ONLINE QUALITY MONITORING OF 3D SURFACE TOPOGRAPHY

By

MEJDAL A. ALQAHTANI

A dissertation submitted to the

School of Graduate Studies

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Industrial and Systems Engineering

Written under the direction of

Elsayed A. Elsayed and Myong K. Jeong

And approved by

____

____

____

____

____

New Brunswick, New Jersey

October, 2020

ABSTRACT OF THE DISSERTATION

Online Quality Monitoring of 3D Surface Topography

By MEJDAL A. ALQAHTANI

Dissertation Directors:

Elsayed A. Elsayed and Myong K. Jeong

In many real-life applications, three-dimensional (3D) surface topography contains rich information about products and manufacturing processes. Different faults commonly appear on the topography of finished products in local and global patterns during manufacturing. Such faults are likely to cause changes in the variance or autocorrelation of the topographic values. Monitoring such changes is challenging due to the unique properties of the topographic surfaces. In particular, the topographic values are spatially autocorrelated with their neighbors and their locations are randomly changing from one surface to another under normal process behavior. The existing online monitoring approaches for the 3D surface topography lack the detection and diagnoses of changes in the topographic surfaces. In this dissertation, we investigate and develop four different online monitoring approaches to accurately characterize, detect, and diagnose various changes in topographic surfaces.

In the first approach, we develop a multi-level spatial randomness approach for online monitoring of global changes the surfaces. We propose a multi-level surface thresholding

algorithm for improving the representation of surface characteristics in which an observed surface topography is sliced into different levels in reference to the characteristics of normal surfaces. The spatial statistical dependencies of surface characteristics at each surface level are accurately captured through a proposed spatial randomness (SR) profile. We then develop an effective monitoring statistic based on the functional principal component analysis for identifying anomaly surfaces with global changes based on their SR profiles.

In the second approach, we propose a multi-label separation-deviation surface model for effective monitoring of local variance changes in 3D topographic surfaces. The approach improves the representation of local topographic changes through a developed multi-label separation-deviation surface (MSS) model, which labels the important surface characteristics and smoothes out the noisy characteristics. We also propose two effective features for monitoring changes in surface characteristics. The MSS feature is introduced for capturing deviations within the label assignments, and the generalized spatial randomness feature is derived for quantifying deviations between the label assignments. These two features are integrated into a single monitoring statistic to detect local variations in topographic surfaces.

In the third approach, we develop a novel approach based on graph theory for accurate monitoring of local autocorrelation changes in 3D topographic surfaces. We enhance the representation of surface characteristics by proposing an in-control multi-region surface segmentation algorithm, which segments the observed surface pixels into clusters according to the information learned from in-control surfaces. The local and spatial topographic characteristics are accurately described through a developed maximum local spatial randomness feature. After representing the surface as a spatially weighted graph, we monitor its connectivity through the developed spatial graph connectivity statistic for accurate detection of local autocorrelation changes in topographic surfaces.

In the fourth and final approach, we investigate a generalized spatially weighted autocorrelation approach for fault detection and diagnosis in 3D topographic surfaces. We develop two algorithms to identify and assign spatial weights to the suspicious topographic regions. The normal surface "hard" thresholding algorithm initially enhances the representation of surface characteristics through binarization, followed by the normal surface connected-component labeling algorithm, which utilizes the obtained binary results to identify and assign spatial weights to the regions with suspicious characteristics. We also develop a generalized spatially weighted Moran (GSWM) index, which exploits the assigned weights to effectively monitor and detect changes in the spatial autocorrelation of each identified region. After an anomaly surface is detected based on its GSWM index, we accurately extract different fault information such as fault size, type, location, magnitude, and the number of faults.

The proposed approaches are validated for their effectiveness, efficiencies, and performance for online monitoring and diagnosis of various changes in 3D topographic surfaces.

ACKNOWLEDGEMENTS

First of all, I would like to present and acknowledge my heartfelt appreciation to my advisors, Professor Elsayed A. Elsayed and Professor Myong K. Jeong, for their excellent guidance and encouragement during my dissertation and research. I am very honored to work with them and been accepted as a member of their academic family. In particular, I am genuinely grateful to Professor Elsayed A. Elsayed, who has provided me extensive personal and academic guidance and taught me a great deal about both scientific research and life in general. I also deeply thank Professor Myong K. Jeong for his sincere support and motivation during all stages of my Ph.D. study. His insightful advices pushed me to sharpen my thinking and brought my work to a higher level. I also would like to thank my committee members, Professor Hoang Pham, Professor Weihong Guo, and Professor Howon Lee, for their insightful and helpful comments and encouragement.

My sincere thanks go to my father, Abdullah, and my mother, Haya, who raised me with such love and with the belief that anything is possible with hard work and dedication. I also would like to express my deep gratitude to my brothers and sisters for supporting me through my studies. Particularly, I would like to thank my great brother, Turki, for his constant support, help, and belief in me.

I would not have been able to finish my Ph.D. degree without the endless support from my wife, Aljawharah, and my wonderful daughters, Haya and Jana, which provided me their constant encouragement, love, and sacrifice. I also want to extend my gratitude to my relatives, especially my uncle, Mubarak, and my mother-in-law, Haya, for their support and prayers.

I want to acknowledge all my friends at Rutgers University for being great colleagues and friends. Especially, I want to thank my faithful friends, Mehmet Turkoz and Sangahn Kim, for their brotherhood and continuous encouragement. I also want to thank my friends, Jaeseungn Baek, Wassem Alemari and Taha Alhindi, for their support and sincere help.

I would like to thank my friends at King Saud University. Especially, I would like to thank my close friend, Abdulmajeed Aljardan, for his sincere brotherhood and constant inspiration. Besides, I would like to thank my friends, Saad Alabdulkarim and Abdullah Almassad, for their support and encouragement.

I would like also to acknowledge King Saud University and Saudi Arabian Cultural Mission for their generous and prestigious scholarship, which enabled me to pursue my Ph.D. degree.

Most importantly, I thank my God for answering my prayers and giving me the strength to complete my Ph.D. degree despite all difficulties and challenges that I have gone through, especially during the COVID-19 pandemic.

Table of Contents

List of Tables

List of Figures

CHAPTER 1

INTRODUCTION

1.1 Importance of Surface Topography

Surface topography, also known as surface roughness, or surface finish, is a key quality characteristic of products in many industries. For instance, in the biomedical industry, when the surface topography of implanted artificial organs joints, such as hip joint, knee replacement, and spine disc, is “rough,” it may result in the acceleration of their wear rates, which may cause unexpected failure (Scholes and Unsworth, 2009, Rengier *et al.*, 2010). Moreover, in the aerospace industry, titanium alloys, which are commonly used in the manufacture of aircraft engines and airframes, require extremely “smooth” surfaces for maintaining high strength and toughness (Ramesh *et al.*, 2008, Aspinwall *et al.*, 2008, Zhang *et al.*, 2017, Liang *et al.*, 2019). Another example from the automotive industry is the piston rings, which are designed to seal the combustion chamber for minimizing the fuel consumption in a car engine. In practice, the piston rings require “smooth” surfaces due to the exposure of lubricants during the combustion process (Brown and Blunt, 2008, Söderfjäll *et al.*, 2017). However, when the surface topography of the piston rings is “rough”, it results in abrasive wear, i.e., the erosion of the material from its original form, which may result in a lower performance of the car engine. Figure 1.1 shows the three-dimensional (3D) surface topography of unworn and worn piston rings measured by atomic force microscope (AFM). Note that the unworn piston ring has a “smooth” surface as shown in Figure 1.1 (a), while the worn piston ring has a “rough” surface as shown in Figure 1.1 (b).

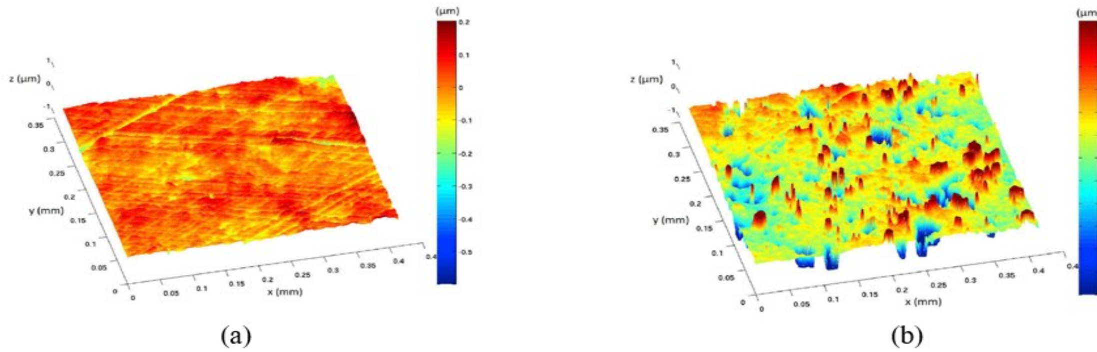
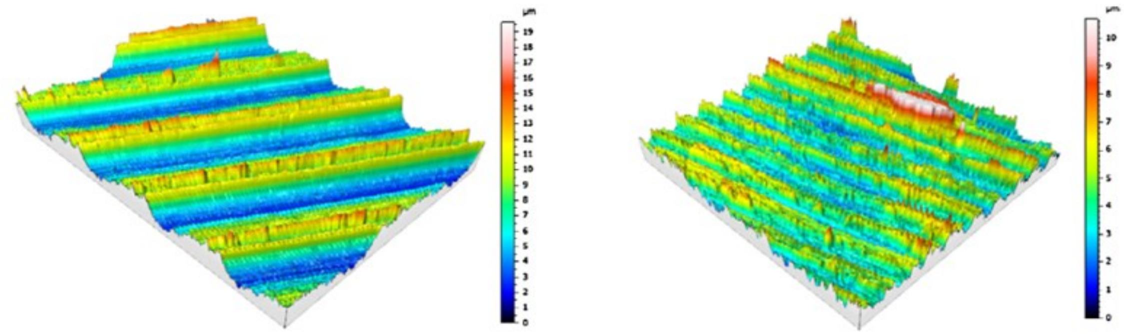


Figure 1.1 3D surface topography of (a) unworn piston ring with a "smooth" surface, and (b) worn piston ring with a "rough" surface (Söderfjäll *et al.* , 2017).

Surface topography also has a major impact on the strength of the material and potential crack propagation. For example, in additive manufacturing processes, studies of as-built 3D printed parts show that the surface roughness can decrease the fatigue strength of finished parts significantly (Wycisk *et al.*, 2014, Jomaa *et al.*, 2014). In particular, Wycisk *et al.* (2014) study the effect of the surface roughness on the fatigue life of laser additive manufactured titanium alloy Ti-6Al-4V. This study shows that after smoothing the surface roughness of as-built parts through a post-processing technique such as polishing, a significant enhancement in the part properties is achieved. More specifically, polishing can introduce compressive residual stresses that improve crack propagation resistance. Polishing is also found to improve the fatigue strength (endurance limit is improved from 210 MPa for non-polished parts to 500 MPa for polished parts) (Wycisk *et al.*, 2014).

Another importance of surface topography is found in the assembly processes where the surface topography has a strong influence on the fitting accuracy of the system components. More specifically, when the finished components contain surface defects, such as ridges, pits, and scratches, they are likely to be assembled inappropriately, which may result in a significant reduction in the functional integrity of system components (Yang *et al.*, 2018). For example, mating surfaces, which are welded together for assembling the vehicle frames in the automotive industry, demand highly "smooth" surfaces (Brown and Blunt, 2008, Yang *et al.*, 2018). This is because inappropriate assembly can accelerate the wear rate between the welded surfaces, which may yield a low performance of the vehicle frame.

Surface topography not only holds information about the surface properties of a product but also holds important information about the manufacturing processes. Particularly, monitoring surface defects, such as cracks, scratches, and bubbles, may assist in detecting shifts in the manufacturing processes and subsequently in identifying the assignable causes of process shifts, such as operator error, tool wear, and defective material (Korkut and Donertas, 2007). Appropriate corrective actions can accordingly be implemented to eliminate these causes. Krolczyk *et al.* (2016) show that the shift in the process parameters in the milling operation, such as the feed rate and cutting speed, affects the structure of the surface topography of finished products. In particular, Figure 1.2 displays the 3D surface topography of two high-strength steels after being exposed to two milling process conditions: the normal milling process where the process parameters do not experience any shift from the nominal values, and the abnormal milling process where the process parameters are shifted. It is observed that the structure of 3D surface topography is considerably changed as shown in Figure 1.2 (a and b).



(a) (b)

Figure 1.2 3D surface topography of two high-strength steels machined via the milling process with (a) no shift in the process parameters, and (b) a shift in the process parameters (Krolczyk *et al.* , 2016).

1.2 Research Background and Challenges

The characteristics of topographic images are typically composed of a complex structure, which makes the monitoring of surface changes challenging. In particular, topographic images of units produced sequentially using the same manufacturing process exhibit high randomness in the locations of their characteristics such as peaks and valleys, although the process operates under normal conditions (Bui and Apley, 2018a, Bui and Apley, 2018b). This makes the monitoring and diagnoses of changes in the topographic image of the new units challenging. Figure 1.3 illustrates the random behavior among image characteristics. In particular, we generate three samples of topographic images with a fixed setting using the Gaussian random model, which is effective in simulating topographic images similar to those obtained during conventional manufacturing processes such as milling, cutting, and turning processes (Garcia and Stoll, 1984, Rao *et al.*, 2015b). We notice that the image characteristics in terms of peaks (red color) and valleys (blue color) occur randomly in different locations for all the three images.

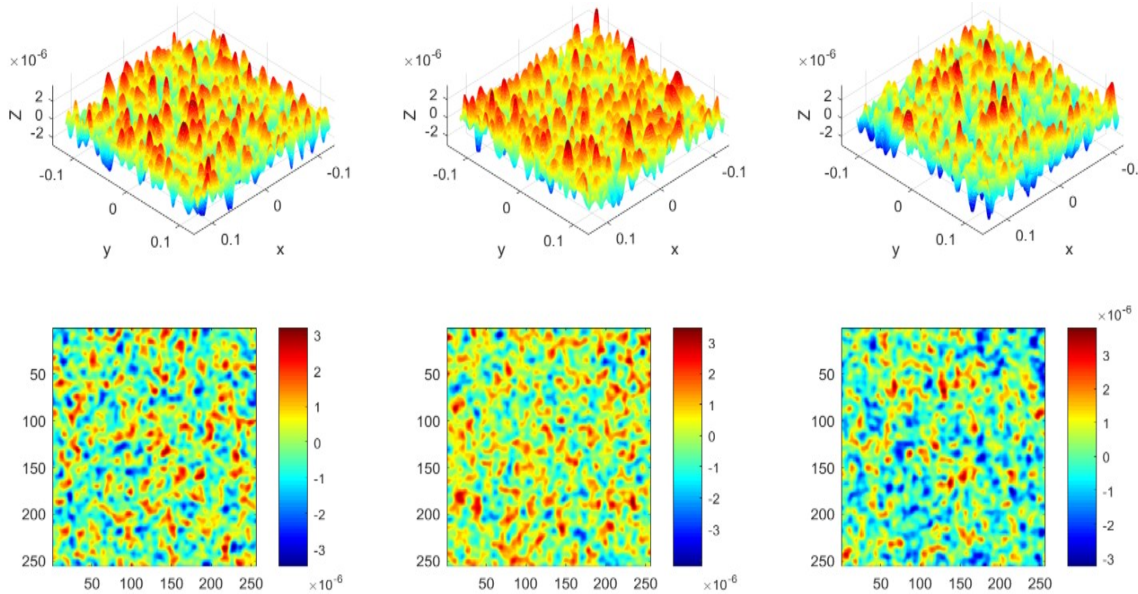


Figure 1.3 3D views of three generated topographic images along with their corresponding 2D views.

Besides, topographic surfaces show high autocorrelation among their close neighboring pixels than the remote neighbors (Jeong *et al.*, 2008, Alqahtani *et al.*, 2020c, Alqahtani *et al.*, 2020b). Subsequently, topographic pixels located at nearby locations have similar values (or colors) since these pixels share similar characteristics as shown earlier in Figure 1.3. In addition, it is expected that the characteristics of the neighborhood of each pixel to be statistically similar under normal process behavior (Bui and Apley, 2018a, Alqahtani *et al.*, 2020c, Alqahtani *et al.*, 2020d). Thus, any change in the characteristics between the topographic values and their neighbors implies the appearance of surface faults (or changes) (Jeong *et al.*, 2008). Therefore, it is critically important to analyze the spatial behavior embodied in the structure of the topographic values for the effective detection of surface faults.

Another challenge in surface monitoring is that there are a variety of surface defects that can be observed in practice, such as broken edges, uneven surfaces, burrs, micro-cracks, and scratches (Yang et al., 2018). These defects may cause two major changes to the characteristics of the topographic values: the variance and autocorrelation changes. The variance change is defined as the deviation of topographic values from the reference plane (or zero plane), whereas the autocorrelation change is defined as the deviation of the autocorrelation structure between the topographic values from the reference structure (Jiang et al., 2007, Jiang and Whitehouse, 2012). Examples of defects that cause the variance change include cracks and pinholes that emerge on the surface topography of manufactured metals (Jolic *et al.*, 1994). In addition, examples of defects related to the autocorrelation change are scratches, pits, and ridges that frequently arise on the surface topography of semiconductor wafers (Rao *et al.*, 2015b).

Figure 1.4 shows examples of normal and two anomaly surfaces with variance and autocorrelation changes, respectively. In Figure 1.4 (b and e), we observe that the variance change mainly impacts the magnitude of topographic values on the z-axis in a way that the topographic values show more deviations from the reference plane, which results in valleys and peaks with darker colors compared to the ones of the normal surface. Moreover, in Figure 1.4 (c and f), we notice that the autocorrelation change mainly affects the structure of topographic values on the x-axis and the y-axis such that the topographic values tend to be more similar to their neighbors, which yields wider valleys and peaks. In general, characterizing and detecting the autocorrelation change is more difficult than the variance change. This is because the autocorrelation change does not impact the magnitude of topographic values, which can be within normal values, but it mainly impacts the spatial relationships between the topographic values, which are difficult to detect. Since these two changes can be observed during manufacturing, developing effective approaches for detecting these two changes is needed.

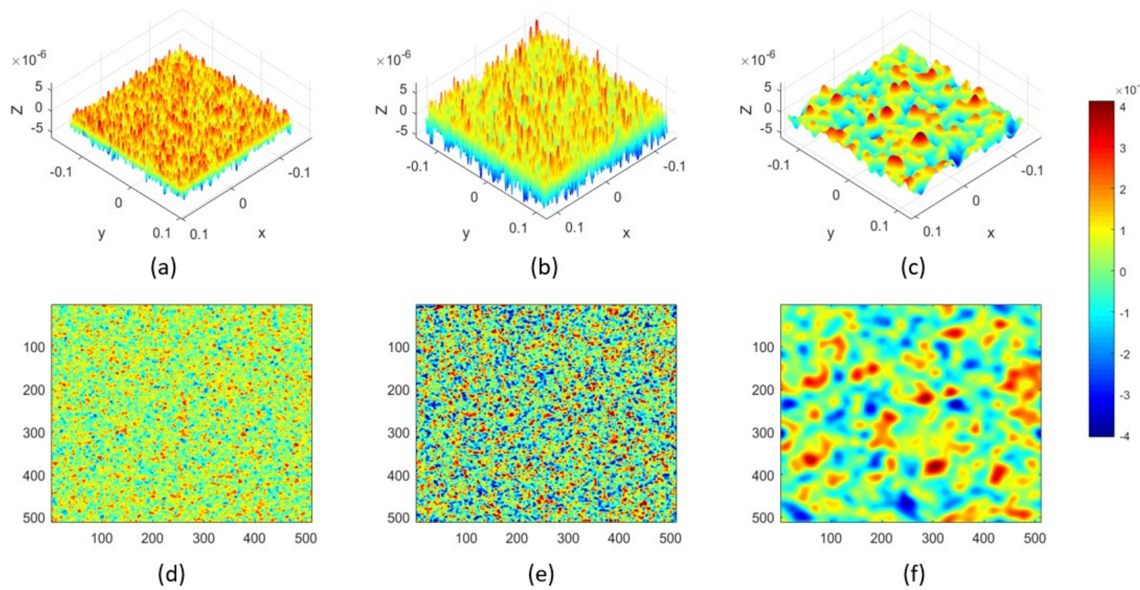


Figure 1.4 Examples of 3D topographic surfaces and their corresponding 2D plan views of (a and d) normal, and two anomaly surfaces with (b and e) variance change, and (c and f) autocorrelation change.

Finally, the locality of surface faults (or change) is a critical issue that should also be addressed in monitoring 3D topographic surfaces (Bui and Apley, 2018a, Alqahtani *et al.*, 2020c, Alqahtani *et al.*, 2020d). In practice, surface faults may occur in a local area where a small portion of a surface contains surface faults, i.e., less than say 10% of the total surface area is changed from the normal surface characteristics, or in a global area where a large portion of the surface is contaminated with faults, i.e., 10% or more of the total surface area is changed. In general, identifying local faults is more difficult than global faults since most of the topographic values of a surface with local faults do not experience a significant change from the norm, which makes the characteristics of local faults to be dominated by the normal characteristics during the assessment process. Figure 1.5 shows the plan view of three simulated anomaly surfaces that include two local faults with 1% and 5% changes and one global fault with a 10% change. Therefore, detecting both local and global changes is important for effective monitoring of changes in topographic surfaces.

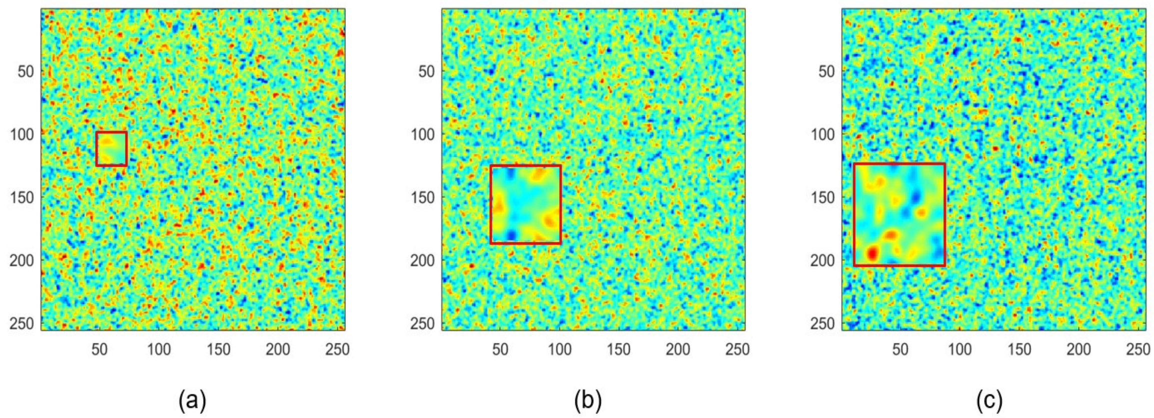


Figure 1.5 Examples of anomaly surfaces with two local faults with (a) 1%, and (b) 5% changes, and one global fault with (c) 10% change.

1.3 Problem Description

Various surface monitoring approaches have been proposed for assessing the quality of products based on their topographic images in the last two decades. Simple monitoring approaches such as the average surface roughness and the root mean square roughness are commonly used to detect changes in topographic surfaces. In addition, advanced monitoring approaches based on signal processing, image processing, graph theory, and machine learning have recently been developed for assessing the characteristics of 3D surface topography. Due to the complex characteristics of topographic surfaces discussed in Section 1.2, the existing monitoring approaches lack the ability to detect local and global changes in the variance and autocorrelation of topographic values. Therefore, in this dissertation, we overcome these limitations by developing different monitoring approaches based on various statistical theories for accurate and efficient detection and diagnosing of different forms of changes in 3D topographic surfaces.

In this dissertation, the input data are a 3D topographic surface (or image), which is composed of a finite number of pixels that are measured from different sites of the surface. In particular, the 3D topographic surface (or image) can be expressed as $Z = \{z_i : i = 1, 2, ..., M\}$, where z_i is the topographic value (or height) of the pixel i and M is the total number of the surface pixels. Therefore, for the given 3D surface topography of N normal products $Z^{(j)}$, $j = 1, 2, ..., N$, the goal of this dissertation is to determine whether the 3D surface topography of a new product $Z^{(new)}$ is an anomaly. In addition, after $Z^{(new)}$ is identified as an anomaly, we extract different fault information, such as fault size, fault type, fault magnitude, fault location, and the number of surface faults.

1.4 Proposed Approaches for Online Monitoring of 3D Surface Topography

In this dissertation, we develop four online monitoring approaches based on different statistical theories for detecting various forms of surface changes (or faults) in 3D topographic surfaces. The proposed approaches are briefly stated in the following sections.

1.4.1 Multi-level Spatial Randomness Approach

We develop a multi-level spatial randomness approach for online monitoring of global changes in 3D topographic surfaces. We propose a multi-level surface thresholding algorithm for enhancing the representation of surface characteristics. The algorithm efficiently slices the 3D surface topography of a new product into different levels in accordance with the characteristics of normal products. The characteristics of the topographic values are then accurately quantified at each surface level through a developed spatial randomness (SR) profile. We then present an effective monitoring statistic based on the functional principal component analysis for detecting anomaly surfaces with different patterns of global changes.

1.4.2 Multi-label Separation-deviation Surface Model

We propose a multi-label separation-deviation surface (MSS) model for effective monitoring of local variance changes in 3D topographic surfaces. The MSS model is effective in enhancing the representation of local topographic defects by segmenting the topographic values into predefined labels. The segmentation results are obtained by minimizing the deviation of the topographic values from their expected values and the separation between the topographic values. This results in an effective representation of surface characteristics where the critical surface characteristics are labeled, and the noisy characteristics are smoothed out. We introduce two effective features for capturing changes in surface characteristics. The MSS feature is presented for quantifying variations within the label assignments, and the generalized spatial randomness (GSR) feature is derived for quantifying deviations in the spatial autocorrelation between the label assignments. The MSS and GSR features are integrated into a single monitoring statistic to detect local variations in topographic surfaces.

1.4.3 Spatially Weighted Graph Theory-based Approach

We present a novel approach based on graph theory for monitoring local autocorrelation changes in 3D topographic surfaces. The representation of surface characteristics is improved by proposing an in-control multi-region surface segmentation algorithm, which segments the observed surface pixels into clusters according to the information learned from in-control surfaces. The algorithm also divides the surface into spatial regions for an effective description of spatial topographic characteristics. The local surface characteristics among obtained clusters are effectively captured through a developed maximum local spatial randomness feature. We then represent the relationships between the characteristics of the new and in-control surfaces as a spatially weighted graph network and subsequently monitor its connectivity through the spatial graph connectivity statistic for accurate detection of local autocorrelation changes in topographic surfaces.

1.4.4 Generalized Spatially Weighted Autocorrelation Approach

We investigate a generalized spatially weighted autocorrelation approach for accurate fault detection and diagnosis in 3D topographic surfaces. The proposed approach utilizes the information learned from normal surfaces to identify and assign spatial weights to the regions with suspicious characteristics of new surfaces. We develop two algorithms to obtain the weight assignments: the normal surface "hard" thresholding algorithm, which improves the representation of surface characteristics through binarization, and the normal surface connected-component labeling algorithm, which exploits the obtained binary results to identify and assign spatial weights to the suspicious topographic regions. We also develop the GSWM index, which exploits the assigned spatial weights to efficiently describe and monitor the spatial autocorrelation structure of each identified region. When an anomaly surface is detected based on its GSWM index, we accurately extract different fault information such as fault location, type, size, magnitude, and the number of faults.

1.5 Dissertation Organization

This dissertation is organized as follows. A comprehensive review, which covers the existing surface metrology systems, the input topographic data structure, the surface simulation model, and the existing surface monitoring approaches, is provided in Chapter 2. In Chapter 3, the multi-level spatial randomness approach is introduced for monitoring global surface changes. Chapter 4 presents the multi-label separation-deviation surface model for monitoring local variance changes. In Chapter 5, the spatially weighted graph theory-based approach is introduced for monitoring local autocorrelation changes. Chapter 6 presents the generalized spatially weighted autocorrelation approach for fault detection and diagnosis. Finally, in Chapter 7, we discuss conclusions and recommendations for future research.

CHAPTER 2

LITERATURE REVIEW

In this chapter, we present a comprehensive review of the research being addressed in this dissertation. We review the recent and common surface metrology systems that are applied for measuring the surface topography of machined parts in Section 2.1. In Section 2.2, we explain the structure of topographic data. We then discuss the simulation model commonly used to generate topographic surfaces in Section 2.3. Finally, in Section 2.4, we review the existing monitoring approaches that are implemented for assessing the quality of products based on their topographic surfaces.

2.1 Surface Metrology Systems

With the emergence of modern technologies, various surface metrology systems have been developed by adopting a wide range of principles and scientific theories. Surface metrology systems are categorized into two types: contact-type and non-contact-type techniques. In the contact-type technique, the surface measurements are obtained by a direct contact between the surface of a workpiece and a detector tip of the measurement tool. For example, the stylus profilometer is one of the most popular tools for measuring the surface topography of products. The stylus profilometer contains a stylus tip, which traces the target surface and electrically identifies the surface measurements based on the observed vertical motion of the stylus tip (Ali, 2012). In addition, scanning probe microscopes (SPMs), such as atomic force microscope (AFM) and scanning tunneling microscope (STM), are a family of the contact-type microscopes that can obtain the 3D images of surface topography at the nanoscale level. More Specifically, the AFM has a small probe, which detects the surface measurements by generating a force between the target surface and the probe (Ali, 2012). Similarly, the STM uses a physical probe that obtains the surface measurements by producing a tunneling current between the target surface and the probe (Townsend *et al.*, 2016).

