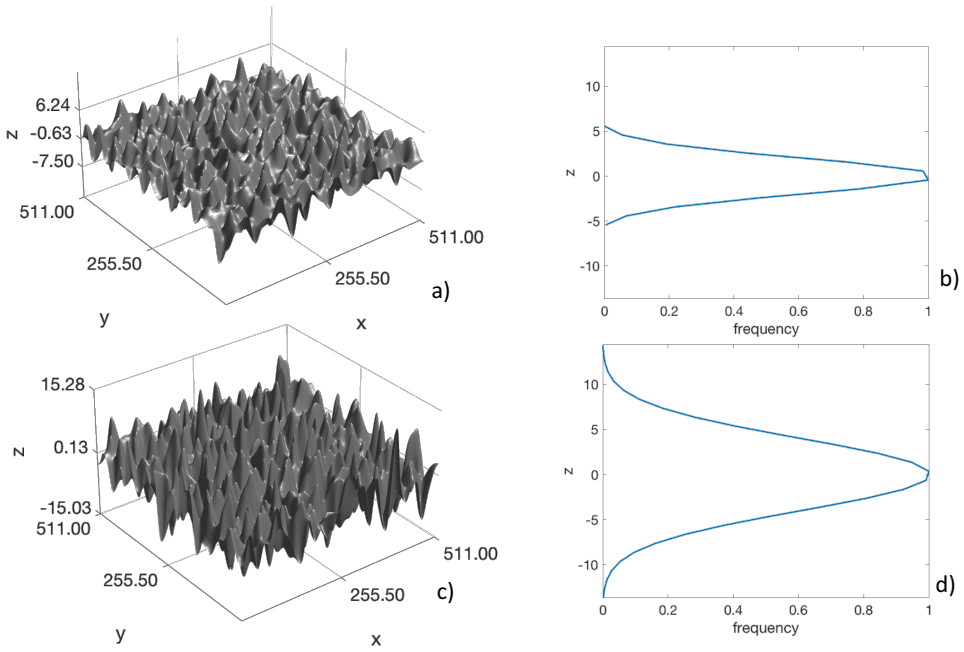


Figure 2.3: Sq values, probability distribution of heights and topographies; a) surface topography with $Sq = 2.0$ (reported as unitless for simplicity); b) corresponding height distribution; c) surface topography with $Sq = 4.0$; d) corresponding height distribution.

profiles and height probability distributions are shown corresponding to negative and positive skewness respectively[4].

Kurtosis of the scale-limited surface - Sku Kurtosis describes how much the shape of the height probability distribution differs from a Gaussian in being either spikier or bulkier. The equation of Sku is the following:

$$Sku = \frac{1}{Sq^4}\frac{1}{A}\iint_A z^4(x,y)\,dxdy. \tag{2.2.3}$$

With respect to the conventional kurtosis as defined in statistics, this one is normalised by dividing by Sq^4, so that again the values of Sq are the same order of magnitude of the range of heights in the probability distribution. Values of $Sku < 3$ indicate a bulkier distribution, $Sku > 3$ indicate a spikier shape whilst $Sku = 3$ corresponds to a Gaussian. In Figure 2.5 two topography profiles and height probability distributions are shown corresponding to kurtosis >3 and <3 respectively[5].

[4]The examples in Figure 2.4 were obtained using Chi probability distributions.
[5]The examples in Figure 2.5 were obtained using Cauchy probability distributions.

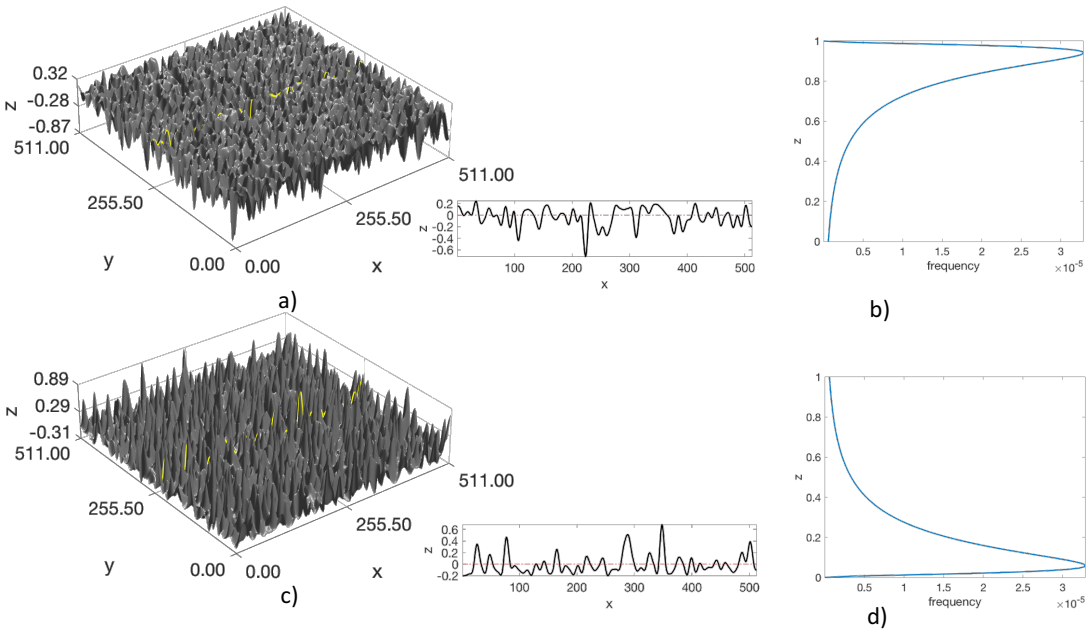


Figure 2.4: *SSk* values, probability distribution of heights and topographies; a) surface with $Ssk < 0$ and extracted cross-section profile (thin yellow line on the surface); b) corresponding probability distribution of heights; b) surface with $Ssk > 0$ and cross-section profile; d) corresponding probability distribution of heights.

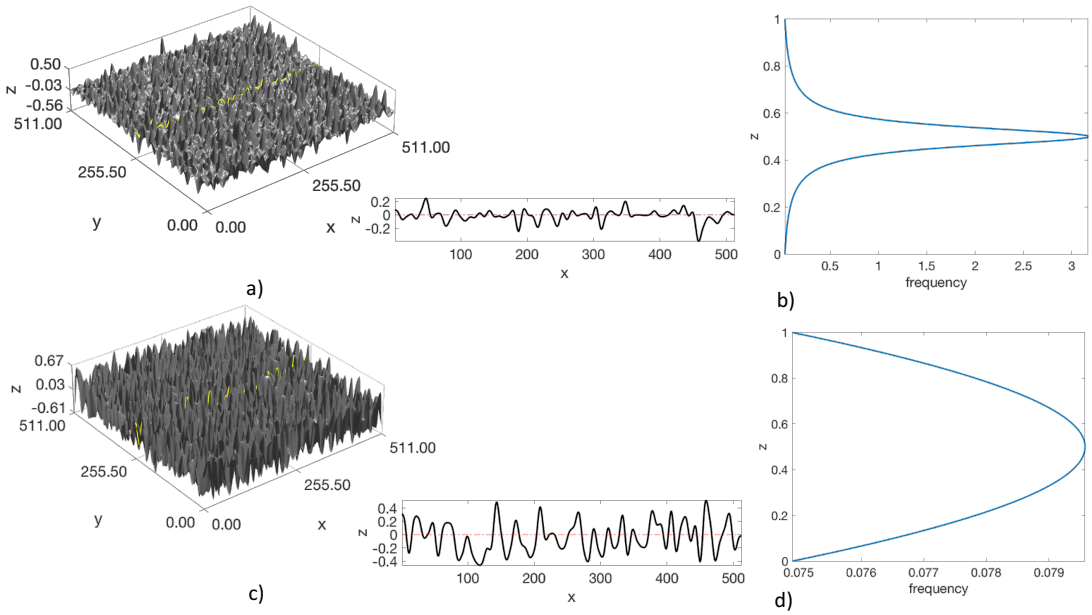


Figure 2.5: *Sku* values, probability distribution of heights and topographies; a) surface with $Sku < 3$ and extracted cross-section profile (thin yellow line on the surface); b) corresponding probability distribution of heights; b) surface with $Sku > 3$ and cross-section profile; d) corresponding probability distribution of heights.

Arithmetical mean height - Sa The arithmetical mean height is an alternative to Sq to quantify dispersion with respect to the mean. Its definition is:

$$Sa = \frac{1}{A} \iint_A |z(x,y)| \, dxdy. \tag{2.2.4}$$

What the equation means is that the original random variable describing heights is transformed into its absolute value (essentially transforming signed height values into distances from the mean), and then the 1^{st}-order moment (the mean) of the new distribution is computed. In short, Sa describes the mean error (i.e. distance) of the topography points with respect to $z = 0$. In Figure 2.6 the procedure for computing Sa is shown.

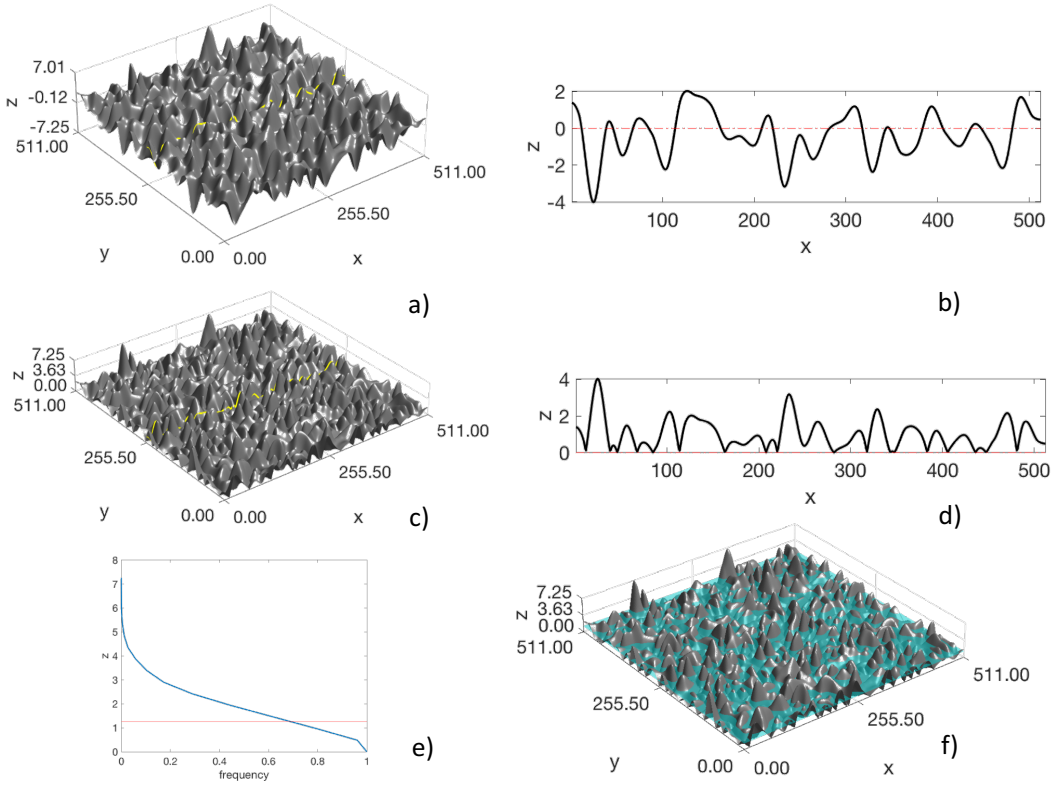


Figure 2.6: Procedure for computing Sa; a) original surface (with cross-section profile represented as a thin yellow line); b) cross-section profile from the original surface; c) absolute heights; d) cross-section profile from the absolute heights; e) probability distribution of the absolute heights (mean indicated in red); f) mean value plotted as a $z = const$ plane in transparent cyan and overlaid to the absolute heights surface.

Maximum peak height, maximum pit height and maximum height of the scale-limited surface - Sp, Sv, Sz These parameters capture the extremes of the probability distribution of heights. Sp represents the height of the highest peak, Sv the depth of

the deepest pit[6]. In truth, since there is no rigorous definition of what a peak or pit really are in the standard ISO 25178-2, the Sp and Sv parameters are commonly accepted as related to the highest and lowest height values in the dataset respectively[7]. Finally, the Sz parameter is defined as the sum of Sp and Sv (in absolute values, since Sv is negative), i.e. the range of height values covered by the probability distribution of heights. Sp,Sv and Sz must be used with care, as they essentially capture the extremes values of a random variable, so they are very sensitive to even single localised peaks or pits, including those cases when these may be due to measurement error. Sp, Sv and Sz become more reliable for those surfaces/profiles that have been subjected to low-pass filtering, due to the higher likelihood that spare peaks/pits having been removed.

Functions and related parameters

If the cumulative probability distribution is taken instead of the probability distribution of heights, then the shape of the former can be characterised through a new set of descriptors, i.e. a new set of field texture parameters.

The most common parameters are those described in the standards ISO 25178-2 and ISO 13565-2 [6, 12] and summarised by the geometric construction in Figure 2.7.

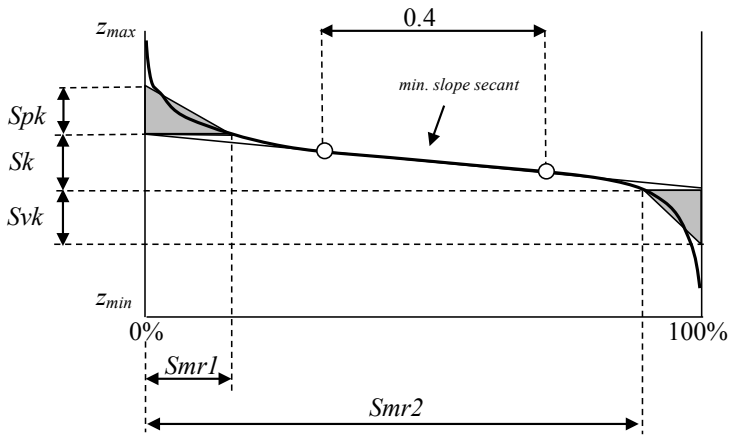


Figure 2.7: Cumulative probability distribution of heights (also known as the areal material ratio curve of the scale-limited surface [6, 12]) and geometric construction leading to the Sk parameter set (Sk, Spk, Svk, $Smr1$ and $Smr2$).

[6] Although it is called *maximum pit height* in the ISO 25178-2, it really is a depth.

[7] For profiles, the definition of peak and pit is easy as they are any local maxima and minima respectively. For surfaces the definition is a little trickier, because the definitions of local maxima and minima depend on how many neighbouring points are considered. The only definition provided by the ISO25178-2 for peaks and pits is used for feature parameters and is not necessarily applicable because it is based on continuous geometry.

The reason these parameters are typically called "functional" or "function-related" is that the cumulative probability distribution of heights can be seen as the percentage of surface area that is being sectioned by intersecting the topography with a $z = const$ plane at increasingly lower heights, starting from the top. As illustrated in Figure 2.8 the erosion of the highest peaks at high c values[8] recalls the "run-in" process experienced by surfaces designed to withstand mechanical contact, during the first period of their functional life (and is quantified by the parameter *reduced peak height - Spk*). The core region of the surface is what is supposedly lost due to wear during the bulk of a surface functional life (quantified by the parameter *core height - Sk*), while the remaining, deepest pits are what is left at the end, where residual fluids can be retained despite most of the surface having been worn off (quantified by the parameter *reduced valley depth - Svk*). Regardless of this specific interpretation, which is intimately tied to mechanical contact, the important point to make here is that Spk, Sk and Svk, together with the corresponding key points in the material ratio axis ($Smr1$ and $Smr2$) contribute to defining the shape of the cumulative probability distribution curve, and thus simply provide a different perspective on the same probability distribution of heights.

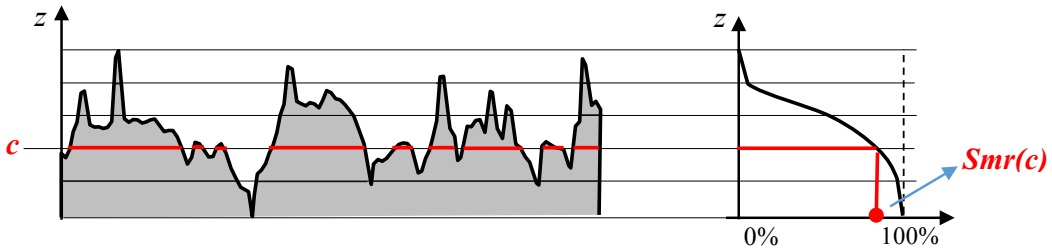


Figure 2.8: $Smr(c)$ indicates the percentage of surface area that have been cut by the sectioning plane at height c (sectioning process shown on profile for better readability).

Considerations on the parameters seen so far

For the texture parameters illustrated so far, it does not matter how height values are spatially organised. Whether they are coming from a profile or a surface, height values are simply binned all together, and considered as observations of the same random variable. This is why height and function parameters are essentially the same thing regardless of whether they have been computed on a profile or surface, so long as equivalent preprocessing (form removal and filtration) is applied. In other words, these parameters do indeed rely on the *uniform random roughness* assumption (see the example in Figure 2.9).

[8]The letter c is used to indicate sectioning height in the ISO 25178-2 [6].

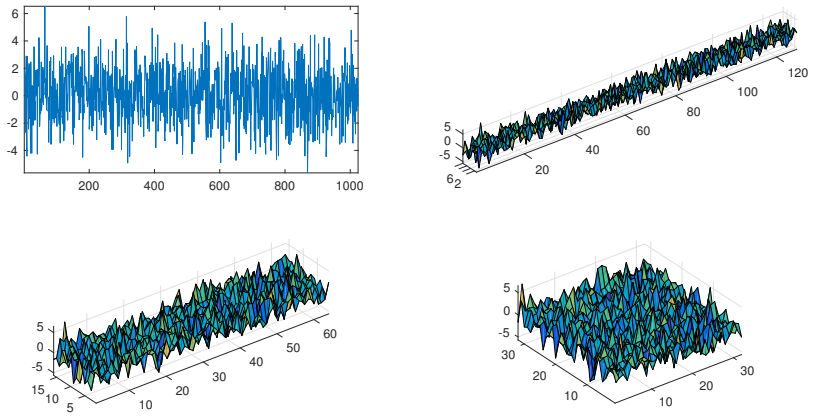


Figure 2.9: The profile and three surfaces in the figure have the same total number of height values and the same x, y point spacing. For each instance, the height values have been randomly rearranged in space, and ultimately lead to the same Sq and Rq values (assuming the same preprocessing conditions).

On the contrary, the parameters discussed in the following sections assume there may be some form of spatial dependency, in particular between close-by or even neighbouring data points. The assumption of "random" topography is relaxed, because spatial correlation often implies dependency[9]. Whilst the randomness assumption is relaxed, the assumption of "uniform" distribution of height-related properties is not. Without this latter assumption, most of the following parameters lose meaning.

Spatial parameters

The following parameters are based on characterising the nature and extents of any spatial correlation possibly existing between close-by data points. Spatial correlation is addressed on both the x and y directions, thus the following parameters are not applicable to profile data. The existence of spatial correlation between height values is captured by the (areal) auto-correlation function $ACF(tx, ty)$ (see Figure 2.10 (top right)). The function is obtained by computing the correlation coefficient between the original surface, and an identical copy that is shifted by tx, ty in the x and y directions respectively. At $tx, ty = (0,0)$ the correlation coefficient will be maximum, because of the identity of the overlaid surfaces (see Figure 2.10 (top right)). As tx and ty increase, correlation will drop, unless the surface has a preponderance of self-similar regions. The *Surface autocorrelation length - Sal* parameter captures how larger than (0,0) tx and ty need to be before the autocorrelation coefficient decays by more than a given percent

[9] Although, strictly speaking, correlation is a necessary but not sufficient condition to prove dependency.

of its maximum value. This is usually computed by truncating the ACF at the chosen reference height, and then by computing the axes of the equivalent ellipse on the truncated region (see Figure 2.10 (bottom))[10]. On the same truncation value (i.e. reference drop), the *Texture aspect ratio - Str* parameter specifies the ratio between minimum and maximum autocorrelation lengths ($tmax$ and $tmin$ in Figure 2.10 (bottom)). Str (being a ratio) indicates whether the maximum autocorrelation length is much larger than the minimum autocorrelation length, implying texture anisotropy (the surface is more deterministic in one direction).

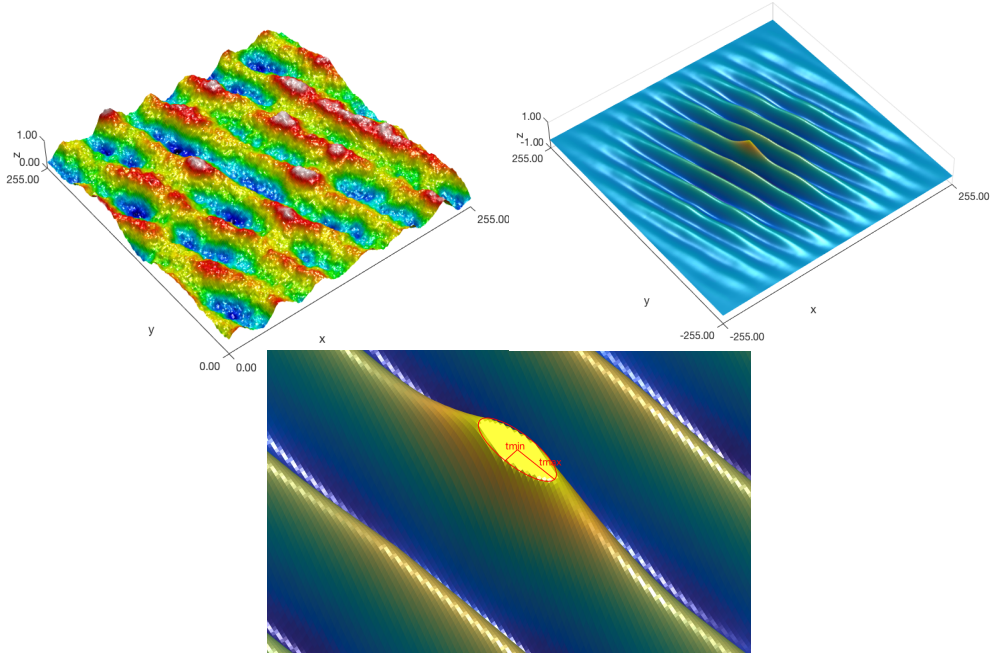


Figure 2.10: Procedure for computing spatial parameters; top left) original topography; top right) ACF (the centre shows the peak corresponding to the 0,0 shift between the surface and its copy); bottom) central peak truncated. Maximum and minimum shifts ($tmax$ and $tmin$) computed on the equivalent ellipse). The Sal parameter corresponds to $tmax$. Str is $(tmin)/(tmax)$

Miscellaneous parameters

The only texture parameter classified by the ISO 25178-2 [6] as miscellaneous is *Texture direction - Std*. Std indicates the direction that contains the higher amount of "textural content" (in simple terms, the largest amount of wavelengths and content at each wavelength). To obtain Std, the magnitude of the complex 2D FFT of the surface is

[10]The function $ACF(tx,ty)$ is typically normalised by dividing by Sq so that the final result is normalised between [1,-1]. This way it is easier to express the reference drop corresponding to the truncation value.

computed first (see Figure 2.11-right)[11]. Then the spectrum is subjected to angular integration with respect to its centre (see Figure 2.11-right). The value of the integral at each direction corresponds to the amount of "content" in that direction. The direction corresponding to the max amount of content is chosen as *Std*. Typically, for a wave-like surface, such direction is orthogonal to the main elongation of the crests (i.e. the direction of *tmax* in the truncated ACF - see previous figure 2.10). What is important to point out, is that in order to talk about prevalent direction of textural content, some degree of spatial correlation between surface points is necessary. A topography exclusively made of pure random, independent height values, would result in a fairly flat FFT magnitude spectrum, and thus originate a uniform angular spectrum, with no clear indication of a local maximum.

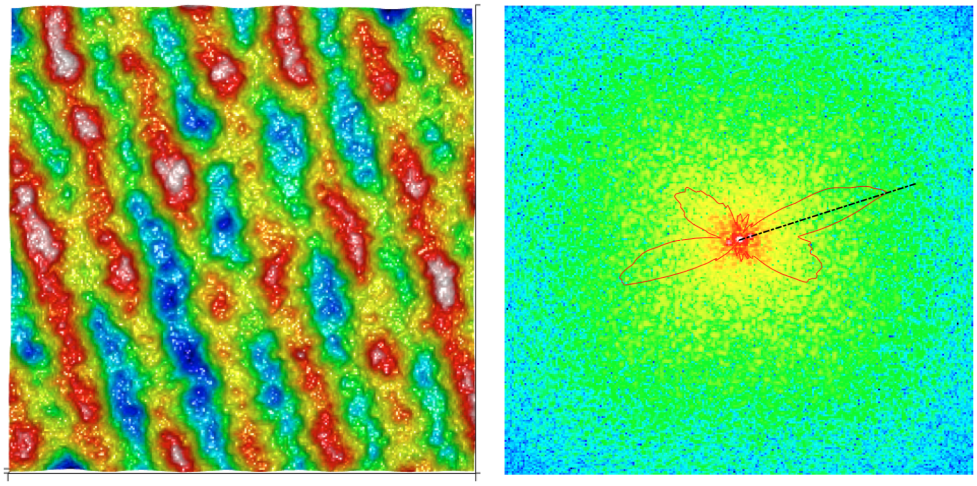


Figure 2.11: Procedure for computing the *Std* parameter; left) original topography (same as previously shown in Figure 2.10); right) magnitude spectrum of the complex 2D FFT with inverted quadrants (visualised in log form). The result of the angular integration (rescaled) is overlaid on top, the direction corresponding to *Std* (maximum value of the angular spectrum) is also indicated.

.

In summary, as anticipated earlier, for spatial and miscellaneous parameters, the assumption of random texture is not necessary any longer, and actually, the parameters are designed to capture aspects that specifically relate to the existence of spatial correlation. However, the assumption of uniform texture is still needed: the presence of a strong deterministic component making texture non-uniform in the x,y space rapidly deprives the ACF and PSD functions of any meaning, thus rendering *Sal*, *Str* and *Std* useless.

[11] Usually the spectrum is plotted with inverted quadrants, so that the contents at larger wavelengths are at the centre.

Hybrid parameters

Hybrid parameters are "hybrid" because they are based on slope, which is a combination of height and "spatial" properties. The two hybrid parameters defined in the standard ISO 25178-2 [6] are the *root mean square gradient of the scale-limited surface: Sdq* and the *developed interfacial area ratio of the scale-limited surface: Sdr*. The former is essentially a measure of mean local slope, the second describes the ratio between the area of the topography and the projected area on the x,y plane. The ratio is of course larger for surfaces with a large amount of high-sloped regions. Both parameters are based on averaging processes involving local slope information, and would not make sense if slopes were not uniformly spread over the surface. Once again, the "uniform" roughness assumption holds.

2.2.2 Feature parameters

Height and functional parameters assume complete independence of surface heights (complete randomness). It is irrelevant where each height point is located, and there is no problem with lumping all height observations into a single bin, in order to treat the entire surface as multiple observations drawn from the same random variable. These parameters fully rely on the uniform, pure random texture paradigm.

Spatial, miscellaneous and hybrid parameters do take into account relative localisation of height values, either by investigating spatial correlation (spatial parameters), or by investigating directions with most wavelength content (miscellaneous parameters), or by considering local slope (hybrid parameters). Spatial, miscellaneous and hybrid parameters drop the "pure random" assumption (because spatial localisation has relevance). However, they still rely on the "uniform topography" paradigm, meaning that they rely on texture properties being uniform over the investigated region.

A new family of ISO 25178-2 parameters, known as *feature parameters* [6, 13] introduce an innovative concept: height points should not be considered individually, but rather as parts of larger aggregates. Such aggregates, each covering a topologically connected region within the measured area, are called *areal features*[12]. According to the ISO 25178-2 [6] two types of areal features are possible: *hills* (regions converging to the same local maximum) and *dales* (regions converging to the same local minimum). The two features are rather elementary, but because of this, simple to describe, and very robust, as they can be found pretty much in any topography[13]. In principle, any topography

[12]The standard ISO 25178-2 defines also point and line features, but their discussion is not covered here.

[13]Not entirely true in case of plateaus and because of issues with the discrete nature of measured areal

can be either partitioned into hills or dales, as shown in the example in Figure 2.12

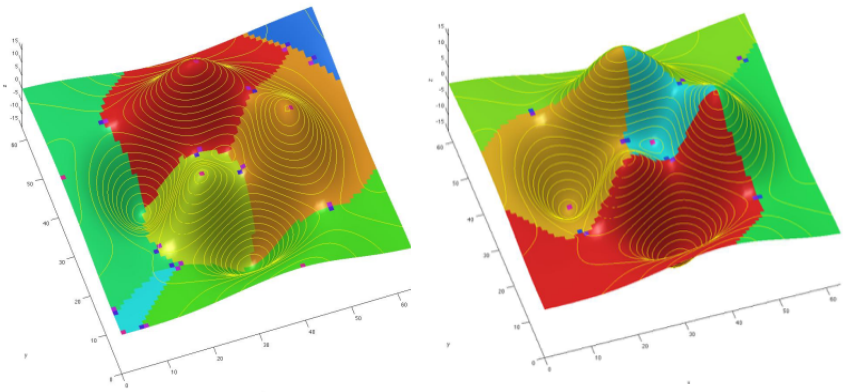


Figure 2.12: The same example topography partitioned into *hills* (left) or *dales* (right). In addition to the areal features (hills and dales, each instance rendered with a different colour) the following point features are indicated as well: *peaks* at the top of each hill (left); *pits* at the bottom of each dale (right), *saddle points* at the intersection of ridge and course lines separating hills and dales (left and right)

Once the partitioning into hills or dale is done, instead of computing texture parameters starting from points, one can describe a surface in terms of the geometrical properties of its higher-level "features". For hills and dales, footprint area, maximum height, perimeter, aspect ratio, etc. can be defined. To characterise the entire surface, statistics can again be calculated (like in conventional parameters that aggregate individual height points), except that the statistics are this time computed on the attributes of the features.

For example, the *Sds - density of summits* feature parameter (ISO 25178-2 [6]) defines the density of hills located within the measured area. Its formal definition is in terms of the procedure that must be followed to compute it:

$$Sds = FC; H; \text{Wolfprune} : X\%; \text{All}; \text{Count}; \text{Density}. \tag{2.2.5}$$

In equation 2.2.5, *FC* means "feature characterisation"; *H* means: "hills" (partitioning into hills as opposed to dales); Wolfprune : X% refers to a hill simplification procedure where smaller hills are merged to larger ones (see ISO 25178-2 [6] for details); *All* means that all the hills found in the areal topography dataset should be considered (as opposed to discarding the boundary ones); *Count* means that the relevant feature attribute is the number hills within the measured region; *Density* means that feature count should be expressed in relation to the unit area.

topography datasets

Feature vs field parameters, and the "uniform random roughness" paradigm

It is clear from the previous sections that the standard ISO 25178-2 introduces a new dichotomy for surface texture parameters. On one hand there are the new *feature parameters*, i.e. parameters dedicated to summarising the properties of the selected features (hills or dales). On the other hand there are *field parameters*, i.e. all the parameters discussed earlier, dedicated to summarising the properties of the measured region (i.e. *the field*) through an investigation that directly targets its measured points (that is, with no higher-level information layers, such as hill/dale features).

Whilst field parameters do indeed rely on the "uniform" paradigm, and in some circumstances also on the "random" paradigm (height and functional parameters), as illustrated earlier, what is the situation with feature parameters? Feature parameters cannot rely on the "random" assumption, as the relative localisation of height points is indeed essential to obtain a partitioning into hills or dales. In principle, feature parameters may not need the assumption of "uniform", if ultimately they are designed to only consider an individual hill or dale. However, it turns out that most feature parameters do indeed rely on the "uniform" assumption, as they capture hill or dale properties that apply to the entire measured area.

2.3 Conventional surface metrology and non-conventional scenarios

In Chapter 1 a series of scenarios involving the measurement of surfaces, were introduced and classified as "non-conventional", as they would escape one or both the assumptions of "uniform" and "random" texture, as prerequisite for the successful computation of texture parameters; or, even if they did comply at least in part to those assumptions, would not satisfactorily be described by texture parameters, because of them needing dimensional characterisation instead. Specifically, these scenarios were:

- multi-modal surfaces with textural transitions,
- surfaces where localised topographic formations (surface features) are of interest,
- multi-faceted surfaces resulting from the aggregation of multiple smaller faces belonging to micro-parts and micro-assemblies, and
- surfaces featuring structured patterns.

After the discussion on the main field and feature parameters defined by the ISO 25178-2, presented earlier in this chapter, it is now clear that all the above "non-conventional" scenarios imply as a minimum a partitioning of the definition area, either to isolate the uni-modal regions of a multi-modal texture, or to isolate the topographic formations (i.e. "areal features") of interest[14], or to isolate specific surfaces of micro-parts/assemblies, or again to isolate the units of structured patterns. Once the relevant "features" are isolated, the needs for characterisation may vary significantly across applications. Presumably, a region within a multi-modal texture could be characterised through conventional texture parameters, as long as - within the region - the conditions of uniform (and possibly random) texture have been established (e.g. via the F operator and the S,L filters). Therefore it is imaginable that in a number of situations involving multi-modal textures, as long as partitioning is possible[15], it might be possible to reduce the characterisation of each individual region to a problem that can be dealt with conventional surface metrology solutions (i.e. texture parameters). However, for the other scenarios, different characterisation routes targeting the feature of interest would be preferable, more akin to what is currently done for dimensional and geometric verification at the macro-scale. For example, it is common to imagine that one may want to know localisation, area, depth/height, aspect-ratio or similar other dimensional/geometric properties from a feature such as a scratch or a pore, originated by functional interaction, or mechanical testing. If the feature has been engineered (e.g. a channel for a microfluidics application, a high aspect-ratio hole obtained by micro-machining, a component of a printed circuit board, etc.) then one may probably want to compare specific dimensions with reference design specifications, either coming from a technical drawing or from a CAD model.

The first thing that must be noted in relation to these new scenarios is that, whilst a great deal of literature, methods and instruments is available to approach surface texture characterisation under the uniform random texture paradigm[16], comparatively little has been done to address the non-compliant scenarios. Solutions available from the literature are sparse, pretty much dedicated to solving specific application scenarios and specific classes of surfaces, usually following a case-by-case approach with no attempt at producing a unified theory.

[14] Not necessarily a hill or a dale, recall the examples provided in Chapter 1

[15] or as long as it is possible to measure the regions separately

[16] not to mention the numerous standards that have been produced on the matter, some of which where cited previously, e.g. ISO 4287 [7], 4288 [14], 25178-2 [6], 25178-3 [11], etc.

2.4 Relationships with the dimensional/geometric inspection and verification of standard-sized parts

Except for the multi-modal texture scenario (which presumably requires texture parameters in the end, to characterise the different regions), one may wonder why the other cases should not be dealt with using established solutions for dimensional inspection and verification of standard-sized parts (see [15] for a review). After all, most issues and methods of dimensional inspection and verification should be basically scale-independent and an areal feature could be considered as any other part, located somewhere in space and in need for inspection. Even in terms of measurement instruments, at least in theory, the transition from the macro to the micro-scale should be possible, given that co-ordinate measuring machines (CMMs), the standard-de-facto for measuring geometry and dimensions of standard-sized parts and assembly in engineering, are now available in variants specifically dedicated to inspecting the micro-scale (i.e. micro-CMMs). The most notable example of this class of machines is the Zeiss F 25 (see [4] for an overview). Less extreme, but often equally viable solutions for inspection if the features to investigate are not too small, involve the use of high-precision, low-force probes mounted on regular CMMs (see [16] for an example application involving the inspection of high-aspect ratio micro-holes).

The problems with using a touch probe on a micro-scale surface feature are unfortunately non-trivial, to the point that such solution has not been particularly successful so far, and has not gained widespread acceptance. The main issues can be summarised as follows:

- The touch-based measurement process is generally driven by an inspection program that specifies where the probe has to go in order to sense the surfaces. Unfortunately, at small scales the uncertainty on where the surfaces are supposed to be can be high, because often the manufacturing process has poor control on the manufactured shape (hence, high variability)[17] and sometime it is also difficult to locate the small-scale feature with great accuracy, starting from a workpiece positioning procedure that necessarily operates at larger scales. High uncertainty on surface location makes it very difficult both to write an inspection program, and to interpret the acquired measurement result [18].
- Surface features may also often be very complex in shape. Imagine a irregular

[17] a typical example is laser micro-machining

[18] for example: it is more difficult to understand what nominal surface should be associated to any acquired point

crater resulting from fretting wear, or a scratch (elongated groove) due to surface damage or produced by mechanical testing: the shape of such features could be so complex that it may require a large number of measurement points to be fully characterised. Unfortunately, touch-probe CMMs are sequential, meaning that measured points can only be acquired one after the other[19], thus a measurement process would not only be very difficult to program, but may also take a very long time to execute.

– Finally, the mechanics of contact must be considered. At small scales, contact easily turns into adhesion, and the small and delicate tips of micro-CMMs are easy to break even under minimum loads. Stick-and-slip events are equally dangerous; also there is the non-negligible risk of the probe unexpectedly hitting the surface at unwanted moments, because of it moving very close to a poorly defined surface. All such issues make touch-based measurement of micro-sized surface features in most cases very hard to accomplish, and in some cases even infeasible.

2.5 Previous work on feature-based surface characterisation

The need to provide solutions for feature identification and characterisation in surface metrology is recognised as an important one to be addressed [17, 18], despite the ISO standards only suggesting morphologic segmentation (i.e. hill/dale partitioning) at the moment [6]. The first natural attempt beyond conventional morphologic segmentation consists of slight modifications of it. For example, as illustrated in [19]), for step-like surface features, a gradient map can be computed from the original height map considering the absolute value of local slope information. The gradient map is then used to turn flat regions into dales surrounded by sharp walls, the walls corresponding to the step-like transitions in the original height map. Once the dales have been created, they can be naturally separated from each other by running a dale partitioning process as defined in the ISO standard [6]. The method has been successfully tested in [19] to isolate bumps obtained by laser zone texturing on the surfaces of hard disk drives. However, it was also recently shown by the author of this thesis [20] that neither this method nor the conventional morphological segmentation options may always work, because a target feature may not always be mappable to a convenient combination of hills or dales, or the feature may not posses sufficiently evident edges to benefit from the application of the modified methods suggested in [19].

Alternatives to ISO morphological segmentation and its research variants are often

[19] the sequential nature applies also to line-scanning strategies

sought for in the domain of computer vision and image processing because of the similarities in the mathematical formulation of height maps and intensity images (see Chapter 3). For example, height-based thresholding is a fairly common technique used to automatically isolate features that are either protruding or recessed within the surface. Height-based thresholding is formally identical to intensity thresholding for 2D images, one of the most common and established techniques for feature detection in computer vision, given its simplicity and execution speed, therefore it is often seen as an appealing choice for feature-based characterisation in surface metrology as well. In [21], binarisation (that is, binary thresholding) is used to isolate the units of a microlens array mould. A comprehensive review of how thresholding can be used to isolate features can be found in work by the author [22]. Approaches other than thresholding are hard to be found: in [23] the autocorrelation function (ACF) of a tessellated surface (an abrasive surface or optical de-pixelator lens) is computed and partitioned into ISO hills [6], each representative of a pattern unit. By joining the hill centroids into a network, the underlying properties of the lattice become visible, and can be encoded through a new set of parameters. However, since the boundaries of each ACF hill do not correspond to the real boundaries of the pattern unit, direct computation of the shape attributes of each hill is challenging. Shape descriptors, that is, mathematical transforms that can turn a shape (in this case, a local topographic feature) into a series of scalar numbers, can be used to identify and encode features. For example, in [24] Fourier descriptors are used to identify defects in micro-optical and micro-mechanical systems. In other literature [25] wavelet decomposition is used to indicate the presence of features (e.g. scratches). Wavelet technologies will be discussed later in Chapter 5 and Chapter 6. Several approaches to feature identification and isolation have been proposed by the author of this thesis [5, 16, 20, 26–37]. The most relevant ones will be discussed in Chapter 5 and 6.

2.5.1 Commercial software

A simple procedure often applied with the tools available in current commercial surface metrology software to compute sizes and distances involving localised features is shown in Figure 2.13. The procedure consists in measuring the distance between markers manually placed onto pixels on screen. This is typically done either using a top-view or on cross-section images.

The procedure illustrated in Fig.2.13 is not robust from a metrological standpoint. The selected pixels may not be the best representatives of the edges; they may be misaligned with respect to the intended measurement direction; or the reference plane selected for

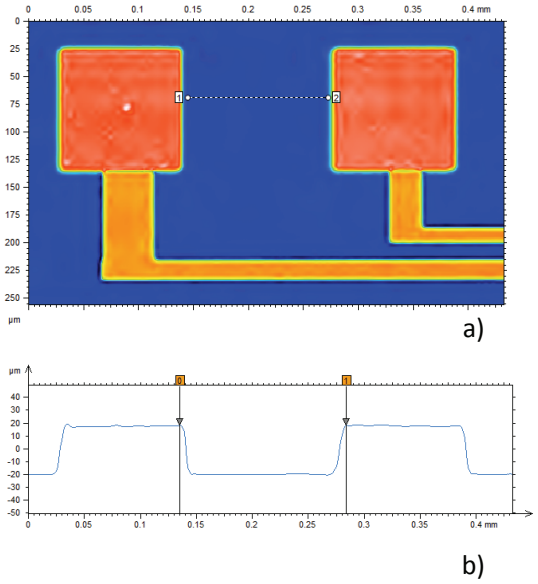


Figure 2.13: Manual placement of markers and computation of Euclidean distances between pixels in current commercial surface metrology software; a) distance on top-view; b) distance on cross-section profile (part of a MEMS device) - from [31]

the measurement procedure may be misplaced. Moreover, while sharp edges make visual identification of feature boundaries simpler (for example, step-like geometries), gradual transitions may make it more difficult (for example, an irregular groove). Ultimately, the manual process is not traceable, repeatable and/or reproducible, and the uncertainty associated with the results is difficult to quantify.

Some commercial surface metrology software products are starting to include algorithmic functionality to trace contours (on planes parallel to the x,y plane) or cross-sectional profiles (on planes containing the z axis). Some packages also provide tools to fit simple 2D geometric elements such as lines, circles, etc. to contours/profiles. This allows a more robust computation of linear dimensions, distances and angles (for example, see Fig. 2.14). Some software packages are also beginning to provide functions compliant with ISO specification standards for profile assessment, for example ISO 12181-2 [38] for roundness assessment and ISO 5436-1 [39] for step height evaluation.

Despite the availability of tracing and fitting algorithms within commercial software, there are still significant issues with accuracy, repeatability, reproducibility and traceability; mostly due to the influence of human judgement in the process. This principally includes the choices made in the initial filtering and levelling of the surface topography, in the extraction of z profiles and x,y contours, and in the identification of the contour/profile point subsets to be used for each fitting. Finally, since only contours and profiles are targeted, only 2D characterization is generally possible. For more complex

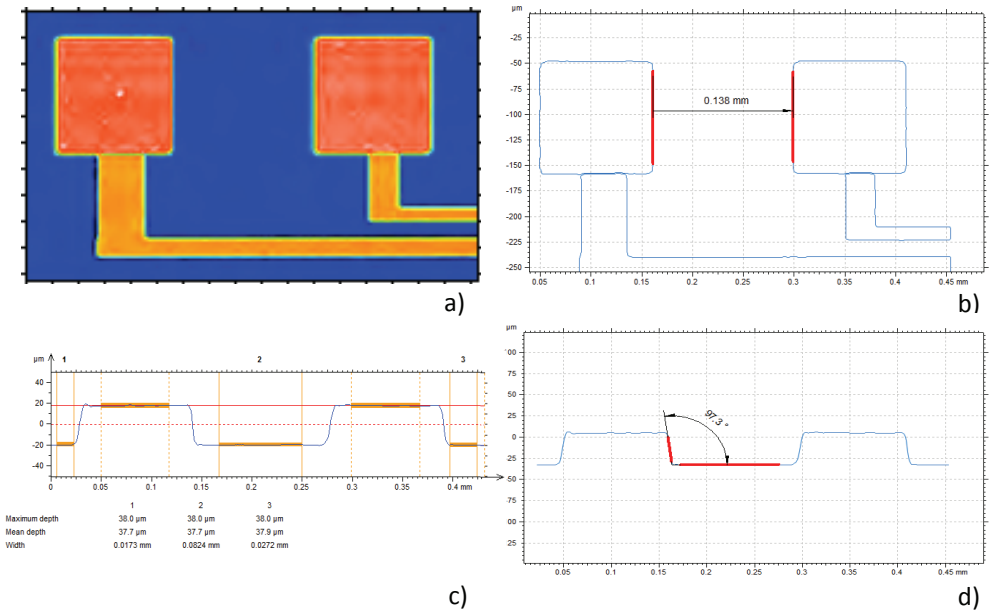


Figure 2.14: Computer-assisted evaluation of linear distances, dimensions and angles in current commercial surface metrology software; a) original topography; b) extraction of contours on the xy plane, fitting to 2D geometry (lines) and computation of linear distances; c) extraction of z profile and computation of step height according to the ISO 5436-1 [39]; d) extraction of z profile, fitting to 2D geometry (lines) and computation of a draft angle - from [31]

types of geometric analysis, areal topography data is usually exported to coordinate metrology/reverse engineering software.

2.6 Summary and outlook

Conventional surface metrology is mostly dedicated to the characterisation of surface topographies which are consistent with the "uniform random roughness" assumption, i.e. topographies which appear random, small-scale with respect to the overall form (shape) of the surface, and uniform throughout the extents of the surface itself. Under the uniform random roughness assumption, characterisation of topography is typically achieved by computing quantitative indicators (texture parameters) that summarise the salient traits of the topography region being observed through simple scalar values. The most important texture parameters currently in use, taken from the standard ISO 25178-2 [6], have been illustrated and discussed in this chapter. The uniform random roughness assumption falls short in a whole series of situations, where either surface topography is not uniform, or one is interested in localised formations, or topography is far from being random. The most common "non-conforming" scenarios have

been introduced in Chapter 1. The reasons why conventional surface texture parameters are often inadequate to characterise these "non-conforming" topographies have been discussed in this chapter. It has been shown that no textbook solution is available to characterise such surfaces, or the relevant topographic formations appearing on them. Instead, users/researchers are typically left to the development of do-it-yourself solutions, lacking generality, re-usability, and more often than not, lacking the mathematical foundations and robustness which instead conventional surface metrology has developed over the year for the characterisation of surfaces which comply to the uniform random roughness assumption. Even the state-of-the-art in terms of research, is currently populated by ad-hoc approaches, and no general theory is available to address scenarios going beyond the uniform random roughness paradigm.

Considering the review of current literature illustrated in this chapter, the logical step forward is the identification of common foundations to all the ad-hoc approaches that have been presented so far, to form the basis for the development of a more general theory of surface metrology, capable of accommodating all the surface characterisation scenarios, conforming and non-conforming ones, under a unified perspective. A more general theory of surface metrology should also provide the users with useful tools to make informed decisions on what characterisation routes should be optimally pursued, based on the characterisation task.

Chapter 3

Measurement of topography and encoding of measurement data

The way information is collected from actual surfaces (by measurement instruments) determines what type of information is available to the characterisation, and how such information is typically encoded and represented mathematically. Type and encoding of surface topography information in turn affect the way information can be manipulated in order to obtain a characterisation result.

In this chapter, the main technologies for acquiring surface topography information and the formats such information is ultimately available are introduced and discussed from the viewpoint of characterisation.

Note that measurement technologies are not classified as commonly done in the literature [4] or in the ISO standards [40–47]. The chosen classification is in line with the need of differentiating topographic content in terms of how it is formally encoded, as this is the main constraint to what type of analysis methods can be applied.

3.1 Two-dimensional imaging

Surface-related information has been collected for a long time as two-dimensional images from cameras, optical microscopes and scanning electron microscopes. In this thesis, images are assumed acquired in digital form.

Optical images, in particular if taken orthogonally to the surface (i.e. orthophotos), allow exploration of a surface topography as it is laid out on the image plane, and form the basis for defining a Cartesian reference system where the image plane is the x,y plane, and the viewing direction corresponds to the z axis. Two dimensional, dig-

ital orthophotos can be used to extract x,y geometric information pertaining surface features. However, the measurement of distance and size on the x,y plane must be reliable, which typically implies that some sort of calibration has been accomplished on the imaging system, so that the mapping between pixel size and actual size covered on the imaged surface is established. The latter implies that any deformation or other aberration in the image reconstruction process must be properly identified and modelled. Regardless of calibration, the main factors governing how surface images are obtained and what use can be made of them, pertain to the interaction of light with the surfaces and with the imaging detector[1]. So for example, self and projected shadows, overly bright or dark regions, multiple reflections, out-of-focus effects, different perception of different colours, are some of the most common factors affecting our behaviour and performance in the visual identification of topography features, in the visual assessment of their sizes and boundaries, in the detection of regularity in patterns. The same complex light-related effects that hamper a reliable, quantitative assessment, on the contrary are often extremely useful for the visual detection of otherwise harder-to-notice elements and patterns. For example, particular reflections may amplify otherwise invisible features, variations in chromatic patterns may help us identify otherwise unnoticeable textural transitions. An example image collected from a laser powder bed fusion surface, by means of an optical microscope equipped with a high magnification objective is shown in Figure 3.1. The image summarises the most important pros and cons of two-dimensional imaging for the investigation of surface topography.

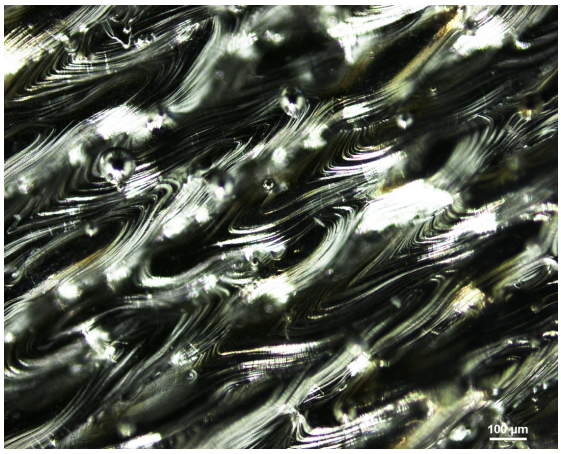


Figure 3.1: Optical image of laser powder bed fusion surface (Ti6Al4V, 10× magnification).

Out-of-focus problems typical of optical imaging at higher magnifications can be reduced by *focus-stacking*, i.e. by taking a series of images at different focal distances,

[1] this discussion refers to optical imaging, analogous considerations can be made for electron imaging, considering secondary or backscattered electrons, instead of light.

and then by extracting in-focus pixels from each and combining them into a final, composite image [48]. Illumination problems (i.e. dealing with overly bright and overly dark regions at the same time) can be similarly approached by multiple techniques, commonly referred to as *high dynamic range* (HDR) imaging [49].

Scanning electron microscope imaging usually presents fewer contrast-related problems, and in general a better depth of focus, thus allowing to obtain wider portions of in-focus data, even at higher magnifications. The difference with optical microscopy is immediately evident by looking at the same laser powder bed fusion surface previously shown in Figure 3.1 as it appears in SEM imaging at different magnifications (Figure 3.2. Shading and shadowing effects due to how secondary or backscattered electrons interact with the surface and with the detector, do however produce imaging artefacts that in turn hamper quantitative assessment of feature boundaries. Calibration of SEM images, so that scale information is reliable at least on the imaging plane is also difficult, which limits SEM imaging to mostly qualitative assessment[2]

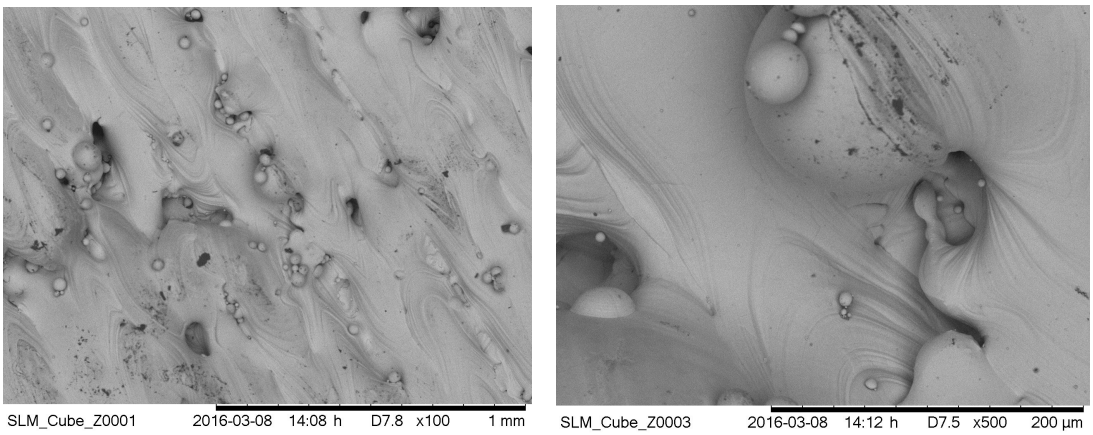


Figure 3.2: Scanning electron microscope images of laser powder bed fusion surface (Ti6Al4V); left) 100×; right) 500×.

It is hard to extract quantitative information from optical and SEM imaging of surface topography, because of the aforementioned calibration challenges. Also, information is predominantly confined to the x,y image plane. Despite both limitations, optical and electron digital imaging of surface topography still plays a prominent role in surface topography investigation and characterisation, although often in combination with other technologies capable of achieving a three-dimensional measurement of surface topography. In order to understand how "pure" two-dimensional imaging data can be com-

[2]Stereo-pair SEM imaging is capable of producing height maps (i.e. 2.5D datasets, see Section 3.2.1) by applying photogrammetry principles to combine 2D SEM images [50]. However, the technology has not reached sufficient maturity yet to be considered for applications of industrial measurement, and is therefore not covered in this thesis.

bined to other surface measurement results, the mathematical solutions for encoding image data must be discussed first.

3.1.1 Encoding of surface data from two-dimensional imaging

A conventional two-dimensional, digital image is typically represented as shown in the top part of Figure 3.3. That is, a matrix (two-dimensional array) of container cells (often referred to as *pixels*, in reference to their physical counterparts), each containing colour information encoded according to one of the commonly recognised schemes (e.g. RGB, YCbCr or HMV for colour, intensity level for grayscale). Each "pixel" is localised by row,column indices (i, j in Figure 3.3). If the imaging system is properly calibrated, then pixel width (or equivalently, the distance between the centres of two consecutive pixels) can be mapped to a reference length on the plane of the measured surface. So pixel widths/distances in x and y are mapped to the spacing parameters dx and dy. This is like saying that the colour information stored in a pixel is representative of an actual area $dx \times dy$-wide on the measured surface. Assuming an arbitrary reference system placed at the corner of the image, then the actual co-ordinates x_i, y_j of the centre of each pixel can be found as: $x_i = (i-1)dx$ and $y_j = (j-1)dy$ (see Figure 3.3-bottom). Notice that for convenience, the indices i, j are kept separate (although one could also describe the entire set of pixels with an individual index). Note also that one needs to choose a convention of whether the row of column image index is associated to the x co-ordinate in the image plane. The dataset comprised of a $n_x \times n_y$ matrix of colour or grayscale values, with associated dx and dy spacing information, is referred to as a *colour map* or *grayscale map* in this thesis. The final size of the map will be $\Delta x \times \Delta y$ where $\Delta x = (n_x - 1)dx$ and $\Delta y = (n_y - 1)dy$.

Notice that a colour or grayscale map implies the definition of a coordinate system local to the map, the x and y axes aligned to the rows and cols of the image matrix, and the origin typically associated to an image corner, as shown in Figure 3.3). The map is usable for many types of analyses regardless of whether the position of the origin (x_0, y_0, z_0) is known. The origin, as well as the direction cosines defining the orientation of the map co-ordinate system are needed only in case multiple maps need to be co-localised in the same global Cartesian space, for example, for comparison or aggregation (see for example Chapter 7).

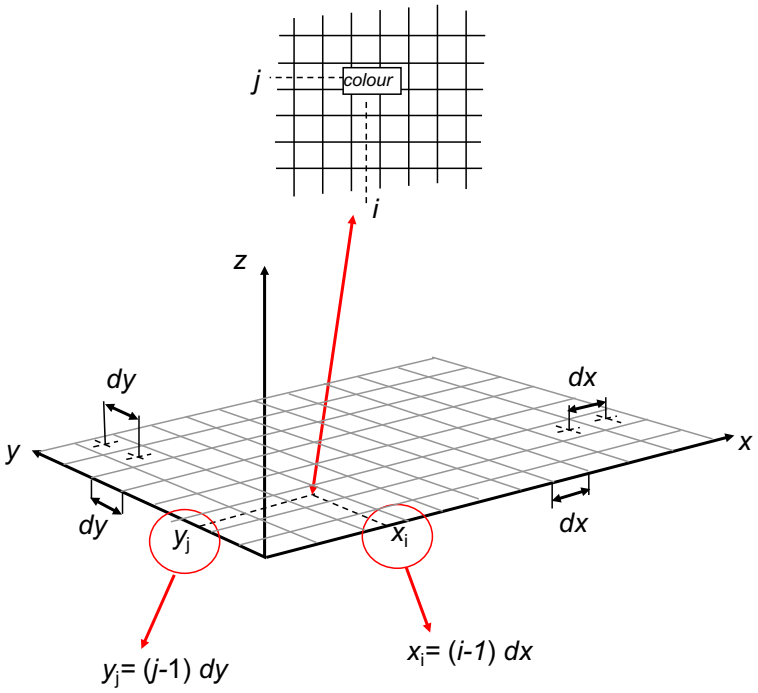


Figure 3.3: Encoding a two-dimensional digital image into an image map.

3.2 Line-scan measurement

In line-scan measurement, an area of the sample is sampled by a single-point height measurement probe traversing the surface along parallel, straight scan lines (here assumed oriented along the x axis). This type of scanning pattern is often referred to as *raster scanning* in this thesis. As the probe traverses a single scan line, the local height z_i co-ordinate is recorded along with the current and corresponding x_i position along the scan line. Depending on the implementation, the x_i co-ordinate may be recorded explicitly as it is read by encoders, or implicitly assumed, by considering a constant traversal speed and constant sampling frequency, which leads to uniform spacing dx, so that $x_i = (i-1)dx$. Single-point height measurement has been historically implemented through contact probing, recording vertical displacements of a shaft with a spherical tip at the end, as it is slides on the surface following the scan line (such probe is typically referred to as *contact stylus*). Other non-contact technologies have been experimented, e.g. in single-point interferometric probes UBM [51]). More recently, point-autofocus measurement has gained popularity (see the standard ISO 25178-605 [45]). Regardless of the implementation of point-height measurement, all the technologies share the same sequential nature intrinsic to measuring along a scan line, and then through multiple, parallel scan lines.

3.2.1 Encoding of surface data from line-scan measurement

If, during the execution of an individual line scan, both the traversal speed of the probe and the frequency of point-height measurement can be assumed constant (see Figure 3.4), then a scanned profile can be simply encoded as a one-dimensional array of height values $(z_1, z_2, \ldots, z_i, \ldots, z_n)$, while the position x_i of the i^{th} point can be simply obtained as $x_i = (i-1)dx$, where dx is the spacing between two consecutively measured points[3]. The result is the equivalent of a one-dimensional map; i.e. a one-dimensional array of scalar height values (the z values) mapped to a one-dimensional co-ordinate axis (the x axis), though clearly the resulting profile exists in a two-dimensional, Cartesian space (the x, z space).

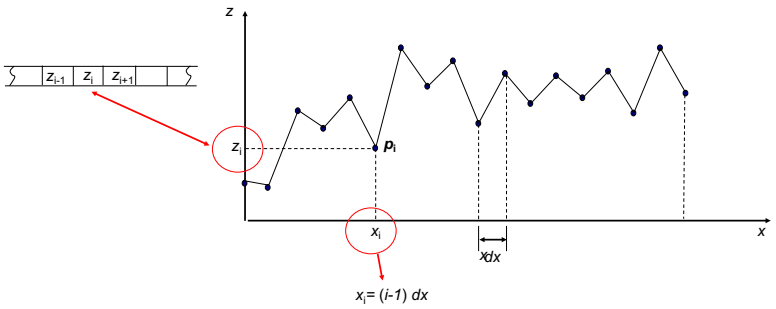


Figure 3.4: Profile scan encoded into a profile map.

For an areal measurement executed through a series of parallel line scans (i.e. by raster scanning), the encoding is as illustrated in 3.5. Notice that the final set of height values (z matrix) is encoded with no difference between rows and columns, while it is known that the dynamics of measurement may influence the results differently along and orthogonally to the scanning direction[4]. Spacing along the x and y axes (referred to as dx and dy respectively in Figure 3.5) may not necessarily be equal, which is often the case if scanning is slow and time-consuming (i.e. a smaller number of scan lines is recorded as opposed to the sampling resolution of each single scan line). Regardless of whether the dx and dy spacing values are equal or not, their knowledge is essential in order to reconstruct the actual x, y coordinates of any $z_{i,j}$ value, being: $x_i = (i-1)dx$ and $y_j = (j-1)dy$[5].

The encoding illustrated in Figure 3.5 provides the first example of a *height map*. That is, a two-dimensional matrix of scalar height (z) values, stored along with either explicit

[3] As stated earlier, the alternative of explicitly storing the x_i coordinates as they are measured by an encoder is sometimes implemented as well

[4] For example, in contact measurement, a stylus probe may "bounce" off the surface, of experience stick-slip phenomena during the scan.

[5] Again, some classes of instruments may record the x_{ij}, y_{ik} co-ordinates explicitly, through encoders.

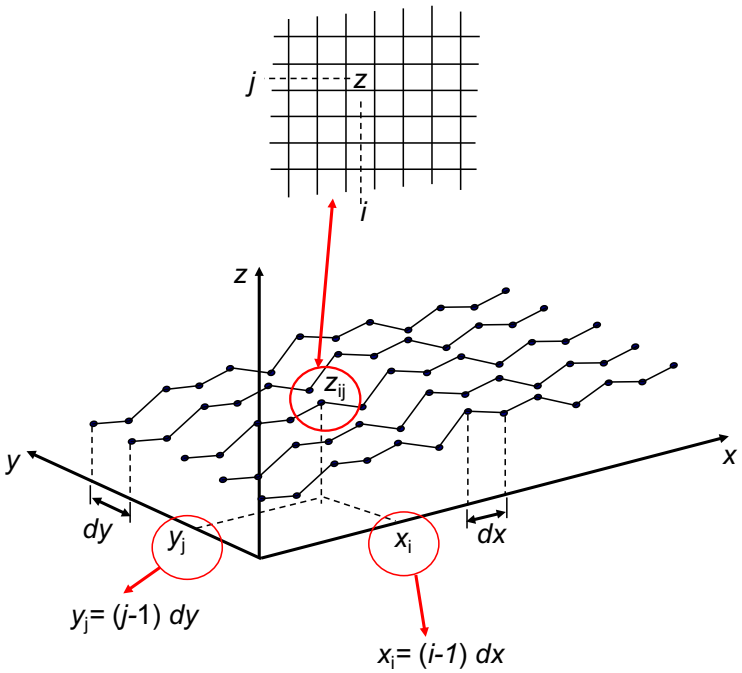


Figure 3.5: Encoding of a raster scan into a height map.

x,y information, or dx,dy spacing. Height maps are sometimes referred to as *range images*, i.e. image-like data structures, whose cells contain height (range) information as opposed to colour or gray level. They are also often referred to as *2.5D geometry* to differentiate them from *full 3D geometry*, because at any x,y co-ordinate only one z value is possible, and therefore no re-entrant geometries can be defined.

3.3 Vertical-scan imaging

The most popular optical technologies currently adopted in surface metrology for the measurement of areal topography are all based on non-contact, optical principles, and include: *coherence scanning interferometry* (see [4] and the standard ISO 25178-604 [44]), *Confocal miscroscopy* (also commonly known as confocal laser scanning microscopy, see [4] and the standard ISO 25178-607 [47]), and *focus variation microscopy* (see [4] and the standard ISO 25178-606 [46]).

All such solutions can be seen as gravitating around a series of common architectural choices which in this thesis are collectively referred to as the *vertical-scan imaging* paradigm. In vertical-scan imaging, some light source is sent to the surface (the technologies differ on whether adopting coherent sources or not). Reflections from the surface are processed in different ways (again according to the specific measurement technology), and the results of such processing are captured by a charge-coupled de-

vice (CCD) or similar matrix-like detector (hence the term *imaging*). The entire process is repeated at multiple vertical distances from the surface (hence *vertical scanning*) to generate a vertical stack of *raw images*, one per each vertical distance between the probing device and the surface. Finally, the stack of raw images is processed (in different ways depending on the measurement technology) until in correspondence of each pixel of the detector an estimate of local height z_{ij} is produced.

Being non-contact, vertical-scan imaging solutions are much faster than any stylus-based raster scanning device. Even when raster scanning is implemented with a non-contact probe (e.g. point autofocus), vertical-scan imaging is still much faster than physically traversing the surface laterally (i.e. on the x,y plane, like instead it is done in all the raster scanning solutions[6].

3.3.1 Encoding of surface data from vertical-scan imaging

Regardless of the adopted height measurement technology, all vertical-scan imaging instruments ultimately produce a $n_x \times n_y$ matrix of height values, whose size corresponds to the native resolution of the detector. The matrix can be mapped to an actual x,y rectangular area on the specimen surface through the knowledge of what unit area $dx \times dy$ in the actual surface is covered by each "pixel" at the detector. Such knowledge is acquired at calibration, and is dependent, amongst other things, on choice of magnification objective. The encoding of surface data as acquired by a vertical-scan imaging device is shown in Figure 3.6. The result is a *height map*, formally identical to what obtained by measurement by raster scanning, except that this time it is not possible to talk about differences between a scan direction and its orthogonal[7]. Since in vertical-scan imaging there is no lateral (x,y) movement between the probe and the surface, the alternative of capturing x,y point information from encoders is not available; the only option is to store the z matrix and the two spacing values dx and dy known from calibration[8].

Height maps as structured datasets

Height maps are an example of *structured dataset*, meaning that they can be seen as collections of points in space (each point is defined by the triplet: $z_{i,j}, x_{i,j} = (i-1)dx, y_{i,j} =$

[6]Confocal microscopy is actually a hybrid, because generally the single raw image is generated by very fast x,y scanning. However, for the purpose of keeping this discussion simple, x,y point acquisition within the single raw image can be considered as simultaneous

[7]except may be for confocal laser scanning microscopy, see [4] for details

[8]again confocal measurement being an exception

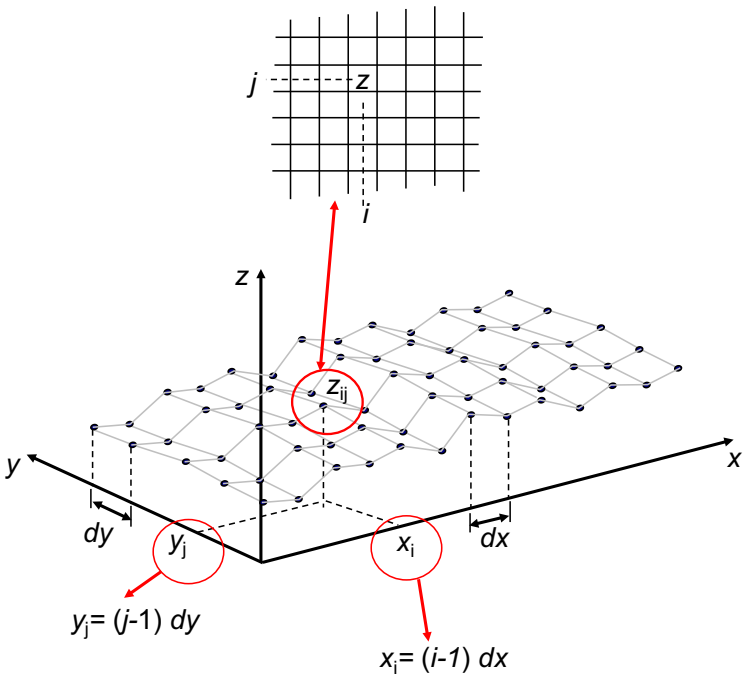


Figure 3.6: Height map from vertical scan imaging

$(j-1)dy$), but topological information (i.e. which points are connected to which by neighbouring relationships) is also implicitly captured by the indices i, j. As illustrated in Figure 3.7, for any given i, j point, neighbours can be found by considering the adjacent indices $i \pm 1, j \pm 1$. This is fundamentally different from having a simple point cloud (i.e. a set of x, y, z triplets), where no topological information is stored[9].

The availability of topological information describing connectedness between height values is of fundamental importance for surface metrology and in particular for the feature-based characterisation framework which will be presented from Chapter 4 onwards.

3.4 Volumetric imaging

The term *volumetric imaging* is used in this thesis to refer to a class of measurement technologies currently embodied by X-ray computed tomography (XCT). In XCT, a series of x-ray images is generated by projecting X-rays through a part placed at different angular orientations, and collecting the rays on the other side of the part, through a detector[10] Once the images (X-ray projections) are generated, they are mathematically combined to generate a three-dimensional matrix of cells (called *voxels*), each encod-

[9] Point clouds are therefore referred to as *unstructured data sets*.

[10] In some configurations, the part rotates, in others, the X-ray source and detector rotate about the part.

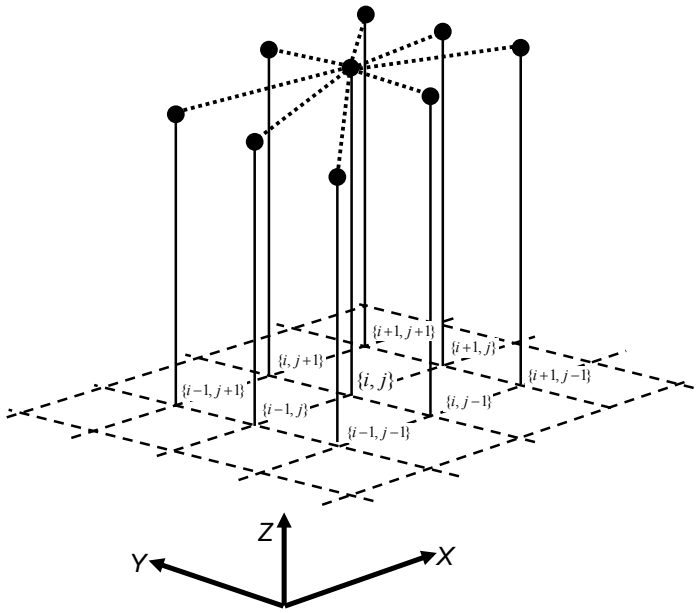


Figure 3.7: Topological information about neighbouring relationships for the i, j point of a height map.

ing local part density at a given position in a Cartesian space [52]. Through a family mathematical procedures, the boundary which separates the inside from the outside of a measured part can be identified within the three-dimensional matrix of voxels, by looking at density transitions, and extracted as a surface [52, 53]. An example volumetric dataset is shown in Figure 3.8. The dataset is sectioned to highlight the inside-outside transition for the feature (high aspect-ratio micro-hole). The extracted surface is also shown overlaid.

In recent work [54–57] it was demonstrated that, as long as proper measurement conditions are set, the surface extracted from XCT volumetric measurement can carry topographic detail comparable to the more established optical technologies for areal topography measurement. Actual quantitative comparisons between measurement technologies including XCT applied to the measurement of metal AM parts were also recently published by the author and collaborators [58, 59].

The boundary surface extracted from a XCT volumetric dataset is typically expressed as a triangle mesh and, because of how it has been extracted from voxels, does not suffer from 2.5D limitations, meaning that undercuts and in general any type of re-entrant surface can be represented natively. In Figure 3.9, a cavity belonging to a laser powder bed fusion surface is shown, as extracted from a volumetric XCT dataset.

A formal representation of a triangle mesh dataset contains as a minimum information

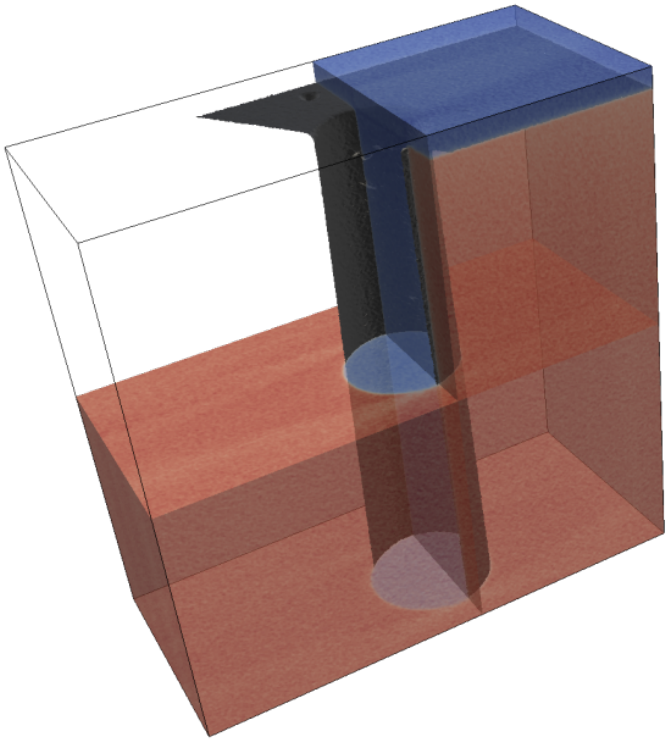


Figure 3.8: Example volumetric dataset containing a high aspect-ratio microhole (shown in cross-section). The areas at different densities are rendered with different colours. The boundary surface of the micro-hole, as extracted from the volumetric dataset, is rendered as well (also in sectioned form). Image from previous work by the author of this thesis [16].

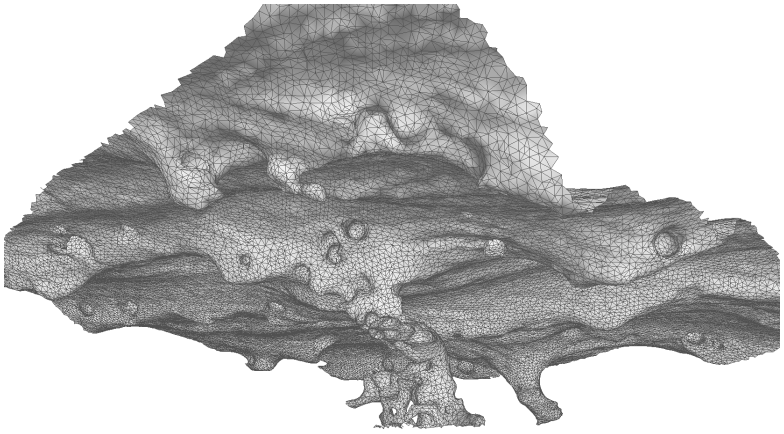


Figure 3.9: Deep cavities under the surface of a laser powder bed fusion part. Triangle mesh from XCT measurement.

about the x, y, z coordinates of the triangle vertices, and topological information about the triangles themselves, each defined by a triplet of vertex identifiers (see Figure 3.10). To facilitate mesh manipulation and navigation, often also information about triangle unit normals, as well as cross-referenced pointers to neighbouring faces, vertices and

edges to each triangle, are commonly stored as well.

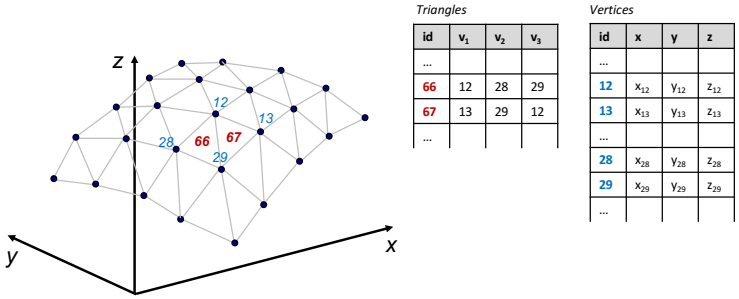


Figure 3.10: Basic encoding scheme for a triangle mesh. Vertices are defined via triplets of scalar co-ordinates; triangle faces are defined by triplets of vertices.

The topography of a surface extracted from XCT volumetric data is heavily dependent on the algorithms that were used for surface determination starting from the volumetric dataset. As mentioned earlier, such algorithms are based on detecting inside-outside transitions across voxels by searching for significant variations of stored density information. Studies by the author of this thesis and collaborators [16] show how different surface determination setups influence the reconstructed topographic detail. In Figure 3.11 the effects of the surface determination algorithm are shown for the entry region of a high aspect ratio micro-hole (from previous work by the author of this thesis [16]).

3.5 Other topography measurement solutions that result in triangle meshes

A measurement system architecture which is gaining popularity for the measurement of surfaces consists of a 2.5D probing solution (typically optical) mounted as the end-effector of a five-axis CMM-like architecture. Multiple 2.5D images (height maps) can be taken from different viewing directions, and then combined into a full 3D shape. The dataset combination methods and challenges are not covered by this thesis. However, it is important to point out that the final result is typically expressed as a triangle mesh. An example of such measurement is shown in Figure 3.12 and pertains a high aspect-ratio micro-hole captured by a focus-variation system equipped with a multi-axis stage capable of tilting. The side walls of the hole are parallel to the hole entry surface (actually, they are slightly re-entrant, as the hole has been realised by water-jet micro-machining); thus, to measure them, the sample is tilted and rotated multiple times, and the height maps stitched together.

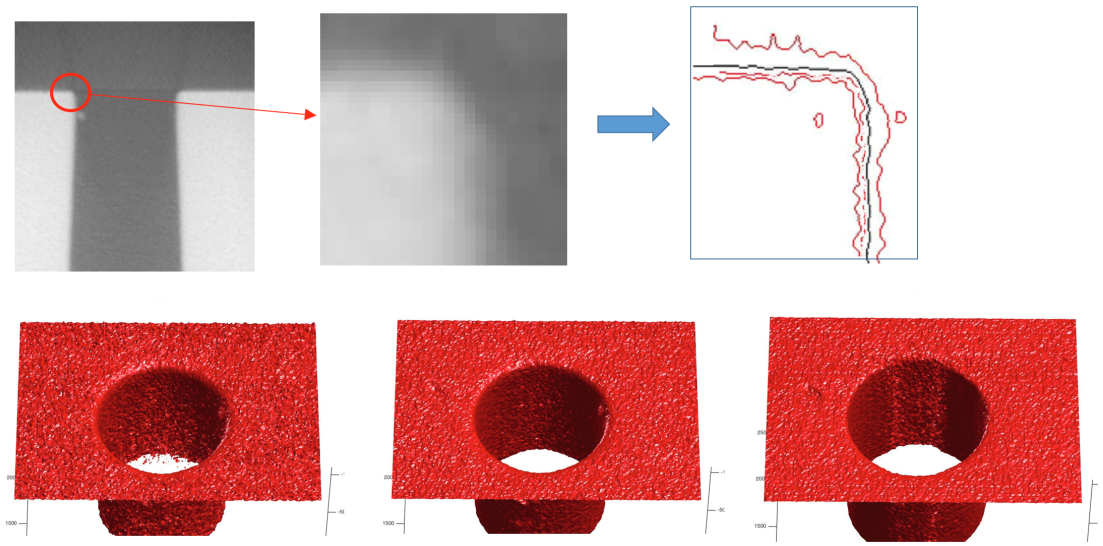


Figure 3.11: Topography as a function of the surface determination algorithm in XCT; the top row shows the entry cross-section of a micro-hole (slice from volumetric dataset); a zoomed-in region showing the transition between material and surroundings; and cross-sections of multiple isosurfaces obtained by using different thresholds on voxel intensity (black line: iso-50 method [53]); the bottom row shows surface topography at the entry region of the microhole, as it has been determined by using different thresholds) (from work by the author of this thesis [16]).

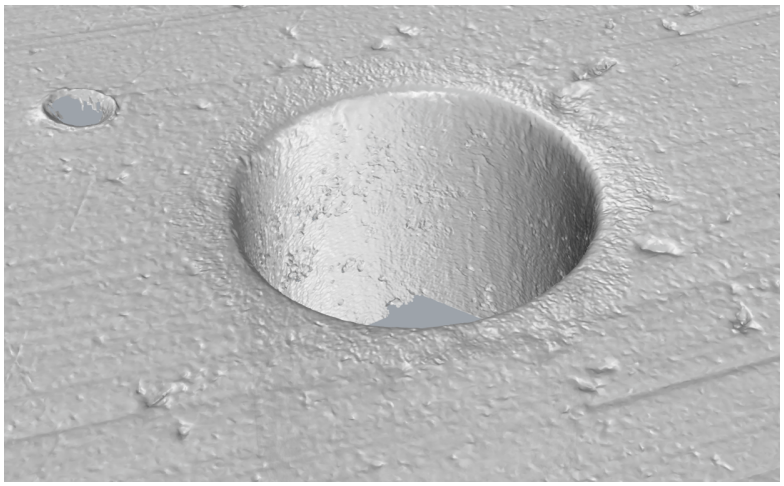


Figure 3.12: Triangle mesh of high aspect-ratio micro-hole (approx 100 μm diameter); obtained by stitching multiple height maps acquired by focus variation (Alicona IF) at 20×, with 45-deg stage-controlled tilting.

3.6 Converting triangle meshes into height maps

Regardless of where the triangle mesh is coming from XCT measurement, or from a combination of multi-directional height map measurement and stitching, it is impor-

tant to point out that, because of the intrinsic capability to represent any type of three-dimensional shape, albeit in approximated form because of the planar nature of the triangle facets, triangle meshes cannot be navigated via $z = f(x,y)$-type formulations, as the concepts of *lateral* (i.e. x,y) and *vertical* (i.e. z) direction cannot be robustly defined. Triangle meshes are not therefore natively suitable to be processed by the conventional and non conventional topography characterisation methods illustrated in this thesis.

The conversion method developed by the author of this thesis and applied in several research projects (see for example [58, 59]) is referred to as *virtual raster scanning*. In virtual raster scanning, single point height measurement is simulated by computing the geometric intersection of a z ray with the surface (triangle mesh) supposed aligned to the x,y axes. With reference to Figure 3.13, a x,y datum plane is first found by principal component analysis of the point cloud made of the vertices extracted from the triangle mesh (principal component analysis is discussed in Section 5.3.2). The x,y plane also implicitly defines the z axis (direction of observation), and allows for the identification of a parallel plane located above the surface. The parallel plane is then used to compute the x,y co-ordinates of a regular scanning grid (hence the term *raster scanning*) and each x,y value of such grid is used as the starting point to project a z ray that intersects with the triangle mesh. The intersection is stored as a measured height value. Since the scanning grid is regular (uniform spacing in x,y), the result of virtual raster scanning can be stored as a height map, i.e. as a matrix of $z_{i,j}$ values attached to dx and dy grid spacing information.

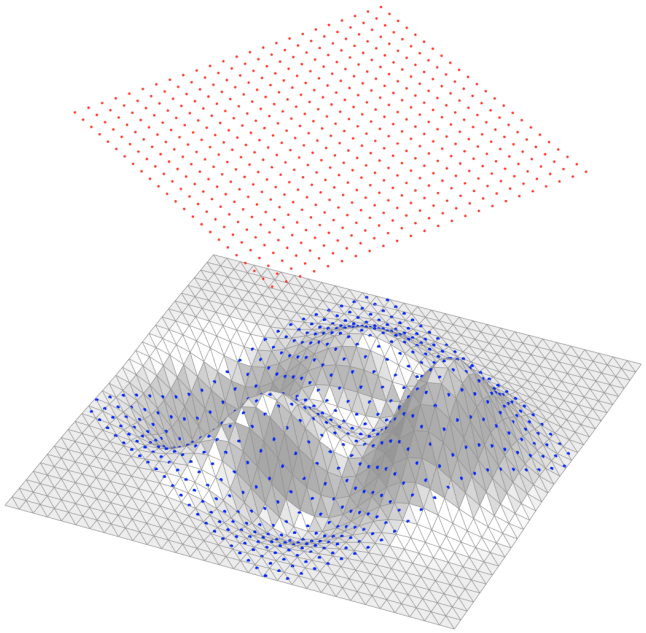


Figure 3.13: Virtual raster scanning example; top) scanning grid (in red); bottom) z rays intersections with the triangle mesh.

3.7 An original approach to map-based encoding of topographic features

Because of how they are defined, maps are rectangular. Topographic features, preliminarily defined as topographic regions of interest located within the measured topography[11] may not necessarily be rectangular. In order to preserve the advantages of map encoding[12], the non-rectangular limits of a topographic feature can be handled without breaking the map data format, via the introduction of additional, special purpose maps, such as classification maps, and masks. A *classification map* is a map whose contents are to be interpreted as *identifiers* (typically, scalar integers). Classification maps can be used to flag the cells of an original height map with different integers corresponding to identifiers of different classes (encoding multiple regions of interest on a surface). Classification maps can be useful in segmentation, i.e. in partitioning a surface topography into regions (see Chapter 6), where each partition can be assigned to a different identifier. *Masks* are a particular type of classification map only allowing boolean values. Masks can be used as selection maps, or to represent any type of binary partitioning of a topography, for example to distinguish between the region occupied by a feature of interest and its surroundings. The use of masking methods to isolate regions of interest (ROIs) is common in digital image processing and is here proposed to identify the regions occupied by any feature. An example is given in Figure 3.14, where a mask is used to discriminate points of a height map belonging to a step-like feature versus points belonging to the feature surroundings.

3.7.1 Implications related to map-based encoding of topographic features

The possibility of retaining a map-based representation for individual topography features provides countless opportunities to apply data analysis and processing methods originally developed in the domain of image analysis and processing, to the characterisation of individual features. The downside of this is that map-based encoding of shapes brings along intrinsic limitations, summarised by the following points:

- *unidirectional probing*,
- *uniform sampling*, and

[11] (more formal definitions will be provided in Chapters 4, 5 and 6

[12] navigational advantages, since it is easy to determine the neighbours of any given point - see Figure 3.7, and processing advantages, since image-like encoding allows for the application of numerous existing algorithms for image analysis and processing

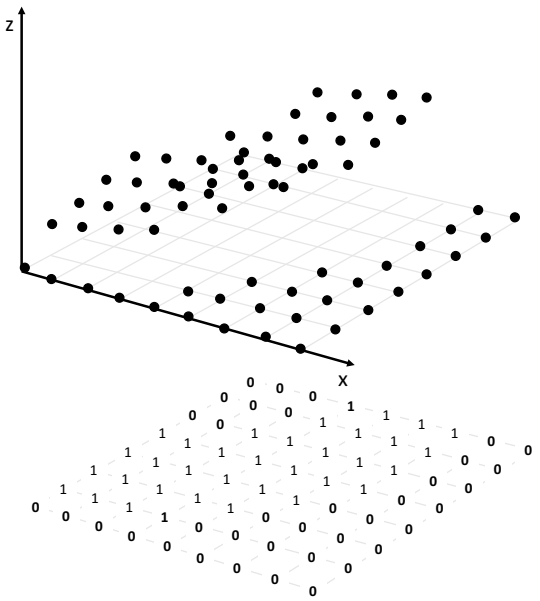


Figure 3.14: Use of a mask (selection map) to discriminate between height map points belonging to a step-like feature from those belonging to the feature surroundings.

– *the blanket effect*.

Unidirectional probing

The map-based encoding of topography information typically implies that surface measurement takes place along a single probing direction, a condition which can be referred to as *unidirectional probing*. When the shape of an individual topographic feature is encoded, the use of a single probing direction means that surface features aligned to the "line-of-sight" direction will not be captured, as well as any re-entrant region,i.e. located in undercut regions with respect to the viewing direction. Similarly, any topographic feature located in a re-entrant region, underneath another one, will not be visible because of occlusion (Figure 3.15).

For large surfaces where the "uniform random roughness" paradigm applies, missing some regions because of local undercuts or high slopes may not be particularly impactful. However, when individual topographic features are concerned, unidirectional probing may lead to significant loss of coverage, in particular if the topographic feature is characterised by a large number of undercuts, or quasi-vertical surfaces (with respect to the viewing direction).

Even if the local orientation of some regions never reaches parallelism with the line-of-sight direction (which, as seen above, implies non-measurability), high slopes will typically lead to a further loss of measurable regions (e.g. see Figure 3.16) because of the

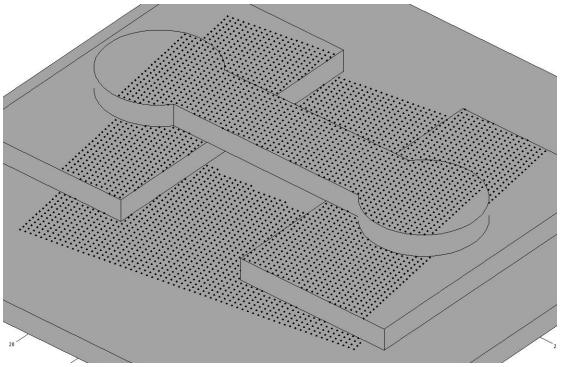


Figure 3.15: Effect of unidirectional probing (simulated): vertical surfaces, undercuts and occluded regions are not captured by the measurement (from [31]).

intrinsic limitations of most measurement technologies currently adopted for surface topography measurement[13].

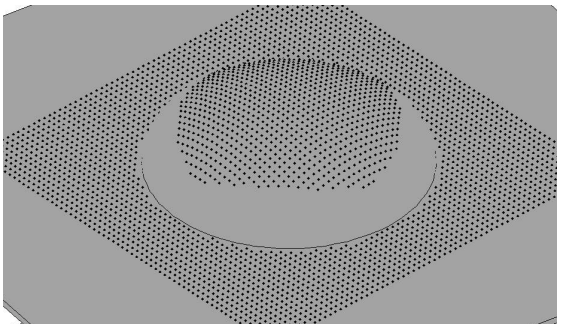


Figure 3.16: Regions beyond the maximum measurable slope may not captured by the measurement process (simulated). From work by the author of this thesis [31].

Whilst in some cases, going beyond the maximum measurable slope will simply result in lack of coverage (as previously shown in Figure 3.16), in other cases, it may result in significant measurement artefacts, such as the batwing effect, typical of CSI measurement [4] (see Figure 3.17).

The adoption of a map-based representation of topography brings additional constraints. As surface sampling takes place along the rows and columns of an imaginary, uniform, rectangular grid, it is generally not possible to optimise position and density of local point-taking, which ideally should be based on the properties of the corresponding portion of the underlying surface. Uniform sampling is therefore often sub-optimal and cannot in principle differentiate between pre-existing surface partitions in the region being subjected to measurement. As opposed to using touch-probe CMMs and micro-CMMs, where it is possible to explicitly program the instrument to perform

[13]Except for XCT and multi-axis probing combined to stitching. In such cases though, the final triangle mesh still needs to be converted back onto a height map via raster scanning, which again leads to a unidirectional probing condition.

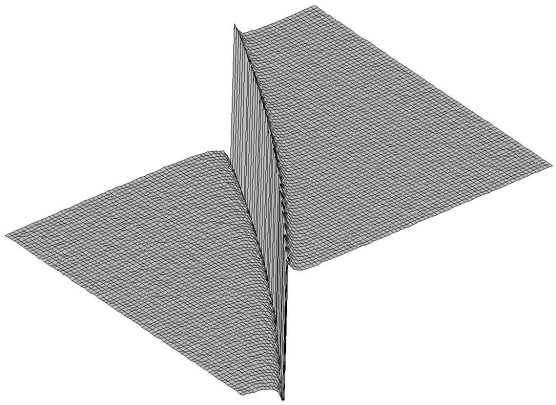


Figure 3.17: Batwing effect observed on the edges of a step-like surface feature (CSI Measurement). From work by the author of this thesis and collaborators [31].

point-taking at specific locations, thus creating pre-defined associations between each measured point and the specific surface it was taken on, in imaging-based technologies (and volume-based CT measurement as well), a phenomenon can be observed that could be called "blanket-effect": uniform sampling covers the entire surface within the field of view of the instrument, regardless of any sub-surfaces or partitioning of other kind which may exist ahead on the surface. The blanket-effect is illustrated with a simulated example in Figure 3.18.

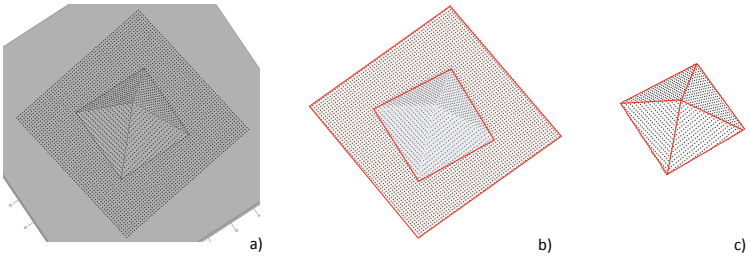


Figure 3.18: The blanket effect generates the need for partitioning the measured point set; a) the original point set covers the target region and its surroundings; b) a partitioning has been made to separate the points belonging to the target region; c) a further partitioning leads to the identification of the points belonging to each specific face of the target region (From previous work by the author and collaborators [31].

3.8 Mathematical notation

Assuming topography information stored in the formats discussed in this chapter, it is useful to introduce some reference notation that will be used throughout the rest of this thesis.

A height map will be always considered as defined by a discrete function $z_{i,j} = f\left(x_i, y_j\right)$

where x_i and y_j are the discrete i^{th} and j^{th} co-ordinates indexed separately, so that recorded observations for the function $z_{i,j}$ with $i = 1,\ldots,n_i$ and $j = 1,\ldots,n_j$ can be represented as a two-dimensional matrix of size $n_i \times n_j$. Sometimes, when convenient for illustrative purposes, height maps will be defined as continuous functions $z = f(x,y)$. Also, topography profiles in the form $z = f(x)$ will be presented to exemplify concepts as opposed to surfaces $z = f(x,y)$. However, this will be done only for the purpose of simplifying the illustration, whilst all the concepts illustrated in this chapter will cover surfaces and profiles indifferently. Analogously, any other information mapped to a surface (e.g. a colour map, or grayscale map, or a mask, etc.) will be represented by a continuous function $\rho = f(x,y)$ or discrete function $\rho_{i,j} = f(x_i,y_j)$ depending on needs, where ρ is either a scalar of vector quantity, depending on contents. In summary, an areal topography feature can be mathematically represented by a pair of functions $z = f(x,y)$ and $\rho = f(x,y)$, the latter being a mask valued *true* at the locations belonging to the feature. The same formalism can be used to represent the entire topography, the only difference being that the mask $\rho = f(x,y)$ is *true* everywhere.

3.9 Summary and outlook

According to the current the state-of-the-art in surface topography measurement and characterisation, the *map* type of data structure is dominant for the representation and storage of topography information. This is because the format is convenient in terms of memory occupation with respect to a point cloud (no need to explicitly memorise the x,y co-ordinates of each measured point) and because data navigation and manipulation is facilitated by the heigh values being organised like the pixels of a digital image, allowing for a trivial mapping of numerous operators originally developed in the domain of digital image analysis and process to the realm of surface metrology. The map data structure has been discussed in this chapter, and a formal, mathematical representation has been given, that can be generalised to accommodate information other than heights (colour, quality, categories,selections, etc.) whilst maintaining the same general structure. Thanks to such generalisation, in this chapter it has been shown that new data structures, such as *classification maps* and *masks* can be introduced and combined with height maps to enable the implementation of powerful concepts related to feature-based characterisation, such as seamless handling of non-rectangular, localised topography formations.

The drawbacks of the 2.5D nature intrinsic to the $z(x,y)$ representations of surfaces enforced by the *map* encoding (i.e. only one z value associated to each x,y location) have been discussed throughout the chapter, and the implications for surface metrol-

ogy in general, and feature-based characterisation in particular, have been discussed. In addition to a largely scale-independent discussion, in the chapter measurement issues which are typical of surface metrology instruments operating at micro and submicrometric scales have been discussed, together with their influence on feature-based characterisation.

The 2.5D nature of height maps reflects the limitations of the main surface measurement technologies currently available, typically based on unidirectional probing, as illustrated in the chapter. However, new measurement technologies are increasingly appearing that allow full-3D capture of surface topography. The most notable ones illustrated in the chapter are X-ray computed tomography and 2.5D optical probes mounted on five-axis measurement instruments. These technologies may soon require a more consistent effort towards the definition of geometric models and data structures capable of supporting full-3D surface topography. Triangle meshes are one of the most likely successful formats for the near future in terms of storing this type of measured data, having topological connectivity as a clear advantage with respect to simpler point clouds. For analytical modelling of topographic information, i.e. a step further than simple storage of measured data, parametric surface models are being given and will be given increasing attention by the research community [60]. Unfortunately, the introduction of full-3D representations means that most current surface topography characterisation methods, first and foremost those based on texture parameters, may need revisiting, as they are deeply ingrained in the 2.5D, $z = f(x,y)$ and $z = f(x)$ paradigm. As it has been shown in the chapter. Until such revisiting effort is justified, full-3D datasets can still be converted into 2.5D ones, for example by *raster scanning*, following the original methods introduced in this chapter.

Chapter 4

Basic concepts underlying the feature-based characterisation framework

4.1 Introduction

In this chapter a new framework for surface metrology is introduced that covers all the application scenarios illustrated in Chapter 1. Conventional surface metrology, i.e. what is commonly considered as the characterisation of surface roughness, covered by the ISO standards and most of the scientific literature, will be shown as representing a special case of a larger theory. The framework can be introduced by discussing three core themes, that together contribute to creating its base pillars. These core themes are:

- determinism and randomness in topography data;
- random processes and random fields to support the representation of topography; and
- feature technology as a means to handle topographic complexity via differentiation.

In the next chapters 5, 6 and 7 further elements of the framework will be discussed, with the introduction of original methods related to the identification and characterisation of topographic formation of interests (features), and the discussion of how to associate uncertainty.

Premise on notation Throughout the chapter and as previously discussed in Chapter 3, surface topography will be assumed available either as profiles $z = f(x)$ or surfaces $z = f(x,y)$. The term *location* will be indifferently used to refer to a x or x,y position, and the term *height* will be used to refer to a z value at any given location. When necessary, the discrete formulations will be adopted $z_i = f(x_i)$ and $z_{i,j} = f(x_i,y_j)$. When needed, a parallel between topography profiles $z = f(x)$ and time histories $z = f(t)$ (also referred to as *signals*) will be used, in particular when the need will arise to introduce theoretical concepts borrowed from signal processing. Analogously, the discrete representation $z_{i,j} = f(x_i,y_j)$ will be used to borrow concepts from image processing. Analogous, but more complex, representations will be introduced when borrowing concepts from spatial statistics.

4.2 Determinism and randomness in topography

The term *randomness* refers to *lack of predictability* in events [61]. Consistently, in the theory of random signals, for a signal $z = f(t)$ to be random, the value z at any point in time t cannot be predicted from past values [62].

While for time series there is a well defined directionality of t, so that it is possible to talk about prediction of an event in the *future*, starting from events in the *past*, for a topography profile $z = f(x)$ there is no implicit directionality of x values (locations). A random topography profile can therefore be defined as a topography where the value of z at any location x cannot be predicted from known z values at other locations.

Let's provide a concrete example of this by assuming a discrete profile acquired through measurement, consisting of a series of height observations $z_0, z_1, \ldots, z_n$ acquired at discrete locations $x_0, x_1, \ldots, x_n$ (Figure 4.1). The profile can be referred to as *random* if the unmeasured value z_{n+1} cannot be predicted with certainty at the new location x_{n+1} (or at any other locations not included in the list of observations), starting from the list of available observations at the other locations.

On the contrary, by looking at the profile in Figure 4.2, it is intuitively easier to imagine what the value z_{n+1} will be at the location x_{n+1}, where no observation is present. The profile shown in Figure 4.2 can be referred to as *deterministic* as long as the z value at x_{n+1} can be predicted with certainty starting from the observed values at other locations.

It is immediately evident that, depending on circumstances, one may have a higher or lower degree of uncertainty associated to the prediction of a value z at an un-probed location x or x,y, meaning that topographies may be "random" to varying degrees. It is

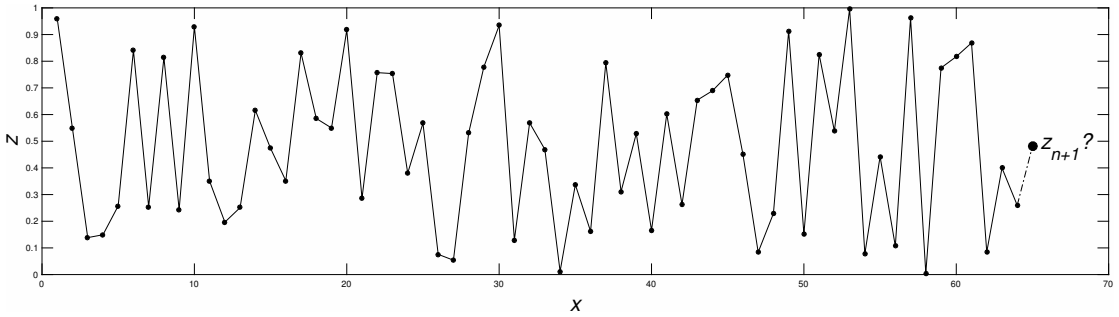


Figure 4.1: Example random profile. The value z_{n+1} at the unobserved location x_{n+1} cannot be predicted with certainty before it is observed.

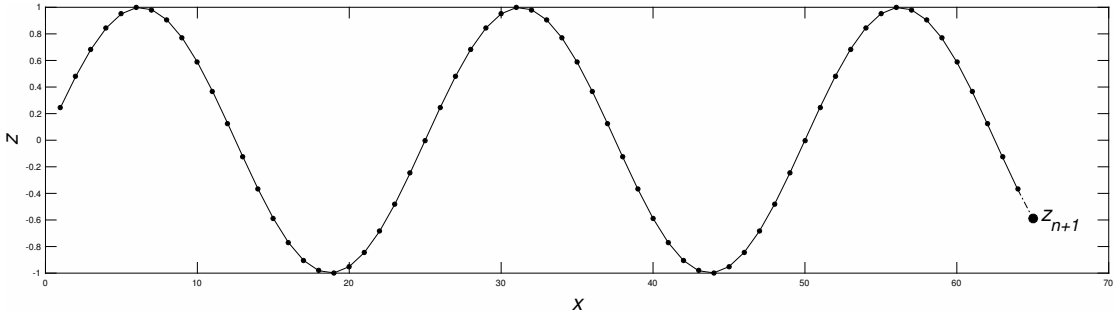


Figure 4.2: This profile is deterministic as long as z_{n+1} can be predicted with certainty from the other available observations.

therefore now possible to provide a first working definition of the *degree of randomness* of a topography profile.

The degree of randomness of a topography profile $z = f(x)$ for which observations $z_1,\ldots,z_n$ are available at locations $x_1,\ldots,x_n$, is related to the level of uncertainty associated to the prediction of any value z at locations x were no observation is available.

An identical definition can be provided for surfaces in the form $z = f(x,y)$ where the one-dimensional location x is replaced with a two-dimensional location x,y[1].

Randomness, patterns and auto-correlation Why does the example profile shown in Figure 4.2 appear presumably easier to obtain a prediction for? Upon visual inspection, the profile immediately elicits association to a simple sine-like, archetypical *pattern*. More or less consciously, our brain associates the observation to a very well known, simple, regular and smooth shape. This immediately lets us "fill the voids", i.e. formulate educated guesses on values where no observation is available. A large role in this pattern-association mechanism, which is at the foundations of the phenomena that regulate our visual perception, is played by how we detect *regularity*. For a profile defined by a function such as $z = f(x)$, regularity is captured amongst other things by

[1] Alternatively, location can be represented as a n-dimensional vector **x** so that the definition applies to any n-dimensional space

spatial auto-correlation. While a full mathematical illustration of auto-correlation will be provided later in this chapter, for the time being it will be sufficient to refer to auto-correlation as an indication that two z values located nearby are somewhat "tied" to each other, so that we could expect that, once one of the two values is known, the other is likely similar, and thus easier to predict[2].

The presence of auto-correlation reduces prediction uncertainty; the higher the auto-correlation, the lower the randomness of the topography profile, or surface.

Spatial auto-correlation between height values of a profile or surface can therefore be elected as a measure of randomness for the topography. So for example, the complete absence of spatial auto-correlation can be assumed as an indication of high randomness (i.e. lowest predictability, maximum prediction uncertainty) of a topography. Albeit it will be shown that absence of auto-correlation is a necessary, but not sufficient, indicator of independence amongst z values at different x locations, it is usually still safe to say that in general if auto-correlation is absent, any z value will be independent of the location it is associated to, and it does not matter where points are located with respect to each other, if heights are to be considered. This is consistent with saying that a random sequence of events has no order.

4.2.1 Random topographies as random functions

Again from the theory of signal processing a simple mathematical formulation to describe random profiles and surfaces can be derived. A random topography profile or surface can be described by a random function of one or two variables $z = f(x)$ or $z = f(x,y)$. That is, a function where, despite the deterministic nature of the input arguments x (or x,y), the output must still be considered a random variable. This means that, the same function, with the same values for the input arguments, may generate different results. Using such a definition, it is possible to reconsider the "irregular" profile shown in Figure 4.1 as the product of a sequence of calls to a random function $z = f(x) = \epsilon$ (ϵ being the observation drawn from a random variable, entirely independent of x), each observation z_i being recorded as it is drawn in correspondence to the input argument value x_i. In contrast, the sine-like profile shown in Figure 4.2 can be now be reconsidered as the result of calling a deterministic function, for example of type $z = sin(x)$ each observation z_i being recorded in correspondence to the input argument value x_i[3].

[2] strictly speaking, this is an example of positive correlation (as opposed to negative, also known as anti-correlation). Positive correlation is the most natural case to perceive

[3] The plot in Figure 4.2 was indeed generated as a sine wave.

4.2.2 Randomness and F-operators

As stated earlier, topographies may exhibit varying degrees of randomness, i.e. they may be characterised by a wide range of different uncertainties associated to the prediction of previously unmeasured points. Like with time series, it is therefore often useful to decompose topography into multiple components, in order to isolate the more deterministic and more random ones, so that each can be treated with more suitable characterisation tools. Let us now consider the profile shown in Figure 4.3a. As for the random profile previously shown in Figure 4.1, it is still reasonable to assume that any estimation of the next value z_{n+1} will be associated to uncertainty. However, a quick visual inspection of the series in Figure 4.3 reveals that this time there is some more knowledge of where the next value will be, in that z will probably increase with larger x values. Visual perception picked up a *pattern* like in the sine-like profile shown in Figure 4.2, it was able to match part of the observation to a known shape pattern (a line slightly tilted upwards to the right of the page) and "see" it in the data. This is one of the simplest types of pattern: it's a trend, i.e. a measure of central tendency.

How can this "trend" be characterised? The simplest route is to use regression, and capture the trend by an equation, for example $z = mx + b$ in case it is safe to assume that the trend can be adequately represented by a linear function. The regression coefficients m and b can be found by solving a least-squares minimisation problem, where the space of the independent variables m, b is searched for a combination of values resulting in the smallest sum of squared residuals (i.e. distances of the observed x_i, z_i points to the $z = mx + b$ line). The result of such minimisation is the best-fit mean line, shown in Figure 4.3b. for the example case.

In the context of this discussion, it is at this point fundamental to point out that the best-fit line is a deterministic function. The function $z = mx + b$ is deterministic because once the regression coefficients m and b are known, the function output z is fully defined by the value of the input argument x, with no uncertainty associated to it. The deterministic function $z = mx + b$ represents a line, and the overall, random profile shown in Figure 4.3a can now be represented as a mixed deterministic/random function: $z = mx + b + \epsilon$, where $mx + b$ is the deterministic part, and ϵ is the residual, where all the randomness is now concentrated.

In Figure 4.3c the best-fit line has been removed, and residual series ϵ can now be considered independently.

The operation of removing the line is called de-trending in statistics and signal processing, but it is clearly an example of *F-operator* (see Chapter 2) in surface metrology.

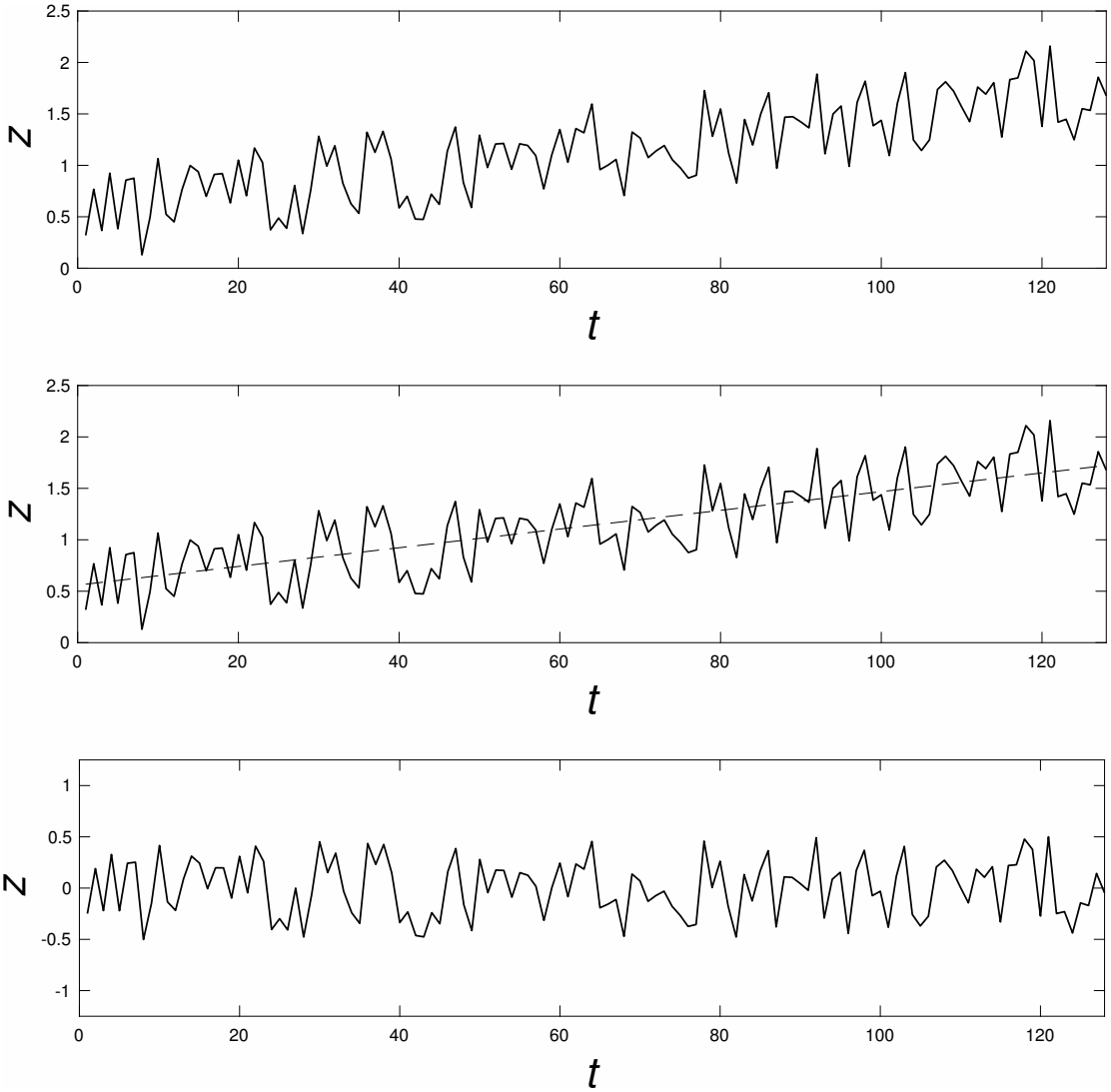


Figure 4.3: Random time series with a trend; a) original series; b) fitting to the least-squares mean line; c) de-trended signal by mean line subtraction.

A more general conclusion can be drawn in regards to the relationship between F-operators and determinism/randomness in topographies.

An F-operator implemented as a regression (to a deterministic function) has the effect of removing a deterministic component from a random topography. The residuals (i.e. what remains after the fitted regression model has been subtracted from the original dataset) is a "less deterministic" topography than the original one.

Usually, the application of an F-operator decreases spatial auto-correlation between topography datapoints (hence reducing determinism). It will be shown later that in many applications the ideal data processing pipeline is capable of removing any trace of spatial auto-correlation, leading to a residual of "pure" random, independent events.

4.2.3 Randomness and S,L-filters

Any operation that results into a transformation of a series of values (time series, topography profile or surface) may increase or decrease the random nature of the dataset, depending on whether or not the operation increases spatial correlation between the observations. Convolution smoothing for example (and as a consequence, low-pass filtering in the frequency domain, in general), increases the determinism of the filtered series with respect to the original, because convolution intrinsically increases spatial correlation of neighbouring points (e.g. see Figure 4.4). The result is also visually intuitive, as a smoother topography is representative of stronger spatial correlation of neighbouring observations, and consequently lower uncertainty in the prediction of values for the series in previously non-probed locations. The same filter, at the same time increases the randomness of the residuals.

The following can therefore be concluded in regards to the application of S-filters and L-filters to topography data:

S-filters (low-pass) implemented via convolution increase spatial auto-correlation and thus increase the determinism of the filtered topography. L-filters (high-pass) also implemented through convolution smoothing, but keeping the residual as the result, remove spatial auto-correlation and thus reduce the determinism of the filtered topography

Randomness and observational scale Spatial auto-correlation increasing determinism in topography is a rather intuitive concept. However, things get a little more complicated when one considers that spatial auto-correlation between two data points is often a function of the relative distance between them (again see later in this chapter for a mathematical illustration of this). If correlation depends on distance, all of a sudden factors such as spacing of the available observations, and distance of the location being predicted from the available observations, start playing a significant role. Typically, i.e. in presence of the most common types of auto-correlation, a prediction made at a small distance from available observations (where "small" is relative to sample spacing - i.e. distance between available observations) is more likely to be associated to a smaller uncertainty than a prediction made much further away. In other words, the degree of randomness of a topography of which we know something because of a point-based measurement process is pretty much dependent on the scale at which we are predicting new values relative to the scale of the sampling. It is therefore safe to say that scale of observation (relative to the scale of the spatially correlated features present on the surface being observed) determines the degree of randomness that can be associated to a topography. For example, a cylindrical surface may look very regular (i.e.

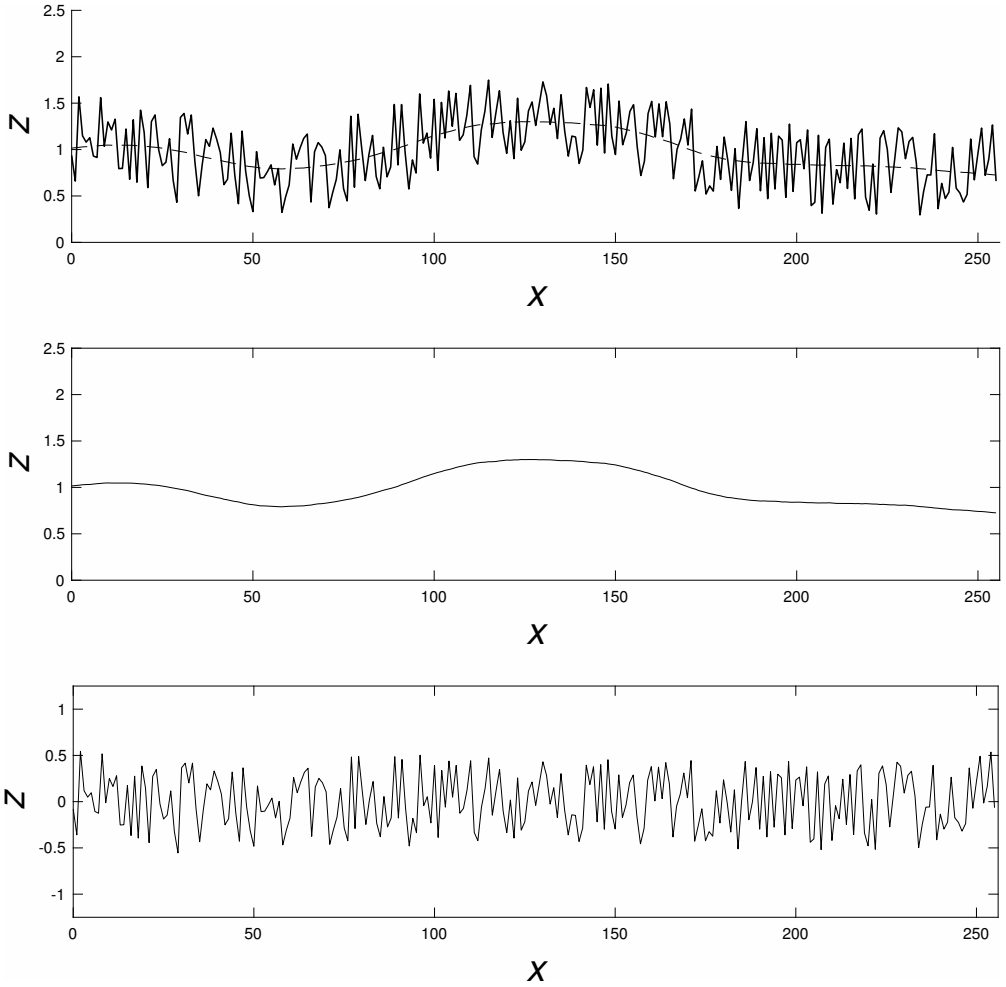


Figure 4.4: Convolution smoothing with a Gaussian kernel leads to a low-pass result more deterministic than the original topography (i.e. higher spatial correlation between neighbouring points), and to a residual more random than the original (lower spatial correlation between neighbours); a) original topography profile (the low pass result is overlaid as a dashed line); b) low-pass result; c) residual topography profile.

deterministic) if observed from far (i.e. large point sampling spacing). However, once observed from much closer (i.e. small point sample spacing) it may reveal previously unseen irregularities, ending up appearing much more random than initially thought.

From this point onwards it will therefore be given for granted that the adopted definition of randomness in topography, related to the uncertainty in the prediction of a new height value in correspondence of a previously non-sampled location, applies only to the specifically selected scale of observation, and is tightly related to the size of the deterministic features that may populate the surface, as well as by the point spacing adopted in the measurement process.

4.2.4 Randomness, determinism and characterisation pipelines

Although this concept will be discussed later in the chapter, for the time being it is already evident that mostly-random topography components and mostly-deterministic topography components are suitable to be characterised through different routes. Deterministic components are more predictable, hence they need fewer variables to be fully defined. Often, such variables are found in the domain of geometrical modelling. On the contrary, random components are more suitable to be characterised via statistical means. As seen in Chapter 2, conventional surface metrology is designed to remove deterministic components from topography, so that the residuals are as random as possible. At that point, if the conditions of "uniform, random roughness" are verified, it is possible to characterise the topography through texture parameters. Up to this point, albeit a formal definition of "uniform, random roughness" has not been given yet, it is clear that such condition refers to absence of deterministic components, and whenever possible, also removal of any residual spatial auto-correlation effect.

4.2.5 A quick recap of randomness vs determinism in topography

Randomness in topography refers to the uncertainty associated to the prediction of z values in locations x or x, y that have not been sampled by measurement. Randomness exists at varying degrees for any given topography, as all surfaces are typically composed of a mixture or more random and more deterministic elements. The presence of spatial auto-correlation between topography points increases determinism, because of the way humans perceive regularity and patterns. This is also reflected by how mathematical models represent geometry, favouring smooth and regular shapes. The applications of operators and filters to modify a topography dataset may alter the degree of determinism/randomness associated to the topography itself if auto-correlation between points is either added or removed. F-operators based on regression to deterministic models, for example, decrease the determinism of the residuals. S-filters implemented through convolution smoothing increase determinism of the filtering result, while L-filters (high-pass) decrease it. Randomness and determinisms are concepts tightly linked to observational scale. The relation between sampling distance and distance of the prediction from available measurements influences prediction uncertainty, and thus the degree of associated randomness.

4.3 Random fields to model topography

The definition of randomness illustrated this far refers to the uncertainty associated to predicting the value of a topography in a previously unprobed location. In such definition, the value of a measurement performed at a location, once obtained, is considered fully determined and immutable. However, if a topography is imagined as generated by a random function, probing at the same location x may result in different values every time. In physical terms, for a measurement performed on the same sample and in the same location, the random element could be traced back to measurement repeatability error. Alternatively, if one is measuring multiple physical instances of the same surface, then local randomness may be related to both measurement and manufacturing process repeatability.

Often times, randomness in surface generation and measurement automatically imply higher uncertainties in prediction. However, less intuitive scenarios could potentially be faced where a manufacturing process would generate surface instances that, if taken individually, are very deterministic, but there could still be high variability across instances. Let's consider for example an imaginary manufacturing process that generates a single hemispheric crater on an ideally flat surface through the mechanical force exerted by a ideally spherical, hard tool. Such an imaginary process would be capable of producing very regular and smooth crater footprints. Let's imagine the process having poor control on tool location and load: there would be randomness associated to crater location, depth and radius, that would vary for each new surface instance (i.e. poor repeatability). In terms of measuring one instance of surface generated by such process, one can easily see that the first measured points would be associated to low predictability (high randomness), because the position of the crater would not be initially known. However, soon the crater would be "discovered" by the measurement process, and after a sufficient number of points taken, the prediction for any new one would be characterised by extremely low uncertainties, due to the regularity of the surface (bar measurement error). In other words, the measured surface would be quickly classified as highly deterministic, despite having been generated by the equivalent of a random function. Randomness due to prediction and randomness due to surface generation would clearly be conflicting in such a scenario.

There is clearly a need to expand our statistical understanding of topography by introducing more general statistical models, that are capable of incorporating both randomness intended as uncertainty in predicting measurement results at unprobed locations, and randomness across multiple surface instances, or across multiple measurements of

the same instance (i.e. in both cases randomness associated to the same location x, independently of it having been already probed or not). In order to achieve that, *random processes* and *random fields* need to be introduced.

4.3.1 Topography as a random field

A (discrete) *random process* is a collection of a finite number of random variables where each is associated to a specific point in time, so that the collection is usually expressed as $\{Z(t_1), Z(t_2), \ldots, Z(t_n)\}$. A *random field* is a random process where time t is replaced with location in space x (for a one-dimensional field) or x, y for a two-dimensional field[4]. A one-dimensional random field $\{Z(x_1), Z(x_2), \ldots, Z(x_n)\}$ can be used to represent a topography profile, where each random variable describes height at the associated location. A two-dimensional one can be used to represent a surface. An example one-dimensional random field of n variables is shown in Figure 4.5. In the example, the random variables are assumed completely independent from each other, thus the entire field is fully described by a complete knowledge of the probability density functions (pdfs) of its variables.

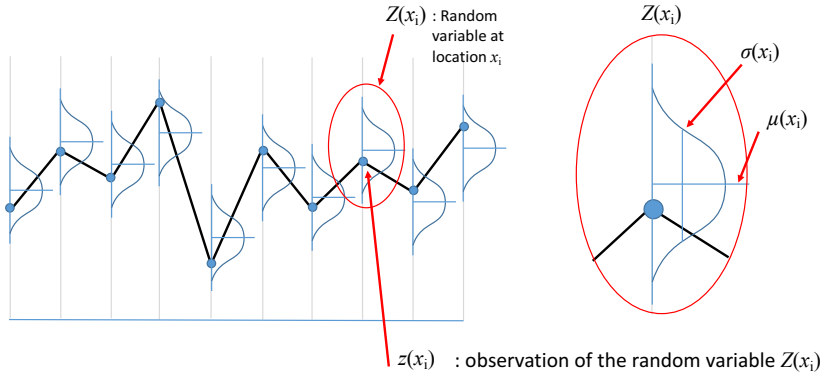


Figure 4.5: Topography profile modelled as one-dimensional random field with independent variables.

A measured topography profile $\{z_1, z_2, \ldots, z_n\}$ can be seen as the observed manifestation of a random field, i.e. a collection of actual height values, each being drawn from the corresponding random variable in the field (Figure 4.5).

In the simple example, each random variable $Z(x_i)$ is assumed characterised by a probability density function (with any shape, depending on the nature of the variable) and can be described for example by its first and second order moments (mean $\mu(x_i)$ and variance $\sigma^2(x_i)$ - the latter represented as standard deviation in figure 4.5). Because both the mean and variance are indexed on location, they can be thought of as functions

[4]the vectorial notation $Z(\mathbf{x})$ is also often used to represent location in a n-dimensional space

of location x as well. Note that obtaining a correct estimate of $\mu(x)$ and $\sigma^2(x)$ at each location x starting from the available observations $\{z_1, z_2, \ldots, z_n\}$ contributes towards solving the random field, but does not define it completely, because the pdfs associated to each $Z(x_i)$ are not fully represented by their first and second-order moments. However, if one could assume the pdfs as normal, then indeed mean and variance would be sufficient for a complete characterisation.

Finally, note that all the above considerations do indifferently apply to both profiles (x locations) and surfaces (x, y locations). Thus, the discussion in the following will refer to profiles to keep the notation simpler.

4.3.2 Introducing dependent random variables in the field

Regardless of whether the assumption of normal pdfs holds or not, one can immediately see that (with reference to Figure 4.5), after a profile has been measured once, the available observations are in any case insufficient to estimate the many unknowns that are necessary to reconstruct a complete model of the random field. Essentially, we are trying to estimate n independent random variables, having only one observation for each. In multivariate statistics, two dependent random variables are fully modelled by their *joint probability density function*. Once the value for one of the two variables is set, the *conditional probability distribution* for the other can be obtained.

A discrete random field of n dependent random variables is a collection of random variables indexed in space $\{Z(x_1), Z(x_2), \ldots, Z(x_n)\}$*, fully defined by a n-dimensional joint probability density function.*

For a discrete random field of n dependent random variables to be fully specified, the n-dimensional joint probability density function must be known.

4.3.3 Covariance in random fields

In a one-dimensional, discrete random field, covariance controls if two random variables at different locations $Z(x_i)$ and $Z(x_j)$ are somewhat "tied", meaning that they are observed as varying together (=0, they do not). Covariance is captured by the following equation:

$$k(x_i, x_j) = E\left[\left(Z(x_i) - \mu(x_i)\right)\left(Z(x_j) - \mu(x_j)\right)\right] \tag{4.3.1}$$

Covariance can be used as an indication of dependency of $Z(x_i)$ and $Z(x_j)$, albeit it is not a sufficient condition. More importantly, covariance can be used as a parameter to

describe some properties of the joint pdf of the two variables, albeit in general it cannot replace it (again because first and second order moments could have the same values for different types of pdfs).

For a finite number of n locations, $k(x_i, x_j)$ can be conveniently collected into a $n \times n$ *covariance matrix*. Note that for $i = j$ covariance becomes variance, so the same covariance matrix can contain also variance terms in its main diagonal.

A couple of considerations must be made:

- covariance and correlation are related concepts. Actually, correlation is a normalised form of covariance, since it is defined as $k(x_i, x_j)$ divided by $\sigma^2(x_i) \times \sigma^2(x_j)$. Differently from covariance, correlation is unit-less because of the ratio.
- covariance is sometimes called *auto-covariance* because it relates two variables belonging to the same field/process, as opposed to two variables belonging to two different fields

4.3.4 Topography profiles as Gaussian random fields

While an individual random variable associated to a generic pdf cannot be fully described by its first and second order moments, one associated to a normal distribution can.

A Gaussian random field is a random field where the n-dimensional joint distribution is normal, any marginal distribution is normal, and any linear combination of random variables extracted form the field has a normal distribution as well

The big advantage of the normality assumption is that the n-dimensional joint pdf can be fully defined by a complete knowledge of the first and second order moments of the random variables, which amounts to their means $\mu(x_i)$, their variances $\sigma^2(x_i)$ and covariances $k(x_i, x_j)$.

In summary, under the assumption that the pdfs of each $Z(x)$ are normal, a topography profile can be represented as a Gaussian field, and thus by:

- a mean function $\mu(x_i) = E[Z(x_i)], \forall i \in 1,2,\ldots,n$
- a $n \times n$ covariance matrix of elements $k(x_i, x_j), \forall i,j \in 1,2,\ldots,n$

4.3.5 Further assumptions to simplify the random field

Again by looking at the example in Figure 4.5, it is clear that by introducing covariance we have reduced the number of unknowns needed to fully define the model. For a one-

dimensional Gaussian field (topography profile) n values of $\mu(x_i)$ and $n \times n$ values of the covariance $k(x_i, x_j)$ matrix [5] are needed to fully define the model. While the estimation problem is simpler than before, there are still too many unknowns, again starting from the assumption that only one observation of the field is available (i.e. a measured profile, consisting of n sampled points).

To further simplify the random field, it is necessary to introduce the concept of *stationarity*.

In simple terms, stationarity means that some statistical properties of the random process/field do not change if observed within a moving window on the field/process, that is, they are shift-invariant.

Two main types of stationarity are typically considered for random fields/processes:

Wide-sense stationarity - WSS . WSS implies that both the first and second-order moments must be shift invariant. That is, two conditions must be verified: a) the mean $\mu(x_i)$ is the same for each location x_i, i.e. it is shift-invariant; b) the co-variance matrix is also shift-invariant, meaning that its formulation does not change if the indices i and j are modified as consequence of a shift along the profile. Usually, in this case the covariance matrix is defined as a function of relative distance between points (i.e. $i - j$), but not as a function of absolute position i or j. Of course shift-invariance of the covariance matrix implies that also variance is shift-invariant. A common formulation for a covariance function that only depends on distance (also referred to as *lag*) and not on absolute position is the follwing:

$$k(x_i, x_j) = \sigma_0^2 \exp\left(-\frac{\|xi - xj\|^2}{2l^2}\right) \tag{4.3.2}$$

The equation is referred to as the squared-exponential formulation for covariance. the quantity σ_0^2 is the common variance shared by all the random variables[6], the numerator contains the Euclidean distance between locations, the denominator contains the quantity l, the *characteristic length* used as reference to indicate that variables located at larger distance will be uncorrelated.

Strict-sense stationarity - SSS . SSS implies that the full n-dimensional joint probability distribution is shift-invariant. This is stronger than WSS because in a generic pdf is not fully defined by its first and second order moments. However, if the pdfs are

[5] the covariance matrix is usually symmetric

[6] it must be the same for all variables because of the shift-invariance of the covariance matrix

normal (like in Gaussian fields), then SSS = WSS (because the first and second order moments are enough to fully define the distributions).

For a Gaussian random field, stationarity means that both the first-order moments (i.e. the mean function $\mu(x_i)$) and the second-order moments (i.e. the covariance function $k(x_i, x_j)$) are shift-invariant, that is, they do not change if computed within a window that moves over the locations across the field. For a Gaussian random field, there is no difference between wide-sense and strict-sense stationarity.

White noise . Stationarity leads to a significant simplification of the problem of fitting an observed profile or surface to a random field, because the number of unknowns to be estimated is significantly reduced. However, things can go even better if one could assume the mean function is zero at all locations ($\mu(x_i) = 0, \forall i, j \in 1,2,\ldots,n$), variance is constant and the same at all locations ($\sigma^2(x_i) = \sigma_0^2, \forall i, j \in 1,2,\ldots,n$), and all the random variables are uncorrelated for all lags ($k(x_i, x_j) = 0, \forall i, j \in 1,2,\ldots,n$, with $i \neq j$). If the above conditions are verified, the random field is commonly referred to as *white noise* in the signal processing literature. The definition holds identical for topography profiles and surfaces[7].

4.3.6 The role of stationarity in surface metrology

Let us go back to the original one-dimensional, discrete random field shown in Figure 4.5 as representative of a measured profile. As a first simplification, let us introduce a stationarity element by assuming that the mean is shift-invariant. For simplicity let's assume that $\mu(x_i) = 0, \forall i, j \in 1,2,\ldots,n$. Second, let's assume that all the random variables $Z(x_i)$ have the same type of pdf (for example, normal, but it does not need to be). Third, let's assume that all the variances are the same $\sigma^2(x_i) = \sigma_0^2, \forall i, j \in 1,2,\ldots,n$. These three assumptions lead to the result illustrated in Figure 4.6.

If the further assumption that all the random variables $Z(x_i)$ are independent is introduced (see Figure 4.7), two important considerations are drawn: a) all the covariances, for all lags, are zero ($k(x_i, x_j) = 0, \forall i, j \in 1,2,\ldots,n$, with $i \neq j$), and neither the location order, nor the distance of the random variables is important any longer.

After all the above assumptions, it can be concluded that all the random variables are identical and independent. Thus, all the available observations, which were originally obtained by drawing one individual values from each individual random variable, can now be safely assumed as n independent draws from a unique random variable. The

[7] note that the definition of white noise does not necessarily imply the pdfs are normal. If they are, then the assumptions of white noise lead to all pdfs being identical

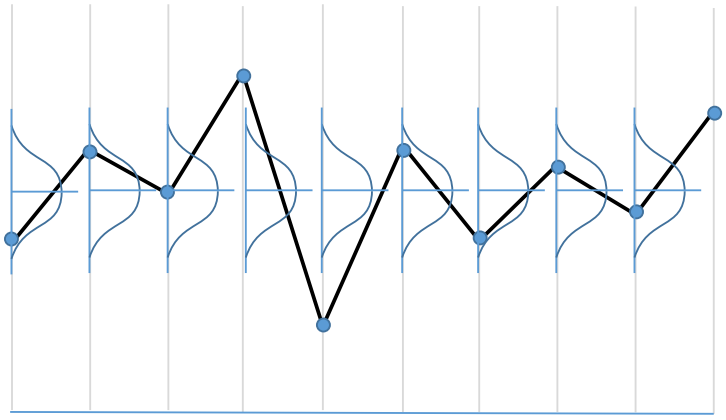


Figure 4.6: Random field under the assumptions of shift invariance of the mean, variance and shape of the pdfs.