Non-contact-type surface metrology systems have also been increasingly applied for measuring the topography of many manufactured products. These systems are mainly designed to obtain the surface measurements by observing the reflective properties of the measured parts without physically touching their surfaces (Blunt and Jiang, 2003). For instance, the 3D laser (or optical) microscope is one of the most popular non-contact-type metrology systems in which a laser beam is emitted to the surface sample, and the reflected beam is then transformed into surface measurements (Rao *et al.*, 2015b, Tootooni *et al.*, 2016). In addition, the scanning electron microscope (SEM) uses the reflection of the electron beam for obtaining the measurements of the target surface (Townsend *et al.*, 2016). Another common non-contact-type metrology system is the white light interferometer (WLI), which uses the reflection of white light waves for detecting the surface measurements (Ali, 2012).

Each type of the aforementioned metrology systems has its advantages and drawbacks. In particular, the contact-type metrology systems have some advantages in terms of its capability of long-distance and high-resolution measurements. However, the use of these systems may cause damage to the surface of the measured workpiece, such as scratches and holes. In addition, when the width of the surface characteristics such as peaks and valleys is smaller than the radius of the detector tip, these characteristics can be ignored and misidentified (Bhushan, 2000). More important, these systems are generally time-consuming due to the setup time for positioning the detector tip and the time for processing the surface measurements, which makes them inappropriate to work under the online monitoring scheme (Townsend *et al.*, 2016). In contrast, the non-contact-type metrology systems are more efficient in acquiring the surface measurements due to the utilization of advanced vision sensors, which makes them more suitable to implement in the online monitoring environment (Townsend *et al.*, 2016). Besides, the measured surface of a workpiece is protected from any damages since these systems do not physically contact with the workpiece. However, the non-contact-type systems show some drawbacks in terms of its high cost and its limitation in measuring the surface characteristics at the nanoscale level (Ali, 2012). Table 2.1 summarizes the advantages and drawbacks of the foregoing surface metrology systems.

Table 2.1 Summary of the advantages and drawbacks of surface metrology systems

N	Type	Metrology systems	Advantages	Drawbacks
1	Contact	Stylus profilometer	Clear wave profile- capable of long-distance- low cost	Time-consuming- can cause scratches on the measured surface- tip wear
2	Contact	AFM- STM	Capable of 3D measurements- high resolution	Time-consuming- not suitable for measuring large sample areas- difficulties in positioning the tip
3	Non-contact	3D laser microscope- 3D optical microscope	Capable of 3D measurements- efficient measurement	Limited to certain surfaces
4	Non-contact	WLI	Capable of 3D measurements- efficient measurement	Limited to certain surfaces- sensitive to vibrations
5	Non-contact	SEM	Easy operation- efficient measurement	Expensive- required more maintenance

2.2 Topographic Data Structure

In this dissertation, the input data are a 3D topographic surface (or image), which is expressed as $Z = \{z_i; i = 1, 2, ..., M\}$, where z_i represents the height information of the pixel i and M is the total number of the surface pixels. Each pixel, $z_i \in Z$, can be a positive or negative value where the positive value represents the positive deviation from the reference plane (i.e., zero plane), while the negative value represents the negative deviation. Figure 2.1 illustrates an example of a generated 3D topographic surface Z.

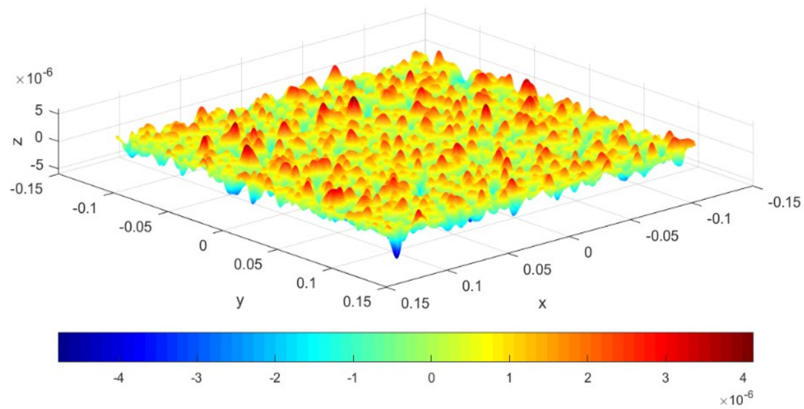


Figure 2.1 Sample of 3D surface topography associated with valleys (blue color), the zero or reference plane (green color), and peaks (red color).

Some existing approaches consider a simple representation of the original topographic image, such as grayscale or binary images. In particular, the surface image can be converted to a grayscale image $G = \{g_i; i = 1, 2, ..., M\}$, where g_i is the grayscale intensity level of the pixel i. The grayscale image is obtained by mapping the topographic values into the grayscale intensity scale, which ranges from "0" (black color) that represents the valley characteristics to "255" (white color) that represents the peak characteristics. Furthermore, the surface image Z can also be converted to a binary image $B = \{b_i; i = 1, 2, ..., M\}$, where

b_i is the assigned binary value of the pixel i. The binary image B is obtained by applying the edge detection filters, such as a Canny filter, in which each pixel z_i is assigned either "0" (black color) that displays the valley characteristics or "1" (white color) that displays the peak characteristics. Figure 2.2 shows the top view of a generated 3D topographic surface and its corresponding grayscale and binary images.

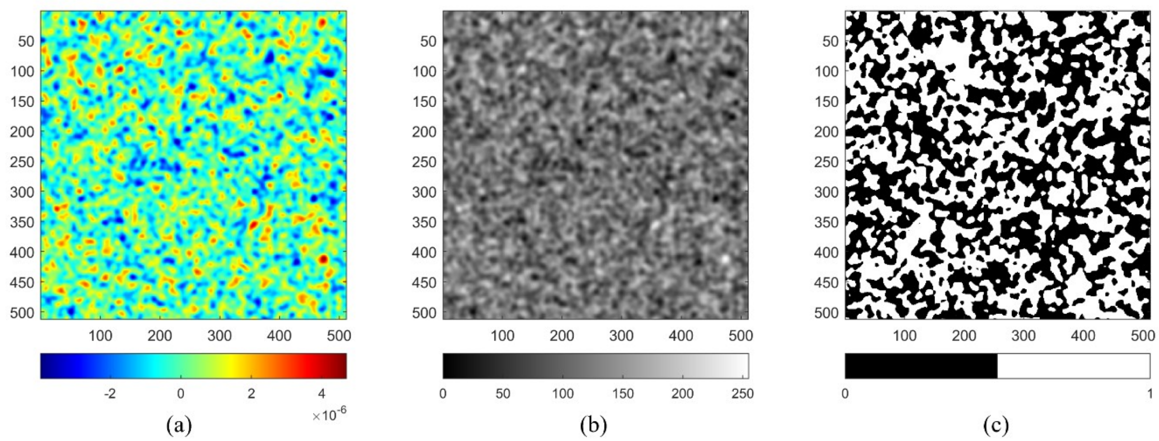


Figure 2.2 Example of (a) a generated topographic surface Z, and its corresponding (b) grayscale image G, and (c) binary image B.

2.3 Surface Simulation Model

The Gaussian random model is commonly used for simulating topographic surfaces similar to those obtained during conventional manufacturing processes such as milling, planning, and cutting processes (Garcia and Stoll, 1984). The simulation model is explained as follows. An uncorrelated Gaussian set is generated $Q = \{q_i; i = 1, 2, ..., M\}$, where M is the total number of the generated pixels and q_i is an independent and identically distributed random variable that has a Gaussian distribution with a mean of zero and a defined standard deviation σ. The Gaussian autocorrelation set C is then obtained based on the Gaussian autocovariance function $C=\{c_i; i=1,2,...,M\}$, where c_i represents the magnitude of the autocorrelation of the pixel i, which is defined as

$$c_i = \exp\left(-[(x_i - x_j)^2 + (y_i - y_j)^2]/[2\times\theta^2]\right),$$

where θ is a scale parameter, (x_i, y_i) is the coordinate of the pixel i, and (x_j, y_j) is the coordinate of the j^{th} neighbor of the pixel i. Next, we convolute the set Q with the set C using the fast Fourier transforms (FFT) algorithm to attain the Gaussian autocorrelated surface as shown in Equation (2.1)

$$Z = (2\times L)/(\sqrt{\pi}\times\sqrt{M}\times\theta)\times F\left[F\left(G\right)\times F(C)\right]^{-1}, \tag{2.1}$$

where L is the length of the topographic surface (e.g., $250\ \mu m$) and $F(.)$ is the FFT of given set. The FFT is applied due to its efficiency in computing convolution. Thereby, the Gaussian autocorrelated surface is obtained $Z=\{z_i; i=1,2,...,M\}$, where M is the size of simulated surface values (e.g., $M=256\times256$). Note that σ controls how strongly the topographic values deviate as the squared distance between the topographic values and their zero mean (the reference plane) increases. Thus, smaller σ yields smaller variation from the reference plane, whereas larger σ yields larger variation. In addition, θ controls how rapidly the correlation decays as the squared distance between the neighboring pixels increases. Therefore, smaller θ leads to fast decay and less autocorrelated topographic values, while larger θ leads to slow decay and more autocorrelated values.

2.4 Review of Existing Monitoring Approaches

Various monitoring approaches have been developed for assessing the "quality" of products based on their topographic surfaces. Most of these approaches extract and monitor a small set of predefined features for detecting anomaly surfaces. These features are extracted based on different statistical approaches, such as topographic distribution, signal processing, image processing, graph theory, and machine learning. In the following section, we review the popular features in each approach and address their advantages and drawbacks. In particular, we evaluate the existing features based on its effectiveness in detecting local and global changes in the variance and autocorrelation of topographic values discussed earlier in Section 1.2.

2.4.1 Features Based on Topographic Distribution Approach

The surface characterization features based on the international organization for standardization (ISO) have been widely applied for assessing the surface topography of a workpiece (Brune *et al.*, 2008, Ali, 2012). These features include the average roughness S_a, the root mean square roughness S_q, the skewness S_{sk}, and the kurtosis S_{ku}, which are respectively derived from the first, second, third, and fourth moments of the distribution of topographic values. In particular, the most commonly used feature for assessing the quality of surface topography is S_a, which is expressed as the average of absolute deviation of all topographic values from the reference plane (Blunt and Jiang, 2003). The second-most used feature is S_q, which is defined as the root mean square deviation of all topographic values from the reference plane (Townsend *et al.*, 2016). In addition, S_{sk} and S_{ku} are introduced to describe the asymmetry and the sharpness of topographic distribution, respectively (Ali, 2012). Therefore, for the given surface topography of a product $Z = \{z_i; i = 1, 2, ..., M\}$, the S_a, S_q, S_{sk}, and S_{ku} are calculated as given in Equations (2.2-2.5), respectively,

$$S_a = \frac{1}{M}\sum_{i=1}^{M} | z_i |, \tag{2.2}$$

$$S_q = \sqrt{\frac{1}{M}\sum_{i=1}^{M} z_i^{\ 2}}, \tag{2.3}$$

$$S_{sk} = \frac{1}{M \times S_q^3}\sum_{i=1}^{M} z_i^{\ 3}, \tag{2.4}$$

$$S_{ku} = \frac{1}{M \times S_q^4}\sum_{i=1}^{M} z_i^{\ 4}. \tag{2.5}$$

These features have received acceptance in both industry and literature due to ease of implementation and explanation (Schmähling, 2006, Rao *et al.*, 2015b). In addition, these features are effective in detecting global variance change because of the way of characterizing the topographic distribution. However, these features lack the detection of local variance changes because local changes are likely to be dominated by other normal values. In addition, these features show a limitation in detecting local and global autocorrelation changes since they are defined as a first-order statistic, which mainly describes the characteristics of the topographic values individually without considering the spatial relationships among them.

2.4.2 Features Based on Signal Processing Approach

Signal processing is applied to monitor changes in topographic surfaces (Schmähling, 2006, Palani and Natarajan, 2011, Kanafi and Tuononen, 2017). In particular, for the given grayscale image of surface topography, which can be represented as a matrix $\mathbf{G} = [g_{xy} : x = 1, 2, ..., N, \ \ y = 1, 2, ..., M]$, where g_{xy} is the grayscale intensity level at the row x and column y, and (N and M) are the number of rows and columns in the image matrix, the fast Fourier transforms (FFT) algorithm is used to improve the representation of surface characteristics by decomposing the surface topography into a few bases where each base explains a certain surface characteristic (Schmähling, 2006, Palani and Natarajan, 2011, Kanafi and Tuononen, 2017). In particular, the FFT of g_{xy} is given as

$$f_{uv} = \frac{1}{N \times M} \sum_{x=1}^{N} \sum_{y=1}^{M} g_{xy} \, e^{-j2\pi(ux/N + vy/M)} \,, \; u = 1,2,\ldots,N, \, v = 1,2,\ldots,M,$$

where u and v are the indexes of the frequency domain and j is an imaginary unit. Note that f_{uv} is a complex number, which is defined in the frequency domain as

$$f_{uv} = r_{uv} + j \, i_{uv} \,, \; u = 1,2,\ldots,N, \, v = 1,2,\ldots,M,$$

where r_{uv} and i_{uv} are the real and imaginary terms, respectively. Consequently, the power spectrum p_{uv} is given by

$$p_{uv} = | f_{uv} |^2 = r_{uv}^2 + i_{uv}^2 \,, \; u = 1,2,\ldots,N, \, v = 1,2,\ldots,M \,.$$

After the surface is transformed into a 2D Fourier transform spectral, the power spectral density feature S_{psd} is obtained by averaging the power spectrum values as given in Equation (2.6)

$$S_{psd} = \frac{1}{N \times M} \sum_{u=1}^{N} \sum_{v=1}^{M} p_{uv} \,. \tag{2.6}$$

Since the FFT algorithm decomposes the surface into a few bases that capture the main characteristics of topographic distribution, S_{psd} is effective in detecting the global variance change of topographic values (Schmähling, 2006). However, the FFT algorithm is defined as a global decomposition filter where local surface defects are likely to be smoothed out, which may result in low detection performance of local variance change (Schmähling, 2006). In addition, the spatial relationships between pixels are ignored during the transformation process, which may yield a low detection performance of the local and global changes in the autocorrelation structure of topographic values.

2.4.3 Features Based on Image Processing Approach

Monitoring approaches based on image processing have also been proposed for describing and evaluating the surface topography of manufactured units (Xiao *et al.*, 2006). In particular, edge detection filters, such as the Canny filter, are applied to convert the surface image $Z=\{z_i; i=1,2,...,M\}$ to a binary image $B=\{b_i; i=1,2,...,M\}$, where b_i is an assigned binary value of the pixel i. Then, the density of summits feature S_{ds} is calculated to capture the ratio of peaks per unit area (Xiao *et al.*, 2006), which is given in Equation (2.7)

$$S_{ds}=\frac{1}{M}\sum_{i=1}^{M}1\{b_i=1\}, \tag{2.7}$$

where $1\{.\}$ is an indicator function defined as $1\{b_i=1\}=1$, otherwise zero. Note that a larger value of S_{ds} indicates that the observed surface is "rough" due to the increase of the number of peaks on the surface and vice versa. Since the variance change can cause a deviation in the topographic values from the reference plane, this is likely to result in an increase in the number of "1" in the obtained binary values. This change can be effectively quantified and detected by S_{ds}, which subsequently yields an effective detection performance of the global variance change. However, the major drawback of S_{ds} is that it heavily depends on a heuristic selection of the edge detection filter parameters. Since surface faults may form in a variety of sizes and shapes, it is challenging to optimize these parameters. Consequently, the fault information located in spatial areas is likely to be lost during the binarization process, which may result in low detection performance of the local and global autocorrelation change. Moreover, the characteristics of normal topographic values are likely to obscure the characteristics of local faults, which may lead to low detection performance of local variance change.

The watershed algorithm is also utilized to quantify and monitor changes in surface characteristics (Yang *et al.*, 2011, Gaetano *et al.*, 2015). The watershed algorithm, which is originally inspired by the mechanism of the flooding process in geography, is used to segment the surface topography into different labels (or regions) (Beucher and Meyer, 1993). More specifically, for the given grayscale image of surface topography $G = \{g_i; i = 1, 2, ..., M\}$, where g_i is the grayscale intensity level of the pixel i, the watershed algorithm begins to fill each separated local minima (or valley) with different colored water (or label) as shown in Figure 2.3 (a). As the water increases, the water from different valleys begins to combine. Subsequently, dams (or watershed lines) are constructed in the coordinates where the water combines to avoid the water merging as shown in Figure 2.3 (b and c). Finally, the algorithm continues the previous process of filling water and constructing dams until all peaks become under the water as shown in Figure 2.3 (d).

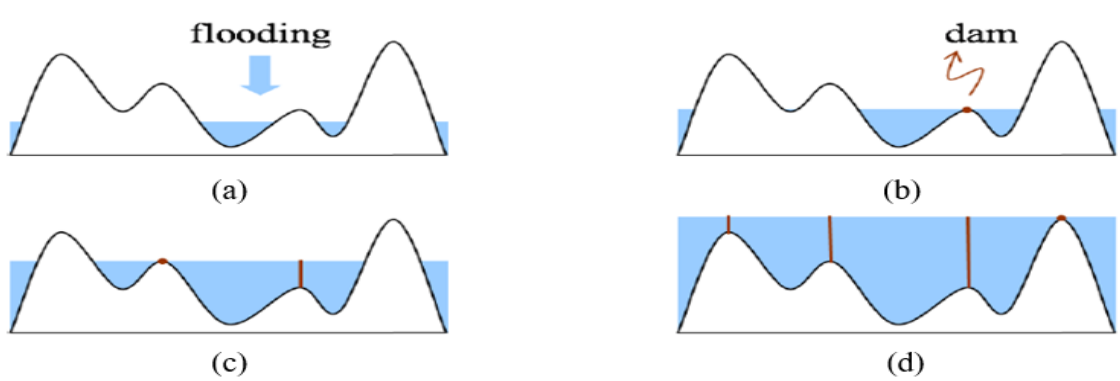


Figure 2.3 Illustration of the watershed algorithm: (a) the water flooding, (b and c) the process of constructing dams (red lines), and (d) the process of labeling surface valleys (Wang, 2010).

As a result, the grayscale surface image G is converted to a discrete surface $D = \{d_i : i = 1, 2, ..., M\}$, where $d_i \in \{1, 2, ...\}$ is an assigned label of the pixel i. Then, the watershed feature, denoted as S_w, is determined by finding the total number of assigned labels (Yang *et al.*, 2011, Gaetano *et al.*, 2015), which is given by Equation (2.8)

$$S_w = \max_{1 \le i \le M} d_i \ . \tag{2.8}$$

Note that a larger value of S_w implies a "rough" surface because of the increase of the number of assigned labels, while a small value of S_w implies a "smooth" surface. The watershed feature S_w is also applied to extract fault information by identifying the fault locations and the number of surface faults (Yang *et al.*, 2011, Gaetano *et al.*, 2015). The watershed algorithm is effective in assessing the surface structure due to the segmentation of surface characteristics. This makes S_w to be effective in monitoring the global autocorrelation change. However, the watershed algorithm does not capture the characteristics of topographic distribution, which may cause a low detection performance of the global variance change. The segmentation performance of local faults can also be impacted by the over-segmentation problem, which is likely to occur due to the noise or any other irregularities in the surface image (Wang, 2010). This may result in low performance in identifying local variance and autocorrelation changes.

2.4.4 Features Based on Graph Theory Approach

Monitoring approaches based on graph theory have been recently developed for describing and monitoring the characteristics of topographic surfaces (Rao *et al.*, 2015a, Rao *et al.*, 2015b). More specifically, edge detection filters (e.g., a Canny filter) are initially applied to convert the surface image into the following binary matrix $\mathbf{B} = [b_{xy} : x = 1, 2, ..., N,$

$y = 1, 2, ..., M]$, where $b_{xy} \in \{0,1\}$ is the assigned binary value in the row x and column y, and (N and M) are the number of rows and columns in the image matrix. Then, the binary matrix $\mathbf{B}$ is represented by its row vectors such as $\mathbf{r}_k, k = 1, 2, ..., N$, where $\mathbf{r}_k$ is the M-dimensional vector of the row k in the matrix $\mathbf{B}$. Next, the following steps are applied:

- **Step 1:** Calculate the distance matrix based on the obtained $\mathbf{r}_k, k = 1, 2, ..., N,$ using the Euclidean distance such as $\mathbf{L} = [l_{ij} : 1, 2, ..., N],$ where l_{ij} represents the Euclidean distance between the i^{th} and j^{th} row vectors and defined as $l_{ij} = \| \mathbf{r}_i - \mathbf{r}_j \|^2$.
- **Step 2:** Calculate the dissimilarity matrix by converting $\mathbf{L}$ to a binary dissimilarity matrix using a predefined hard threshold C such as

$$\mathbf{A} = [a_{ij}],\ a_{ij} = \begin{cases} 1, & \text{if } l_{ij} > C \\ 0, & \text{otherwise} \end{cases}, \text{ for } \forall i, j \in \{1, 2, ..., N\},$$

  where a_{ij} represents the dissimilarity value between the i^{th} and j^{th} row vectors.
- **Step 3:** Construct an undirected graph based on the obtained $\mathbf{A}$, where the graph nodes are equal to the number of row vectors (N nodes), and the graph edge between the i^{th} and j^{th} nodes is represented by the calculated dissimilarity value a_{ij}, $i, j \in \{1, 2, ..., N\}$.
- **Step 4:** Calculate the normalized Laplacian matrix to quantify the characteristics of the obtained graph as

$$\mathbf{O} = \mathbf{D}^{-1/2} \times (\mathbf{D} - \mathbf{A}) \times \mathbf{D}^{-1/2},$$

  where $\mathbf{D}$ is the degree matrix obtained as $\mathbf{D} = diag(d_j),$ where d_j is the number of edges connected to the j^{th} node and is defined as $d_j = \sum_{i=1}^{N} a_{ij},\ j = 1, 2, ..., N$.
- **Step 5:** Solve the Eigen spectrum using Equation (2.9)

$$\mathbf{O} \times \mathbf{K} = \boldsymbol{\lambda} \times \mathbf{K}, \tag{2.9}$$

where $\boldsymbol{\lambda}$ is a vector that includes the eigenvalue of each node and $\mathbf{K}$ is a $N \times N$ matrix such that each column in $\mathbf{K}$ represents the eigenvector of each node.

- **Step 6:** Obtain the graph connectivity by finding the second smallest non-zero eigenvalue in $\boldsymbol{\lambda}$, which is known as the Fiedler connectivity feature (or number) S_c.

The Fiedler connectivity feature S_c is used to quantify the "smoothness" of surface topography such that larger connectivity value indicates a more "rough" surface and vice versa. Due to the quantification of the pairwise distance between all nodes of the graph network, the Fiedler connectivity feature S_c is effective in detecting the global variance change. However, S_c lacks the identification of local variance change because faulty characteristics can be dominated by normal characteristics during the calculation of S_c. In addition, S_c has a limitation in detecting the local and global autocorrelation changes since the spatial autocorrelation between the topographic values is not considered.

2.4.5 Features Based on Machine Learning Approach

Machine learning has been recently applied for monitoring and diagnosing faults in topographic surfaces (Bui and Apley, 2018a, Bui and Apley, 2018b). In particular, for the given surface topography of a part $Z = \{z_i; i = 1, 2, ..., M\}$, each topographic value z_i is subtracted from an estimated topographic value $\hat{f}(z_i)$, which is obtained by a supervised machine learning model (regression tree), to calculate the residual at each surface pixel such as

$$r_i = \hat{f}(z_i) - z_i,\ i = 1, 2, ..., M,$$

where r_i is the calculated surface residual of the pixel i. The behavior of the obtained surface residuals is subsequently monitored using the one-sample Anderson–Darling (OAD) feature, which examines whether the cumulative distribution function (CDF) of the observed residuals is similar to the CDF of the normal residuals. Particularly, after calculating the OAD feature for each pixel a_i, $i = 1, 2, ..., M$, based on a squared moving window with a defined width w, the maximum OAD feature S_{ad} is obtained as provided in Equation (2.10)

$$S_{ad} = \max_{i=1,2,...,M} a_i \ . \tag{2.10}$$

After detecting an anomaly surface based on its S_{ad}, the fault locations are identified by comparing the calculated a_i, $i = 1, 2, ..., M$, for each pixel to a predefined diagnostic threshold obtained based on a defined quantile of the probability density function (PDF) of the normal values of a_i, $i = 1, 2, ..., M$. Note that S_{ad} is effective in detecting the global variance change because S_{ad} captures the deviation of the CDF of the observed residuals from the CDF of the normal residuals. However, S_{ad} yields a low performance in detecting the local and global autocorrelation changes since S_{ad} fails to utilize the spatial autocorrelation among topographic pixels. Moreover, S_{ad} has a low detection performance of the local variance change because local faults are likely to be dominated by the normal topographic values. In addition, there are no clear guidelines on the selection of the width of the moving window parameter w, which may also impact the detection performance of S_{ad}.

In summary, the aforementioned monitoring approaches do not fully characterize the complex behavior among topographic pixels, which renders them ineffective in detecting and diagnosing local and global changes in the variance and autocorrelation of topographic values. Therefore, in this dissertation, we overcome these limitations by introducing four accurate and efficient statistical approaches for online monitoring and diagnosis of various changes in topographic surfaces.

CHAPTER 3

MULTI-LEVEL SPATIAL RANDOMNESS APPROACH FOR MONITORING GLOBAL CHANGES IN 3D TOPOGRAPHIC SURFACES

3.1 Introduction

Effective online monitoring approaches based on the 3D surface topography of manufactured products are needed in many industries as discussed in Section 1.1. Most of the existing monitoring approaches consider a single statistic for assessing the characteristics of surface topography, such as the average roughness and the root mean square roughness discussed in Section 2.4. However, using only a single statistic may not be sufficient for quantifying the spatial and complex structure of surface characteristics, which may result in a low assessment of global surface changes. Thus, representing the 3D surface topography as an 1D profile that describes the spatial structure of surface characteristics at different levels of surface heights is a novel approach for accurately characterizing and monitoring different forms of global changes in 3D topographic surfaces.

In this chapter, we develop a multi-level spatial randomness approach for online monitoring of global changes in 3D topographic surfaces. Specifically, we enhance the representation of surface characteristics by slicing the 3D surface topography into different levels in accordance with the characteristics of the in-control surfaces through a proposed multi-level surface thresholding algorithm. We also introduce the spatial randomness (SR) profile, which captures different patterns of surface changes by quantifying the spatial characteristics of topographic values at each surface level. After obtaining the SR profile, an effective monitoring statistic based on the functional principal component analysis is developed for accurate detection of global changes in topographic surfaces. Note that Chapter 3 is mostly based on the following published paper: ALQAHTANI, M. A., JEONG, M. K. and ELSAYED, E. A. 2020c. Multilevel spatial randomness approach for monitoring changes in 3D topographic surfaces. *International Journal of Production Research*, 58, 5545-5558.

The remainder of this chapter is organized as follows. Section 3.2 presents the proposed approach in detail. Then, Section 3.3 shows the performance comparison between the proposed and the existing monitoring approaches. Section 3.4 demonstrates the performance of the proposed approach using a case study from the semiconductor industry. Conclusions are discussed in Section 3.5.

3.2 Proposed Approach

The proposed monitoring approach includes three stages; Stage 1: multi-level surface thresholding, Stage 2: feature extraction, and Stage 3: anomaly detection. In Stage 1, we divide the surface into different levels for an effective representation of topographic characteristics. Stage 2 extracts a statistical feature from each surface level for accurate quantification of spatial surface characteristics. Finally, Stage 3 monitors and detects anomaly surfaces based on their extracted features. The following sections explain each stage in detail.

3.2.1 Stage 1: Multi-level Surface Thresholding

The spatial and random characteristics of 3D topographic surfaces make the detection of surface changes difficult. Thus, it is crucial to simplify and enhance the representation of topographic characteristics for achieving accurate detection performance of surface changes. Hard thresholding (HT) algorithms based on edge detection filters, such as Canny and Sobel filters, are commonly used for binarizing images based on a single threshold. The main goal of these filters is to identify any discontinuity or sharp change in the pixel intensities. This subsequently simplifies and improves the representation of image characteristics for effective surface analysis (Xiao *et al.*, 2006). Due to the complex structure of topographic values, identifying only a single threshold is likely to remove the important characteristics of topographic values, which may yield a low representation of surface changes.

The multi-level thresholding algorithms, on the other hand, are developed to achieve an improved representation of image characteristics (Liao *et al.*, 2001, Vala and Baxi, 2013) For example, Otsu's multi-level thresholding (OMT) algorithm is commonly used for image segmentation (Otsu, 1979, Liao *et al.*, 2001). Particularly, the OMT algorithm divides the pixel intensities into "*K*-non-overlapping intervals" based on selected $(K+1)$ thresholds. The selection of thresholds is obtained by minimizing the variance of each interval such that each pixel intensity is strictly assigned to one of the "*K*-non-overlapping intervals". However, the algorithm does not consider the spatial autocorrelation among adjacent pixels during the segmentation process, which may result in a lower representation of changes in topographic surfaces.