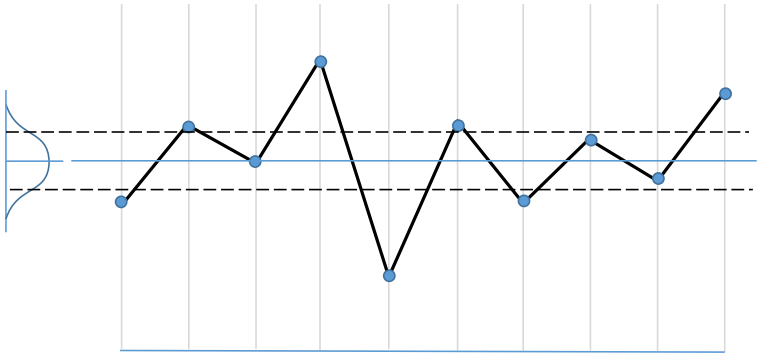


Figure 4.7: Random field under the further assumption of independence of the individual random variables.

random field is now showing properties of *ergodicity*.

A random process/field is ergodic if a sample obtained by collecting observations at different times/locations has identical statistical properties of a sample obtained by collecting multiple observations at the same time/location. So we can use a single "observation" of the process to infer properties of the entire process even without repeated realisations of the entire process.

The following conclusions can be drawn about texture parameters as they are defined in the ISO standards.

Height and functional parameters versus ergodicity Height and functional parameters are defined by implicitly assuming the random field is turned into a state where it can be defined as white noise, plus the additional assumption of independence between random variables. This state is implicitly an ergodic form, and texture parameters work because they exploit the ergodic assumption, that is they can reach to a full characterisation of the random field by simply aggregating observations at different

locations, belonging to the same individual realisation of the field. Of course, to reach to this state, suitable regions must be measured, by careful selection of range and resolutions, and appropriate combinations of F-operator and S,L filters must be applied in order to remove any source of spatial auto-correlation (i.e. removing determinism, see earlier in this chapter).

Spatial, miscellaneous and hybrid parameters, and stationarity Spatial, miscellaneous and hybrid parameters require stationarity within the field of view to produce meaningful results. Without constancy of the mean, there would be a trend that must be removed. Without constancy of the standard deviation, the results would not be interpretable. Without shift-invariance of the covariance function, the results would again be not interpretable. Concerning covariance, it is particularly important to point out that spatial, miscellaneous and hybrid parameters usually provide meaningful results only if covariance is shift-invariant, but different than zero. Without covariance, point location would not be important any longer, and only height and functional parameters would be safely assumed as carrying some meaning.

4.4 Characterisation of non-stationary fields

In a non-stationary field, the mean function $\mu(x_i)$, the covariance function $k(x_i, x_j)$ and/or the shape of the joint pdf (in case of non-Gaussian fields) may not be position-independent. When this happens the amount of unknowns in the estimation problem (i.e. trying to estimate the parameters of all the random variables involved in the field, starting from a single observation for each of them) makes the fitting problem too hard to solve. Simplifications can be made by assuming specific deterministic models for the mean function (for example, a linear relationship like $\mu(x_i) = mx_i + b$ for profiles) and one of the most common formulations for the covariance function $k(x_i, y_j)$, such as the squared-exponential previously illustrated in Equation 4.3.2[8].

However, in terms of the typical application scenarios of surface metrology, that would imply topography as a non-stationary random field, two principal situations would usually arise:

– one is interested in isolating regions where topography can be deemed stationary, so that the conventional route of computing texture parameters would still be viable;

[8]Which, strictly speaking, would apply to a stationary co-variance, i.e. function of the lag between x_i and x_j, but not of absolute position

– one is really interested in studying the non-stationarity of the topography.

In the first case, measurement should be planned to focus on the stationary regions as much as possible from the beginning; the F-operator and S,L filters should also be designed to remove any residual source of non-stationarity (compatibly with the required observational scope of the application) so that texture parameters could be computed. Specifically, the F-operator would be in charge to de-trend the field by removing a sizeable chunk of determinism from the $\mu(x)$ function, while the S and L filters would primary be in charge of removing any residual hint to co-variance if possible (ideal scenario for height and functional parameters), or at least remove enough to allow for reasonable handling of the residual co-variance via spatial, miscellaneous and hybrid parameters. However, if the interest is really on the non-stationary part of the topography, or in the transition between stationary regions, a different strategy than texture parameters must be pursued.

Common non-stationary scenarios

The non-conventional application scenarios previously illustrated in Chapters 1 and 2 can now be re-introduced from the viewpoint of randomness/determinism, and using random fields.

Most of the scenarios involving localised topography features of interest (grooves, scratches, individual pores, regular structures, etc.) can be classified as random fields where the relevant non-stationary elements are described by local changes in the mean function $\mu(x)$ (or $\mu(x,y)$ of course). Such changes may be abrupt (sharp feature edges) or gradual (smoother transitions); usually affect medium to large-sized regions (i.e. happen at comparatively large wavelengths, i.e. they produce features that are large enough to be visible within the FoV), and appear quite deterministic, i.e. characterised by significant auto-correlation (i.e. regions of the feature surfaces appear relatively smooth). An example of this type of random field is illustrated in Figure 4.8 where a topography profile is illustrated whose mean function $\mu(x)$ slowly transitions between two stationary conditions. Scenarios where non-stationarity is represented by a smoothly (or sharply) changing $\mu(x)$ function, suggest characterisation routes that focus on the deterministic part of the random field, the $\mu(x)$ function, treating the remainder as a disturbance that must be properly handled.

The other family of application scenarios involves transitions between textural states that are not necessarily interpretable as a change of $\mu(x)$, but could be characterised by a change in higher-order moments. A simple example is shown in Figure 4.9 where the

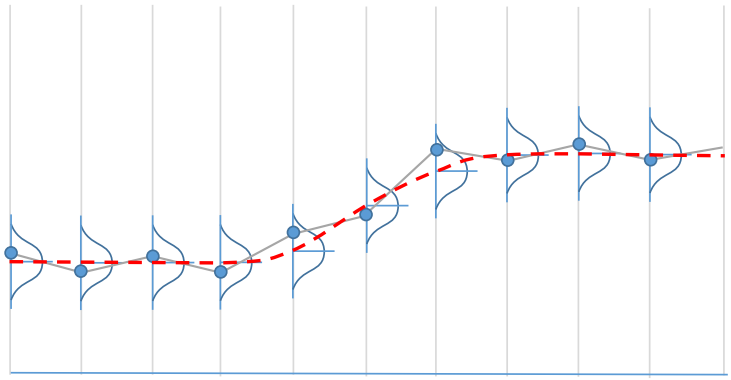


Figure 4.8: Example topography profile depicted as a random field of non-stationary mean

second-order moment variance smoothly transitions between two stationary states.

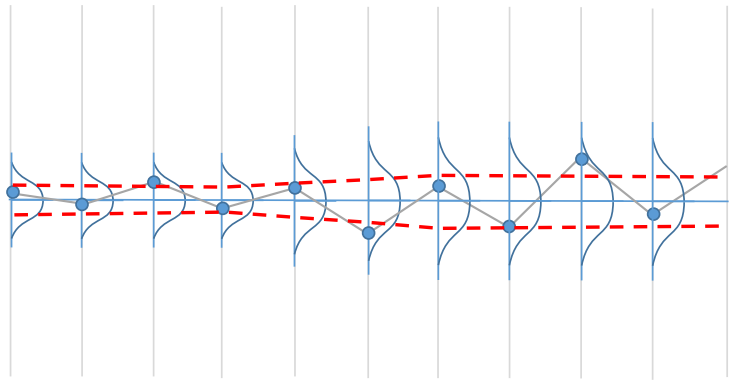


Figure 4.9: Example topography profile depicted as a random field of non-stationary variances

Typical real-life scenarios that can be thought of as described by non-stationarity of higher-order moments have been hinted at in Chapter 1, and would usually involve multi-modal textures, albeit typically, even for multi-modal texture, non-stationarity is likely to involve also the first-order moment (i.e. also the mean function $\mu(x)$).

As hinted at earlier, predominantly deterministic topographies favour the use of specific families of characterisation tools, mostly based on deterministic shape encoding and characterisation, while predominantly random ones appear more suitable to be pursued via statistics-derived descriptors. Such two routes will be explored in the next two chapters (i.e. Chapter 5 and 6). For the time being, it is important to point out that they share an important, common aspect: the need for spatially partitioning the random field (i.e. the topography) in order to successfully pursue the characterisation of the transition, and/or of the regions it divides. This is the subject of the third and

last core theme discussed in this chapter.

4.4.1 Linking random fields to randomness intended as uncertainty in prediction

A problem that is often dealt with random fields/processes is function estimation, which in this case becomes a topography estimation problem. In such case, the goal is to find the complete topography of which we only know the available observations $z(x_i)$, using the random field to estimate the missing observations z at unprobed locations x. So, the goal is to estimate one of the *realisations* of the random field. Of course, to achieve such goal, knowledge of the random field (and thus of its random variables $Z(x_i)$ and joint pdf) is still required. In case of Gaussian fields, both the mean function $\mu(x_i)$ and the covariance function $k(x_i, x_j)$ must be estimated before a prediction can be attempted at unprobed locations. As illustrated earlier, this is a complex task that is generally approached by making assumptions on the nature of both the mean and the covariance function, and then solving them as an optimisation problem, for example as a maximum-likelihood estimation, where the best result is the one that maximised the likelihood of observing the available data points. Whilst a full coverage of fitting to a random field is beyond the scope of this text, a simple example of the results that can be obtained is illustrated in Figure 4.10. The grey band represents the *prediction interval*, i.e., the interval of z values within which the topography may fall with a given confidence level (e.g. 95% confidence). The width of the prediction interval is therefore a measure of the uncertainty associated to predicting $z(x)$ at an unprobed location x and therefore ties directly into the concepts of randomness and determinism illustrated in the first section of this chapter.

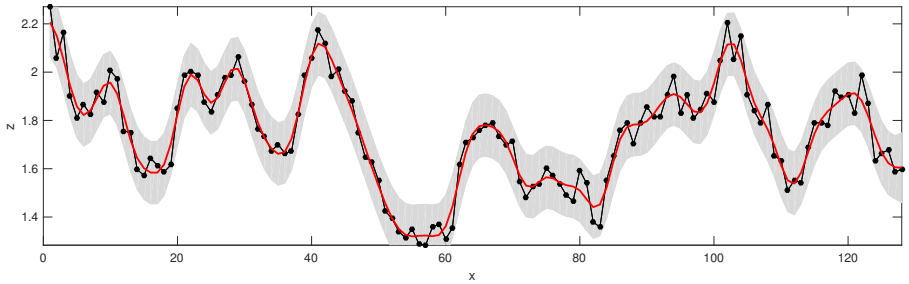


Figure 4.10: Gaussian field fitted to an observed profile dataset to predict height points at previously unprobed locations; black: measured profile; red: mean function of the Gaussian field; gray: band of the prediction interval (at 95 % confidence level).

4.4.2 Random fields versus moving statistics

At this point, one may wonder whether the increased complexity of introducing random fields provides a clear advantage over simply computing moving statistics. For example, stationarity of first or higher-order moments could be simply computed using moving means, moving variance, etc., starting from a single series of observations $\{z(x_1),z(x_2),\ldots,z(x_n)\}$. That is, once defined a fixed-size window of adjacent observations, one could compute any order statistics by simply moving the window across the entire field. The same technique can be effectively used also to separate more deterministic components (the moving mean is a convolution operation) from more random residuals. As a matter of fact, in many of the topography analysis and processing techniques introduced in the next two chapters, local function estimators based on observations captured within moving windows will be frequently used as opposed to the more complex statistical estimators routinely used in fitting random processes and fields. There are a couple of situations where random fields are particularly useful though:

- Manufacturing processes are often quasi-periodic in nature in terms of the surface signatures they generate, that is, it is often possible to identify a characteristic length over which the topography of a surface repeats its salient traits. A typical example are machining marks in a turning process. A more general representation of such surfaces with respect to simply measuring a portion of them, may consist in measuring multiple regions over the same characteristic length, and them overlaying the observations as multiple realisations of the same random field. In this way, a more comprehensive statistical model of the topography could be obtained (e.g. see Figure 4.11).

- Measurement processes are characterised by multiple sources of error, some of which are random in nature. By measuring the same region of a surface multiple times, in repeatability or reproducibility conditions, and by overlaying the resulting datasets, one could see the measurements as multiple realisations of the same random field, the analysis of which, would shed new light on the metrological properties and behaviour of the measurement process (e.g. see Figure 4.12). This later concept is explored in detail in Chapter 7.

An example of the second scenario is shown in Figure 4.13 (on simulated data)

Finally, even if replicate observations at each location x are not available, that is, we only have one observation $z(x_i)$ at each location x_i (which is often the case for profile

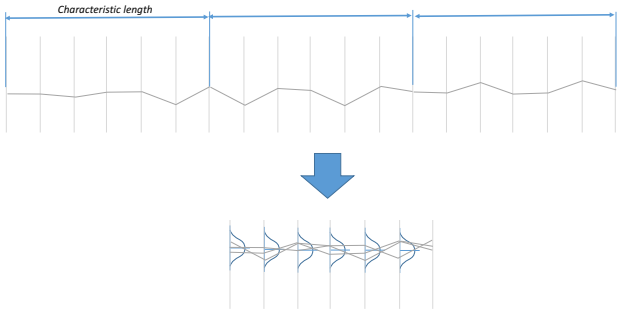


Figure 4.11: Random field built from overlaid measurements of different regions each covering the characteristic length of the topography

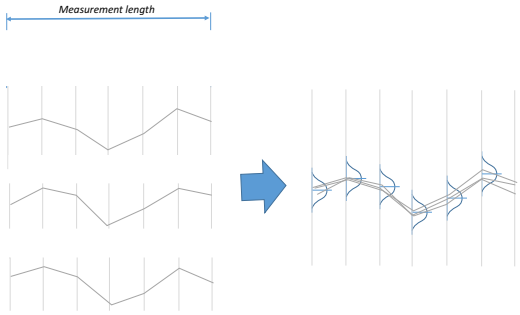


Figure 4.12: Random field built from overlaid measurements of the same region.

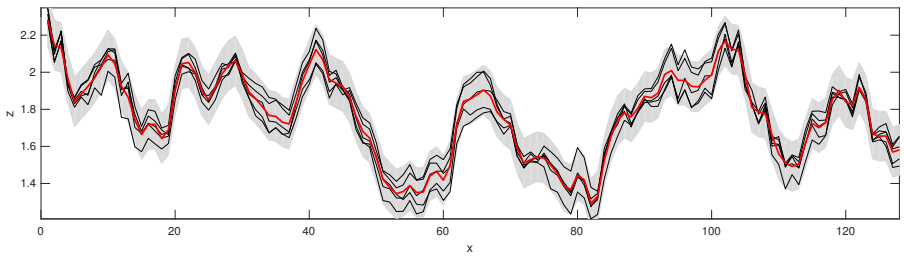


Figure 4.13: Replicate profile measurements (simulated) fitted to a Gaussian field; original profiles shown in black; mean function shown in red, prediction band shown in grey (at 95 % confidence level).

or areal topography measurement), random fields have the additional advantage of being capable of estimating unknown z values at unprobed locations x with associated estimation uncertainty, something that simple "deterministic" interpolation cannot do. This possibility was already illustrated in Figure 4.10.

4.5 Feature technology

4.5.1 Shape complexity, spatial differentiation and feature technologies

Surface topographies usually provide quite striking examples of geometric complexity. The intricate networks of hills, dales, crests and other formations that typically populate a surface, also usually at multiple scales, make topography one of the most complicate geometric entities imaginable.

It is fairly easy to describe and fully specify a regular shape. A sphere only needs the co-ordinates of its centre, and a radius; a cylinder is fully specified by an origin point, a segment specifying axis orientation, a length and a radius, etc. As the shape becomes more complex, it gets increasingly more difficult to fully describe it. A generic, irregular surface topography is usually too complex to be suitable for full, comprehensive description. Since topography is usually available as raw measurement data (point clouds, height maps or triangle meshes), it is rather common to hold onto the raw data, as the safest and most robust form of representation of the shape itself. This is technically called a low-level representation of geometric knowledge, which - whilst safe and robust - does not offer that many options to do shape-based reasoning, including simple qualification of shape properties, or comparison. Texture parameters can be seen as an attempt to handle such complexity by introducing summary indicators mapping simple scalar values to topographic complexity. As illustrated previously, texture parameters are only useful for a very specific subclass of topographies, i.e. stationary ones.

Spatial differentiation is a common procedure adopted by the human brain to handle complexity. For example the light patters collected through the retina in our vision system allow for the low-level image data to be decomposed into regions; scene understanding is then aided by the analysis of region shape and spatial arrangement [63]. Studies on perceptual organisation date back to the Gestalt principles and related disciplines [64]. Like for humans, spatial differentiation has also been recognised as a fundamental milestone of machine understanding of shape, and as the gateway to machine reasoning based on shape information. The concept of spatial differentiation exists in various forms in many disciplines involving processing and analysis of low-level information, such as in signal processing, image processing, and even geometric modelling. Differentiation between surfaces and regions of CAD geometry is a key element in many computer-aided technologies that support product design and manufacturing [65], such as CAD/CAM, and CAE ranging from finite-element analysis to multi-body modelling.

In geometric modelling, the term *feature technology* is often used to indicate any conceptual approach where low-level information is spatially partitioned and re-organised so that higher-level understanding is possible [65]. Feature-technologies have come to indicate, over the last decades, any approach that, in partitioning the geometric model (low-level information) into regions, and in assigning tags to such regions, allows for augmenting the original dataset with multiple higher-level information layers, each related to a specific, relevant application context.

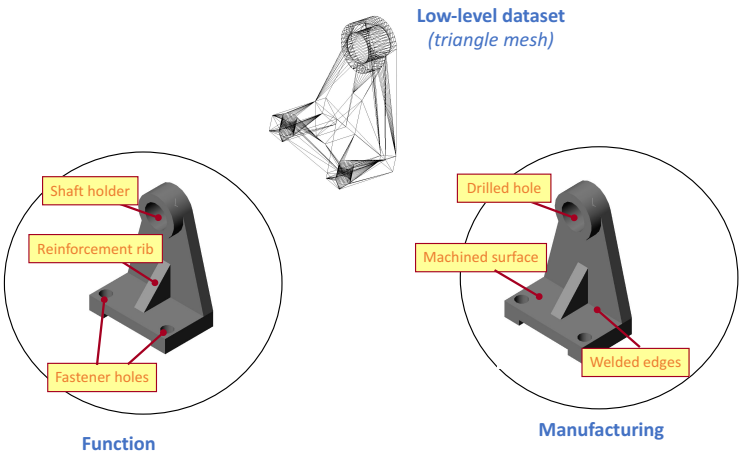


Figure 4.14: Feature technology as envisioned in the world of geometric modelling. The same low-level geometric model (e.g. a triangle mesh) is augmented by the addition of higher-level information layers mapped to portions/regions of the geometry.

4.5.2 Definitions

As the example in Figure 4.14 illustrates, there may be multiple higher-level layers associated to the same lower-level geometric content (triangles, edges, vertices). Within each layer, portions of the low-level dataset are isolated from the remainder of the geometry, and associated to a "meaning", or "role" related to a specific analysis context. The term *feature* or *surface feature* or *topographic feature* is used in this work to refer to any of such regions, i.e. a topologically connected subset of the original topography dataset, of some interest in the context of some application. Each feature (region in the topography dataset) is mapped to a higher-level role or meaning because it has been recognised as *instance* or *manifestation* of a more general *class* or *category* of regions that, because of some of their size/shape/texture properties, has been defined as relevant within a specified application context, and therefore has been defined as part of a specific higher-level knowledge representation associated to such surface.

So for example, within an user-defined knowledge model (i.e. an *ontology*, [66]) de-

scribing all possible damages a specific class of surfaces may undergo within a given application context, a user-defined class/category may be that of *scratches*. Within a measured surface dataset, local regions, because of their shape/size or textural properties, may be recognised as instances of such class, and thus mapped to it. Together, such regions may form a higher-level information layer that lives along the original low-level dataset. The new layer would be referred to as a *feature-based description* of that topography, and would be applicable, and spendable, in the specific application context it is referred to.

The following considerations must be made:

- the concept of topographic feature is not entirely new in surface metrology. Even the standard ISO 25178-2 talks about "features" (see Chapter 2). However, ISO features are specific to only one viewpoint, or *knowledge model* as it is being referred to in this work. So the only possible features are hills and dales (areal features), ridge and course lines (line features), peak, pit and saddle points (point features);.
- mapping topography regions to features implies a spatial partitioning of the dataset. This is also referred to as *subdivision*, or more commonly *segmentation* from the discipline of image processing, a *segment* being a subset of topologically connected surface points (i.e. a region);
- a partitioning process may imply different relevance of the resulting partitions: for example when a feature is discriminated from its surroundings. Alternatively, multiple regions may have the same relevance.

Feature identification is the act of recognising that a region of topography is an instance of a class being sought for. That is, feature identification is the act of establishing a mapping between a region of the topography and a class belonging to a knowledge model.

Feature extraction is a purely geometric operation, consisting of isolating the topography region representing the feature from its surroundings, and representing it as an independent topography/geometry dataset.

Feature characterisation is the act of describing the shape and size properties of the feature. Notice how, according to the above definition, nothing stops the whole measured region to be considered as a feature itself. So characterisation is the act of describing any topography region, including the originally measured one. The main differentiating factor when describing regions is the amount of deterministic vs random content, which influences the preferential characterisation routes.

Since a proper description of a region is fundamental prerequisite for identification, chapter 5 covers characterisation. Identification and extraction will be covered in Chapter 6, whilst Chapter 7 will cover uncertainty.

4.6 Summary and outlook

In this chapter an original approach has been presented, where the characterisation of surface topography is approached as a statistical modelling problem. More specifically, topography characterisation is depicted as the estimation of a random field, a problem which is in general difficult to solve starting from available observations (measured datasets) because of the large number of unknowns. In addition, a new perspective on the meaning of randomness in surface topography has been presented in this chapter, where randomness is described both as uncertainty of estimation at unprobed locations of the same observation for which some data points are known (individual dataset), and as variability between entirely new surfaces (topography instances - observations) extracted from the same random field. It is also shown that the two aspects are correlated. The concepts of stationarity and ergodicity, properties of a random field, have been introduced to establish an original correlation between the nature of the surface being investigated, and the degree of applicability and meaningfulness of ISO texture parameters. Stationarity and ergodicity estimation can now be used to infer applicability of conventional characterisation routes to any type of surface topography. Types of non-stationary behaviours and their relationships with the non-conforming topographies originally introduced in Chapter 1 have been discussed as well. The need to manage the complexity of non-stationary random fields has been used in this chapter as the original motivation behind justifying the need for spatial partitioning of the random field, one of the fundamental, conceptual pillars of feature-based characterisation. Finally, a formal definition of what is meant by feature has been provided based on the concept of partitioning.

The overall theoretical apparatus proposed in this chapter is able to accommodate both conventional surface characterisation via texture parameters and feature-based characterisation, showing that characterisation based on texture parameters can be actually seen as addressing a special case of a more general problem. Despite such accomplishments, additional research efforts are needed to dig further into understanding how non-stationary topographies could be handled in a more comprehensive and statistically-sound manner, in addition to only defining the conditions that imply loss of stationarity. Another issue that may require increased attention in the future is related to the computational burden that is typically associated to fitting a random field

model to available observations (measured datasets), in particular since the current technological trend sees the introduction of areal topography measurement systems capable of collecting an ever growing number of data points.

Chapter 5

Feature characterisation

Shape characterisation is the act of describing a shape/geometry in mathematical terms. Characterisation is fundamental for storage and retrieval, and essential to those activities where similarity or difference between shapes must be quantified, e.g. in comparison and classification.

In this chapter, the problem of how to characterise shape/geometry is discussed in the specific context of topographic features. This problem is here referred to as *feature characterisation*.

Since in Chapter 4 a feature has been defined as a generic portion of a measured topography (i.e. a region)[1], there is formally no difference between characterising any region located within a measured topography, and the complete measured region itself. In any case, topography, must be generally assumed as a complex geometric entity. Thus, the problem of feature characterisation is a problem of characterising complex shapes.

5.1 Approaches to the characterisation of topography

Elementary, deterministic geometric entities are easy to characterise. An ideal circle (Figure 5.1.a) fully defined by its centre co-ordinates and a radius. Three scalar values (in a two-dimensional space) are therefore fully sufficient to entirely describe the set of infinite points laying on the circle. This is the scenario typically encountered in *geometric modelling*, where geometries are described in terms of simple deterministic archetypes, easily summarized through a finite, short set of scalar descriptors (e.g. dimensions). On the other hand of the spectrum, a completely random topography profile (i.e. stationary random field of independent variables) can only be defined in

[1] only *areal* features are considered in this thesis, i.e. no *point* or *line* features

statistical terms (e.g. via statistical moments such as mean and variance, for a Gaussian field) (see Figure 5.1.c). And yet, as discussed earlier, despite having a full model of the random field, there will still be uncertainty associated to the prediction of an infinite number of topography points located between the available observations. This is the scenario typically covered by surface topography modelling via texture parameters, as discussed earlier in Chapter 2.

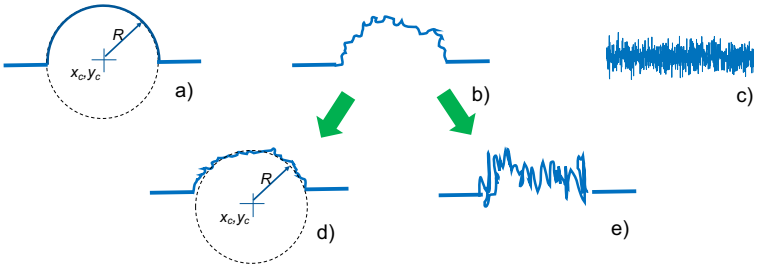


Figure 5.1: Example shapes (topography features) with different degrees of determinism/randomness; a) fully deterministic, archetypical shape (circle); c) pure random profile (stationary random field with independent variables); b) intermediate state encompassing all the situations where the feature is comprised of a part-deterministic and part-random topography. The intermediate state encompasses all the possible variants between the extremes: c) strong deterministic component, the random component is seen as small-scale "error"; d) strong random component, the deterministic component is barely appreciable.

.

Between fully-deterministic and fully-random topographic features there lays an intermediate class of shapes that are part-deterministic and part-random (Figures 5.1.d and 5.1.e). This is where topography features typically are, with characterisation choices depending on the balance between deterministic and random components, verging towards one or the other depending on circumstances and characterisation needs. The field of *dimensional metrology* for example, considers shapes affected by very little deviation from nominal (Figure 5.1.d), so most characterisation efforts are based on developing descriptors of deterministic geometry with robustness to smaller-scale random deviations. This is legit in typical industrial inspection and verification scenarios involving standard-sized parts, as manufacturing process performance is comparatively good, and errors are typically at least one order of magnitude smaller than the features being fabricated. Most of the target features that are covered in this work though, would still be comprised of a deterministic part, but would often be affected by much larger random components, leading to the necessity of developing shape characterisation solutions that cannot be entirely based on paradigms typical of dimensional metrology (because of the larger error components - Figure 5.1.e is an extreme situa-

tion). The presence of much larger error and variability across feature instances has already been discussed in Chapter 1, where it was already mentioned how error could in some circumstances be the same order of magnitude of the size of the feature being investigated.

In summary though, the "non-conventional" application scenarios originally introduced in Chapter 1 are comprised of topography features that fall in the wide range of situations encompassed by the extreme cases depicted by the *d* or *e* categories in Figure 5.1. The following original topography feature characterisation approaches have been devised:

- *CAD-compare methods*. Inspired to the CAD-compare methods developed for the inspection and verification of standard-sized parts.
- *Datum-based methods*. Based on computing reference datums, and the defining topography features through size and localisation properties of the datums.
- *Shape resampling and integration methods*. Based on resampling the topography in the surroundings of a centre-point and on integrating the resulting height values according to some scheme.
- *Multi-scale decomposition methods*. Based on processing the dataset in order to extract specific information at multiple scales.
- *Spatial occupancy enumeration methods*. Based on subdividing the volumetric space of existence of the topography feature and then encoding the full and empty partitions.

The most important take-away point from the previous section is that the amount of randomness/determinism in a topography feature affects the characterisation route that may be undertaken, and thus the choice of what is important to define, in order to describe a given topography region. The above-listed methods are best suited to operate with different amounts of randomness associated to the feature topography. Also, whilst some methods aim at reducing the feature information content to a short series of (summary) descriptors for ease of classification and comparison, other methods actually multiply the amount of information available by simply producing multiple datasets each highlighting specific aspects of the original one.

5.1.1 Data preparation for feature characterisation

What topography data should feature characterisation start from?

One could start from the available, individual measurement data, i.e. a dataset in the form $z_i = f(x_i)$ or $z_{i,j} = f(x_i, y_j)$. Or one could consider such dataset as an observation drawn from a statistical model, and could base the characterisation on the extracted deterministic part, or on the $\mu(x_i)$ function extracted from the resolution of the random field. Moving towards more "random" topographies, for example assuming some form of stationarity where $\mu(x_i)$ is constant, then one could build a characterisation solution starting from the analysis of the covariance matrix $k(x_i, x_j)$, thus probably ending up electing some statistics as descriptor, instead of analysing an actual geometric model.

It is unlikely that one predefined solution may work in any case. The choice of what dataset to start from may depend on the nature of the topography itself (see previous discussion and Figure 5.1), and on the characterisation objectives. Moreover, as demonstrated in Chapter 4, the introduction of feature technology allows one to partition a complex dataset into regions, so that each one may be suitable to undergo characterisation following a different route.

The amount of possibilities increases dramatically if one considers that, whatever information is elected as the starting point (e.g. the observation $z(x_i)$, a fitted deterministic model, the mean of a random field $\mu(x_i)$, etc.), it could be further tweaked into more deterministic or more random forms by the application of F-operators and/or S,L-filters, as shown in Chapter 4. For example, as previously illustrated, both fitting to deterministic models and smoothing to achieve higher spatial correlation, contribute towards generating a more deterministic reference that can be used as the basis of characterisation. Applicable techniques range from least-squares fitting to low-order polynomial curves and surfaces and fitting to splines or b-splines, or similar parametric representations. Even non-parametric regression techniques, such as LOESS and LOWESS [67, 68] can be used to represent surface information, as discussed in previous work by the author of this thesis [69]. Non-parametric regression has recently been proven useful to obtain a more deterministic reference to rely upon as the basis for feature characterisation [37] and in general to better capture the deterministic part of freeform topographies in a range of applications [69]. To an extent, even the low-pass result of any convolution smoothing operation, including a simple moving average, may be used as reference for characterisation.

In all those approaches where a more deterministic component is extracted from the available dataset for use in the characterisation pipeline, the more random residual may not necessarily be discarded. In Chapter 7, simple examples will be given on how it can be used as a source to estimate error associated to characterisation, and consequently useful in the construction of an uncertainty budget.

5.2 CAD-compare methods

When the topography feature is highly deterministic and the product of explicit design intent (that is, there is a nominal specification describing what shape and size the topography feature is supposed to have), then an equivalent approach to what is colloquially referred to as "CAD-compare" in the inspection of standard-sized parts may be an option. In such a scenario, shape is described as the composition of a deterministic component matching the nominal geometry (e.g. a CAD model, provided for example as a triangle mesh, to represent the design specification) and a "residual" component capturing the discrepancies with the nominal geometry as spatially mapped "error vectors", locally representing the distance between the measured and nominal surfaces. The entire feature shape can therefore be described by the combination of the nominal model and an error map encoding the discrepancies of the measured dataset and the nominal one. An example application of such technique is illustrated in Figure 5.2) (from original work by the author [36]). A topography dataset measured from a surface tile obtained by micro-embossing is shown. The CAD model (triangle mesh) representing the nominal design for the tile is aligned to the measured dataset by converting the latter (originally a height map) intro a triangle mesh, and then by performing a global alignment operation in 6 dof (Figure 5.2.a). The vertical distance (i.e. z distance) between the measured and nominal surface, computed by projecting z-rays from the measured points onto the triangle mesh (as explained in Section 3.6), is used as a local indication of average discrepancy within each height map "pixel" (of $dx \times dy$ width, see Chapter 3), and rendered in volumetric form (Figure 5.2.b).

The CAD-compare example approach shown in Figure 5.2 applied to a topography feature is slightly different from how CAD-compare typically works in the inspection and verification of standard-sized parts. In the latter case, measured data is provided as either a triangle meshes, or as unstructured point sets, geometries are fully three-dimensional, and error can be computed from distances measured orthogonally to the local normals. In the characterisation of surface topography features instead, the discrepancy between the measured feature and its nominal counterpart is preferentially computed by considering the vertical (i.e. z-oriented) distances only[2].

Challenges of the CAD-compare approach to the characterisation of topography features are for the most part associated to the alignment of the CAD model (e.g. a triangulated mesh) and the measured feature topography (a height map, turned into a

[2]because of the nature of height maps and the surface topography measurement process, distances other than z-oriented would not be reliable to obtain

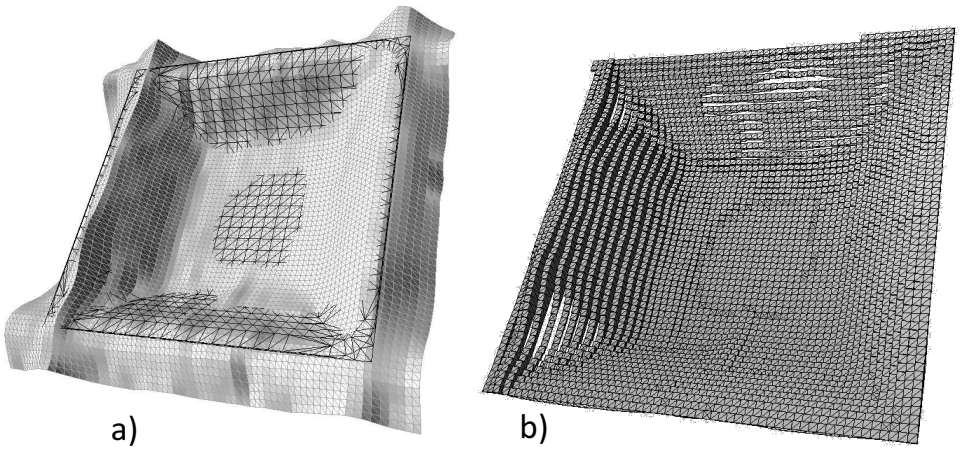


Figure 5.2: Characterisation based on the CAD-compare approach for a micro-embossing dale ($(280 \times 280 \times 50)\mu m$ approximate size, measured by a single-point interferometric probe); a) the measured topography of the feature is aligned to the reference CAD model (triangle mesh); b) height differences are computed on each point of the height map, and rendered as local, volumetric errors. In the end, the topography feature can be represented as the aggregation of the CAD model and the error map (from [36]).

triangulated mesh for alignment purposes - see Chapter 3 for details) because alignment error will evidently affect the quantitative assessment of discrepancy. Alignment is usually performed by global least-squares minimisation of Euclidean distances between the CAD and measured surfaces in six-degrees of freedom. However, if not all the regions of the feature have the same amount of "error" with respect to the CAD model, and if error is not globally small with respect to feature size, global alignment will more likely produce a wrong alignment result for lack of proper alignment references [69]. Another interesting challenge can be traced back to the limiting nature of height map datasets, and ultimately back to the whole unidirectional nature of the surface measurement process (except for XCT data, see Chapter 3 for further details). This makes the CAD-compare method proposed here substantially limited, as neither undercuts nor very steep surfaces can be successfully captured by the characterisation process. Finally, sampling resolution of measurement also plays a fundamental role. CAD-compare approaches work in the inspection of standard-sized parts when they start from high-density point sets or meshes. As density decreases the meaningfulness of the comparison rapidly degrades. In areal surface measurement, in particular if optical or XCT technologies are implied (see Chapter 3 for details), sampling is relatively dense for surface features of suitable sizes, but decays rapidly for very small features (compared to sampling resolution).

5.3 Datum-based methods

CAD-compare methods should be seen as a simplified version of a general procedure for the inspection of and verification of standard-sized parts that is based on comparing measured surfaces (shape and location) to ideal geometric entities acting as references. The conceptual approach is very well known in the disciplines dedicated to the inspection and verification of standard-sized parts, and its basic terminology and concepts can be found in the standard ISO 17450-1 ([2]). The application of the same concepts to the inspection of highly deterministic topographic features fabricated at micrometric and sub-micrometric scales has been recently investigated by the author [30]. Instead of having the entire nominal (CAD) model fitted to the measured dataset, like in the more basic CAD-compare methods, the idea is that both the nominal model and the measured dataset can be decomposed to isolate individual surfaces. The nominal model gives origin to a series of ideal surfaces referred to as *datums*, each meant to act as reference to assess localisation and shape discrepancy with a measured region obtained from the corresponding decomposition of the measured surface. Individual datums may represent the nominal expectations for a region of geometry, or they may appear in pairs to represent the ideal boundaries constraining maximum negative and positive allowed variation. Each measured region (obtained from partitioning the measured dataset) would therefore be fit to the associated datum or datums, in order to assess whether discrepancies would fall within acceptable ranges or not. Similarly to the CAD-compare methods, the main differences with the macro-scale version of the method are related to the limited selection possible for the choice of datums, as anything occluded or in undercut with respect to the measurement direction could not be inspected anyway[3]. The other main differences can be discussed with the help of Figure 5.3 (further details can be found in work by the author [5] and [31]).

The micro-pocket shown in Figure 5.3 has been manufactured via pulsed-laser machining. The nominal specification is given in the form of a cylindrical pocket with flat bottom located onto a flat plane. Given the poor control exerted by this specific laser machining operation, both side and bottom surfaces of the manufactured pocket are so irregular that it is difficult to identify the edges that are supposed to allow the dataset to be split into relevant faces. Also, the shape of the bottom and sides of the pocket is so irregular that it is difficult to see how they could be fitted to a cylinder and and a plane respectively, using conventional means such as least-squares minimisation. Furthermore, because of the unique probing direction (the z axis is supposed orthogonal to the surface the pocket has been machined onto), the side walls of the pocket are far less de-

[3] under the assumption of $z = f(x)$ or $z = f(x,y)$ geometry

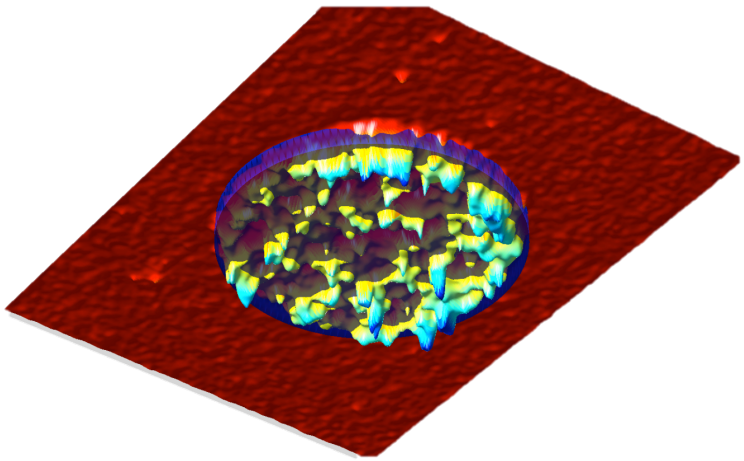


Figure 5.3: Datum-based characterisation. The topography feature (micro-pocket, about $100\mu m$ diameter, measured via focus variation) is shown overlaid to a cylindrical datum and two planar datums, one at the bottom of the pocket, the other aligned to the pocket surroundings (from [5]).

.

fined because of the scarcity of measured points, and large influence of measurement error on surfaces close to the maximum measurable slope (the pocket has been measured with a focus variation optical instrument). With respect to the CAD-compare example shown in the previous section thus, a larger amount of randomness affects the topography dataset, both because of the laser machining operation, and because of measurement error. CAD-compare itself would not be particularly meaningful given the large differences between the actual pocket and the nominal one, but an acceptable compromise can be achieved by adopting a datum based method, where the cylindrical datum needed to investigate the pocket side walls, and the planar datums needed to investigate the pocket bottom and surroundings external to its outer rim, are obtained as separate and independent entities, by separate fitting operations (as opposed to global fitting to a unique CAD model, like in the CAD-compare method). The fitting operations are different from each other, each one being specifically dedicated to the portion of measured dataset it is applied on. With reference to Figure 5.4, an initial partitioning operation is executed to identify the region in the x, y plane occupied by the pocket (see the next Chapter 6 and Chapter 9 for details on how the region is identified).The outer region is then fitted to a plane using least-squares minimisation of distances computed exclusively in the z direction[4]. The fitted plane is used to align the entire topography to a $z = 0$ reference. A material ratio curve (see Section 2.2.1) is computed only within the region containing the pocket. Using the geometric construction for the *Sk* family of

[4]this is essentially the application of an F-operator based on subtraction of the least-squares mean plane

parameters (see Section 2.2.1 and [5]) a reference depth z_{datum} is defined corresponding to the height value where the transition between the core and valley zones of the material ratio curve takes place. The z_{datum} reference height is used to define the second datum, representative of the pocket bottom, and parallel to the previous one aligned with the support surface. Finally, the *void volume* of the pocket is computed using the two reference datums, and a cylinder datum is created whose axis is at the x,y centroid of the pocket region, and whose radius R_{datum} is as such that the cylinder volume is equivalent to the void volume.

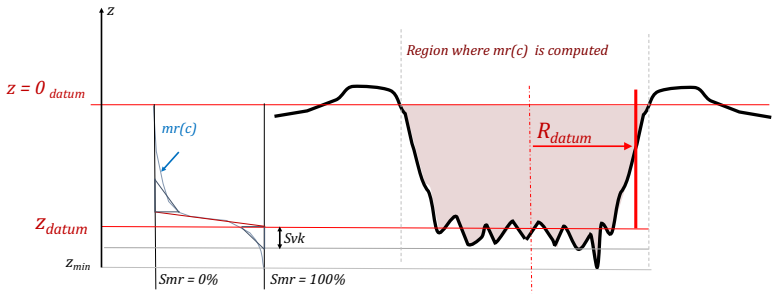


Figure 5.4: Construction of the datums for the characterisation of the laser-machined pocket previously shown in Figure 5.3. The construction is from original work by the author [5].

The pocket illustrated in the Figures 5.3 and 5.4 is finally characterised by the x,y *location* of the cylinder axis, by a *depth* defined as the z distance between the two planar datums, and by a *radius* R corresponding to the cylinder radius.

In summary, with respect to CAD-compare methods illustrated in Section 5.2, datum-based methods allows for splitting the measured dataset into regions that can then be separately fitted to datums using dedicated approaches developed to suit the specific requirements of each region. Once the datums are obtained, they can be used to define the feature in terms of the properties of the datums themselves. Of course two feature instances characterised via datum-based methods are comparable only if the same datum construction procedures have been adopted. As an example of this, in Figure 5.5 three different datum construction options are given to compute the thickness of a single-track artefact manufactured to study an ink-jetting process [70] (the application is described in more detail in Section 9.1). The feature (single track) is separated from its surroundings via a dedicated segmentation method (see [70], Chapter 6 and Section 9.1). The surroundings are fitted to a datum plane used as base reference for thickness computation. Thickness is then obtained as the distance between the bottom datum and a parallel, top datum obtained by means of three different construction procedures (see Figure 5.5 for details and again Section 9.1).

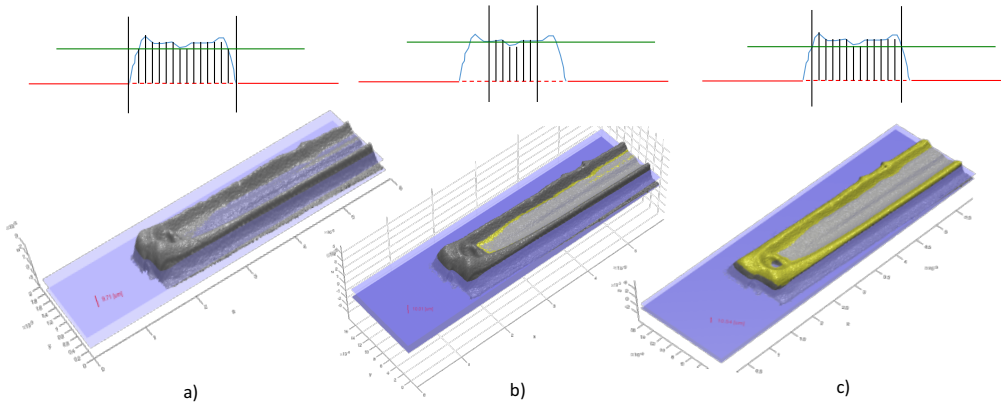


Figure 5.5: The thickness of the ink-jetted single track in the figure depends on the datums selected to define it; a) top datum obtained from the mean of the height values of the entire track; b) top datum obtained as the mean of the height values in the plateau region on the top surface of the track (highlithed in yellow); c) top datum obtained as the mean of the heights of both the plateau region and the surrounding crests (test case courtesy of J. Vaithilingam and the Additive Manufacturing and 3D Printing Research Group, University of Nottingham).Further details in [70] and Section 9.1.

5.3.1 Slicing and 2D datums

Planar or otherwise elementary datums cannot successfully follow the complexities of many topography shapes. If it does not make much sense to fit to higher-order polynomial surfaces or other parametric types of freeforms (e.g. spline surfaces), then an alternative option that can be pursued consists of slicing the feature with multiple planes (e.g. parallel to the main Cartesian axes). Examples of this approach are shown in Figure 5.6. With slicing, each slice results in a two-dimensional cross-section of the feature, that at least in principle could be easier to subject to datum-based characterisation, this time using an entirely two-dimensional approach consisting of 2D datums such as lines, circles, etc. Again, the datums of each cross-section can be used as the starting point to define the feature cross-section through simple measures of size and distance between the datum elements.

The slicing approach works also well if the aim is to describe how a shape varies along a specific direction. Such direction can be identified as orthogonal to the slicing planes, and plane spacing can be used to control sampling of shape properties along the chosen direction.

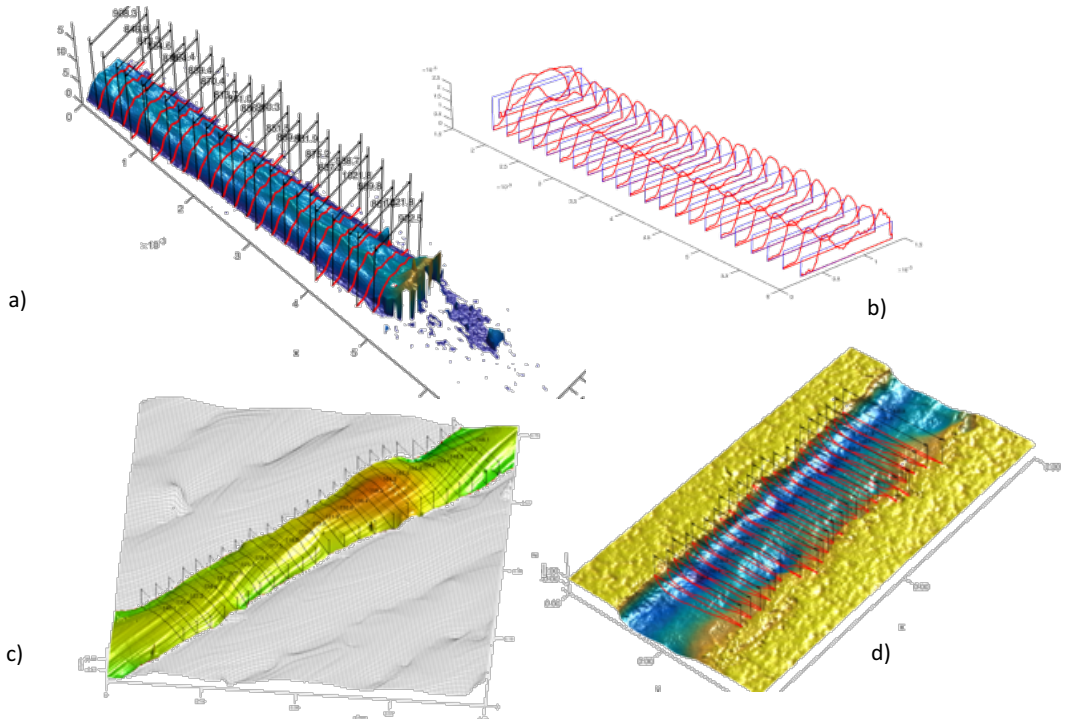


Figure 5.6: Example characterisations based on slicing and use of two-dimensional datum systems; a) width assessment for ink-jetted single track (from [70]); b) schematics of the cross-sections of another ink-jetted track with indication of the main two-dimensional datum lines (in blue) fitted to each cross-section; c) single weld track from laser powder-bed fusion of metals (from [37]; d) width assessment of groove obtained by electrolyte jet machining (dataset, courtesy of Jonathon Mitchell-Smith and Adam Clare, University of Nottingham).

5.3.2 Datum axes

Before the actual characterisation takes place, it is often convenient to refer the feature geometry to a local co-ordinate system. Cartesian configurations are typically preferred, with the x_f, y_f feature axes aligned to the directions of maximum and minumum elongation of the feature in the x, y plane (so that they can be used to encode orientation), and the origin corresponding to the feature centroid (so that it can be used to encode localisation within a global co-ordinate system). Since feature characterisation in this work is exclusively considered for feature topographies definable as $z(x, y)$ functions, the feature z_f axis datum is always considered as parallel to the global z axis of the entire topography (usually corresponding to the measurement direction), so that the x, y feature axes lay parallel to the x, y surface plane. Features with no visible or relevant, prevalent orientation, may also be encoded in a local, cylindrical coordinate system, defined by the same z axis and centroid used for Cartesian configurations, but then relying on an angle, radius and height to describe feature points. In the latter case,

the intrinsically Cartesian nature of the initial topography dataset (see Chapter 3 for details) implies the need for resampling the feature. Whilst examples for both Cartesian and cylindrical datum systems will be illustrated in the following sections, it is now important to discuss how the feature centroid (axes origin) and axes orientations can be determined in feature-based characterisation. It is worth pointing out right away that the choice of a reference datum system for feature characterisation may have very significant consequences on feature encoding, and thus in the assessment and comparison of feature instances. Two otherwise identical topography features may appear completely different at characterisation just because of a different positioning of origin and axes orientation.

Computation of feature centroid and axes via image moments

By relying on the formal equivalence between height maps and intensity images (see Chapter 3) it is possible to use the theory of *image moments* [71] to define a feature centroid and axes. In order to achieve this, all the $z(x,y)$ points (equivalent to pixels) external to the feature boundaries shall be set to zero (to avoid them influencing the moments). Then, a generic image moment of orders p,q can be defined as:

$$m_{p,q} = \sum_{i=1}^{n}\sum_{j=1}^{m} x_i^p y_j^q z\left(x_i, y_j\right). \tag{5.3.1}$$

The centroid of the feature is defined as:

$$x_c = \frac{m_{1,0}}{m_{0,0}}, \qquad y_c = \frac{m_{0,1}}{m_{0,0}} \ . \tag{5.3.2}$$

The feature centroid can be used for localisation, and is thus set as the origin of the datum coordinate system, Knowledge of the centroid also allows for defining the central moments. The general formulation of a central moment of orders p,q is:

$$\mu_{p,q} = \sum_{i=1}^{n}\sum_{j=1}^{m} \left(x_i - x_c\right)^p \left(y_j - y_c\right)^q z\left(x_i, y_j\right). \tag{5.3.3}$$

From the central moments, a tensor of moments J for the feature can be defined:

$$J = \begin{bmatrix} \mu_{2,0} & -\mu_{1,1} \\ -\mu_{1,1} & \mu_{0,2} \end{bmatrix}. \tag{5.3.4}$$

The eigenvectors of J correspond to the principal axes of an ellipse fitted to the feature x,y footprint (see Figure 5.7). The orientation of the major axis, expressed by an angle

with respect to the x axis of the entire topography, can be computed as:

$$\theta = \frac{1}{2}\arctan\left[\frac{2\mu_{1,1}}{\mu_{2,0} - \mu_{0,2}}\right]. \tag{5.3.5}$$

The angle θ can be used to define the feature datum axis x_f, whilst the y_f axis is orthogonal to it (see Figure 5.7).

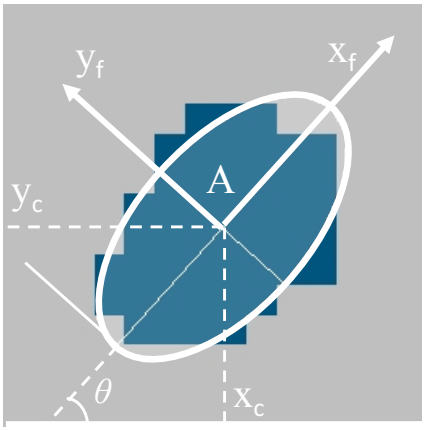


Figure 5.7: Geometric interpretation of intensity image moments applied to a topography feature.

The same equations can be applied to a binarised version of the feature, i.e. by assigning the value of one to all the pixels belonging to the feature, and zero to all the pixels falling outside. Moments on binarised features will be entirely dependent on the feature footprint on the x,y plane, with no "weighting" effect given by the different height values $z(x_i,y_j)$ of the feature itself. For example, the area of a binarised feature (in pixel coordinates) can be obtained as $m_{0,0}$, because :

$$m_{0,0} = \sum_{i=1}^{n}\sum_{j=1}^{m} z\left(x_i,y_j\right). \tag{5.3.6}$$

The actual area in topographic units can be obtained by multiplying the area (in pixel units) by pixel width/spacing dx,dy. The length of the ellipse semi-axes is can be obtained from the eigenvalues λ_1 and λ_2 of the tensor of moments J:

$$\lambda_{1,2} = \sqrt{\frac{1}{2}\left(\mu_{2,0} + \mu_{0,2}\right) \pm \sqrt{4\mu_{1,1}^2 - \left(\mu_{2,0} - \mu_{0,2}\right)^2}}. \tag{5.3.7}$$

Also, the eccentricity ϵ of the equivalent ellipse can be used as shape descriptor, since it is an indicator of feature elongation ratio on the x,y plane. The eccentricity is defined as:

$$\varepsilon = \sqrt{1 - (\lambda_2/\lambda_1)}. \tag{5.3.8}$$

Finally, higher-order moments can also be used to extract additional shape-related properties, turning the theory of image moments into a viable alternative for feature shape encoding.

In Figure 5.8 two example uses of image moments to characterise topography features are given. In Figure 5.8.a a multi-modal topography obtained by laser texturing has been segmented to isolate "pockets" characterised by increased determinism relative to their surroundings. Details on the segmentation process will be given in Section 6.4.1. As the only interest for this application is the characterisation of the x,y properties of the pockets and local heights within the pockets are roughly uniform, binary image moments are used to identify centroids, orientations, areas and x,y aspect ratios. In Figure 5.8.b image moments are used to characterise clusters of spatter formations identified on a laser powder bed fusion surface [37] (details of the identification process will be provided in Section 6.4.7, whilst the test case is discussed in Section 9.1.1). In this case, intensity image moments are used to keep into account the different height values comprised within the features.

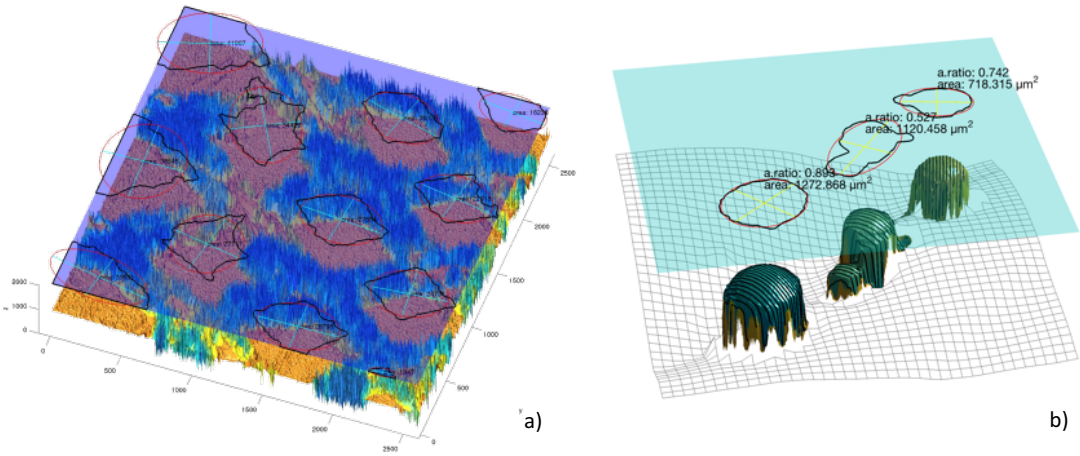


Figure 5.8: Image moments applied to feature characterisation; a) binary image moments are used to describe the x,y properties of pockets obtained by laser texturing (dataset courtesy of L. Blunt, University of Huddersfield); b) intensity image moments (i.e. with height-based pixel weighting) are used to characterise clusters of spatter formations in laser powder bed fusion (from original work by the author [37]).

Computation of feature centroid and axes via principal component analysis

A datum centroid and reference axes for topography features can also be obtained by turning the height map into a point cloud. In the conversion, only height map points belonging to the feature are kept, whilst the remaining ones are masked away as not belonging to the feature. On the obtained point cloud, a statistical method known as *principal component analysis* (PCA) can then be performed.

In statistics, given the realisation of a set of n possibly correlated random variables, PCA is a method that allows to generate a new set of n random variables, called *principal components*, that are linear combinations of the original variables, and are uncorrelated instead. In addition, the first principal component has the largest variance, accounting for most of the variance of the entire dataset, while the other components are characterised by decreasing variance. Translated to a Cartesian, three-dimensional space, PCA can be applied to point clouds, and the result can be interpreted geometrically: a point cloud of m points can be seen as the realisation of three random variables: X (associated to the observation of m values of x co-ordinate), Y and Z, defined analogously. The PCA method identifies three new random variables, the principal components, that identify new co-ordinates laying onto a new set of Cartesian axes, that are potentially rotated with respect to the originals (because the principal components are linear combinations of the original variables) and orthogonal to each other, because by definition they are uncorrelated). Like in statistics, the first principal component accounts for most of the variance of the original dataset, which, in geometric terms, means that the first principal axis is aligned to the direction of maximum spread (maximum elongation) of the point cloud, whilst the third and last one is aligned to the direction of minimum spread.

Principal components can therefore be used as datum axes to describe a topography feature orientation in space. As origin of the new set of axes, the centroid of the point cloud is typically chosen, i.e.:

$$x_c = \frac{1}{m}\sum_{i=1}^{m} x_i, \quad y_c = \frac{1}{m}\sum_{i=1}^{m} y_i, \quad z_c = \frac{1}{m}\sum_{i=1}^{m} z_i \ . \tag{5.3.9}$$

The original point cloud co-ordinates are then re-written relative to the centroid position. Then the 3×3 covariance matrix is built:

$$C = \begin{bmatrix} k(x,x) & k(x,y) & k(x,z) \\ & k(y,y) & k(y,z) \\ & & k(z,z) \end{bmatrix} . \tag{5.3.10}$$

Similarly to image moments, the eigenvectors of the covariance matrix represent the directions of the new axes (the principal components), whilst the eigenvalues allow for ranking the components from the largest (where most of the variance is found) to the smallest. The principal components are therefore elected as the new datum axes for the feature, a typical convention being that the x_f direction is associated to the largest component between the two closest to the global x,y surface plane.

Some considerations must be made to differentiate the PCA method from the method of image moments to determine centroid and datum axes for a topography feature:

- Image moments produce x_f and y_f datum axes always aligned to the x,y surface plane, and the z_f datum axis is always by definition aligned to the global z direction. On the contrary, PCA operates with full six degrees-of-freedom on the point cloud, and may not necessarily result in such an alignment;
- because of limitations intrinsic to the 2.5D nature of height maps and topography datasets in general, there is generally no interest in allowing full three-dimensional freedom in axes placement, whilst having the z_f axis aligned to the z global axis is usually favoured (leading to the x_f and y_f axes being parallel to the global x,y plane);
- to avoid PCA finding a solution that is not aligned with the global x,y plane, the freedom of the method is forcefully limited by projecting the point cloud on the x,y plane (essentially, by temporarily setting all the z coordinates to the same value). Once this is achieved, the PCA is reduced to operating in conditions that are equivalent to image moments operating on binary images, because in both cases the "weights" resulting from feature points at different heights are eliminated.

As an example, consider the PCA axes obtained on the full-3D feature shown in Figure 5.9.a (spatter formation in triangle mesh obtained from XCT measurement). These axes are freely oriented in space, depending on the shape characteristics associated to the topography feature. On the contrary, the PCA axes shown in Figure 5.9.b are computed on a 2.5D height map converted into a point cloud, with the points projected onto the x,y plane to force PCA into an aligned result.

Since this text focuses primarily on $2.5D$ topographies (i.e. definable as $z = f(x,y)$ functions), there is no real advantage to using PCA instead of image moments, the latter actually providing the possibility to take into account height-based weighing, which on the contrary cannot be provided when projecting the point cloud onto the x,y plane in order to achieve 2.5D behaviour in PCA.

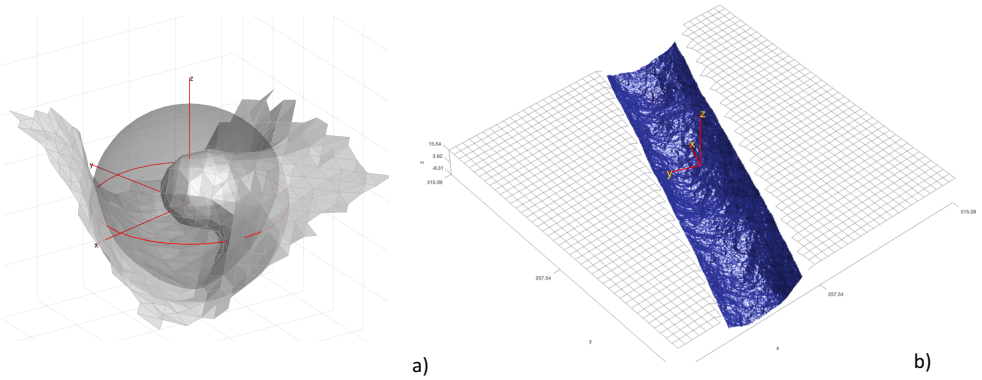


Figure 5.9: Application of principal component analysis (PCA) to identify reference axes for a topography feature; a) full-3D feature (spatter formation) represented as a triangle mesh: reference axes computed via full-3D PCA; 2.5D feature on height map, and 2D-constrained PCA using point projection onto the x,y plane.

5.3.3 Similarity metrics for datum-based descriptions

If two feature instances have been associated to the same structured set of datums, then the comparison of the features is straightforward, and reduced to comparing scalar values (e.g. the distance between two planar datums to determine a size such as width or depth). As anticipated earlier, the evident weak point is the sensitivity to datum misplacement. Datum placement errors may happen for any surface datum, but also for point and axis datums such as those defining the origin and the x_f,y_f and z_f axes controlling feature orientation. Evidently even a slight difference between placements of such datums will lead to different encoding, and thus different comparison results.

5.4 Shape resampling and integration methods

The methods collectively referred to as *shape resampling and integration methods* are all based on the concept that the feature topography (height map) can be resampled, typically in polar, cylindrical or spherical co-ordinates, in the surroundings of the feature centroid. The result of such resampling can then be aggregated in different ways, typically by some form of integration over radiuses, circles, spheres, etc., ultimately obtaining a series of scalar values which form a shape descriptor that can be used for assessment and comparison with other features. Resampling is typically designed as a form of control on the range of scales covered by the descriptor. The integration step is usually designed to reduce a complex shape to a simpler and shorter series of scalar values. In some cases, integration is also performed as a means to achieve some form of

invariance for the final descriptor, most typically, invariance to angular orientation on the x, y plane, so that two otherwise identical features, expect for their orientation, will result into the same shape descriptor values. Finally, integration is also a method for reducing noise via averaging effects, and thus makes these methods suitable for features whose shape is predominantly deterministic, and the descriptor is expected to achieve robustness to noise. Several shape resampling and integration methods are illustrated in the next sections, all derived from transforms that have been originally developed by the image and video processing research communities and here redesigned to operate on topography data. The transforms that will be discussed are:

- spin images;
- the ring projection transform (RPT);
- the angular radial transform (ART);
- multi-scale descriptors; and
- the scale-invariant feature transform (SIFT).

5.4.1 Spin images

The shape descriptors known as *spin images* were introduced originally in [73] to describe the shape of fully three-dimensional geometric models defined by surface meshes [73, 72]. Their first use as descriptors of local surface topography features is found in work by the author of this thesis [26]. In layman terms, a spin image describes the 360° panorama viewed by an hypothetical observer located at any given point of a topography. Two features (topography regions) will be classified as similar if similar panoramas are observed from their respective centre points. The geometric construction procedure followed to build a spin image descriptor for topography is illustrated with the help of Figure 5.10. The topography is assumed as available as a triangle mesh (see Chapter 3). Once a centre point v_0 has been chosen, information about the neighbouring locations is collected and used to compute the local normal direction n_0 and the associated reference plane π (usually via PCA, see Section 5.3.2 for details). Differently from the originally proposed spin image model, in the version dedicated to topography [26] the surface surrounding v_0 is resampled in order to gain more control on the window of investigated scales. By assuming the hypothetical observer as located on v_0 (Figure 5.10.a) and being n_0 the "up" direction (Figures 5.10.b,c), the panorama image is built by scanning the topography all around the observer and encoding each point in spherical co-ordinates (Figures 5.10.d,e). The spherical space is usually binned

so that its pixels can show accumulation of similar values (Figure 5.10.f). The final panorama result, referred to as "spin image" can be compared to analogous images using typical similarity metrics for images, such as normalised cross-correlation. The cross-correlation function is particularly suitable because it makes the comparison independent on the initial orientation of the observer in v_0.

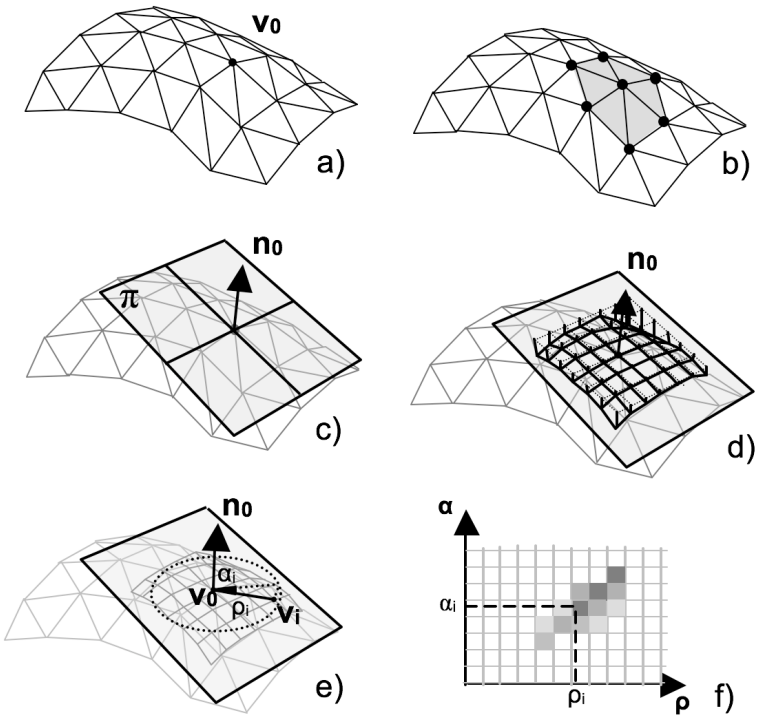


Figure 5.10: Construction of a spin image for a given point in a surface topography (from [26]); a) the point v_0 in the mesh is selected; b) surrounding points up to a predefined radial distance from v_0 are selected as well; c) the local normal n_0 is computed using the surrounding points. The normal identifies the locally tangent plane π; d) a square reference grid is created on the π plane and projected to the local surface to resample it; e) the surface resampled points are encoded in spherical co-ordinates through the radius ρ and the angle from π plane α; f) ρ, α pairs are binned into a two-dimensional grid, equivalent to an intensity image. The final image "pixel" intensities are representative of the frequency of ρ, α pairs falling within each pixel.

In Figure 5.11 the example computation of spin images on the surroundings of, and inside a smooth, elongated dale is shown [26]. Though the example surface is fairly smooth (thus making the spin images visually similar to each other), different patterns corresponding to locally convex and locally concave regions are visually perceivable.

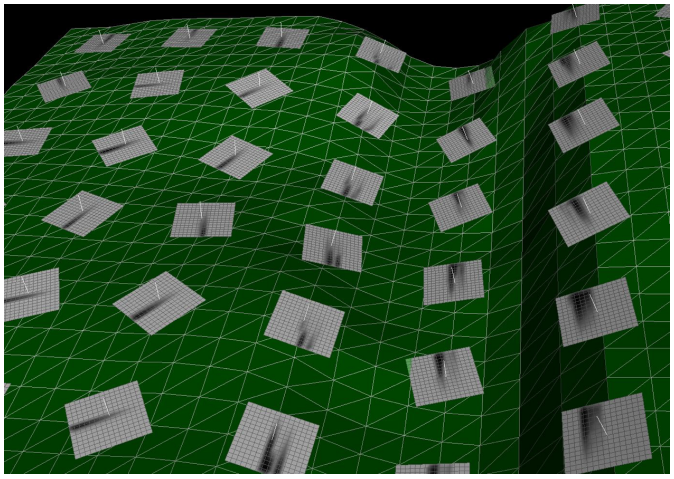


Figure 5.11: Example computation of spin images on the surroundings and inside a smooth, elongated dale. Each spin image is shown along with its local surface normal, using darker colour for higher intensity pixels (from [26]).

5.4.2 The ring projection transform - RPT

The ring projection transform (RPT) was originally introduced in [74] for the analysis of digital images and was later converted to operate on topography data (in the form of height maps) by the author of this thesis [32, 33]. Starting from the feature centre point, the RPT considers local topography in cylindrical co-ordinates, and integrates over 2π all the surface heights sampled at a given radius as shown in Equation 5.4.1, thus resulting in a function or radius r:

$$RPT(r) = \frac{1}{2\pi r} \int_0^{2\pi} z(r,\theta)\, d\theta\,. \tag{5.4.1}$$

In the formulation reported in this text (from [33]), each angular integral is normalised by the circle perimeter, to avoid portions of the features further away from the centre to excessively influence the descriptor. The radius r in equation 5.4.1 is typically discretized over a range of values extending from zero to r_{max} so that the RPT function is turned into a finite and discrete series of scalar values. For ease of interpretation, such values can be linearly interpolated and shown as plots. An example involving the computation of the RPT over a region featuring an isolated protrusion (grit from super-abrasive surface [32, 33]) is shown in Figure 5.12.

In terms of similarity metrics, computing the area between two different RPT curves as a measure of cumulative difference was found as the best alternative (see [33] for details, and Chapter 6 for examples).

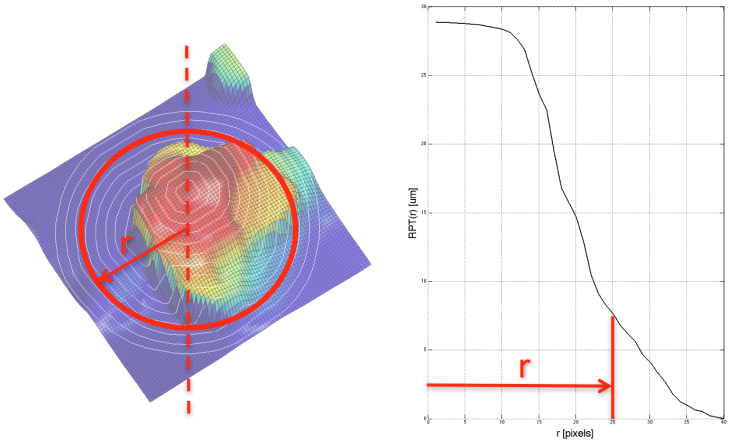


Figure 5.12: Example computation of the RPT over an isolated protrusion; left) topography region containing the isolated protrusion (grit on a super-abrasive surface). The integration circles used by the RPT descriptor are shown overlaid. Right) the curve resulting from computing the RPT descriptor. The RPT result at the example radial distance r is shown highlighted.

5.4.3 The angular radial transform - ART

The angular radial transform (ART) is a moment-based description method originally developed for digital images and video frames, and adopted by the MPEG-7 standard [20–23]. An original version of the ART redesigned to operate on surface topography data (height maps) is presented here [36]. The ART is defined on a unit disk and based on complex orthogonal sinusoidal basis functions in polar coordinates. For surface topography data, the ART coefficients $F_{n,m}$ of order n and m are given by:

$$F_{n,m} = \int_0^{2\pi} \int_0^1 V_{n,m}(\rho,\theta)\, z(\rho,\theta)\, \rho d\rho d\theta \tag{5.4.2}$$

where the basis functions $V_{n,m}$ are defined as:

$$\begin{aligned} V_{n,m}(\rho,\theta) &= A_m(\theta)\, R_n(\rho) \\ A_m(\theta) &= \tfrac{1}{2\pi} e^{jm\theta} \\ R_n(\rho) &= \begin{cases} 1 & \text{if} n = 0 \\ 2\cos(n\pi\rho) & \text{if} n \neq 0 \end{cases} \end{aligned} \tag{5.4.3}$$

An example series of basis functions for $n = 0,\ldots,2$ and $m = 0,\ldots,11$ is shown in Figure 5.13 (from [36]). Note that n controls the number of radial "oscillations", whilst m controls the number of angular oscillations.

A shape descriptor coefficient $S_{n,m}$ can be defined from normalised ART coefficients as

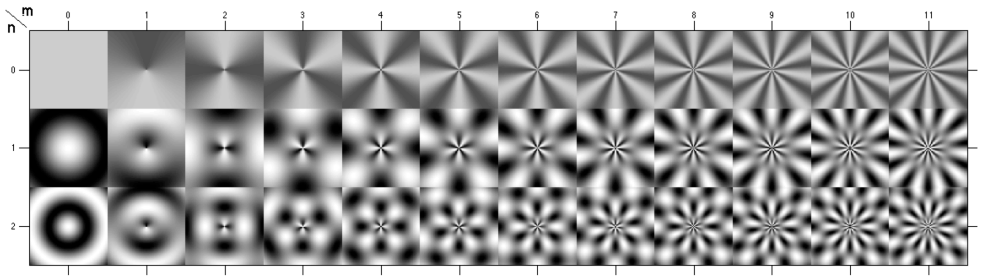


Figure 5.13: ART basis functions for $n = 0,\ldots,2$ and $m = 0,\ldots,11$ (from [36]).

follows:

$$S_{n,m} = \frac{|F_{n,m}|}{|F_{0,0}|}. \tag{5.4.4}$$

Normalisation ensures that $S_{n,m}$ is independent from the size of the window it has been computed from. Finally, the topography feature descriptor is defined as the $(n+1) \times (m+1)$ matrix of coefficients from $S_{0,0}$ to $S_{n,m}$, each coefficient bearing a scalar value. In Figure 5.14 an example application is shown where the ART-based descriptor is used to encode the shape of a tile feature, part of a micro-embossing pattern fabricated by laser micro-machining [36] (same example shown in Figure 5.2). The $(n+1) \times (m+1)$ matrix of coefficients $S_{0,0},\ldots,S_{n,m}$ is represented as a three-dimensional bar diagram.

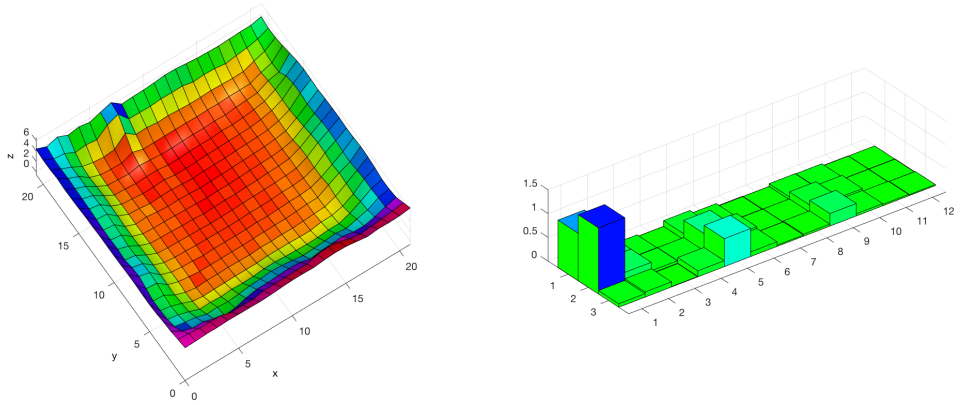


Figure 5.14: Example encoding via the ART-based shape descriptor; left) original feature (tile of micro-embossing pattern); right) 3×12 matrix of shape descriptor components represented as a three-dimensional bar diagram (from [36]).

5.4.4 Multi-scale descriptors

The idea shared by this familty of descriptors is that the feature can be described by multiple versions of the same height-map, each presenting a "view" of the feature at a different scale of observation. The final descriptor itself may be several times larger

than the initial dataset. Thus, in a sense, the descriptor does not do good job at summarising the shape of the feature. On the contrary, these descriptors are designed to highlight shape-related properties under multiple perspectives, thus effectively multiplying the amount of available information. The methods differ at what information is highlighted, but all share the concept that such information pertains shape content at multiple scales. The following methods are illustrated in the next sections:

- Pyramid decompositions
- Gabor filter banks
- Wavelet decompositions

All such methods were originally developed in the field of signal or image processing. However, they were recently redesigned to operate on height maps, exploiting the equivalent formulation of height maps and intensity images, discussed in Chapter 3. Wavelet decompositions for surface topography were first explored in [25] and then further explored by the author [35]. Gabor filter banks and Pyramid decompositions for surface topography where first introduced by the author [35]. In the domain of image processing, the multi-scale descriptors introduced here are frequently referred to as "texture descriptors". In image processing, "texture" is a term used to define any visible pattern in an image[5]. Clearly, even from an intuitive standpoint, the term "texture" in images refers to patterns that due to their irregular nature, are hardly comparable to elementary shapes such as circles, squares, etc. It follows that multi-scale descriptors like those described here are more suitable to describe topographies with a significant degree of randomness.

Pyramid decompositions

The concept of pyramid decomposition was first introduced in digital images [75, 76] as a means to represent image content at different scales of observation. Instead of actually acquiring multiple images at different magnification, the idea is to take only one image, at a given resolution, and then increasingly downsample it (typically by a factor of two, in a process called *dyadic decomposition*) while also processing the pixels in some way, in order to replicate observation at a different scale. Pyramid decompositions of surface topographies expressed through height maps were first introduced by the author [35]. The first layer of a *Gaussian pyramid* can be obtained by processing the

[5]different from the definition of texture adopted by surface metrology and illustrated in Chapter 2.

original height map through a Gaussian smoothing operation with a convolution kernel of given width and σ. The smoothed map is then downsampled by a factor of two (dyadic decomposition) to obtain the second layer. The second layer is convolution-smoothed and downsampled again, to obtain the third layer, and so on. A Gaussian pyramid decomposition applied to an example height map (human mole) is shown in Figure 5.15.a (from original work by the author [35]). A *Laplacian pyramid* is similarly obtained by processing the original height map through a Laplacian operator (to obtain the local curvature map) before downsampling it to the next level, and so on. However, because the Laplacian operation is computationally expensive, an approximation of the Laplacian pyramid is commonly accepted, consisting of obtaining each Laplacian layer as the subtraction of two consecutive Gaussian layers (resampled to the same size). Such a pyramid is referred to as *difference-of-Gaussians (doG)*. Notice that the last level of a doG pyramid is the level of a Gaussian pyramid because there is no further level to perform the subtraction operation. Used as shape descriptors, Gaussian pyramids are capable to extract the more deterministic components from a random pattern, as they appear at different scales. Invariance across scales can be used to investigate shape properties, as explored in the next chapter (6). On the contrary, Laplacian pyramids highlight high-curvature regions, and allow to investigate how curvature changes across scales.

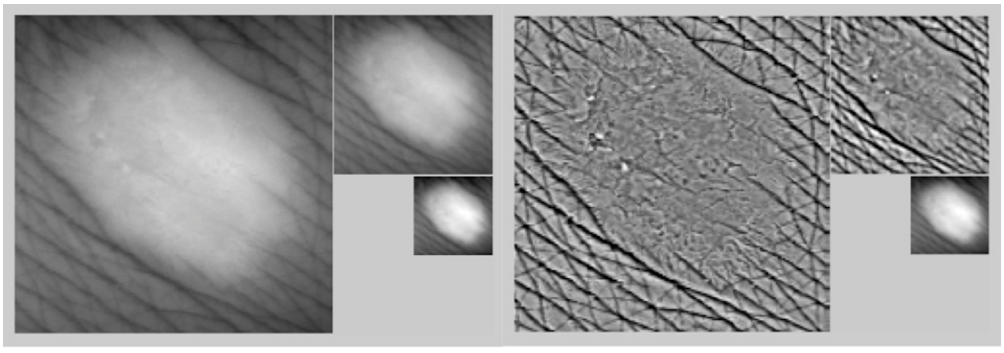


Figure 5.15: Example three-level pyramid decompositions for a human mole (height map); a) Gaussian pyramid; b) Laplacian pyramid (approximated by difference-of-Gaussians (from [35]).

Gabor filter bank decompositions

A Gabor filter bank is a series of Gabor kernel functions, which can be separately convolved with the a digital image in order to highlight textural patterns with specific orientations and wavelengths [77–79]. The first investigation on using Gabor filter banks to process areal topography data was proposed in a work by the author et al. [35].

Each Gabor kernel function is obtained by modulating an oriented planar wave with a Gaussian envelope. The result of convolving each kernel with the original areal dataset is a response map where texture components of specific orientation and wavelengths, "resonating" with the kernel, will be highlighted. A bank of n kernels will result in an equivalent number of response maps, A typical choice of filter bank covers at least six directions (between 0 and 180°) and a varying number of wavelengths depending on application. Some example kernels and resulting response maps are shown in Figure 5.16 for the same mole areal dataset previously shown in the pyramid decomposition example (Figure 5.15).

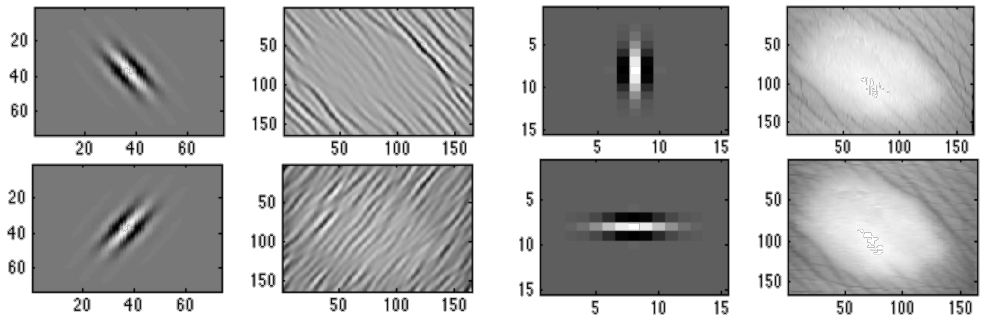


Figure 5.16: Gabor kernels and corresponding response maps for the mole test case previously shown in Figure 5.15 (from [35]).

Wavelet-based decompositions

In wavelet-based decomposition, an image is convolved with variants of the same mother wavelet function, where a mother wavelet is simply a modulated wave-like function[6]. The variants are obtained by differently "stretching" and orienting the mother wavelet, so that it resonates with different elements of an image texture. Wavelet decompositions for areal surface topography were first explored in [25], with particular attention to the lifting wavelet scheme . A range of different mother wavelets and decompositions for areal topography data, including the dual-tree complex wavelet transform [81], were later investigated by the author of this thesis [35]. An example decomposition based on a Daubechies 3 mother wavelet is shown in Figure 5.17.

5.4.5 The Scale Invariant Feature Transform - SIFT

The Scale Invariant Feature Transform (SIFT) is a popular method developed for image processing [82], that allows one to collate multiple digital images of the same scene into larger panoramas. The original SIFT method is comprised of the following steps (for a

[6] multiple mother wavelets are available [80]

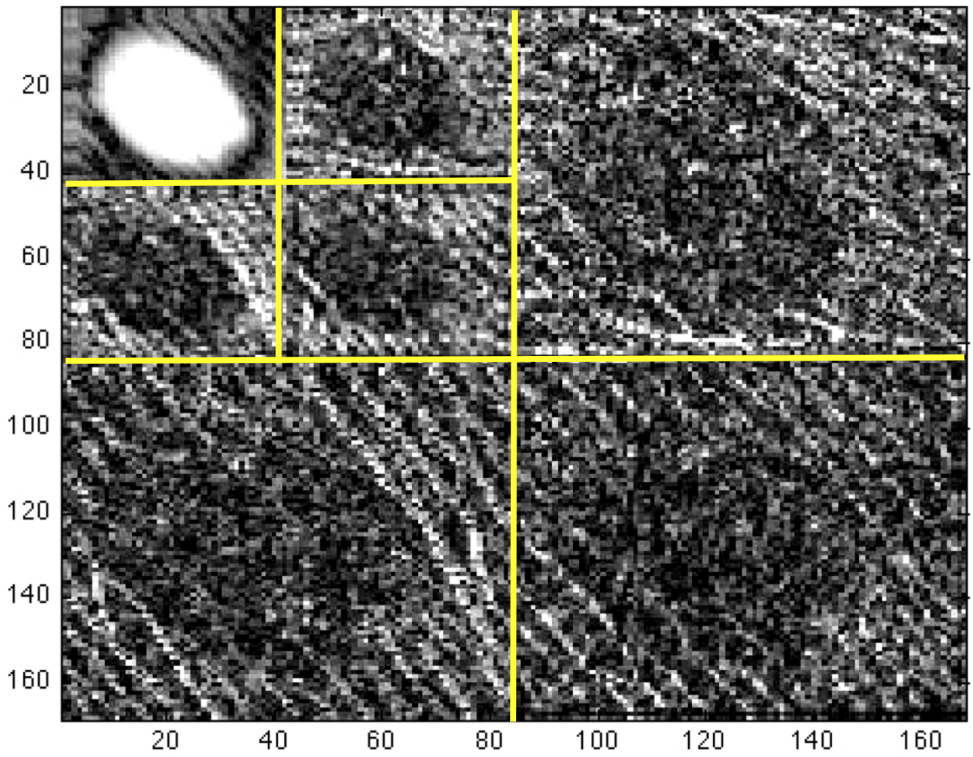


Figure 5.17: Example wavelet decomposition of the mole dataset. Mother wavelet: one-dimensional Daubechies 3 (db3) applied along horizontally (x), vertically (y) and diagonally (x,y). The three larger quadrants (bottom row and rightmost upper quadrant) contain the 1st level horizontal, vertical and diagonal detail, the upper left quadrant contains the results of the 2nd level decomposition (four smaller quadrants) applied to the 1st level approximation (not shown). These are the 2nd level approximation (upper left), and the horizontal, vertical and diagonal details (the remaining three small quadrants). From [35]

pair of images to be collated): a) the identification of keypoints, i.e. points identifying recognisable landmarks in each image, b) the identification of matches between pairs of keypoints, representative of correspondences across two images; c) the identification of a transform that aligns the matched keypoints, thus co-localising the images in the same reference frame[7]. The SIFT method has been redesigned during the doctorate to operate on topography data, given the formal equivalence between digital images and height maps, illustrated in Chapter 3. In the specific context of feature topography characterisation, the subject of this chapter, some steps of the SIFT method can be redesigned to generate descriptors for topography features. The interesting steps of the SIFT method, i.e. the identification of keypoints and the construction of local texture

[7]In the original SIFT collating transform [82], homorphic transformations are allowed, i.e. transformations comprised of scale changes and deformations, to compensate for distorsions due to the different viewpoints the images have been taken at

descriptors surrounding the keypoints, are illustrated in the following sections.

Construction of the pyramids

Pyramids are needed to identify the landmarks acting as keypoints for the SIFT method. Firstly, Gaussian pyramids are constructed starting from the original height map. The standard procedure for obtaining a Gaussian pyramid was initially introduced in Section 5.4.4, where a Gaussian pyramid was constructed by increasingly smoothing and downsampling the original topography dataset. For the SIFT method, the procedure is more complicated as multiple pyramids are constructed instead, to cover all the combinations of downsamplings and Gaussian smoothing σ values. The updated procedure is described in the following and with the help of Figure 5.18. The original height map is first downsampled multiple times by following a dyadic series. Then for each octave (referred to as *scale* in the SIFT method [82]), a separate pyramid is built by smoothing the height map by Gaussian convolution filtering with increasingly larger σ values. Assuming $z(x,y)$ as the height map obtained by downsampling at each scale (octave), the levels of the corresponding Gaussian pyramid can be obtained by convolving $z(x,y)$ by a Gaussian kernel $G(x,y,\sigma_L)$ where σ_L increases according to a predefined progression. Thus the L^{th} level of the Gaussian pyramid at a given scale (octave) is defined as:

$$L(x,y,\sigma_L) = G(x,y,\sigma_L) * z(x,y)\,. \tag{5.4.5}$$

On each Gaussian pyramid, the difference-of-Gaussians (doG) method can applied (identically to what shown in Section 5.4.4) to obtain the corresponding Laplacian pyramid. Thus, any level of the Laplacian pyramid at a given scale (octave), referred to as $D(x,y,\sigma_L)$ is defined as:

$$D(x,y,\sigma_L) = L(x,y,\sigma_L) - L(x,y,\sigma_{L-1}) = [G(x,y,\sigma_L) - G(x,y,\sigma_{L-1})] * z(x,y)\,. \tag{5.4.6}$$

Identification of tentative keypoints

Tentative keypoints are defined as high-curvature points, thus in principle they could be spotted on a Laplacian map. However, curvature is scale-dependent, so high-curvature points may exist at different scales. Thus, all the pyramids built at different scales (octaves) are investigated, and a first set of tentative keypoints is found by identifying local maxima in each single doG layer $D(x,y,\sigma_L)$. Then, only those which are also local

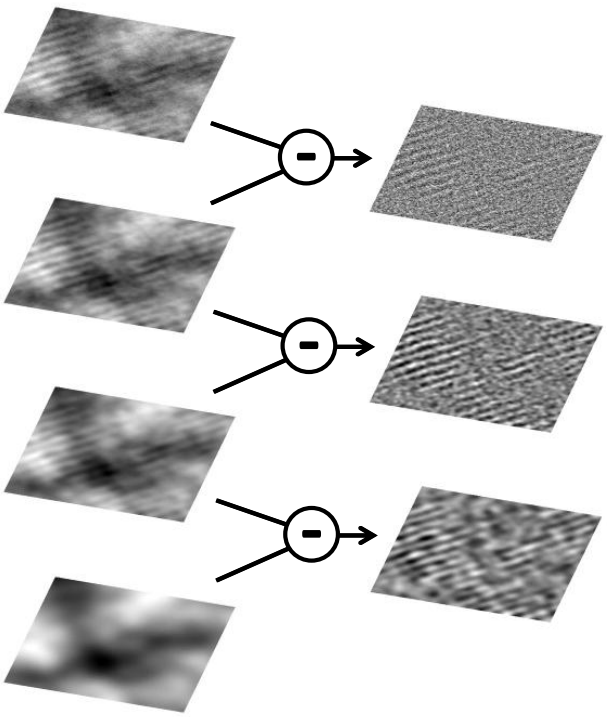


Figure 5.18: Computation of SIFT pyramids on example topography dataset; left) four levels of a Gaussian pyramid corresponding to a single scale (octave); right) Laplacian approximation levels obtained by doG.

maxima if the same location is considered across three consecutive layers of the same doG pyramid (i.e. $D(x,y,\sigma_{L-1})$,$D(x,y,\sigma_L)$ and $D(x,y,\sigma_{L+1})$ are kept (see Figure 5.19).

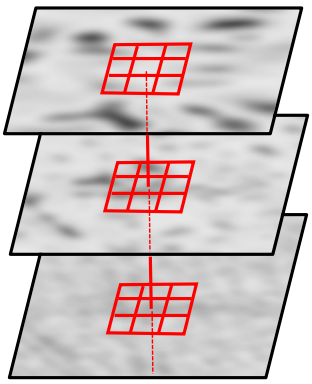


Figure 5.19: Identification of a SIFT keypoint; zoomed-in view of the regions on three consecutive layers of the doG pyramid, where the local maximum is found across three consecutive maximum curvature points.

The first set of tentative keypoints is then subjected to a series of post-processing operations:

– *The keypoint positions are refined to sub-pixel accuracy* by locally fitting the region surrounding each keypoint (in the $D(x,y,\sigma_L)$ dataset where the keypoint is located) by Taylor series expansion, and then by finding the new local maximum.

– *Low contrast keypoints are eliminated* by thresholding the derivatives of the curvature computed across scales.

– *Keypoints corresponding to edges are eliminated* by checking curvature across the two principal directions: an edge point will have larger curvature across the edge, and almost zero along the edge.

.

Computation of the descriptor around each keypoint

The SIFT descriptor is needed as a means to compute similarity between regions surrounding different keypoints. The descriptor is computed as illustrated in Figure 5.20. Firstly, the local gradient vector at the x_i,y_j position where the keypoint is located[8] is computed on the $L(x,y,\sigma_L)$ map of the Gaussian pyramid. The gradient vector magnitude m and orientation θ are defined by means of finite differences, as:

$$\begin{aligned} m\left(x_i,y_j\right) &= \sqrt{\left[L\left(x_{i+1},y_j\right)-L\left(x_{i-1},y_j\right)\right]^2+\left[L\left(x_i,y_{j+1}\right)-L\left(x_i,y_{j-1}\right)\right]^2} \\ \theta\left(x_i,y_j\right) &= \tan^{-1}\left(\frac{\left[L\left(x_i,y_{j+1}\right)-L\left(x_i,y_{j-1}\right)\right]}{\left[L\left(x_{i+1},y_j\right)-L\left(x_{i-1},y_j\right)\right]}\right). \end{aligned} \tag{5.4.7}$$

The result is the identification of a gradient vector on the keypoint, as shown in Figure 5.20.a. The surroundings of the keypoint in the $L(x,y,\sigma_L)$ map are then resampled over a grid of specified width and resolution, oriented as the gradient vector (see again Figure 5.20.a). Within each cell of the resampling grid, local gradient vectors are computed (magnitude and angle) as shown in Figure 5.20.b. The magnitudes are then weighted by a Gaussian weighing function covering the entire resampling grid, so that gradient vectors located far from the keypoint are attenuated (see again 5.20.b). Finally, a three-dimensional binning of the gradient vectors is performed as shown in Figure 5.20.c. Binning is on the x and y co-ordinates of the gradient vector in the resampling grid (a 4×4 configuration is shown in Figure 5.20.c), and in its orientation value (8 bins is the default, as shown in Figure 5.20.c). The final three-dimensional histogram is built with the average intensities (magnitudes) of the binned gradient vectors (in Figure 5.20.c,

[8]The position x_i,y_j is indicated for simplicity. In reality, the keypoint has been located with sub-pixel accuracy during one of the previous steps (see the previous section)

this leads to a total of $4 \times 4 \times 8$ histogram values). The histogram values are stored as the final SIFT descriptor value, together with the local orientation of the resampling grid computed at the keypoint[9].

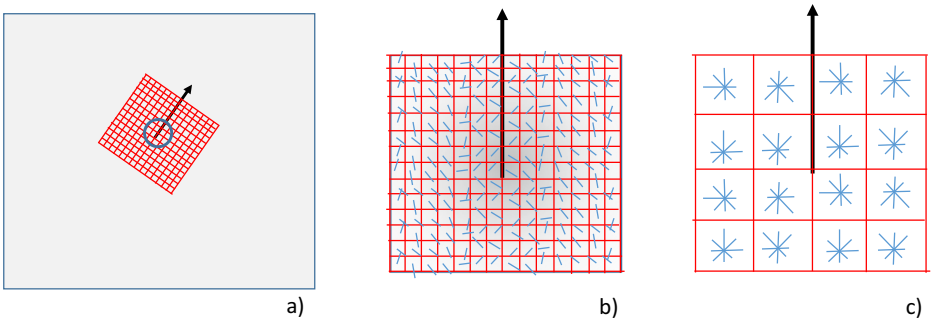


Figure 5.20: Construction of the SIFT descriptor; a) local orientation vector computed on the keypoint in $L(x,y,\sigma_L)$ and associated resampling grid; b) local gradient vectors (magnitude and orientation) computed within the cells of the resampling grid; c) binning to obtain the 3D histogram of orientations. In the example, 4 horizontal $\times$ 4 vertical $\times$ 8 angular bins for the gradient vector are used. In each bin the magnitudes are averaged amongst the values falling within that bin.

Using the SIFT descriptor to characterise local topography features

A SIFT descriptor in its native formulation (i.e. as described in the previous sections) is very useful to identify correspondences amongst pairs of datasets, for alignment and stitching purposes. However, prior to this doctorate, it has never been used as shape descriptor of an arbitrary topographic feature. This is because of two main reasons, the first being that the descriptor is computed on a resampling window whose size and resolution are dependent on the σ_L of the pyramid level $L(x,y,\sigma_L)$ where the keypoint is located, meaning that there is poor control on making the resampling window the appropriate size to cover the entire feature of interest. Second, the location of the descriptor itself is governed by where the keypoint is located, again out of the user's direct control, being dependent on local maxima computed on the doG pyramid levels. However, one could compute the SIFT descriptor at arbitrary locations x,y and using arbitrary σ values in order to control the width of the resampling window. The SIFT descriptor has been found compatible with some of the mechanisms underlying human perception (i.e. the capability of reacting to gradient patterns at multiple scales - see [82]); however, its thorough investigation as a general purpose shape descriptor is yet to be performed.

[9]The actual SIFT descriptor usually undergoes further normalisation and post-processing steps [82], only the most important steps have been reported here.

5.5 Spatial occupancy enumeration and cell decomposition methods

Spatial occupancy enumeration methods describe the shape of a geometric entity by subdividing the Cartesian, three-dimensional space into adjacent "blocks" or "cells", and then by counting how many and which cells are occupied by the shape. Spatial occupancy enumeration methods were first introduced in the research field of solid modelling [65] and haven't been applied to the characterisation of areal topography data yet [10]. Spatial occupancy enumeration methods for encoding feature shapes in areal topography data have been briefly discussed in previous work by the author [31]. The main drawback of the method is that any discrete spatial subdivision in cells will suffer from sampling error (if not completely empty or completely occupied). Increasing the subdivision resolution can reduce the issue, but at a price of increased memory occupancy and complexity in data processing: an example involving the encoding of a step-like feature is shown in Figure 5.21. In this case, spatial subdivision is only performed on the x, y plane and cells are filled with a scalar representative of average surface height within the cell.

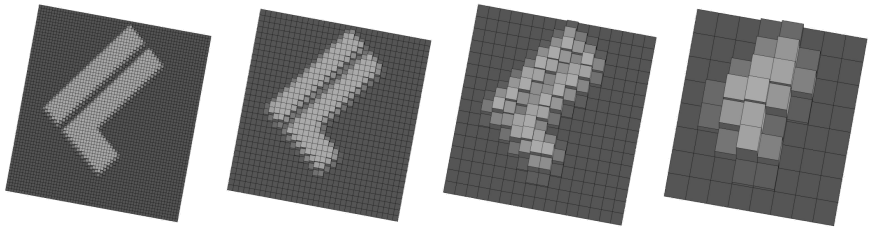


Figure 5.21: Spatial occupancy enumeration method for the encoding of step-like features. As the resolution of cell subdivision decreases, the model has a smaller memory footprint and it is easier to process; however, resolution is increasingly lost.

Spatial occupancy enumeration methods require the cells to be adjacent (i.e. topologically connected and not intersecting). However, spatial occupancy enumeration methods can be seen as a special case of the more general *cell decomposition* methods where such assumption is removed. More promising cell decomposition methods in terms of encoding efficiency seem to be those making use of recursive subdivision approaches (i.e. cells inside other cells) such as those based on *quadtrees* (for two-dimensional de-

[10] In computed tomography, the surface of an object is initially encoded within the voxels of a raw volumetric dataset. This is technically a spatial occupancy enumeration (because the surface is defined by a series of voxels which are either "empty" or "occupied" by material), however so far it has not been used to describe the shape of topographic features yet, as volumetric data is only used as a starting point to extract a triangle mesh (see again Chapter 3).

compositions in the x, y space) and *octrees* (for fully three-dimensional decompositions). An example octree subdivision of a test areal topography dataset investigated by the author during the doctorate is shown in Figure 5.22.

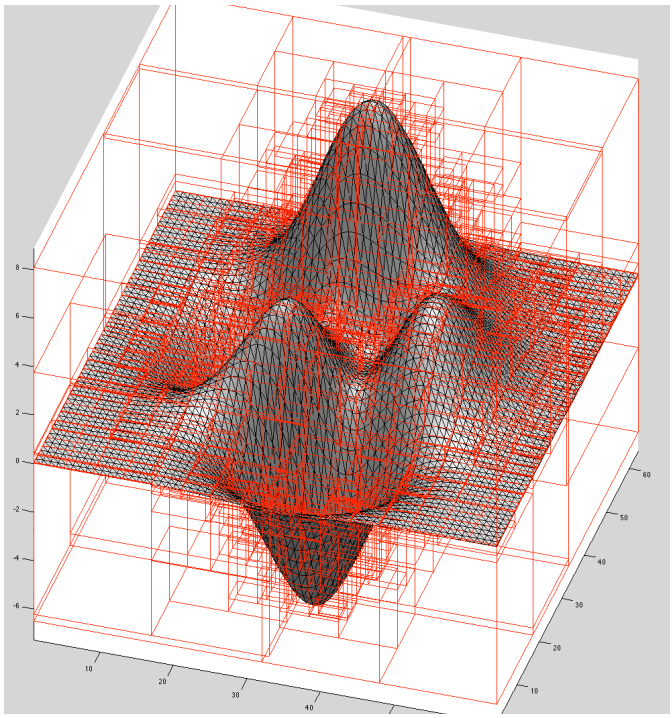


Figure 5.22: Octree subdivision of example topography.

Both quadtrees and octrees can be represented as directed graphs (i.e. *trees*, hence their names) with repercussions on similarity metrics, i.e. on how shapes can be compared. The comparison of two features encoded through a quadtree or octree starts from the top node in the tree, and recursively proceeds downwards, towards smaller subdivisions, while checking for correspondences in patterns of empty or occupied nodes. As the comparison between trees proceeds downwards, increasingly smaller-scale shape details are involved, making this a multi-scale comparison method capable of recognising both shapes that are only similar at the large-scale feature level, and shapes that are similar down to the smallest detail. The big weakness of these methods is the strong dependency on the reference frame for the spatial subdivision, as even the smallest shift of boundary surfaces in the subdivision may make a large cell either empty or occupied in the tree. For octrees and quadtrees, explored subdivision strategies consider the feature centroid to define the origin, and make use of either the PCA axes (see Section 5.3.2) or the x, y, z axes of the entire dataset, to implement the subdivision.

5.6 Characterisation of feature patterns

This chapter has been dedicated so far to the illustration of approaches for the characterisation of individual features. However, some of the most successful classes of surfaces whose topography has been engineered (i.e. highly-deterministic topographies) are based on the presence of a large number of features arranged according to some type of regular pattern. These surfaces are commonly known as *tessellated surfaces*; popular examples include retro-reflector surfaces, low-friction surfaces and aerodynamic surfaces [18, 83–86]. Tessellated surfaces achieve their functional role through the concerted action of many features acting at once. In such context, the characterisation of the individual feature is usually not relevant without the characterisation of the properties of the entire pattern. Regularity of the entire pattern is typically sought for by design. That is, local properties of the pattern must be replicated with minimal variation throughout the surface. From a nominal standpoint, the pattern itself is usually defined by a *unit lattice* replicated multiple times over the surface. The characterisation of a pattern of features is comprised of the characterisation of its units (i.e. individual features), for which the solutions discussed in the previous sections of this chapter apply, and of the characterisation of the layout, which is essentially an analysis of the layout properties, and how much they deviate from the nominal lattice. This latter aspect is discussed here.

Not much is available in the surface metrology literature concerning the assessment of regularity in tessellated surfaces. The most notable approaches consists in segmenting the areal correlation function (ACF) in search for evidence of loss of regularity (these methods have been discussed in Section 2.5). Alternatively, as also discussed in Section 2.5, the individual features can be identified, their pattern reconstructed and encoded into a graph, and regularity assessed through the investigation of graph properties. This latter method was applied in [21]. Original variants developed by the author of this thesis [5], are discussed in the following sections.

5.6.1 Reconstruction of feature patterns

A pattern of surface features can be thought of as a graph-like geometric entity, made of *connected nodes*. Nodes are typically assigned to centroids of the features, or to notable keypoints between features. The concept is illustrated in Figure 5.23. The positions and relative distances between nodes determine the geometric content of the pattern, whilst connectivity determines the topological content. Characterisation of a pattern,

and consequently the assessment of its conformance to design specifications, should be based on both a geometric and topological evaluation.

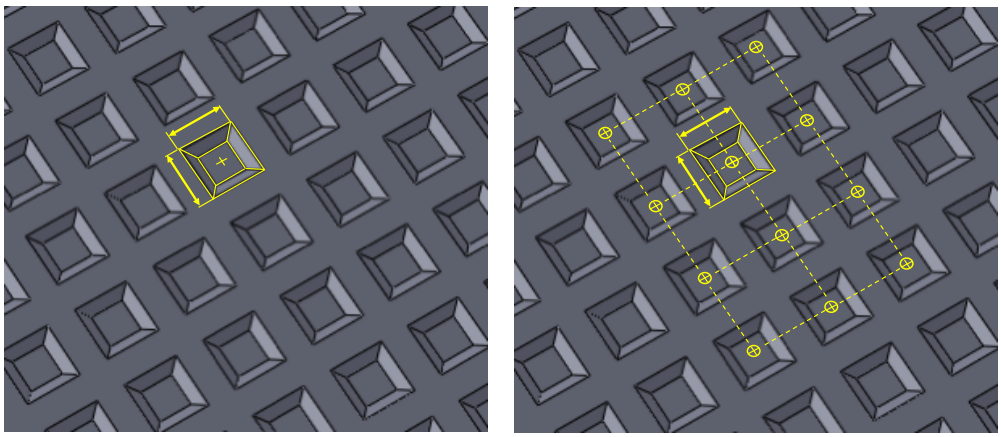


Figure 5.23: A regular pattern of topography feature instances can be seen as a graph connecting feature centroids or other notable keypoints; a) individual feature with its defining dimensions; b) graph formed by connecting the feature centroids.

Patterns can be *explicit* or *implicit*. An explicit pattern is included in the design of the tessellated surface, and thus available to the characterisation process in some nominal form. Pattern characterisation can be based on CAD-compare methods equivalent to those previously illustrated to characterise the shape of topography features (see Section 5.2), except that this time they are applied to compare the shape and connectivity of graph-like data structures. On the contrary, an implicit pattern has not been provided with the specification of the tessellated surface, or more commonly it has been provided in an incomplete form. In such circumstances, a new lattice can be constructed and overlaid to the feature instances, creating connectivity for example by using Delaunay triangulation. An example of this latter scenario is illustrated in Figure 5.24, where a pattern of micro-pockets designed for friction reduction is reconstructed starting from the feature centroids (from previous work by the author of this thesis [5]).

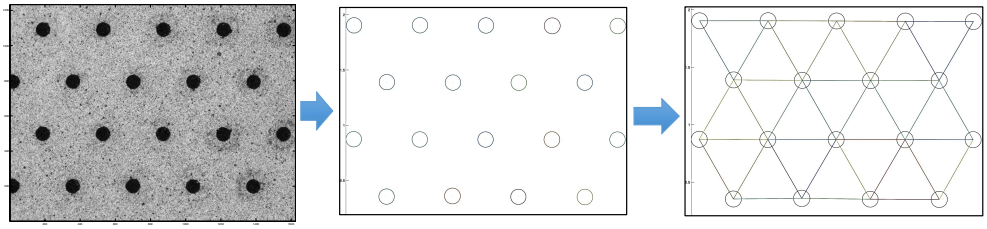


Figure 5.24: Pattern reconstruction from a measured dataset; a) array of micro-pockets for a low-friction surface; b) identified feature instances (feature boundaries shown); c) lattice reconstructed from Delaunay triangulation of feature centroids (from [5]).

As mentioned earlier, lattice (and hence pattern) reconstruction may not necessarily op-

erate on the centroids, but rather focus on keypoints located between the features, for example by identifying points demarcating the transition from the region occupied by a feature to its immediate neighbour. This alternative route is illustrated in Figure 5.25 where the Delaunay reconstruction is shown side by side with the Voronoi reconstruction from the same set of feature instances of Figure 5.24. The Delaunay and Voronoi graphs are complementary, and can be generated from each other, so that the transition is seamless. Centroid-based representations, such as what obtained by means of the Delaunay reconstruction, are more useful when regularity of properties related to a feature centroid are important. For example, the mean distance between neighbouring features, or the alignment of feature instances along some reference direction, etc. On the contrary, keypoint-based lattices and patterns, such as what obtained from Voronoi reconstruction, are more useful when the regularity must be assessed on properties such as the uniformity of the areas occupied by a feature instance and its immediate surroundings.

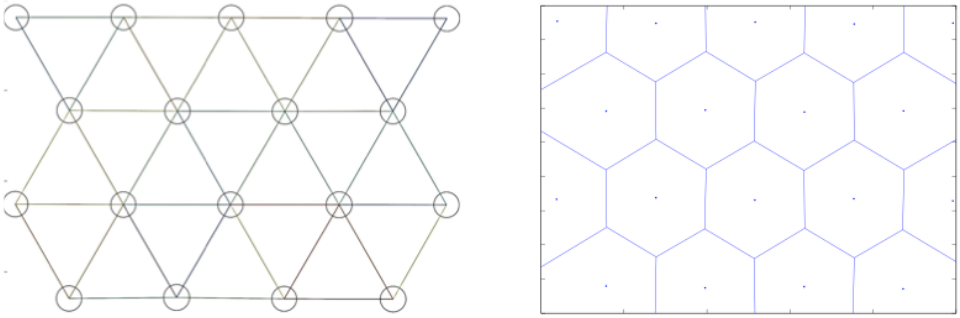


Figure 5.25: Centroid-based and keypoint-based lattice and pattern reconstruction of the same set of feature instances; a) Delaunay reconstruction (centroid-based); b) Voronoi reconstruction (keypoint-based).

5.6.2 Characterisation of feature patterns

Once the pattern (i.e. the graph) is available, it can be characterised through the investigation of its topological and geometric attributes. Usually, the most imporant property of any attribute of a regular pattern is *uniformity*. Uniformity can be defined as the constancy of local pattern properties, as an observational window is moved across the pattern. Special-purpose visual representations can be generated that help obtain a qualitative assessment of uniformity of any pattern attribute. For example, the images in Figure 5.26 show that the nodes in the graph reconstructed from the circular pockets belonging to the same surface used for Figures 5.24 and 5.25 can be coloured proportionally to the diameter of the pockets (Figure 5.26.b), or proportionally to the

length of the segments connecting feature centroids (Figure 5.26.c) thus providing a quick overview of uniformity, and of the existence of any spatial trends. This in turn can be used to identify defects, as illustrated in the dedicated section 6.5 of Chapter 6 (dedicated to feature identification and extraction).

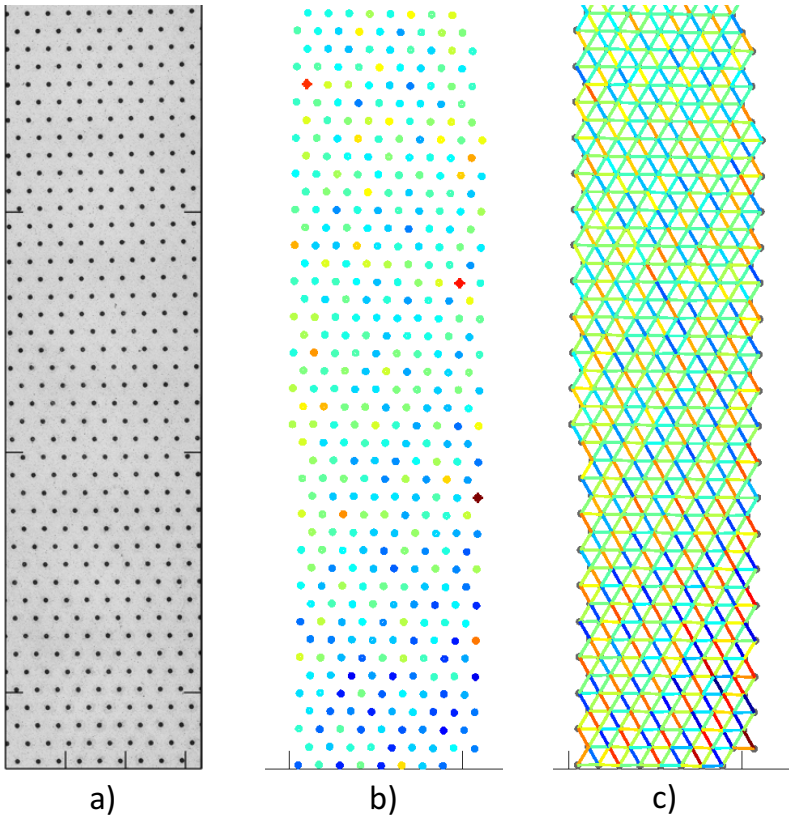


Figure 5.26: Characterisation of pattern uniformity via specialised visual representation; a) original strip of micro-pockets (test case described in [5]); b) pockets coloured proportionally to diameters; c) graph segments coloured proportionally to Euclidean distance between pocket centroids.

Once a mathematical representation of the graph representing the pattern is available, sample statistics and related population parameters can be estimated as indicators of central tendency and scatter for any interesting attribute of the pattern. Typically targeted attributes may be: spacing between neighbouring feature instances (mean and dispersion), footprint area occupied by each feature instance (again mean and dispersion), alignment of feature instances within individual, nominal rows (mean orientation and dispersion), etc. to cover almost endless series of potentially interesting attributes to investigate.

5.7 Summary and outlook

The development of characterisation solutions effective and efficient at encoding the properties of surface topography features is of critical importance as it lays at the foundations of many tasks involving identification, quantitative assessment and comparison of surface features. In this chapter, an original approach has been presented, where the characterisation of topographic features is seen as the problem of how to capture shape in mathematical terms when feature shape may be very complex, and a potentially associated to significant amounts of randomness. Randomness in particular may vary significantly depending on the type of feature, thus a wide range of alternative options have been proposed in the chapter for feature characterisation, ranging from those more suitable to capture strongly deterministic geometries, to those more oriented towards the use of statistics to handle large variability. Moreover, for any shape characterisation solution adopted, i.e. for any *shape descriptor*, suitable *similarity metrics* must be adopted as well (i.e. mathematical models describing the extent of similarity/difference between two shapes encoded using the same descriptor). In this chapter it has been illustrated that some shape variability may be captured within the shape descriptor, whilst other sources of variability may be handled by a robust similarity metric (i.e. to a certain degree, invariant to the "noise" filtered through the descriptor). Finally, for many surfaces, after the individual features have been characterised, it may be also important to characterise their networked layout. Tessellated surfaces have been discussed as one of the most typical examples of this type of challenge and methods have been presented to handle such cases.

Despite several solutions to shape characterisation have been proposed in this chapter that are capable of handling both the deterministic and the random part of a feature shape, a complete and comprehensive theory supporting the decomposition of shapes into their deterministic and random components suitable for application in surface metrology has not been fully developed yet. As none of the solutions discussed in this chapter may be considered consistently superior to the others across the entire range of application scenarios where feature characterisation may be applied, expert assessment is still needed in order to select the best line of action.This is definitely a challenge that will need addressing.

Chapter 6

Feature identification and extraction

6.1 Introduction

The concept of *topography feature* as a topographic element of interest within a surface was already introduced in Chapter 4. A preliminary, working definition of the concept of *feature identification* was provided in Section 4.5.2, where it was referred to as the act of recognising a region of topography as representative of a particular type of topographic formation being searched for. In this sense, feature identification aims to establish a direct mapping between regions of a surface and a class defining a type of topographic formations of interest. The mapping results in the region being recognised as containing an instance of the class being searched for. The class itself is assumed as containing a general definition of the feature shape and properties acting as reference, but a class could be as general as saying "scratch", meaning any scratch of any size, shape, etc. In some cases, a definition for the reference shape which should be associated to a class may not even be provided, as defining complex shape-related information content is challenging. In those cases where deemed possible, a class may also be defined within a larger set of organised knowledge surrounding a surface, a concept which could be conveniently (but not necessarily) captured through an *ontology* [66].

On the contrary, the term *feature extraction*, also preliminarily introduced in Section 4.5.2 consists of isolating the topography region representing the feature from its surroundings, and turning it as an independent topographic dataset. Feature extraction is particularly critical for those applications interested in obtaining an accurate measurement of a feature shape properties, in particular if such properties are heavily influenced by accurate feature boundary identification. On the contrary in "yes/no" scenarios of feature identification (for example if one is solely concerned with determining if the feature is present or not), feature extraction may not be needed, as later discussed in

Section 6.6.

6.2 An overview of feature identification methods

In this thesis two main families of feature identification approaches are presented and discussed, namely:

- template matching; and
- template-less, segmentation-based identification.

In *template matching*, candidate regions are compared to a reference topography that has been predefined to represent the feature being searched for. Identification is based on recognising a region as sufficiently similar to the template. In *template-less segmentation-based identification* there is no template to rely upon. Instead, properties computed locally within candidate regions are retrieved and compared across regions in order to achieve a segmentation of the topography dataset, that is, its partitioning into regions/segments. Once the regions are found, rule-based methods operating on segment attributes are used to identify those partitions that, alone or in combination, can be finally recognised as instances of the feature being searched for.

6.3 Template matching

The family of feature identification approaches here referred to as *template matching* revolves around the assumption that a *shape template* must be available that describes the shape of the feature being searched for. Identification is then about comparing the template with any candidate region selected within the measured topography dataset, in the search for good matches. Matching is in turn thought of as based on quantifying both the shape of the template and the shape of any candidate region through the same type of shape descriptor (that is, any characterisation solution amongst those discussed in Chapter 5) and by adopting an adequate similarity metric to compare shape descriptor results. The conceptual pipeline underlying feature identification by template matching is summarised in Figure 6.1.

It is important to point out that the template must be encoded in the same format used for the source topography. Thus, typically as a height map (see Chapter 3 for a discussion on topography data formats). This is a necessary precondition to being able to apply the same shape descriptor to both the template and any candidate region being

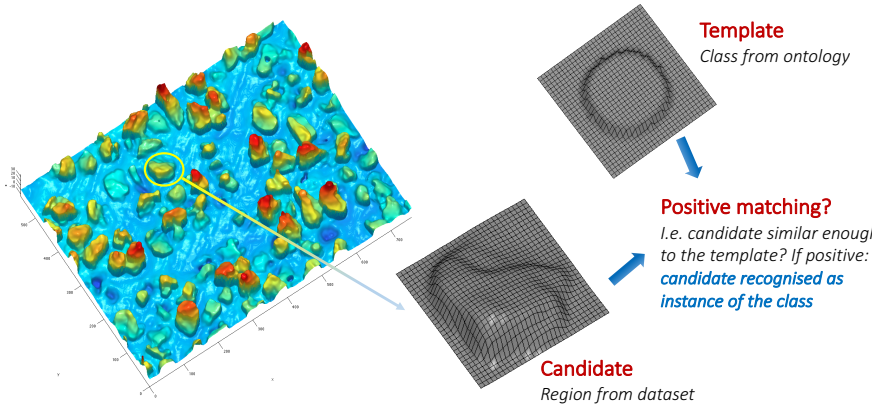


Figure 6.1: The conceptual pipeline of feature identification by template matching. Example identification of grits on super-abrasive surfaces, from previous work by the author of this thesis [32].

.

inspected for matching purposes. Thus, if the template is available as a triangulated mesh (e.g. in the STL format), it must be converted into a height map first, possibly with the same sampling resolution of the original topography to facilitate comparability of the results of the encoding process via shape descriptor.

6.3.1 Template matching by scanning

The topography being explored in the search for regions being good matches to the template is here referred to as the *source topography*. As template matching is based on identifying good matches between the template and any region of the source topography, a method for exploring the source topography thus producing candidate regions for comparison must be developed. The simplest solution consists of creating a moving window, the size and shape of the template, and then sliding it over the source topography, extracting its contents as a tentative candidate region, at any position of the window. When the window reaches the border of the dataset, either padding, mirroring or wraparound solutions can be implemented to take care of the missing data points, the choice being dependent on how the selected method for handling missing values affects the shape descriptor chosen for the identification problem[1]. The most typical scanning pattern for a moving window is *raster scanning* with *one-to-one sampling*, i.e. with steps of the moving window being one pixel-wide both in the x and y directions, to ensure complete coverage[2]. The shape of the moving window depends

[1]Padding, mirroring and wraparound reflect analogous techniques being used in image and signal processing to handle missing points in applications involving moving windows, like for example in convolution filtering

[2]the term *pixel* is here used to refer to a single position of the height map, as discussed in Chapter 3

primarily on the choice of descriptor. Descriptors operating in polar co-ordinates (see for example the RPT and ART descriptors discussed in Chapter 5) prefer circular windows. However for algorithmic convenience, usually moving windows are chosen of rectangular or square shape, and unused points are simply masked out when computing the descriptor.

In Figure 6.2 an example template matching approach is illustrated based on a moving window operating in raster scanning mode (from previous work by the author of this thesis[32, 33]). The template shape is a cylindrical protrusion, meant to represent a generic grit on top of a super-abrasive surface. The actual surface, measured by a focus variation instrument, is populated by an unknown amount of grits, each one with different shape an size. The "grit" feature is evidently characterised by a deterministic component (the fact that any grit is a protrusion). However, significant randomness is added, because grits will vary in shape and size. For this application, the template shape (a cylinder) is defined in parametric form (parameters being the height and diameter of the cylinder), and at each position of the moving window an optimisation process is performed to identify the parametric template instance that better matches the topographic content currently being found within the moving window. The ring projection transform (RPT), previously illustrated in Section 5.4.2 is chosen as shape descriptor, whilst as similarity metric, the area comprised between the RPT plots of the region and the parametric template is used.

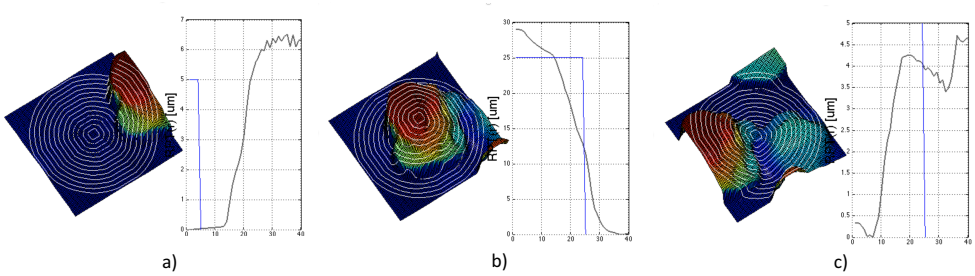


Figure 6.2: Template matching with raster scanning for the identification of grits in the super-abrasive surface (from [32, 33]); Top row, from left to right: the moving window scans the surface along a rectilinear portion of trajectory; Bottom row: corresponding profiles of the RPT shape descriptor (see Section 5.4.2) and comparison to the RPT computed on the parametric template.

6.3.2 Scan-free template matching

If the source topography allows it, i.e. knowledge is available about the source that can help directing the search for candidate regions towards specific locations, then candidates for template matching can be identified with the need for complete scanning.

For example, In Figure 6.3, from previous work by the author of this thesis [36] feature identification has been performed on micro-embossed surfaces (the test case has been already shown in Figure 5.2 and Figure 5.14). The template shape is encoded as a dale-like topography. Differently from the grit example case [33], there is no need for a parametric template as feature identification is dealing with a much more deterministic topography (the micro-embossing units are more similar to each other), and variations across pattern units can be handled by the robustness built-in in the shape descriptor and associated similarity metric. In [36], the ART shape descriptor (see Section 5.4.3) has been selected. Being the micro-embossing units arranged along the rows and columns of a regular pattern which is also significantly deterministic (see Figure 6.3.a), candidate windows can be found by simply applying a partitioning process that splits the source dataset into single dales (see Figure 6.3.b). In [36], ISO 25178-2 morphologic segmentation into dales [6] has been used. The resulting individual partitions (i.e. the coloured regions shown in Figure 6.3.b) are directly considered as candidate regions for template matching, with no need for raster scanning.

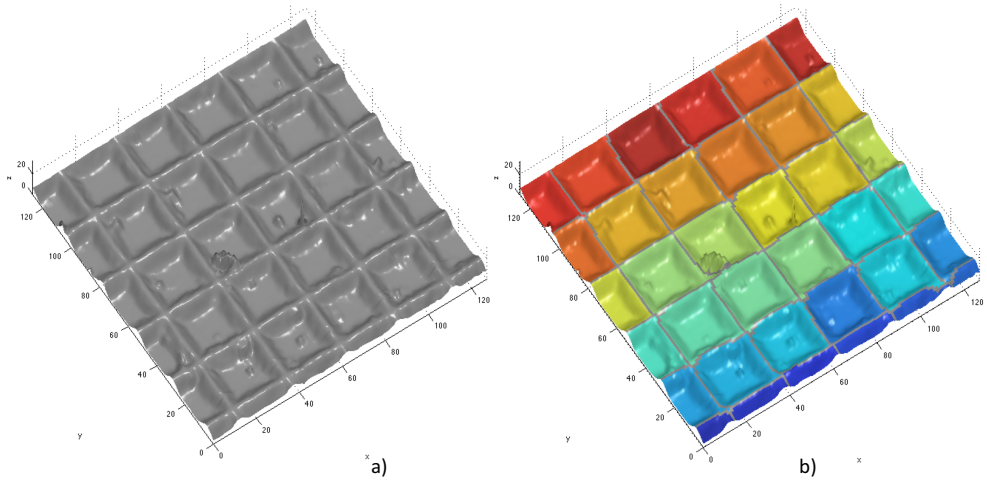


Figure 6.3: Scan-less template matching on micro-embossing surface; a) original topography dataset; b) result of ISO 25178-2 [6] morphologic segmentation into dales (each dale is rendered with a different colour).

Some descriptors may be defined as part of a more complex procedure which is comprised also of a step dedicated to identifying where the descriptor should be computed, thus also removing the need for raster scanning. In the previous chapter, in Section 5.4.5, the SIFT descriptor was introduced, which is computed only on *keypoints* recognised on a doG pyramid decomposition (see Section 5.4.5 for details).

6.3.3 Template matching, randomness and determinism

Template matching approaches to feature identification seem to work well when the target feature is characterised by a significant amount of determinism[3]. This is because it must be possible to define a reference shape acting as a template for the identification. Clearly, the intrinsic variations associated to the real instances of the feature, regardless of whether due to measurement error or to actual variability across instances on the surface, must be accounted for by the more random components in the feature model (again random intended as less predicable). The solution adopted so far to handle such variability is the use of variation-robust shape descriptors and associated similarity metrics, for example playing with the averaging effect of height integration (like exploited both by the RPT and the ART shape descriptors). In more extreme cases, like the grit identification example (see the previous sections and [33]), the additional use of parametric templates is suggested, that adapt to the current observation by changing their size and shape properties through control parameters in the attempt to maximise similarity according to the adopted metric.

6.4 Template-less, segmentation-based identification

This family of methods is characterised by not there being a template explicitly describing the reference shape of the feature being searched for. Instead, a two-steps approach is pursued:

- the topography is partitioned into regions, based on the analysis of some lower-level local topographic property; and
- feature instances are recognised by rule-based assessment of region properties.

This two-step approach mimics a possible interpretation of some of the mechanisms underlying visual perception [63]. Lower-level processing of the image reconstructed by our vision system (for example, processing brightness, contrast, gradients, etc.) results in the identification of uniform regions, transitions and in general, patterns, leading to an initial partitioning of the visual field. Further processing pertains the interpretation of such partitions through the analysis of their properties, and the final creation of links between some of such partitions and classes of interest, meaning that some regions have been "recognised" as instances of predefined classes. So for example, a blue

[3] determinism and randomness are intended as related to uncertainty of prediction at unprobed locations, as discussed in Section 4.2.

and coarsely uniform partition perceived within our visual range may be recognised as "the sky" because of its properties such as colour, extensions and localisation, whilst a smaller partition within the same field of view, but with different shape and colour properties may be recognised as "a person".

In a sense, it could be stated that this two-step approach is not really template-less, because the definition of recognition rules associated to the analysis of the created partitions is an implicit way to define another type of reference template. However, for the sake of simplicity in the discussion, in this thesis being template-less will specifically refer to the absence of an explicit definition of a *reference shape* for the feature being searched for, which does not necessarily mean that the feature may not be defined in other indirect ways (for example, through rules on its properties).

6.4.1 Introduction to segmentation

The first aspect that must be discussed about segmentation-based identification is indeed *segmentation*. Segmentation is the process of partitioning a surface topography into topologically connected *regions*, each occupying a portion of the measured dataset, and recognised as a partition because of some reason, such as uniformity of some properties, or because it is surrounded by strong transitions of some other properties, etc. The idea of segmentation was already introduced together with the fundamentals of the proposed framework (see Section 4.5.2), because the decomposition of surface topography into relevant features more or less explicitly implies a partitioning of the dataset, i.e. a segmentation process. As stated in Section 4.5.2, the term "segmentation" is not new to surface metrology, as in the standard ISO 25178-2, the decomposition of a topography into hills or dales is described as a segmentation process [6]. The term has actually been borrowed from the image processing literature [71], where it equivalently refers to partitioning an image into regions, also called *segments* because of the topological connectedness of the contained points. Here it is important to point out that, given the formal equivalence of digital images and topography maps (see Chapter 3), basically all the segmentation ideas, methods and algorithms originally developed for images can be mutuated, with proper adaptation, to operate on height maps, and thus can be used to generate partitions in surface topography datasets.

The necessary precondition to segmentation is the identification of topographic properties that can be computed locally and later used as local descriptors to compare topography points and determine whether they should belong to the same partition or not. Notably, all the shape descriptors discussed in Chapter 5 are potentially applicable. However, the preference goes towards the application of simple descriptors to

capture lower-level properties, such as local slope, curvature, or even local roughness, both because of computational issues (see next section) and because more complex descriptors are more suited to be used to encode complex shapes, thus instead preferable for template matching approaches.

- The first phase of segmentation is the computation of low-level, local shape descriptors associated to the original topography. This results in a series of *response maps*, each containing the value of one descriptor as obtained at any point x,y of the original topography.
- The second phase of segmentation consists of analysing the response maps (there may be more than one, if more than one local shape descriptor is computed) in order to identify an initial set of partitions.
- The third phase of segmentation (if needed) consists of post-processing the partitions, for example by removing small or malformed ones, or by merging some partitions into larger ones. The result of the third phase is the final set of partitions (final segmentation result).

The three phases are discussed in the following.

6.4.2 Phase I of segmentation - obtaining response maps

Several methods to obtain response maps are presented in this section. A response map is a map containing the value of a local shape descriptor as computed at any x,y location of a topography dataset.

Response maps by scanning

One way to obtain a response map is to compute a shape descriptor locally within a moving window, which slides over the entire topography, for example via a raster scanning approach [27]. The method is identical to what already defined in the section dedicated to template matching by scanning (see Section 6.3.1), except that in this case the shape descriptors computed within the moving window are only intended to capture local properties of the topography, instead of aiming to encode an entire shape in the attempt to match it to a template, as it was the case in Section 6.3.1.

Thus, assuming the shape descriptor produces a single scalar quantity to describe the region it has been computed onto (i.e. the contents of the moving window), then a descriptor value will be computed for each position of the window within the source

dataset. An example is shown in Figure 6.4, where the *Sku* parameter defined in the ISO standard 25178-2 [6] (see also Chapter 2) is computed within a moving window that travels over an example topography. In case of raster scanning and one-to-one sampling, if the original surface topography dataset (typically, a height map) has size $n_x \times n_y$, then a complete scan will produce a map of descriptor values (a *descriptor response map* as defined earlier in Section 6.4.1) the same size as the original topography. Smaller response maps will be generated if the window is stopped before reaching the borders of the source dataset, a common occurrence for those shape descriptors that are not particularly suited to handle missing values and are negatively affected by padding and/or mirroring solutions. In any case, the value at position i,j of the descriptor response map corresponds to the moving window whose centre is at position i,j of the source dataset; that is, each descriptor value is conventionally associated to the centre point of the moving window, albeit the value itself is the result of the processing of the entire window, which in turn means that two descriptor values adjacent in the response map will have been generated within moving windows with a large degree of overlapping.

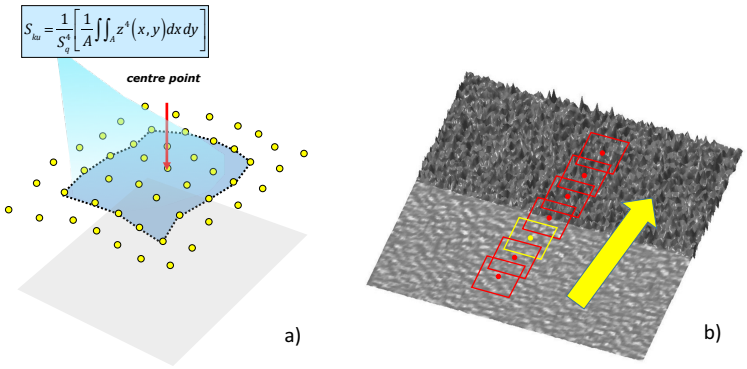


Figure 6.4: Example computation of a local texture descriptor (the ISO 25178-2 *Sku* parameter [6]) within a moving window that raster-scans a topography dataset (from original work by the author of this thesis [27]); a) the *Sku* parameter is computed within a square window, the result is assigned to the window centre point; b) the current window is highlighted in yellow, as part of a series of windows computed along a scan path.