We propose a multi-level surface thresholding (MST) algorithm for improving the representation of topographic values. Let the input data be represented as a matrix

$\mathbf{Z} = [z_{xy}; x, y = 1, 2, .., M]$, where z_{xy} is a real number that represents the height or topographic value located in the (x, y) coordinate and $(M \times M)$ is the matrix size. The surface matrix is then decomposed into K binary matrices $\mathbf{A}^{(r)} = [\, a_{xy}^{(r)}]$, $r = 1, 2, ..., K$, according to the sign of predefined thresholds $\delta^{(r)}$, $r = 1, 2, ..., K$. Thus, each r^{th} binary matrix (or surface level), $\mathbf{A}^{(r)}$, $r = 1, 2, ..., K,$ is obtained as given in Equation (3.1)

$$
\begin{aligned}
&\delta^{(r)} < 0 : \mathbf{A}^{(r)} = [\, a_{xy}^{(r)}] \ , a_{xy}^{(r)} = \begin{Bmatrix} 1\, , \, z_{xy} < \delta^{(r)} \\ 0\, , \text{otherwise} \end{Bmatrix} \\
&\delta^{(r)} \geq 0 : \mathbf{A}^{(r)} = [\, a_{xy}^{(r)}], \ a_{xy}^{(r)} = \begin{Bmatrix} 1\, , \, z_{xy} \geq \delta^{(r)} \\ 0\, , \text{otherwise} \end{Bmatrix},
\end{aligned}
\tag{3.1}
$$

where $a_{xy}^{(r)}$ is a binary indicator that represents whether there is a material or a void located in the (x, y) coordinate. Note that $a_{xy}^{(r)}$ is assigned “one”, if the slicing plane hits a “material”, and it is assigned “zero” if the slicing plane hits a “void”. More specifically, if the threshold is negative at the r^{th} surface level (i.e., $\delta^{(r)} < 0$), then $a_{xy}^{(r)}$ is assigned “one” when z_{xy} is less than $\delta^{(r)}$. However, if the threshold is positive at the r^{th} surface level (i.e., $\delta^{(r)} \geq 0$), then $a_{xy}^{(r)}$ is assigned “one” when z_{xy} is greater than or equal $\delta^{(r)}$.

In the proposed MST algorithm, the selection of thresholds $\delta^{(r)}$, $r = 1, 2, ..., K$, is important for binarizing the topographic values at each surface level. Thus, for the given N in-control surfaces $\mathbf{Z}_i$, $i = 1, 2, ..., N$, the total number of surface levels K (the determination of K is discussed later in the chapter) and a defined type I error rate (e.g., $\alpha = 0.001$), the following steps are taken for estimating the threshold for each surface level:

- **Step 1:** Rank the topographic values in an increasing order for each surface $\mathbf{Z}_i$, $i = 1, 2, ..., N$,

$$F_i = \{z_{i(1)}, z_{i(2)}, ..., z_{i(M \times M)}\},$$

where $z_{i(1)}, z_{i(2)}, ..., z_{i(M \times M)}$ are the order statistics of the topographic values in the i^{th} surface.

- **Step 2:** Obtain the α and $1-\alpha$-quantiles for each F_i, $i = 1, 2, ..., N$, as shown in Equation (3.2)

$$\Pr[F_i \le q_i^{(\alpha)}] = \alpha, \ \Pr[F_i \le q_i^{(1-\alpha)}] = 1 - \alpha, \tag{3.2}$$

where $q_i^{(\alpha)}$ and $q_i^{(1-\alpha)}$ are the α and $1-\alpha$ -quantiles obtained from the i^{th} surface.

- **Step 3:** Calculate the r^{th} threshold for each obtained $q_i^{(\alpha)}$ and $q_i^{(1-\alpha)}$, $i = 1, 2, ..., N$, by equally dividing the interval between $q_i^{(\alpha)}$ and $q_i^{(1-\alpha)}$ as given in Equation (3.3)

$$\delta_i^{(r)} = q_i^{(\alpha)} + (r-1)\frac{(q_i^{(1-\alpha)} - q_i^{(\alpha)})}{(K-1)}, \text{ for } r = 1, 2, ..., K, \tag{3.3}$$

where $q_i^{(\alpha)} \ge \delta_i^{(r)} \ge q_i^{(1-\alpha)}$ is the obtained threshold value at the r^{th} surface level of the i^{th} surface.

- **Step 4:** Calculate the average threshold for each r^{th} surface level using Equation (3.4)

$$\delta^{(r)} = \sum_{i=1}^{N} \delta_i^{(r)} / N, \text{ for } r = 1, 2, ..., K. \tag{3.4}$$

Figure 3.1 is an example of a 3D surface topography before and after being sliced into 20 levels (K=20) using the proposed MST algorithms. Note that surface level 1 is the lowest surface level, whereas surface level 20 is the highest level. In addition, the white areas represent the “material” pixels (i.e., “ones” in $\mathbf{A}^{(r)}$), while the black areas represent the “void” pixels (i.e., “zeros” in $\mathbf{A}^{(r)}$). Figure 3.1 shows that the MST algorithm constructs the surface levels in a cumulative manner where each surface level is calculated based on the information obtained from the previous levels. Accordingly, each topographic value may be assigned to one or multiple surface levels. Figure 3.1 (b) displays the first surface level $\mathbf{A}^{(1)}$, which only contains the material pixels below the first threshold $\delta^{(1)}$. However, Figure 3.1 (d) shows the tenth surface level $\mathbf{A}^{(10)}$, which contains the material pixels below the tenth threshold $\delta^{(10)}$, including the ones in the previous levels $\{\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, ..., \mathbf{A}^{(9)}\}$.

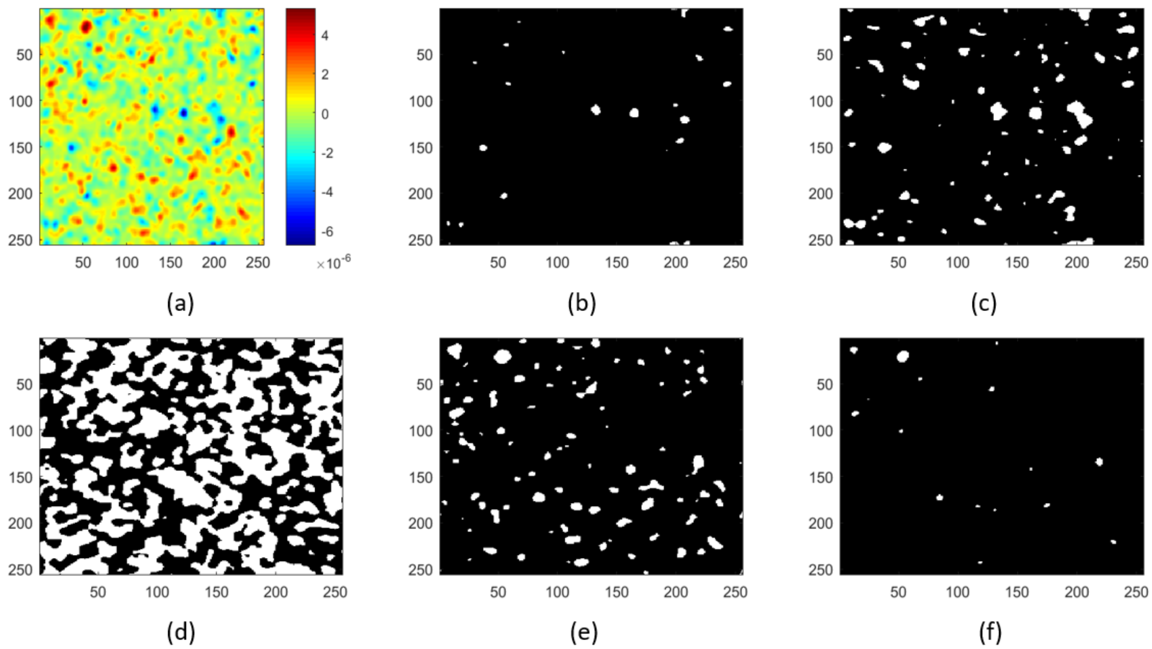


Figure 3.1 Example of (a) a topographic surface $\mathbf{Z}$, and its corresponding surface levels $\mathbf{A}^{(r)}$ with (b) r =1, (c) r =5, (d) r =10, (e) r =15, and (f) r =20, obtained the MST algorithm.

There are several advantages of applying the MST algorithm in monitoring topographic surfaces. In particular, the MST algorithm improves the representation of topographic values by discretizing the continuous topographic matrix into *K* binary matrices. The MST algorithm is also utilized for online monitoring such that the surface levels of a new surface are constructed based on predefined thresholds extracted from the characteristics of in-control surfaces. In addition, the MST algorithm is effective in preserving the autocorrelation structure of topographic values. Particularly, the MST algorithm shows a specific pattern of the material pixels at each surface level where a less clustered pattern of the material pixels is achieved at the top and bottom surface levels as shown in Figure 3.1 (b and f), and a more clustered pattern is achieved at the central levels as shown in Figure 3.1 (c, d, and e). This is because most of the material pixels typically appear around the reference plane, while fewer material pixels appear farther away. Thus, it is expected normal surfaces to have a similar pattern of their material pixels since they share similar surface characteristics. Subsequently, when a new surface contains any abnormality in its characteristics, this is likely to impact the pattern of its material pixels, which yields better monitoring of anomaly surfaces.

In the following section, a spatial statistic is presented for quantifying the pattern of material pixels at each surface level.

3.2.2 Stage 2: Feature Extraction

The spatial randomness test (SRT) is well-known for detecting the existence of spatial autocorrelation in observed binary images (Taam and Hamada, 1993). The SRT determines whether the surface pixels are dependent on the pixels located at their nearby locations. The SRT is computed based on the join-count (JC) statistic, which counts the frequency of the appearance of the 0-to-0 join (i.e., between void pixels) and the 1-to-1 join (i.e., between material pixels) based on a specified neighborhood construction rule (Hansen and Thyregod, 1998). Since the topographic surface is classified as continuous data, the current SRT is deemed inapplicable for quantifying the spatial autocorrelation that exists between topographic values. Therefore, the proposed MTS algorithm is utilized to overcome this limitation. Particularly, let the obtained binary matrices $\mathbf{A}^{(r)}$, $r=1,2,...,K$, be represented as a set of binary values $A^{(r)}=\{a_j^{(r)}; j=1,2,...,P, r=1,2,...,K\}$, where $a_j^{(r)}$ is the assigned binary value of the j^{th} pixel at the r^{th} surface level and $P = M \times M$ is the total number of the image pixels. Then, the JC statistics of the 0-to-0 and 1-to-1 joins are computed as given in Equations (3.5 and 3.6), respectively,

$$J_v^{(r)} = \sum_{j=1}^{P} \sum_{i \in \Omega} \left(1 - a_i^{(r)}\right)\left(1 - a_j^{(r)}\right), \tag{3.5}$$

$$J_m^{(r)} = \sum_{j=1}^{P} \sum_{i \in \Omega} a_i^{(r)} a_j^{(r)} , \tag{3.6}$$

where r is the index of each surface level, $r = 1, 2, ..., K$, $(J_v^{(r)}$, $J_m^{(r)})$ are the JC statistics of void and material pixels, respectively, and Ω is a defined neighborhood construction rule. In this chapter, we utilize the popular king-move neighbor (KMN) rule, which considers the eight neighbors located vertically, horizontally, and diagonally adjacent to the examined pixel. Consequently, we obtain the spatial randomness (SR) profile by computing the spatial randomness level at each surface level as given in Equation (3.7)

$$T(r) = w^{(r)} \, J_v^{(r)} + (1 - w^{(r)}) \, J_m^{(r)} \text{, for } r = 1, 2, ..., K , \tag{3.7}$$

where $w^{(r)}$ is an assigned weight for the JC statistic of material pixels at the r^{th} surface level, which is obtained as $w^{(r)} = v^{(r)} / P$, where $v^{(r)}$ is the total number of "ones" in the r^{th} binary set $A^{(r)}$. Note that the value of $T(r)$ represents how the topographic values are spatially distributed and structured at the r^{th} surface level, $r = 1, 2, ..., K$. A lower value of $T(r)$ means that the topographic values at the r^{th} surface level are more random and distant from each other, whereas a higher value indicates that the topographic values are more autocorrelated and clustered. Therefore, the 3D surface topography is converted into a K -

dimensional SR profile, where each value in the SR profile represents the spatial characteristics at each surface level.

The SR profile has the advantages of capturing the spatial and random characteristics between the topographic values. In particular, the SR profile quantifies the spatial autocorrelation structure of topographic values by analyzing the JC statistics of void pixels (i.e., 0-to-0 join) and material pixels (i.e., 1-to-1 join) at each surface level based on a defined neighborhood construction rule. The SR profile is also effective in quantifying the random characteristics of topographic values in which the SR of topographic values at each surface level is expected to be statistically similar under the in-control process behavior. As a result, the shapes of the corresponding SR profiles of in-control surfaces are expected to be similar even though the locations of topographic values randomly change from one surface to another. This is because the calculation of the SR profile is independent of the actual locations (or coordinates) of the topographic values. This makes the SR profile effective in characterizing and monitoring both spatial and random characteristics of topographic values.

3.2.3 Stage 3: Anomaly Detection

In this section, the topographic surfaces with abnormal characteristics are effectively detected based on their SR profiles. Stage 3 includes two phases: Phase 1 (offline) and Phase 2 (online). In Phase 1, we learn and extract in-control parameters that capture the characteristics of in-control SR profiles $T_i(r)$, $i = 1, 2, ..., N$. Then, in Phase 2, we identify whether the SR profile of a new surface $T_{new}(r)$ is an anomaly based on the in-control parameters.

3.2.3.1 Phase 1: Estimating Parameters from In-control SR Profiles

The functional principal component analysis (FPCA) is commonly applied to capture the prominent modes of variation in functional data (Ramsay and Silverman, 2005). Therefore, the FPCA is utilized to capture most of the variation between the grid points of in-control SR profiles. In particular, the FPCA transforms the in-control SR profiles (i.e., input data), which has a size $N \times K$, where N is the sample size and K is the total number of surface levels (or grid points), to a reduced set of uncorrelated variables, called as principal component (PC) scores, which has a size $N \times D$, where D is significantly less than K (i.e., $D << K$). Therefore, each in-control SR profile $T_i(r), i = 1,2,...,N$, can be represented by the first D eigenfunctions using Equation (3.8)

$$T_i(r) \approx \bar{T}(r) + \sum_{j=1}^{D} \upsilon_{i,j}\, e_j(r)\,,\ \text{for } r \in \{1,2,...,K\}\,, \tag{3.8}$$

where $\bar{T}(r)$ is the mean SR profile, $e_j(r)$, $j = 1,2,...,D$, are the eigenfunctions, $\upsilon_{i,j}$, $i = 1,2,...,N$, $j = 1,2,...,D$, are the principal component (PC) scores, and (K, D) are the original and reduced dimensional spaces, respectively. Note that the index "r" refers to the r^{th} grid point in the SR profile, $r \in \{1,2,...,K\}$, while the index "j" refers to the j^{th} mode of variation in the FPCA, $j = 1,2,...,D$.

The eigenvalues λ_j, $j = 1,2,...,D$, and eigenfunctions $e_j(r)$, $j = 1,2,...,D$, are both calculated based on the covariance operator shown in Equation (3.9)

$$\mathrm{c}(t,s) = \frac{1}{N-1}\sum_{i=1}^{N} \{T_i(t) - \bar{T}(t)\}\,\{T_i(s) - \bar{T}(s)\}\,, \tag{3.9}$$

where t and s are two grid points in the in-control SR profile, $t, s \in r = \{1, 2, ..., K\}$, $\left(T_i(t),\ T_i(s)\right)$ are the SR value at the t^{th} and s^{th} grid points, and $\left(\bar{T}(t),\ \bar{T}(s)\right)$ are the mean test values of the SR at the t^{th} and s^{th} grid points, which are obtained as $\bar{T}(t) = \frac{1}{N}\sum_{i=1}^{N} T_i(t)$, and $\bar{T}(s) = \frac{1}{N}\sum_{i=1}^{N} T_i(s)$. After $\mathrm{c}(t,s)$ is calculated, the eigenvalues and eigenfunctions $\{\lambda_j\ , e_j(r),\ j = 1, 2, ..., D\}$ are obtained by the spectral decomposition (SD). Specifically, the SD of the covariance function $\mathrm{c}(t,s)$, $t, s \in r = \{1, 2, ..., K\}$, is given as $\sum_{j=1}^{K} \lambda_j e_j(t) e_j(s)$, where $\lambda_1 \geq \lambda_1 \geq ... \geq \lambda_K$ are the ordered nonnegative eigenvalues and $\left(e_j(t), e_j(s)\right)$ are the corresponding eigenfunctions (Paynabar *et al.*, 2016). Subsequently, we select the first D eigenvalues and eigenfunctions $\{\lambda_j\ , e_j(r),\ j = 1, 2, ..., D\}$ as $\sum_{j=1}^{D} \lambda_j / \sum_{j=1}^{K} \lambda_j \geq \omega$, where ω is a prespecified percentage of the cumulative target variance (Yu *et al.*, 2012).

The PC scores are calculated based on the inner product of two profiles: the centered profile (i.e., the i^{th} in-control SR profile is subtracted from the mean SR profile, $\Delta_i(r) = T_i(r) - \bar{T}(r)$, $i = 1, 2, ..., N$), and the first D eigenfunctions $e_j(r)$, $j = 1, 2, ..., D$. Note that it is reasonable to use summation instead of integral to calculate the inner product since each SR profile is discretized into an equally spaced grid of points (Paynabar, Zou, and Qiu 2016). Thus, the PC scores are obtained using Equation (3.10)

$$\upsilon_{i,j} = \left\langle \Delta_i(r), e_j(r) \right\rangle = \sum_{r=1}^{K} \Delta_i(r) e_j(r), \text{ for } i = 1, 2, ..., N, \text{ and } j = 1, 2, ..., D, \tag{3.10}$$

where $\upsilon_{i,j}$ is the j^{th} PC score of the i^{th} in-control SR profile. In the following section, we present how to monitor the quality of a new surface based on the parameters obtained from in-control SR profiles, including the mean SR profile, PC scores, eigenvalues, and eigenfunctions.

3.2.3.2 Phase 2: Identifying Anomaly Surfaces

The PC scores are key indicators that reflect the deviation between the new and mean profiles (Yu *et al.*, 2012). Therefore, after observing a new surface Z_{new}, the corresponding SR profile $T_{new}(r)$ is calculated. Then, the deviation between $T_{new}(r)$ and $\bar{T}(r)$ is computed in the original K-dimensional space as given in Equation (3.11)

$$\Delta_{new}(r) = T_{new}(r) - \bar{T}(r). \tag{3.11}$$

Then, the inner product of the obtained deviation $\Delta_{new}(r)$ and the first D eigenfunctions $e_j(r)$ is calculated for projecting $\Delta_{new}(r)$ into the reduced D-dimensional space, which results in the values of the PC scores given in Equation (3.12)

$$\upsilon_{new,j} = \left\langle \Delta_{new}(r), e_j(r) \right\rangle = \sum_{r=1}^{K} \Delta_{new}(r) e_j(r), \text{ for } j = 1, 2, ..., D. \tag{3.12}$$

Next, the obtained PC scores are squared and scaled by their respective eigenvalues for emphasizing and standardizing the projected deviation in the reduced D-dimensional space, which yields the monitoring statistic provided in Equation (3.13)

$$Q_{new} = \sum_{j=1}^{D} \frac{(\upsilon_{new,j})^2}{\lambda_j}, \tag{3.13}$$

The proposed monitoring statistic Q_{new} captures the total deviation between the SR profile of a new surface and the mean SR profile of in-control surfaces in the reduced D-dimensional space. Consequently, any deviation in the characteristics of the topographic values at any surface level can be effectively monitored and captured by Q_{new}. Finally, the new surface is identified as an anomaly, if $Q_{new} > C$, where C is a prespecified critical limit determined based on a bootstrapping method (Febrero *et al.*, 2008). For the obtained monitoring statistics of N in-control surfaces $\{Q_1, Q_2, ..., Q_N\}$ and a defined type I error rate (e.g., $\alpha = 0.05$), the steps for determining the critical limit C are as follows:

- **Step 1:** Estimate the empirical distribution function of the sampled monitoring statistics by randomly sampling R observations with replacement from the obtained N in-control monitoring statistics such that we obtain the following set $\{Q_1, Q_2, ..., Q_R\}$.
- **Step 2:** Rank the R sampled observations in increasing order, $F = \{Q_{(1)}, Q_{(2)}, ..., Q_{(R)}\}$, where $Q_{(1)}, Q_{(2)}, ..., Q_{(R)}$ are the order statistics, which are defined as $Q_{(1)} \le Q_{(2)} \le .... \le Q_{(R)}$.
- **Step 3:** Calculate the α-quantile of the ranked observations $\Pr[F \le h\ ] = \alpha$.
- **Step 4:** Repeat steps 1, 2, and 3 for T replicates and obtain the α-quantile of each replicate such that we obtain the following set $H = \{h_1, h_2, ..., h_T\}$.
- **Step 5:** Obtain the critical limit C by taking the median over the values in the set H as given in Equation (3.14)

$$C = \underset{t=1,2,...,T}{\text{median}} \ (\ h_t\ ). \tag{3.14}$$

Figure 3.2 shows a summary of the proposed approach. In the following section, different simulation studies are conducted to validate the performance of the proposed approach.

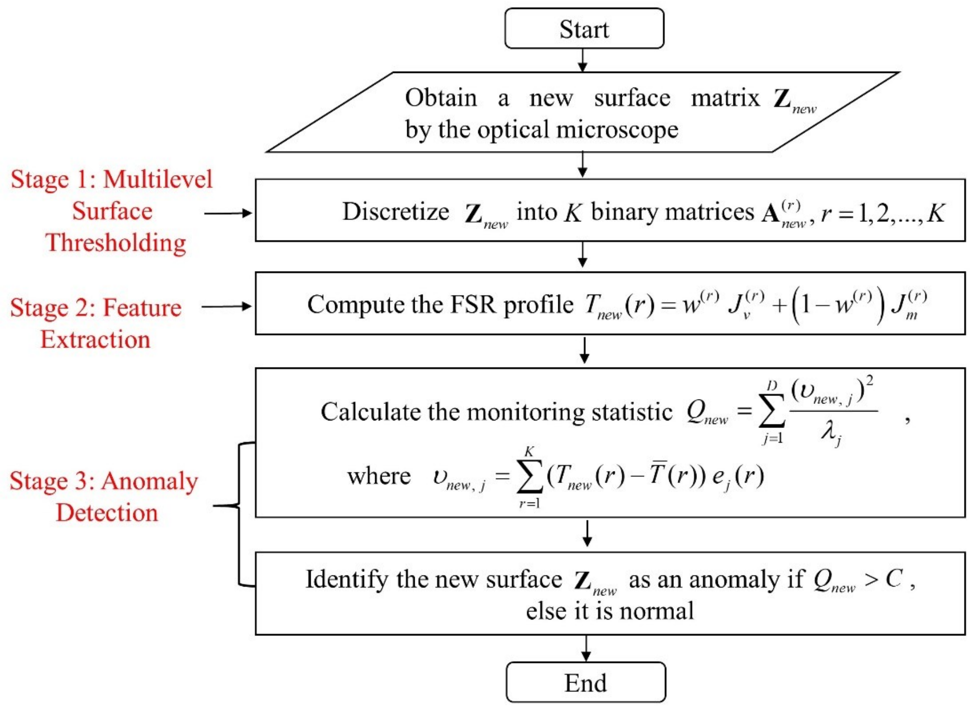


Figure 3.2 Summary of the proposed approach.

3.3 Performance Study

3.3.1 Surface Settings for In-control and Anomaly Surfaces

In the simulation experiments, we apply the Gaussian random model discussed in Section 2.3 to generate topographic surfaces. We consider different fault scenarios to address various faults in topographic surfaces. In particular, we address different surface changes (or fault types), including variance and autocorrelation changes. Note that the variance and autocorrelation changes are generated by shifting the parameters σ and θ in the Gaussian random model, respectively. We also examine different fault sizes where the sizes of the parameters σ and θ are shifted from their normal values. Besides, different fault areas are considered where a local fault area that allocates a specific percentage of the total surface area is randomly superimposed on the surface.

Figure 3.3 shows examples of normal surface and two anomaly surfaces with local variance and autocorrelation changes, respectively, along with a selected region obtained from the bottom-left corner of each surface. Note that the three surfaces initially are generated under the following normal settings: $\sigma = 1\ \mu m$ and $\theta = 1\ \mu m$. We also superimpose a local change that covers 2% of the total surface area (i.e., 36×36 pixels) on the bottom-left corner of each anomaly surface for illustration. In particular, a local variance change with $\sigma = 2\ \mu m$ and $\theta = 1\ \mu m$, and a local autocorrelation change with $\sigma = 1\ \mu m$ and $\theta = 2\ \mu m$ are superimposed on the surfaces shown in Figure 3.3 (b and c), respectively. We observe that the scale of the topographic values of the normal surface is about $4 \pm\ \mu m$, and its characteristics (e.g., peaks and valleys) are appeared to be less autocorrelated. However, when the surface experiences the variance change, we notice that the scale of the topographic values is increased to about $6 \pm\ \mu m$, but the pattern of peaks and valleys does not differ from the normal surface. Furthermore, when the surface has the autocorrelation change, we observe that the scale of topographic values does not change from the scale of the normal surface. However, a significant change in the valley and peak characteristics is observed in a way that the valleys and peaks become more similar and clustered.

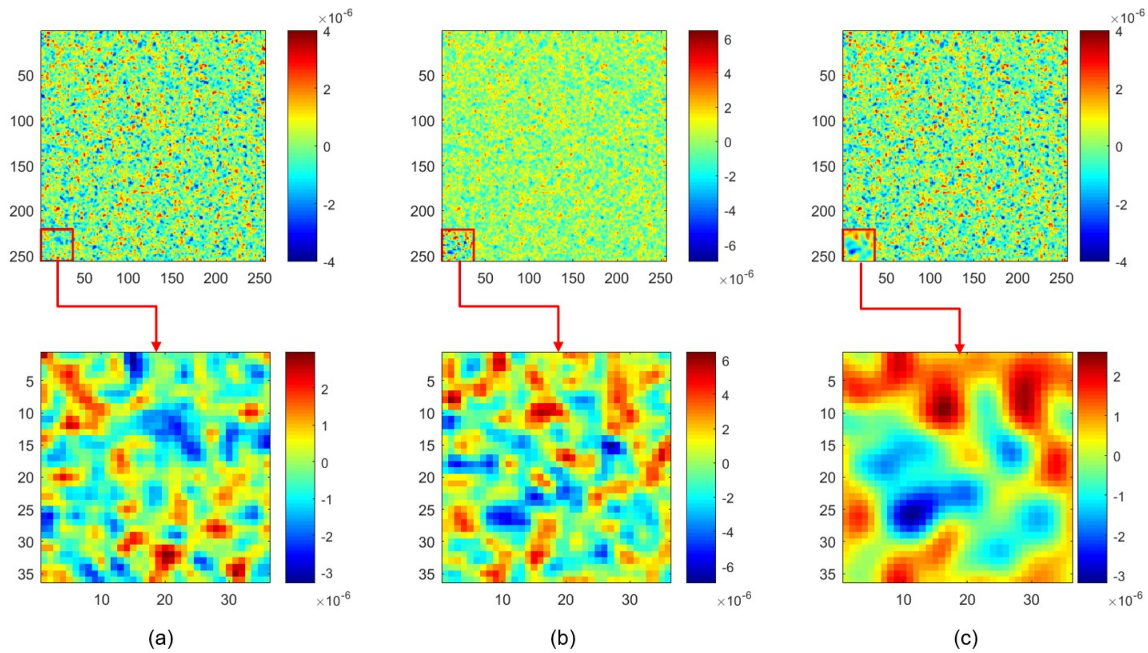


Figure 3.3 Three generated surfaces (first row) along with three selected regions (second row) of (a) normal surface, and two anomaly surfaces with (a) local variance change, and (c) local autocorrelation change.

3.3.2 Effectiveness of the Proposed Monitoring Approach

3.3.2.1 Advantage of the SR Profile in Characterizing Surface Changes

The patterns of the in-control SR profiles are generally characterized by a bell-shaped curve in which the lowest values of the SR profile (i.e., close to "zero") are located at the top and bottom surface levels because few random peaks and valleys are likely to appear at these levels. In addition, the highest values of the SR profile are positioned at the central levels since most of the peaks and valleys are likely to cluster at these levels. Therefore, any significant change from the pattern of the bell-shaped curve can lead to identifying different surface changes, including both variance and autocorrelation. In particular, the variance change mainly impacts the magnitude of topographic values at the top and bottom levels, which results in the appearance of more clustered areas at these levels. Furthermore, the autocorrelation change primarily impacts the structure of the topographic values at the central levels, which subsequently yields a change in the cluster pattern of topographic values at these levels.

Figure 3.4 illustrates the effectiveness of the SR profile obtained by the proposed MST algorithm in capturing different forms of surface changes. The figure shows the SR profiles for 10 in-control and anomaly surfaces with different fault sizes and types where each surface is sliced into 20 levels ($K = 20$). Note that the x -axis represents the index of each surface level, $r = 1, 2, ..., 20$, and the y -axis represents the corresponding test value of the SR. We observe that the obtained anomaly SR profiles are clearly distinguished from the in-control profiles. This is due to the effectiveness of the SR profiles in capturing the spatial pattern of the material pixels at each surface level. Furthermore, when the fault size becomes larger, the SR profiles of anomaly surfaces become more evident from the in-control profiles.

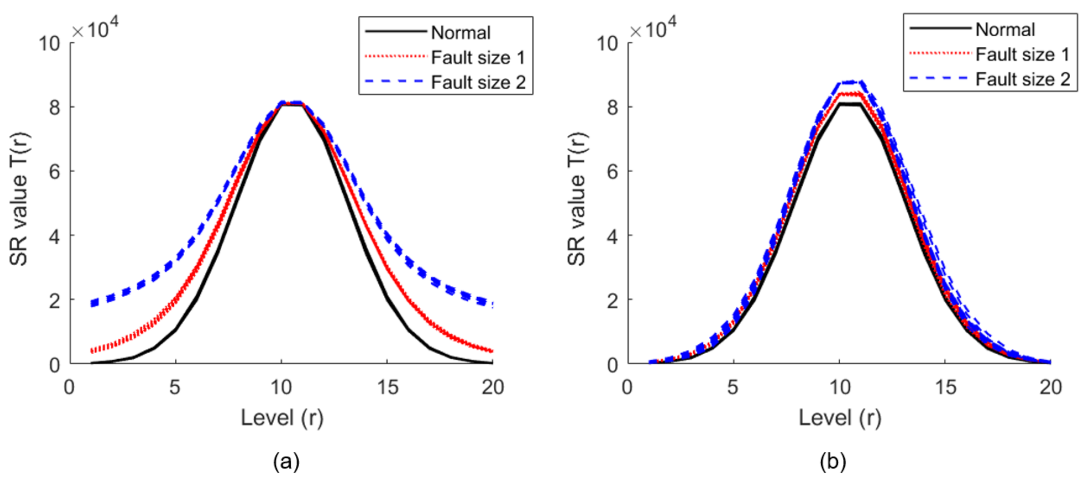


Figure 3.4 Calculated SR profiles of in-control and anomaly surfaces with (a) variance change, and (b) autocorrelation change, along with two fault sizes: fault size 1 with 1 μm change and fault size 2 with 4 μm change.