If multiple local texture descriptors are computed within the moving window, the result is multiple response maps. When all the windows are the same size and resolution, the values of descriptors corresponding to the same position within each response map can be stacked up into an *attribute vector* (see Figure 6.5)

Past studies by the author of this thesis have investigated how size, shape and step spacing of the moving window affect the local variations of shape descriptors com-

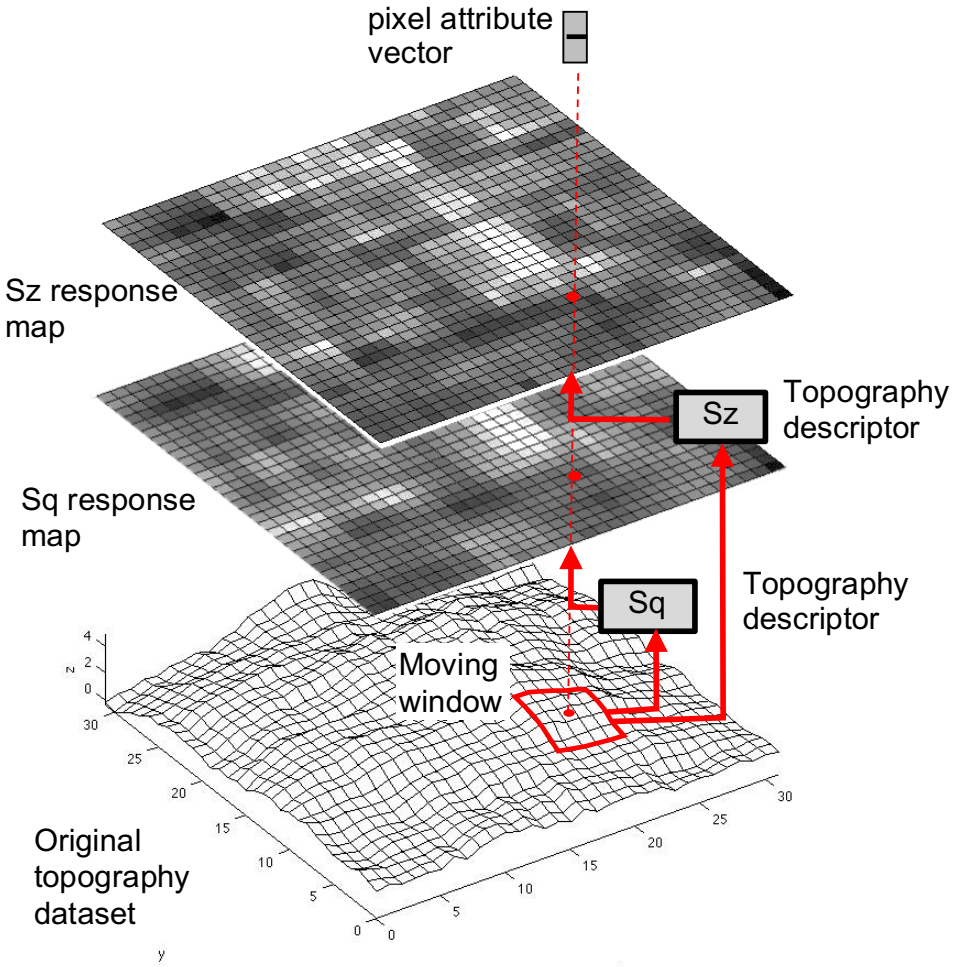


Figure 6.5: Example computing multiple response maps using ISO 25178-2 [6] texture parameters and construction of a multi-valued attribute vector. In the example, the *Sq* and *Sz* parameters are used to build a two-valued attribute vector associated to each topography map point ("pixel").

puted within it [27], and how different shape descriptors react differently to the scanning parameters. The main issues intrinsic to the scanning approach are related to the potentially long times required to scan the entire source dataset, if all the positions must be explored (e.g. by raster scanning), and considering that at all the positions the shape descriptor must be computed. The issue of computational time is exacerbated with larger moving windows. The similarity with convolution and correlation operations in digital image processing, also based on performing computations within moving windows, suggest that in some circumstances, specifically related to using shape descriptors represented by linear transforms, and exhibiting properties of orthogonality (i.e. rows can be computed independently of columns), computational times can be greatly reduced by performing computations in parallel, and using several other precalculation tricks. However, for the method to be general enough, any shape descriptor

could in principle be computed within the moving window, including non-linear ones, which will surely result into significant hits in computational time. In past work by the author of this thesis, the possibility to use ISO texture parameters as local shape descriptors computed within a moving window (in some case these parameters being significantly non-linear and non-orthogonal) has been explored [27]. Nowadays, this option is provided also in commercial software [87], though only as a qualitative exploration tool, to investigate variations of local texture, i.e. not for segmentation purposes.

The second main issue of using scanning methods is related to how shape descriptors react to local topographic changes happening within the moving window as it slides over the source dataset. Clearly, both how long it takes for the descriptor to react to a change, and how accurately is the transition detected in the x,y space, are factors significantly influencing the performance of the feature identification process, because these factor determine how long will it take to the scanning method to recognise that a transition has been identified, and how accurate the position of the detected transition will be.

Response maps from multi-scale analysis

Another family of methods which can be used to obtain response maps is referred to as *multi-scale analysis* in this thesis. Multi-scale analysis refers to the application of multi-scale filters to decompose a topography into its salient elements at different scales and (usually) orientations. Example of multi-scale filters include pyramid decompositions, wavelet filtering and the use of Gabor filter banks. Such approaches were originally described in Section 5.4.4 as techniques to obtain descriptors for topographic features, but can also be used to obtain response maps useful for segmentation (see [35] by the author of this thesis). Gabor filter banks produce response maps the same size and resolution of the original topography dataset (see Figure 5.16 for an example filtering result). However, wavelet filters and pyramid decompositions usually follow a dyadic progression, producing maps each downsampled to a half[4] of its predecessor (see Figure 5.17 for an example involving wavelets, and Figure 5.15 for an example involving pyramids). Thus, pyramid and wavelet-originated response maps need to be upsampled, usually by bilinear or bicubic interpolation. The concept is illustrated in Figure 6.6. After upsampling, an attribute vector can be constructed at each x_i,y_j location, carrying the contents of the associated cell in each response map. This method of constructing attribute vectors was already illustrated in relation to using texture parameters in combination to moving windows (see Section 6.4.2 and Figure 6.5) and is

[4] A quarter to be exact, since each side is halved.

similarly shown for multi-scale analysis in Figure 6.6.

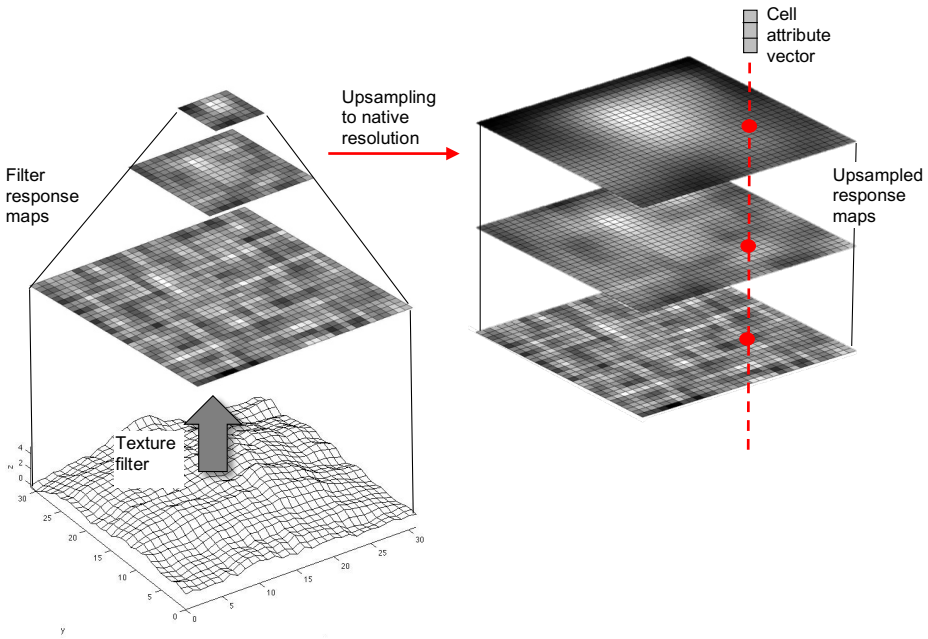


Figure 6.6: Response maps obtained by using a multi-scale filter; left) original topography dataset and three-levels Laplacian pyramid (computed through the difference-of-Gaussians method); right) the pyramid levels are upsampled to match the resolution of the original dataset. An attribute vector for each x_i, y_j position is obtained by stacking the values of all the response maps at the same location.

6.4.3 Phase II of segmentation: obtaining the partitions

Once the response maps have been generated, several methods have been suggested by the author of this thesis in order to use response map information to generate an initial set of partitions. Most of the methods described here more or less implicitly assume that partitions should be recognised as such because of some invariance in the local values of the associated response maps. This means implicitly assuming that partitions should be recognised as such because of constancy of some local properties. An alternative approach sees partitions recognised as such not because of internal constancy of properties, but because of being bounded by appreciable transitions. This latter approach is conceptually associated to *edge detection* and will be discussed later in Section 6.4.7.

The main methods that have been explored by the author to study for constancy and variation of local texture descriptors are summarised in Figure 6.7 and discussed in the following sections.

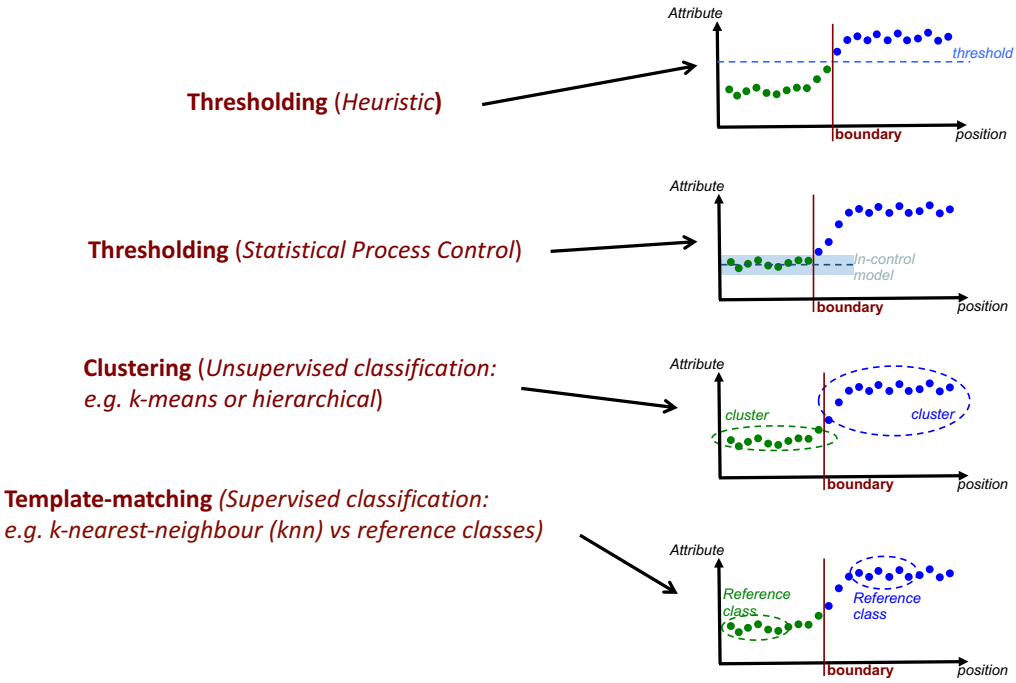


Figure 6.7: Overview of the main partitioning methods assuming a one-valued attribute vector and assuming a response map computed on a single profile.

Thresholding (heuristic) The *thresholding (heuristic)* method consists in setting a threshold value on a selected attribute value in the response map. For example, one may imagine to set a reference threshold value on the ISO 25178-2 texture parameter *Sa* [6] and then classify each topography point based on whether the associated *Sa* value (computed in a window centred on the point) is above or below the threshold. Similarly, if the attribute vector contains multiple attributes, then a consistent series of threshold values should be produced. The threshold value (or values) is chosen through heuristics, i.e. based on experimentation or past experience on the same class of problems/topographies. It must be pointed out that, despite the topography points are classified as belonging to either one of two classes (the "above" and "below" sets), this may not necessarily correspond to a partition of the topography into two regions, as the classification method is not influenced by whether the points are topologically connected with each other or not, when checking for the threshold. Also, it should be pointed out that in its simplest implementation, the thresholding method may directly consider the simple height of each point (i.e. the z value) as the only attribute, thus reducing the approach to a simple heigh-based thresholding. Thresholding may be global (i.e. only one value for the entire topography dataset), or local (also known as *adaptive*) meaning that the threshold value may change with x,y location on the topography, based on some parallel considerations about the nature of the local topography.

Additional considerations on partitioning based on thresholding can be found in previous work by the author of this thesis [22, 31]

Thresholding (statistical process control) This family of methods is ultimately still about thresholding, so the methods work equivalently and produce a binary classification (but not necessarily a binary partition). The only difference is how the threshold value is computed. This methods are mutuated from the literature on statistical process control [88]: the idea is to build a statistical model that represent the "in-control" state, and use such model to determine the threshold value that would trigger the detection of an "out-of-control" situation. This implies that one of the two "states" of a topography dataset must be investigated in order to build the model of the "in-control", the other being implicitly defined as any significant departure from the "in-control" state. To the purpose of defining the "in-control" state and associated thresholds to detect significant departures, there are as many statistical modelling routes that can be followed, each potentially giving origin to a new thresholding method for topography. The illustration in Figure 6.7 only uses one attribute (one-dimensional attribute vector) computed onto a moving window scanning a topography profile: this directly maps onto statistical process control applied to univariate series. The blue band drawn in Figure 6.7 indicates symbolically the boundaries of the "in-control" state, as determined by the model. In general though, for a proper application onto a complete topography dataset (height map in the form $z = f(x,y)$), possibly involving multiple attributes per data point, the adoption of multivariate statistical process control techniques is required. This venue of research is currently being investigated by the author of this thesis.

Clustering Clustering methods do not need a threshold, as the classes are defined by statistical clustering (univariate if one-dimensional attribute vectors are considered, multivariate otherwise). Common options for clustering are *hierarchical clustering* and *k-means clustering*, the latter having been used multiple times in several publications by the author of this thesis [27, 30, 35]. Not needing a threshold, which is intrinsic to clustering operating as an *unsupervised classifier*, is a major advantage, as there is no need to dedicate resources to investigating boundary values of the attributes that determine the transition from a class to the next. Secondly, clustering allow for multiple classes to be encoded, whilst thresholding only allows two[5]. The main difference between hierarchical and k-means clustering is that the latter requires the number of

[5]Strictly speaking, one could define multiple thresholds to isolate bands of attribute values mapped to multiple classes. However, this leads to a multiplication of the challenges related to threshold determination.

classes (clusters) to be defined in advance (meaning that some knowledge about what we are expecting to find on the surface is required). However, techniques exist to determine the optimal number of clusters that analyse the goodness of separation between the identified clusters (see for example [89]). Like for thresholding, the classification of each topography data point is ultimately independent of any topological relationship between the points themselves. Thus, two topography points may end up being assigned to different clusters only because of their attribute vectors, irrespective of where the two points are located, or even if they are neighbours. In Figure 6.8, the conceptual schema of how k-means clustering is applied starting from the response maps, in order to obtain a final partition of a topography dataset is shown. The example shown in Figure 6.8 assumes a two-dimensional attribute vector, leading to a clustering process in a two-dimensional attribute space.

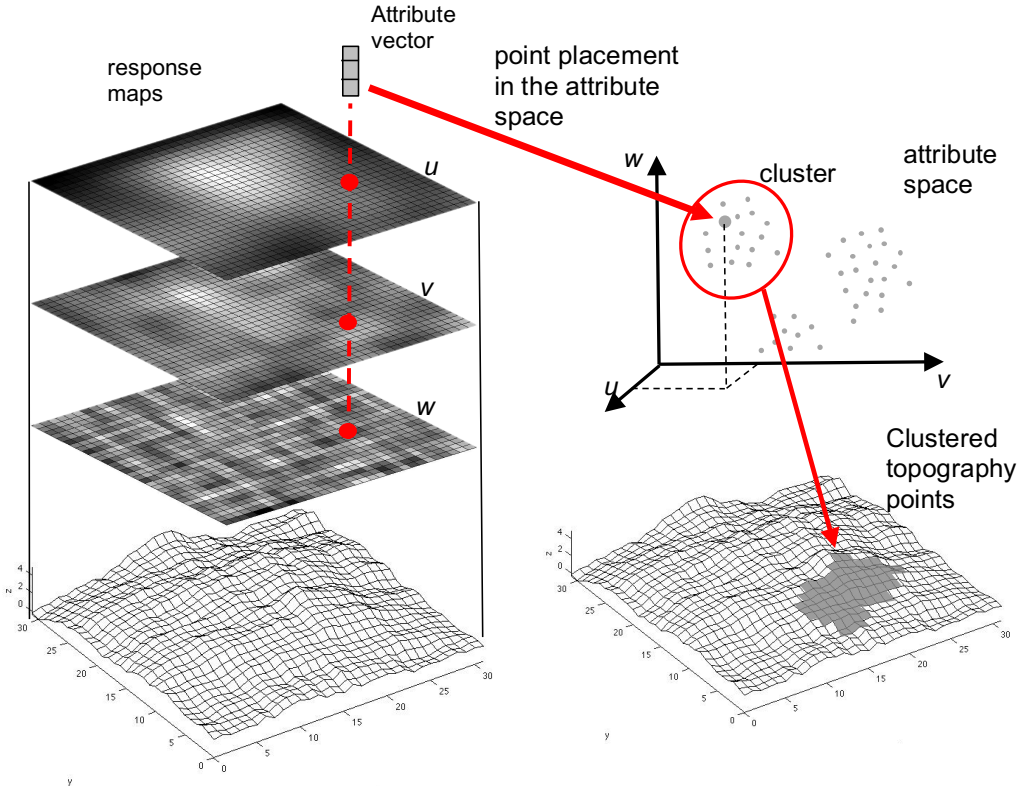


Figure 6.8: Schema of segmentation by using k-means clustering with multiple response maps (from work by the author of this thesis [35]).

Template matching The last family of methods goes under the name of *template matching*, and conceptually resembles the template matching methods that were discussed earlier in the chapter (see Section 6.3) to identify feature instances as appropriate matches to the reference shape (template) of any given class. However, the template here is not a feature shape, but rather a pattern of the attributes contained in the response maps.

Partitions are formed by checking the attributes of each point against all the template patterns, and determining which is the "closest" one. The issue of how to define distance of a point (defined by its n attributes) from a template pattern (a set of m points, each defined by n attributes) is a challenging one, but essentially the same problem of defining a similarity metric within a n-dimensional attribute space already faced by the *clustering* family of methods. Template matching and associated similarity metrics are currently being studied by the author of this thesis as examples of *supervised classification problems*, for example by using *K-D trees* and *k-nearest-neighbour (knn)* methods [90].

6.4.4 Phase III of segmentation: post-processing of the partitions

Regardless of the adopted partitioning method, the result of partitioning is referred to as a *segment map*, i.e. a type of classification map where map points carry the *id* of the segment they have been associated to (see Chapter 3 for an illustration of classification maps). Segment maps often need post-processing: because the transitions between segments are too irregular, or because the transition do not match one's expectations of where the boundaries between regions are supposed to be, or because of "noise" factors in the response maps altering the result of the partitioning process. A complete sequence of post-processing operations done on a segment map is shown through an example application (detection of regions with delaminated PVD coating) in Figure 6.9 (from previous work by the author of this thesis [91]). As shown in Figure 6.9, by far the most frequent reason for post-processing is due to clustering and thresholding methods not keeping into account topological connectivity of the points when they are assigned to the different classes or clusters solely because of their attribute vectors (see works by the author of this thesis discussing this issue [27, 31, 35].). The most typical post-processing operation is thus the test for topological connectivity (commonly referred to as *blob detection* in image processing [71]) that assigns different identifiers to regions that have been classified as belonging to the same class/cluster, but are actually topologically disconnected, and therefore should generate two separate partitions (see Figure 6.9.c). Other frequent post-processing operations are based on well known *morphological operators* for intensity and binary images (see [71]), such as *dilation* and *erosion* (and consequently, *opening* and *closing*). Morphological operators can be used to slightly alter the borders of each partition, in some cases also cleaning them by removing smaller, insignificant blobs most likely due to noise in the response maps. Partitions that are anomalous in size or shape can be automatically detected using *image moments*. The technique has been already introduced in Section 5.3.2 as a

means to compute datum axes for features. *Blob area* (i.e. size), *aspect ratio*, and other shape properties such as *circularity*, all computed from image moments, can be used to identify blobs that need to be removed (see Figure 6.9.d and Figure 6.9.e).

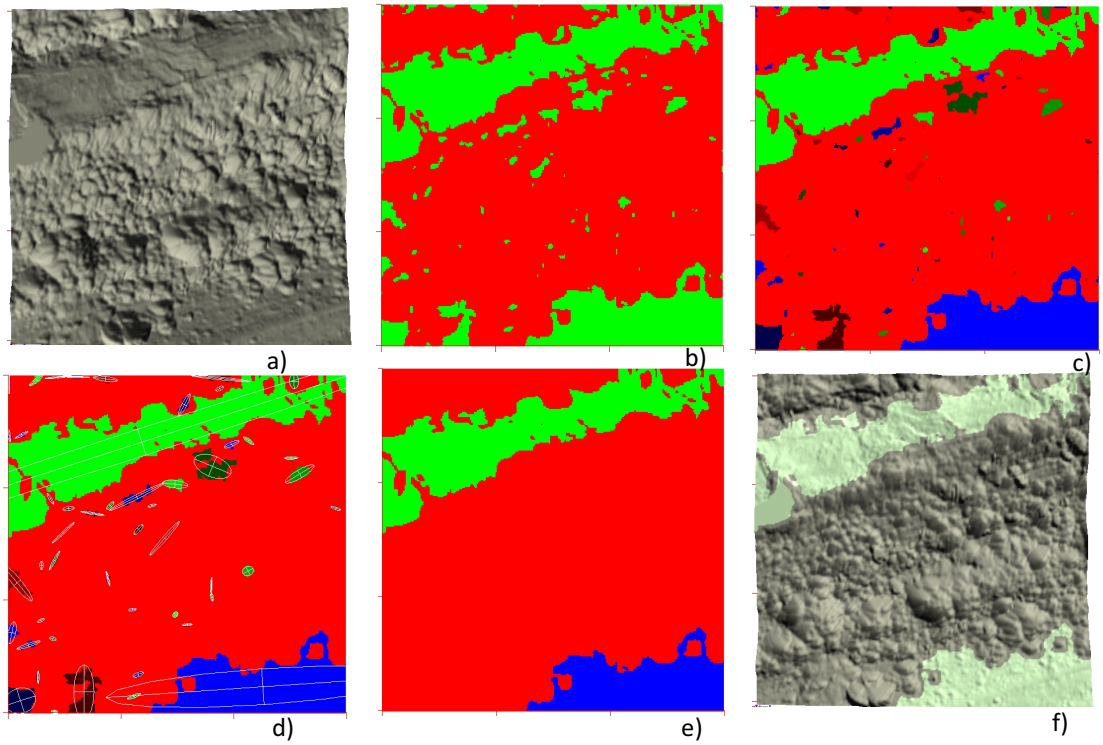


Figure 6.9: Example segmentation showing post-processing steps; a) original topography dataset: PVD coated surface measured with atomic force microscope (20×20) μm; b) initial segment map obtained by k-means clustering of Sa values computed on a moving window ($k = 2$); c) blob detection to identify topologically separate regions (shown in different colours); d) computation of image moments on the blobs; e) result of region elimination (based on area and aspect ratio, as computed from image moments); f) the remaining segments result as aligned to the uncoated regions in the original topography (from work by the author of this thesis [91]).

The most significant downside in post-processing is that at the moment the design of the operations is entirely trial and error, based on testing different post-processing parameters and pipelines in the attempt to match a desired result. The process can be quite inefficient and there is no robust way to indicate whether some choices will lead to an improved or degraded result. Also there is the problem of how to define the goodness of a segmentation result. The issue has been dealt with several times by the author of this thesis [22, 27, 31–33, 91]. Yet, at the moment there is still no general solution to obtain reference segmentation results on real measured data. One approach to evaluate the goodness of different segmentation solutions is to rely on simulated to-

pographies acting as reference benchmarks. A complete simulation framework for the generation of simulated topographies to validate segmentation and other topography data analysis and processing methods is presented in Chapter 8.

6.4.5 From segment maps to feature identification

The segmentation process only results in a partitioning of the topography dataset into regions (segments). The last step needed, is the actual recognition that one or more of the found segments are actually instances to the features being searched for. Strictly speaking, the establishment of a link (mapping) between the region and the class representative of the feature is the real identification process. In this context, segmentation is really only a form of lower-level pre-processing to facilitate identification[6]. At current state, the template-less, segmentation-based feature identification methods that have been investigated by the author of this thesis have always relied onto rule-based approaches to solve the last step. A region is recognised as instance of the feature class being searched for because of some compliance in region size, or shape, or position, or a combination of the above. In a sense, therefore, the methods for blob analysis which were discussed in Section 6.4.4 are also the methods that can be used at this step. Thus, image moments computed on specific blobs (regions) may lead to rule-based identification of a region as the target feature, because of conformance of area, position, aspect ratio, etc. Figure 6.9, which was already discussed as presenting an example of post-processing operation, can also be seen from the viewpoint of identification, as the two remaining segments visible in Figure 6.9.e are recognised as instances of PVD coating-deprived regions because of size checks (performed on the equivalent ellipsoids shown in Figure 6.9.d).

Most of the downsides of trial and error, rule-based post-processing of segment maps do apply also to the last identification step. Again, having an identification process that heavily relies on thresholds on region properties implies that each and every new application needs to be carefully studied, so that the identification algorithms can be properly tuned to it, in order for them to operate at the required efficiency.

6.4.6 Revisiting ISO 25178-2 morphological segmentation

ISO 25178-2 morphological segmentation [6], i.e. the partitioning of a surface topography into hills (regions whose points converge to the same local maximum) or dales

[6]Lower-level because it is based on local topographic properties which may not necessarily be directly related to the overall properties of the feature being searched for

(regions whose points converge to the same local minimum) was already discussed in Chapter 2 as a part of the current state of the art methods for surface metrology. Under the new general perspective introduced in the previous sections on surface topography segmentation, ISO morphological segmentation can be considered as a special case. In fact, in morphological segmentation the local attribute that is used to establish if a point belongs to a specific partition is the local minimum or maximum the point is connected to via the maximum downwards or upwards path [6]. Thus points are automatically assigned to partitions, each being identified by a different local minimum or maximum. Thresholding or clustering techniques are not needed, because the distance of each point from all the partitions is not distributed on a continuous space (a point either belongs to a partition, or it does not). In previous work by the author [20, 31] it has been discussed that, despite morphological segmentation being ideal to partition a topography into adjacent hills or dales, when applied to generate partitions that separate topography features of any other shapes, the method may perform less than ideally, in particular if the target feature cannot be easily obtained by recombination of identified hills or dales. An example, from original work by the author [31], is shown in Figure 6.10. The original dale map obtained by ISO morphological segmentation (Figure 6.10.a) is simplified by dale merging methods based on Wolf and area pruning (see [6] for details on the methods) (Figures 6.10.b-c). However, the final result (Figure 6.10.d) is still not capable of accurately capturing the boundaries of the groove.

6.4.7 Edge detection methods

A special mention in this work is reserved to edge detection methods. This family of methods belongs to the template-less, segmentation-based family of feature identification approaches. However, the mode of operation is slightly different, as described in the following sections.

Edges in surface topography

The definition of what is meant by "edge" in a topography dataset (height map) can be obtained by adaptation from image processing [71]. An edge is an arbitrarily thin and elongated region[7], which can be recognised as marking the transition between two additional, neighbouring regions located at both sides of the edge. "Conventional" edges are typically recognised as such because they may appear as abrupt changes (e.g. see the step-like feature in Figure 6.11.a). However, our visual perception may "see" an

[7] arbitrarily because there are no real rules on aspect ratio that set a threshold between what can be considered an edge, and what not.

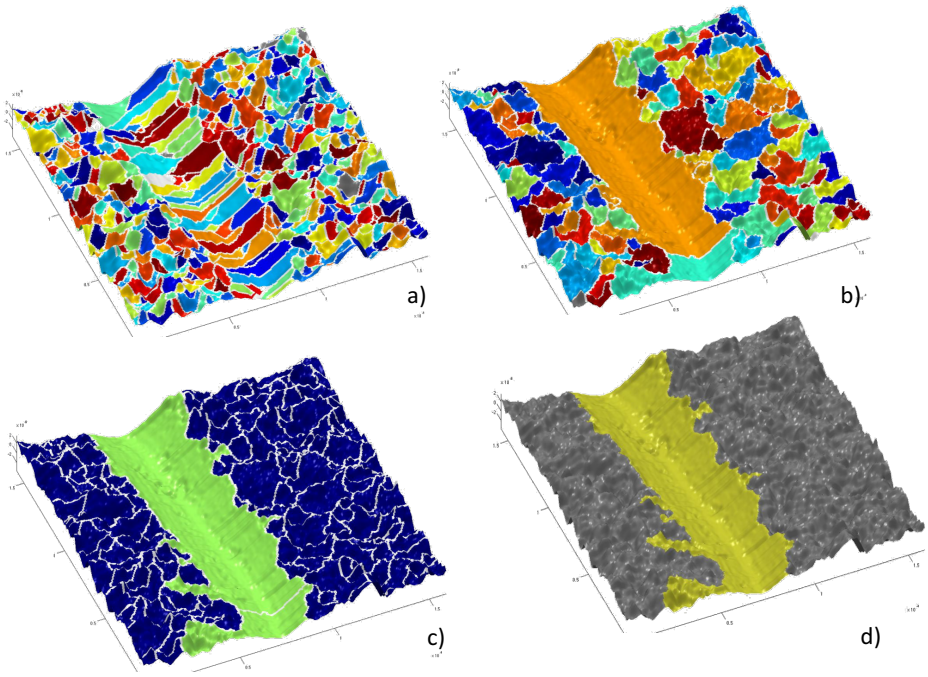


Figure 6.10: ISO 25178-2 morphological segmentation used to isolate a groove (from [31]). Despite area and Wolf pruning, the groove boundaries cannot be accurately captured.

edge even in absence of such properties, simply because two neighbouring regions may present significantly different textures (e.g. see Figure 6.11.b).

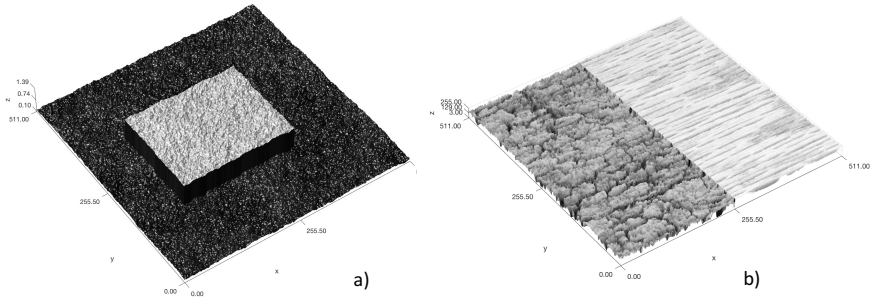


Figure 6.11: Types of edges which could be found in topography datasets (height maps); a) sharp transition between regions at different mean height; b) transition between two different textural patterns (implicit edge).

The case shown in Figure 6.11.b, is here referred to as *implicit edge*. Implicit edges are typically better dealt with, by making use of the segmentation methods introduced in Section 6.4.1. That is, by focusing on detecting the uniform regions on the sides of the edge, instead of focusing on the edge itself. The edge is then implicitly identified, as

the boundary of the identified regions.
On the contrary, this section will focus on the more "conventional" edges, defined through the assessment of changes in local slope or curvature.

In Figure 6.12 and Figure 6.13 two of the most common approaches for edge detection developed for digital images are shown, redesigned to operate on height maps. The figure uses cross-sectional profiles for simplification. In Figure 6.12, thresholding on the absolute value of the gradient is used to identify the edge. The threshold value determines whether an edge is captured (sharp transition of height leading to large gradient value) or not (smoother height transition leading to smaller gradient).

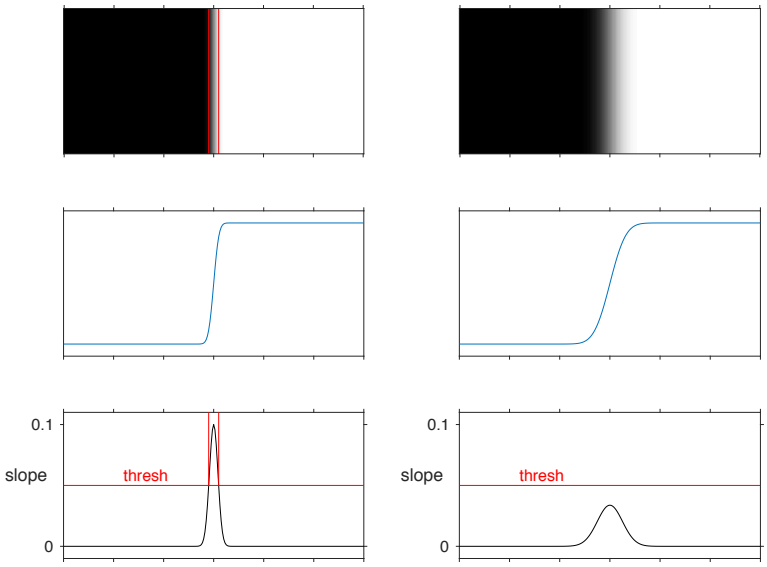


Figure 6.12: Gradient-based approach for detecting an edge; top row: height maps seen from above: sharp transition (left) vs. smoother transition (right); middle row: cross sectional profiles from the height maps; bottom row: plot of the absolute gradient and threshold for edge identification. In the left case, the edge is found and delimited between the two red, vertical lines. In the right case, the edge is not found (gradient below the threshold).

In Figure 6.13 the edge for the same profile is found via identification of the zero crossing point on the curvature plot (again shown in cross-sectional profiles for simplification). The method still needs some kind of threshold to detect whether the "crossing" point is worth being considered as edge.

Figures 6.12 and 6.13 provide a first hint at there being significant, conceptual differences between our understanding of what an edge is within a geometric shape, and what instead is classified as edge by using redesigned image processing methods. Let us consider the case of the sharp step in both Figures 6.12 and 6.13 (i.e. the left columns). Evidently, the actual step has two edges, one at the bottom and one at the

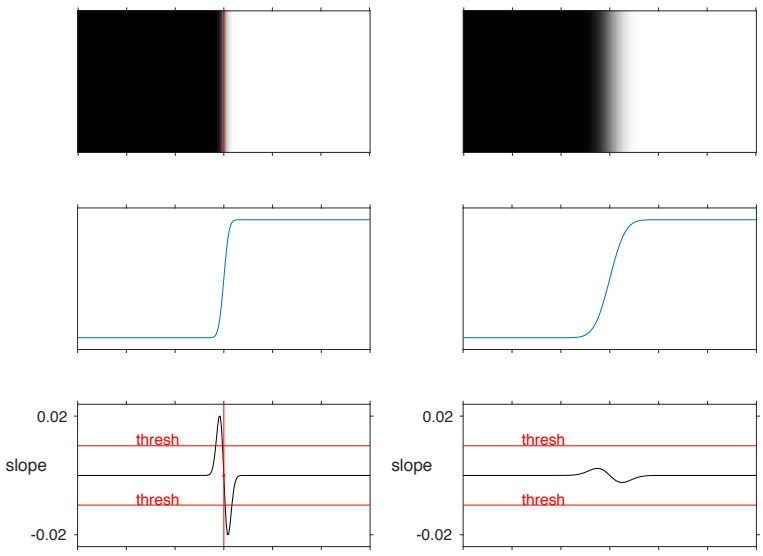


Figure 6.13: Curvature-based approach for detecting an edge (on the same test steps shown in Figure 6.12); bottom row,left: threshold crossing is verified. The edge is identified at the zero crossing; right: the threshold is not crossed. No edge is found.

top. These two edges would be aligned in x,y, if measurement had ideal, infinitesimal lateral sampling capabilities. In reality, in the measured dataset the two edges will actually be slightly pulled apart from each other, because of sample spacing dx,dy[8]. However, when using either gradient or curvature to identify the edges and assess their location, both gradient and curvature-based approaches result in a single edge, which will be roughly located between the two actual edges being searched for[9]. This fundamental discrepancy between the concept of edge in geometric modelling and the concept of edge from image processing is often overlooked when conventional edge detection methods are applied to height maps, leading to underestimation or overestimation of feature boundaries (and of any feature shape attribute, if based on such boundaries). This problem has been addressed in previous publications by the author of this thesis [5, 29, 31] and is particularly relevant for step-like features often found in highly deterministic surfaces (e.g. microfluidics components, micro-parts, micro-patterns). Some examples are given in Chapter 9.

[8]Topography aliasing due to averaging within each x,y position plays a role as well in determining where the step edges are located. Aliasing and averaging effects are discussed in Chapter 3.

[9]The gradient-based approach results in a region, which can be morphologically thinned [71] to a sigle-"pixel" edge, which will still be in the middle of the two actual edges

Edge detection via contour stability analysis

Local topographic irregularities make it difficult to detect edges with curvature and gradient-based methods. Determining a threshold value needed to filter gradient and curvature maps in order to isolate edges may be challenging, leading either to an overestimation or to an underestimation of the edges which may be present on a surface. In this section, an original method is presented, called *contour stability analysis*, developed by the author of this thesis [37] in the attempt to alleviate such issues. The method is illustrated with the help of Figure 6.14, where it is applied to identify the edges of spatter formations on a laser powder-bed fusion surface (Figure 6.14.a). The test case is described in Section 9.1.1. The method begins by sectioning the topography with planes parallel to the x,y plane where the height map is supposed to be laying. Starting from the top and proceeding downwards. Each sectioning operation is actually implemented as a height-based thresholding of the height map, resulting in a binary map (Figure 6.14.b). Contours can be extracted from the binary map by application of morphological erosion followed by subtraction [71], leading to a contour map (Figure 6.14.c).



Figure 6.14: Edge detection by contour stability analysis: example application to a laser powder bed fusion surface. Contour maps are generated by truncating the original height maps with places at increasing depths from the top; a) original height map; b) thresholded map (result of truncation); c) contour map (result of the application of morphological operators to the thresholded map)

.

As sectioning plane is lowered by a finite, vertical step, new contour maps are generated. An accumulating grid is defined (Figure 6.15), whose cells are as large as the desired maximum allowed variation of contour position which would still allow to consider it as a "stable" contour. The accumulation grid acts as a binning solution, each cell containing the count of contour points falling within it from the entirety of the available contour maps. A highly populated cell in the accumulation grid indicates a contour that, within the range of x,y positions captured by the cell, has remained

fairly stable across different sectioning depths. In the contour stability method, valid edges are defined (by the user, depending on application) as those contours that have not moved beyond a given x, y range (defined by the width of the accumulation cells), over a given z range (i.e. defined by a number of sectioning planes and their vertical spacing). Thus, to obtain the final edge map, it is sufficient to select the bins (in the accumulation map) whose count contents are compatible with the vertical range.

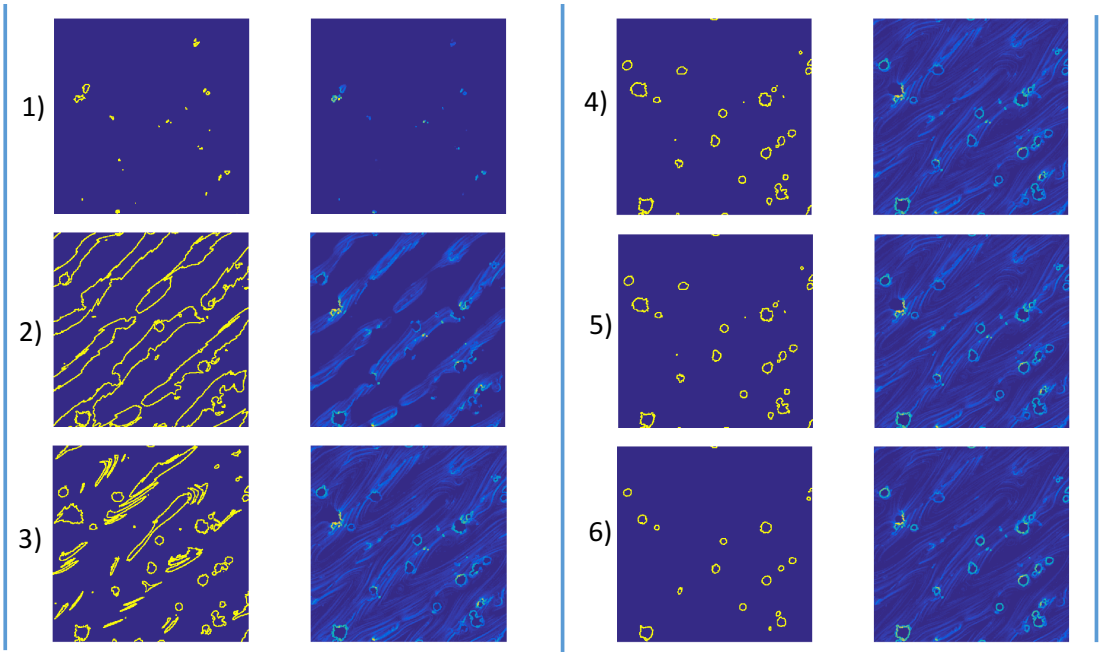


Figure 6.15: Edge detection by contour stability analysis: example application for spatter detection on a laser powder bed fusion surface. Contour maps vary with sectioning plane, and contribute to populate an accumulation grid. Stable contours eventually dominate the grid. Example images divided in two main columns, numbered rows indicate increasing sectioning depth. Within each main column, the two sub-columns show the contour map at each sectioning depth (left), and the complete accumulation map, obtained by adding all the contour maps up to the same depth (right)

.

It should be also noted that defining a maximum lateral (x, y) variation over a given vertical (z) variation, which is what the contour stability method does to identify edges, is conceptually equivalent to defining a threshold on slope (since the gradient is defined as a vertical variation over a lateral length). However, it may often be easier to specify boundaries of topography features by explicitly indicating how much the side surfaces are allowed to move on the x, y plane over a given height range, as opposed to defining a threshold on local slope. Moreover, the binning mechanism implemented by the contour stability method ensures that most small-scale disturbances in edge localisation (which would be typical in conventional edge detection because of the irregularity of

surface topography) are captured within the same bin and averaged out by the accumulation mechanism. The edge map resulting from the contour stability method suffers from the same intrinsic issues of any other edge map, primarily considering that edge detection is only a preliminary step towards a complete feature identification result. The issues and methods for turning an edge map into a complete feature identification result are discussed in the next section.

From edges to features

Regardless of whether it has implemented via contour stability analysis method or any other more conventional approach, edge detection can be defined as segmentation process. This is because, in a sense, identified edges are partitions that differentiate themselves from the other "non-edge" regions. However, the ultimate goal of edge detection is, in the context of this chapter, not the identification of edges per se, but rather,the identification of the topography features that are delimited by such edges. The problem is far from trivial, because in real-life topography data, surface irregularities greatly affect edge-detection methods, often giving origin to fragmented edges, at times entirely missing large portions of the boundaries of the target features, or mistakenly recognising as edges regions which should not be identified as such. This phenomenon is sometimes less evident in the contour stability analysis method, but nevertheless present, as illustrated in Figure 6.16.a, where the same edge map previously obtained in Figure 6.15 with contour stability analysis is shown. Figure 6.16.b shows the result of another popular edge detection approach, based on the Canny algorithm [71]. Albeit looking remarkably different, the two edge maps share the same type of problems.

Whilst humans are good at visually extrapolating missing information and at reconstructing lost connectivity via inference, the same cannot be said about algorithmic procedures. An edge map, the result of edge-detection, is a binary map, formally equivalent to a binary image. Thus the result of edge detection can be further manipulated using methods developed for binary images. As previously discussed in Section 6.4.4 for post-processing segments, the same morphological operators of dilation, erosion, opening and closing [71] can be therefore applied in the trial and error attempt to "clean" the edges, with the aim to reconstruct the topological connectedness necessary to isolate the closed boundaries that delimit each feature of interest. If edge are found that form closed boundaries around the features of interest, then a simple blob detection process [71]) applied to the complementary of the edge image will lead to a useful segmentation result (see Figure 6.17). After segmentation, the proper feature identifi-

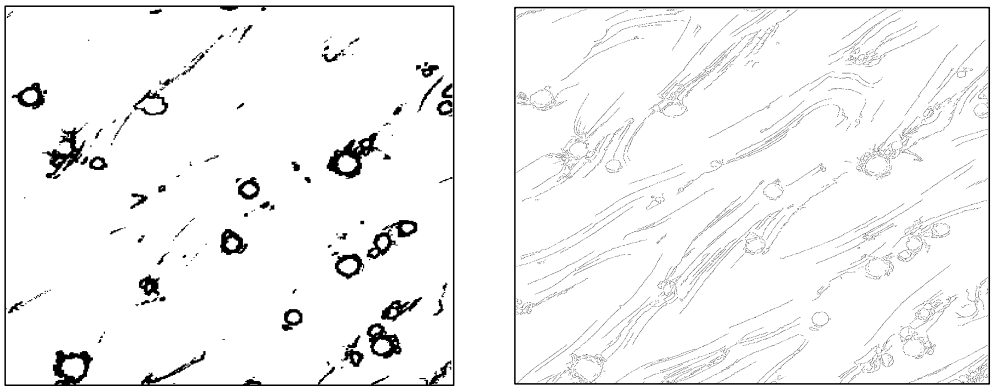


Figure 6.16: Edge maps for the laser powder bed fusion surface previously shown in Figures 6.14 and 6.15; a) edge map by contour stability analysis; b) edge map from the Canny algorithm (note that the edges in the Canny algorithm are automatically thinned, whilst no thinning has been performed in the edges obtained by the contour stability analysis method)

cation result can be obtained by applying heuristics on segment area, size, aspect-ratio etc., with all the issues previously discussed in Section 6.4.5.

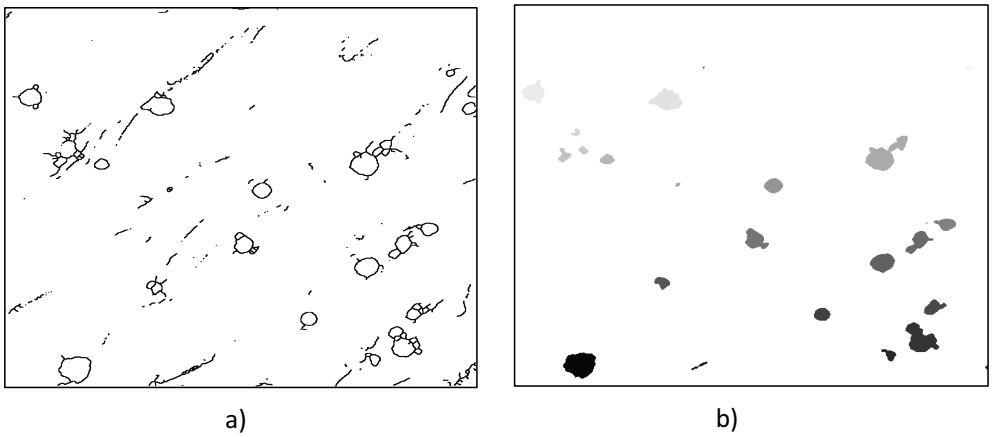


Figure 6.17: From edges to segments. Example on the edge map obtained by contour stability analysis; a) original edge map (thinned); b) result of blob detection on the complement of the edge image.

6.5 Defect identification in regular patterns

The characterisation of feature patterns was discussed in Section 5.6. If a pattern is nominally regular, i.e. its geometric and topological properties are constant within an observational window which traverses the surface, then any departure from such regularity can be seen as a "defect", and thus treated as a feature identification prob-

lem. As previously discussed in Section 5.6, feature patterns may be conveniently represented through graph-like data structures. Under such assumption, all the feature identification techniques described in this chapter and targeting individual features in the conventional sense (i.e. regions whose topography makes them notable) may be also redesigned to operate on graphs, and thus made capable to recognise "features" intended as local anomalies (i.e. "defects") in a graph. The techniques to obtain partitions starting from computation of local properties, as discussed in Section 6.4.1 and summarised in Figure 6.7, can be adapted to operate on the geometric and topological attributes of graph-like structures. In Figure 6.18, the example low-friction surface already discussed in Section 5.6, made of a regular array of micro-pockets, is used to demonstrate how simple thresholding can be used to remove graph nodes corresponding to malformed feature instances, thus resulting in anomalies detectable in the graph.

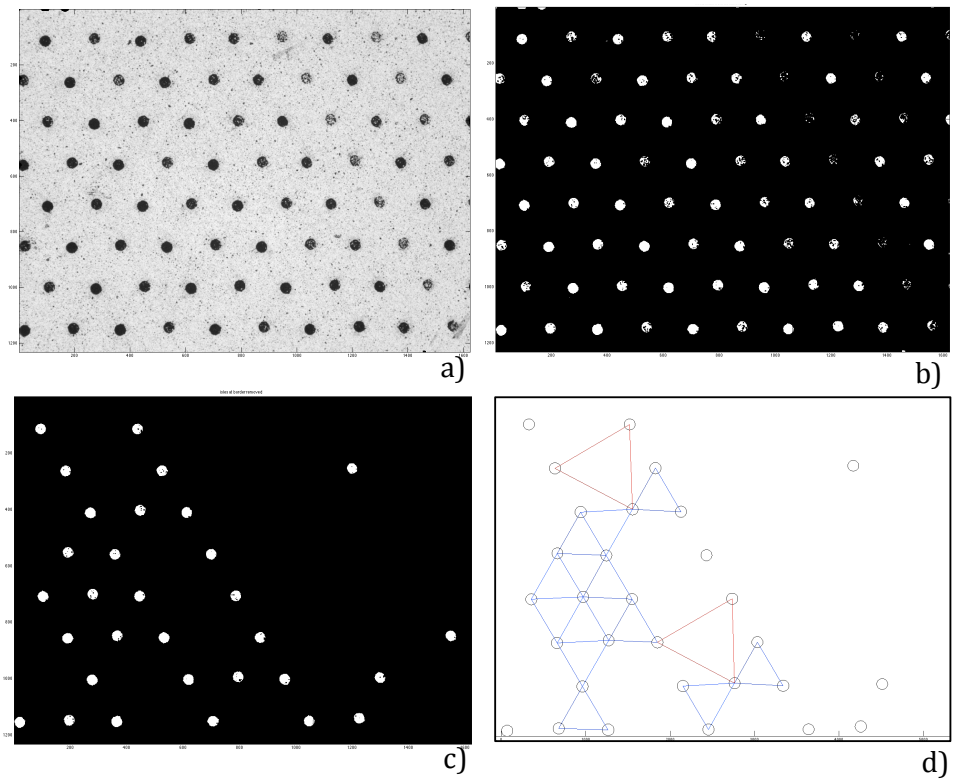


Figure 6.18: Identification of pattern defects on a low-friction surface populated of micro-pockets; a) original surface topography (height map); b) feature identification by height thresholding; c) removal of malformed features (through checks on footprint shape and size; d) missing features resulting in missing nodes and altered connectivity in the graph.

6.6 Feature extraction

Feature identification terminates when successfully establishing a map between a x,y region of the original topography dataset and the class of the feature being searched for. The result of feature identification is a classification or selection map, where the x,y locations having been identified as belonging to the target feature have been appropriately flagged as different from the other, surrounding locations.
The next step (if needed) is *feature extraction*, to which the last part of this chapter is dedicated. Feature extraction consists of separating the identified region to form an independent topography dataset. As illustrated in chapter 3, such dataset is encoded as any other topography dataset (i.e. as a height map, with a number of additional maps possibly associated to it, for example, a colour map, or a quality map (see again Chapter 3)). In addition, the new dataset is associated to a mask, defining the boundary of the feature that most likely will not correspond to the rectangular region defined by the extracted dataset.

6.6.1 Feature boundary determination

Feature identification usually terminates with an indication of the x,y region corresponding to an instance of the feature being searched for. In template-less, segmentation-based identification, the contours of the segments that have been recognised as instances of the features being searched for, are commonly assumed as the being the boundaries of the features themselves. For example, in Figure 6.19 it is shown that the segments previously shown in Figure 6.17.b can be recognised as instances of spatter formations, directly mapped back onto the original height map, and used to delimit the feature boundaries (the application is described in detail in Section 9.1.1).

However, in template-based identification, there is no segment map to provide contours for feature boundaries, and often such boundaries cannot be easily derived from knowledge of the template. An example of the issue is shown in Figure 6.20, from previous work by the author of this thesis [33]: template-based identification of the grits in the super-abrasive surface results in several positive identifications (Figure 6.20.a). The identification is based on using the ring projection transform (RPT) - see Section 5.4.2) together with a cylinder-like template (further details have been given in Section 5.4.2). The identified features are all quite different from the template shape, as clearly shown in Figure 6.20.b, where a feature instance is shown overlaid with the template. The alignment corresponds to the position of maximum similarity between the template and the feature, however, the boundaries of the two geometries are clearly different.

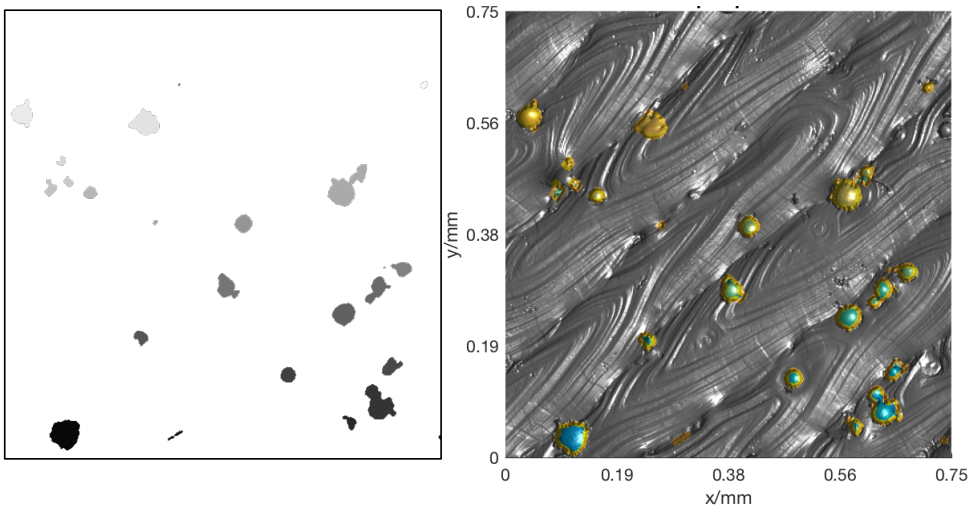


Figure 6.19: Feature boundary identification from segmentation contours. Example on spatter formations in laser powder bed fusion (from [37]); a) segment map (same as shown in Figure 6.17; b) segment map overlaid on top of the original height map. Segment contours (highlighted in yellow) are used to identify the feature boundaries.

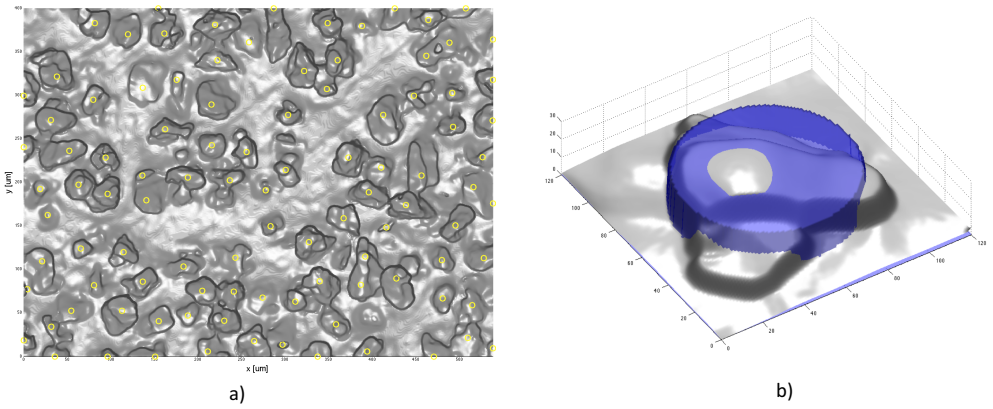


Figure 6.20: Template-based identification and the challenge of feature boundary determination (example from previous work by the author of this thesis [33]); a) result of template-based identification: the position of each region corresponding to a match is highlighted in yellow; b) because of the intrinsic variability of each feature instance, the template shape, shown in transparent colour, aligned to the feature instance, may present significant boundary differences.

The problem of finding the boundaries of identified features in template-based identification has been tackled in previous work by the author of this thesis [28, 32, 33]. Albeit at the moment a general solution has not been found, an approach that seems to work sufficiently well for a range of cases consists in adapting *region growing* algorithms [71] in order to automatically search for the feature boundaries starting from the seed points corresponding to the location of maximum match between each feature and the tem-

plate (i.e. the yellow markers shown in Figure 6.20.a). An example application of a region growing algorithm is shown in Figure 6.21, from [33]. Following a conceptually similar approach to the *contour stability analysis* method of edge detection, previously illustrated in Section 6.4.7, an increasingly larger region in the x,y topography plane is considered, starting form the x,y point corresponding to the template-matching position (any yellow point in Figure 6.20.a). Growing is obtained by horizontally slicing the feature with increasingly lower parallel planes (the feature is supposed as protruded, since it is representative of a grit). Stability of cross-sectional area is used to determine the growing termination condition. Essentially, if the cross-sectional area does not increase significantly any longer, it means that the slicing has encountered very steep side walls, which is what we would normally expect to find at the sides of a grit[10]. Note that this method only works for feature with steep walls.

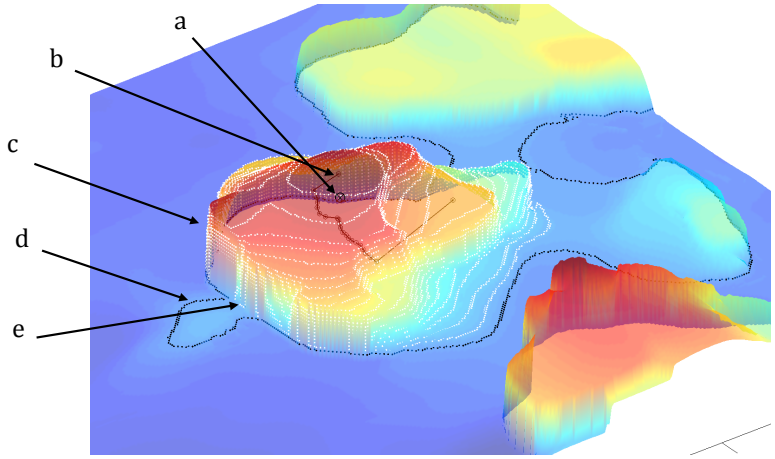


Figure 6.21: Feature boundary determination by region growing. Example application on grit from super-abrasive surface (from [33]; b) marker from template-based identification, used to determine the first slicing height; a) feature centroid of the first slice; c) one of the cross-section as slicing proceeds at increasing depths; d) slice corresponding to a change in cross-sectional area beyond acceptable threshold; e) last cross-sectional contour within the threshold (adopted as feature boundary).

6.6.2 Feature separation

The final step of feature extraction consists of the actual separation of the feature region and the creation of an independent dataset. As the creation of an independent dataset implies that the feature is referred to a new co-ordinate system, feature separation may not always be needed. For example, one may be interested in knowing the relative position between two features, in which case their extraction would imply the loss of

[10] note that the check is not explicitly done on the gradients to avoid the condition where on one side the grid smoothly merges into the surrounding substrate.

a common coordinate system. However, when characterisation is aimed at extracting shape properties intrinsic to the feature itself (i.e. not dependent on any external datum), then separation helps creating a simpler and smaller dataset that can later be handled, using the large variety of methods illustrated in Chapter 5. In terms of actual implementation, feature separation is possibly the simplest of the operations on topography data amongst those illustrated this far. Assuming the feature has been identified within a height map, then separation implies the creation of a new height map, which - by definition - will be defined on a rectangular footprint like any other height map (see Chapter 3 for a discussion on surface topography data representation formats). If the original height map is accompanied by additional maps (for example, colour maps, quality maps, other masks, etc., see the definitions reported in Section 3.7), then the corresponding portions of such maps may be extracted as well, depending on application. As the feature will most likely not be of rectangular footprint, typically the smallest rectangular footprint which the feature is inscribed to is computed. Together with such rectangular region, a new mask is generated to more accurately capture the feature boundaries within the cropped region. An example of feature separation involving a pre-existing mask and a new mask for accomodating non-rectangular feature contours is shown in Figure 6.22.

Again depending on needs, the feature may be resampled within the new coordinate system. In some cases, finally, instead of creating a rectangular crop from the original dataset (i.e. aligned with the x,y axes), Principal component analysis or image moments (see Sections 5.3.2) and 5.3.2 can be used to re-orient a local co-ordinate system more suitable to capture the feature shape (useful for example when features are supposed to be stored in a database for future comparisons).

6.7 Summary and outlook

The concept of *feature identification* has been presented. Feature identification methods have been formally defined and organised according to an original taxonomy made of *template matching* and *template-less, segmentation-based* methods. For both classes of methods, formal definitions and subclasses have been proposed to accommodate existing feature identification methods as well as original ones developed during the doctoral studies. Feature extraction and boundary identification operators have been presented and discussed as well.

A fundamental open issue that remains to be addressed related to how to better investigate how shape variability across instances (introduced by the random elements

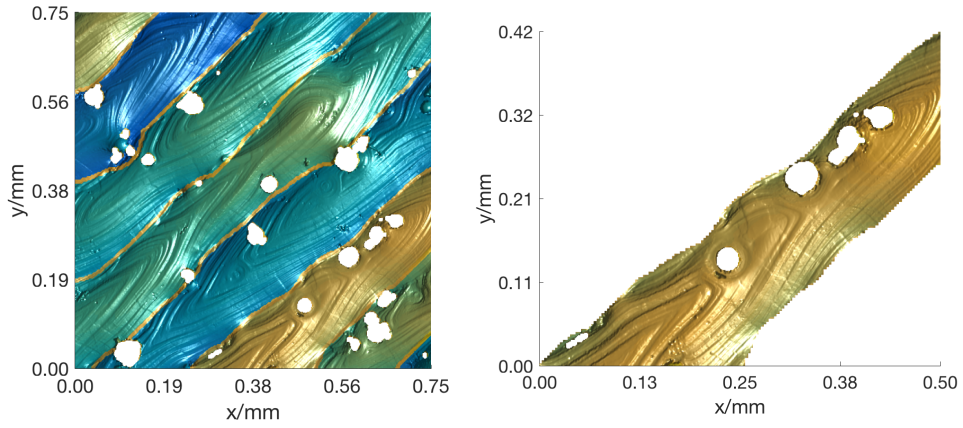


Figure 6.22: Example feature separation with masking and no resampling (from original work by the author [37]); a) weld tracks are identified on laser powder bed fusion surface (shown in different colours). Note that spatter formations, recognised at a previous step of the analysis are removed from the topography; b) separation of individual weld track as independent dataset. The surrounding regions have been voided by a dedicated mask.

present in the surfaces themselves, and in their measurement processes) may influence feature identification. The concept of robustness of the shape descriptor and associated similarity metrics has been discussed both in the previous and in this chapter. However, a more comprehensive investigation is needed, delving deeper into how to make a better use of solutions for statistical shape modelling. Analogous randomness-related issues plague feature boundary determination. In this, case, the concept ties directly back to how to perform the spatial partitioning of a random field, and deserves further research efforts.

Chapter 7

Error and uncertainty in feature-based characterisation

7.1 From uncertainty in measurement to uncertainty in the characterisation of surfaces

According to the *International Vocabulary of Metrology (VIM)* [92], measurement is defined as the process of experimentally obtaining one or more quantity values that can reasonably be attributed to a measurand. The simplest measurement process one can think of may be imagined as resulting in single scalar number, located on a one-dimensional axis: for example, a temperature value, as read by a thermometer. A measurement result should be associated to a *measurement uncertainty* whenever possible. Measurement uncertainty is meant to express lack confidence in the measurement result representing the true value of the measurand. Lack of confidence derives from awareness of measurement processes and systems being imperfect, leading to results which will differ from the true value. Measurement uncertainty is commonly expressed as an interval: the interval indicates the region within which one would expect to find the true value, starting from the measurement result. Details about the available approaches to estimate uncertainty can be found in the *Guide to the expression of uncertainty in measurement (GUM)* [93] and supplement [94]; and in the NPL good practice guide series [95].

7.1.1 Uncertainty in surface metrology

The problem of how to consider and compute uncertainty in the evaluation of surface topography is challenging, and subject of current research in surface metrology. A

recent review of the main issues can be found in [96]. A typical surface metrology analysis session is a complex pipeline, consisting of data acquisition, followed by multiple data processing and analysis steps. Can it be considered as a "measurement process"? It depends on what we mean by "measurement".
To begin with, we are comfortable with measurement processes resulting in single scalar values, like a temperature value, or a length. After all, uncertainty is defined as an one dimensional interval, thus it is applicable to single scalar values only. Thus, if the surface metrology pipeline leads to a single scalar result, for example the value of a texture parameter, or the value of a geometric property of a feature (in case of feature-based characterisation) such a depth, area, length, etc., then one could (at least in principle) be comfortable with the general idea that an uncertainty budget could be associated to such results (texture parameter value, attribute of a feature), exactly how one can associate an uncertainty interval to a temperature value.

The real challenges arise when considering the following issues.

- A surface characterisation pipeline may not terminate with a scalar value as a result. Some pipelines, in particular in feature-based characterisation, may terminate with a "yes/no" answer, related to the detection of presence/absence of a target feature (e.g. defect). Defining uncertainty for a non-numeric (in this case, categorical) result is a non-trivial problem.

- The main challenge in the study of uncertainty is the estimation of the amount of uncertainty that should be associated to the final result. A typical pathway for the estimation of uncertainty goes through the decomposition of the pipeline, the analysis of the error sources, and the formulation of uncertainty budgets, or at least error models, for all the intermediate results generated at the different stages of the process. However, many of the steps of any surface metrology pipeline do produce very complex datasets. far from simple scalar quantity values. Even the initial step of a typical surface topography analysis pipeline, i.e. the actual acquisition of surface topography data through the use of topography measurement instruments, results in complex datasets (height maps, colour maps, triangulated meshes. etc. as shown in Chapter 3), that are significantly more complex than a single temperature readout. How should uncertainty be defined for such complex and structured bits of information? For example, how should uncertainty be defined for the co-ordinates of the measured surface points, and their topological relationships? An established solution to such questions is not currently available.

Many of the above issues are not of novel introduction by surface metrology, and have been studied in different domains. For example the problem of what uncertainty may mean for image-like data, and how uncertainty budgets should be formulated in applications of computer vision has been dealt with in the literature (see for example the work by Yi et al. [97], Zhou et al. [98] and Weibensee et al. [99]). Computer vision and feature-based topography characterisation share a lot of commonalities, as shown in the previous chapters, basically because of the mathematical similarity of the underlying datasets (see Chapter 3). Clearly though, conventional optical cameras and surface topography measurement devices may have points in common, but also significant differences. Thus, further studies on the internal mechanisms underlying the major established areal topography measurement technologies to understand error sources are in progress (see for example [100] for a study on focus variation technologies, and [101–103] for the problem of how to calibrate the axes of areal topography measurement devices).
Then again, a surface topography characterisation pipeline, whether conventional (i.e. addressing the computation of texture parameters) or feature-based, is not only comprised of the surface data acquisition step, and numerous further data processing phases need to be considered. Assuming that an appropriate representation (and estimation) of uncertainty can be carried on for a surface topography dataset, it is still necessary to investigate how uncertainty propagates through the data analysis and manipulation routines, and how it combines with new sources of error originated within the algorithmic data processing pipeline, all the error components ultimately converging into the uncertainty budget to associate to the final result. At the moment, studies targeting uncertainty on the entire surface topography characterisation pipeline are scarce. Amongst those dedicated to surface metrology in general, a recent overview of the issues can be found in [96]. Fewer publications specifically target feature-based characterisation, such as [104], and [105].

In the following sections, original (published) work by the author of this thesis during the doctorate is illustrated. In Section 7.3, an original method to associate repeatability error to surface topography datasets is presented, published in [58, 59]. The possibility of adding further information to estimate also bias with respect to a more accurate measurement is also discussed. In Section 7.4 an original method developed during the doctorate to associate reproducibility error to dimensional characterisation results for single features is illustrated, published in [106]. In this latter case, no attempt is done to investigate systematic components, thus no attempt is made at estimating actual uncertainties. Before introducing the methods, however, some considerations are

presented about the issues of estimating uncertainty by following purely experimental routes based on repeated experiments (i.e. the approach followed in the methods illustrated in the following sections).

7.2 Estimating uncertainty through replicate experiments

The repeated application of the same pipeline dedicated to topography measurement, feature identification and feature characterisation (for example aimed at computing the depth of a pocket starting form a given, physical specimen) will result in a *dispersion* of the targeted feature attribute (for example: depth value). Dispersion is a challenging concept: in the theory of measurement uncertainty, dispersion is considered as general indication of discrepancy of measured quantity values with respect to the true quantity value. Uncertainty should be estimated starting from this kind of dispersion, i.e. spread about the truth. Clearly, such a dispersion carries information on systematic and random measurement error components, and thus is affected by both *accuracy* and *precision* of a measurement. However, usually when performing replicate experiments, and under the condition that the true quantity value is unknown, the only appreciable "dispersion" is that between the quantity values themselves. In such conditions, without any further knowledge of the problem except for what one has from replicate experiments, one would build an uncertainty interval based solely on the dispersion of the values between themselves. In terms of error components, clearly this new type of dispersion can only capture random error and systematic (i.e. deterministic) non-constant components. That is, it cannot detect constant bias.

It follows that, if one is to embark in a purely replicate-based approach to the evaluation of uncertainty, care should be taken to make sure any systematic effect is removed, considering in particular constant bias, the most dangerous offender[1]. Removing constant bias implies the availability of a more accurate result to compare experimental observations with, another challenging concept for surface metrology in general, and feature-based characterisation in particular. In fact, it is not clear where more accurate results could be taken from. If one were to concentrate only on the surface topography acquisition step of the pipeline, then a more accurate reference may be represented by another measurement taken by an instrument a-priory recognised as more performing. However, as shown later in Section 7.3.1, this is seldom possible with current state-of-the-art technologies. If one were to concentrate on the data analysis and processing steps of the pipeline, then a more accurate reference may be formed by different classes

[1] variable systematic effects are easier to capture through regression modelling, whilst constant bias is basically invisible in replicate experiments.

of algorithms, or implementations. In the latter case, synthetic geometry would likely be needed to separate the algorithmic-induced error from uncharacterised error in the input datasets. In Chapter 8, the generation of simulated topographies is discussed, along with the issue of incorporating simulated measurement error. In absence of further knowledge of the truth, based solely on the results of repeated experiments, one can only assess errors (and thus the uncertainty components) related to the conditions that define the repeated experiments. The concepts of *repeatability* and *reproducibility* from the VIM [92] can be adopted, but should be handled with care. Such concepts are only loosely defined for complex pipelines such as those of conventional and feature-based surface metrology.

7.3 Estimating uncertainty in surface topography measurement via experimental evaluation of repeatability error and local bias

An original method has been proposed by the author of this thesis [57–59] that can serve the dual purpose of evaluating surface topography measurement error in repeatability conditions, and evaluating local bias with respect to a (supposedly) more accurate measurement. The two evaluations combined can lead to a first estimate of uncertainty that could be associated to a surface topography dataset.

The method works for any type of replicate measurements, for example in repeatability conditions and reproducibility conditions. The choice of conditions influences the interpretation of the final results. In this section the method is shown assuming repeatability conditions. Thus, multiple measurements are performed on the same region of the sample, in sequence, without sample repositioning and without any change of operator, procedure, or instrument set-up (number of replicates indicated as k).

To simplify the illustration and the notation, the procedure is explained on profiles ($z = f(x)$), with the help of Figure 7.1. However, the method is indifferently applicable also to surfaces ($z = f(x,y)$). In Figure 7.1.a, the measured datasets are relocated in the same co-ordinate system (a procedure known as *registration*, see work by the author of this thesis [34, 69]). This is typically done by global minimisation of the squared Euclidean distances of the points of one dataset to the closest pairwise matches of the other[2]. Registration is necessary because, despite the sample not having been moved between measurement (repeatability conditions are assumed), some slight misplace-

[2]In absence of an external reference for co-localisation

ments may still exist because of repeatability error in the axes of the measurement system. After registration, the datasets are resampled in space, so that, for each location (x_i for profiles), k replicate z values $(z_1(x_i), z_2(x_i), \cdots, z_k(x_i)$ are available. In Figure 7.1.b the resampling is shown for three registered profiles. The use or replicate measurements as a starting point to build a random field was discussed in Chapter 4. In this case, a slightly simpler approach is adopted where complete independence between z values located at different x_i positions is assumed. A partial estimation of the random field can be obtained through the first and second-order moments. Thus, local sample mean $\bar{z}(x_i)$, and local sample standard deviation $S(x_i)$ are computed at each x_i profile location[3]. Because of the assumption of independence, the covariance matrix of terms $k(x_i, x_j), \forall i, j = 1, 2, \cdots, n$ is not estimated. Then, a confidence interval for the local population mean $\mu(x_i)$, referred to as $CI_\mu(x_i)$ is built at each x_i location of the profile assuming a t-distribution with $k-1$ degrees of freedom at each location (see Figure 7.1.c)[4]) Thus:

$$CI_\mu(x_i) = \bar{z}(x_i) \pm t\left(S(x_i)/\sqrt{k}\right) \tag{7.3.1}$$

where t is the critical value computed from the t-distribution with $k-1$ degrees of freedom, at 95% confidence level. The CI values can be assumed as an estimate of local repeatability error. They can be interpolated to obtain CI estimates related to previously unsampled locations. The upper and lower limits of the interpolated CIs define the upper and lower surface boundaries within which the population mean of the local heights $\mu(x)$ can be found at the given confidence level.

An example computation of the local repeatability error is shown in Figure 7.2.a, in this case directly calculated on profiles. (generated from data published in work by the author [58]). In Figure 7.2.b an entire surface is shown with its upper and lower surface boundaries (generated using data from [57]).

7.3.1 Estimating local bias

The investigation of replicate $z(x_i)$ values alone does not allow one to capture any bias existing with respect to the local true value (i.e. the true z value in x_i). Consequently, any assessment of uncertainty solely based on the local confidence interval on the mean $CI_\mu(x_i)$, would fail to capture a potentially significant error contribution related to local lack of accuracy. Ideally, a zero-error, perfect measurement system should be available, to measure the same portion of surface so that the discrepancy of local $z(x_i)$ values with

[3] "Partial" as no assumption has been made about the field being Gaussian, see Chapter 4

[4] A t-distribution given the scarce number of replicates, hence z values in the sample at each x_i location.

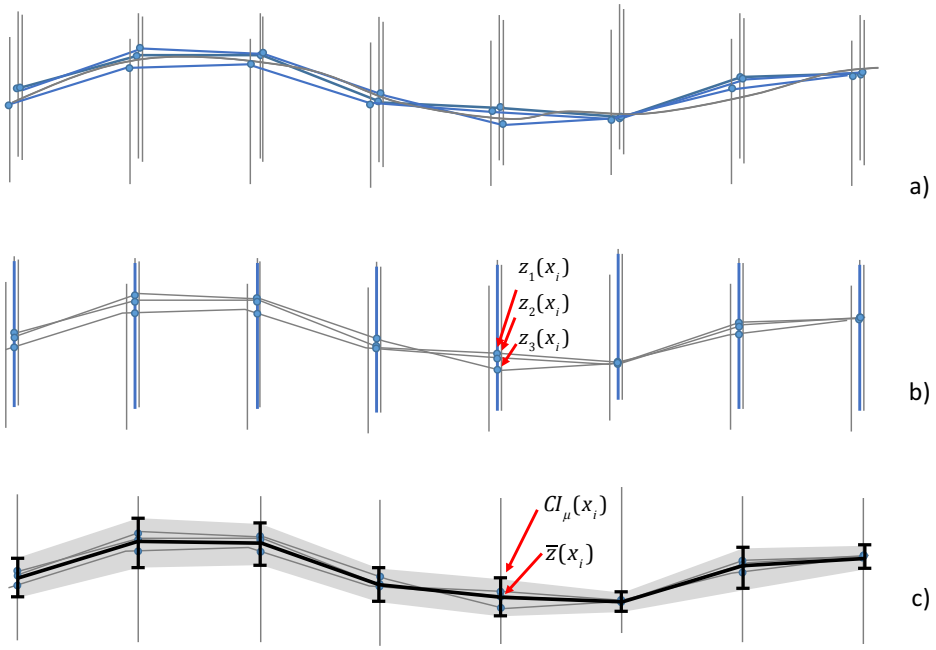
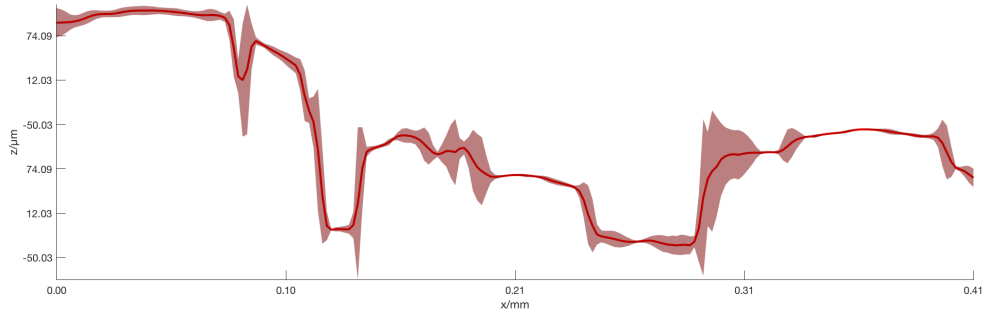


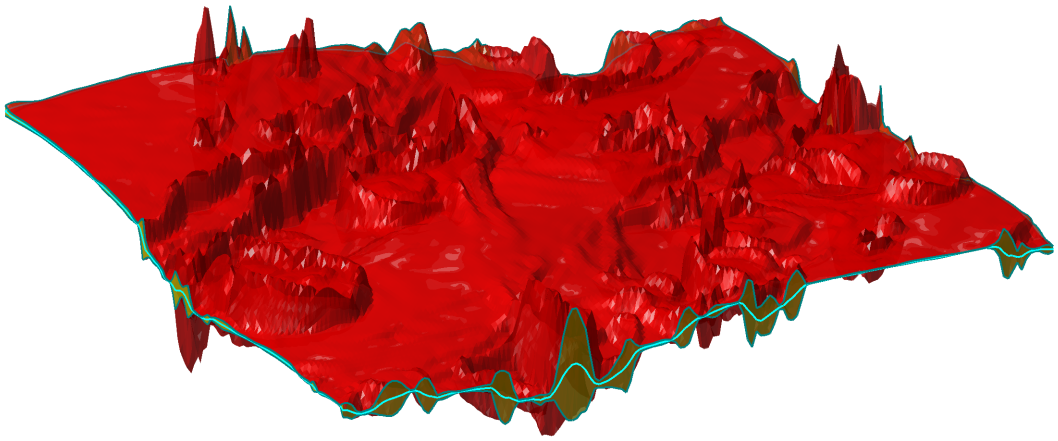
Figure 7.1: Assessement of local repeatability error from replicate surface measurements (visualised in simplified form, as profiles); a) the datasets are relocated in space (this is likely to result in the original measured points not being aligned in x any longer); b) the relocated datasets are resampled at specified locations x_i, so that a new set of aligned, replicate height values are obtained $z(x_i)$; c) the local replicate height values are used to compute the sample mean $\bar{z}(x_i)$, and the confidence interval of the population mean $CI_\mu(x_i)$. The latter can be used as an estimate of repeatability error.

respect to the ideal result $z_{true}(x_i)$ could be used as an estimate of local bias and incorporated into an estimate of measurement uncertainty together with the estimate of local scatter between the $z(x_i)$ points. In reality, given the lack of a perfect measurement solution, in principle one could rely at least on a higher-class measurement instrument acting as *reference instrument*, meaning, an instrument with significantly higher accuracy and precision than the one subjected to evaluation, so that its own error can be neglected[5]. If such a reference instrument could be found, then local bias could be assessed by expanding the previously illustrated method as shown in Figure 7.3. In the figure, the mean profiles and associated confidence intervals are shown for two measurement solutions. For sake of clarity, the hypothetical "real" profile is shown also overlaid, to indicate that one of the two measurement solutions has a much greater accuracy so that it can act as reference $z_{ref}(x_i)$ for computing local bias. Local bias is then simply estimated as the distance between the mean of the replicates of the instrument under evaluation, and the mean of the reference. In the simplification, the local disper-

[5]As further alternative, a higher-class instrument with known uncertainty could be used.



a)



b)

Figure 7.2: Local repeatability error in focus variation measurement of metal additive surface; a) upper, mean and lower boundary shown as profiles; b) uppen, mean and lower boundry surfaces.

sion of the reference is neglected, assuming that it is orders of magnitude smaller: this leads to considering the sample mean of the reference $\bar{z}_{ref}(x_i)$ as coincident with the population mean of the reference $\mu(x_i)$ and also coincident with the true value $z_{true}(x_i)$ because of the previous assumption of lack of bias in the reference measurement. It should be noted also that, if $\bar{z}_{ref}(x_i)$ is assumed as computed on a large sample, then local bias can be seen as an indication of *trueness* (see [92]).

It should be noted that the method, despite being capable of capturing local bias (as long as the reference measurement is a reliable approximation of the true measurement), would still not be capable of detecting a *constant bias* component, if present (i.e. a constant z-shift with respect to the reference, along the entire extents of the profile (or surface). This is because the profile/surface co-localisation procedure, being based on the global minimisation of the distances between the profiles (including the refer-

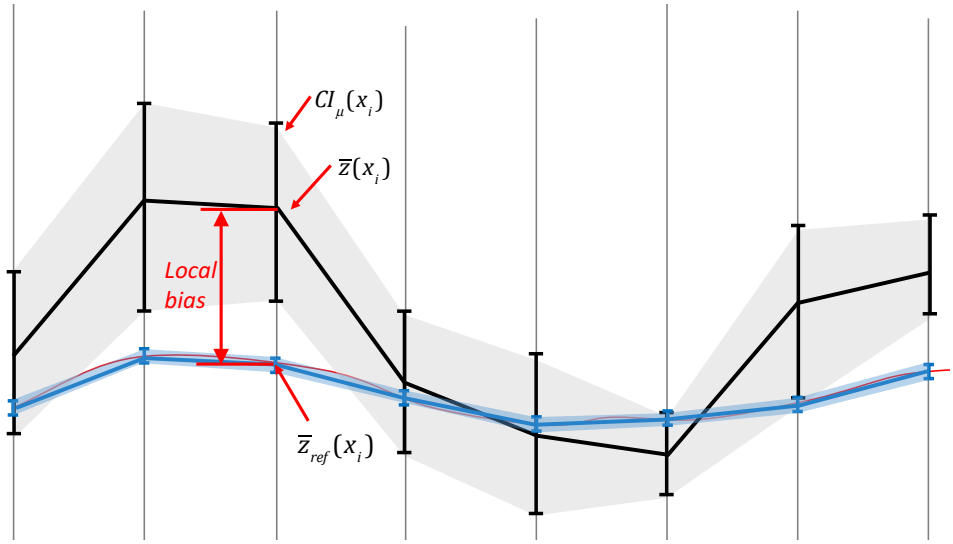


Figure 7.3: Estimation of local bias using a reference measurement instrument (in blue). The proportions are exaggerated. The true measurement is indicated in red underneath the more accurate ($z_{ref}(x_i)$) measurement.

ence as well), would end up absorbing any constant z shift in the minimisation process. What the method is good for, is the detection of interesting, non-constant bias-related behaviour, for example in correspondence of specific features, as illustrated in the following. Aside from that, both the reference and the measurement under assessment must be assumed as "calibrated", i.e. deprived of any constant bias term.

In relation to actually using the method discussed above in real-life applications, the main issue is that it is currently unclear whether there exists any measurement instrument, or more in general, any measurement technology, that is incontrovertibly superior to the others in terms of measurement performance, in any type of application, so that it could act as (accurate) reference. Comparison studies by the author of this thesis [58, 59] clearly show that, even within a very specific domain such as the measurement of metal, additive surfaces, it is difficult to pinpoint a measurement solution that outperforms the others, and rather all solutions show individual strengths and weaknesses in relation to specific types of topography features encountered during measurement (in [58] and [59] measurement solutions based on focus variation, confocal microscopy, coherence scanning interferometry and X-ray computed tomography have been compared).

So, whilst the actual applicability of the method as a means to estimate uncertainty is fundamentally tied to the availability of a "true" reference, at least the same method can be used to compare the metrological performance of surface measurement instruments in relative terms. Going back to Figure 7.3, any region on the x axis where the confidence intervals ($CI_\mu(x_i)$) of the two measurements do not overlap are regions where

the hypothesis concerning two measurements having the same mean (null hypothesis H_0) can be safely rejected[6]. In layman terms, for such regions it can be said with confidence that the measurements differ, given the available observations. Measurements may differ also in the location x where the CIs overlap, however, this cannot be stated with enough confidence, with the current set of observations (not enough evidence to reject the null hypothesis). Such type of comparison based on CI interval overlapping, basically a test of hypothesis repeated at each x_i location, can provide a percentage of profile (or surface) points where it can be stated that two measurement solutions differ, given the current set of observations for both. The idea is illustrated in Figure 7.4 where the same plots used for Figure 7.3 are reported, but the regions where discrepancy can be asserted with confidence are highlighted.

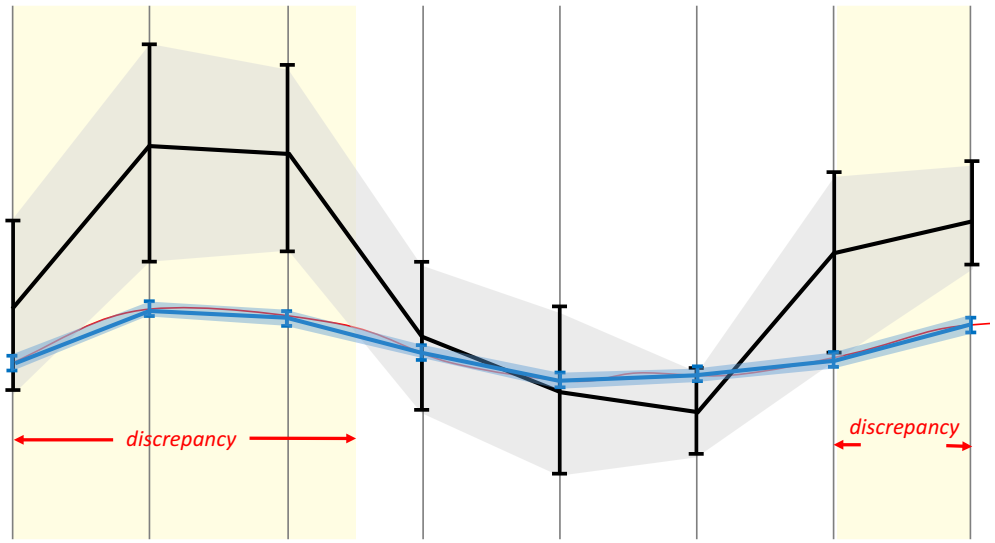


Figure 7.4: Regions where the assertion of difference between the mean z measurements can be stated with confidence

Uses of the illustrated method that go beyond the idea of estimating uncertainties, range from the comparison of measurement technologies, to the comparison of measurement set-ups for the same technology. Two examples are provided in Figure 7.5, from work by the author of this thesis ([58] and [57]).

7.4 Experimental evaluation of repeatability and reproducibility error in feature-based dimensional assessment

A simple experimental approach has been developed to associate repeatability and reproducibility error to individual dimensions of topography features. The method has

[6] at the confidence level set when constructing the CIs, e.g. 95%

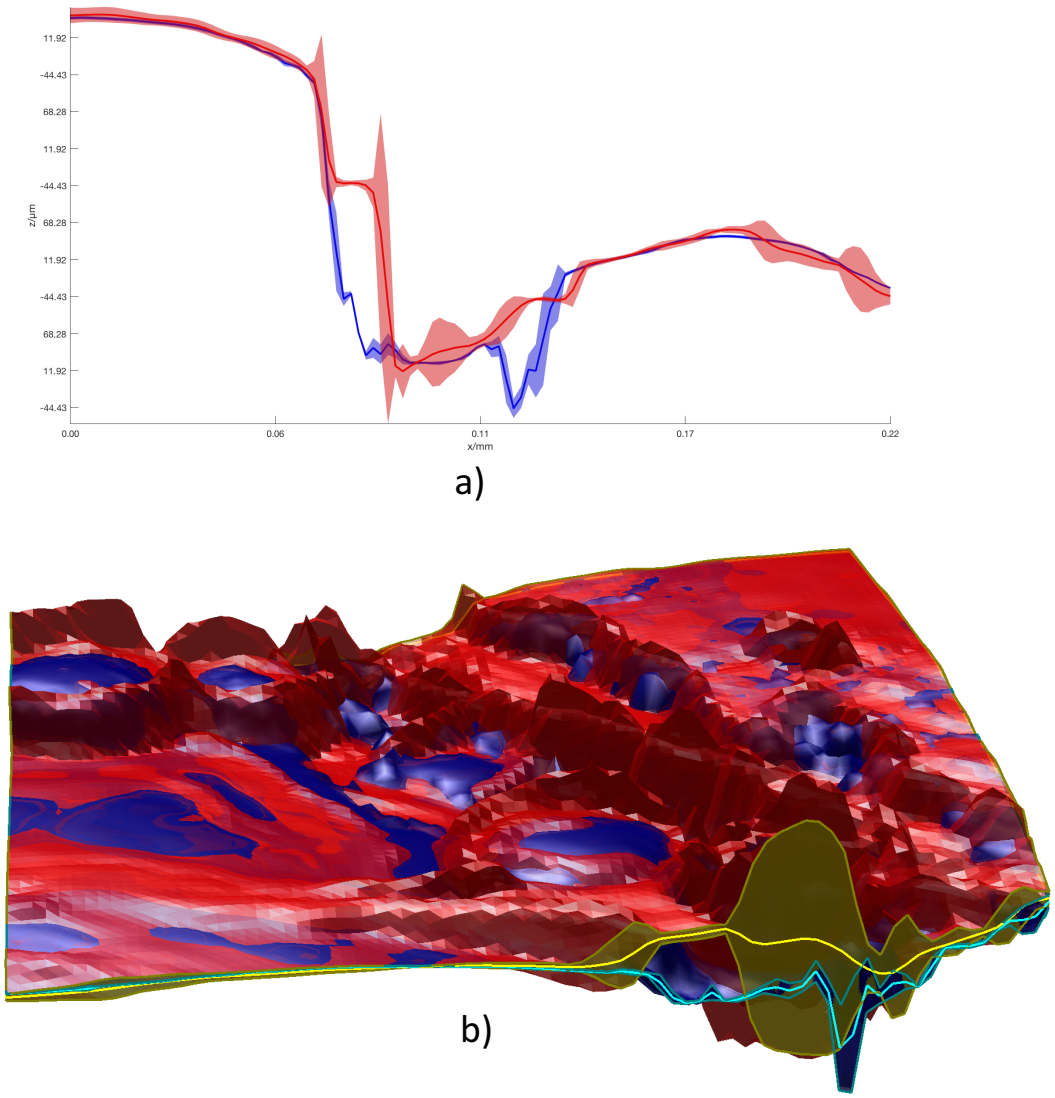


Figure 7.5: Focus variation and coherence scanning interferometry results from measurements on the same region of metal additive surface; a) confidence intervals on profiles; b) confidence intervals on surfaces.

been published in [106] by the author of this thesis.

As for the previous method, because of the lack of truer references, the method can only estimate scatter of the measurement result as a consequence of replicate measurements. It does not incorporate any solution for estimating constant bias or any other systematic components referred to a truer reference, thus it can only provide a portion of what would be needed to estimate an actual uncertainty. The method consists of performing replicate measurements of the same sample. Depending on the elements that are allowed to vary between replicate measurements, a number of scenarios ranging from simpler repeatability to increasingly more general forms of reproducibility can be achieved. For example, the same sample can placed in multiple positions and

orientations under the instrument, but aside from placement, it may be subjected to acquisition with the same operator, protocol, and instrument set-up. As an alternative, laboratory, instrument, operator and protocol may be allowed to change. In [106], the same algorithmic pipeline is applied to all the measurement datasets, consisting of topography data preprocessing (noise removal, levelling, filtering, depending on the application), feature identification and extraction, and finally, computation of the target dimensional attribute of the feature, e.g. a depth, width, distance, area or volume, or any other interesting property of the feature.
Interestingly, whilst it is relatively common practice to think of multiple scenarios covering repeatability and reproducibility for surface topography "measurement", it appears more challenging to find a way to consider similar conditions from the side of data processing. Reproducibility in data processing would likely imply considering a number of changes ranging from the variation of algorithm control parameters, to the implementation of different computational solutions for obtaining the same attribute of the topography feature.
The result of the replicate application of the pipeline is a sample of scalar values for the feature attribute, whose scatter behaviour can provide indication of associated repeatability and reproducibility error (depending on the conditions for the replicate experiments).
As discussed in [106], an interesting type of experiment where the above approach can be applied consists of performing K repeated measurements on the sample, without any change of operator, instrument set-up or protocol (the simplest repeatability conditions), for each one of H pre-set positions (i.e. *poses*) of the sample under the measurement probe (a pose consists of the combination of a lateral and an angular displacement). Measurements at different poses would be considered as taken in reproducibility conditions.

Assuming y as the feature attribute (e.g. a height, or a footprint area) a series of values $y_{h,k}$ is available where $h = 1,2,\cdots,H$ and $k = 1,2,\cdots,K$, being H the number of poses, and K the number of replicates per pose. The error in repeatability conditions $E_{y,repeat}$ can be obtained as the mean across the H poses of the population standard deviations $\hat{\sigma}_h$ estimated at each pose. That is:

$$\hat{\sigma}_h = \sqrt{\frac{\sum_{k=1}^{K} \left(y_{h,k} - \bar{y}_h\right)^2}{1-K}}, \tag{7.4.1}$$

$$E_{y,repeat} = \frac{\sum_{h=1}^{H} \hat{\sigma}_h}{H} \cdot \frac{1}{c_4(H)}. \tag{7.4.2}$$

Where the mean is corrected by the unbiasing constant $c_4(H)$ to compensate for estimation bias due to H not being infinite[106]. The error in reproducibility conditions $E_{y,reprod}$ is computed by rearranging the available $H \times K$ observations of y. K new sets are created, each obtained by randomly drawing one replicate from those available at each pose, with no repetitions. Assuming the k^th new set as:

$$Y_k = \{y_{1,u(k,1)}, y_{2,u(k,2)}, \cdots, y_{h,u(k,h)}, \cdots, y_{H,u(k,H)}\} \tag{7.4.3}$$

where $u(k,h)$ is a random non-repeating index with values in the interval $\{1,2,\cdots,K\}$ (i.e. a random permutation), drawing from the h^{th} dataset, then, being $\bar{y}_k$ the sample mean of such new set, the population standard deviation is estimated as:

$$\hat{\sigma}_k = \sqrt{\frac{\sum_{h=1}^{H}\left(y_{h,u(k,h)} - \bar{y}_k\right)^2}{1-H}} \tag{7.4.4}$$

and the error in reproducibility conditions is obtained as the mean of the population standard deviations, again corrected by the unbiasing constant c_4:

$$E_{y,reprod} = \frac{\sum_{k=1}^{K}\hat{\sigma}_k}{K} \cdot \frac{1}{c_4(K)}. \tag{7.4.5}$$

The approach allows to associate random error components to a result of a feature-based characterisation pipeline by taking into account reproducibility and repeatability error at the measurement stage, but without producing a model of how the error is associated to the surface topography datasets, as instead illustrated in the method discussed in Section 7.3. As stated earlier though, in absence of a truer reference no bias can be estimated, meaning that the method discussed in this section could be used to estimate uncertainty only in case of demonstrated umbiasedness of the measurement process.

An example application of the method is shown in Figure 7.6, where repeatability and reproducibility error is associated to height, footprint area and lateral size of a test feature artefact fabricated by material jetting (from previous work by the author of this thesis [106]).

7.5 Summary and outlook

The construction of an uncertainty budget to associate to the results of a surface topography characterisation pipeline, whether it targets the computation of texture parameters, or the computation of feature attributes (i.e. a feature-based characterisation

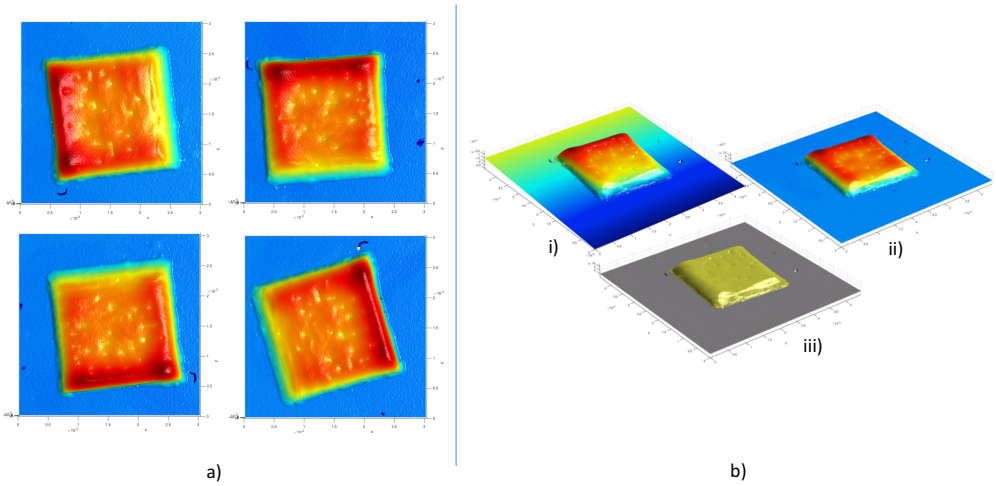


Figure 7.6: Example application of the experimental method for the evaluation of repeatability and reproducibility error associated to a feature-based characterisation pipeline (test case courtesy of Amanda Hûesler and Ricky Wildman, University of Nottingham, from [106]); a) sample measurements at different angular orientations and lateral displacements; b) steps of feature identification in the feature-based characterisation pipeline: i) original dataset; ii) preliminary segmentation and selective levelling on the feature surroundings; iii) final segmentation and separation of the feature (in yellow).

pipeline) remains one of the most challenging, open issues in surface metrology. Complex datasets are handled all throughout the pipeline, even more so in a feature-based approach given the additional data structures related to partitioning, and representation of individual features, and many error sources should be investigated and incorporated in statistical models that are likely far more complex than what shown in this chapter. The original methods proposed in this thesis are primarily based on obtaining information pertaining random error components through repeated experiments. However, the reach of such a class of approaches is limited by the generalisability of results which in turn clearly depends on the intrinsic limitations introduced within the experimentation. In addition, the means of estimation of systematic components are limited, since in principle there is the need to rely on more accurate reference measurements, and these in turn are a real challenge to identify. It is likely that more successful attempts to the estimation of uncertainty may stem from a thorough decomposition of the surface topography characterisation pipeline, and a more systematic investigation of all the potential error sources.

Chapter 8

Generation of simulated topographies

8.1 The role and importance of generating topographies

The term *generation of simulated topographies* refers in this thesis to the act of generating artificial topography datasets by means of algorithmic procedures. The generation of simulated topographies plays a fundamental role in the development of novel methods for surface topography data analysis and processing, as primarily it allows testing of novel methods onto a much larger pool of datasets. Simulated topography is much quicker and much cheaper to obtain than real, measured topography; a wider array of situation and cases can be covered thanks to simulation.

In addition, simulated topographies give the opportunity for generating test datasets of known properties. For example, one may generate a simulated topography matching a very specific value of the areal texture parameter *Sa*, so that it could be used to validate an algorithm for the computation of *Sa*, etc. More complex cases pertain the generation of topographies with known partitions, to validate areal segmentation methods, or with the presence of known feature instances, to validate feature identification, extraction and characterisation methods.

In this chapter, a comprehensive set of methods is illustrated, that can be used to generate a wide array of simulated topography datasets destined to testing and validation of the data analysis and processing methods described in the previous chapters of this thesis.

8.2 Fundamentals of map generation

Goal of the topography generation methods described in this chapter is the generation of *maps* (see Chapter 3 for the definition of a map). In particular, the generation of *height maps* is discussed. Marginally, the discussion will cover the generation of *masks* (i.e. boolean maps) and *classification maps* (both covered in Chapter 3). Analogous methods and ideas could be used to generate *colour maps*, or other types of scalar or vector maps (e.g. quality maps). However, such methods are not discussed here.

All the maps covered in this chapter are scalar maps, therefore generation needs to cover the choice of the pixel spacing/width parameters dx and dy and how to populate the matrix of scalar values, referred to as the z matrix for simplicity (to cover the most common case of it containing height values), assumed of size $n_x \times n_y$. As the setting of dx and dy is trivial once scale is defined for the topography, this chapter will concentrate on how to populate the z matrix.

8.2.1 Layer-based generation

The topography generation pipeline introduced in this chapter considers the simulated topography as comprised of multiple layers, which can be generated independently and combined in multiple ways. Each layer contains a different z matrix of values, as long as the matrices of all the layers have all the same size $n_x \times n_y$. The same spacing dx, dy is automatically assumed for all of them, so that they can be combined.

Combination operators allow for great flexibility in the range of obtainable final results. For scalar, numerical values (e.g. heights) the simplest combination is the algebraic sum. Since the typical operator is considered as applied to all the cells of the matrix independently, the application of the sum operator can be described as $z_{res}(i,j) = z_1(i,j) + z_2(i,j), \forall i,j$. Subtraction, multiplication, division are also possible, as well as non linear operators such as *max* (for example $z_{res}(i,j) = max(z_1(i,j), z_2(i,j))$ or *min* ($z_{res}(i,j) = min(z_1(i,j), z_2(i,j))$). For boolean values (i.e. when combining boolean layers to create masks) the typical set operations such as AND and OR can be used. Integer values such as those used for classification maps must be combined using special logic operators, because integers are used as IDs (identifiers) in classification maps, that is *categories*; thus sum, subtraction, etc. usually have no meaning. A typical logic operator to combine classification maps would create a new identifier for a matrix i,j position, in correspondence to finding specific identifiers in the corresponding i,j cells belonging to the matrices that are being combined.

8.3 Deterministic topographies

Deterministic topographies have been defined and discussed in Chapter 4. The main data processing pipelines that can be used to generate an artificial deterministic topography layer are: interrogating a continuous function of the format $z = f(x,y)$ at specific locations, interrogating a discrete function $z_{i,j} = f(x_i,y_j)$ at the x_i,y_j locations where the function is defined (or by generating interpolated values for the locations where it is not), or finally by resampling an existing geometric model (i.e. a CAD model, in the following assumed as expressed in the form of a triangle mesh[1]).

8.3.1 Deterministic topographies from continuous and discrete functions

The first way to obtain the height map for a deterministic topography is either to interrogate a continuous function of the type: $z = f(x,y)$, or to interrogate a discrete function of the type $z_{i,j} = f(x_i,y_j)$, in the latter case, by using interpolation amongst known values, if the requested location does not coincide with any x_i,y_j pair the discrete function is defined onto. For example, the following set of equations defines a continuous function to model an ellipsoid protrusion in x,y plane, centred in x_c,y_c, with axes a,b and c and β angle with respect to the x-axis.

$$z = f(x,y,x_c,y_c,a,b,c,\beta), \qquad f = \begin{cases} x' = (x - x_c)\cos(\beta) + (y - yc)\sin(\beta) \\ y' = -(y - y_c)\sin(\beta) + (y - yc)\cos(\beta) \\ z' = 1.0 - (x'/a)^2 - (y'/b)^2 \\ z'' = c\sqrt{z'} \\ z = z'' \text{ when } z'' > 0 \\ z = 0 \text{ when } z'' < 0\,. \end{cases} \tag{8.3.1}$$

The following set of equations defines a cylinder in the x,y plane, with x_c,y_c centre, *beta* angle with respect to the x-axis, r radius, infinite length of the cylinder axis, and protruding by h from the $z = 0$ plane:

[1] See Chapter 3 for a discussion on the triangle mesh format.

$$z = f(x,y,x_c,y_c,r,h,\beta), \qquad f = \begin{cases} v_1 = (x - x_c)\sin(\beta) + (y - yc)\cos(\beta) \\ v_2 = 1.0 - (v_1/r)^2 \\ z' = -r + h + r\sqrt{v_2} \\ z = z' \text{ when } z' > 0 \\ z = 0 \text{ when } z' \leqslant 0\,. \end{cases} \tag{8.3.2}$$

The use of multiple layers allows for the generation of complex topographies obtained as combination of much simpler analytic functions. For example, by combining one instance of ellipse and one instance of cylinder, so that $z_{res} = max(z_{ellips}, z_{cylind})$ a result such as what is shown in Figure 8.1 can be obtained.

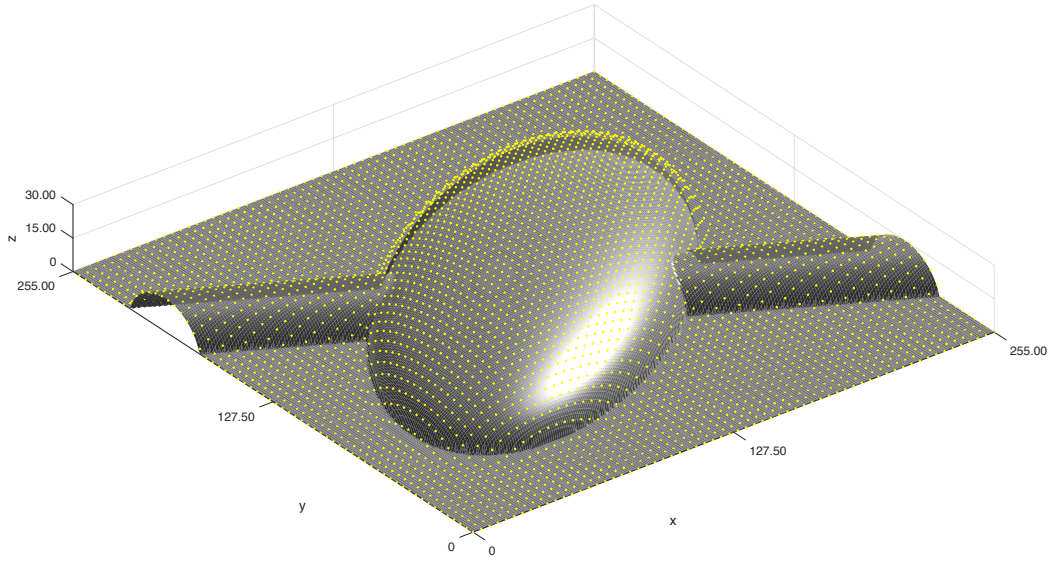


Figure 8.1: Combination of a cylinder and an ellipse instance, both defined from continuous functions; yellow points: result of function interrogation at predefined locations; gray surface: continuous reference built from interpolation of a denser interrogation grid, provided for visualisation purposes.

8.3.2 Deterministic topographies from triangle meshes

Deriving deterministic topography from triangle meshes is the suggested option in particular for those scenarios where topography is easily captured by CAD models (e.g. surfaces of micro-parts, micro-electromechanical systems, etc.). Alternatives to the triangle mesh representation are fairly common in geometric modelling, such as NURBS (non-uniform, rational b-splines). However, as the geometry gets more and more complicated, NURBS and similar representations require the combination of mul-

tiple patches, quickly becoming harder to handle. Triangle meshes have the same representational complexity regardless of the complexity of the model they are approximating[2]. Since any other surface models can be typically exported as a triangle mesh, considering only triangle meshes is a good enough choice for most applications.

In order to obtain a height map from a triangle mesh, *raster scanning* can be used. Raster scanning was already introduced in Chapter 3 and is shown applied to a triangle mesh with undercuts in Figure 8.2.

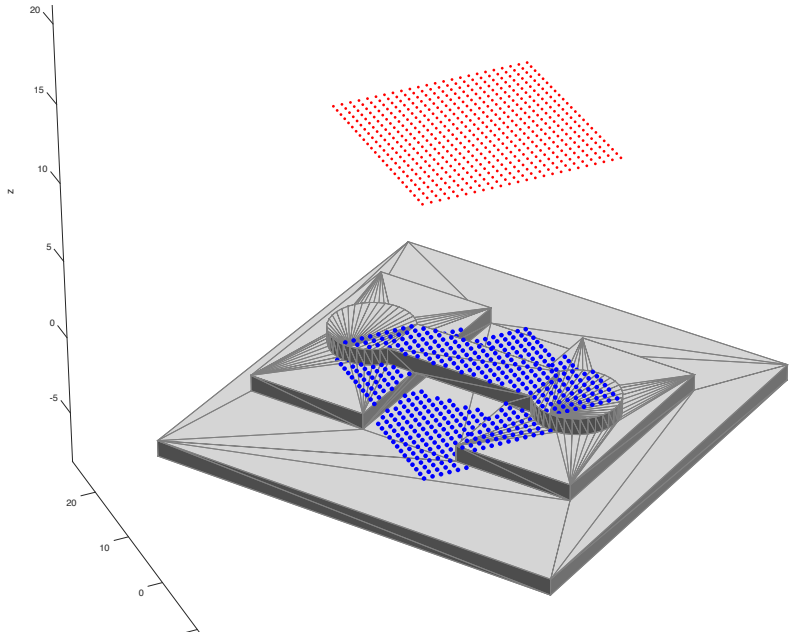


Figure 8.2: Raster scanning applied to example triangle mesh. In red: x,y virtual grid overlay; in grey: triangle mesh; in blue: intersections between the z-rays and the triangle mesh

8.4 Random topographies

This chapter is about topography generation, so *randomness* is here primarily intended from the perspective of generation, not prediction (see Chapter 4 for details on the difference). However, design choices pertaining how to use randomness to generate topographies may give origin to datasets that are more or less deterministic, this time intended from the viewpoint of prediction. To summarise this concept, for example, if only a few height points are determined randomly, and all the other points in between are obtained by interpolation, then the resulting topography may be seen as highly deterministic from a predictability standpoint. Conversely, if all the height points are

[2]approximating, because curved surfaces and freeforms in general cannot be fully captured by piecewise linear representations

determined from a random number generator, supposedly acting independently for each point (see the pure random topographies discussed in the next section), then the generated topography would be seen as highly random, again from a predictability standpoint.

Because the ultimate motivation for topography generation is validation and testing of data analysis and processing methods, reproducibility is essential also for random topographies. Luckily, regardless of how the output of random number generators is used (there are different ways, as illustrated in the following sections) as long as one stores the random seed used to initialise the generators, reproducibility can be ensured.

8.4.1 Pure random topographies

The simplest form of random topography is obtained by assigning a random number to the height value stored in each cell of a $z(x_i, y_j)$ height map. The result is a height map where each height value can be considered as independent of its neighbours[3]. The random values may derive from a uniform distribution (e.g. a 0-1 distribution), which is typically the case for random number generators implemented in most software development platforms, or they may be drawn from different probability distributions. The latter case is typically implemented through *Monte Carlo rejection sampling [107]*: random values are drawn from a uniform probability distribution, and used to populate a sample. The probability distribution of the sample is checked at each iteration. Each newly drawn value is rejected if it moves the sample probability distribution away from the target shape and statistics (e.g. Gaussian with zero mean and given standard deviation), or it is kept if it is consistent with said target distribution. In Figure 8.3 two pure random profiles are shown, generated with random values extracted from a uniform and Gaussian distribution respectively. Pure random topography layers with Gaussian random values are useful for example to simulate measurement repeatability error in single-point height measurement, assuming the error associated to each measured point is independent from the error associated to any of its neighbouring points, and assuming (rather simplistically) normal distribution of the error.

8.4.2 Interpolated random topographies

Interpolated random topographies are obtained by generating random height values only at specific locations (typically, uniformly spaced), and then by interpolating the

[3] in truth, not strictly independent, if the random values are extracted in sequence from the same random number generator, but can be assumed as such in first approximation.

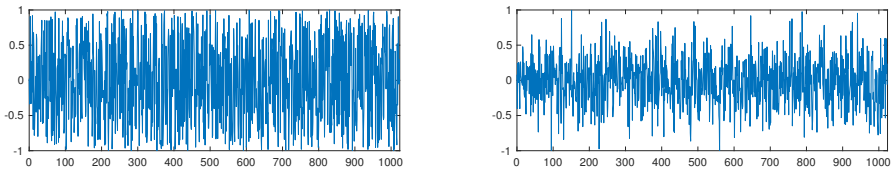


Figure 8.3: Pure random topographies; left) values drawn from a uniform distribution in the 0-1 interval; right) values drawn from a Gaussian distribution with Monte Carlo rejection sampling (mean = 0.0, sigma = 0.33).

height values in between. Thus, while pure random topographies are intrinsically discrete geometric entities, interpolated random topographies are continuous in the x,y space by definition, and need to be interrogated at discrete locations in order to finally result into height maps. In an interpolated random topography, the ratio between the z range in random height generation, and x,y spacing of randomly generated height values determines the predictability of the resulting topography, and thus the degree of randomness/determinism in the primary sense it was illustrated in Chapter 4 (e.g. see Figure 8.4). Concerning interpolation, linear of higher-order solutions can be used, cubic interpolation being preferred as the lowest interpolation order ensuring a smooth result even with local slopes changing sign.

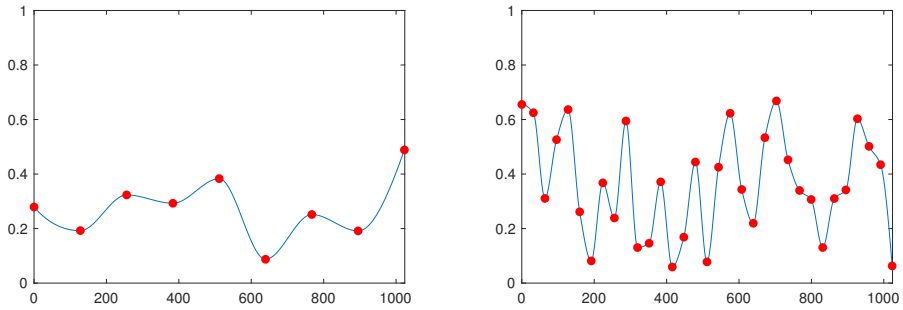


Figure 8.4: Interpolated random topography profiles; left) cubic interpolation, large spacing between random points and small range of random height values (resulting in lower randomness in prediction); right) smaller spacing between random points, larger range of random height values (resulting in higher randomness in prediction). The random points are highlighted in red.

8.4.3 Multi-scale interpolated random topographies

Multiple interpolated random topography layers can be of course combined in many different ways. When the layers are generated with different spacing between random height values, and different amplifications of the random height range, their summation leads to what is here referred to as *multi-scale interpolated random topographies*; scale referring both to spacing and amplitude of the summed oscillation components. The

most common arrangement is when spacing is halved at each new component (dyadic decomposition of x,y space, each element referred to as an *octave*). A wide range of topographies is then obtained by simply changing the way amplitude decays with octaves. Amplitude decay can be captured by *persistence*, i.e. a multiplying coefficient p so that $Ampl_{nextoctave} = p \times Ampl_{curroctave}$. A couple of example multiscale topography profiles are shown in Figure 8.5, featuring the same dyadic decomposition into eight octaves, but different persistences. A lower persistence value generally leads to smaller-scale noise decaying more rapidly (i.e. influencing less the final result).

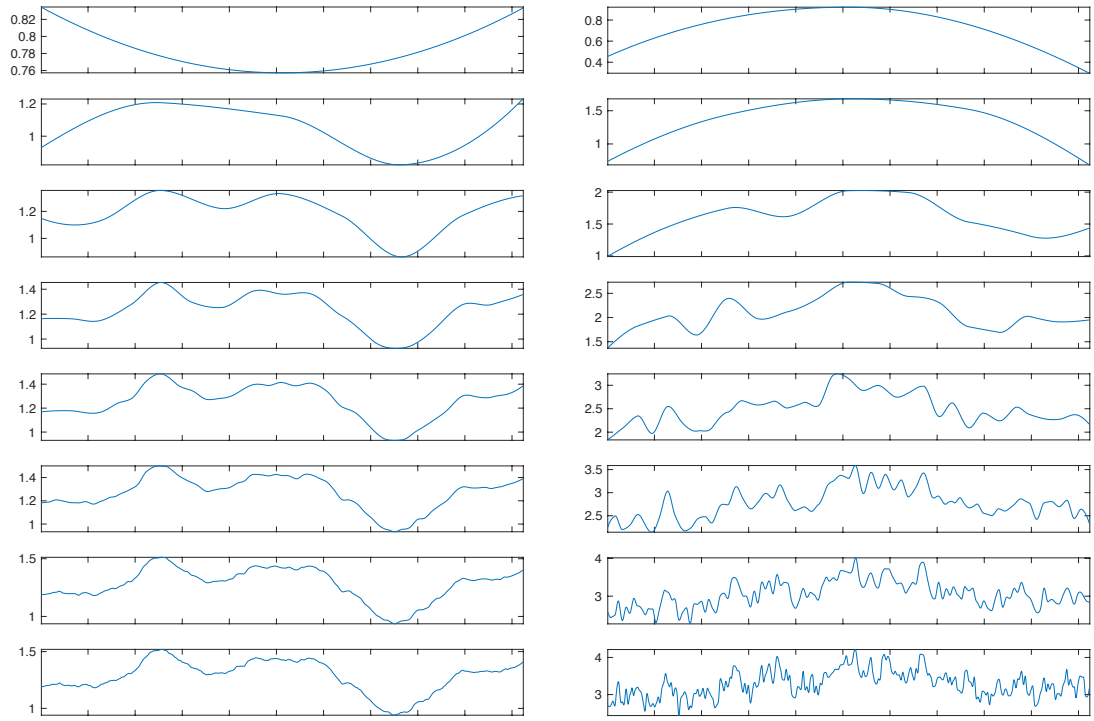


Figure 8.5: Multi-scale, interpolated random profiles with same dyadic decomposition in eight octaves, but different persistences; left column: $p = 0.5$; right column: $p = 0.9$. From top to bottom: each row shows the addition of a new octave.

In Figure 8.6 an equivalent example is provided using surfaces (i.e. $z = f(x,y)$ functions), with persistence $p = 0.5$, and eight octaves.

8.4.4 Parametric topographies with random parameters

Multi-scale interpolated random functions may not necessarily be mapped to heights directly, they could also be used to feed parameter values to parametric topographies. For example, a deterministic sine-wave profile can be seen as a parametric function $z = Asin(\omega x + \phi)$ where the parameters are the amplitude A, wavelength ω and phase

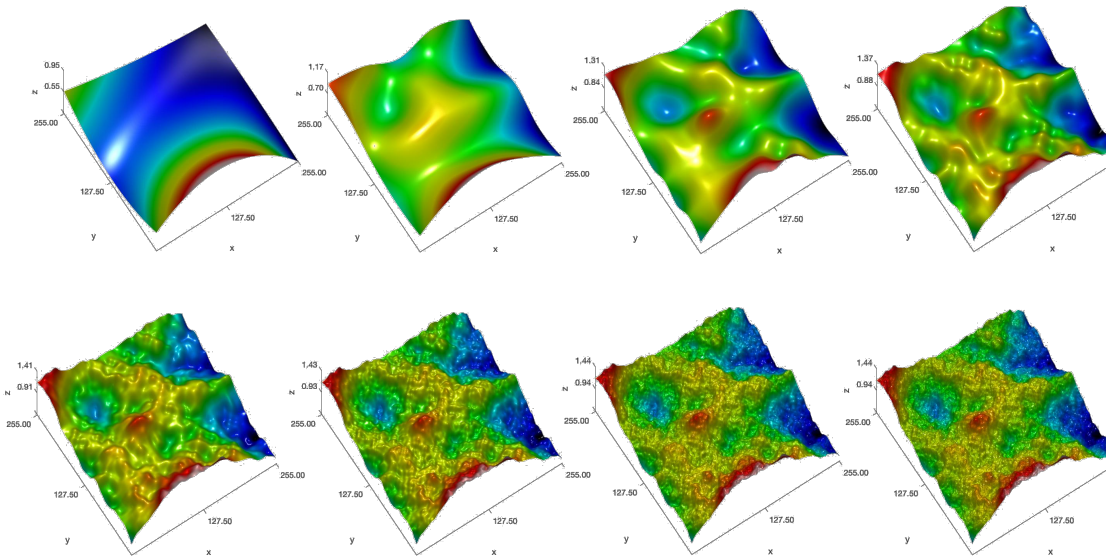


Figure 8.6: Multi-scale, interpolated random surface with eight octaves and $p = 0.5$ persistence. Each image shows the addition of a new octave (from left to right, from top to bottom).

ϕ, and are usually intended as constants. However, A, ω and ϕ may as well be random, and even continuously changing in space following a multi-scale interpolated random function such as those illustrated in section 8.4.3. An example of this approach is illustrated in Figure 8.7, where a sine wave 2D parametric function has been disturbed by defining the sine amplitude parameter as an interpolated random function in x, y.

8.5 Matching topography specifications

One of the goals of generating simulated topographies is to create topographies that match given specifications, for example, a target value for a texture parameter. Combined with the generation tools illustrated above, the two approaches that have been investigated are Monte Carlo rejection sampling and genetic optimisation.

8.5.1 Monte Carlo rejection sampling

The approach of Monte Carlo rejection sampling can be applied to pure random topographies, i.e. those ascribed to the classes illustrated in Section 8.4.1. In such case, every topography value $z_{i,j} = f(x_i, y_j)$ can be assumed as drawn from the same random variable, whose probability density function is what we want to control in order for the resulting topography to match target values with its statistics.

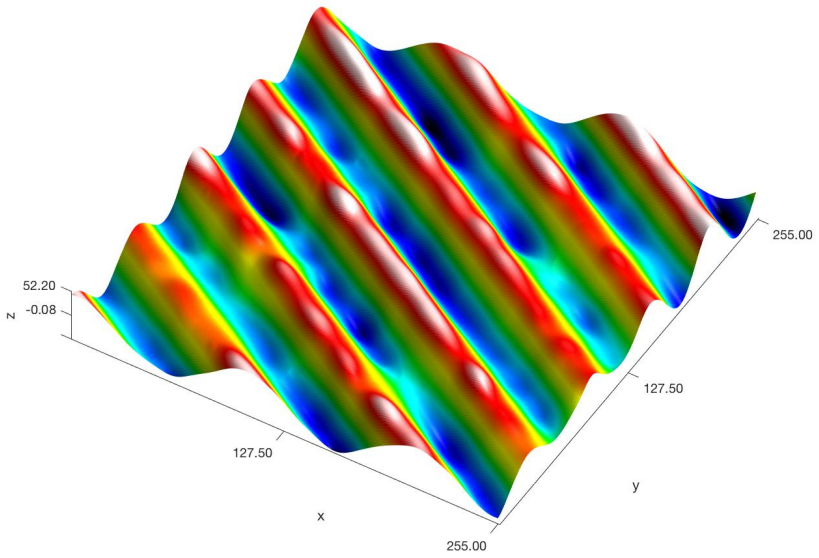


Figure 8.7: Sine-wave surface where the amplitude parameter has been defined as an interpolated random function of x,y (persistence = 0.7, 4 octaves). Wavelength and phase have been left constant.

A simple case of how Monte Carlo rejection sampling works is provided in this section. The goal is to generate a topography that matches a given value for the texture parameter Sq (ISO 25178-2 [6]). The definition of Sq was already provided in Chapter 2, but is reported here for convenience in its discrete form:

$$Sq = \sqrt{\frac{1}{n_x n_y} \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} z(x_i, y_j)^2}. \tag{8.5.1}$$

To implement Monte Carlo rejection sampling, the traversal sequence for the discrete set of $n_x \times n_y$ locations x_i, y_j is randomised first (by randomisation of the i,j indices). Then for each element of the sequence, a value $z_{i,j}$ is extracted from a uniform random distribution. As all the x,y locations are traversed, the polled z values are binned and their frequency distribution checked against an upper limit defined by the shape of the desired probability distribution. In case of a target Sq value to achieve, the upper limit can be modelled as a Gaussian with known standard deviation / variance (since Sq is essentially capturing the second-order moment of the distribution of heights). While the first z polled values will be most likely accepted in any case, as the number of polled

values grows, an increasing amount will be rejected due to exceeding the upper limit. For each rejected value, a new one will need to be extracted until it is accepted. The loop terminates where all i, j locations have been explored, and all $z_{i,j}$ values have been assigned to accepted values. An example application of Monte Carlo rejection sampling is shown in Figure 8.8, where a pure random topography is generated consisting of two regions, with different target Sq values.

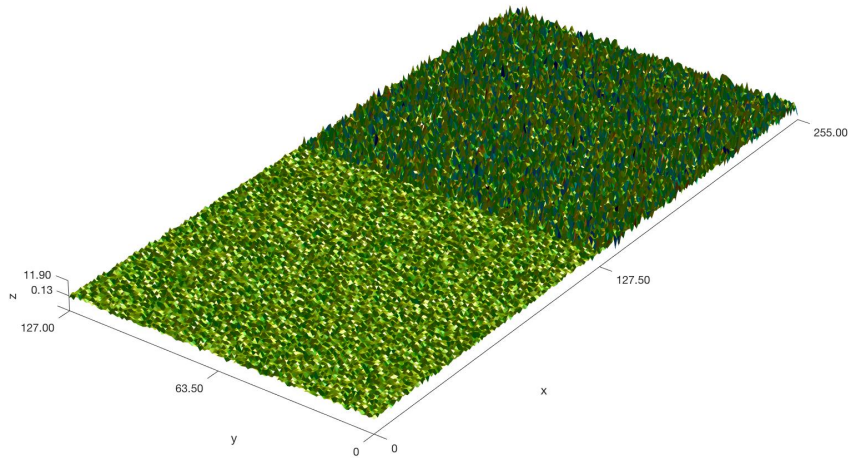


Figure 8.8: Monte Carlo rejection sampling applied to the generation of two regions with different Sq values.

8.5.2 Genetic optimisation

When the topography generation pipeline is more complicated, like in many of the cases illustrated in the previous sections, rejection sampling is harder to apply, mostly because the strategy of polling random values for each x, y location cannot be applied any longer (polling of random values only works if complete independence across x, y locations is assumed).

In such cases, an alternative approach can be pursued, consisting in seeing topography generation as solving an optimisation problem, where the objective function is to obtain a topography whose descriptor values are as close as possible to the specification targets. The first step of such approach consists of creating a pipeline that is actually the concatenation of two processes: the process that leads to topography generation (for example, the combination of multiple deterministic and interpolated random datasets), and the process that leads to the computation of the indicators whose values must be matched by the simulated surface (for example a target Sq value). The second step is

the isolation of a selected set of parameters controlling the generation process, and that are turned into the independent variables of an optimisation problem. The objective function is set as the minimisation of the difference between the current value of the indicators, and the desired target.

The concatenated topography generation/indicator computation pipeline is often significantly complex, and it is hard to solve the optimisation problem through conventional gradient-based methods. In Figure 8.9 an example is illustrated where genetic algorithms have been used. The surface is generated as a multi-scale interpolated random function, with amplitude and persistence selected as the independent variables in the optimisation problem (number of octaves in dyadic decomposition = 8, cubic interpolation). The random seed has been set to a constant value to allow reproducibility. The target is a set value of the Sq parameter, so the fitness function is defined as the absolute difference of the current and target Sq values. Genetic algorithm parameters: number of generations: 50, population size: 100, crossover: single-point, mutation rate: 0.01.

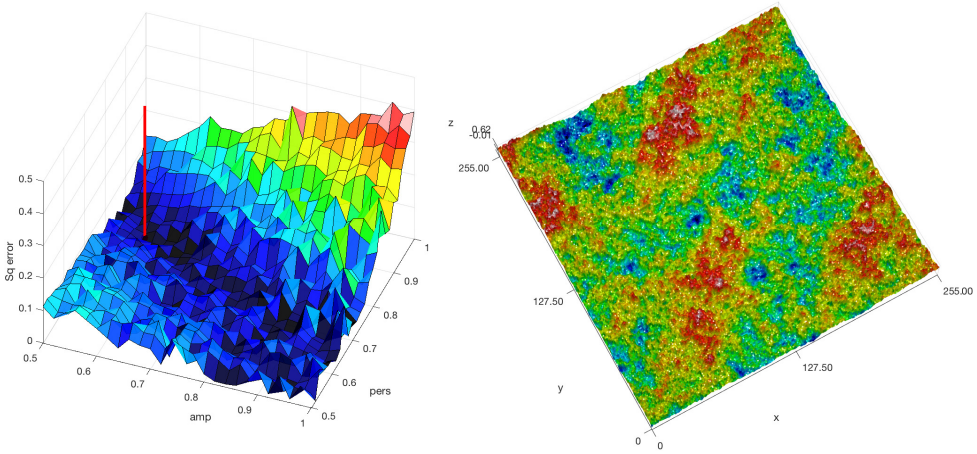


Figure 8.9: Genetic optimisation of a multiscale interpolated random topography with target Sq value; left) fitness function (difference of current and target Sq values) computed by sampling the search space (amplitude and persistence) for visualisation purposes. The optimal value found by the genetic algorithm is highlighted by the red vertical line; right) generated topography corresponding to the optimal amplitude and persistence values.

8.6 Generating partitions

Since a simulated topography dataset can be obtained as a combination of multiple layers, it is possible to generate a mask together with each new layer, that specifies

what x, y coordinates of the layer should contribute to the final combination result. To generate smoother transitions, a weighting map (0-1 real-valued map) can be generated instead of a mask (0-1 binary-valued map). In Figure 8.10 artificial craters simulating erosion are added to a simulated substrate. A mask has been generated together with the craters to indicate the regions affected by the modification.

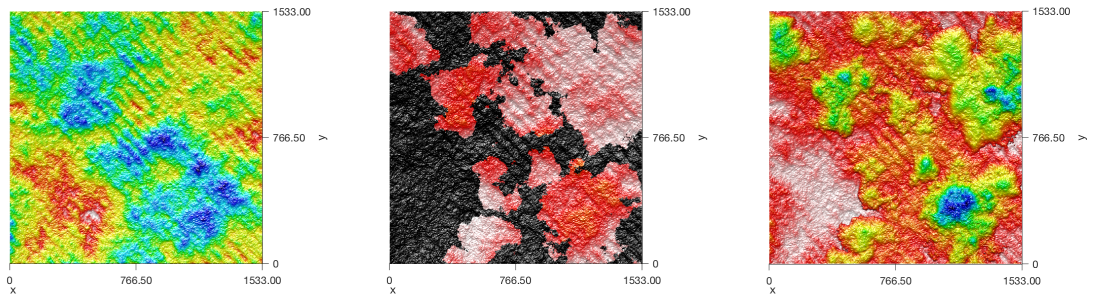


Figure 8.10: Use of masks to specify partitions; left) original topography; b) mask specifying the regions that are affected by erosion-like modifications; c) eroded topography.

8.7 Examples

In this section two examples are provided that summarise the techniques presented in the previous sections of this chapter. The first is about simulating topography modifications due to functional operation, the second is about including measurement error components in the simulation.

8.7.1 Simulating the effects of operational life of a surface

A common task for the generation of artificial topography is reproducing the topographic features that would naturally appear on a surface as a consequence of its working life. Examples are scratches from mechanical contact, crater-wear (e.g. from fretting phenomena), crest erosion and basin filling. Depending on whether the signature features for a surface modification process are of the protruded or the eroded type, generation pipelines can be set that generally consist of modelling the individual feature instances as parametric shapes controlled by random parameters, and further disturbed by higher-frequency noise-like components, for example generated by multi-scale interpolated random functions. For example, in Figure 8.11 an artificial surface featuring machining marks has been modified with the addition of simulated scratches (8.11.left), and then also with the addition of simulated crater-like formations (8.11.right). Scratches are obtained as severely elongated ellipsoids (random length,

depth, position, orientation and number); craters are dale-like features deformed by the addition of several layers of multi-scale interpolated random noise. Layers are combined as: $z(x_i,y_j)_{res} = min(z(x_i,y_j)_{orig}, z(x_i,y_j)_{scratches}, z(x_i,y_j)_{craters})$ to simulate a sequence of topographic overwrites by subtractive processes.

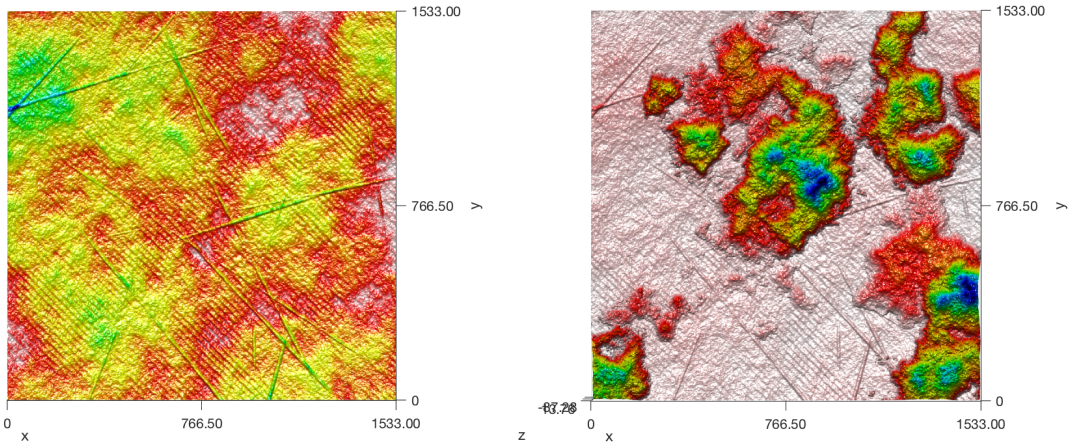


Figure 8.11: Simulation of damage; left) scratches; right) scratches and crater wear.