3.3.2.2 Effect of the Number of Surface Levels in the Detection Performance

We study the effect of selecting different values of surface levels K, which is a parameter in the MST algorithm, in the detection performance. In practice, K is determined based on the characteristics of the topographic values learned from the in-control surfaces. A smaller K may impact the detection performance of the variance change because few surface levels are insufficient to describe the change in the magnitude of topographic values. Similarly, a larger K may affect the detection performance of the autocorrelation change because adding more levels can cause the loss of the spatial autocorrelation structure of topographic values. Thus, a reasonable selection of K is recommended where different patterns of surface changes can be effectively captured.

Table 3.1 shows examples of the effect of selecting different values of K in the detection performance using the proposed monitoring statistic Q. Note that the anomaly surfaces are generated with a 1% local fault under different surface changes (i.e., variance and autocorrelation). It is shown that selecting a few surface levels (e.g., 5 and 10 levels) impacts the detection performance of variance change. However, the detection performance of autocorrelation change is improved because fewer surface levels can lead to higher detection of changes in the spatial autocorrelation of topographic values. In contrast, selecting a larger K (e.g., K=50 and K=100 levels) results in a higher detection of the variance change as shown in Table 3.1. However, the detection performance of autocorrelation change is decreased since a larger K can cause the loss of the detailed structure of surface characteristics. Moreover, a larger K (e.g., $K = 100$) can increase the average computation time of the proposed monitoring statistic Q as shown in Table 3.2.

Therefore, we suggest selecting a reasonable number of surface levels (e.g., 25 levels) to enhance the overall detection performance for variance and autocorrelation and the detection time.

Table 3.1 Detection performance of the monitoring statistic Q under different K

Fault Type	Fault Size		Number of surface levels (K)				
	σ	θ	5	10	25	50	100
Variance	2	1	0.370	0.465	0.583	0.657	**0.683**
	4	1	0.955	0.999	**1.000**	**1.000**	**1.000**
	6	1	0.993	**1.000**	**1.000**	**1.000**	**1.000**
Autocorrelation	1	2	**0.119**	0.118	0.116	0.115	0.111
	1	4	**0.294**	0.260	0.269	0.258	0.267
	1	6	**0.420**	0.395	0.407	0.397	0.408

Table 3.2 Average computation time of the monitoring statistic Q under different K

Number of surface levels (K)	5	10	25	50	100
Time in seconds (standard deviation)	0.066 (0.003)	0.154 (0.002)	0.395 (0.016)	0.783 (0.012)	1.550 (0.008)

3.3.3 Performance Comparisons

This section assesses the performance of the proposed monitoring approach with the existing approaches stated in Section 2.4. In particular, normal surfaces are generated based on the following surface parameters: $\sigma = 1\ \mu m$ and $\theta = 1\ \mu m$. We also address different fault scenarios by considering two fault types (i.e., variance and autocorrelation), two fault sizes (i.e., 2 and 4), and three fault areas (i.e., two local faults with 0.5% and 1% changes and one global fault with 10% change). Note that we calculate the critical limit for all approaches based on a fixed type I error rate (i.e., $\alpha_3 = 0.05$) to obtain a fair comparison. Moreover, the detection performance is assessed based on the performance measure provided in Equation (3.15), which is defined as the average percentage of accurate detection of anomaly surfaces (Bui and Apley, 2018a)

$$\text{Power of detection } (P) = \frac{1}{V}\sum_{j=1}^{V}\frac{T_j}{N_j}, \tag{3.15}$$

where V is the size of simulation replicates ($V = 100$ replicates), T_j is the number of accurate detection of anomaly surfaces obtained at the j^{th} replicate, and N_j is the size of anomaly surfaces generated at the j^{th} replicate (i.e., $N_1 = N_2 = \ldots = N_V = 100$ samples).

Table 3.3 shows the detection performance of anomaly surfaces with the variance and autocorrelation changes. We notice that S_a and S_q generally show low performance in detecting surface faults since they fail to capture the spatial characteristics among topographic values. Next, S_{psd} has a low detection performance of surface faults because S_{psd} is based on the FFT algorithm, which is ineffective in decomposing the information about local surface faults. Similarly, S_{ds} shows a low performance in detecting surface faults due to the heuristic approach of assigning binary values to the topographic values. Moreover, S_c results in low detection performance of surface faults because the topographic values with abnormal characteristics can be dominated by the in-control values during the computation of S_c.

In contrast, it is shown that the SR profile is effective in capturing the variance change at the top and bottom surface levels and the autocorrelation change at the central surface levels. Thus, the proposed statistic Q, which captures the deviation of the observed SR profile from the mean SR profile, reveals a superior detection performance for both variance and autocorrelation changes under all defined fault scenarios and outperforms the existing monitoring approaches as shown in Table 3.3. We also notice that when the fault area increases, we achieve a better detection performance. Finally, the proposed approach is more sensitive to detect global faults associated with 5% or more surface charge than local faults associated with less than 5% surface charge.

Table 3.3 Power of detection of anomaly surfaces under different surface changes

Fault type	Fault area	Fault size		Monitoring approaches					
		σ	θ	S_a	S_q	S_{psd}	S_{ds}	S_c	Q
Variance	0.5%	2	1	0.058	0.060	0.060	0.059	0.072	**0.178**
		4	1	0.067	0.110	0.110	0.064	0.070	**0.889**
	1%	2	1	0.062	0.070	0.070	0.056	0.068	**0.600**
		4	1	0.076	0.241	0.240	0.062	0.066	**1.000**
	10%	2	1	0.100	0.982	0.979	0.054	0.053	**1.000**
		4	1	0.166	0.999	0.980	0.056	0.058	**1.000**
Autocorrelation	0.5%	1	2	0.069	0.066	0.065	0.064	0.075	**0.085**
		1	4	0.079	0.075	0.075	0.075	0.092	**0.127**
	1%	1	2	0.068	0.068	0.067	0.066	0.083	**0.107**
		1	4	0.118	0.117	0.117	0.102	0.134	**0.243**
	10%	1	2	0.097	0.117	0.117	0.080	0.135	**1.000**
		1	4	0.241	0.252	0.252	0.191	0.206	**1.000**

In Table 3.4, we show the effectiveness of applying single and multiple thresholding algorithms on the detection performance of topographic faults. In the single thresholding algorithm, the surface is initially binarized using a single threshold obtained by an edge detection filter (e.g., Canny filter), and then the spatial randomness test (SRT), denoted as T, is computed based on Equation (3.7) (Taam and Hamada, 1993). In the multi-level thresholding algorithms, the proposed monitoring statistic is calculated based on both the existing OMT algorithm, denoted as R, and the proposed MST algorithm Q. It is noticed that R shows a better detection performance than T in detecting changes in topographic surfaces. As discussed earlier in Section 3.2.1, selecting only a single threshold may lead to the removal of important characteristics of topographic values, which can yield low detection performance of surface changes. In addition, we observe that the proposed statistic Q outperforms T and R under both variance and autocorrelation changes. As presented earlier in Figure 3.1, the proposed MST algorithm is shown to be effective in preserving the autocorrelation structure of topographic values by constructing accumulative surface levels where each level contains the spatial information about the topographic values of the previous levels, which results in effective monitoring of topographic faults.

Table 3.4 Effect of the proposed MST algorithm in the detection performance

Fault type	Fault area	Fault size		Monitoring approaches		
		σ	θ	T	R	Q
Variance	0.5%	2	1	0.048	0.069	**0.178**
		4	1	0.049	0.100	**0.889**
	1%	2	1	0.050	0.299	**0.600**
		4	1	0.045	0.917	**1.000**
Autocorrelation	0.5%	1	2	0.051	0.053	**0.085**
		1	4	0.106	0.113	**0.127**
	1%	1	2	0.100	0.106	**0.107**
		1	4	0.138	0.143	**0.243**

3.4 A Case Study of Copper Wafer Surfaces

In semiconductor manufacturing, hundreds of integrated circuits (ICs) or chips are assembled on a single wafer made of different materials such as copper and silicon (Zhang *et al.*, 2016). These ICs have different functions based on the type of electronic devices to be used in. However, surface faults such as pits, ridges, and scratches commonly appear on the surface topography of wafers during semiconductor manufacturing. These faults can form nucleation sites for corrosion and cracks that can cause low functional integrity of ICs, which yields a low performance of electronic devices (Rao *et al.*, 2015a, Rao *et al.*, 2015b). Therefore, monitoring the surface topography of wafers is needed for enhancing the quality of its manufacturing process.

There are various faults that might be observed during the manufacture of semiconductor wafers. In particular, pits and ridge faults may appear on the surface topography of wafers because of the chemical aging and uneven temperatures, whereas scratch faults may be observed due to the material handling and shipping of wafers (Jeong *et al.*, 2008). Such faults are important indicators that should be addressed for accurate analyses of process changes. Figure 3.5 presents an example of the surface topography of a non-smooth and smooth copper wafer obtained using a laser interferometer (Rao *et al.*, 2015a). Note that the non-smooth copper wafer is associated with different faults such as pits, ridges, and scratches.

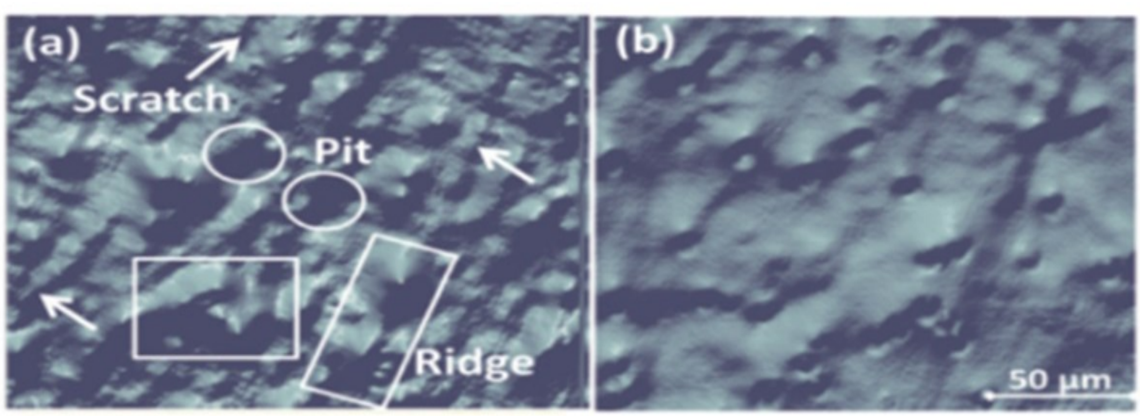


Figure 3.5 Example of the surface topography of (a) a non-smooth and (b) smooth copper wafer obtained using a laser interferometer (Rao *et al.*, 2015a).

Following Rao *et al.* (2015b), we obtain the topography of one normal and three anomaly wafer surfaces associated with the following three types of wafer defects: scratches, pits, and ridges as shown in Figure 3.6. In particular, Figure 3.6 (a) illustrates the surface topography of a normal wafer. In Figure 3.6 (b), we show the surface topography of an anomaly wafer with random scratch defects (i.e., inclined lines with different lengths) that cover 1% of the total surface area. Figure 3.6 (c) depicts an anomaly wafer surface associated with random pits (i.e., scattered points) that allocate 1% of the surface area. Finally, in Figure 3.6 (d), we present an anomaly surface with random ridge defects (i.e., clustered points with different sizes) that cover 2% of the surface area.

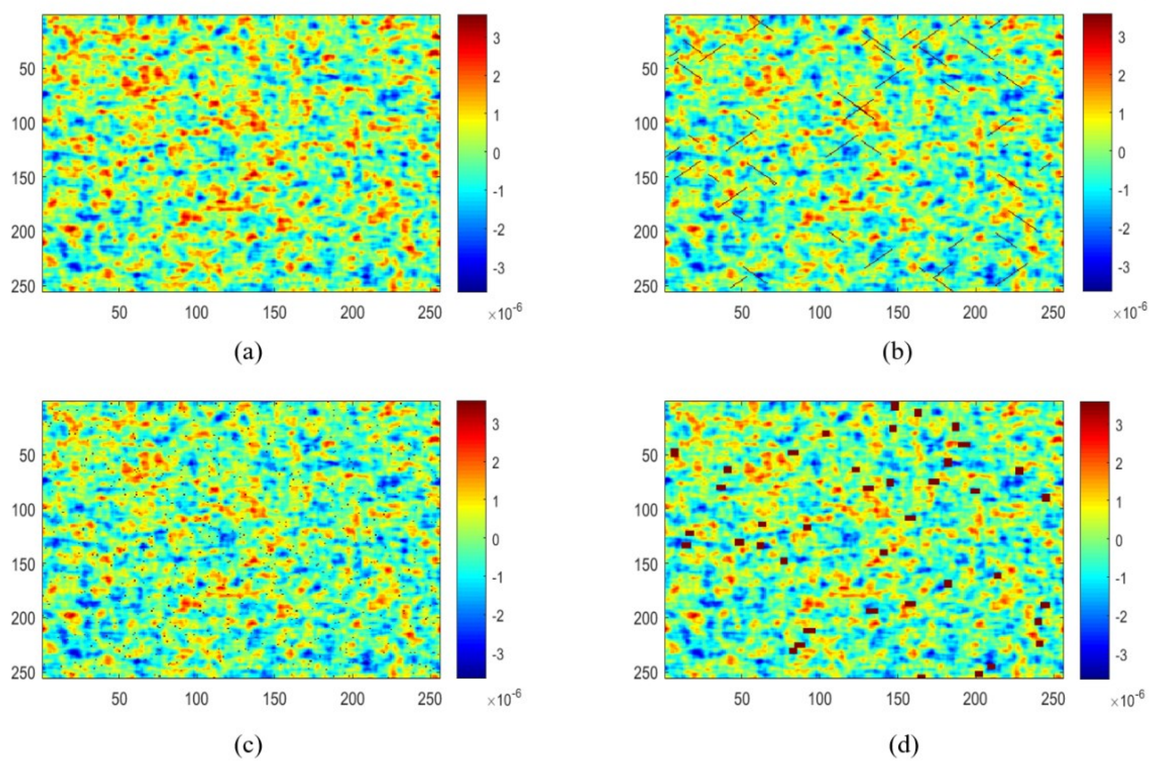


Figure 3.6 Illustrations of (a) normal wafer surface, and three anomaly wafer surfaces with (b) scratch defects, (c) pit defects, and (d) ridge defects.

Table 3.5 shows the performance comparison between the proposed and the existing monitoring approaches discussed in Section 2.4. The performance comparison includes different fault types (i.e., scratches, pits, and ridges) and fault areas (i.e., 0.3%, 0.6%, and 1% surface changes). Note that the scratches, pits, and ridges can be classified as an autocorrelation change because they mainly impact the spatial structure of topographic surfaces. Since the SR profile is shown to be powerful in quantifying this change at the central surface levels, Q results in superior detection performance and outperforms the existing monitoring approaches at all defined fault scenarios. Therefore, Q is proven to be effective in monitoring the quality of wafer surfaces.

Table 3.5 Power of detection of anomaly copper wafers

Fault type	Fault area	Monitoring approaches					
		S_a	S_q	S_{psd}	S_{ds}	S_c	Q
Scratches	0.3%	0.091	0.100	0.100	0.053	0.063	**0.377**
	0.6%	0.176	0.208	0.208	0.078	0.062	**0.911**
	1%	0.355	0.438	0.437	0.115	0.061	**1.000**
Pits	0.3%	0.090	0.095	0.095	0.059	0.065	**0.241**
	0.6%	0.164	0.191	0.191	0.071	0.064	**0.771**
	1%	0.320	0.400	0.398	0.105	0.067	**0.997**
Ridges	0.3%	0.091	0.097	0.097	0.058	0.060	**0.543**
	0.6%	0.176	0.203	0.203	0.074	0.058	**0.991**
	1%	0.344	0.419	0.419	0.109	0.061	**1.000**

3.5 Conclusions

In this chapter, we propose the multi-level spatial randomness approach for characterizing different changes in 3D topographic surfaces. It is shown that the developed MST algorithm enhances the representation of topographic values by decomposing the surface image into different binary images. Besides, quantifying the spatial characteristics of topographic values at each surface level through the proposed SR profile is shown to improve the description of surface changes, including both variance and autocorrelation. The SR profile is also proven to be effective in characterizing the spatial and random properties of topographic values. Moreover, the proposed monitoring statistic Q shows high performance in detecting changes in different types of topographic surfaces and subsequently outperforms the existing monitoring approaches. The proposed statistic Q is also shown to be more superior in detecting global changes.

CHAPTER 4

MULTI-LABEL SEPARATION-DEVIATION SURFACE MODEL FOR MONITORING LOCAL VARIATIONS IN 3D TOPOGRAPHIC SURFACES

4.1 Introduction

Topographic data hold critical information about the quality of finished products and manufacturing processes as discussed in Section 1.1. Topographic defects commonly appear in local patterns during manufacturing. Due to the complex structure of topographic data, existing monitoring approaches lack the detection of local defects as discussed in Section 2.4. In Chapter 3, although the proposed approach is shown to be effective in detecting both variance and autocorrelation changes with global changes, the approach appeared to be less effective in detecting local changes. This is because the local surface characteristics are likely to be dominated by the global characteristics during the calculation of spatial randomness at each surface level. The approach also has a limitation in dealing with noisy characteristics such as measurement errors, which may impact its detection performance. Therefore, in this chapter, we address these limitations by proposing an effective monitoring approach for accurate detection of local defects in topographic surfaces.

Boykov *et al.* (2001) develop an effective model, which is called as separation-deviation (SD), to improve the representation of image characteristics. The SD model segments the image pixels into predefined labels by minimizing an objective function that includes two terms: the deviation term, which captures the similarity within image pixels, and the separation term, which captures the similarities between adjacent image pixels. However, applying the SD model to surface monitoring has some limitations that may impact its performance in segmenting local surface defects. In particular, the model does not consider the normal surface characteristics during the segmentation of new surfaces, and the assigned labels are not accurately mapped to the value (or height) of topographic pixels. Thus, we overcome these limitations by proposing a multi-label separation-deviation surface (MSS) model for effective monitoring of local defects in topographic surfaces. Note that Chapter 4 is mostly based on the following published paper: ALQAHTANI, M. A., JEONG, M. K. and ELSAYED, E. A. 2020b. Multi-label separation-deviation surface model for detecting spatial defects in topographic surfaces. *IEEE Transactions on Industrial Informatics*. Accepted.

The proposed MSS model enhances the representation of local defects by assigning each topographic pixel of a new surface into predefined labels according to the characteristics of normal surfaces (Alqahtani *et al.*, 2020b). This results in an effective representation of topographic characteristics due to the labeling of important characteristics and the smoothing of noisy characteristics. Two statistical features are also presented for characterizing various local surface changes. The MSS feature is presented for quantifying variations within the assigned labels, and the generalized spatial randomness (GSR) feature with optimal weights is derived for quantifying variations in the spatial autocorrelation between the assigned labels. These two features are integrated into a single monitoring statistic to assess local changes in topographic surfaces.

The remainder of this chapter is as follows. The existing SD model and its limitations to surface monitoring are reviewed in Section 4.2. The developed approach is presented in

Section 4.3. The performance of the proposed approach is validated by conducting extensive simulation studies in Section 4.4 and analyzing a case study from the semiconductor industry in Section 4.5. Conclusions are discussed in Section 4.6.

4.2 Separation-deviation Model for Surface Monitoring

Image segmentation algorithms are effective tools for improving the representation of image characteristics. These algorithms mainly aim to segment image pixels into homogenous and similar labels to capture the meaningful and important characteristics within an image (Hochbaum, 2011). Some image segmentation algorithms, such as *k*-means and *k*-medoids, segment image pixels based on the similarities of the intensity of each pixel where pixels with similar intensities are assigned to the same label. Other algorithms, such as the watershed algorithm, segment image pixels based on the similarity between neighboring pixels where nearby pixels are assigned to the same label (Wang, 2010). Segmenting image pixels based on only the similarities within pixel intensities or the similarities between adjacent pixels may yield a low representation of image characteristics. Thus, the separation-deviation (SD) model is introduced to simultaneously capture the similarities within and between image pixels for the effective representation of image characteristics (Boykov *et al.*, 2001, Hochbaum, 2011). The SD model includes the deviation term, which captures the similarity within image pixels, and the separation term, which captures the similarities between adjacent image pixels.

The objective of the SD model is to convert a topographic surface $Z = \{z_i; i = 1, ..., m\}$ to a discrete surface $X = \{x_i; i = 1, 2, ..., m\}$, where x_i is a discrete number that belongs to one of the k predefined labels, $x_i \in K = \{1, 2, ..., k\}$ (Boykov *et al.*, 2001, Hochbaum, 2011).

The assigned labels are obtained by minimizing the sum of deviation cost $D_i(.)$ and the separation cost $S_{ij}(.)$ over all topographic values as shown in Equation (4.1)

$$X = \underset{x_i \in K}{\operatorname{argmin}} \sum_{i \in Z} D_i(x_i) + \alpha \sum_{i \in Z} \sum_{j \in \Omega(i)} S_{ij}(x_i, x_j), \tag{4.1}$$

where α is a smoothing parameter that controls the smoothness of segmentation results and $\Omega(i)$ is a specified neighborhood construction rule. Two popular neighborhood construction rules are commonly used to define $\Omega(i)$: the rock-move neighborhood (RMN) and the king-move neighborhood (KMN). The RMN construction rule considers the four neighbors positioned vertically and horizontally adjacent to a topographic pixel, while the KMN construction rule considers the eight neighbors positioned in the vertical, horizontal, and diagonal of a topographic pixel. In the SD model, the deviation term captures the similarity of topographic values by minimizing the deviation between the observed topographic values z_i and their expected values. The separation (or smoothing) term captures the similarities between assigned labels by minimizing the separation between the assigned labels x_i and their neighbors x_j.

The use of the SD model in surface topography monitoring has some limitations. In particular, the SD model ignores the utilization of the characteristics of the normal surfaces during the segmentation process. Besides, the model does not consider the relationships between the assigned labels and the topographic values (heights). Such limitations may lead to a lower performance of segmenting changes in topographic surfaces. Therefore, we overcome these limitations by developing an effective monitoring approach based on the proposed multi-label separation-deviation surface (MSS) model for accurate monitoring of changes in topographic surfaces.

4.3 Developed Approach

The developed approach has the following three stages; Stage 1: multi-label surface segmentation, Stage 2: feature extraction, and Stage 3: anomaly detection. In Stage 1, we enhance the description of local topographic defects by segmenting the topographic values into predefined labels through the developed MSS model. Stage 2 extracts two features from the obtained labels for quantifying changes in surface characteristics. Finally, in Stage 3, anomaly surfaces are detected using the extracted features. Figure 4.1 shows a summary of the developed approach (Alqahtani *et al.*, 2020b).

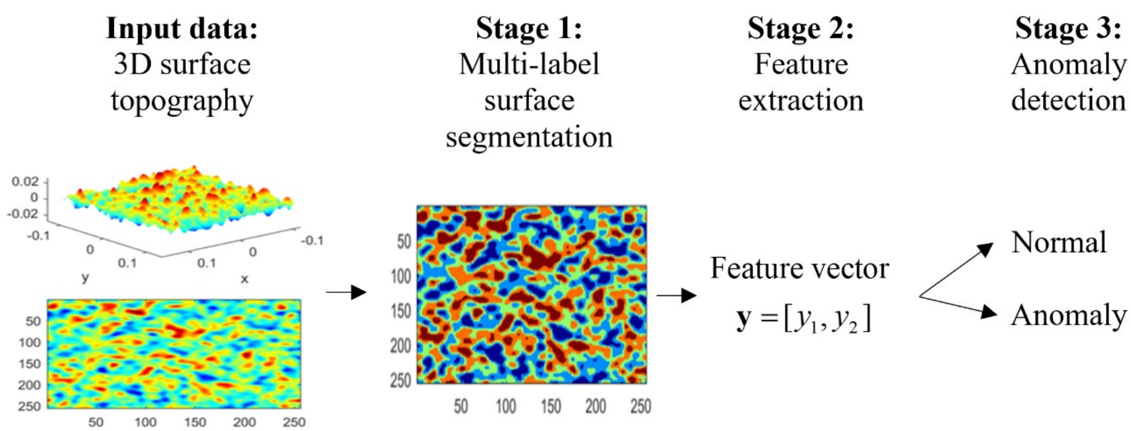


Figure 4.1 Flow diagram of the developed approach

4.3.1 Stage 1: Multi-label Surface Segmentation

We overcome the limitations of the existing SD model by proposing a multi-label separation-deviation surface (MSS) model for online monitoring of 3D topographic surfaces. The MSS model assigns each topographic value z_i in a new (or observed) surface into one of the k predefined labels, $x_i \in K = \{1, 2, ..., k\}$, by minimizing the sum of the deviation and separation costs as given in Equation (4.2)

$$X = \underset{x_i \in K}{\operatorname{argmin}} \sum_{l \in K} \sum_{i \in C_l} D_i(z_i, \mu_l^0, \sigma_l^0) + \alpha \sum_{i \in Z} \sum_{j \in \Omega(i)} S_{ij}(x_i, x_j), \tag{4.2}$$

where α is a smoothing parameter, (μ_l^0, σ_l^0) are the normal mean and standard deviation of the label l, $l = 1, 2, ..., k$, C_l is the set of values whose label is l, and $\Omega(i)$ is a specified neighborhood construction rule. In the MSS model, the Mahalanobis function is utilized as a deviation function, which is defined in Equation (4.3)

$$\sum_{l \in K} \sum_{i \in C_l} D_i(z_i, \mu_l^0, \sigma_l^0) = \sum_{l \in K} \sum_{i \in C_l} \frac{(z_i - \mu_l^0)^2}{\sigma_l^0}. \tag{4.3}$$

From Equation (4.3), the distance between each observed topographic value z_i and the normal mean value of each label μ_l^0, $l = 1, 2, ..., k,$ is calculated and then scaled by the normal standard deviation of each label σ_l^0, $l = 1, 2, ..., k$. Subsequently, each topographic value, $i \in Z$, is assigned to a label set that has the closest distance $\{C_1, C_2, ..., C_k\}$ to minimize the deviation cost.

We propose Algorithm 4.1 to estimate the parameters of the normal mean and standard deviation of each label $\{\mu_l^0, \sigma_l^0; l = 1, 2, ..., k\}$. Specifically, we consider n normal topographic surfaces, denoted as $Z^{(j)} = \{z_i^{(j)}; i = 1, ..., m\}$, $j = 1, ..., n$, where $z_i^{(j)}$ is the topographic value of the pixel i measured from the surface j and m is the total number of pixels as shown in the for-loop in steps 1 and 8. In step 2, we utilize the k-means algorithm to efficiently obtain the segmentation results. Particularly, the k-means algorithm has two main steps: the assignment and update steps. In the assignment step, the algorithm assigns each topographic value of the pixel i to a label set, $i \in R_l^{(j)}$, $l = 1, 2, ..., k$, that has the nearest mean by minimizing the sum of the squared Euclidean distance between each topographic value and a predefined mean of each label set. Then, in the update step, the algorithm updates the mean of each label set $\eta_l^{(j)}$, $l = 1, 2, ..., k$. The k-means algorithm repeats these two steps until the assigned labels do not change. Consequently, the mean and standard deviation are calculated from the topographic values of the final obtained label sets $\{\mu_l^{(j)}, \sigma_l^{(j)}; l = 1, 2, ..., k\}$ as explained in steps 3 to 5. In step 6, we rank the mean of each label in increasing order to connect each assigned label set to the magnitude of its topographic values. As a result, $\mu_{(1)}^{(j)}$ represents the mean of the set with the lowest topographic values, whereas $\mu_{(k)}^{(j)}$ represents the mean of the set with the highest topographic values. In step 7, we map the standard deviation of each label to its ranked mean. Finally, the normal mean and standard deviation of each label are calculated by taking the mean over the values obtained from given normal surfaces as presented in steps 9 to 11.

Algorithm 4.1: Estimating the normal mean and standard deviation for each label $\{\mu_l^0, \sigma_l^0; l = 1,2,...,k\}$

Given the total number of the assigned labels k and normal topographic surfaces $Z^{(j)}$, $j = 1,..., n$ (e.g., n=1000), we apply the following steps:

1: **for** $j = 1,2,...,n$ **do**

2: assign each topographic value into one of the k labels by minimizing the following function

$$\min \sum_{l=1}^{k} \sum_{i \in R_l^{(j)}} (z_i^{(j)} - \eta_l^{(j)})^2 ,$$

where $\eta_l^{(j)}$ is the mean of topographic values with the label l obtained from the surface j and $R_l^{(j)}$ is the set of assigned values with the label l obtained from the surface j.

3: **for** $l = 1,2,...,k$ **do**

4: calculate the mean and standard deviation of the topographic values that corresponds to the label l

$$\mu_l^{(j)} = \frac{1}{n_l^{(j)}} \sum_{i \in R_l^{(j)}} z_i^{(j)}, \text{ and } \sigma_l^{(j)} = \sqrt{\frac{\sum_{i \in R_l^{(j)}} (z_i^{(j)} - \mu_l^{(j)})^2}{(n_l^{(j)} - 1)}} ,$$

where $(\mu_l^{(j)}, \sigma_l^{(j)})$ are the mean and standard deviation of the topographic values with the label l obtained from the surface j and $n_l^{(j)}$ is the total number of the assigned values with the label l obtained from the surface j.

5: **end for**

6: rank the mean of each label in increasing order

$$\mu_{(1)}^{(j)} < \mu_{(2)}^{(j)} < ... < \mu_{(k)}^{(j)} ,$$

where $\{\mu_{(1)}^{(j)}, \mu_{(2)}^{(j)}, ..., \mu_{(k)}^{(j)}\}$ are the order statistics of the mean of each label.

7: map the standard deviation of each label to its corresponding ranked mean

$$\{\sigma_{(1)}^{(j)}, \sigma_{(2)}^{(j)}, ..., \sigma_{(k)}^{(j)}\} .$$

8: **end for**

9: **for** $l = 1,2,...,k$ **do**

10: calculate the normal mean and standard deviation

$$\mu_l^0 = \frac{1}{n} \sum_{j=1}^{n} \mu_{(l)}^{(j)} \text{ and } \sigma_l^0 = \frac{1}{n} \sum_{j=1}^{n} \sigma_{(l)}^{(j)} .$$

11: **end for**

The Potts function is defined as a discontinuity function, which indicates whether the assigned pixels are disconnected with its neighbors (Boykov *et al.*, 2001). More specifically, the Potts function penalizes each assigned pixel x_i by a defined smoothing parameter α if it does not match its neighbors x_j. Therefore, we utilize the Potts function as a separation function to capture the similarities between the assigned labels and their neighbors as given in Equation (4.4)

$$\alpha\sum_{i\in Z}\sum_{j\in\Omega(i)} S_{ij}(x_i, x_j) = \alpha\sum_{i\in Z}\sum_{j\in\Omega(i)} I(x_i \neq x_j), \tag{4.4}$$

where $I(.)$ is an indicator function defined as $I(x_i \neq x_j) = 1$, otherwise zero. Therefore, the MSS model segments the topographic values of a new surface into one of the k predefined labels with respect to the characteristics of normal surfaces. In addition, the obtained labels are accurately mapped to the magnitude of topographic values due to the use of the ranked normal means in the deviation function. As a result, the assigned values with the label 1 $(x_i = 1)$ constantly represent the region with the lowest topographic values (or lowest valley), and the assigned values with the label k $(x_i = k)$ always represent the region with the highest topographic values (or highest peak). Finally, we utilize the graph cut algorithm to efficiently obtain the solution of the MSS model (Boykov *et al.*, 2001).

The developed MSS model has several advantages in monitoring topographic surfaces. The model is effective in enhancing the representation of topographic defects (Alqahtani *et al.*, 2020b). Specifically, when an anomaly surface is observed, we expect the structure of that surface to be distinguishable from the structure of the normal surfaces. This is because any abnormal topographic value is assigned to the nearest normal mean for minimizing the objective function of the MSS model. This causes labels 1 and k to be assigned more than other labels to minimize the objective function, which subsequently yields a significant change in the structure of the assigned labels. In addition, the model is superior in dealing with noisy values by controlling the smoothing parameter α in which a higher value of α results in smoother segmentation results and vice versa. The model is also efficient for online monitoring since the computation time of the graph cut algorithm is linear with the number of assigned labels $O(k)$ (Boykov *et al.*, 2001).

4.3.2 Stage 2: Feature Extraction

In this section, we introduce two features for quantifying various topographic defects based on the label assignments obtained by the developed MSS model. Specifically, the multi-label separation-deviation surface (MSS) feature is suggested for capturing the deviation of the observed topographic values from their respective normal mean values. The MSS feature, denoted as y_1, is defined as the objective value of minimizing the MSS model as given in Equation (4.5)

$$\begin{aligned} &\text{Minimize} \quad \sum_{l\in K}\sum_{i\in C_l}\frac{(z_i-\mu_l^0)^2}{\sigma_l^0}+\alpha\sum_{i\in Z}\sum_{j\in\Omega(i)} I(x_i\neq x_j) , \\ &\text{Subject To} \quad x_i\in K \qquad \forall i\in Z \end{aligned} \tag{4.5}$$

The MSS feature reflects how the characteristics of an observed surface are similar to those of normal surfaces. A higher value of the MSS feature implies that the characteristics of observed topographic values deviate from the normal characteristics and vice versa. Therefore, it is expected the corresponding MSS values of normal surfaces to be statistically similar since they share similar surface characteristics. However, when some defects appear on the topography of an observed surface, this is likely to cause a deviation in the characteristics of the topographic values. As a result, a significant increase in the value of the MSS feature is expected due to the increase in the total cost of minimizing the deviation and separation functions. Thus, the MSS feature is powerful in capturing deviations in surface characteristics.

The spatial randomness (SR) feature is applied for identifying the presence of the spatial autocorrelation in observed binary surfaces (Hansen and Thyregod, 1998). The SR feature enables the evaluation of surface uniformity by examining whether an observed binary value of a pixel at one location is independent of the values of that pixel at neighboring locations. The SR feature is measured using the concept of join-count (JC) statistic, which computes the actual count of the 0-to-0 and 1-to-1 joins under a defined neighborhood construction rule. Therefore, for the given binary surface of a product $X = \{x_i; i = 1, 2, ..., m\}$, $x_i \in \{0,1\}$, the JC statistics of the 0-to-0 and 1-to-1 joins are calculated using Equation (4.6), respectively,

$$J_{00} = \sum_{i \in X} \sum_{j \in \Omega(i)} (1 - x_i)(1 - x_j), \text{ and } J_{11} = \sum_{i \in X} \sum_{j \in \Omega(i)} x_i x_j, \tag{4.6}$$

where (J_{00}, J_{11}) are the JC statistics of label 0 and label 1, respectively, and $\Omega(i)$ is a defined neighborhood construction rule of the pixel i. The SRT is given in Equation (4.7)

$$\mathrm{S} = \beta_0 \, J_{00} + \beta_1 \, J_{11}, \tag{4.7}$$

where β_0 and β_1 are the assigned weights for the 0-to-0 and 1-to-1 joins, respectively. Note that β_0 and β_1 are defined as the probability of not observing the labels 0 and 1 in the given binary surface, respectively. Specifically, β_0 and β_1 are calculated as $\beta_0 = m_1 / m$ and $\beta_1 = m_0 / m$, where m_0 and m_1 are the number of pixels with the labels 0 and 1, respectively, and m is the total number of the image pixels (Jeong *et al.*, 2008). The SRT is commonly used for detecting the presence of the spatial autocorrelation among values in binary surfaces. Since the 3D topographic surfaces naturally contain values that are spatially autocorrelated with their neighbors, the SR feature can be powerful not only in detecting but in monitoring the level of the spatial autocorrelation between the topographic values. However, the SR feature is applicable for binary surfaces only. Since the topographic surfaces are expressed as continuous data, the existing SR feature is unsuitable in monitoring these surfaces.

We overcome the aforementioned drawback by generalizing the existing SR feature from the binary case to a more general case (segmented case) in which the spatial randomness of each assigned label obtained by the MSS model is quantified and monitored. Thus, we develop a generalized spatial randomness (GSR) feature to quantify and capture the spatial characteristics among each assigned label. In particular, we obtain the GSR feature by first generalizing the JC statistics for all k assigned labels. Specifically, the JC statistic for each assigned label is obtained by measuring the actual count of the l-to-l join, $l = 1, 2, ..., k$, according to a defined neighborhood construction rule $\Omega(i)$. Therefore, for the obtained segmented surface $X = \{x_i; i = 1, 2, ..., m\}$, $x_i \in \{1, 2, ..., k\}$, we propose the generalized JC statistic of the label l, $l = 1, 2, ..., k$, which is given in Equation (4.8)

$$J_{ll} = \sum_{i \in X} \sum_{j \in \Omega(i)} I(x_i = l)\, I(x_j = l), \tag{4.8}$$

where $I(.)$ is an indicator function defined as $I(x_i = l) = 1$, and $I(x_j = l) = 1$, otherwise zero. Subsequently, the spatial randomness feature of the l^{th} label, $l = 1, 2, ..., k$, is derived using Equation (4.9)

$$S_l = d_l J_{ll} + d_r J_{rr}, \tag{4.9}$$

where (d_l, d_r) are the assigned weights of label l and label r, J_{ll} is the JC statistic of the label l, and J_{rr} is the JC statistic of the label r, which is defined as the total sum of the JC statistics without the label l, $J_{rr} = \sum_{j=1|j\neq l}^{k} J_{jj}$. We develop Lemma 4.1 to obtain the optimal weights by minimizing the variance of S_l (the proof of Lemma 4.1 is given in Appendix A).

Lemma 4.1: *For the spatial randomness feature of the label* l, S_l , $l = 1, 2, ..., k$, *the optimal weights which minimize the variance of* S_l *subjected to* $d_l + d_r = 1$ *are obtained* $(d_l^*, d_r^*) = (\beta_r, \beta_l)$, *where* $\beta_l = n_l / m$ *is the probability of assigning the label* l, *which is measured by dividing the number pixels with the label* l, n_l, *over the total number of surface pixels,* m, *and* $\beta_r = \sum_{j=1|j\neq l}^{k} \beta_j = 1 - \beta_l$ *is the probability of not observing the label* l.

According to Lemma 4.1 and Equation (4.9), we propose the generalized spatial randomness (GSR) feature by integrating the spatial randomness feature of each assigned label as given in Equation (4.10)

$$y_2 = \sum_{l=1}^{k} \sum_{j=1|j\neq l}^{k} \beta_j J_{ll}, \tag{4.10}$$

Note that when the total number of labels is "two" ($k = 2$), the GSR is reduced to the existing SR feature.

The developed GSR feature is robust to the random behavior among topographic values. In particular, the corresponding values of the GSR feature of the normal surfaces are expected to be similar even though the topographic values of these surfaces do not match their locations. This is because the GSR feature is measured independently from the actual locations of the topographic values. In addition, the GSR feature is powerful in monitoring deviations in the spatial structure between the assigned labels. Specifically, when an observed surface contains defects (anomaly surface), then the spatial structure of the assigned labels of that surface is expected to change from the structure of normal surfaces. This change can be effectively captured by quantifying the spatial randomness level of all assigned labels through the developed GSR feature.

4.3.3 Stage 3: Anomaly Detection

The goal of Stage 3 is to identify whether the new surface $Z^{(new)}$ is an anomaly based on given n normal surfaces, $Z^{(j)}$, $j = 1, 2, ..., n$. Specifically, after a new surface $Z^{(new)}$ is observed, the surface is converted to a segmented surface $X^{(new)}$ by applying the MSS model. After the MSS feature $y_1^{(new)}$ and the GSR feature $y_2^{(new)}$ are both computed from $X^{(new)}$, we integrate them into a single feature vector $\mathbf{y}^{(new)} = [y_1^{(new)}, y_2^{(new)}]$. Finally, the monitoring statistic is calculated using Equation (4.11)

$$T^{(new)} = (\mathbf{y}^{(new)} - \overline{\mathbf{y}}_0)' \, \hat{\mathbf{\Sigma}}_0^{-1} (\mathbf{y}^{(new)} - \overline{\mathbf{y}}_0) , \tag{4.11}$$

where $\overline{\mathbf{y}}_0$ and $\hat{\mathbf{\Sigma}}_0^{-1}$ are the normal mean and covariance matrix of the feature vector computed from normal surfaces. Therefore, when the monitoring statistic of a new surface $T^{(new)}$ exceeds a defined critical value C, then, we identify the new surface $Z^{(new)}$ as an anomaly. Note that C is defined as $C = [2(n-1)/n-2]\ F_{\delta,2,n-2}$, where n is the total number of normal surfaces and δ is a prespecified false positive alarm rate, which is defined as the percentage of incorrectly identifying the normal surfaces as anomaly surfaces (Montgomery, 2007).

4.4 Performance Analysis

In the simulation studies, we utilize the Gaussian random model discussed in Section 2.3 to generate topographic surfaces. In addition, anomaly surfaces are generated under different defect types, sizes, and areas to mimic various real-life defect scenarios as described earlier in Section 3.3.1.

4.4.1 Analysis of the Developed MSS Model

In this section, we study the effect of the smoothing parameter α and the number of assigned labels k on the segmentation results. The smoothing parameter α is defined as a trade-off parameter that controls the effect of the deviation and separation terms in the MSS model. If the deviation term dominates the separation term, each topographic value tends to be assigned to the label set with the nearest normal mean. In contrast, if the separation term is weighted more than the deviation term, then a stronger influence is set on the boundary of each assigned label, which results in smoother segmentation results. The value of α is selected according to the behavior of the topographic surfaces under normal process behavior. A lower value of α (e.g., $\alpha = 0.25$) is recommended for surfaces with few noisy values, while a larger value of α (e.g., $\alpha = 15$) is recommended for surfaces with many noisy values.

The number of assigned labels k is another important parameter of the MSS model. In practice, the selection of k is determined according to the characteristics of normal surfaces. A larger k (e.g., $k=10$) is recommended for surfaces composed of dense characteristics, whereas a smaller k (e.g., $k=3$) is recommended for surfaces composed of sparse characteristics. Figure 4.2 displays the result of specifying different values of α and k on segmenting a topographic surface. In the simulation studies, we select $\alpha=0.25$ and $k=3$ since the generated normal topographic data contain sparse characteristics and less noise.

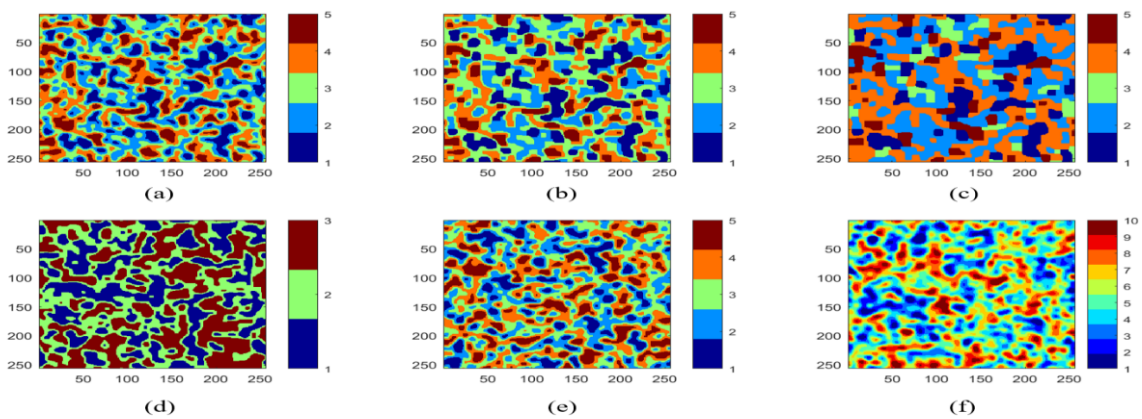


Figure 4.2 Segmentation results under different values of α and k: (a-c) $\alpha=$ 0.25, 5, 15, and (d-f) $k=3, 5, 10$.

We also study the effectiveness of the developed MSS model in capturing topographic defects. Figure 4.3 shows examples of normal and anomaly surfaces before and after being segmented into three labels $(k=3)$ using the developed MSS model. We observe that the structure of the assigned labels of the two generated anomaly surfaces can be clearly distinguished from the structure of the normal surface. Particularly, when the variance change occurs, labels 1 and 3 tend to be assigned more than the other labels as shown in Figure 4.3 (b). This is because the MSS model minimizes the deviation function by assigning the topographic values with abnormal negative values (or abnormal valleys) to the label set with the lowest normal mean (label 1) and the topographic values with abnormal positive values (or abnormal peaks) to the label set with the highest normal mean (label 3). Moreover, when the autocorrelation change occurs, the topographic values become more similar to their neighbors, which yields the labels assigned by the MSS model to be more grouped and clustered as presented in Figure 4.3 (c).

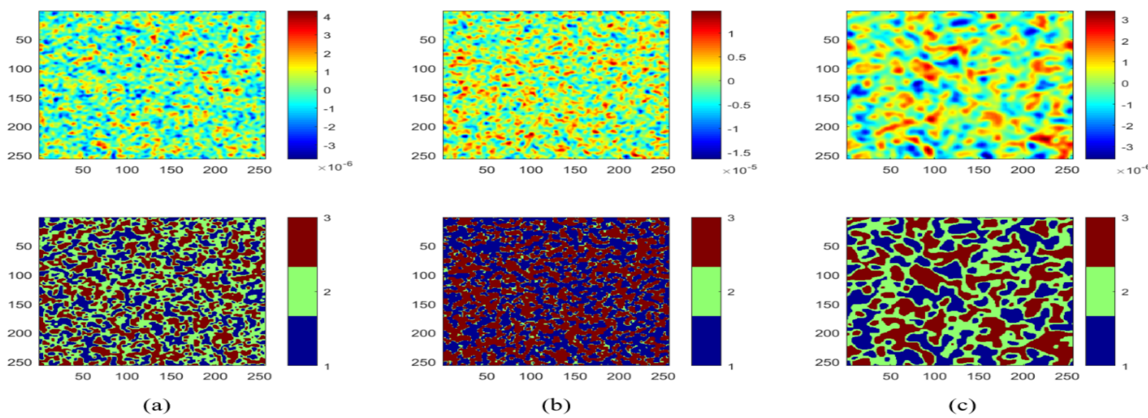


Figure 4.3 Examples of (a) normal surface, and two anomaly surfaces with (b) variance change, and (c) autocorrelation change (first row), along with their corresponding label assignments obtained by the MSS model (second row).

4.4.2 Detection Performance Comparisons

This section compares the detection performance of the developed approach with the traditional monitoring approaches discussed in Section 2.4. In the performance comparisons, we use the performance measure described earlier in Equation (3.15). In particular, Table 4.1 presents the power of detection P of the developed and traditional monitoring approaches using anomaly surfaces obtained by the Gaussian random model. Overall, we notice that the detection performance of the variance change is lower than the detection performance of the autocorrelation change. This is because the variance change primarily affects the variance of topographic values, which is not difficult to capture.

However, the autocorrelation change does not affect the variance of topographic values, but it mainly affects the spatial relationships among topographic values, which is difficult to detect. In addition, we notice that the power spectral density S_{psd} and the density of summits S_{ds} show low detection performance due to the use of the FFT filter as in S_{psd} and edge detection filters as in S_{ds}, which are likely to smooth the local defect information. Similarly, when the Fiedler number S_c is calculated, the topographic values of normal regions are likely to dominate the values of local defects, which results in lower power of detection of S_c. We also observe that the one-sample Anderson–Darling approach S_{ad} shows better performance in detecting the variance change compared to the other traditional approaches due to the utilization of information learned from normal surfaces during the fitting process of the supervised learning model (regression tree). However, S_{ad} yields low performance in detecting the autocorrelation change because the spatial characteristics of local faults can be obscured by the spatial characteristics of normal topographic values.

In contrast, the developed MSS feature is shown to be sensitive in capturing the variance change by quantifying the deviation between the assigned labels and their corresponding normal mean values. In addition, the developed GSR feature is effective in detecting the autocorrelation change by calculating the spatial randomness level for all assigned labels. Therefore, the monitoring statistic T, which calculates the deviation of a vector that includes both features from the normal mean vector, shows superior performance in detecting local topographic defects caused by either variance or autocorrelation change and outperforms the traditional monitoring approaches as presented in Table 4.1. In addition, the proposed approach is shown to be more sensitive to detect the variance change than the autocorrelation since T yields superior detection performance with a small change in the fault size of σ (0.5 μm changes from the normal value) compared to the change in the fault size of θ (2 μm changes from the normal value).

Table 4.1 Power of detection of the developed and traditional monitoring approaches using anomaly surfaces generated by the Gaussian random model

Defect type	Defect area	Defect size		Monitoring approaches				
		σ	θ	S_{psd}	S_{ds}	S_c	S_{ad}	T
Variance	2.0%	1.5	1	0.242	0.273	0.189	0.740	**0.802**
		1.6	1	0.281	0.336	0.223	0.910	**0.964**
	2.25%	1.5	1	0.246	0.281	0.197	0.750	**0.943**
		1.6	1	0.276	0.335	0.222	0.960	**0.994**
Autocorrelation	2.0%	1	3	0.095	0.095	0.112	0.120	**0.681**
		1	3.2	0.096	0.097	0.120	0.200	**0.746**
	2.25%	1	3	0.094	0.095	0.113	0.190	**0.863**
		1	3.2	0.098	0.097	0.116	0.300	**0.910**

We examine the significance of differences between the power of detection results obtained by the developed and traditional monitoring approaches. We apply the Wilcoxon signed test, which is effective in examining the median differences between two data sets. Particularly, we obtain the performance measure (power of detection) for 100 replicates for the developed and traditional monitoring approaches using the fault scenarios considered in Table 4.1. Consequently, after we obtain a vector of size 100 for each approach, we calculate the *p*-values for all pairwise comparisons between the developed and traditional approaches as shown in Table 4.2.

We observe that the power of detection results obtained by the developed approach are significantly different from the traditional approaches under a defined positive false rate (0.05). For example, for the variance change with the defect area of 2% and defect size of 1.5, the power of detection results obtained by T and S_{psd} are statistically different since its calculated *p*-value is significantly less than the defined positive false rate (*p*-value =3.74E-18 <0.05).

Table 4.2 Pairwise comparisons between the power of detection results of the developed and traditional approaches using their p -values

Defect type	Defect area	Defect size		Monitoring approaches			
		σ	θ	S_{psd}	S_{ds}	S_c	S_{ad}
Variance	2.0%	1.5	1	3.74E-18	3.73E-18	3.68E-18	0.011
		1.6	1	3.76E-18	3.75E-18	3.69E-18	0.042
	2.25%	1.5	1	3.72E-18	3.73E-18	3.67E-18	0.024
		1.6	1	3.74E-18	3.78E-18	3.68E-18	0.045
Autocorrelation	2.0%	1	3	3.70E-18	3.73E-18	3.73E-18	3.89E-16
		1	3.2	3.74E-18	3.74E-18	3.78E-18	1.48E-11
	2.25%	1	3	3.67E-18	3.67E-18	3.73E-18	3.10E-16
		1	3.2	3.69E-18	3.71E-18	3.79E-18	1.27E-11

We also perform additional experiments to investigate the effectiveness of the proposed approach under the statistical process control (SPC) environment, i.e., sequential testing. Table 4.3 shows the ARL_1 performance of the developed and traditional monitoring approaches using surfaces generated by the Gaussian random model. Note that we set the in-control ARL_0 for all approaches to be 200 and then calculate the out-of-control ARL_1 under different defect types, sizes, and areas. In addition, we consider one surface at each sampling point (i.e., sample size $n = 1$), and the ARL_1 is obtained based on 1000 replicates. We observe that the developed approach is significantly quicker in detecting both variance and autocorrelation changes than the other monitoring approaches. This is more apparent in the autocorrelation change than in the variance change due to the difficulty of detecting the change in the spatial relationships among topographic data. For example, in the variance change with the defect area of 2% and defect size of 1.5, the ARL_1 of the developed approach is about 1.97, whereas the ARL_1 of the other monitoring approaches are at least around (≥4.34). However, in the autocorrelation change with the defect area of 2% and defect size of 3, the ARL_1 of the developed approach is around 2.78, while the ARL_1 of the other approaches are at least about (≥100.2). This implies the effectiveness and efficiency of the proposed approach in monitoring different topographic changes under the SPC environment.

Table 4.3 ARL_1 performance of the developed and traditional monitoring approaches using anomaly surfaces generated by the Gaussian random model

Defect type	Defect area	Defect size		Monitoring approaches				
		σ	θ	S_{psd}	S_{ds}	S_c	S_{ad}	T
Variance	2.0%	1.5	1	22.42	17.12	19.00	4.34	**1.97**
		1.6	1	16.45	11.16	12.57	1.42	**1.17**
	2.25%	1.5	1	21.07	15.52	18.78	2.69	**1.25**
		1.6	1	15.31	10.88	12.56	1.17	**1.03**
Autocorrelation	2.0%	1	3	166.4	170.2	163.7	100.2	**2.78**
		1	3.2	163.1	161.6	148.4	16.08	**2.39**
	2.25%	1	3	154.2	152.3	137.0	54.27	**1.66**
		1	3.2	147.0	142.4	134.4	12.10	**1.43**

We study the effect of the sample size n, i.e., the number of observed surfaces at each sampling point in the ARL_1 performance. Table 4.4 shows the ARL_1 performance of the developed approach under different sample sizes ($n = 1, 2, 3$). We notice that the ARL_1 performance is dependent on n. More specifically, when n increases, better performance of ARL_1 is achieved. This is because when the sample size of the anomaly surfaces increases at each sampling point, this improves the accuracy of identifying anomaly surfaces.

Table 4.4 Effect of different sample sizes in the ARL_1 performance of the developed approach

Defect type	Defect area	Defect size		Sample size		
		σ	θ	n=1	n=2	n=3
Variance	2.0%	1.5	1	1.967	1.087	**1.008**
		1.6	1	1.167	1.001	**1.000**
	2.25%	1.5	1	1.253	1.005	**1.000**
		1.6	1	1.025	1.000	**1.000**
Autocorrelation	2.0%	1	3	2.783	1.203	**1.014**
		1	3.2	2.388	1.144	**1.016**
	2.25%	1	3	1.663	1.039	**1.002**
		1	3.2	1.426	1.015	**1.001**

Table 4.5 presents the average computation time of obtaining the statistics of the developed and traditional monitoring approaches for surfaces generated by the Gaussian random model. Note that the computation time for each approach is obtained by taking the average over 1000 samples, and the standard deviation for each approach is presented in the parentheses. We observe that the computation time of the developed approach T is larger

than the other approaches (e.g., S_{ds}, S_{psd}, and S_c) due to its modeling complexity. However, the computation time of T is still efficient and smaller than the computation time of other complex approaches such as S_{ad}. Particularly, T takes an average of 0.121 seconds to detect an anomaly surface, whereas S_{ad} takes an average of 5.363 seconds. This makes the developed approach suitable for the online monitoring of topographic surfaces.

Table 4.5 Average computation time of the developed and traditional monitoring approaches

Monitoring approaches	S_{psd}	S_{ds}	S_c	S_{ad}	T
Time in seconds (standard deviation)	0.001 (7.54E-05)	5.18E-05 (6.22E-06)	0.012 (0.002)	5.575 (0.067)	0.121 (1.32E-03)

4.5 A Case Study on Monitoring the Topography of Wafer Surfaces

This section presents the detection performance of the developed approach using the case study from the semiconductor industry explained earlier in Section 3.4. Particularly, Table 4.6 demonstrates the power of detection P of the developed and traditional monitoring approaches using anomaly wafer surfaces. We observe that the appearance of the scratch, pit, and ridge defects can be categorized as an autocorrelation change since these defects mainly affect the spatial autocorrelation among topographic values (Rao *et al.*, 2015b). We also notice that the traditional monitoring approaches fail to identify these defects due to their ignorance of the spatial relationships between the topographic values. In contrast, the proposed approach T yields superior detection performance for all defined defect scenarios due to the effective modeling of the spatial relationships among topographic data and subsequently outperforms the traditional monitoring approaches. Therefore, T shows its robustness in detecting different types of defects that appear on the topography of wafer surfaces.

Table 4.6 Power of detection of anomaly wafer surfaces

Defect type	Defect area	Monitoring approaches			
		S_{psd}	S_{ds}	S_c	T
Scratches	1%	0.105	0.099	0.110	**0.867**
	1.25%	0.108	0.100	0.112	**0.972**
	1.5%	0.110	0.104	0.113	**0.997**
Pits	1%	0.103	0.096	0.112	**0.874**
	1.25%	0.104	0.097	0.115	**0.977**
	1.5%	0.109	0.102	0.114	**0.998**
Ridges	2%	0.113	0.103	0.107	**0.680**
	3%	0.117	0.104	0.111	**0.955**
	5%	0.127	0.106	0.116	**0.999**

Other manufacturing processes may experience direction changes in the surface topography of the finished products due to the wear of machine tools (Zhang *et al.*, 2018, Bui and Apley, 2018b). For example, using a rotating grinder tool produces circular patterns while vertical or horizontal grinder produces completely different patterns. Figure 4.4 (a) shows an example of directional patterns in a ground surface when using a horizontal grinder, while Figure 4.4 (b) shows the effect of the tool wear on the surface finish when using a rotating grinding (Zhang *et al.*, 2018).

(a) (b)

Figure 4.4 Directional patterns on ground surfaces obtained from (a) horizontal grinder, and (b) rotating grinder (Zhang *et al.*, 2018).

We examine the effectiveness of the developed approach in monitoring directional changes. Particularly, we generate directional changes by convoluting a non-Gaussian distribution function (e.g., beta distribution) with a non-Gaussian autocorrelation function (e.g., linear motion). Note that the linear motion includes a shape parameter θ, which controls the direction of the surface characteristics and takes values from 0 to 360 degrees. Figure 4.5 shows three generated non-Gaussian surfaces with one normal sample with no directional change ($\theta = 0$), and two anomaly samples with mild and severe directional changes ($\theta = 20$ and $\theta = 60$), respectively.

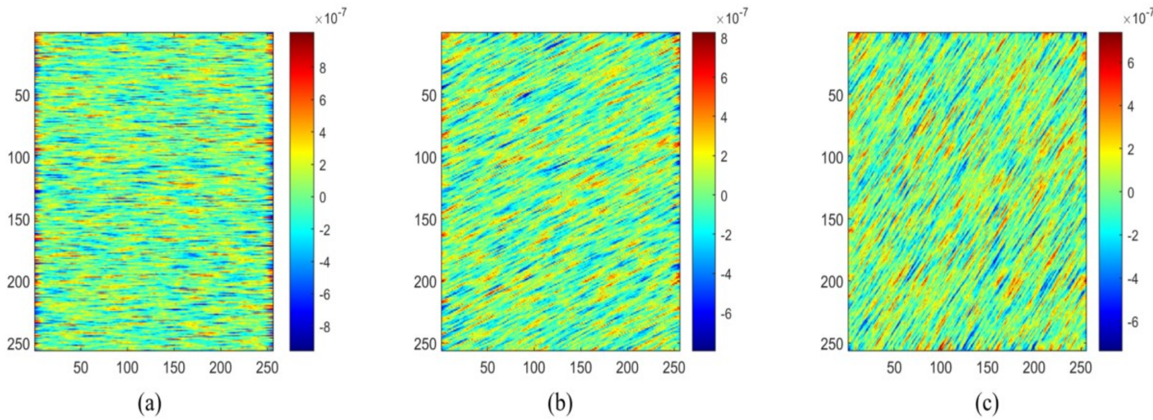


Figure 4.5 Examples of three generated surfaces with (a) normal characteristics, and two anomaly characteristics with (b) mild, and (c) severe directional changes.