8.7.2 Simulating measurement

It is often useful to simulate also the behaviour of a specific surface measurement technology/instrument, when applied to an already simulated, artificial surface. Measurement simulation is usually comprised of two steps: in the first, the raster sampling nature of the topography acquisition process is captured by implementing a z-ray scanning process, along the rows and columns of a regular x,y grid. The process has already been illustrated as applied to triangle meshes in Section 8.3.2; in this case bilinear or bicubic interpolation among known values of the previously simulated, discrete $z_{i,j} = f(x_i,y_j)$ function, is used to obtain the values at x,y locations corresponding to the scanning grid. The "deterministic" measurement result (deterministic because it derives from a deterministic interpolation of known z values) can then be disturbed by the introduction of models that describe the instrument behaviour in correspondence to the specific topography point. The simplest model consists of assuming normally distributed random noise in the form $N(0,\sigma)$, independent at each measurement location. More complex models may use local topography information (e.g. measures of local roughness, local slope, local curvature or even colour, if available) to derive a more personalised measure of noise. An example is shown in figure 8.12. A simulated topography is generated first (same model as what previously shown in Figure 8.10, but different random seed). The height map has x,y spacing of $1\mu m$. Then measurement is simulated over a central region of the topography, assuming raster sampling

with 1.5μm point spacing. An independent noise component is finally added to each measured height point, assuming a normal distribution $N(0,\sigma)$ source, multiplied by a 0-1 weighting factor computed from local slope in the original, simulated map (higher slope corresponding to larger noise).

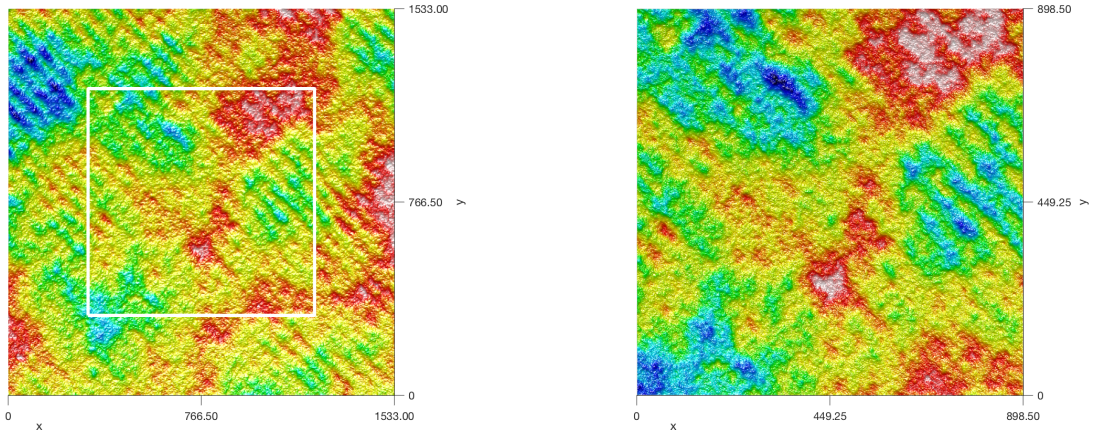


Figure 8.12: Measurement simulation; left) simulated topography as generated, the measured area is shown as a white rectangle; right) measurement result showing the effects of raster sampling and measurement noise)

8.8 Summary and outlook

in this chapter, a toolbox of original methods developed during the doctorate for the generation of artificial topographies of known properties has been presented. The generation of artificial topographies plays a relevant role in feature-based characterisation (and in general in any surface topography characterisation approach) as a fundamental tool for validation and quantative performance assessment of new methods and algorithms. In fact, topography generation allows the creation of test surfaces of known properties, and methods/algorithms can be evaluated in terms of how they achieve a correct assessment of such properties, or in terms of how they react to them. In addition to serving as a tool for data analysis/processing validation, the generation of artificial topographies may also indirectly support the investigation of measurement error sources, through their estimated effects on the reconstructed topographies. The quality and information content of an artificial topography is determined by the amount of work that goes into generating it, meaning that very simple and inaccurate topographies may be generated, as well as very complex, detailed and information-rich ones, depending on the efforts at identifying and modelling the components that contribute to creating the final result. One of the most critical aspects is the identification and incorporation of proper measurement error models, capable of accurately simulating the

artefacts that are typical of each measurement technology, for example, the generation of artificial batwing formations in simulated coherence scanning interferometry, or the generation of non-measured points in conditions of scarce amounts of reflected light (or on the contrary, because of saturation due to excessive reflection) in optical surface topography measurement technologies based on focus variation. Models capable of predicting how each measurement technology may react in correspondence to specific topographic formations and material properties are being studied, for example,by means of the statistical modelling methods introduced in this thesis and discussed in Section 7.3.

A better understanding of how to generate artificial topographies that match specification requirements is needed. While at the moment it is possible to generate artificial topographies that match target texture parameter values, a general method to generate topographies which conform to specifications describing the actual shape and distribution of the features that should be present on the surface may be useful. Whilst this is already partially possible with the original tools presented in this thesis (i.e. one can create a scratch, a cavity, a ridge, or even a structured (i.e. deterministic) feature or pattern, a more general approach, which for example allows better control on accepted and rejected levels of shape/size variability for each feature instance would be useful.

Chapter 9

Applications

A series of applications is illustrated in this chapter, covering research themes in the fields of advanced manufacturing, product and material development. The cases have been selected amongst those to which feature-based characterisation solutions have been applied. Most of the test cases presented here were investigated during the years 2016-2018, in collaboration with researchers at the University of Nottingham. Those that have generated publications are accompanied by the proper citations. In some circumstances, images from the test cases discussed here were also used in previous chapters to illustrate specific challenges or aspects related to the feature-based approach to the investigation of surface topography, for example to illustrate feature characterisation approaches (in Chapter 5), feature identification and extraction solutions (in Chapter 6), or the assessment of measurement error and uncertainty (in Chapter 7).

The general aspects, challenges and goals of the scenarios the selected set of applications refer to, have been already introduced and discussed in Chapter 1. Thus, in this chapter, only basic information specifically related to each specific test case will be provided.

Most importantly, all the examples illustrated in the following sections contribute to showing the importance and potential of moving beyond the characterisation of surface roughness, targeting surface topography elements more directly related to the element of interest within the specific research application.

9.1 Research in advanced manufacturing processes

Feature-based characterisation of surface topography can provide useful information to researchers working at the understanding and optimisation of novel, advanced man-

ufacturing processes. During research on manufacturing processes, information may be extracted from multiple sources. For example, from *artefact specimens*, i.e. physical samples whose geometry has been designed to allow the investigation of specific properties and behaviour of the manufacturing process being investigated (e.g. process capability, sensitivity). The examination of entire surfaces may also be pursued with the intent to extract a manufacturing process *signature* or *fingerprint*, that is, a series of selected topographic properties that are typical of the process, and allow to reverse-engineer the physical phenomena that happened during surface fabrication. Finally, feature-based characterisation can be used to compare manufacturing process configurations/set-ups, compare surfaces before/after the execution of surface modification processes (e.g. finishing operations) or to focus on understanding how performance of the process and results may vary in correspondence to different, nominal product surface geometries.

Inkjet artefacts

The measurement of the geometric and dimensional properties of *manufacturing artefacts*, physical samples specifically fabricated to highlight aspects of the manufacturing process under investigation, plays a fundamental role in process understanding and optimisation. Datasets obtained from focus variation measurement of specimens intended to study the behaviour and performance of an advanced additive process consisting of inkjet printing and infrared sintering are shown in Figure 9.1. The samples are part of a research project by members of the Additive Manufacturing Research Group at the University of Nottingham. Details on the project can be found in a recent publication [70] to which the author of this thesis has participated.

Feature-based characterisation can be used to assess geometric and dimensional properties of specimen artefacts such as those shown in Figure 9.1. For the specific application, the variable of interest was thickness (height of the artefact with respect to the plane it lays onto).

As opposed to taking a few measurements interactively using surface metrology software that allows for computing distance between two surface points selected on screen (the conventional procedure implemented in commercial surface metrology software, as illustrated in Chapter 2, Section 2.5.1), the advantage of feature-based characterisation is two-fold: on one hand, one can afford to execute many more measurements than when proceeding manually, for example by generating many cross-sections of the artefact and measuring thickness onto each one. Secondly, results obtained algorithmically are more reproducible than anything produced manually, i.e. by an operator

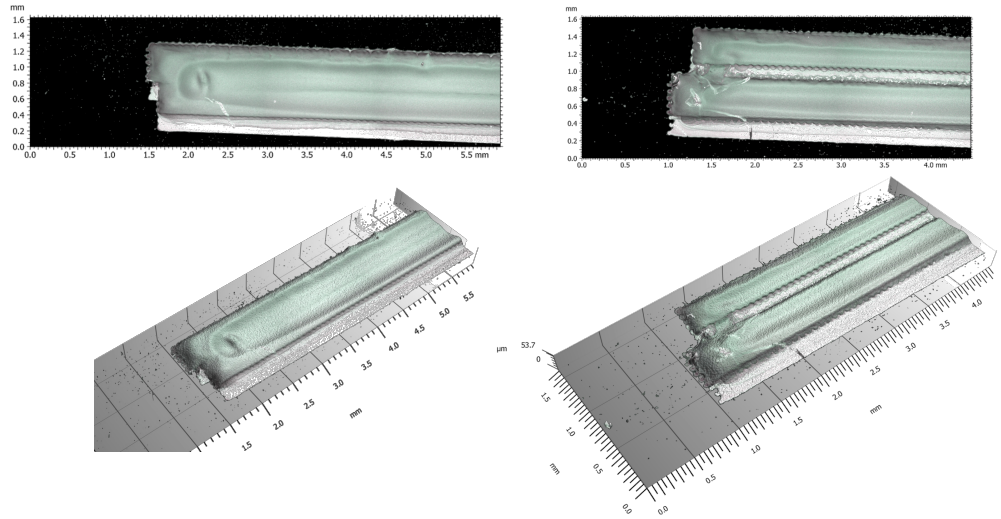


Figure 9.1: Example datasets obtained by focus variation measurement of inkjet artefacts (10×, stitched mode, polarised light). Rendering generated via MountainsMap software [87]); top) RGB colour maps; bottom) height maps with RGB overlays.

following an interactive procedure. The adopted procedure for the automated identification of the inkjet artefact and the definition of a $z = 0$ datum plane is shown in Figure 9.2. Since the artefact is manufactured on a transparent support, knowing that focus variation cannot resolve height on transparent regions because of insufficient contrast information[1], a first segmentation is performed by isolating topologically disconnected regions containing valid height points (i.e. the complement of *non-measured points*, see Section 3.7 for an illustration and discussion on validity masks). The largest region, corresponding to the artefact, is temporarily removed whilst the smaller ones (dust, spattered material, other singularities) are taken as references for plane-fitting to create the $z = 0$ datum[2]. The largest segmented region (the artefact) can finally be levelled with respect to the datum so that its thickness can be analysed with respect to the established $z = 0$ reference.

Thickness is computed by identifying another reference datum to act as the top surface, so that thickness computation can be implemented similarly to step-height evaluation, as prescribed in the standard ISO 5436-1 [39]. As discussed in Section 5.3, there are several options to select the top datum. Alternative choices for inkjetted artefacts were briefly illustrated in Figure 5.5. In Figure 9.3 two alternatives to compute thickness are illustrated in more detail. In Figure 9.3.a, the plateau region on the top surface of the artefact is identified by plane fitting a selected subset of surface points (high-

[1] in the implementation available for this project

[2] In this application, only the smallest singularities are used, to make the datum as close as possible to the actual, non-measurable, transparent support

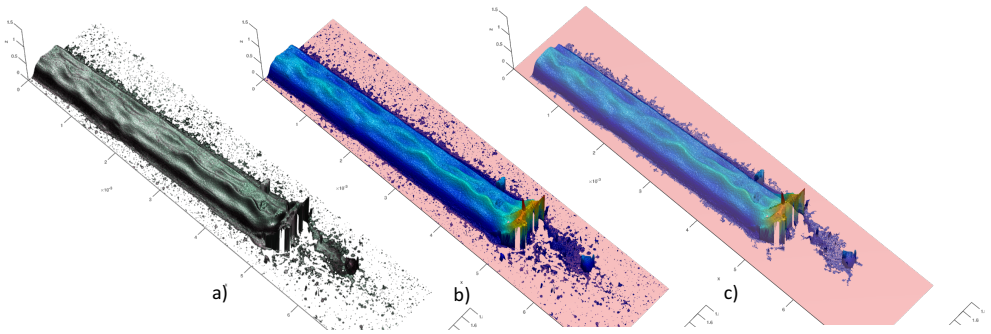


Figure 9.2: Identification of the feature and definition of the $z = 0$ datum plane; a) original height map with RGB colour overlay (under polarised illumination); b) $z = 0$ datum plane fitted to the small material regions surrounding the artefacts; c) identified and extracted artefact feature shown on the $z = 0$ datum plane.

lighted in yellow), in turn found by isolating a subset of x, y coordinates by erosion of the feature footprint. The mean distance between the fitted plane and the $z = 0$ datum (mean because the two planes may not necessarily be parallel) is used as a global measure of artefact thickness (see again Figure 9.3.a). In Figure In Figure 9.3.b, multiple cross-sections of the artefact are computed by slicing the feature orthogonally to its main direction (main direction computed via image moments, see Section 5.3.2, slicing discussed in Section 5.3.1), then, the height of a rectangle of equal width as the cross-section profile and equal area, is used as a measure of local thickness (see again Figure 9.3.b).

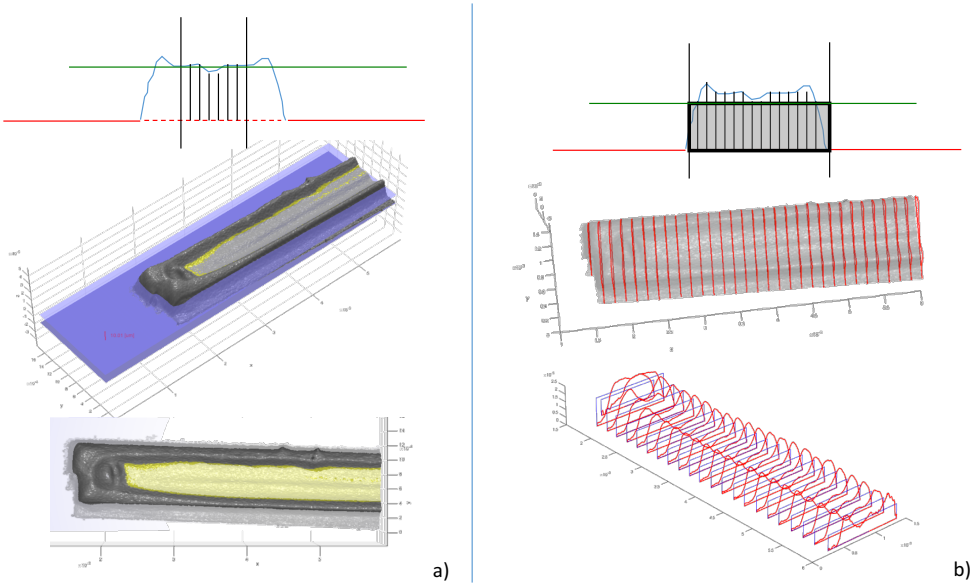


Figure 9.3: Two methods for artefact thickness computation; a) mean distance between $z = 0$ datum and datum plane fitted to the plateau region on top of the artefact; b) local height of the rectangle with equivalent area to the local artefact cross-section.

9.1.1 Signature topography of laser powder bed fusion of metals

The interactions between the laser source and the powder bed in the realisation of a layer in laser powder bed fusion (L-PBF) leaves a signature in the surface of the fabricated layer, which contains useful information to understand the process and the history of physical events it underwent during fabrication. This application is discussed in work by the author of this thesis [37].

In Figure 9.4 an example dataset obtained by coherence scanning interferometry is shown. Weld tracks, spatter formations and weld ripples are visible.

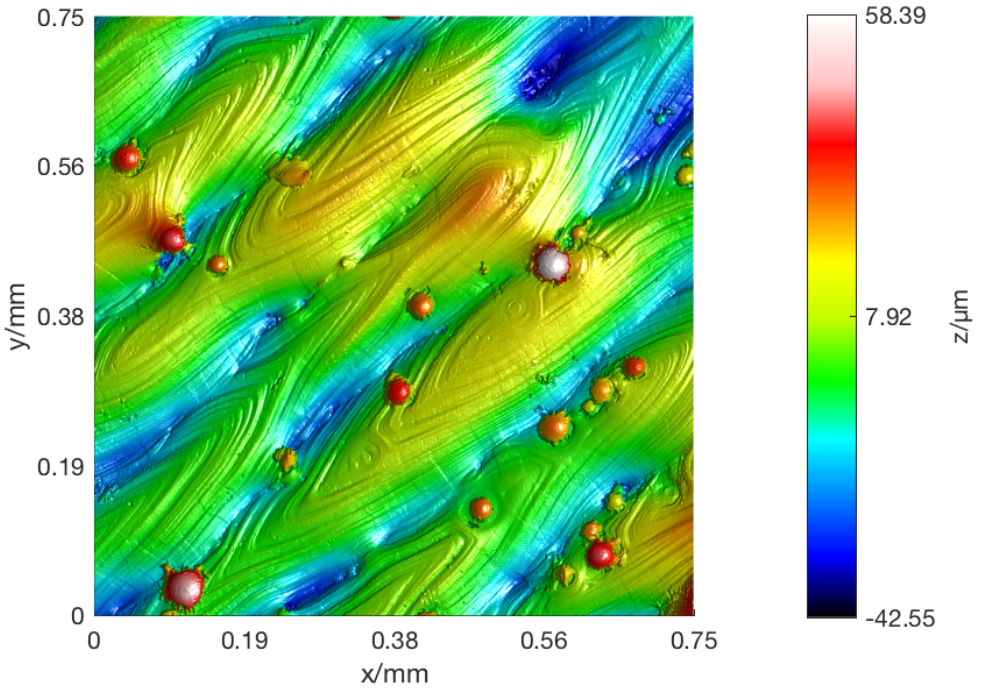


Figure 9.4: Example L-PBF dataset obtained by coherence scanning interferometry with a 20× objective.

Feature-based characterisation can be used to isolate and characterise some of the signature formations: spatter, weld tracks and weld ripples. In Figure 9.5, spatter formations are identified and separated from the substrate, after larger wavelength components have been temporarily removed[3]. In this example, the identification of the spatter features has been performed by means of the contour stability method (see Section 6.4.7).

Once separated, the spatter features can be characterised in terms of their dimensions

[3]Larger wavelength components are removed to facilitate the identification and separation of the spatter features. the components are reinstated once the partitioning is complete, so that after separation, the spatter features do not experience any shape alteration.

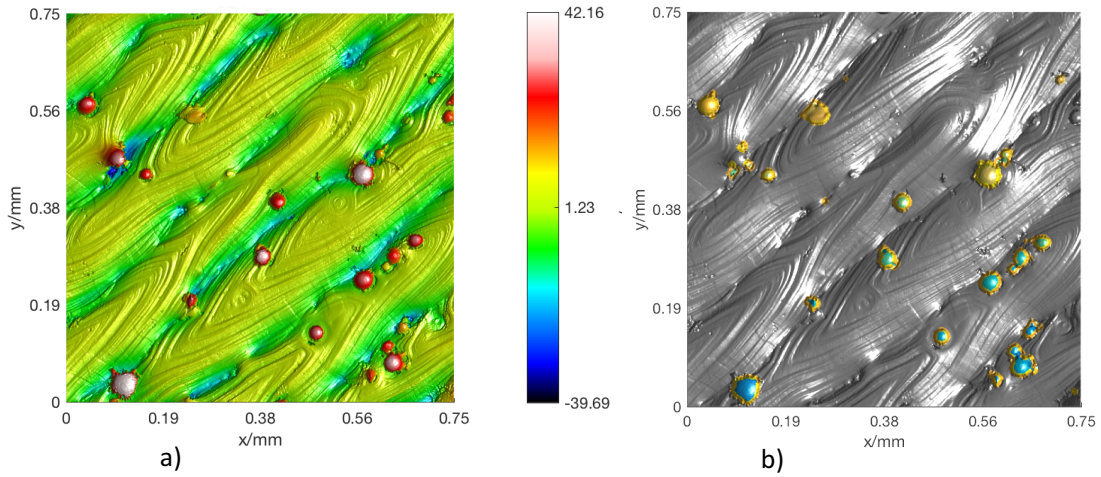


Figure 9.5: Identification of spatter formations; a) original topography after removal of large-wavelength components; b) identified spatter formation are shown coloured over the original topography in grayscale (from [37].)

and other shape properties. In Figure 9.6, two examples are reported: in Figure 9.6.a, image moments (see Section 5.3.2) computed on the features footprints are used to determine several x,y properties, such as area, centroid localisation, aspect ratio. In Figure 9.6.b, the thickness of the features is computed through a datum-based approach (see Section 5.3), as the mean distance between a datum plane fitted to the region immediately surrounding the base of the feature, and a datum plane fitted to the topmost portion of the feature. Both regions are highlighted in yellow in Figure 9.6.b.

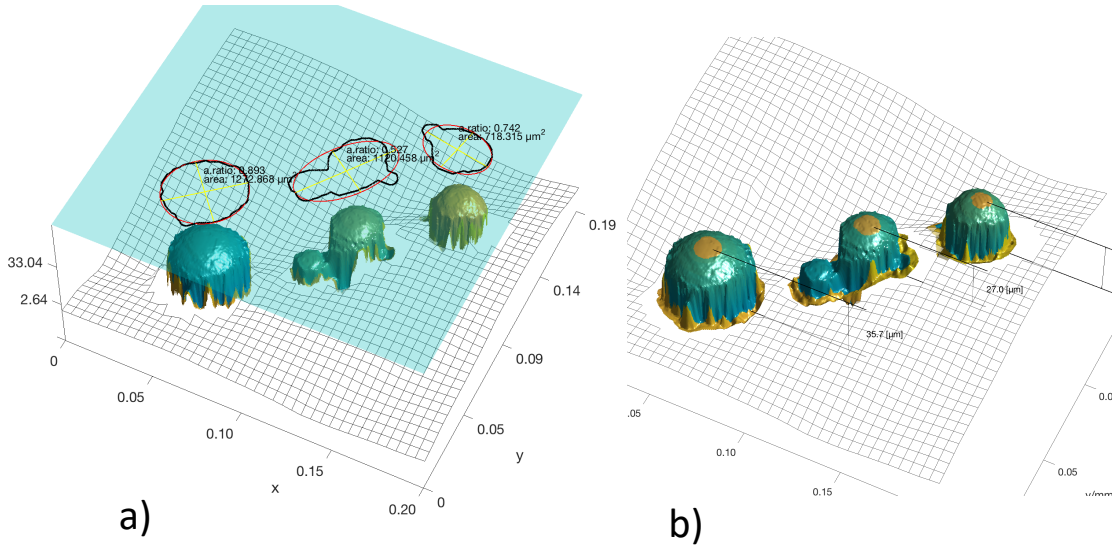


Figure 9.6: Dimensional characterisation of spatter formations; a) computation of x,y properties using image moments; b) computation of z properties (thickness) using reference datums (from [37]).

To obtain the individual weld tracks, the surface topography is first deprived of the

identified spatter formations. Then the remaining topography is subjected to a series of filtering operations aimed at removing any source of "noise" which may make the partitioning into weld tracks more challenging. The two main filtering steps are illustrated in Figure 9.7: in Figure 9.7.a, the topography is shown after extraction of the identified spatter particles (result referred to as *starting dataset*). In Figure 9.7.b the result of the weighted non-parametric, local regression (LOESS) fitting ($span 0.02\%$, 2^{nd}-order polynomial, see Section 5.1.1) is shown, applied to the fourth level of a Gaussian pyramid decomposition ($\sigma = 0.8\mu m$, see Section 5.4.4) of the starting dataset.

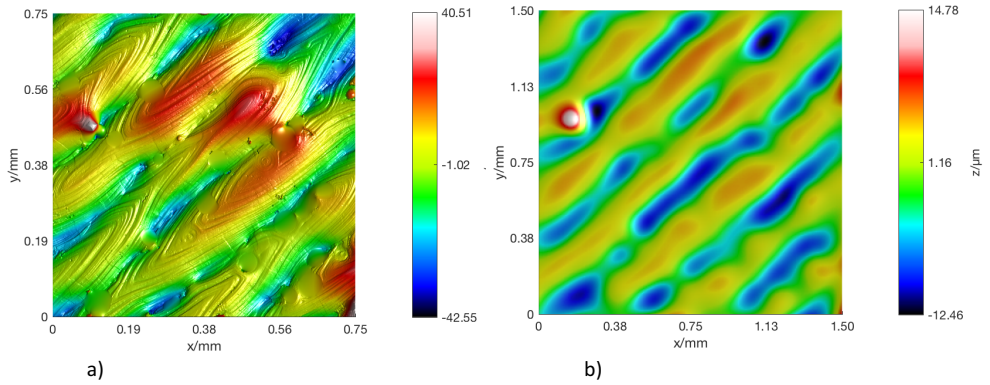


Figure 9.7: L-PBF topography filtering to facilitate separation of weld tracks; a) topography after removal of identified spatter formations; b) result of Gaussian pyramid decomposition and LOESS filtering.

The weld tracks can then be obtained by morphological segmentation into hills, according to the standard ISO 25178-2 [6], and as discussed in Section 2.2.2. Because of residual waviness left despite of filtering, hills that are aligned along their principal direction[4] are merged into individual partitions, as shown in Figure 9.8.

As done for spatter formations, the individual weld tracks, once isolated, can be characterised in terms of their width, thickness, etc. In order to investigate uniformity along the weld track, dimensional properties are computed in cross-sections, with similar methods as previously discussed for the inkjetted artefacts (see Figure 9.9). In particular, weld track width (see Figure 9.9.a) is computed as the width of each cross-sectional profile obtained by cutting the weld track across its main direction of elongation (see Section 5.3.1 for slicing methods and Section 5.3.2 for image moments, needed to find the direction of elongation). Weld track height is computed on the same cross-sections, but by using two section-related datum lines to identify the local "height" of the feature (corresponding to a given percentage of summit points) and local "base" of the feature, corresponding to surface points surrounding the feature at the sides of the cross-section, up to a given distance (See Figure 9.9.b).

[4]Found by using image moments, like described in Section 9.1

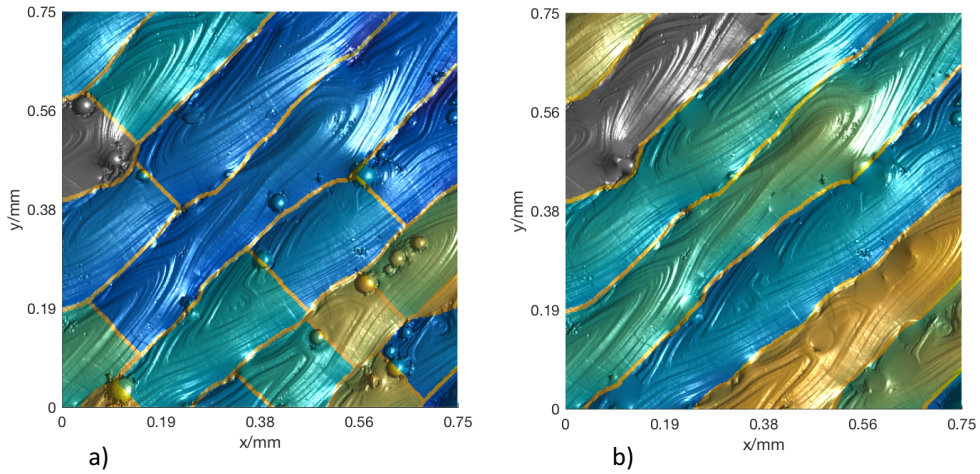


Figure 9.8: Morphological segmentation and hill-merging to obtain the weld tracks.

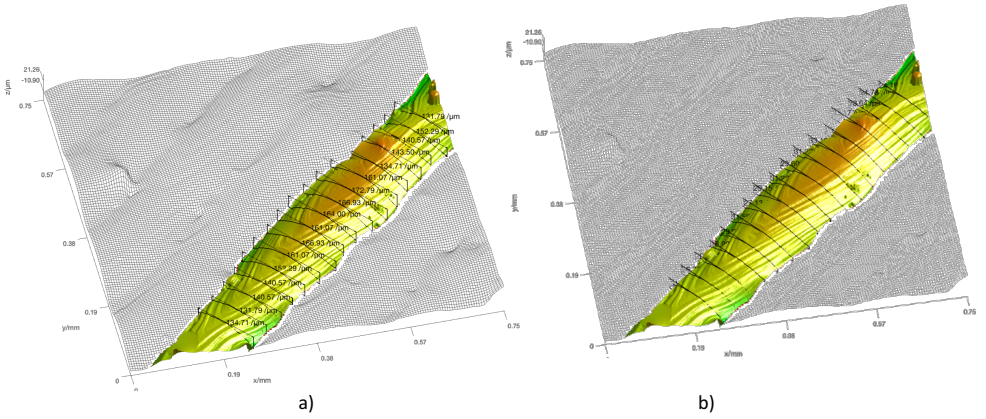


Figure 9.9: Width and thickness of weld tracks; a) width computation by slicing; b) height computation by slicing (from [37])

Finally, weld ripples can be studied by subjecting the topography to a different series of filtering operations, this time aimed at removing any component which may disturb the identification of the crests and valleys forming the high-frequency weld ripples (see Figure 9.10). The filtering is a combination of high-pass Gaussian filtering (residual from subtraction of Gaussian smoothing result) and LOESS fitting with 1^{st}-order polynomial, span=0.04 mm (see Section 5.1.1). Once such components are removed (see Figure 9.10.b), edge detection techniques based on the Canny algorithm, as discussed in Section 6.4.7, can be applied to identify the crests (see Figure 9.10.c).

A simple check on local crest orientation, performed by means of a moving window (see Section 6.4.1) and k-means clustering can then be used to separate regions where the ripples are parallel. Within such regions, the mean spacing between crests (a measure of ripple spacing) can be computed as shown in Figure 9.11.

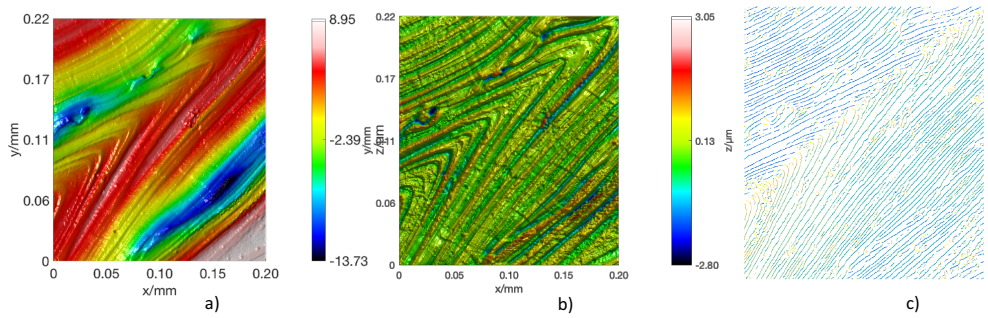


Figure 9.10: Identification of weld ripples; a) example topography dataset (cropped from a 20× CSI dataset as the one shown in Figure 9.4); b) filtered topography (high-pass Gaussian filtering and LOESS fitting ; c) ripple crests identified via Canny edge detection (from [37]).

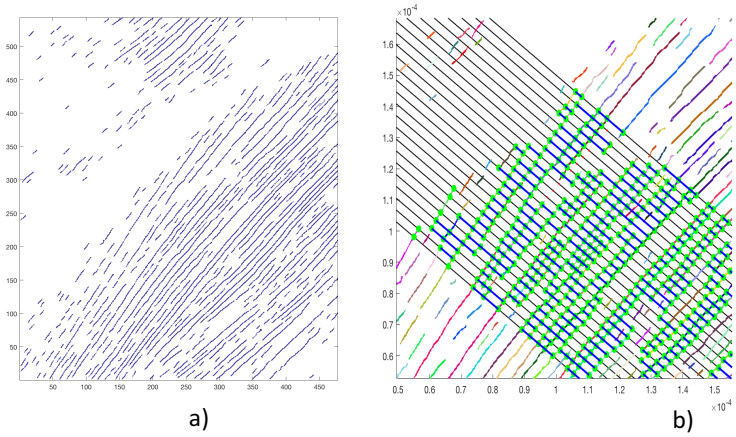


Figure 9.11: Spacing computation on the crests of weld ripples; a) subset of crests with uniform direction; b) spacing computation (from [37]).

9.1.2 Spatter as function of tilt in electron powder bed fusion

Once solutions for the automatic detection of spatter and other protruding singularities (i.e. particles) are in place, they can be used to evaluate how the quantity and spatial distribution of such features varies when varying process parameters or conditions. For example, in Figure 9.12 surfaces built by electron powder bed fusion (E-PBF) at different tilt angles are compared in terms of the amount of spatter formations and unmelted/partially melted particles that have remained attached to the surface. Conventional surface metrology would only allow one for comparing roughness states via texture parameters.

9.1.3 Effects of post-processing on electron powder bed fusion surfaces

The understanding of the behaviour and performance of post-processing operations applied to additively manufactured parts is a fundamental milestone towards the achieve-

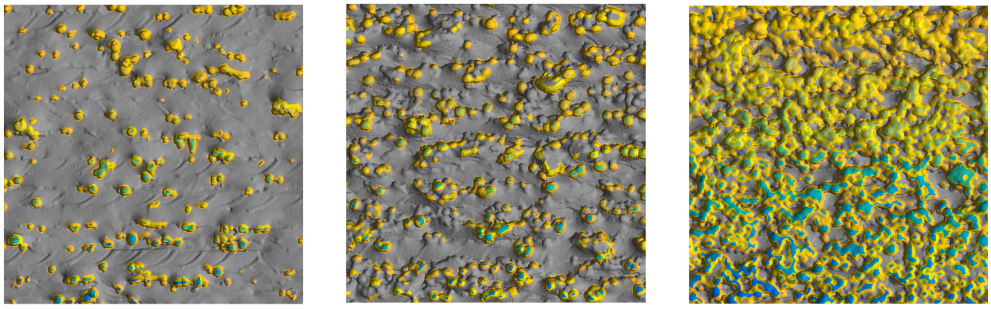


Figure 9.12: Spatter and unmelted/partially melted particles attached to electron powder bed fusion surfaces manufactured at different tilt angles. The automatically identified features are highlighted in yellow, a) 0° tilt; b) 45° tilt; c) 90° tilt.

ment of industrially applicable and effective additive manufacturing processes. By using conventional surface metrology means, one can typically compare the state of a surface before/after a modification process via the computation of texture parameters. A lower value of the *Sa* value for example (see ISO 25178-2 [6]) computed on a properly filtered dataset (in order to isolate the spatial frequencies of interest) indicates lower "roughness", i.e. a smoother surface. On the contrary, through the application of a feature-based pipeline, one can focus on individual topographic features, and better understand how the specific post-processing operation may influence their modification. In Figure 9.13 an additive surface produced by electron powder-bed fusion (E-PBF) is measured before and after a shot peening operation. The datasets shown in the figure would normally constitute the starting point to compute texture parameters to determine variations in roughness.

In Figure 9.14, feature-based characterisation is used to isolate spatter formations that have survived the shot-peening modification process. The features have been recognised by means of contour stability analysis (a method discussed in Section 6.4.7. With such results, one can study the survival of individual features, and modification events altering their shape and size.

9.1.4 Material removal by electrolyte jet machining

Electrolyte jet machining (EJM) processes are being studied in the Advanced Manufacturing Research Group at the University of Nottingham, as an innovative means for inducing material removal and surface topography modifications. As part of a series of investigations aimed at understanding behaviour and performance of such processes, for example in terms of material removal rates and controllability of the fabricated topographies, a series of artefacts were generated, consisting of single grooves produced

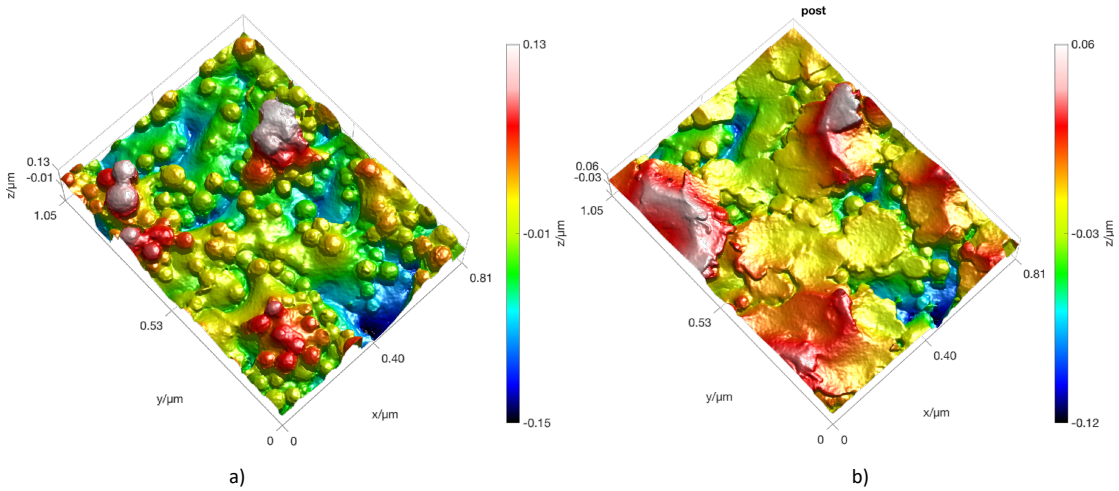


Figure 9.13: Focus variation measurements of a surface fabricated via E-PBF and subjected to shot-peening (20*times* magnification); a) result from the E-PBF process; b) same region after shot-peening (datasets courtesy of Lewis Newton (University of Nottingham) and the Manufacturing Technology Centre, UK).

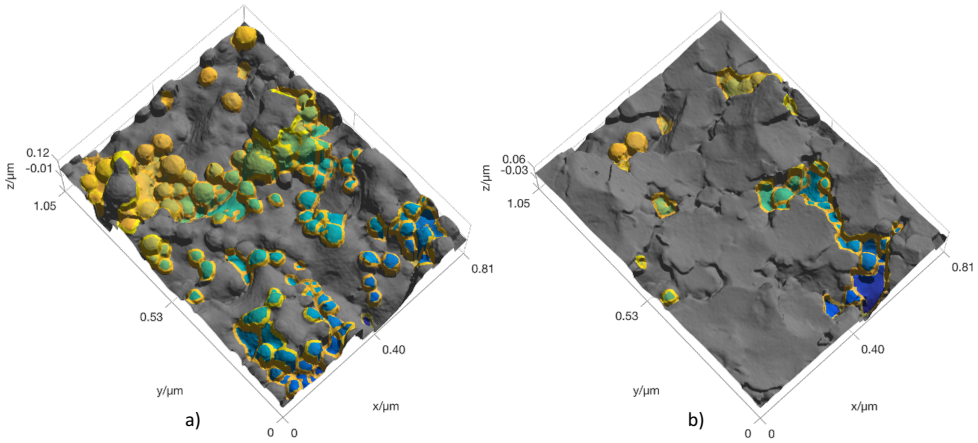


Figure 9.14: Feature-based characterisation of E-PBF surface pre and post surface modification; a) as-built; b) after shot-peening (datasets courtesy of Lewis Newton (University of Nottingham) and the Manufacturing Technology Centre, UK).

with varying manufacturing process parameters. Example datasets obtained by focus variation measurement of sample grooves are shown in Figure 9.15.

Feature-based characterisation can be used to identify the grooves and characterise their geometric and dimensional properties. An example application of a feature-based characterisation pipeline is illustrated in Figure 9.16. The groove is identified by k-means segmentation on heights (see Section 6.4.1). The result of identification is visible in figure 9.16.b. The dimensional properties of the groove are then investigating by slicing along its main direction of elongation (see Section 5.3.2 and Section 5.3.1). In Figure 9.16.c the results of width computation are shown.

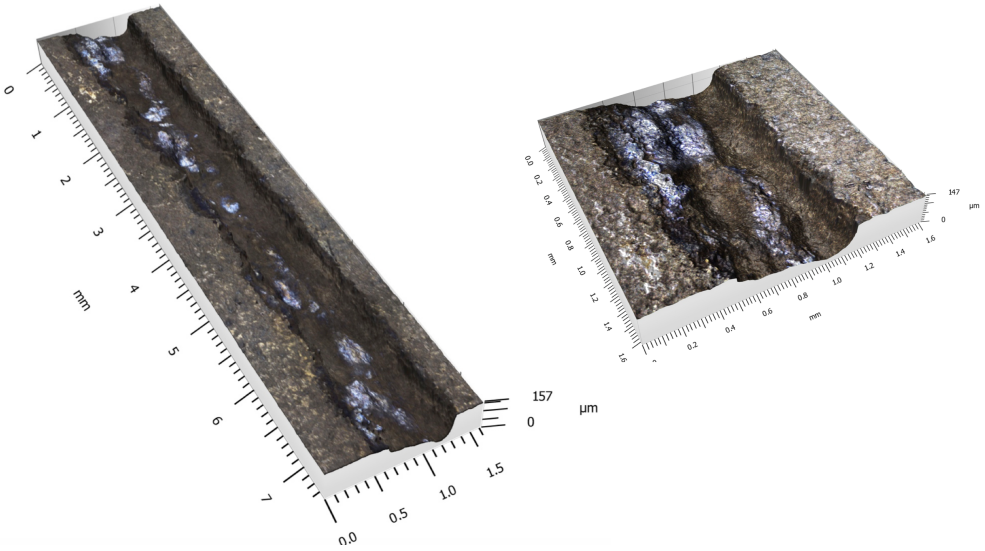


Figure 9.15: Focus variation measurement of EJM datasets at 5× magnification; left) stitched mode; right) single field of view. height maps shown with RGB colour overlays. Rendering by MountainsMap software [87]. Samples courtesy of J. Mitchell-Smith and A. Clare, University of Nottingham.

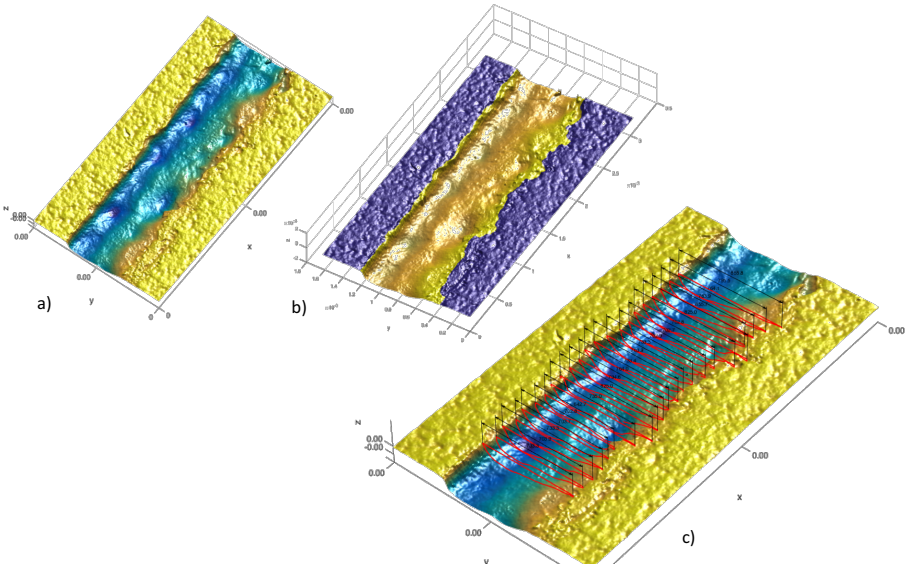


Figure 9.16: Feature-based characterisation of EJM groove; a) original height map (after levelling); b) result of feature identification; c) computation of groove width in multiple cross-sections. Images a) and c) do not report units, but the groove is approximately 800μm wide (samples courtesy of J. Mitchell-Smith and A. Clare, University of Nottingham).

9.2 Engineered topographies

Many applications where feature-based characterisation can play a useful role involve surfaces whose topography has been specifically designed and engineered to play an

assigned functional role. Tessellated surfaces, i.e. surfaces comprised of regular patterns of repeating, deterministic features, have numerous applications in aerodynamics, friction reduction, enhanced lubrication, and optics. An equally important category of applications consists of micro electro-mechanical systems and many other micro-parts that are often conveniently measurable by means of areal surface topography measurement instruments, and need to be inspected and verified against imposed dimensioning and tolerancing schemes, like it happens for normal, standard-sized parts. In such scenarios, feature-based characterisation can enable the identification and separation of the relevant surfaces, so that dimensional and geometric inspection can take place.

9.2.1 Micro-targets

A thin foil micro-target manufactured by STFC (Science and Technology Facilities Council, UK) for laser driven ion acceleration experiments is shown in Figure 9.17. The figure refers to a circular micro-target, suspended by three spokes. Manufactured variants include triangular and hexagonal targets, suspended by a variable number of spokes. Dimensional verification of these devices is at the moment impossible because they are too small to be measured by conventional means. Like for many products at this scale, inspection relies solely on qualitative assessment via SEM images like what shown in Figure 9.17. Without any quantitative estimate of manufacturing error it is difficult to identify ways to improve the manufacturing process, and thus reduce the scrap rates.

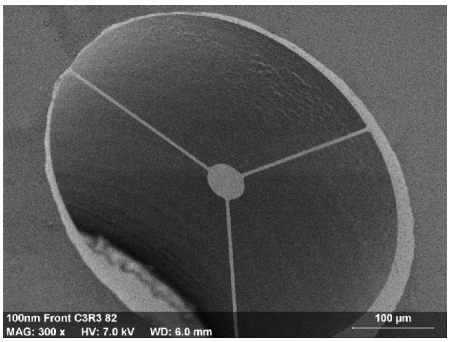


Figure 9.17: SEM image of a micro-target for ion acceleration experiments (from [29]).

An attempt was made by the author of this thesis to develop an inspection and verification solution based on areal topography measurement and feature-based characterisation [29, 30]. The use of coherence scanning interferometry for topography measurement leads to results like what shown in Figure 9.18. Whilst evidently the undercut sections and vertical walls of the micro-part cannot be captured, information may be extracted pertaining the top surfaces, and about the x, y geometric properties of the

micro-part.

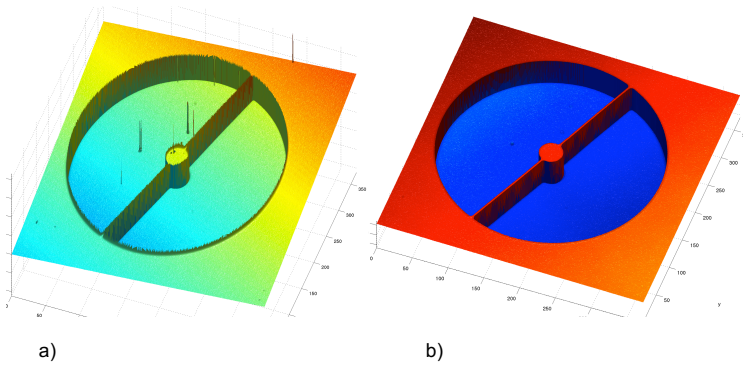


Figure 9.18: Height map resulting from CSI measurement of the micro-target (from [29]); a) raw dataset; b) dataset after form removal (subtraction of the least-squares mean plane) and reduction of measurement error artefacts (batwings).)

In Figure 9.19, the procedure for identifying and separating the central hub and spokes of the micro-target is illustrated for a two-spokes configuration with a circular, centre target. Given the step-like nature of the device, the procedure is mainly based on thresholding via k-means clustering on heights (as illustrated in Section 6.4.1), with region-based post-processing and checks for topological connectivity (see Section 6.4.4) to separate the main components of the micro-target.

Once the components have been separated, their geometric and dimensional properties can be characterised by identification of reference datums for dimensional measurement (see Section 5.3) as illustrated in Figure 9.20).

Variants of the illustrated procedure for the inspection of triangular and hexagonal target shapes, with multiple spokes, are shown in Figure 9.21.

9.2.2 Micro-fluidics channels for lab-on-a-chip devices

Lab-on-a-chip devices are comprised of micro-parts designed for delivering fluids to different parts of the device, or to act as flow-control systems, e.g. micro-valves. Current micro-fabrication technologies for this type of application, e.g. predominantly multilayer soft lithography, can achieve the desired geometries, but often with poor geometric and dimensional control. Like for many other micro-parts, verification is mainly via functional testing, with consequent generation of high scrap rates, and poor understanding of how fabrication processes should be optimised. However, many devices for micro-fluidics applications consist of surface features of sizes compatible with the observational scales of the areal topography measurement instruments available for surface metrology. Thus it makes sense to investigate the possibility of applying

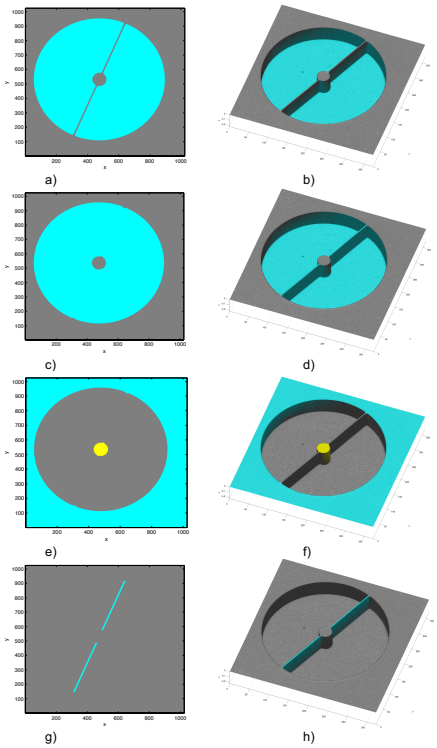


Figure 9.19: Data processing steps to identify and isolate the relevant features (target, spokes and outer rim); a,b) thresholding by k-means clustering on height; c,d) extraction of central target and outer rim by morphological erosion; e,f) discrimination between central target and outer rim by test for topological connectedness; g,h) identification of the spokes by manipulation of the previous classification maps (from [29])

feature-based characterisation pipelines for the inspection and verification of such devices. In Figure 9.22, the height map resulting from a micro-channel belonging to microfluidics valve, measured by means of a confocal laser scanning microscope, is shown. Because the sides of the channel are beyond the maximum measurable slope, considerable error artefacts are present at the channel edges.

Two segmentation processes are executed (see Figure 9.23): one to isolate the planar surfaces far from the channel (Figure 9.23.a), to perform selective levelling with the $z = 0$ plane; the other to isolate the bottom portion of the channel (see Figure 9.23.b) where measurement is more reliable, as being further away from the maximum measurable slope limit. The first segmentation is executed by k-means clustering on local slope (see Section 6.4.1). The second segmentation is executed by thresholding on height (see again Section 6.4.1).

The bottom region of the channel is then fitted to a parametric cylindrical surface by least-squares minimisation (see Figure 9.24). The axis of the fitted cylinder defines the channel position and orientation, whilst a CAD-compare approach (see Section 5.2) can

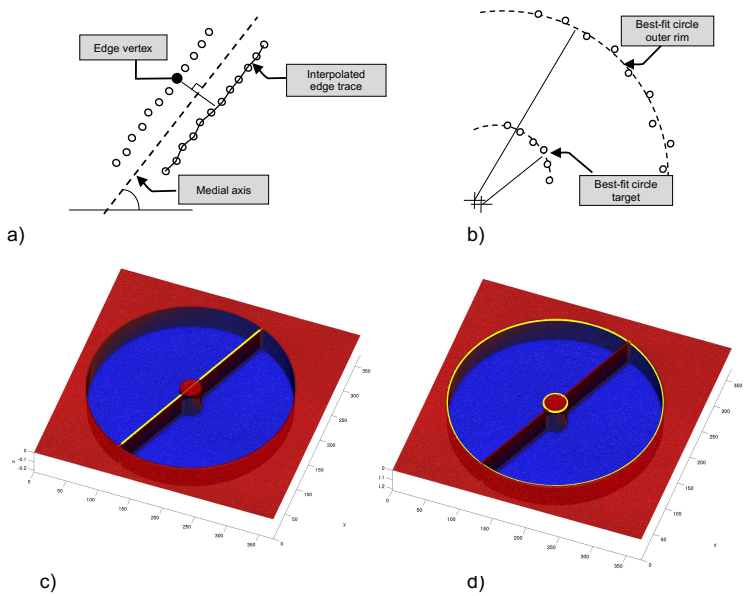


Figure 9.20: Identification, separation and characterisation of the main micro-target components; a,c) identification of the spoke medial axis (to obtain spoke orientation) by fitting the spoke edge points; b,d) identification of the outer rim and central target boundaries (to compute diameters) by circle fitting of edge points (from [29]).

be used to obtain a quantitative map of differences between the channel surface and its nominal counterpart.

9.2.3 Arrays of micro-pockets for friction reduction

The use of regular arrays of micro-pockets to reduce friction in surfaces destined to bearing applications is currently subject of a considerable amount of research efforts (several references can be found in the work by the author of this thesis [5]). As for any other tessellated surface, the problem of characterising regular patterns is a multi-scale one, since function is exerted at large scales due to the concerted action of many feature instances (e.g. by influencing the friction coefficient), but the manufacturing process responsible for generating the features of the pattern usually operates at smaller scales, for example by fabricating the unit features in sequence. Several examples of how pattern units can be identified and characterised, and how the characterisation can be extended to check for regularity of the entire pattern have been illustrated and discussed in the chapters dedicated to feature characterisation (Chapter 5, Section 5.6) and feature identification (Chapter 6, Section 6.5). In Figure 9.25 example datasets are shown belonging to a silicon nitride disk patterned with circular micro-pockets ($100\mu m$ nominal diameter, fabricated by nano-second pulsed laser machining), highlighting the multi-scale nature of the surface, and the need for areal topography measurement to cover

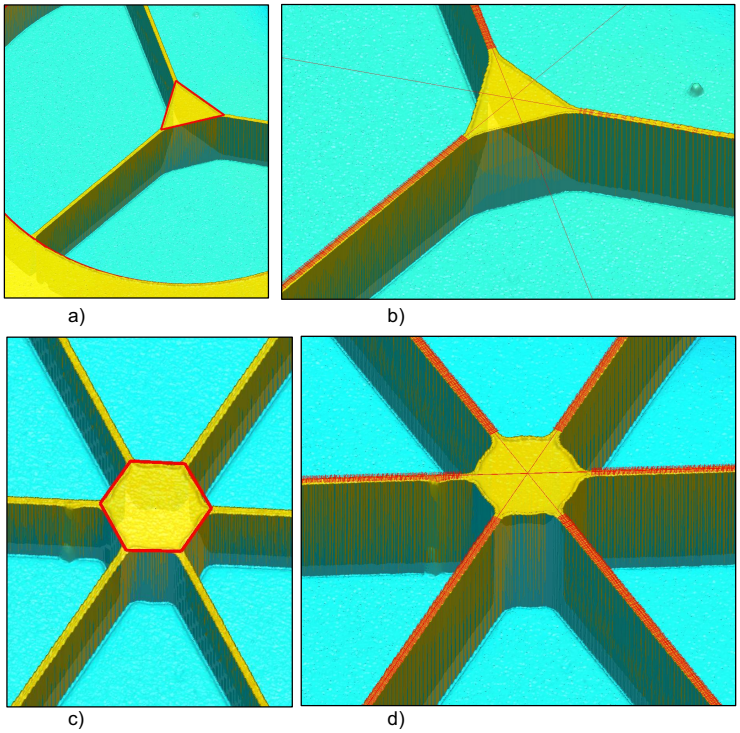


Figure 9.21: Identification of reference datums for the characterisation of multiple target shapes and variable numbers of spokes; a) best-fit edges for the triangular target; b) best-fit medial axes for spoke orientation, and cross-sectional segments for evaluation of uniformity of spoke width; c) best-fit edges for hexagonal target; d) medial axes and cross-sectional segments for the six-spokes configuration. (from [29]).

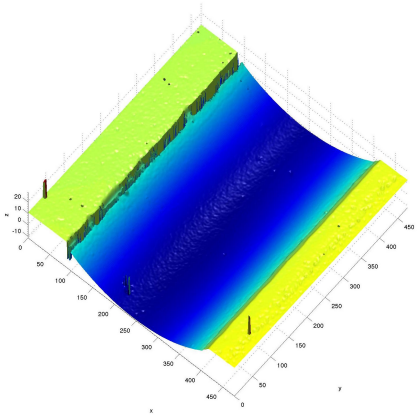


Figure 9.22: Dataset resulting from areal topography measurement of micro-channel by means of a confocal laser scanning microscope (dataset courtesy of Manuel Balcon and Enrico Savio, University of Padua, Italy). Field of view approximately $(637 \times 477)\mu m$.

multiple ranges and resolutions in order to capture any potentially relevant elements of the pattern (from [5]).

In Figure 9.26, two maps are shown, one highlighting the spatial distribution of pocket

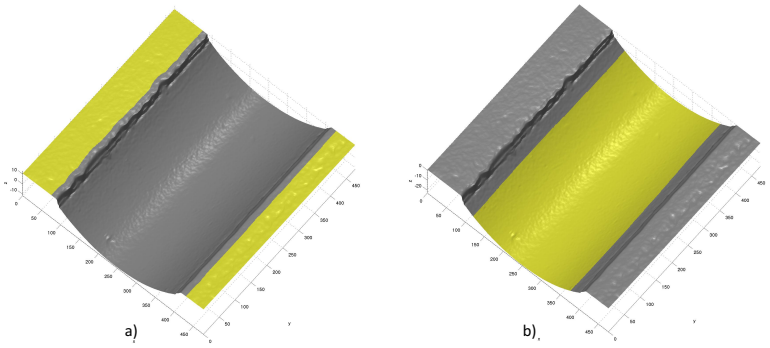


Figure 9.23: Results of micro-channel topography segmentation; a) isolation of regions sufficiently far away from the channel to act as references for selective levelling; b) identification of the bottom portion of the channel (selected regions are highlighted in yellow).

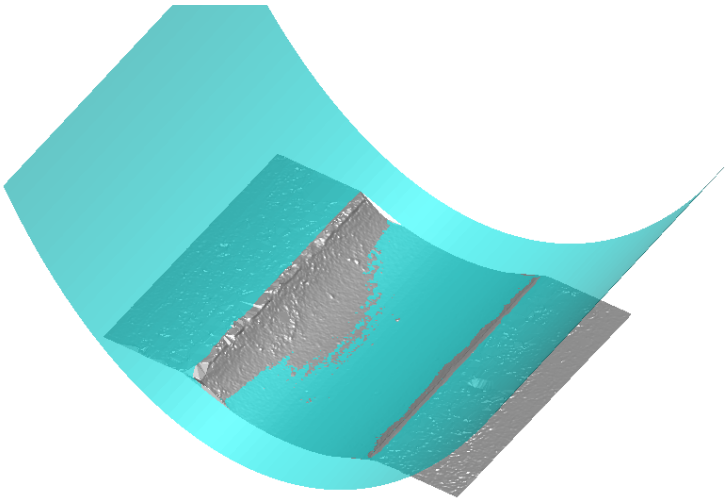


Figure 9.24: Result of fitting the bottom portion of the micro-channel to a parametric cylindrical surface. The cylinder axis can serve as an indication of channel position and orientation. The cylinder radius and the local discrepancies between the cylinder itself and the local channel topography provide information on the cylinder shape and dimensions.

diameters, the other highlighting horizontal spacing between neighbouring pockets (both generated using the methods described in Section 5.6 and Section 6.5).

Results such as those shown in Figure 9.26 may act as an important starting point to investigate how functional performance of the surface is related to its multi-scale topographical content.

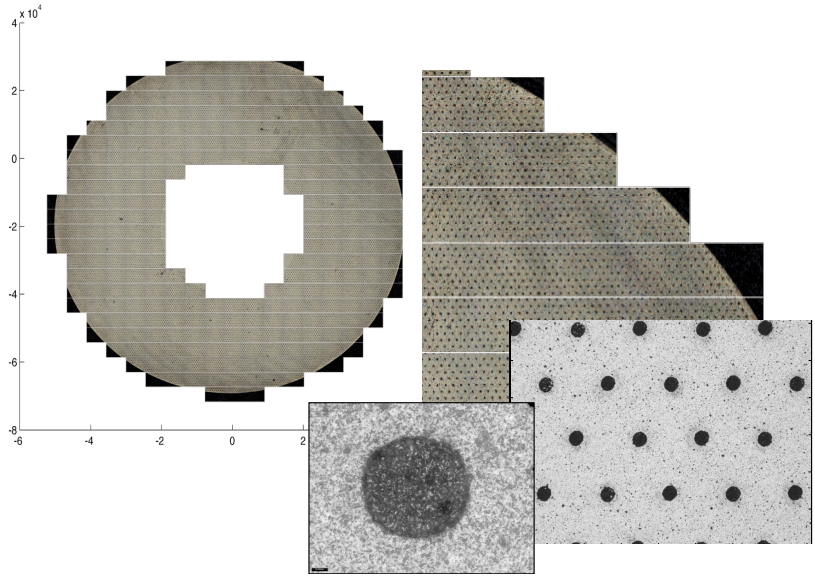


Figure 9.25: Focus variation measurement at multiple magnifications to capture the multi-scale nature of a regular pattern of micro-pockets for friction reduction; from 2.5× with lateral stitching to 50× single field-of-view.

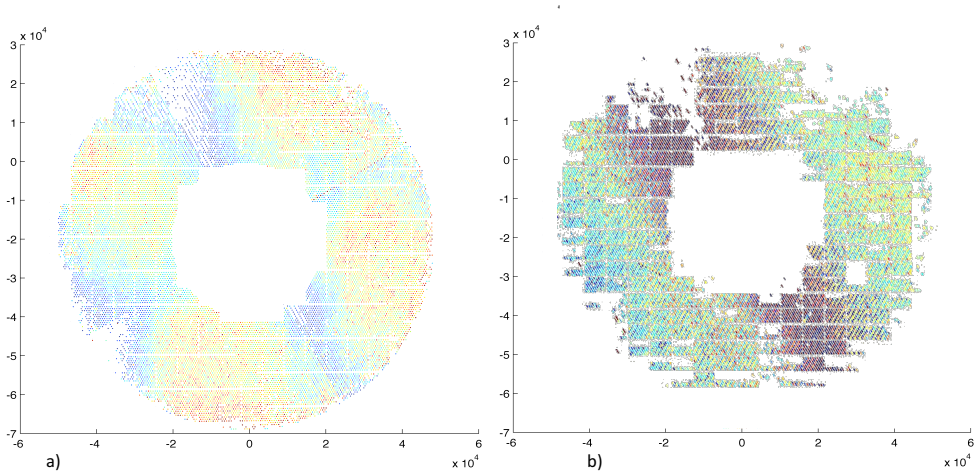


Figure 9.26: Maps of the entire disk highlighting the spatial distribution of pocket attributes; a) colour proportional to pocket diameter; b) colour proportional to horizontal spacing of neighbouring pockets. In both cases, values above the acceptability thresholds have been removed from the maps to highlight defected regions (from [5]).

9.2.4 High aspect-ratio micro-holes

The problem of how to measure high aspect-ratio micro-holes is relevant to many industrial applications, but is currently unsolved. CMMs equipped with touch probes cannot penetrate beyond a certain depth, and are limited by probe size. Optical tech-

nologies have trouble receiving enough light back from deep recesses, and are thus equally limited to the immediate entry or exit regions of the holes. X-ray computed tomography seems to be the only industrially-viable measurement technology currently capable of measuring high aspect-ratio holes throughout, but the metrological performance and behaviour of XCT is still under investigation, with significant research going on in the attempt to understand measurement error and traceability. In previous work by the author of this thesis [16], touch probe measurement by CMM, optical measurement by focus variation and XCT measurement have been compared in correspondence to the entry and exit regions of high aspect-ratio holes manufactured by abrasive water jet cutting. A feature-based characterisation pipeline has been implemented to enable the identification of the hole from areal topography datasets (from both the focus variation and the XCT measurements), and the computation of diameter and hole cross-sectional shape as a function of depth. The procedure for computing diameter as a function of depth is illustrated in the following on an example dataset from XCT measurement (entry region of a $100\mu m$ diameter hole). The performance comparison between the different measurement technologies is discussed in [16]. The surface extracted from the volumetric, raw dataset produced by XCT measurement is shown in Figure 9.27. The nature of XCT volumetric datasets and how surfaces can be extracted from them is discussed in Chapter 3, Section 3.4.

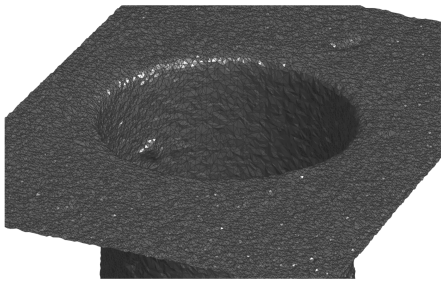


Figure 9.27: Surface (triangle mesh) resulting from XCT measurement of high aspect-ratio micro-hole (from [16]).

The procedure for identifying the hole and computing diameter as a function of depth is illustrated in Figure 9.28. An initial segmentation process (k-means clustering on local slope and local height - see Section 6.4.1) is used to isolate portions of the surface surrounding the hole, to act as reference for selective levelling, so that a $z = 0$ datum plane can be defined as a reference for depth computation. A series of parallel z planes are used to slice the hole at increasing depth, thus obtaining cross-sectional profiles useful to assess circularity and hole diameter. The diameter itself is obtained by circle fitting each cross-sectional profile[5]

[5]It should be noted that this is the only example reported in this thesis where the entire procedure

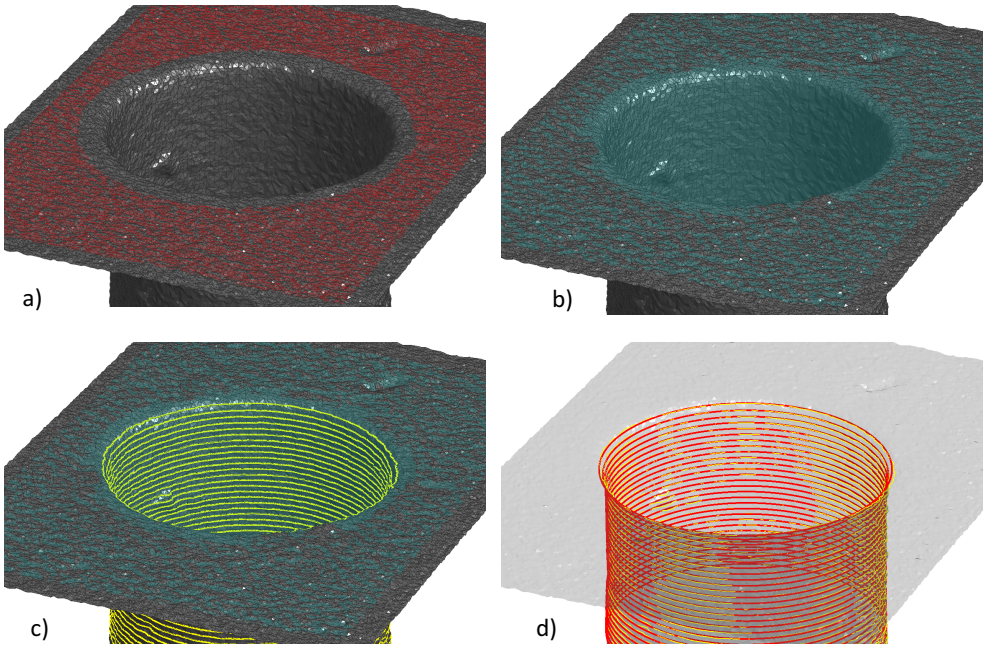


Figure 9.28: Procedure for computing diameter as a function of depth for the high aspect-ratio micro-hole; a) points belonging to the top surface surrounding the hole entry region are found by segmentation (k-means clustering on local slope and height); b) a plane is fitted to the segmented region to establish a $z = 0$ datum reference for depth computation; c) parallel planes at decreasing z are used to slice the triangle mesh; d) the hole cross-sections (in yellow) are fitted to circles (in red) to obtain diameter as a function of depth (from [16]).