Table 4.7 shows the detection performance results of the developed and traditional monitoring approaches using the power of detection under different directional changes (θ = 0.2, 0.3, 0.4, and 0.5). Note that when a surface experiences a directional change, the autocorrelation structure between the topographic values is likely to be impacted (Bui and Apley, 2018b). As the traditional monitoring approaches do not fully characterize the spatial relationships among topographic values, they show lower detection performance of such changes. In contrast, the developed approach outperforms the traditional approaches under all specified directional changes due to the accurate quantification of the spatial autocorrelation structure of the surface. Besides, the developed approach does not assume any distribution in advance, which makes the approach suitable in monitoring different types of topographic surfaces, including both Gaussian and non-Gaussian surfaces.

Table 4.7 Power of detection of the developed and traditional monitoring approach of non-Gaussian anomaly surfaces

Theta (θ)	Monitoring approaches			
	S_{psd}	S_c	S_w	T
0.2	0.114	0.085	0.250	**0.379**
0.3	0.116	0.091	0.299	**0.725**
0.4	0.123	0.100	0.336	**0.933**
0.5	0.125	0.103	0.368	**0.990**

4.6 Conclusions

This chapter introduces a novel approach for monitoring local variations in topographic surfaces. The approach effectively improves the representation of local defects through the developed MSS model, where important characteristics are labeled, and noisy characteristics are smoothed out. The MSS and GSR features are introduced for capturing variations within and between the label assignments, respectively. After integrating the two presented features into a single monitoring statistic, effective monitoring of anomaly surfaces is accomplished. The developed approach is shown to be effective in detecting various surface changes, including variance, autocorrelation, and directional changes, and outperforms the traditional monitoring approaches. In particular, the proposed approach is more sensitive in detecting the local variance change than the local autocorrelation due to the effective segmentation of abnormal values achieved by the developed MSS model. In addition, different types of topographic surfaces, including both Gaussian and non-Gaussian surfaces, are effectively monitored and assessed. Finally, the developed approach is proven to be efficient in monitoring topographic surfaces under the SPC environment.

CHAPTER 5

SPATIALLY WEIGHTED GRAPH THEORY-BASED APPROACH FOR MONITORING LOCAL AUTOCORRELATION CHANGES IN 3D TOPOGRAPHIC SURFACES

5.1 Introduction

Surface topography is a key characteristic in monitoring the quality of finished products and manufacturing processes as discussed in Section 1.1. Due to the complex structure of topographic surfaces, existing minoring approaches lack the detection of local surface faults (Alqahtani *et al.*, 2020c). In Chapter 4, the proposed approach is shown to be superior in detecting local variance change. However, the approach is appeared to be less effective in detecting local autocorrelation change, even though it outperforms the existing monitoring approaches. Examples of local autocorrelation change include pits, ridges, and scratches that commonly appear on the topography of wafer surfaces during semiconductor manufacturing (Alqahtani *et al.*, 2020c). Detecting local autocorrelation change is challenging since this change is likely to cause a local change in the spatial structure of topographic values. Subsequently, this change cannot be simply captured by analyzing the distribution of topographic values where the spatial relationships among surface pixels are ignored. In this chapter, we address this challenge by developing a novel approach for accurate monitoring of local autocorrelation change in topographic surfaces.

Rao *et al.* (2015a) propose a graph-based monitoring approach for characterizing and monitoring changes in surface characteristics. Although the approach yields an improved description of surface characteristics, it shows a limitation in detecting local autocorrelation changes for not considering the spatial relationships among topographic pixels as discussed earlier in Section 2.4.4. Thus, we overcome these limitations by proposing a spatially weighted graph theory-based approach. The proposed approach initially improves the representation of surface characteristics by introducing an in-control multi-region surface segmentation algorithm, which segments the observed surface pixels into clusters according to the information learned from in-control surfaces. The local and spatial topographic characteristics are accurately described through a proposed maximum local spatial randomness (MLSR) feature. We improve the description of the topographic structure by representing the surface as a spatially weighted graph network where the graph nodes represent the obtained clusters, and the graph edges represent the similarity between clusters in terms of their MLSR feature. The local changes in the spatial autocorrelation of topographic values are subsequently detected by monitoring the connectivity of the obtained graph network through a developed spatial graph connectivity statistic. Note that Chapter 5 is mostly based on the following published paper: ALQAHTANI, M. A., JEONG, M. K. and ELSAYED, E. A. 2020d. Spatially weighted graph theory-based approach for monitoring faults in 3D topographic surfaces. *International Journal of Production Research*. Accepted.

The remainder of this chapter is organized as follows. The proposed monitoring approach is presented in Section 5.2. Extensive simulation studies to validate the proposed approach are conducted in Section 5.3. In Section 5.4, a case study of semiconductor copper wafers is presented to evaluate the performance of the proposed monitoring approach. Finally, conclusions are discussed in Section 5.5.

5.2 Proposed Monitoring Approach

The proposed approach is composed of the following stages; Stage 1: multi-region surface segmentation, Stage 2: feature extraction, and Stage 3: anomaly detection as shown in Figure 5.1. In Stage 1, we enhance the representation of local surface characteristics through segmentation. Then, in Stage 2, we extract features that capture the spatial structure among the obtained segmentation results. Finally, in Stage 3, we utilize the graph theory to represent and monitor changes in topographic surfaces. Each stage is discussed in detail in the following sections.

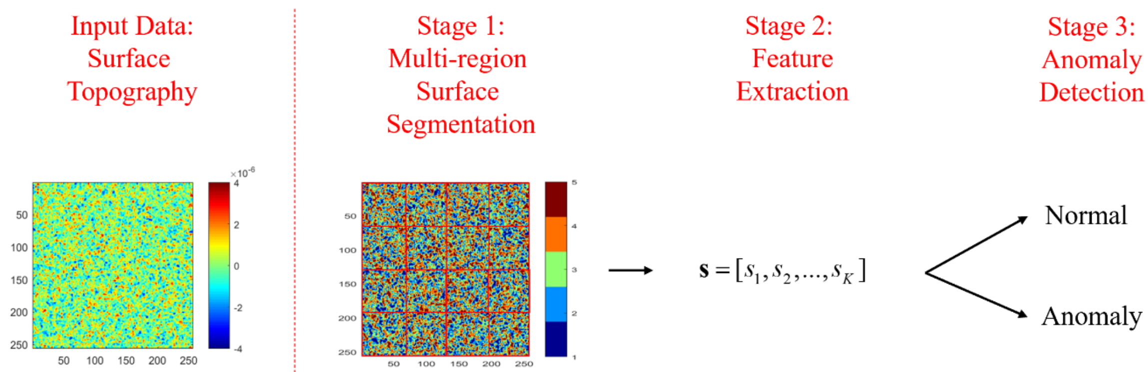


Figure 5.1 Overview of the proposed monitoring approach.

5.2.1 Stage 1: Multi-Region Surface Segmentation

In many real-life applications, image segmentation is an important tool for improving the representation of image (or surface) characteristics by clustering the important characteristics and smoothing out the noisy ones. Image segmentation is typically defined as the process of assigning the image pixels into distinct clusters with similar attributes (Patil and Deore, 2013). Various algorithms, such as *K*-means, watershed, and edge detection, have been developed for obtaining accurate segmentation results (Wang, 2010). However, applying these algorithms for monitoring changes in surface characteristics has some limitations. In particular, the information about in-control surfaces is ignored and not fully utilized during the segmentation process. Furthermore, the cluster labels are not connected to the values of topographic pixels. For example, if we apply these algorithms (e.g., *K*-means) to segment two surface samples generated under the same settings, we may observe that cluster 1 in sample 1 represents the pixels with the lowest topographic values, whereas cluster 1 in sample 2 represents the pixels with the highest topographic values. This inconsistency is likely to cause difficulty in monitoring changes in the characteristics of each cluster during the online monitoring of topographic surfaces.

We overcome the above limitations by proposing an in-control multi-region surface segmentation (IMSS) algorithm, which is based on the *K*-means algorithm. Particularly, the IMSS algorithm considers the information learned from in-control surfaces during the segmentation of new surfaces (Alqahtani *et al.*, 2020d). Accordingly, if a new surface contains characteristics similar to those of in-control surfaces, we expect the segmentation results of the new and in-control surfaces to be similar. However, if the new surface contains abnormal characteristics, we expect a significant change in the obtained segmentation results. In addition, the segmentation results are connected to the values of topographic pixels to track deviations in topographic characteristics. The IMSS algorithm also divides the segmented surface into regions that preserve the local surface characteristics.

The IMSS algorithm is implemented by applying the following procedure for the given in-control surfaces, denoted as $Z^{(j)} = \{z_i^{(j)} : i = 1, 2, ..., M\}$, where $z_i^{(j)}$ is the topographic value of the i^{th} pixel located in the j^{th} surface, $j = 1, 2, ..., N$, and M is the total number of surface pixels. Initially, we randomly assign the topographic pixels into K clusters at the first iteration, $r=1$, and calculate the corresponding means of each cluster $\mu_k^{(r)}, k=1,2,...,K$. Then, we repeat the following two steps for each r^{th} iteration: the segmentation and update steps. In the segmentation step, we segment the topographic pixels into K clusters such that the sum of the squared Euclidean distance between the topographic values and the defined cluster means is minimized as given in Equation (5.1)

$$X^{(j)} = \arg\min_{k} \sum_{k=1}^{K} \sum_{i \in k} \| z_i^{(j)} - \mu_k^{(r)} \|^2 \text{ , } i=1,2,...,M \text{ ,} \tag{5.1}$$

where $\mu_k^{(r)}$ is the mean of the pixels located in the k^{th} cluster obtained from the r^{th} iteration, $r=1,2,...$, and $X^{(j)} = \{x_i^{(j)}, i=1,2,...,M\}$, $x_i \in k = \{1,2,...,K\}$, is the obtained segmentation results. In the update step, we update the mean of each cluster using Equation (5.2)

$$u_k^{(r)} = \frac{\sum_{i=1}^{M} 1\{x_i^{(j)} = k\} \times z_i^{(j)}}{\sum_{i=1}^{M} 1\{x_i^{(j)} = k\}} \text{ , } k=1,2,...,K, \tag{5.2}$$

where $1(.)$ is an indicator function, which is expressed as $1\{x_i^{(j)} = k\} = 1$, otherwise "zero". The algorithm stops at iteration R when the cluster assignments do not change. Subsequently, the means of the cluster obtained in the final iteration $(r=R)$ are specified as the cluster means of the j^{th} surface $(\mu_k^{(R)} = \mu_k^{(j)},\ k=1,2,...,K)$. Since the obtained cluster labels do not reflect the magnitude of topographic values, we rank the mean of each cluster in ascending order $\mu_{(1)}^{(j)} < \mu_{(2)}^{(j)} < ... < \mu_{(K)}^{(j)}$, and then calculate the in-control mean of each cluster as given in Equation (5.3)

$$\mu_k^{(0)} = \frac{1}{N}\sum_{j=1}^{N}\mu_{(k)}^{(j)} \ , \ k = 1,2,...,K \ , \tag{5.3}$$

where $\mu_1^{(0)}$ is the mean of the cluster with the smallest in-control topographic values and $\mu_K^{(0)}$ is the mean of the cluster with the largest in-control topographic values.

When $Z^{(new)} = \{z_i^{(new)} : i = 1,2,...,M\}$ for a new surface is obtained, we assign each topographic pixel into one of the K clusters by minimizing the sum of the squared deviation of the values of topographic pixels from their corresponding in-control cluster means $\mu_k^{(0)}, k = 1,2,...,K,$ as expressed in Equation (5.4)

$$X^{(new)} = \arg\min_k \sum_{k=1}^{K}\sum_{i\in k} \| z_i^{(new)} - \mu_k^{(0)} \|^2 \ , \ i = 1,2,...,M \ , \tag{5.4}$$

where $X^{(new)} = \{x_i^{(new)}; i = 1,2,...,M\}$ is the obtained segmentation results of the new surface. We also enhance the representation of local surface information by dividing the obtained segmented surface $X^{(new)}$ into T regions. More specifically, we divide $X^{(new)}$ into T-nonoverlapped regions with equal size $X_t^{(new)} = \{x_{t,q}^{(new)}; \ q = 1,2,...,Q_t\}, \ t = 1,2,...,T,$ where Q_t is the total number of topographic pixels located in the t^{th} region, which is assumed to be the same for all T regions (i.e., $Q_1 = Q_2 = ... = Q_T = Q$). Guidelines on the choice of the number of clusters K and the number of regions T are discussed later in the chapter.

Figure 5.2 displays a generated surface topography before and after being segmented into seven clusters $(K = 7)$ using the IMSS algorithm. Note that the algorithm divides the segmented surface into 64 spatial regions $(T = 64)$ with an equal size (i.e., $Q = 32\times 32$

pixels). We also show one of the obtained regions located at the top-left corner of the segmented surface. Note that the colors range from the dark blue that represents the lowest valley ($k = 1$) to the dark red that represents the highest peak ($k = 7$). In the following section, we introduce a new feature for quantifying the characteristics of the obtained segmentation results.

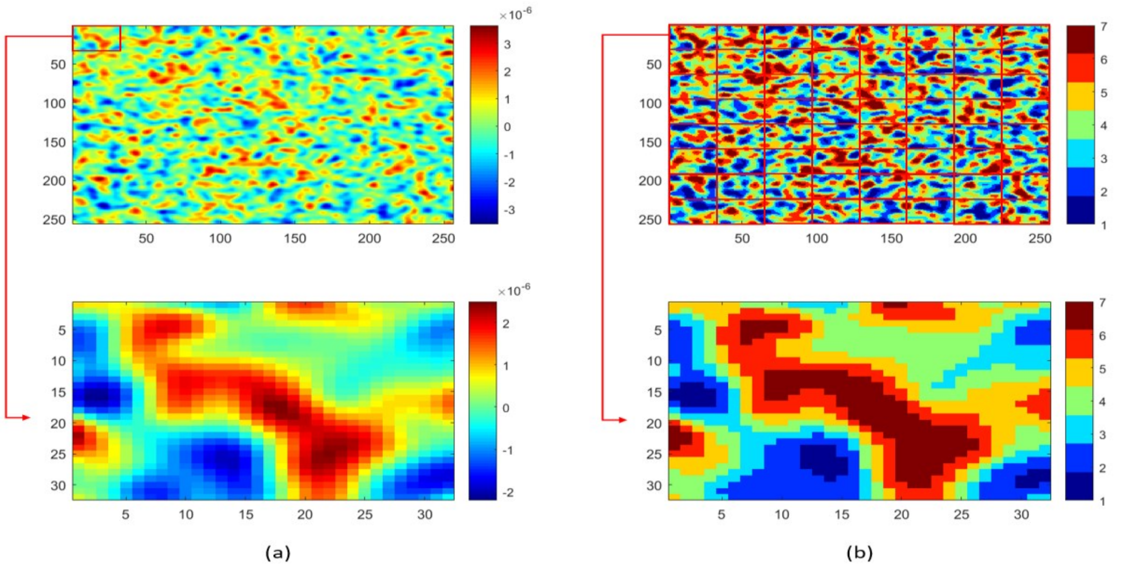


Figure 5.2 Example of a generated surface topography and its corresponding top-left corner region (a) before and (b) after segmentation using the proposed IMSS algorithm.

5.2.2 Stage 2: Feature Extraction

Surface faults commonly occur in spatial and local patterns on the topography of anomaly products during manufacturing processes (Jeong *et al.*, 2008, Bui and Apley, 2018a). Therefore, it is crucial to quantify the local change in the spatial relationships between topographic values for an effective identification of surface faults. The spatial randomness test (SRT) is powerful in detecting the existence of the spatial autocorrelation among pixels in binary images as discussed earlier in Section 4.3.2. However, the current version of the SRT is mainly designed for capturing spatial autocorrelations in binary surfaces. Thus, the

SRT is not applicable for topographic surfaces, which are composed of continuous values. In addition, the SRT is defined as a global measure in which the local faulty values are likely to be dominated by the other normal values, which may lead to low quantification of local changes in the spatial characteristics of local faults.

We address the above drawbacks by generalizing the SRT feature to work with discrete (or segmented) surfaces. Mainly, after obtaining the segmentation result of a topographic surface using the IMSS algorithm $X_t = \{x_{t,q} : q = 1,2,...,Q\}$, $t = 1,2,...,T,$ we initially propose a local joint count (local JC) statistic of the *k*-*k* join, which counts the frequency of occurrence of the cluster k to be neighbored to another cluster k based on a defined neighborhood construction rule. Thus, the local JC statistic of the *k*-*k* join located in the t^{th} region is defined using Equation (5.5)

$$J_{t,k} = \sum_{q=1}^{Q} \sum_{p \in \Omega(q)} 1\{x_{t,q} = k\}\, 1\{x_{t,p} = k\}, \text{for } t = 1,2,...,T,\; k = 1,2,...,K, \tag{5.5}$$

where $\Omega(q)$ is defined as the eight neighbors that are positioned horizontally, vertically, and diagonally to the pixel q and $1(.)$ is an indicator function, which is expressed as $1\{x_{t,q} = k\} = 1,\; 1\{x_{t,p} = k\} = 1,$ otherwise "zero". Then, we propose a local spatial randomness (LSR) feature to quantify the spatial autocorrelation level of the k^{th} cluster located in the t^{th} region as shown in Equation (5.6)

$$s_{t,k} = \lambda_{t,k} J_{t,k}, t = 1,2,...,T,\; k = 1,2,...,K, \tag{5.6}$$

where $\lambda_{t,k} = \sum_{p=1|p\neq k}^{K-1} \frac{1}{Q} \sum_{q=1}^{Q} 1\{x_{t,q} = p\}$ is an assigned weight, which is defined as the probability of not observing the k^{th} cluster in the t^{th} region. Subsequently, we describe the spatial autocorrelation structure of each cluster by selecting the LSR feature from the most spatially autocorrelated region for effective monitoring of local surface faults, which tend to be spatially clustered in local areas (Bui and Apley, 2018a). Particularly, we obtain the maximum local spatial randomness (MLSR) feature for each k^{th} cluster as given in Equation (5.7)

$$s_k = \max_{t=1,2,\dots,T} s_{t,k}\,,\ k=1,2,\dots,K\,. \tag{5.7}$$

As a result, we obtain the MLSR feature vector for each in-control surface $\mathbf{s} = [s_k; k=1,2,\dots,K]$. After obtaining the MLSR feature for N given in-control surfaces $\mathbf{s}^{(j)} = [s_k^{(j)}; k=1,2,\dots,K]$, $j=1,2,\dots,N$, we capture the in-control characteristics of each cluster by calculating the referenced MLSR vector $\mathbf{s}^{(0)} = [s_p^{(0)}; p=1,2,\dots,K]$, where $s_p^{(0)}$ is the referenced MLSR feature of the p^{th} cluster obtained by taking the mean over the N given in-control surfaces as given in Equation (5.8)

$$s_p^{(0)} = \frac{1}{N}\sum_{j=1}^{N} s_p^{(j)}\,,\ p=1,2,\dots,K\,. \tag{5.8}$$

The proposed MLSR vector is an accurate representation of surface characteristics. In particular, the complex characteristics of a surface are simplified by converting the 3D surface topography into a 1D-dimensional vector with size K. The spatial characteristics of a surface are also captured by the MSLR vector such that each element in the MSLR vector represents the spatial characteristics of each cluster calculated from the suspicious region in the surface. More important, the MSLR vector is effective in monitoring changes in surface characteristics. Specifically, we expect the in-control surfaces to share similar

MSLR vectors $\mathbf{s}^{(j)}, j = 1, 2, ..., N$, since such surfaces are composed of similar spatial characteristics. However a new surface that contains faults is likely to cause a change in the spatial characteristics of the surface (Jeong *et al.*, 2008). This change can subsequently be captured in the MLSR vector of the surface $\mathbf{s}^{(new)}$ because of the accurate description of the spatial surface characteristics.

In the next section, we quantify and monitor the similarity between the new and referenced MLSR vectors through graph theory for accurate detection of anomaly surfaces.

5.2.3 Stage 3: Anomaly Detection

Representing the surface topography as a graph network is a powerful technique for describing and monitoring changes in topographic surfaces. Existing graph-based monitoring approaches such as Rao *et al.* (2015b) and Tootooni *et al.* (2016) represent the characteristics of an observed surface as a graph network without considering the information about in-control surfaces. In addition, the spatial characteristics among topographic values are ignored during the construction of the graph network. These drawbacks can cause low characterization and detection of local and spatial topographic faults. Therefore, we present a novel graph network that fully describes the relationships between the spatial characteristics of new and in-control surfaces in terms of their MLSR feature for effective detection of local and spatial surface faults. Therefore, this stage includes two phases; Phase 1: graph representation and Phase 2: graph monitoring. In Phase 1, we represent the relationship between the new and in-control surfaces in terms of their extracted MLSR feature as a spatially weighted graph network. Then, in Phase 2, we monitor the connectivity of the obtained graph to detect anomaly surfaces.

5.2.3.1 Phase 1: Graph Representation

In this phase, we represent the spatial characteristics of new and in-control surfaces as a graph network based on two types of similarities: the between-neighboring-cluster and within-cluster similarities. Due to the fact that neighboring clusters are more spatially autocorrelated than the distant clusters (Jeong *et al.*, 2008), we consider the between-neighboring-cluster similarity to capture the change in the relationship between the obtained clusters and their neighbors. We also consider the within-cluster similarities to capture the deviation of the characteristics of the obtained clusters from their in-control characteristics. Thus, any abnormal change in the values of these two types of similarities can lead to the detection of anomaly surfaces.

We measure the within-cluster similarity by calculating the pairwise similarity between each k^{th} element in the j^{th} in-control MLSR vector, $\mathbf{s}^{(j)} = [s_k^{(j)}; k = 1,2,...,K]$, $j = 1,2,...,N$, to its p^{th} corresponding element in the referenced MLSR vector, $\mathbf{s}^{(0)} = [s_p^{(0)}; p = 1,2,...,K]$, using the Gaussian similarity measure $w_{k,p}^{(j)}$, $k, p = 1,2,...,K$, $k = p$, as described in Equation (5.9). Moreover, we quantify the between-cluster similarity by calculating the pairwise similarity between each k^{th} element in $\mathbf{s}^{(j)}$, $j = 1,2,...,N$, and its p^{th} adjacent element in $\mathbf{s}^{(0)}$ using the Gaussian similarity measure $w_{k,p}^{(j)}$, $k, p = 1,2,...,K$, $k \neq p$, as given in Equation (5.9)

$$w_{k,p}^{(j)} = \beta_{k,p} \exp(-\| s_k^{(j)} - s_p^{(0)} \|^2 / \sigma_{k,p}^{(0)}), \quad j = 1,2,...,N, \tag{5.9}$$

where $\beta_{k,p} = \{1 \text{ if } (\| k - p \|^2 = 1 \text{ or } 0, \text{otherwise } 0\}$ is an assigned spatial weight that indicates whether the k^{th} and p^{th} clusters are adjacent to each other (i.e., $\| k - p \|^2 = 1$) or the same clusters (i.e., $\| k-p \|^2 = 0$), and $\sigma_{k,p}^{(0)}$ is a prespecified in-control standard deviation of the Euclidean similarity between the k^{th} and p^{th} clusters in terms of their MLSR values. We propose Algorithm 5.1 to calculate $\sigma_{k,p}^{(0)}$, $k, p = 1, 2, ..., K$, from in-control surfaces. Note that $w_{k,p}^{(j)}$ indicates the level of the Gaussian similarity value, which takes values from "zero", which implies that $s_k^{(j)}$ and $s_p^{(0)}$ are totally dissimilar, to "one", which indicates that $s_k^{(j)}$ and $s_p^{(0)}$ are totally similar. Subsequently, we obtain the following similarity matrix $\mathbf{W}^{(j)} = [w_{k,p}^{(j)}; k, p = 1, ..., K]$, $j = 1, 2, ..., N$, where the diagonal elements represent the within-cluster-similarities, and the off-diagonal elements represent the between-cluster similarities as illustrated in Figure 5.3.

$$\mathbf{W}^{(j)} = \begin{array}{cc} & \text{Referenced surface } (p) \\ & \begin{array}{cccccc} 1 & 2 & 3 & 4 & \cdots & K \end{array} \\ \begin{array}{c} j\text{th in-control surface } (k) \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ \vdots \\ K \end{array} \end{array} & \begin{bmatrix} w_{1,1}^{(j)} & w_{1,2}^{(j)} & 0 & 0 & \cdots & 0 \\ w_{2,1}^{(j)} & w_{2,2}^{(j)} & w_{2,3}^{(j)} & 0 & \cdots & 0 \\ 0 & w_{3,2}^{(j)} & w_{3,3}^{(j)} & w_{3,4}^{(j)} & \cdots & 0 \\ 0 & 0 & w_{4,3}^{(j)} & w_{4,4}^{(j)} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 0 & w_{K,K-1}^{(j)} & w_{K,K}^{(j)} \end{bmatrix} \end{array}$$

Figure 5.3 Illustration of the structure of the obtained similarity matrix.

Algorithm 5.1: Extraction of in-control similarity characteristics

Given the MLSR vector for N in-control surface $\mathbf{s}_j = [s_k^{(j)}; k = 1,2,...,K]$, $j = 1,2,...,N$, and the referenced MLSR vector $\mathbf{s}^{(0)} = [s_k^{(0)}; k = 1,2,...,K]$, we apply the following steps:

1: **for** $k = 1,2,...,K$ **do**
2: **for** $p = 1,2,...,K$ **do**
3: **for** $j = 1,2,...,N$ **do**
4: calculate the within-cluster similarity value between the k^{th} and p^{th} clusters using the squared Euclidean similarity measure

$$w_{k,p}^{(j)} = \| s_k^{(j)} - s_p^{(0)} \|^2$$

5: **end for**
6: calculate the in-control standard deviation of similarities between the k^{th} and p^{th} clusters over N surfaces

$$\sigma_{k,p}^{(0)} = \sqrt{\sum_{j=1}^{N} (w_{k,p}^{(j)} - w_{k,p}^{(0)})^2 / (N-1)}$$

where $w_{k,p}^{(0)} = \frac{1}{N}\sum_{j=1}^{N} w_{k,p}^{(j)}$ is the mean value of similarities between the k^{th} and p^{th} clusters
7: **end for**
8: **end for**

We present the obtained similarity matrix as a weighted and directed graph $G^{(j)} = (V, E)$, where V is the graph nodes that represent the obtained clusters of observed and referenced surfaces and E is the graph edges that represent the similarity (or weight) between the graph nodes. Note that each node is defined by two indexes $V(j,k)$, where j is the index of each surface and k is the index of each cluster as illustrated in Figure 5.4. Thus, the total number of obtained nodes is $2 \times K$ nodes. In addition, three possible edges are created from each node of the j^{th} surface to the nodes of the referenced surface. Specifically, two possible edges are connected from each node of the j^{th} surface to the neighboring nodes of the referenced surface, where the edge weights are obtained based on the between-neighboring-cluster similarity values $w_{k,p}^{(j)}, k,\ \ p = 1,2,...,K, k \neq p$. In addition, one edge is connected from each node of the j^{th} surface to the same node of the referenced surface, where the edge weights are specified based on the within-cluster similarity values $w_{k,k}^{(j)},\ k = 1,2,...,K$. Thus, we create a total of $(K + 2 + (K-2)\times 2)$ edges.

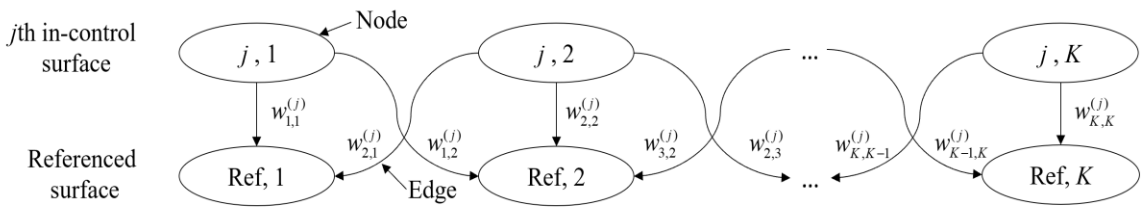


Figure 5.4 Illustration of the proposed spatially weighted graph network.

The graph network is effective in describing the relationship between the spatial characteristics of the in-control surfaces by quantifying the between-neighboring-cluster and within-cluster similarities. Thus, we expect the structure of the graph network of in-control surfaces to be similar to each other since they share similar characteristics. Accordingly, when there is a change in the structure of the graph network of a new surface $G^{(new)} = (V, E)$, this indicates the "abnormality" of that surface. Thus, in the next section, the connectivity of the proposed graph is monitored for detecting anomaly surfaces.

5.2.3.2 Phase 2: Graph Monitoring

In this section, we monitor the connectivity of the obtained graph network for an effective identification of anomaly surfaces. Particularly, we propose the intra- and inter-connectivity measures, which are both integrated to quantify the overall connectivity of the obtained graph network. The intra-connectivity measure quantifies the connectivity within the same node of the observed and referenced surfaces for accurate monitoring of changes in the characteristics of the topographic values (variance change). Moreover, the inter-connectivity measure quantifies the connectivity between the neighboring nodes of the observed and referenced surfaces for effective monitoring of changes in the relationships between the topographic values (autocorrelation change). Thus, after obtaining the similarity matrices of in-control surfaces $\mathbf{W}^{(j)} = [w_{k,p}^{(j)}; k, p = 1, ..., K],\ \ j = 1, 2, ..., N$, the intra- and inter-connectivity measures of these surfaces are respectively calculated using Equations (5.10 and 5.11)

$$C_w^{(j)} = \frac{1}{D_w} \sum_{k=1}^{K} w_{k,k}^{(j)},\ \ j = 1, 2, ..., N, \tag{5.10}$$

$$C_b^{(j)} = \frac{1}{D_b} \sum_{k=1}^{K} \sum_{p=1|p \neq k}^{K} w_{k,p}^{(j)},\ j = 1, 2, ..., N, \tag{5.11}$$

where D_w is the total number of the within-cluster edges (i.e., $D_w = K$) and D_b is the total number of the between-neighboring-cluster edges (i.e., $D_b = 2 + (K-2) \times 2$). Subsequently, we integrate these two measures into a single monitoring statistic called spatial graph connectivity for effective monitoring of surface changes, which is defined in Equation (5.12)

$$C^{(j)} = C_w^{(j)} + C_b^{(j)},\ j = 1, 2, ..., N. \tag{5.12}$$

In Equation (5.12), we assume an equal contribution of the intra- and inter-connectivity measures because the fault may yield a change in the characteristics of the topographic pixels, which can be captured by the intra-connectivity measure, or the relationships between the topographic pixels, which can be detected by the inter-connectivity measure. Since the type of change is mostly unknown in advance in real-life applications, we assume an equal contribution of both intra- and inter-connectivity measures to achieve an overall high detection performance of both changes. However, if there is prior knowledge about the occurrence distribution of certain types of changes, then we can easily incorporate this ratio in Equation (5.12) to improve the detection performance. In addition, the value of $C^{(j)}$ ranges from "zero", which indicates that the nodes of the observed surface are weakly connected to the nodes of the referenced surface, to "two", which implies the strong connectivity between the nodes of the observed and referenced surfaces. Accordingly, a larger connectivity value means that the characteristics of the observed and referenced surfaces are similar (i.e., "smooth" surface) and vice versa. Therefore, after we obtain the spatial graph connectivity statistic for in-control surfaces $C^{(j)}$, $j = 1, 2, ..., N,$ we determine the lower critical limit H for identifying anomaly surfaces as explained in Algorithm 5.2. Specifically, we utilize a bootstrapping method with a larger number of replicates for obtaining an estimate of the empirical distribution of the in-control statistics $C^{(j)}$, $j = 1, 2, ..., N$. Subsequently, we determine the lower critical limit H from the estimated distribution for a prespecified type I error, which is defined as the probability of incorrectly identifying the surfaces with in-control characteristics as anomaly surfaces (Febrero *et al.*, 2008). Subsequently, after the spatial graph connectivity statistic of a new surface $C^{(new)}$ is obtained, and if $C^{(new)} < H$, then the new surface is classified as an anomaly.

Algorithm 5.2: Calculation of the lower critical limit H
Given the in-control spatial graph connectivity statistics $C^{(j)}$, $j=1,2,...,N$, the number of bootstrapping sampling observations U (i.e., $U=1000$), the number of replications R (i.e., $R=1000$), and a defined type I error α (i.e., $\alpha=0.05$), we apply the following steps: 1: **for** $r=1,2,...,R$ **do** 2: calculate the empirical distribution from the obtained N in-control spatial graph connectivity statistics by randomly sampling U observations with replacement $$A=\{C_u, u=1,2,...,U\}$$ 3: rank the collected samples in an ascending order $$E=\{C_{(u)}, u=1,2,...,U\}$$ where $C_{(1)},C_{(2)},....,C_{(U)}$ are the order statistics of the selected samples (i.e., $C_{(1)} \le C_{(2)} \le .... \le C_{(U)}$) 4: Obtain the α -quantile of the ranked samples for the r^{th} replicate $$\Pr[E \le t_r]=\alpha$$ 5: **end for** 6: calculate the lower critical limit as $$H=\underset{r=1,2,...,R}{\text{median}}\ (t_r)$$

Finally, the proposed approach is effective in extracting the fault locations for fault diagnosis. Particularly, the locations of surface faults are identified by first finding the node (or cluster) with the most abnormal characteristics in terms of its MLSR value, and then locating the region where the abnormal node belongs to. More specifically, after identifying a new surface as an anomaly, we find its fault locations by applying the following diagnosis algorithm:

- **Step 1:** Given the calculated intra-connectivity measures of the anomaly surface $w_{k,p}^{(new)}$, $k,p=1,2,..,K$, $k=p$, which is defined earlier in Equation (5.9), we find the index of the abnormal node that has the minimum similarity to its corresponding node of the referenced surface as provided in Equation (5.13)

$$\omega = \underset{k=1,2,\dots,K}{\arg\min}\left(w_{k,k}^{(new)}\right), \tag{5.13}$$

where $\omega \in \{1,2,\dots,K\}$ is the index of the identified abnormal node.

- **Step 2:** Given ω, we locate the index of the abnormal region in the surface by finding the maximum MLSR value over all T regions as given in Equation (5.14)

$$\eta^{(new)} = \underset{t=1,2,\dots,T}{\arg\max}\left(s_{t,\omega}^{(new)}\right), \tag{5.14}$$

where $\eta^{(new)} \in \{1,2,\dots,T\}$ is the identified index of the abnormal region and $s_{t,\omega}^{(new)}$ is the MLSR value of the abnormal node ω calculated from the region t as given earlier in Equation (5.7).

In the proposed diagnosis algorithm, we consider the intra-connectivity measure to locate abnormal nodes since this measure is effective in analyzing each cluster independently, which allows us to accurately determine the abnormal node with the least connectivity. This makes the proposed approach not only effective in identifying anomaly surfaces but also in locating surface faults.

5.3 Performance Study

In this section, we utilize the Gaussian random model discussed in Section 2.3 for simulating topographic surfaces. Besides, we generate anomaly surfaces by following the same fault scenarios explained earlier in Section 3.3.1

5.3.1 Analysis of the Proposed Approach

5.3.1.1 Effect of the Choice of *K* and *T* in the IMSS Algorithm

In this section, we study the detection performance under different values of the number of clusters K and the number of regions T, which are parameters of the proposed IMSS algorithm. In Table 5.1, we present the detection power of the proposed approach with different choices of K (i.e., $K = 3$, 7, and 10) and T (i.e., $T = 4$, 64, and 256) under various fault scenarios. Figure 5.5 also illustrates the segmentation results obtained by the IMSS algorithm under the defined values of K and T. We observe that when the number of clusters K is large (e.g., $K = 10$), we approach the complex structure of the original surface. This loses the advantage of the IMSS algorithm in simplifying the surface structure and subsequently yields lower detection power of surface faults as shown in Table 5.1. Similarly, when K is small (e.g., $K = 3$), this loses the detailed structure of the surface characteristics and results in lower detection power. We also notice that a larger number of regions T (e.g., $T = 256$) may cause a low representation of the surface structure, which yields lower detection power of surface faults. Likewise, a small number of regions T (e.g., $T = 4$) may cause the local fault areas to be dominated by the normal areas, which results in lower detection power. Thus, a "moderate" selection of K and T is recommended (e.g., $K = 7$ and $T = 64$) to achieve the best detection performance as presented in Table 5.1.

Table 5.1 Power of detection of the proposed approach with different K and T

Fault type	Fault area	Fault size σ	Fault size θ	T \ K	3	7	10
Variance	2.5%	1.5	1	4	0.094	0.727	0.844
				64	0.115	**0.963**	0.946
				256	0.074	0.573	0.505
		1.75	1	4	0.143	0.970	0.998
				64	0.285	**0.999**	0.998
				256	0.073	0.973	0.978
	3%	1.5	1	4	0.133	0.867	0.944
				64	0.141	**0.991**	0.990
				256	0.084	0.684	0.626
		1.75	1	4	0.214	0.994	0.997
				64	0.374	**0.999**	0.998
				256	0.088	0.984	0.988
Autocorrelation	2.5%	1	2	4	0.207	0.680	0.660
				64	0.799	**0.949**	0.873
				256	0.559	0.839	0.723
		1	2.25	4	0.348	0.786	0.779
				64	0.932	**0.991**	0.954
				256	0.828	0.957	0.887
	3%	1	2	4	0.316	0.815	0.795
				64	0.890	**0.983**	0.951
				256	0.638	0.900	0.826
		1	2.25	4	0.476	0.894	0.893
				64	0.978	**0.996**	0.984
				256	0.906	0.981	0.938

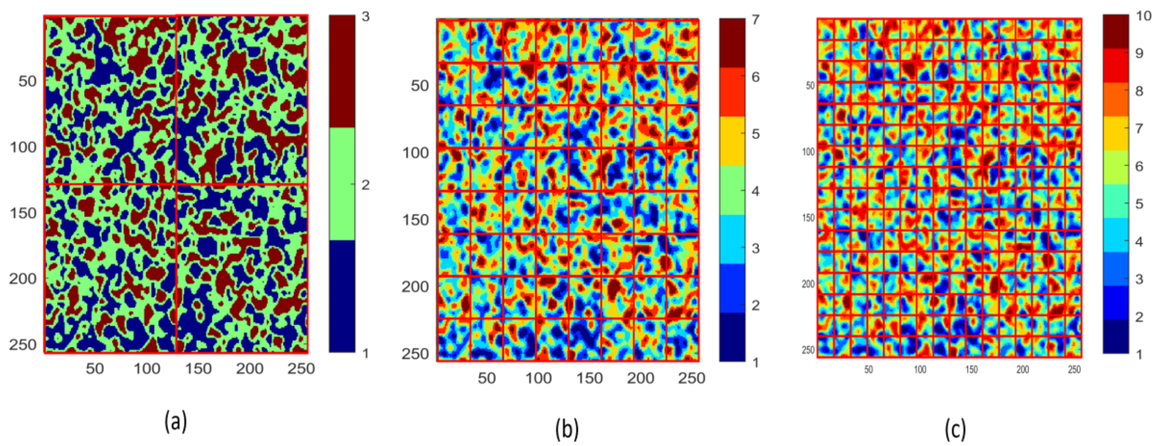


Figure 5.5 Segmentation results obtained by the IMSS algorithm with: (a) $K=3$, $T=4$, (b) $K=7$, $T=64$, and (c) $K=10$, $T=256$.

5.3.1.2 Effectiveness of the IMSS Algorithm

In this section, we study the effectiveness of the IMSS algorithm in representing different changes in the surface characteristics. In general, when the characteristics of a new surface are different from the in-control characteristics, we expect the obtained segmentation results to capture this change. For example, when the standard deviation parameter σ is shifted, we expect to observe more abnormal positive and negative topographic values. Subsequently, the IMSS algorithm assigns these abnormal values to their closest defined in-control cluster means for minimizing their sum of squared Euclidean distances. This causes the abnormal negative values to be assigned to the cluster with the lowest mean (cluster 1) and the abnormal positive values to be assigned to the cluster with the highest mean (cluster K), which results in a significant change in the spatial structure of these two clusters. In addition, when the scale parameter θ is shifted, we expect the observed values of topographic pixels to be more similar to their neighbors. Since the IMSS algorithm assigns the pixels with similar topographic values to the same cluster, this results in a considerable change in the structure of all assigned clusters.

Figure 5.6 illustrates the segmentation results of three generated surfaces: one in-control and two anomaly surfaces with variance and autocorrelation changes, using the proposed IMSS algorithm. Note that we select $K = 7$ and $T = 64$ such that each region has the same size (i.e., $Q = 32 \times 32$ pixels). We also show the top-left region of the three defined surfaces. We observe that the structure of assigned clusters obtained from the two anomaly surfaces experiences a significant change from the structure of the in-control surface. In particular, the structure of the top-left region of the in-control surface shows a less autocorrelated pattern for all assigned clusters. However, the structure of the top-left region of the anomaly surface with the variance change shows that clusters 1 and 7 are assigned more than the other clusters due to the minimization of the sum of the squared Euclidean distance between the topographic values and their closest in-control cluster means. In addition, the structure of the top-left region of the anomaly surface with the autocorrelation change exhibits a significant change in the pattern of all clusters due to the deviation of the autocorrelation structure from the in-control structure. This shows the effectiveness of the IMSS algorithm in representing abnormal surface characteristics that are resulted from either variance or autocorrelation change.

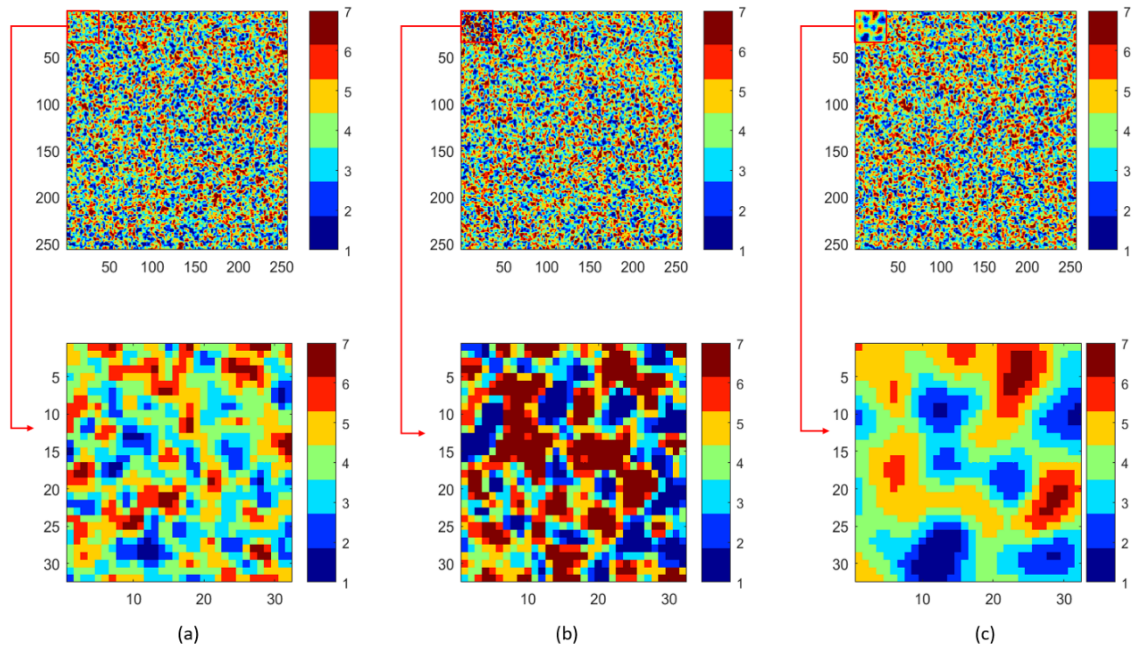


Figure 5.6 Segmentation results of three generated surfaces using the IMSS algorithm with (a) in-control characteristics, and anomaly characteristics with (b) variance change, and (c) autocorrelation change.

5.3.1.3 Effectiveness of the Proposed Graph Network

This section analyzes the effectiveness of the proposed graph network in monitoring different topographic changes. Figure 5.7 shows the structure of the proposed graph network for the three surfaces illustrated earlier in Figure 5.6. Note that the red circles represent the nodes of the observed surface, and the black circles represent the nodes of the referenced surface. In addition, the thickness of the edges represents the magnitude of the similarity (or weight) between the graph nodes, which ranges from zero (totally dissimilar) to one (totally similar). In the graph network of the in-control surface, the edge weights between the nodes of the observed surface and their corresponding and neighboring nodes of the referenced surface are high as shown in Figure 5.7 (a) (e.g., $w_{1,1} = 0.99$, $w_{2,1} = 1.00$). This is because these nodes share similar characteristics in terms of their MLSR values. However, when we obtain an anomaly surface with the variance change, we expect clusters

1 and 7 to appear more than other clusters as shown earlier in Figure 5.6 (b). Accordingly, Figure 5.7 (b) depicts a less connected network in which the edge weights between the nodes 1 and 7 of the observed surface and their respective nodes of the referenced surface are significantly low ($w_{1,1} = 0.29, w_{7,7} = 0.25$). In addition, in the graph network of the anomaly surface with the autocorrelation change, the edge weights between the nodes of the observed surface and their respective and neighboring nodes of the referenced surface are decreased as shown in Figure 5.7 (c) (e.g., $w_{7,7} = 0.11, w_{5,4} = 0.80$). This is due to the change in the pattern of the clusters as presented earlier in Figure 5.6 (c). This makes the proposed graph network effective in representing different changes in the surface, including both variance and autocorrelation changes.

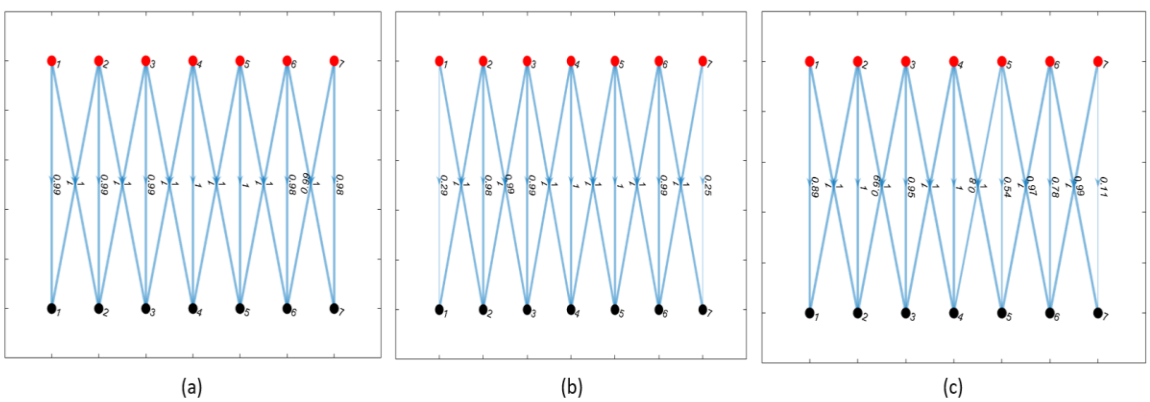


Figure 5.7 Examples of the proposed graph network of three generated surfaces with (a) in-control characteristics, and anomaly characteristics with (b) variance change, and (c) autocorrelation change.

5.3.2 Performance Comparison

This section compares the detection performance of the developed spatial graph connectivity statistic C with the existing monitoring approaches presented earlier in Section 2.4. We also use the power of detection as a performance measure, which is given earlier in Equation (3.15) under the same type I error (i.e., $\alpha = 0.05$) for all approaches to conduct a fair comparison. In particular, Table 5.2 shows the power of detection P of the proposed and the existing monitoring approaches under various fault sizes, areas, and types. We observe that the average roughness S_a fails to identify local variance and autocorrelation changes since S_a ignores the deviation in the spatial relationships among topographic values. Similarly, the power spectral density S_{psd} shows lower detection of local faults because it mainly relies on the FFT algorithm, which is expressed as a global decomposition filter where local faults are likely to be smoothed out during the transformation of topographic values to the frequency domain. Next, the watershed measure S_w shows lower performance in identifying both local variance and autocorrelation changes because S_w is mainly based on the watershed algorithm, which suffers from the over-segmentation problem that may cause the local faults to be misidentified. The Fiedler connectivity feature S_c shows better performance in detecting the variance change compared to the other existing approaches. This is because S_c captures the variance change by calculating the dissimilarity between each row of the image matrix. However, a lower detection performance of the autocorrelation change is observed for not considering the spatial relationships among topographic values.

In Table 5.2, we also obtain the power of detection for the approach based on the existing SRT feature Q. In particular, after binarizing the topographic values using the edge detection algorithms (e.g., Canny filter), we calculate Q as explained earlier in Equation (4.7). We observe that Q fails to detect the anomaly surfaces with local variance and autocorrelation changes. This is because Q is obtained based on the binary result of the original surface, which may yield the loss of the detailed structure of topographic values. In addition, the local fault area can be dominated by other normal areas during the calculation of Q.

On the other hand, we show that anomaly surfaces associated with either variance or autocorrelation change result in a significant change in the structure of clusters obtained by the IMSS algorithm. This change is adequately quantified by the proposed MLSR feature, which measures the local spatial autocorrelation level of each obtained cluster. The proposed spatially weighted graph is also shown to be effective in describing both variance and autocorrelation changes by quantifying the similarity within and between each cluster of the new and referenced surfaces in terms of their MLSR feature. Subsequently, the proposed spatial graph connectivity statistic C, which monitors the connectivity of the obtained graph, reveals a superior performance in detecting both variance and autocorrelation changes and outperforms the existing monitoring approaches under all specified fault scenarios as shown in Table 5.2. In addition, we outperform the existing graph approach (Rao *et al.*, 2015b), which is based on the Fiedler connectivity feature S_c, by improving the detection performance of the variance and autocorrelation changes by 56.04% and 308.95%, respectively.

Table 5.2 Power of detection of the proposed and the existing monitoring approaches using anomaly surfaces obtained by the Gaussian random model

Surface type	Fault area	Fault size		Monitoring approaches					
		σ	θ	S_a	S_{psd}	S_w	S_c	Q	C
In-control surfaces		1	1	0.051	0.052	0.051	0.050	0.051	0.050
Anomaly surfaces with variance change	2.5%	1.5	1	0.247	0.254	0.091	0.473	0.253	**0.963**
		1.75	1	0.275	0.287	0.094	0.704	0.288	**0.999**
	3%	1.5	1	0.250	0.259	0.090	0.474	0.258	**0.991**
		1.75	1	0.290	0.294	0.096	0.550	0.292	**0.999**
Anomaly surfaces with autocorrelation change	2.5%	1	2	0.097	0.097	0.270	0.210	0.095	**0.949**
		1	2.25	0.107	0.110	0.405	0.245	0.108	**0.991**
	3%	1	2	0.110	0.113	0.380	0.215	0.110	**0.983**
		1	2.25	0.113	0.116	0.544	0.220	0.117	**0.996**

We also show the effectiveness of the proposed diagnosis algorithm in identifying the locations of anomaly regions. Figure 5.8 shows two examples of anomaly surfaces with variance and autocorrelation changes and their corresponding anomaly regions identified by the proposed algorithm. Since the surface is divided into local regions, we can immediately locate the region with the most abnormal characteristics using Equations (5.13 and 5.14). In particular, the algorithm accurately identifies regions 22 and region 44 as anomalies, where these two regions are truly associated with variance and autocorrelation changes as shown in Figure 5.8 (a and b), respectively.

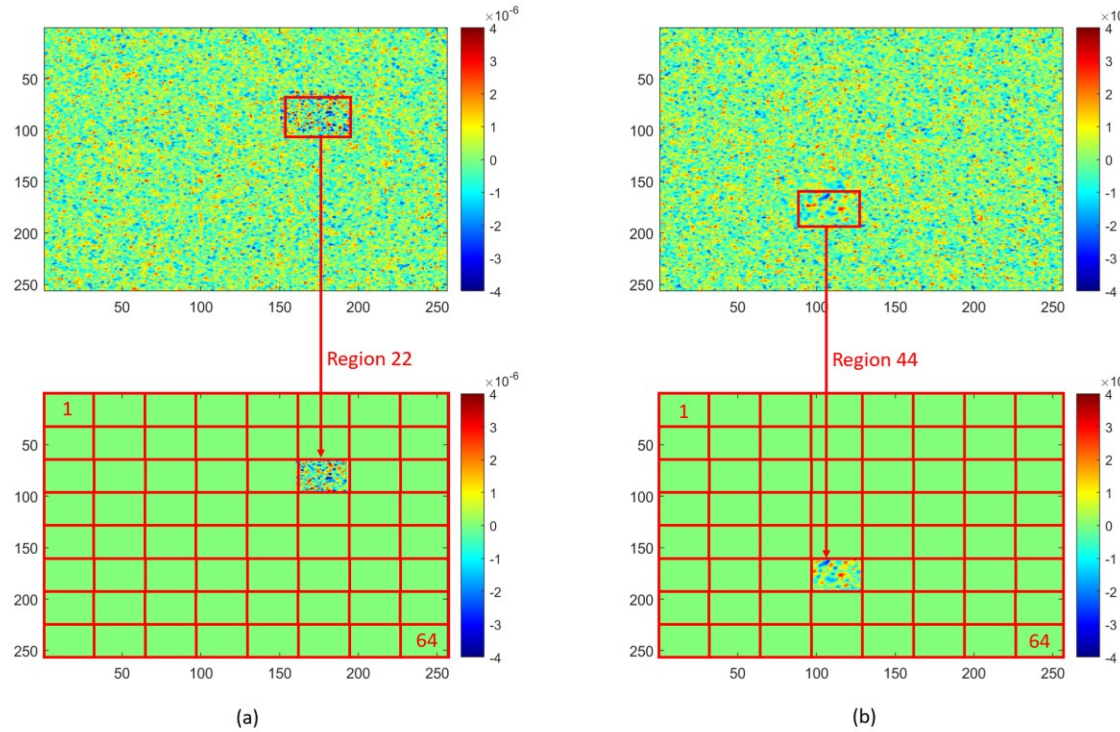


Figure 5.8 Examples of two anomaly surfaces with (a) variance change, and (b) autocorrelation change, and their corresponding anomaly regions identified by the proposed diagnosis algorithm.

5.4 A Case Study of Wafer Surfaces

We demonstrate the effectiveness of the proposed approach by using the case study of wafer surfaces described earlier in Section 3.4. Particularly, in Table 5.3, we show the detection power P of the proposed and the existing monitoring approaches for anomaly semiconductor wafers. Note that P is obtained for all approaches under a prespecified type I error (i.e., $\alpha = 0.05$). We observe that the existing monitoring approaches result in lower detection power of wafer faults because of the lack of characterization of the spatial autocorrelation among topographic pixels. In contrast, the proposed spatial graph connectivity statistic C is shown to be more sensitive in detecting deviations in the spatial autocorrelation structure of the surface that may result from the appearance of faults, such as pits, ridges, and scratches, during semiconductor manufacturing. Therefore, C yields higher performance in identifying anomaly wafers with different faults sizes and areas and outperforms the existing monitoring approaches.

Table 5.3 Power of detection of the proposed and the existing monitoring approaches using anomaly semiconductor wafers

Surface type	Fault area	Monitoring approaches					
		S_a	S_{psd}	S_w	S_c	Q	C
In-control surfaces		0.050	0.050	0.053	0.051	0.050	0.051
Anomaly surfaces with pits	6%	0.084	0.083	0.802	0.212	0.081	**0.818**
	7%	0.095	0.095	0.910	0.262	0.101	**0.913**
	8%	0.104	0.096	0.952	0.289	0.116	**0.967**
Anomaly surfaces with ridges	1.5%	0.081	0.081	0.110	0.110	0.161	**0.849**
	2%	0.086	0.084	0.114	0.111	0.181	**0.973**
	2.5%	0.089	0.087	0.120	0.112	0.202	**0.996**
Anomaly surfaces with scratches	6%	0.079	0.079	0.821	0.200	0.074	**0.907**
	7%	0.091	0.088	0.911	0.267	0.113	**0.972**
	8%	0.097	0.094	0.965	0.275	0.116	**0.993**

5.5 Conclusions

In this chapter, we propose a novel spatially weighted graph theory-based approach for monitoring local autocorrelation changes in 3D topographic surfaces. We propose the IMSS algorithm, which segments the topographic values of a new surface into predefined clusters based on the information learned from in-control surfaces. We also propose the MLSR feature, which measures the local spatial autocorrelation level of each obtained cluster. After calculating the similarity values within and between clusters of the new and in-control surfaces using their MLSR feature, a novel spatially weighted graph network is proposed with nodes that represent the segmented clusters and edges that represent the similarity between the graph nodes. Surfaces with abnormal characteristics are effectively detected by monitoring the connectivity of the obtained graph network through the developed spatial graph connectivity statistic.

The proposed IMSS algorithm shows its effectiveness in enhancing the representation of local changes in surface characteristics. In addition, the spatial statistical dependencies among obtained clusters are adequately quantified and captured through the proposed MLSR feature. We accurately describe the spatial relationship between the characteristics of the new and in-control surfaces through the proposed graph network. Finally, the developed spatial connectivity statistic is proven to be robust in detecting different forms of local autocorrelation changes in topographic surfaces.

CHAPTER 6

GENERALIZED SPATIALLY WEIGHTED AUTOCORRELATION APPROACH FOR FAULT DETECTION AND DIAGNOSIS IN 3D TOPOGRAPHIC SURFACES

6.1 Introduction

Detecting and diagnosing surface faults are crucial tasks in many applications. The existing monitoring approaches do not fully characterize the complex behavior among topographic pixels, which renders them ineffective in detecting and diagnosing faults in topographic images as discussed in Section 2.4. Although the proposed approach in Chapter 5 is effective in locating surface faults, it has a limitation in extracting fault diagnostic information, such as the fault size, and the fault magnitude and the number of faults. The approach also has a limitation in finding faults in multiple locations on the surface. Therefore, in this chapter, we overcome these limitations by introducing an effective and efficient online monitoring approach for detecting and diagnosing faults in topographic images.

Moran (1950) develops Moran's index, which quantifies the spatial autocorrelation between the image pixels based on their similarities. However, applying Moran's index to online surface monitoring is ineffective because it lacks the quantification of local surface changes, is limited to the first-order neighbor, and is computationally expensive. Therefore, in this chapter, we overcome these limitations by introducing a generalized spatially weighted autocorrelation approach based on a developed generalized spatially weighted

Moran (GSWM) index for online monitoring and diagnosis of local faults in topographic images (Alqahtani *et al.*, 2020a).