9.3 Tribological testing and functional life of surfaces

The final, relevant category of applications illustrated in this chapter pertains to how feature-based characterisation can support tribological testing and in general, the investigation of surface modifications as a consequence of functional life. Modifications due to testing or functional life can be quantified in terms of variation of texture parameters, for example, by computing the *Sa* parameter before and after the modification process. However, when the modification is local, texture parameters may fail to capture it properly, being influenced also by the surroundings of the modification. In addition, in many applications one may be interested in the size and shape properties of the local modification (e.g. a micro-indentation mark, a crater, etc.), again something which cannot be properly captured by texture parameters, but is instead suitable to be investigated by means of feature-based characterisation methods.

for feature identification and characterisation has been designed to operate on a triangle mesh instead of on height maps, as usually done. See Chapter 3 and in particular Section 3.4 for a discussion on the implications of working on full-3D triangle meshes.

9.3.1 Fretting wear (testing)

Fretting wear testing performed on cylindrical specimens results in material removal, and in the formation of a recessed region whose shape and volume can help understand how the material reacted to the modification events. At the moment there is no consolidated solution to quantify the geometric and dimensional properties of the recessed region. However, such region is often of size compatible with the areal topography measurement solutions available in surface metrology. In Figure 9.29, the height map resulting from the measurement of a specimen after it has been subjected to testing is shown.

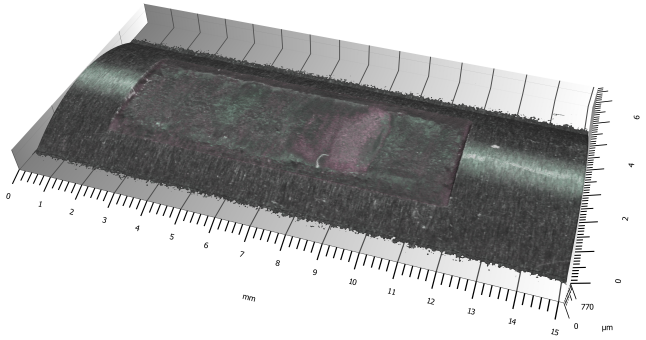


Figure 9.29: Height map and overlaid RGB colour map resulting from focus variation measurement of a specimen subjected to fretting wear testing (specimen courtesy of Z. Tengtuo and P. Shipway, University of Nottingham. Note: the image has been generated with commercial software).

In Figure 9.30, the application of feature-based characterisation is shown for the automated identification of the worn region and the computation of the volume of removed material. Segmentation uses thresholding on local slope (recognising that the worn region is much flatter than its surroundings), using the methods illustrated and discussed in Section 6.4.1.

9.3.2 Fretting wear (functional life)

Wear modifications due to functional operation of a surface may be characterised by large variability in shape and size, due to increased randomness and lesser known circumstances, with respect to the experimental conditions that may be created in tribology testing scenarios. In Figure 9.31.a, the height map resulting from focus variation measurement on a worn metallic surface is illustrated. Computing a texture parameter such as *Sa* (ISO 25178-2 [6]) may provide indications on overall roughness, but would fail to capture the visible evidence of modifications not being uniformly distributed

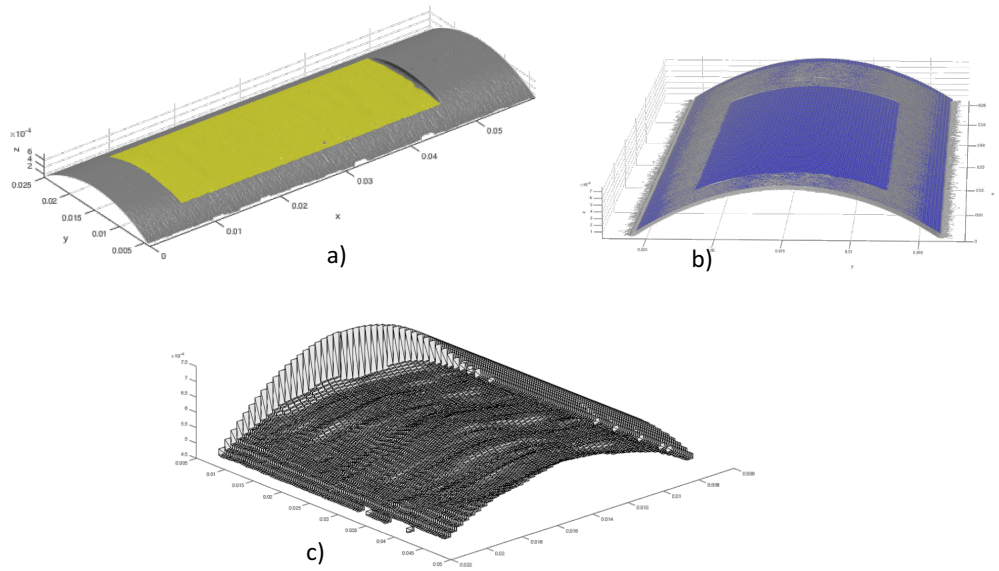


Figure 9.30: Identification of the worn region and computation of the volume of removed material; a) identification of the worn region (shown in yellow) by segmentation; b) identification of a surface datum (shown in blue) via selective least-squares fitting of cylindrical surface to the unworn regions; c) reconstructed void volume between the datum and the worn surface (voxelised, altered proportions).

over the surface. In Figure 9.31.b, a feature-based characterisation pipeline has been designed where the topography is segmented by k-means clustering using height variance within a moving window as local texture descriptor (as discussed in Section 6.4.2 and Section 6.4.1). The largest worn region (selected by post-processing of the classification map, as illustrated in Section 6.4.4) is highlighted. Once the region has been identified, it can be characterised in terms of size, shape, relative spacing (if multiple). depending on application requirements.

9.4 Summary and outlook

A series of test cases have been presented in this chapter, that demonstrate the application of the original solutions developed during the doctoral studies and the usefulness of feature-based characterisation in general. The test cases cover a series of industrially-relevant scenarios involving manufacturing process understanding and optimisation, product development, material and tribological testing. The cases provide an indication how feature-based characterisation could applied within diverse scenarios, and what types of benefits it could provide, both in the domain of industrial inspection and in terms of answering specific research questions in more research-oriented scenarios.

It is likely that the original topography analysis methods developed during this doctor-

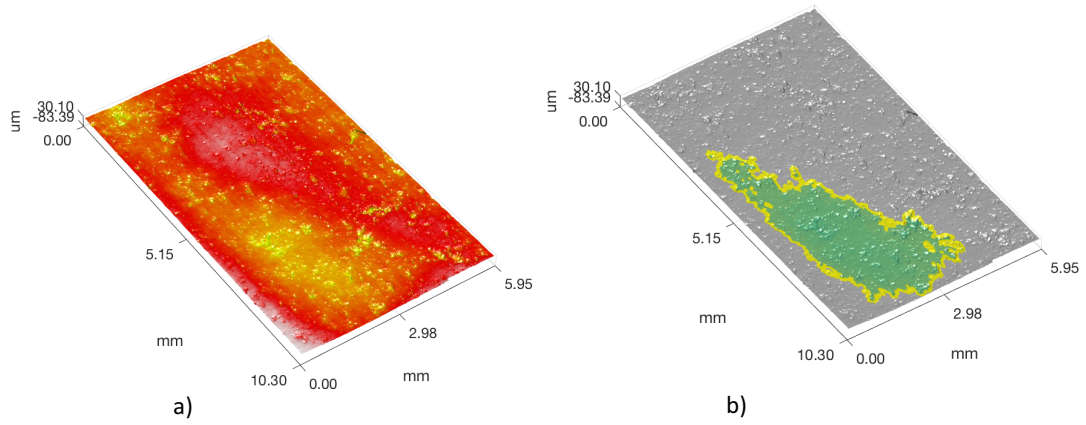


Figure 9.31: Identification of the largest worn region on a sample measured after functional operation (undisclosed application); a) original dataset (focus variation $10\times$, with lateral stitching, field of view approximately $6\mu m \times 10\mu m$; b) the largest worn region is highlighted (original sample courtesy of K.Barman and K. Voisey, University of Nottingham)

ate, and feature-based surface characterisation in general, may provide benefits to further application scenarios which have not been covered in this thesis. A fundamental verification that should always be performed on any of such new scenarios is whether the application of the methods proposed in this thesis may bring actual improvement with respect to the state-of-the-art in each specific field of application.

Chapter 10

Conclusions

In this thesis a series of original methods for *feature-based characterisation* has been illustrated and discussed. A novel framework encompassing such methods and providing a theoretical foundation to a broader view of surface metrology has been presented. As opposed to the conventional tools available for the characterisation of surface topography based on the computation of texture parameters and dedicated to the assessment of roughness-related properties, feature-based characterisation lets the user focus on individual topographic formations, or networks of them, providing means to identify, extract and characterise features of interest in terms of their shape, size, position and orientation.

10.1 Achievement of research objectives

In this section, the achievement of the research objectives set for the doctoral studies is discussed, with reference to the list of objectives previously reported in Chapter 1, Section 1.2.

– *Identify a mathematical formalism and associated digital representation of surface topography which can handle both conventional and non-conventional (i.e. feature-based) characterisation approaches.*

  Mathematical representations based on $z = f(x,y)$ and $z = f(x)$ functions discretised over a regular, equispaced x,y grid have been elected as the format of choice for observed instances of surfaces, reflected in terms of digital encoding into *height map* data structures, storing measured height values in matrix-like fashion. This solution is illustrated in Chapter 3. In the research work presented in this thesis, the use of multiple maps storing additional information other than

height to encode selection, classification, colour and other local attributes, is presented. The idea of using multiple overlays is not new in surface metrology and implemented in most state-of-the-art surface metrology software. The same idea underlies the much more mature concept of layer-based editing in image processing. However, the approach is originally used here to seamlessly represent entire topographies, as well as the results of segmentation (classification maps), and the results of feature selection and extraction (boolean maps), demonstrating to be suitable to both conventional and non-conventional (i.e. feature-based) characterisation paths. Furthermore, an original, more general formal model of surface topography, which incorporates variability (across spatial localisations within the same topography instance, or across multiple instances) has also been presented and discussed in Chapter 4. The model is based on spatial statistics and in particular on the theory of random fields. The basic concept is not new, as it is mutuated from the representation of random time series as random processes in the signal processing literature. However it is here formally redefined to apply to spatial profiles and spatial areal datasets. Moreover, parallels have been drawn between the most common operators used in surface metrology to remove form and the purposes and effects of statistical model fitting. Furthermore, filtering as a means to separate more random and more deterministic components from topography model and the consequent effects on spatial correlation has been discussed.

– *Identify the conditions of applicability of conventional characterisation methods (i.e. based on texture parameters) in formal/mathematical terms, and describe conditions of non-applicability.*

The adoption of the theory of random fields to model topography affected by variability (randomness) has allowed to identify the conditions of stationarity and ergodicity, in particular applied to Gaussian fields, as useful for determining the degree of applicability of ISO texture parameters and the meaningfulness of the results (Chapter 4). The role and influence of spatial correlation has been discussed as well in Chapter 4. The results hint at the possibility to use statistical tests for stationarity or ergodicity as a means to assess the suitability of parameter-based characterisation. Furthermore, again starting from random fields, various types of non-stationary behaviour have been identified (Chapter 4), and matched to the classes of "non-conventional" test cases identified in Chapter 1. Clearly, characterisation of topography based on texture-parameter can be seen as a special case of a broader problem, as stationarity is a special condition within a much larger set of possible non-stationary variants of a random field.

– *Introduce a formal definition of* feature *and* feature-based characterisation*, and formalise the phases of the feature-based characterisation process, providing a consistent definitional and taxonomic framework that can accommodate existing and new methods dedicated to each phase.*

  A formal treatment of the concepts of feature and feature-based characterisation, the definition of the main phases of the feature-based characterisation procedure as well as a formal characterisation and taxonomy of methods belonging to each phase, have been provided through Chapters 4,5 and 6. The main phases of feature characterisation, feature identification and feature extraction have been defined and discussed in the context of the global feature-based characterisation framework.

– *Develop new methods dedicated to the various phases of feature-based characterisation, compensating and complementing the existing ones, and define how they fit within the framework.*

  Original methods developed during the doctorate for feature characterisation include: slicing and 2D datums (Section 5.6), principal component analysis on full-3D geometries (Section 5.3.2), the Angular Radial Transform for height maps (Section 5.4.3) and of Scale-invariant Feature Transform for height maps (Section 5.4.5). Original methods for feature identification and extraction include: scan-free template matching (Section 6.3.2), all the partitioning methods described in Section 6.4.3, except simple thresholding and k-means clustering, the contour stability analysis method (Section 6.4.7) and the solutions for transitioning from edge maps to feature maps (Section 6.4.7).

– *Demonstrate the methods and the overall framework of feature-based characterisation through the application to a selected set of test cases.*

  All the applications investigated during the doctorate, except for the micro-target test case (Section 9.2.1) are reported in Chapter 9. For each test case, the scenario, the applied feature-based characterisation procedure, the issues encountered and the final benefits obtained with respect to conventional surface topography characterisation, are illustrated and discussed. To further demonstrate the developed methods for feature-based characterisation and quantify their performance, original simulation models and tools have also been developed for the generation of synthetic topographies. These are presented and discussed in Chapter 8. Finally, two original methods to assess topography measurement error components and their propagation through feature-based characterisation pipeline have been

presented in Chapter 7, and discussed in context with the challenge of assessing uncertainty.

– *Provide a critical overview of accomplishments as well as pointers to currently unsolved challenges.*

A section named "summary and outlook" has been provided at the end of each chapter, to discuss the accomplishments, as well as the main open challenges related to each specific subject covered in section of this thesis. In Section 10.2 a more general discussion is reported, addressing issues pertaining feature-based characterisation in general.

10.2 Discussion

A series of considerations must be made, in relation to the general role of feature-based characterisation, its main challenges, open issues, and other aspects worth further investigation.

10.2.1 Usefulness (and relations to surface function)

One of most relevant open issues with conventional surface metrology is the elucidation of the relationships between surface texture parameters and surface "function". In other words, do texture parameters actually convey information that is useful to assess or compare functional performance of surfaces? This problem has been discussed for a long time, in particular in periods when there has been a proliferation of texture parameters and no one would be sure of what to use and when [3, 17]. Several studies have reached to the conclusion that certain subsets of texture parameters may be more useful in specific application domains (e.g. see the examples reported in [1]). However, the general problem of deciding what texture parameters better match specific functional needs is far from being solved.

Under this perspective, one may wonder whether feature-based characterisation may bring any improvement. As this work has hopefully shown, feature-based characterisation encourages the adoption of a broader perspective when referring to the concept of "function", because not everyone involved with the characterisation of a surface may necessarily be interested in studying how a surface behaves during its operational life. Manufacturing researchers would likely be more interested in studying a surface to understand and improve the process that generated it; tribologists may want to focus on specific types of damage, or individual results of mechanical testing (e.g. indenta-

tion marks); researchers working on the development of micro-products may be more interested in dimensional and geometric verification at the micro-scale. In a sense thus, feature-based characterisation entirely removes the problem of associating topography characterisation results to "function", because it is designed to target surface properties which the end-users themselves have selected as relevant to their application. So "function" is really what the end-user is interested in. Instead, from the perspective of the feature-based approach, the problem of linking characterisation to function has originated from conventional metrology because of the adoption of "generalistic" texture parameters. If on one hand texture parameters are of wide applicability and application independent, on the other hand they often fail to establish a direct link with the requirements being targeted within each application, and thus usually require further investigation to establish correlations with the variables of interest.

Perhaps the real question is whether feature-based characterisation, as it has been intended in this work (i.e. a collection of methods and algorithms for feature identification, extraction and characterisation), is really justified when in many cases one may achieve similar results by simple point and click of features on screen, and by manually computing sizes, distances, angles, and other shape properties in a few simple passages, using the interactive data manipulation tools that nowadays most commercial surface metrology software applications offer (examples where given in Section 2.5.1). Surely, in many situations, including for some of the test cases discussed in previous chapters, simple manual investigation via interactive tools may suffice. However, several important scenarios come to mind, where an algorithmic approach such as what proposed in this work, would be preferable:

- when there are many instances of the target feature within the dataset;
- when there are many datasets to be processed;
- when the nature or properties of the feature may be debatable, or uncertain, so that different operators may reach to different conclusions about the exact location of a boundary, or about whether a region represents an actual instance of the feature or not;
- when there is any other reason of concern regarding aspects which may undermine repeatability and reproducibility of the approach; and
- when any quantification of uncertainty must be associated to the result.

In all the above scenarios, approaches based on algorithmic feature-based characterisation can be considered fundamental. At least to ensure consistency of the results

through multiple iterations, and to ensure comparability. Uncertainty is a much bigger issue, as discussed later in Section 10.2.5.

10.2.2 Usability

How usable is really feature-based characterisation as illustrated and discussed in this work? Meaning, how difficult would it be to apply feature-based approach to a test surface, and how likely for one to be successful? How skilled should an operator be to adopt this type of approach, and how knowledgeable of the specific application?

One of the advantages of conventional characterisation of surface topography by means of texture parameters, is that the procedures have been refined over decades, and are now essentially fool-proof, almost entirely automatically implemented by any surface metrology instrument and software manufacturer, with a minimal amount of effort required, (primarily to properly set-up the measurement conditions, and to choose the proper filter setting[1]). Most importantly, conventional surface topography characterisation can be carried out in most cases with a minimal knowledge of the nature of the surface, and the nature of the processes or events one is trying to investigate. After all, texture parameters such as *Sa* (see Chapter 2) do provide an indication of scatter of height data which is independent of any application scenario, or even of the nature of the surface from which measurement data has been obtained from. With feature-based characterisation, the situation is different. Feature-based characterisation is better seen as a toolbox, a set of tools to choose from depending on needs and preferences. So, one has typically a multitude of alternative solutions available for feature identification, extraction, characterisation. The price to pay for such flexibility is complexity, and a higher entry point in terms of knowledge and skills required on the user's side (on the specific application requirements, on the nature and peculiar aspects of the involved surfaces, on the methods and algorithms).

Evidently, feature-based characterisation may also be likely to require longer set-up times, a larger number of variables to consider before an appropriate processing pipeline can be implemented, and a clearer view of the characterisation objectives, meaning that one needs to settle on what dimensional and geometric variables are really of interest for any given specific application.

However, one may say that most of these activities are not really spared to users embracing conventional approaches based on computing texture parameters. Lack of

[1] It is amazing how often things may go wrong anyway.

knowledge on how to set cut-offs for filters, or lack of knowledge on what texture parameter should be chosen to better capture application-relevant properties of the surface, and in any case lack of justification for choices made when setting up data processing pipeline, are frequently observed in the application of conventional surface metrology solutions both in industry and research. Which is likely to be one of the fundamental reasons for the current disconnect between results of surface topography characterisation and the reasons surfaces may be characterised in the first place (as discussed in Section 10.2.1), i.e. to analyse their functional properties, or to analyse the processes that created them.

10.2.3 The set-up challenge

The large amount of optional routes and tools available to the user pursuing a feature-based characterisation approach generates an additional problem, that is, how should the optimal pipeline be found for any given application. This does not only imply the choice of what tools to use, but also how to configure each one of them, given the multitude of controlling parameters which could be varied for most of the feature identification, extraction and characterisation methods illustrated in the previous chapters. This issue is evidently deeply tied with the previous one concerning usability and is currently unsolved.

At the moment, feature-based characterisation pipelines must be set-up and configured via trial & error, by continuously tweaking the numerous control parameters in the attempt to fine tune behaviour and performance to match the application requirements. So, for example, if for an application a 95% success rate in the identification of a target feature is acceptable, then multiple test surfaces must be explored, and the feature identification part of the pipeline must be tuned to achieve that 95%. Analogously, if the result of a dimensional characterisation process must have a repeatability error of $\pm 5\mu m$ on a specific dimension, then the pipeline must be tuned and validated to make sure it achieves such goal.

An automated solution to directly create an optimised pipeline, starting only from information about the characterisation objectives and from a sample of measured surfaces, would be highly desirable. At the moment though, the harsh reality is that one needs to rely on heuristics coming from multiple attempts. Expertise and experience of the operator called to design the feature-based pipeline may also help speed up the process, but it is undeniable that any application may have its peculiar, previously unseen aspects, making the time needed to reach to an acceptable result often hard to predict.

10.2.4 Procedural definitions

Because of the wide array of optional routes available to achieve a feature-based characterisation result, and because of the complexity often accompanying the shapes of the features targeted by the investigation, concepts apparently absolutely elementary such as for example a feature "height" may be really difficult to define. As discussed in Chapter 5, everyone is clear on the meaning of "diameter" for a simple and archetypical shape such as a circle, or a cylinder, but what would "diameter" mean for a more irregular geometry? As discussed throughout Chapter 5, the meaning of "diameter" would be tied to the procedure used to compute it, including how the surface data was filtered, how reference datums where associated to it, how the datums were associated to each other in order to define the dimension. So, two "diameter" attributes may mean completely different things, despite sharing the same name, because of how they have been procedurally obtained. Much as *Ra* or *Sa* values do not mean anything if not accompanied by information on the entire procedure of levelling, filtering and selection of measurement lengths, so any result of a feature-based characterisation pipeline is of little value if not accompanied by a fully documented computational procedure, needed to replicate it. Because of this reasons, feature attributes should always be considered as *procedural attributes*, meaning that they should be defined in terms of the procedure applied to compute them. All considered though, this should not be seen as an additional burden of feature-based characterisation, and providing information for replication of a result should be common practice in any measurement-related activity.

10.2.5 Uncertainty and validation

As discussed at length in Chapter 7, the problem of how to define and associate an uncertainty budget to the results of a feature-based characterisation pipeline is a challenging one, and far from having an complete solution. As previously seen, this problem is actually shared by surface metrology as a whole, as it directly maps onto how to associate uncertainty to the results of surface measurement in general. A thorough investigation of error is needed, covering all aspects that go from surface data acquisition, to its algorithmic processing. whilst some of the error components can be reconstructed via replicate experiments, as shown in Section 7.3 and Section 7.4, the problem of how to estimate local bias remains a fundamental one, in particular given the current lack of measurement solutions of better metrological performance to elect as comparison references (see Section 7.3.1).

10.2.6 Shape and randomness

At the heart of feature-based characterisation, lies the issue of how to represent shape in mathematical terms; how to represent complexity that often goes well beyond archetypical shapes such as cubes and cylinders, and often also beyond what can be captured by simple mathematical formulations. How to represent variability of shape, with its deterministic and random components, across instances, and/or across time-frames; and how to represent variability due to measurement error, often the same order of magnitude of the shapes themselves. As an aggravating factor, there should really be no difference between an entire surface and a localised topography feature laying onto it. In both cases we should be dealing with "shape", despite working with varying complexity, and varying randomness. There should really be a continuum connecting the characterisation of the regular, macroscopic shape of an industrial part, and the characterisation of the irregular, microscopic shape of its surface topography features, approaches changing only based on shape complexity and randomness, independently of scale[2]. This seamless transition has not been fully implemented yet. In the theory presented in this work, there is currently a disconnect between the advanced statistical models of topography which one may generate starting form measured data (Chapter 4) and the fact that most feature identification and characterisation methods presented in Chapter 5 and Chapter 6 actually operate on single observations of profiles and surfaces. The possibility of better integrating the two concepts as a means to study the propagation of measurement error and estimating uncertainty has been explored in Chapter 7, but we are currently far from a complete and seamless integration.

10.2.7 Moving to fully three-dimensional geometric models

Our current approaches to statistical modelling are also intrinsically limited, mostly because the starting point is still geometric information defined by $z = f(x,y)$ equations, thus unable to capture fully three-dimensional geometric information. The mathematical and statistical tools whose application has been explored in this work, are all based on the $z = f(x,y)$ assumption, and this is justified by most surface metrology instruments being still unidirectional probing devices (see Chapter 3). The increasing introduction of volumetric measurement solutions, and multi-axis CMM-like architectures equipped with optical 2.5D imaging probes (see Chapter 3) suggest that $z = f(x,y)$ representations may soon be proven insufficient. When moving towards fully three-dimensional representations of shapes, all of a sudden many solutions derived from

[2]In the ideal conditions of scale-independent measurement error

computer vision, and consequently from image processing, fall apart. Many solutions derived from signal processing also become inapplicable, as the concept of there existing a "lateral" and a "vertical" direction needs rethinking. Statistical signal modelling, and all the ramifications in terms of random fields, need to be redefined. The challenge is both in terms of the need to radically redesign the methods, but also in terms of needing to accommodate increasingly larger datasets, when, even now, fitting a Gaussian field to a point cloud may prove daunting, given the high densities that can be achieved with the current generation of measurement instruments.

10.2.8 The role of machine learning

Feature-based characterisation touches upon subjects such as visual perception, and makes frequent use of solutions involving statistical modelling and data mining for classification, clustering and in general to support decision making. Evidently, all such subjects constitute fertile ground for the application of machine learning (ML), and an increasing number of ML-related technologies are becoming conveniently available thanks to the resurgence and popularity of ML in many scenarios involving ICT and big data. Despite this, the role of ML throughout this thesis has been somewhat muted. Despite clustering playing a dominant role in several of the segmentation techniques illustrated in Section 6.4.1, clustering itself being considerable as a type of unsupervised machine learning, in the research solutions described in this thesis there has been so far a clear lack of methods based on the application of supervised learning technologies, such as artificial neural networks.

There is a fundamental issue at play here, based on the distinction between approaches that focus on processing low-level information to extract transformed data to feed into ML engines, and approaches that rely entirely on the capability of the ML engine to identify itself the best way to transform the raw data. Deep learning technologies, and in particular convolutional neural networks for image classification are an example of the second approach, as the network itself develops the most suitable way to process pixel data in order to improve its classification performance. However, so far, the methods and solutions investigated for feature-based characterisation may all be considered as belonging to the first approach; i.e. explicit processing of raw measurement data to create transformed information (i.e feature-based characterisation results). Such results then may as well be fed into a ML engine for prediction, classification, etc., but this possibility has not yet been considered as priority in the research so far. The multiple, possible roles of ML in feature-based characterisation are thus worth further

discussion, and are a promising venue for future investigation.

10.2.9 In-process metrology

Most of the methods and algorithms that have been illustrated and discussed in this work have relatively slow execution times, meaning that they were not originally designed to operate in real-time environments where response time is important. For this reason, the applications presented throughout this thesis are all concerned with R&D type of inspection scenarios, where measurement is imagined as performed off-line, in laboratory environments, with no particularly stringent time constraints.

However, the importance of metrology for in-process monitoring is undeniable, in particular in relation to the fabrication of high value-added parts where in-situ measurement may mean the difference between early detection of issues and the costly rejection of non-conforming parts at the end of the fabrication process.

In-process and in particular in-situ measurement often imply the need for very rapid executions, both in term of data acquisition and data processing. The first problem is that most of the surface measurement technologies that have been assumed as the basis of the data analysis and processing methods discussed in this work are simply not quick enough for in-process application, even for relatively slow processes like many additive manufacturing ones[3]. Many surface measurement technologies have also troubles with being placed close to the process to perform in-situ measurement (because of obstruction, limited manoeuvrability, hostile environments). Thus, for a future in-process application, feature-based data processing methods would need to either adapt to new data formats, compatible with different measurement technologies suitable for in-process application, or marry new measurement technologies and solutions capable of providing data in the same formats currently accepted by feature-based characterisation (see Chapter 3). For the time being, the only viable application of feature-based characterisation for in-process measurement is based on the use of two-dimensional digital imaging, which means though that the technologies used are essentially the ones already established in computer vision. Solutions capable of generating hight maps and compatible with in-situ application are slowly emerging, though mostly in prototype form.

The second problem pertains data processing speed. Again, most of the methods illustrated in the previous chapters have not necessarily been designed to be fast. But

[3] slower if compared to high-production volume processes such as forging.

for many applications of in-process metrology, data processing speed is essential. A similar problem has been dealt with in the evolution of computer vision for in-process inspection. The most typical trends followed so far in computer vision consist of converting the algorithms into hardware-embedded forms (i.e. the development of smart camera devices), and in preferring the adoption of very simple data processing methods (e.g. simple thresholding on pixel intensity levels for segmentation, and use of image moments for feature characterisation). Hardware-embedded ML solutions are also being investigated. It remains to be seen whether these design trends can be embraced by feature-based characterisation of surface topography.

10.3 Final remarks

Feature-based characterisation of surface topography provides ways to answer metrological questions that currently cannot be answered by a mere evaluation of surface roughness, and offers new opportunities for exploring original pathways in the development of advanced manufacturing processes, materials and products.

Feature-based characterisation sets its mathematical foundations in the problem of how to analyse and represent complex shapes, affected by varying amount of variability, and measured under non-negligible noise. It delves deep into pre-existing knowledge and technologies from computer vision, image processing, geometric modelling, and statistical modelling.

The results obtained throughout the doctorate have been illustrated and discussed in this thesis. Original methods for the identification and characterisation of topography features have been presented, together with a complete theoretical foundation supporting feature-based analysis and framing conventional characterisation via texture parameters as a special case, subordinate to the verification of properties of the topography under investigation. The methods and framework have been demonstrated via application to a series of test cases originated from applications of industrial interest, spread across the domains of manufacturing process research, inspection of micro-products and micro-structures, material characterisation and tribological testing.

Despite many having already attempted to develop application-specific, ad-hoc approaches where measured topography data is processed with the intent to isolate and quantify topographic formations of interest, the work illustrated in this thesis represents the first attempt at unifying methods of feature-based characterisation into a comprehensive framework, and the first attempt to develop a formal apparatus to support

it. Amongst the original contributions of this doctoral work, the introduction of random fields to model topography and its associated variations, and the identification of conditions of applicability of conventional characterisation routes (texture parameters) as well as feature-based methods, promotes the viewpoint that sees both feature-based characterisation and parameter-based characterisation as facets of a common, wider-breadth discipline of surface metrology, encompassing all the types of characterisation tasks and scenarios explored so far. Further endeavours may pertain the strengthening of this new formal foundation for surface metrology, as well as the expansion of the surface characterisation toolset with the addition of new methods to support existing and new problems.

Bibliography

[1] C. Brown, H. N. Hansen, X. J. Jiang, F. Blateyron, J. Berglund, N. Senin, T. Bartkowiak, B. Dixon, G. L. Goic, Y. Quinsat, W. J. Stemp, M. K. Thompson, P. S. Ungar, and E. H. Zahouani, "Multiscale analyses and characterizations of surface topographies," *CIRP Annals*, vol. 67, no. 2, pp. 839–862, 2018.

[2] "ISO 17450-1, Geometrical product specifications (GPS) – General concepts – Part 1: Model for geometrical specification and verification," International Organization for Standardization, Geneva, Switzerland, Standard, 2011.

[3] R. K. Leach, *Characterisation of areal surface texture*. Heidelberg: Springer Verlag, 2013, pp. 353–353.

[4] R. Leach, Ed., *Optical measurement of surface topography*. Berlin: Springer, 2011.

[5] N. Senin, G. MacAulay, C. Giusca, and R. K. Leach, "On the characterisation of periodic patterns in tessellated surfaces," *Surface Topography: Metrology and Properties*, vol. 2, no. 2, p. 25 005, 2014.

[6] "ISO 25178-2 Geometrical Product Specification (GPS) - surface texture: areal - part 2: terms, definitions and surface texture parameters," International Organization for Standardization, Geneva, Switzerland, Standard, 2012.

[7] "ISO 4287 Geometrical Product Specifications (GPS) – Surface texture: Profile method – Terms, definitions and surface texture parameters," International Organization for Standardization, Geneva, Switzerland, Standard, 1997.

[8] "ISO/DIS 25178-607, Geometrical product specifications (GPS) – Extraction," International Organization for Standardization, Geneva, Switzerland, Standard, 2010.

[9] "ISO 16610-71, Geometrical product specifications (GPS) – Filtration – Part 71: Robust areal filters: Gaussian regression filters," International Organization for Standardization, Geneva, Switzerland, Standard, 2014.

[10] "ISO 16610-61, Geometrical product specification (GPS) – Filtration – Part 61: Linear areal filters – Gaussian filters," International Organization for Standardization, Geneva, Switzerland, Standard, 2015.

[11] "ISO 25178-3, Geometrical product specifications (GPS) – Surface texture: Areal – Part 3: Specification operators," International Organization for Standardization, Geneva, Switzerland, Standard, 2012.

[12] "ISO 13565-2, Geometrical Product Specification (GPS) - Surface texture: Profile method: Surfaces having stratified functional properties - part 2: height characterisation using the linear material ratio curve," International Organization for Standardization, Geneva, Switzerland, Standard, 1998.

[13] P. J. Scott, "Pattern analysis and metrology:the extraction of stable features from observable measurements," *Proc. R. Soc. Lond. A*, vol. 460, pp. 2845–64, 2004.

[14] "ISO 4288, Geometrical Product Specifications (GPS) – Surface texture: Profile method – Rules and procedures for the assessment of surface texture," International Organization for Standardization, Geneva, Switzerland, Standard, 1998.

[15] R. Hocken and P. Periera, *Coordinate measuring machines and systems*, 2nd ed. Bosa Roca: CRC Press, 2012.

[16] A. Rabani, N. Senin, F. Helmli, P. Butler-Smith, and R. Leach, "An integrated measurement method for complex micro-scale geometries," in *Proceedings of the 16th EUSPEN International Conference*, May 10, 2016.

[17] X. Jiang, P. Scott, D. Whitehouse, and L. Blunt, "Paradigm shifts in surface metrology. part ii. the current shift," *Proceedings of the Royal Society, A*, vol. 463, no. 2085, pp. 2071–2099, 2007.

[18] X. J. Jiang and D. J. Whitehouse, "Technological shifts in surface metrology," *CIRP Annals - Manufacturing Technology*, vol. 61, pp. 815–836, 2012.

[19] L. Blunt and S. Xiao, "The use of surface segmentation methods to characterise laser zone surface structure on hard disc drives," *Wear*, vol. 271, no. 3-4, pp. 604–609, 2011.

[20] N. Senin, L. A. Blunt, R. K. Leach, and S. Pini, "Morphologic segmentation algorithms for extracting individual surface features from areal surface topography maps," *Surface Topography: Metrology and Properties*, vol. 1, no. 1, pp. 15 005–15 005, 2013.

[21] L. B. Kong, C. F. Cheung, X. Q. Jiang, W. B. Lee, S. To, L. Blunt, and P. Scott, "Characterization of surface generation of optical microstructures using a pattern and feature parametric analysis method," *Precision Engineering*, vol. 34, no. 4, pp. 755–766, 2010.

[22] G. D. MacAulay, N. Senin, C. L. Giusca, R. K. Leach, and I. Atanas, "Review of feature boundary identification techniques for the characterisation of tessellated surfaces," *Surface Topography: Metrology and Properties*, vol. 3, no. 013002, 2015.

[23] X. Jiang, P. Scott, and D. Whitehouse, "Freeform surface characterisation - a fresh strategy," *CIRP Annals - Manufacturing Technology*, vol. 56, no. 1, pp. 553–556, 2007.

[24] A. Burla, "Fourier descriptors for defect indication in a multiscale and multisensor measurement system," *Optical Engineering*, vol. 50, no. 4, pp. 043 603–043 603, 2011.

[25] X. Q. Jiang, L. Blunt, and K. J. Stout, "Development of a lifting wavelet representation for surface characterization," *Proceedings of the Royal Society of London*, no. 456, pp. 2283–2313, 2000.

[26] N. Senin and R. Groppetti, "Automated identification and extraction of dale-like features from the three-dimensional topography of manufactured surfaces: An application to scratch testing," in *Proceeding of the 8th AITeM conference*, Montecatini terme, Italy, 2007.

[27] N. Senin, M. Ziliotti, and R. Groppetti, "Three-dimensional surface topography segmentation through clustering," *Wear*, vol. 262, no. 3-4, pp. 395–410, 2007.

[28] N. Senin, S. Pini, and R. Groppetti, "Identification of microtopographic surface features and form error assessment," in, B. M. Colosimo and N. Senin, Eds. London: Springer, 2010, pp. 159–187.

[29] N. Senin, L. Blunt, and M. Tolley, "The use of areal surface topography analysis for the inspection of micro-fabricated thin foil laser targets for ion acceleration," *Measurement Science and Technology*, vol. 23, no. 10, pp. 105 004–105 004, 2012.

[30] N. Senin, L. A. Blunt, and M. Tolley, "Dimensional metrology of micro parts by optical three-dimensional profilometry and areal surface topography analysis," *Proceedings of the Institution of Mechanical Engineers Part B: Journal of Engineering Manufacture*, vol. 226, no. 11, pp. 1819–1832, 2012.

[31] N. Senin and L. Blunt, "Characterisation of individual areal features," in *Characterisation of Areal Surface Texture*, R. K. Leach, Ed. Heidelberg: Springer, 2013, pp. 179–216.

[32] N. Senin, M. Moretti, and L. Blunt, "Identification of individual features in areal surface topography data through template-matching," in *14th international conference on Metrology and Properties of Engineering Surfaces*, Taipei, Taiwan, Jun. 2013.

[33] N. Senin, M. Moretti, and L. A. Blunt, "Identification of individual features in areal surface topography data by means of template matching and the ring projection transform," *Surface Topography: Metrology and Properties*, vol. 2, no. 1, pp. 14 007–14 007, 2014.

[34] M. Moretti, N. Senin, and R. K. Leach, "Geometric registration of areal topography data for comparative surface metrology," in *Proceedings of the AITEM Conference*, Palermo, Italy, 2015.

[35] N. Senin, R. Leach, S. Pini, and L. Blunt, "Texture-based segmentation with gabor filters, wavelet and pyramid decompositions for extracting individual surface features from areal surface topography maps," *Measurement Science and Technology*, vol. 26, no. 9, p. 095 405, 2015.

[36] N. Senin, M. Moretti, and R. K. Leach, "Shape descriptors and statistical classification on areal topography data for tile inspection on tessellated surfaces," *Measurement*, vol. 95, pp. 82–92, 2017.

[37] N. Senin, A. Thompson, and R. Leach, "Feature-based characterisation of signature topography in laser powder bed fusion of metals," *Measurement Science and Technology*, vol. 29, no. 4, 2018.

[38] "ISO 12181-2, Geometrical product specifications (GPS) – Roundness – Part 2: Specification operators," International Organization for Standardization, Geneva, Switzerland, Standard, 2011.

[39] "ISO 5436-1, Geometrical Product Specifications (GPS) – Surface texture: Profile method; Measurement standards – Part 1: Material measures," International Organization for Standardization, Geneva, Switzerland, Standard, 2000.

[40] "ISO/DIS 25178-600, Geometrical product specifications (GPS) – Surface texture: Areal – Part 600: Metrological characteristics for areal-topography measuring methods (under development)," International Organization for Standardization, Geneva, Switzerland, Standard, 2017.

[41] "ISO 25178-601, Geometrical product specifications (GPS) – Surface texture: Areal – Part 601: Nominal characteristics of contact (stylus) instruments," International Organization for Standardization, Geneva, Switzerland, Standard, 2010.

[42] "ISO 25178-602, Geometrical product specifications (GPS) – Surface texture: Areal – Part 602: Nominal characteristics of non-contact (confocal chromatic probe) instruments," International Organization for Standardization, Geneva, Switzerland, Standard, 2010.

[43] "ISO 25178-603, Geometrical product specifications (GPS) – Surface texture: Areal – Part 603: Nominal characteristics of non-contact (phase-shifting interferometric microscopy) instruments," International Organization for Standardization, Geneva, Switzerland, Standard, 2013.

[44] "ISO 25178-604, Geometrical product specifications (GPS) – Surface texture: Areal – Part 604: Nominal characteristics of non-contact (coherence scanning interferometry) instruments," International Organization for Standardization, Geneva, Switzerland, Standard, 2013.

[45] "ISO 25178-605, Geometrical product specifications (GPS) – Surface texture: Areal – Part 605: Nominal characteristics of non-contact (point autofocusing) instruments," International Organization for Standardization, Geneva, Switzerland, Standard, 2014.

[46] "ISO 25178-606, Geometrical product specification (GPS) – Surface texture: Areal – Part 606: Nominal characteristics of non-contact (focus variation) instruments," International Organization for Standardization, Geneva, Switzerland, Standard, 2015.

[47] "ISO/DIS 25178-607, Geometrical product specifications (GPS) – Surface texture: Areal – Part 607: Nominal characteristics of non-contact (confocal microscopy) instruments (under development)," International Organization for Standardization, Geneva, Switzerland, Standard, 2017.

[48] S. F. Ray, *Applied Photographic Optics: Lenses and Optical Systems for Photography, Film, Video, Electronic and Digital Imaging*. Focal, 2002.

[49] E. Reinhard, G. Ward, S. Pattanaik, and P. Debevec, *High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting (The Morgan Kaufmann Series in Computer Graphics)*. Morgan Kaufmann Publishers Inc., 2005.

[50] P. Bariani, L. DeChiffre, H. N. Hansen, and A. Horsewell, "Investigation on the traceability of three dimensional scanning electron microscope measurements based on the stereo-pair technique," *Precision Engineering*, vol. 29, pp. 219–228, 2005.

[51] U. Breitmeier, "Optical follower optimized, technical note," UBM, Tech. Rep., 1990.

[52] S. Carmignato, D. W, and R. Leach, *Industrial X-ray computed tomography*. New York: Springer, 2017.

[53] "Vgstudio max," Tech. Rep., 2017.

[54] A. Townsend, L. Pagani, P. Scott, and L. Blunt, "Areal surface texture data extraction from x-ray computed tomography reconstructions of metal additively manufactured parts," *Precision Engineering*, vol. 48, pp. 254–264, 2017.

[55] A. Townsend, L. Pagani, L. Blunt, P. J. Scott, and X. J. Jiang, "Factors affecting the accuracy of areal surface texture data extraction from x-ray ct," *Ann. CIRP*, vol. 66, pp. 547–550, 2017.

[56] A. Thompson, N. Senin, I. Maskery, L. Körner, S. Lawes, and R. Leach, "Internal surface measurement of metal powder bed fusion parts," *Additive Manufacturing*, vol. 20, pp. 126–133, 2018.

[57] A. Thompson, N. Senin, I. Maskery, and R. Leach, "Effects of magnification and sampling resolution in x-ray computed tomography for the measurement of additively manufactured metal surfaces," *Precision Engineering*, vol. 53, pp. 54–64, 2018.

[58] A. Thompson, N. Senin, C. Giusca, and R. K. Leach, "Topography of selectively laser melted surfaces: A comparison of different measurement methods," *CIRP Annals*, vol. 66, no. 1, pp. 543–546, 2017.

[59] N. Senin, A. Thompson, and R. K. Leach, "Characterisation of the topography of metal additive surface features with different measurement technologies," *Measurement Science and Technology*, vol. 28, no. 9, p. 095 003, 2017.

[60] L. Pagani, Q. Qi, X. Jiang, and P. J. Scott, "Towards a new definition of areal surface texture parameters on freeform surface," *Measurement*, vol. 109, pp. 281–291, 2017.

[61] D. C. Montgomery and G. C. Runger, *Applied Statistics and Probability for Engineers*, 3rd edition. John Wiley & Sons, Inc., 2003.

[62] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing (3rd Ed.): Principles, Algorithms, and Applications*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1996.

[63] I. E. Gordon, *Theories of Visual Perception*. New York: Taylor Francis, Psychology press, 2004.

[64] M. Wertheimer, *Über gestalttheorie*, an address before the Kant Society, Berlin, 7th December 1924, translated by Willis D. Ellis, 1924.

[65] M. E. Mortenson, *Geometric modeling*. Wiley, 1997.

[66] R. Weber, *Ontological Foundations of Information Systems*, ser. Coopers & Lybrand accounting research methodology monograph. Coopers & Lybrand, the Accounting Association of Australia, and New Zealand, 1997.

[67] W. S. Cleveland, "Robust locally weighted regression and smoothing scatterplots," *Journal of the American Statistical Association*, vol. 74, no. 368, pp. 829–836, 1979.

[68] W. S. Cleveland and S. J. Devlin, "Locally-weighted regression: An approach to regression analysis by local fitting," *Journal of the American Statistical Association*, vol. 83, no. 403, pp. 596–610, 1988.

[69] N. Senin, B. M. Colosimo, and M. Pacella, "Point set augmentation through fitting for enhanced icp registration of point clouds in multisensor coordinate metrology," *Robotics and Computer Integrated Manufacturing*, vol. 29, pp. 39–52, 2013.

[70] J. Vaithilingam, M. Simonelli, E. Saleh, N. Senin, R. D. Wildman, R. J. M. Hague, R. K. Leach, and C. J. Tuck, "Combined inkjet printing and infrared sintering of silver nanoparticles using a swathe-by-swathe and layer-by-layer approach for 3-dimensional structures," *ACS Applied Materials & Interfaces*, vol. 9, no. 7, pp. 6560–6570, 2017.

[71] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. 2002.

[73] A. E. Johnson and M. Hebert, "Using spin images for efficient object recognition in cluttered 3d scenes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 5, pp. 433–449, 1999.

[72] A. E. Johnson and M. Hebert, "Surface matching for object recognition in complex three-dimensional scenes," *Image and Vision Computing*, vol. 16, pp. 635–651, 1998.

[74] Y. H. Lin and C. H. Chen, "Template matching using the parametric template vector with translation, rotation and scale invariance," *Pattern Recognition*, vol. 41, pp. 2413–2421, 2008.

[75] E. H. Adelson, C. H. Anderson, J. R. Bergen, P. J. Burt, and J. M. Ogden, "Pyramid methods in image processing," *RCA Engineer*, vol. 29, no. 6, pp. 33–41, 1984.

[76] E. E. H. A. Peter J. Burt, "The laplacian pyramid as a compact image code," *IEEE Transactions on Communications*, vol. 31, pp. 532–540, 1983.

[77] S. Marcelja, "Mathematical description of the responses of simple cortical cells," *Journal of the Optical Society of America*, vol. 70, no. 11, pp. 1297–1297, 1980.

[78] J. G. Daugman, "Two-dimensional spectral analysis of cortical receptive field profiles," *Vision research*, vol. 20, no. 10, pp. 847–56, 1980.

[79] A. K. Jain and F. Farrokhnia, "Unsupervised texture segmentation using gabor filters," *Pattern Recognition*, vol. 24, no. 12, pp. 1167–1186, 1991.

[80] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674–693, 1989.

[81] N. G. Kingsbury, "Complex wavelets for shift invariant analysis and filtering of signals," *Journal of Applied and Computational Harmonic Analysis*, vol. 10, no. 3, pp. 234–253, 2001.

[82] D. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 91, 2004.

[83] A. Malshe, K. Rajurkar, A. Samant, H. N. Hansen, S. Bapat, and W. Jiang, "Bio-inspired functional surfaces for advanced applications," *CIRP Annals - Manufacturing Technology*, vol. 62, no. 2, pp. 607–628, 2013.

[84] A. A. G. Bruzzone, H. L. Costa, P. M. Lonardo, and D. A. Lucca, "Advances in engineered surfaces for functional performance," *CIRP Annals - Manufacturing Technology*, vol. 57, no. 2, pp. 750–769, 2008.

[85] K. J. Stout and L. Blunt, "A contribution to the debate on surface classifications – random, systematic, unstructured, structured and engineered," *International Journal of Machine Tools and Manufacture*, vol. 41, no. 13-14, pp. 2039–2044, 2001.

[86] C. J. Evans and J. B. Bryan, "Structured,textured, or engineered, surfaces," *CIRP Annals-Manufacturing Technology*, vol. 48, no. 2, pp. 541–556, 1999.

[87] D. Surf, "Digital surf mountains map software," Tech. Rep., 2018.

[88] G. Stephanopoulos, *Chemical Process Control: An Introduction to Theory and Practice.* Englewood Cliffs, N.J.: Prentice-Hall, 1984.

[89] D. Pfitzner, R. Leibbrandt, and D. Powers, "Characterization and evaluation of similarity measures for pairs of clusterings," *Knowledge and Information Systems*, vol. 19, no. 3, p. 361, 2008.

[90] J. L. Bentley, "Multidimensional binary search trees used for associative searching," *Commun. ACM*, vol. 18, no. 9, pp. 509–517, 1975.

[91] N. Senin and R. Groppetti, "Application of recursive segmentation and identification techniques to the analysis of the three-dimensional topography of manufactured surfaces," in *Proceedings of the AITEM Conference*, 2005.

[92] "JCGM 200:2012, International vocabulary of metrology - Basic and general concepts and associated terms (VIM)," JCGM, Paris, Standard, 2012.

[93] "JCGM 100:2008, Evaluation of measurement data – Guide to the expression of uncertainty in measurement (GUM)," JCGM, Paris, Standard, 2008.

[94] "JCGM 101:2008, Evaluation of measurement data – Supplement 1 to the "Guide to the expression of uncertainty in measurement" - Propagation of distributions using a Monte Carlo method," JCGM, Paris, Standard, 2008.

[95] A. Bell Stephanie, "A beginner's guide to uncertainty of measurement," National Physical Laboratory, Report, 2001.

[96] H. Haitjema, "Uncertainty in measurement of surface topography," *Surface Topography: Metrology and Properties*, vol. 3, no. 3, p. 035 004, 2015.

[97] S. Yi, R. M. Haralick, and L. G. Shapiro, "Error propagation in machine vision," *Machine Vision and Applications*, vol. 7, pp. 93–114, 1994.

[98] Z. Ling-Xiang and G. Wei-Kang, "Evaluation, evolution, and optimal manipulation of uncertainties in computer vision," in *1997 IEEE International Conference on Intelligent Processing Systems (Cat. No.97TH8335)*, vol. 2, 1473–1477 vol.2.

[99] K. Weißensee, O. Köhn, and G. Linß, "Knowledge-based uncertainty estimation of dimensional measurements using visual sensors," *Journal of Physics: Conference Series*, vol. 238, no. 1, p. 012 023, 2010.

[100] C. L. Giusca, J. D. Claverley, W. Sun, R. K. Leach, F. Helmli, and M. P. J. Chavigner, "Practical estimation of measurement noise and flatness deviation on focus variation microscopes," *CIRP Annals*, vol. 63, no. 1, pp. 545–548, 2014.

[101] L. Giusca Claudiu, K. Leach Richard, F. Helary, T. Gutauskas, and L. Nimishakavi, "Calibration of the scales of areal surface topography-measuring instruments: Part 1. measurement noise and residual flatness," *Measurement Science and Technology*, vol. 23, no. 3, p. 035 008, 2012.

[102] L. Giusca Claudiu, K. Leach Richard, and F. Helery, "Calibration of the scales of areal surface topography measuring instruments: Part 2. amplification, linearity and squareness," *Measurement Science and Technology*, vol. 23, no. 6, p. 065 005, 2012.

[103] L. Giusca Claudiu and K. Leach Richard, "Calibration of the scales of areal surface topography measuring instruments: Part 3. resolution," *Measurement Science and Technology*, vol. 24, no. 10, p. 105 010, 2013.

[104] G. D. MacAulay, N. Senin, C. L. Giusca, and R. K. Leach, "Study of manufacturing and measurement reproducibility on a laser textured structured surface," *Measurement*, vol. 94, pp. 942–948, 2016.

[105] G. Macaulay and C. Giusca, "Assessment of uncertainty in structured surfaces using metrological characteristics," vol. 65, 2016.

[106] G. Musso, N. Senin, M. Galetto, and R. K. Leach, "Towards uncertainty in dimensional metrology of surface features for advanced manufacturing," in *euspen 16th International Conference & Exhibition*, 2016.

[107] G. Casella, C. P. Robert, and M. T. Wells, "Generalized accept-reject sampling schemes," in *A Festschrift for Herman Rubin*, A. DasGupta, Ed., ser. Lecture Notes–Monograph Series. Beachwood, Ohio, USA: Institute of Mathematical Statistics, 2004, vol. Volume 45, pp. 342–347.

Author's publications

This section lists the publications in the field of surface metrology which feature me as primary author or co-author. The publications are listed in reverse chronological order and with separate numbering. Those that have been cited in the thesis are also reported in the main Bibliography section.

The journal articles [J1 - J9, J11] where entirely conceived during the period of the doctorate. The article [J10] despite being published in 2016, was drafted earlier. The book chapters were all published before the doctoral studies. The conference papers [C1 - C11] were conceived and published during the doctorate.

A description of my contributions to each journal paper published during the doctorate is reported at the end of publication list.

Journal papers

[J1] C. Brown, H. N. Hansen, X. J. Jiang, F. Blateyron, J. Berglund, N. Senin, T. Bartkowiak, B. Dixon, G. L. Goic, Y. Quinsat, W. J. Stemp, M. K. Thompson, P. S. Ungar, and E. H. Zahouani, "Multiscale analyses and characterizations of surface topographies," *CIRP Annals*, vol. 67, no. 2, pp. 839–862, 2018.

[J2] N. Senin, A. Thompson, and R. Leach, "Feature-based characterisation of signature topography in laser powder bed fusion of metals," *Measurement Science and Technology*, vol. 29, no. 4, 2018.

[J3] A. Thompson, N. Senin, I. Maskery, L. Körner, S. Lawes, and R. Leach, "Internal surface measurement of metal powder bed fusion parts," *Additive Manufacturing*, vol. 20, pp. 126–133, 2018.

[J4] A. Thompson, N. Senin, I. Maskery, and R. Leach, "Effects of magnification and sampling resolution in x-ray computed tomography for the measurement of additively manufactured metal surfaces," *Precision Engineering*, vol. 53, pp. 54–64, 2018.

[J5] Z. Pala, E. Shaw, J. W. Murray, N. Senin, and T. Hussain, "Suspension high velocity oxy-fuel spraying of tio2: A quantitative approach to phase composition," *Journal of the European Ceramic Society*, vol. 37, no. 2, pp. 801–810, 2017.

[J6] N. Senin, M. Moretti, and R. K. Leach, "Shape descriptors and statistical classification on areal topography data for tile inspection on tessellated surfaces," *Measurement*, vol. 95, pp. 82–92, 2017.

[J7] N. Senin, A. Thompson, and R. K. Leach, "Characterisation of the topography of metal additive surface features with different measurement technologies," *Measurement Science and Technology*, vol. 28, no. 9, p. 095 003, 2017.

[J8] A. Thompson, N. Senin, C. Giusca, and R. K. Leach, "Topography of selectively laser melted surfaces: A comparison of different measurement methods," *CIRP Annals*, vol. 66, no. 1, pp. 543–546, 2017.

[J9] J. Vaithilingam, M. Simonelli, E. Saleh, N. Senin, R. D. Wildman, R. J. M. Hague, R. K. Leach, and C. J. Tuck, "Combined inkjet printing and infrared sintering of silver nanoparticles using a swathe-by-swathe and layer-by-layer approach for 3-dimensional structures," *ACS Applied Materials & Interfaces*, vol. 9, no. 7, pp. 6560–6570, 2017.

[J10] G. D. MacAulay, N. Senin, C. L. Giusca, and R. K. Leach, "Study of manufacturing and measurement reproducibility on a laser textured structured surface," *Measurement*, vol. 94, pp. 942–948, 2016.

[J11] A. Townsend, N. Senin, L. Blunt, R. Leach, and J. Taylor, "Surface texture metrology for metal additive manufacturing: A review," *Precision Engineering*, vol. 46, pp. 34–47, 2016.

[J12] B. M. Colosimo, M. Pacella, and N. Senin, "Multisensor data fusion via gaussian process models for dimensional and geometric verification," *Precision Engineering*, vol. 40, pp. 199–213, 2015.

[J13] G. D. MacAulay, N. Senin, C. L. Giusca, R. K. Leach, and I. Atanas, "Review of feature boundary identification techniques for the characterisation of tessellated surfaces," *Surface Topography: Metrology and Properties*, vol. 3, no. 013002, 2015.

[J14] N. Senin, R. Leach, S. Pini, and L. Blunt, "Texture-based segmentation with gabor filters, wavelet and pyramid decompositions for extracting individual surface features from areal surface topography maps," *Measurement Science and Technology*, vol. 26, no. 9, p. 095 405, 2015.

[J15] D. M. Gavin, S. Nicola, L. G. Claudiu, and K. L. Richard, "Comparison of segmentation techniques to determine the geometric parameters of structured surfaces," *Surface Topography: Metrology and Properties*, vol. 2, no. 4, p. 044 004, 2014.

[J16] N. Senin, G. MacAulay, C. Giusca, and R. K. Leach, "On the characterisation of periodic patterns in tessellated surfaces," *Surface Topography: Metrology and Properties*, vol. 2, no. 2, p. 25 005, 2014.

[J17] N. Senin, M. Moretti, and L. A. Blunt, "Identification of individual features in areal surface topography data by means of template matching and the ring projection transform," *Surface Topography: Metrology and Properties*, vol. 2, no. 1, pp. 14 007–14 007, 2014.

[J18] A. Casoli, P. Cremonesi, C. Isca, R. Groppetti, S. Pini, and N. Senin, "Evaluation of the effect of cleaning on the morphological properties of ancient paper surface," *Cellulose*, pp. 1–17, 2013.

[J19] N. Senin, B. M. Colosimo, and M. Pacella, "Point set augmentation through fitting for enhanced icp registration of point clouds in multisensor coordinate metrology," *Robotics and Computer Integrated Manufacturing*, vol. 29, pp. 39–52, 2013.

[J20] N. Senin, L. A. Blunt, R. K. Leach, and S. Pini, "Morphologic segmentation algorithms for extracting individual surface features from areal surface topography maps," *Surface Topography: Metrology and Properties*, vol. 1, no. 1, pp. 15 005–15 005, 2013.

[J21] N. Senin, L. A. Blunt, and M. Tolley, "Dimensional metrology of micro parts by optical three-dimensional profilometry and areal surface topography analysis," *Proceedings of the Institution of Mechanical Engineers Part B: Journal of Engineering Manufacture*, vol. 226, no. 11, pp. 1819–1832, 2012.

[J22] N. Senin, L. Blunt, and M. Tolley, "The use of areal surface topography analysis for the inspection of micro-fabricated thin foil laser targets for ion acceleration," *Measurement Science and Technology*, vol. 23, no. 10, pp. 105 004–105 004, 2012.

[J23] P. Zurlini, A. Lorenzi, I. Alfieri, G. Gnappi, A. Montenero, S. N, R. Groppetti, and P. Fabbri, "Titanium and zirconium hard coatings on glass subtrates prepared by the sol-gel method," *Thin Solid Films*, vol. 31, pp. 5881–5887, 2009.

[J24] N. Senin, M. Ziliotti, and R. Groppetti, "Three-dimensional surface topography segmentation through clustering," *Wear*, vol. 262, no. 3-4, pp. 395–410, 2007.

[J25] N. Senin, R. Groppetti, L. Garofano, F. Paolo, and M. Pierni, "Three-dimensional surface topography acquisition and analysis for firearm identification," *Journal of Forensic Sciences*, vol. 51, 2006.

[J26] N. Senin and R. Groppetti, "Surface microtopography design and manufacturing through topography descriptors: An application to prosthetic implant surfaces," *Computer Aided Design*, vol. 37, pp. 1163–1175, 2005.

Book chapters

[B1] N. Senin and L. Blunt, "Characterisation of individual areal features," in *Characterisation of Areal Surface Texture*, R. K. Leach, Ed. Heidelberg: Springer, 2013, pp. 179–216.

[B2] N. Senin and G. Campatelli, "Quality inspection of microtopographic surface features with profilometers and microscopes," in *Geometric tolerances: impact on product design, quality inspection and statistical process monitoring*, B. M. Colosimo and N. Senin, Eds. Springer, 2010, pp. 356–372.

[B3] N. Senin, S. Pini, and R. Groppetti, "Identification of microtopographic surface features and form error assessment," in, B. M. Colosimo and N. Senin, Eds. London: Springer, 2010, pp. 159–187.

Conference proceedings

[C1] R. Leach, N. Senin, P. Bointon, X. Feng, S. Lawes, S. Piano, D. Sims-Waterhouse, P. Stavroulakis, R. Su, W. Syam, and M. Thomas, "Information-rich manufacturing metrology," in *Precision assembly technologies and systems*, IFIP International Federation for Information Processing, vol. 372, 2018.

[C2] R. Leach, W. Elmadih, S. Piano, N. Senin, D. Sims-Waterhouse, W. Syam, R. Su, and M. Thomas, "Enriching micro-scale metrology with an all-optical dimensional measuring system," in *Proc. EUSPEN Micro/Nano Manufacturing Workshop*, Strathclyde, UK, Nov. 2017.

[C3] R. Leach, N. Senin, and T. Widjanarko, "Revisiting optical scattering with machine learning," in *16th International Conference on Metrology and Properties of Engineering Surfaces*, Gothenburg, Sweden, Jun. 2017.

[C4] R. Leach, A. Thompson, and N. Senin, "A metrology horror story: The additive surface," in *ASPEN/ASPE Spring Topical Meeting, Manufacture and Metrology of Structured and Freeform Surfaces for Functional Applications*, Hong Kong, China, Mar. 2017.

[C5] L. Newton, N. Senin, and R. Leach, "Focus variation measurement of electron beam melted surfaces," in *16th International Conference on Metrology and Properties of Engineering Surfaces*, Gothenburg, Sweden, Jun. 2017.

[C6] A. Thompson, L. Korner, N. Senin, S. Lawes, R. Leach, and I. Maskery, "X-ray computed tomography of additively manufactured metal parts: The effect of magnification and reconstruction sampling on surface topography measurement," in *16th International Conference on Metrology and Properties of Engineering Surfaces*, Gothenburg, Sweden, Jun. 2017.

[C7] A. Thompson, L. Korner, N. Senin, S. Lawes, I. Maskery, and R. Leach, "Measurement of the internal surface texture of additively manufactured parts by x-ray computed tomography," in *Conference on Industrial Computed Tomography (iCT)*, Leuven, Belgium, Feb. 2017.

[C8] A. Townsend, R. Racasan, P. Bills, A. Thompson, N. Senin, R. Leach, and L. Blunt, "Results from an interlaboratory comparison of areal surface texture parameter extraction from x-ray computed tomography of additively manufactured parts," in *17th Int. EUSPEN Conference*, May 2017.

[C9] G. Musso, N. Senin, M. Galetto, and R. K. Leach, "Towards uncertainty in dimensional metrology of surface features for advanced manufacturing," in *euspen 16th International Conference & Exhibition*, 2016.

[C10] A. Rabani, N. Senin, F. Helmli, P. Butler-Smith, and R. Leach, "An integrated measurement method for complex micro-scale geometries," in *Proceedings of the 16th EUSPEN International Conference*, May 10, 2016.

[C11] A. Thompson, N. Senin, and R. Leach, "Towards an additive surface atlas," in *ASPE Conference on Dimensional Accuracy and Surface Finish in Additive Manufacturing*, Rayleigh, NC, USA, Jun. 2016.

[C12] G. MacAulay, N. Senin, C. Giusca, and R. Leach, "Marker-based watersheds for segmentation of structured surfaces," in *Proceedings of the 15th Conference on Metrology and Properties of Engineering Surfaces*, Charlotte, NC, Mar. 2015.

[C13] M. Moretti, N. Senin, and R. K. Leach, "Geometric registration of areal topography data for comparative surface metrology," in *Proceedings of the AITEM Conference*, Palermo, Italy, 2015.

[C14] G. MacAulay, N. Senin, C. Giusca, and R. Leach, "Investigation of micro-manufacturing process performance via innovative surface characterisation methods," in *11th Laser Metrology for Precision Measurement and Inspection in Industry (IMEKO)*, Japan, Sep. 2014.

[C15] N. Senin, M. Moretti, and L. Blunt, "Identification of individual features in areal surface topography data through template-matching," in *14th international conference on Metrology and Properties of Engineering Surfaces*, Taipei, Taiwan, Jun. 2013.

[C16] N. Senin, L. Blunt, and M. Tolley, "Evaluation of critical dimensions and geometric errors of thin foil laser targets through areal surface topography," in *12th CIRP Conference on Computer Aided Tolerancing*, Apr. 2012.

[C17] P. Zurlini, A. Lorenzi, I. Alfieri, G. Gnappi, A. Montenero, S. N, R. Groppetti, and P. Fabbri, "Hard coatings on glass substrates by sol-gel method," in *AIAr 2009*, Ravenna, 2009.

[C18] N. Senin and R. Groppetti, "Automated identification and extraction of dale-like features from the three-dimensional topography of manufactured surfaces: An application to scratch testing," in *Proceeding of the 8th AITeM conference*, Montecatini terme, Italy, 2007.

[C19] ——, "Application of recursive segmentation and identification techniques to the analysis of the three-dimensional topography of manufactured surfaces," in *Proceedings of the AITEM Conference*, 2005.

[C20] R. Groppetti, M. Malpeli, A. Scrivani, and N. Senin, "Contribution to the development and characterization of biomimetic and bioactive hydroxyapatite coating for orthopaedic and dental implants," in *Proc. of the 9th annual seminar and meeting: Ceramics Cells and Tissues*, Faenza, Italy, 2004.

[C21] N. Senin and R. Groppetti, "On the application of segmentation techniques to three-dimensional surface topography analysis of engineered surfaces," in *4th CIRP International Seminar on Intelligent Computation in Manufacturing Engineering (ICME)*, Sorrento, Italy, 2004.

[C22] R. Groppetti and N. Senin, "Bioceramic materials surface topography measurement and analysis for biomedical applications," in *8th Annual Seminar and Meet-*

ing: Ceramics Cells and Tissues, Faenza, Italy: Consiglio Nazionale delle Ricerche (CNR), 2003.

[C23] ——, "Contribution to the development of three-dimensional nano and micro-topography measurement and analysis techniques and systems," in *11th International Metrology Congress*, Montpellier Cedex 3/France: Collège Français de Métrologie, Toulon, France, 2003.

[C24] S. Nicola and R. Groppetti, "On three-dimensional surface topography analysis process planning and documenting for engineering applications," in *6th AITeM International Conference*, Gaeta, Italy, 2003.

Contribution to the journal papers published during the doctorate

J1 Contributed to: paper structure, text editing, definitions of scale, multi-scale analysis, scale-limited measurement. Written the section on segmentation.

J2 Idea, methods, implementation of algorithms, execution and analysis of results. Primary writer.

J3 Methods ideation and algorithmic implementations for topography comparison and feature-based characterisation. Participation to ideation and planning of the research project. Draft editing.

J4 Method ideation and algorithmic implementation for topography comparison and feature-based characterisation. Participation to ideation and planning the research project. Draft editing.

J5 Selection of existing methods, execution and analysis of the results for part on surface topography characterisation. Written the corresponding sections (methods and results) of the draft.

J6 Idea, methods, algorithmic implementation, execution and analysis of results. Primary writer.

J7 Idea, methods, algorithmic implementation. Collaborated to the analysis of results. Primary writer.

J8 Idea, methods, algorithmic implementation. Collaborated to the analysis of the results. Draft editing.

J9 Method ideation, algorithmic implementation, execution and analysis of the results for the part of surface topography analysis. Written the corresponding sections (methods and results) of the draft.

J11 Contributed to paper structure, text editing.