The proposed approach utilizes the information learned from normal surfaces to assign spatial weights to the regions with suspicious characteristics of new surfaces. We propose two algorithms to obtain the weight assignments: the normal surface hard thresholding algorithm, which improves the representation of surface characteristics through binarization, and the normal surface connected-component labeling algorithm, which utilizes the obtained binary results to identify and assign spatial weights to the topographic regions with suspicious characteristics. We also develop the GSWM index, which exploits the assigned weights for characterizing the spatial autocorrelation structure of each identified region. The GSWM index calculates the cross-products of the spatially weighted deviations of the topographic values and their different order neighbors from the mean of normal topographic values. When an anomaly surface is detected based on its GSWM index, we accurately extract different fault information. Note that Chapter 6 is mostly based on the following working paper: ALQAHTANI, M. A., JEONG, M. K. and ELSAYED, E. A. 2020a. Generalized spatially weighted autocorrelation approach for monitoring faults in 3D topographic surfaces with application to wafer surface monitoring. Working paper.

This chapter is organized as follows. In Section 6.2, we review the existing Moran's index and its limitations to online surface monitoring. We then introduce the proposed generalized spatially weighted autocorrelation approach in Section 6.3. In Section 6.4, we compare the performance of the proposed approach with other existing approaches. A real-life case study of the topography of semiconductor wafers is assessed in Section 6.5. Conclusions are stated in Section 6.6.

6.2 Moran's Index for Surface Monitoring

Moran's index is commonly applied to identify the existence of spatial autocorrelation among spatial data. Moran's index is expressed as the sum of cross-products between the deviation of observations and their neighbors from the observation mean (Moran, 1950). Particularly, for an observed topographic image (or surface) $Z = \{z_i : i = 1, 2, ..., M\}$, where z_i is the intensity value of the pixel i and M is the size of image pixels, Moran's index is given in Equation (6.1)

$$I = \frac{\sum_{i=1}^{M}\sum_{m=1}^{M} w_{i,m}(z_i - \bar{z})(z_m - \bar{z})}{\sigma^2 \sum_{i=1}^{M}\sum_{m=1}^{M} w_{i,m}}, \tag{6.1}$$

where $(\bar{z}, \sigma)$ represent the mean and standard deviation of the values of topographic pixels, respectively, and $w_{i,m}$ is an assigned binary weight that specifies whether the pixels z_i and z_m are neighbors, i.e., $w_{i,m} = 1$ if (z_i, z_m) are neighbors, $w_{i,m} = 0$ otherwise. The weight assignments in Moran's index are commonly calculated based on the first-order neighbors, i.e., the first closest neighbors to a pixel of interest. There are two popular neighborhood construction rules to identify the neighbors of a pixel of interest: the king-move neighborhood (KMN) and the rook-move neighborhood (RMN). The KMN includes the neighbors located in the vertical, horizontal, and diagonal axes of a pixel of interest, while the RMN includes the neighbors located in the vertical and horizontal axes of a pixel of interest (Jeong *et al.*, 2008).

Applying Moran's index for monitoring changes in topographic surfaces has some limitations. The characteristics of normal surfaces are ignored when evaluating new surfaces, which may yield low quantification of changes in surface characteristics. Moreover, Moran's index is defined as a global index where local surface characteristics are likely to be dominated by global characteristics, which may yield the mischaracterization of local faults. The weight assignments in Moran's index are typically limited to capture the spatial autocorrelation at the first-order of neighboring pixels. Generalizing the weight assignments to multiple neighboring orders is crucial for an accurate evaluation of spatial autocorrelation structure. However, this task is challenging since the topographic characteristics are not identical under normal process conditions. In addition, the calculation of Moran's index is computationally expensive since its computation time is polynomial, which is given as $O(R \times M)$, where R is the size of neighboring pixels and M is the size of image pixels. This makes the current Moran's index not suitable for online surface monitoring.

In the following section, we overcome the above limitations by presenting the generalized spatially weighted autocorrelation approach, which is based on the developed generalized spatially weighted Moran index, for effective monitoring of changes in surface characteristics (Alqahtani *et al.*, 2020a).

6.3 Proposed Generalized Spatially Weighted Autocorrelation Approach

The proposed approach is composed of four consecutive stages; Stage 1: suspicious topographic region identification, Stage 2: feature extraction, Stage 3: anomaly detection, and Stage 4: fault identification and diagnoses as illustrated in Figure 6.1. In Stage 1, we identify and assign spatial weights to the topographic regions with suspicious characteristics. We then exploit the assigned weights to extract features that accurately describe the spatial autocorrelation of each identified region in Stage 2. Stage 3 detects anomaly surfaces based on their extracted features. Finally, in Stage 4, we extract fault information from the detected anomaly surfaces for fault diagnosis. In the following sections, we discuss each stage of the proposed approach in detail.

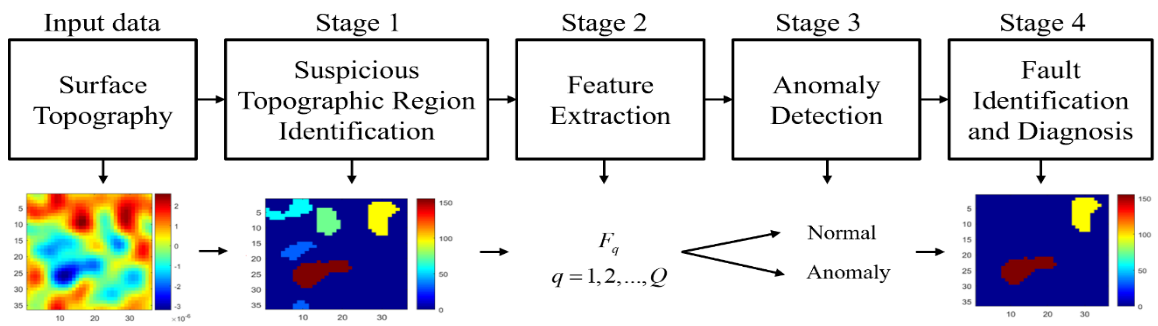


Figure 6.1 Overview of the proposed approach.

6.3.1 Stage 1: Suspicious Topographic Region Identification

Connected component labeling (CCL) is an important process for finding the connected regions, also called components or objects, in binary images (Št and Beneš, 2011, He *et al.*, 2017). A connected region is a region that contains pixels with the same intensity (i.e., pixels with the intensity of "1"), and all pixels in that region are connected to each other by a path obtained based on a specified neighborhood construction rule (He *et al.*, 2017). The flood-fill algorithm is one of the most popular CCL algorithms, which aims to find the connected regions in a binary image and label them uniquely as separate regions. The flood-fill algorithm obtains the connected regions by representing the binary image as a graph network with nodes and edges (Silvela and Portillo, 2001). Each pixel is described as a node and an edge between two nodes is constructed if they are neighbors. The algorithm requires the following information: a starting node, a target value (i.e., a foreground value "1"), and a replacement value (i.e., an assigned label for each identified connected region k, $k = 1, 2, \ldots$) (Silvela and Portillo, 2001). The algorithm consists of two main steps: the identification and assignment steps. The identification step identifies all neighboring nodes that are connected to the starting node by a path of the target value "1" and then the assignment step assigns the identified nodes to a replacement value k. The algorithm repeats these two steps until all nodes associated with the target value are assigned to a replacement value.

Applying the flood-fill algorithm to surface monitoring has some limitations. Particularly, the characteristics of normal topographic images are not utilized during the labeling process, which may lead to low performance of labeling fault information. In addition, the algorithm is limited to binary images, which makes it inapplicable for topographic images that consist of continuous values. The algorithm is also typically applied to a single binary image, which may lead to a low identification of surface characteristics such as peaks and valleys. Therefore, we overcome these limitations by introducing two algorithms for identifying the regions with suspicious characteristics of new surface based on the information learned from normal surfaces: the normal surface hard thresholding algorithm, which improves the surface representation through binarization and the normal surface connected-component labeling algorithm, which exploits the obtained binary results to identify and assign spatial weights to the regions with suspicious characteristics. The two algorithms are explained below.

6.3.1.1 Normal Surface Hard Thresholding Algorithm

Hard thresholding (HT) algorithms are effective in representing the important characteristics of an image as discussed earlier in Section 3.2.1. In particular, in Section 3.2.1, we propose the multi-level surface thresholding (MST) algorithm to binarize a surface into different levels, and each surface level is calculated based on the information obtained from the previous levels. However, in this section, we propose the normal surface hard thresholding (NSHT) algorithm, which only considers two surface levels, and each surface level is obtained independently from the other levels for accurate analysis of the valley and peak characteristics. More specifically, the NSHT algorithm assigns each pixel, $z_i \in Z,\ i = 1, 2, ..., M$, into a binary value ("0" or "1") based on a prespecified normal surface hard threshold h_1 as given in Equation (6.2)

$$\begin{aligned} z_i < 0:\ B_1 &= \{b_i^{(1)} : i = 1, 2, ..., M\}, \text{ where } b_i^{(1)} = \{1 \text{ if } |z_i| > h_1, 0 \text{ otherwise}\}, \\ z_i \geq 0:\ B_2 &= \{b_i^{(2)} : i = 1, 2, ..., M\}, \text{ where } b_i^{(2)} = \{1 \text{ if } |z_i| > h_1, 0 \text{ otherwise}\}, \end{aligned} \tag{6.2}$$

where (B_1, B_2) are the output of binarizing the negative and positive topographic values, respectively. We propose Algorithm 6.1 to determine the normal surface hard threshold h_1 based on normal surfaces. Particularly, h_1 is calculated based on the median value of α_1-quantile of absolute normal topographic values, where α_1 is a specified type I error rate, i.e., the rate of incorrectly identifying the normal topographic values as abnormal values. Figure 6.2 shows a topographic image and its corresponding negative and positive binary results obtained by the NSHT algorithm.

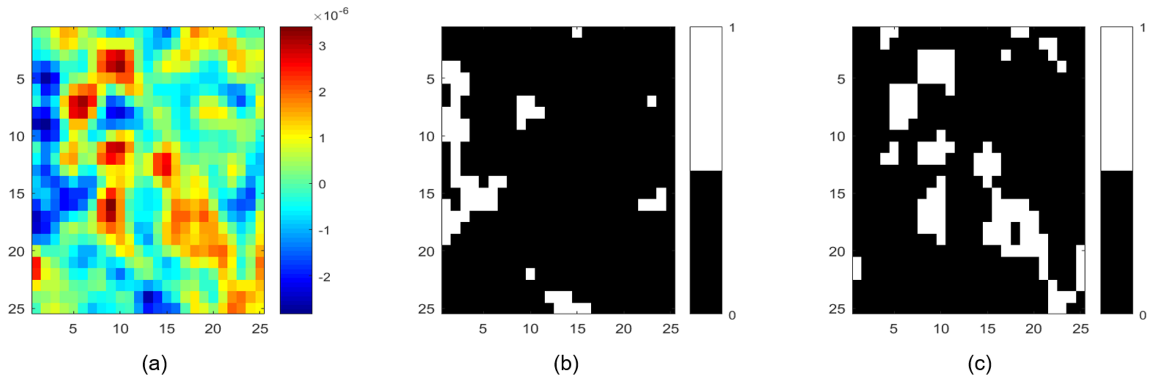


Figure 6.2 Example of (a) an observed surface z, and its corresponding binary results of (b) negative topographic values B_1, and (c) positive topographic values B_2 obtained by the NSHT algorithm.

Algorithm 6.1: Determining the normal surface hard threshold h_1
Given N normal topographic images $Z^{(j)} = \{z_i^{(j)} : i = 1,2,...,M\}$, $j = 1,2,...,N$, and a specified type I error rate α_1 (e.g., $\alpha_1 = 0.1$), we implement the below steps: 1: **for** $j = 1,2,...,N$ **do** 2: obtain the absolute topographic values $E_j = \{e_i^{(j)} : i = 1,2,...,M\}$, $e_i^{(j)} = \lvert z_i^{(j)} \rvert$ 3: rank the obtained values in increasing order such that $E_j^{'} = \{e_{(i)}^{(j)} : i = 1,2,...,M\}$, where $e_{(1)}^{(j)}, e_{(2)}^{(j)}, ..., e_{(M)}^{(j)}$ are the order statistics (i.e., $e_{(1)}^{(j)} \le e_{(2)}^{(j)} \le .... \le e_{(M)}^{(j)}$) 4: obtain the α_1-quantile of the values in the set $E_j^{'}$ $$\Pr[E_j^{'} \ge f_j] = \alpha_1$$ 5: **end for** 6: calculate the normal surface hard threshold as $h_1 = \underset{j=1,2,...,N}{\text{median}} (f_j)$

6.3.1.2 Normal Surface Connected-Component Labeling Algorithm

We overcome the limitations of the existing flood-fill algorithm by proposing the normal surface connected-component labeling (NSCCL) algorithm. The NSCCL algorithm utilizes the information learned from normal surfaces to effectively identify the regions with suspicious characteristics using the binary results obtained by the proposed NSHT algorithm. In particular, the NSCCL algorithm includes two sub-algorithms: Sub-algorithm

6.2-a and Sub-algorithm 6.2-b. In Sub-algorithm 6.2-a, we exploit the flood-fill algorithm to identify the connected regions in the obtained binary results of normal surfaces. In particular, we iterate between the identification and assignment steps of the flood-fill algorithm until all pixels with the value of "1" are assigned to a replacement value $k = 1, 2, \ldots$. Note that $b_m^{(p,j)} \in \Omega(i)$ shown in step 9 indicates that the pixel m is connected to the pixel i in the binary result $B_p^{(j)}$ by a path determined according to a specified neighborhood construction rule Ω. Moreover, in step 10, we set each identified m^{th} neighbor to "zero" ($b_m^{(p,j)} = 0$) to avoid selecting this pixel again in the for-loop in step 5. As a result, we obtain the following weight assignments set $C_k^{(j)} = \{c_i^{(k,j)} : i = 1, 2, \ldots, M\}$, $k = 1, 2, \ldots, K$, where $c_i^{(k,j)} \in \{0,1\}$ is an assigned binary weight that specifies whether the pixel i is located in the k^{th} identified region of the surface j, and K is the number of identified regions. In steps 13 and 17, we obtain the size of each identified region and the total number of identified regions, respectively. Finally, in steps 18 to 21, we obtain the threshold of the normal region size h_2 by calculating the median value of α_2-quantile of normal region sizes, where α_2 is a specified type I error rate, i.e., the rate of incorrectly identifying the normal region sizes as anomaly sizes.

In Sub-algorithm 6.2-b, we exploit the characteristics obtained from Sub-algorithm 6.2-a to effectively identify the topographic regions with suspicious sizes in the binary results of a new surface. In general, Sub-algorithm 6.2-b consists of steps that are similar to the ones presented in Sub-algorithm 2-a. However, in step 13, we identify Q suspicious regions that are associated with region sizes that exceed the threshold of the normal region size h_2.

Consequently, we obtain the following weight assignments set $S_q = \{s_i^{(q)} : i = 1, 2, ..., M\}$, $q = 1, 2, ..., Q$, where $s_i^{(q)} \in \{0,1\}$ is an assigned binary weight that specifies whether the pixel i is located in the q^{th} suspicious region and Q is the number of identified suspicious regions such that $Q \le K$.

Sub-algorithm 6.2-a: learning the characteristics of normal surfaces

Given the binary results of negative and positive topographic values for normal surfaces $B_p^{(j)} = \{b_i^{(p,j)} : i = 1, 2, ..., M\}$, $p = 1, 2$, $j = 1, 2, ..., N$, a neighborhood construction rule of the pixel i, $\Omega(i)$, and a type I error rate α_2 (e.g., $\alpha_2 = 0.05$), we implement the below steps:

for $j = 1, 2, ..., N$ **do**
    set the replacement value $k = 0$
    for $p = 1, 2$ **do**
        for $i = 1, 2, ..., M$ **do**
            if $(b_i^{(p,j)} = 1)$ **do**
                count $k = k + 1$ and set $C_k^{(j)} = \{c_e^{(k,j)} : e = 1, 2, ..., M\}$, $c_e^{(k,j)} = 0$
                set $c_i^{(k,j)} = 1$, update $C_k^{(j)} = \{c_i^{(k,j)}\}$, and set the region size $n = 1$
                for $m = 1, 2, ..., M$ **do**
                    if $(b_m^{(p,j)} = 1)$ and $(b_m^{(p,j)} \in \Omega(i))$ **do**
                    set $c_m^{(k,j)} = 1$, update $C_k^{(j)} = \{c_m^{(k,j)}\}$, set $b_m^{(p,j)} = 0$, and count $n = n + 1$
                    end if
                end for
                Obtain the region size of the k^{th} region $r_k^{(j)} = n$
            end if
        end for
    end for
    obtain the total number of identified regions $K = k$
    rank the obtained region sizes in increasing order $R_j = \{r_{(k)}^{(j)}; k = 1, 2, ..., K\}$, where $r_{(1)}^{(j)}, r_{(2)}^{(j)}, ..., r_{(K)}^{(j)}$ are the order statistics (i.e., $r_{(1)}^{(j)} \le r_{(2)}^{(j)} \le .... \le r_{(K)}^{(j)}$)
    calculate the α_2-quantile of the values in the set R_j as $\Pr[R_j \ge n_j] = \alpha_2$
end for
calculate the threshold of the normal region size as $h_2 = \underset{j=1,2,...,N}{\text{median}} (n_j)$

Sub-algorithm 6.2-b: identifying the topographic regions with suspicious characteristics

Given the binary results of the negative and positive topographic values of a new surface $B_p = \{b_i^{(p)} : i = 1,2,...,M\}$, $p = 1,2,$ a neighborhood construction rule of the pixel i, $\Omega(i)$, and the threshold of the normal region size h_2, we implement the below steps:

set the replacement value $k = 0$ and set the number of suspicious regions $q = 0$
for $p = 1,2$ **do**
    for $i = 1,2,...,M$ **do**
        if $(b_i^{(p)} = 1)$ **do**
            count $k = k + 1$ and set $C_k = \{c_e^{(k)} : e = 1,2,...,M\}$, $c_e^{(k)} = 0$
            set $c_i^{(k)} = 1$, update $C_k = \{c_i^{(k)}\}$, and set the region size $n = 1$
            for $m = 1,2,...,M$ **do**
                if $(b_m^{(p)} = 1)$ and $(b_m \in \Omega(i))$ **do**
                    set $c_m^{(k)} = 1$, update $C_k = \{c_m^{(k)}\}$, set $b_m^{(p)} = 0$, and count $n = n + 1$
                end if
            end for
        Obtain the region size of the k^{th} region $r_k = n$
        if $(r_k \geq h_2)$ **do** count $q = q + 1$ and update $S_q = C_k$ **end if**
        end if
    end for
end for
obtain the total number of identified suspicious regions $Q = q$

The proposed NSCCL algorithm improves the description of local and spatial surface characteristics by identifying the regions with suspicious region sizes and removing the regions with normal sizes. This can improve the detection performance of topographic faults, which are likely to associate with a change in the region size (Rao *et al.*, 2015b). The algorithm is also effective in identifying the surface peaks and valleys with suspicious characteristics because it is implemented separately to the binary results of the negative and positive topographic values (i.e., B_1 and B_2). Figure 6.3 shows a topographic image and its corresponding suspicious regions identified by the NSCCL algorithm. In particular, we apply the NSCCL algorithm to the binary results obtained earlier in Figure 6.2 (b and c) to identify the regions with suspicious sizes.

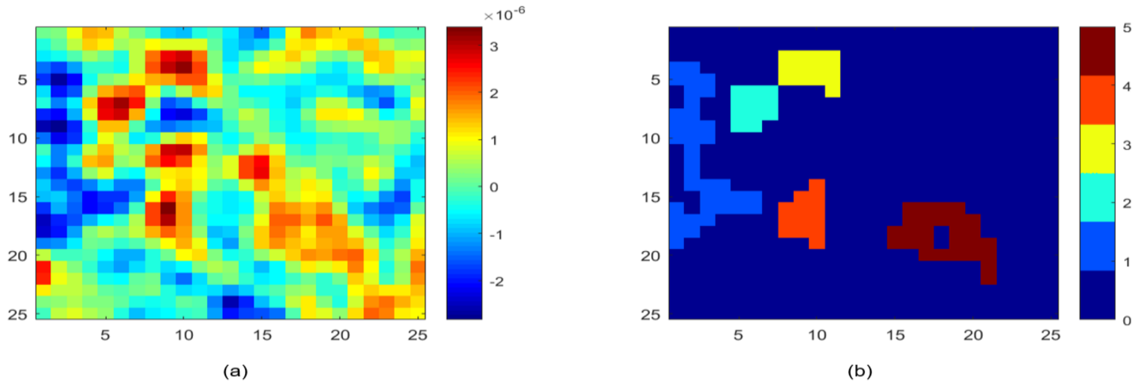


Figure 6.3 Example of (a) an observed topographic image Z, and (b) its corresponding suspicious regions identified by the NSCCL algorithm.

6.3.2 Stage 2: Feature Extraction

We propose the generalized spatially weighted Moran (GSWM) index to accurately quantify the level of the spatial autocorrelation structure of each suspicious region identified by the NSCCL algorithm. The GSWM index is defined as the cross-products of the spatially weighted deviations of the topographic values and their different neighboring orders from the normal mean. More specifically, the GSWM index generalizes the weight assignments in the Moran's index from the first-order neighbor to multiple-order neighbors for adequate characterization of the spatial autocorrelation structure of each identified region. In other words, we utilize the weight assignments obtained by the NSCCL algorithm to define the neighborhood of each pixel at different order. The GSWM index also includes the characteristics of normal surfaces during the quantification of the spatial autocorrelation structure of new surfaces for effective monitoring of surface changes. The GSWM index is also insensitive to the random behavior between the surface characteristics because the spatial structure of each identified region is quantified regardless of the actual coordinate of its pixels on the surface. More important, the GSWM index is suitable for online monitoring since its computation time is expressed as $O(R \times Y)$, where R is the size of neighboring pixels and Y is the size of suspicious surface pixels, which is considerably less than the size of the surface pixels (i.e., $Y << M$).

For a given topographic image $Z = \{z_i : i = 1, 2, ..., M\}$ and its weight assignments set identified by the NSCCL algorithm $S_q = \{s_i^{(q)} : i = 1, 2, ..., M\}$, $q = 1, 2, ..., Q$, the GSWM index of the q^{th} suspicious region is formulated using Equation (6.3)

$$F_q = \frac{\sum_{i=1}^{M}\sum_{m=1}^{M} s_i^{(q)} w_{i,m}^{(q)} (z_i - \bar{z}_0)(z_m - \bar{z}_0)}{\sigma_0^2 \sum_{i=1}^{M}\sum_{m=1}^{M} s_i^{(q)} w_{i,m}^{(q)}}, \; q = 1, 2, ..., Q, \tag{6.3}$$

where $s_i^{(q)}$ is an assigned binary weight that specifies whether the pixel z_i is located in the q^{th} suspicious region, i.e., $s_i^{(q)} = 1$ if z_i is located in the q^{th} suspicious region, $s_i^{(q)} = 0$ otherwise, $w_{i,m}^{(q)}$ is an assigned binary weight that specifies whether the pixels z_i and z_m are neighbors in the q^{th} suspicious region, i.e., $w_{i,m}^{(q)} = 1$ if $\{s_i^{(q)} = 1\}$ and $\{s_m^{(q)} = 1\}$, which indicates that the pixels z_i and z_m are connected to each other by a path of any neighboring order determined based on a specified neighborhood construction rule, $w_{i,m}^{(q)} = 0$ otherwise, and ($\bar{z}_0, \sigma_0$) are the mean and standard deviation of normal topographic values; obtained based on N normal surfaces $Z^{(j)} = \{z_i^{(j)} : i = 1, 2, ..., M\}$, $j = 1, 2, ..., N$, as given in Equations (6.4 and 6.5), respectively,

$$\overline{z}_0 = \frac{1}{N \times M} \sum_{j=1}^{N} \sum_{i=1}^{M} z_i^{(j)} \text{ , and} \tag{6.4}$$

$$\sigma_0 = \frac{1}{N} \sum_{j=1}^{N} \sqrt{\sum_{i=1}^{M} (z_i^{(j)} - \overline{z}_0)^2 / (M-1)} \text{ .} \tag{6.5}$$

6.3.3 Stage 3: Anomaly Detection

In this stage, we determine whether a new surface $Z^{(new)}$ is an anomaly according to given N normal surfaces $Z^{(j)}$, $j = 1, 2, ..., N$. Specifically, after we calculate the GSWM index for all suspicious regions of the j^{th} normal surface $F_q^{(j)}$, $q = 1, 2, ..., Q^{(j)}$, and $j = 1, 2, ..., N$, we obtain the maximum generalized spatially weighted Moran (MGSWM) index as given in Equation (6.6), which describes the spatial autocorrelation level of the most suspicious region of a surface

$$A^{(j)} = \max_{q=1,2,...,Q^{(j)}} F_q^{(j)}, \; j = 1, 2, ..., N \text{ ,} \tag{6.6}$$

where $A^{(j)}$ is the MGSWM index of the surface j. Note that we expect the MGSWM index of topographic images obtained under normal process conditions to be statistically similar since these surfaces are composed of similar characteristics. However, when a new surface contains faults, these faults are likely to cause a local change in the structure of that surface (Alqahtani *et al.*, 2020d). Subsequently, we expect the MGSWM index to detect this change by quantifying the spatial autocorrelation level of the most suspicious region of that surface.

After the MGSWM index is calculated for all given normal surfaces $A^{(j)}$, $j = 1, 2, ..., N$, we obtain the upper critical limit U using Algorithm 6.3. In particular, Algorithm 6.3

determines U based on a bootstrapping method with a prespecified type I error rate, i.e., the rate of incorrectly detecting the surfaces with normal characteristics as anomaly surfaces (Alqahtani *et al.*, 2020c). Thus, when a new surface $Z^{(new)}$ is observed, we calculate its MGSWM index $A^{(new)}$. We subsequently compare $A^{(new)}$ with the upper critical limit U. If $A^{(new)} > U$, the new surface is then identified as an anomaly.

Algorithm 6.3: Determining the upper critical limit U

Given the MGSWM index set of N normal surfaces $A_0 = \{A^{(j)}, j = 1,2,...,N\}$, the size of bootstrapping samples C (e.g., $C = 1000$), the size of simulation replicates V (e.g., $V = 1000$), and a specified type I error rate α_3 (e.g., $\alpha_3 = 0.05$), we implement the below steps:

7: **for** $v = 1,2,...,V$ **do**

8: obtain the empirical distribution of the set A_0 by randomly selecting C samples with replacement from the set A_0 such as $Y_v = \{y_c, c = 1,2,...,C\}$, $y_c \in A_0$

9: rank the values in the set Y_v in increasing order $Y_v^{'} = \{y_{(c)}, c = 1,2,...,C\}$, where $y_{(1)}, y_{(2)},...., y_{(C)}$ represent the order statistics (i.e., $y_{(1)} \le y_{(2)} \le .... \le y_{(C)}$)

10: compute the α_3-quantile of the values in the set $Y_v^{'}$ for the replicate v

$$\Pr[Y_v^{'} \ge p_v] = \alpha_3$$

11: **end for**

12: determine the upper critical limit by taking the median over the obtained replicates

$$U = \underset{v=1,2,...,V}{\text{median}} \, (p_v)$$

6.3.4 Stage 4: Fault Identification and Diagnosis

In this stage, we extract different fault information from each identified anomaly surface (Alqahtani *et al.*, 2020d). In particular, we obtain the number of anomaly regions, denoted as T_1, by counting the number of regions with the GSWM index that exceeds the specified critical limit U such as $T_1 = \sum_{q=1}^{Q} 1\{F_q > U\}$. After identifying the anomaly regions, we extract the following fault information from each anomaly region: fault size, location, magnitude, autocorrelation, and type. The fault size, denoted as T_2, is defined as the unit area of each identified anomaly region, which is calculated by counting the number of pixels in each anomaly region. In particular, given the weight assignments set of the q^{th} anomaly region $S_q = \{s_i^{(q)} : i = 1, 2, ..., M\}$, we calculate its fault size by counting the pixels with the value of "1" as $T_2 = \sum_{i=1}^{M} 1\{ s_i^{(q)} = 1 \}$. In addition, the fault location, denoted as T_3, is defined as the locations of pixels in each anomaly region. Therefore, the fault location of the q^{th} anomaly region is given as $T_3 = S_q = \{s_i^{(q)}, i = 1, 2, ..., M\}$.

The fault magnitude, denoted as T_4, measures the variation of topographic values from the reference plane. More specifically, T_4 is commonly quantified based on the standard deviation (SD) of topographic values (Rao *et al.*, 2015b). However, the SD may overlook local variance changes since the values of local changes can be masked by the normal values. Thus, we overcome this limitation by proposing the spatial standard deviation (SSD) index, which quantifies the fault magnitude of each identified anomaly region by calculating the square root of the average of spatially weighted squared differences of the topographic values from the normal mean. Particularly, given the weight assignments set of the q^{th} anomaly region $S_q = \{s_i^{(q)} : i = 1, 2, ..., M\}$, the fault magnitude of the q^{th} anomaly region is calculated based on its SSD index as given in Equation (6.7)

$$T_4 = D_q = \sqrt{\sum_{i=1}^{M} s_i^{(q)} (z_i - \overline{z}_0)^2 / (\sum_{i=1}^{M} s_i^{(q)} - 1)}, \qquad (6.7)$$

where D_q is the SSD index of the q^{th} anomaly region and $\bar{z}_0$ is the mean of normal topographic values defined earlier in Equation (6.4). The fault autocorrelation, denoted as T_5, measures the level of spatial autocorrelation among pixels in each anomaly region. Specifically, T_5 is obtained based on the proposed GSWM index, which is effective in quantifying the spatial autocorrelation among topographic pixels. Thus, the fault autocorrelation of the q^{th} anomaly region is defined as $T_5 = F_q$.

The fault type, denoted as T_6, identifies the type of surface change (i.e., variance or autocorrelation) that occurred in each identified anomaly region. The proposed SSD and GSWM indexes are simultaneously used to identify T_6. Specifically, when an anomaly region experiences a variance change, we expect both SSD and GSWM indexes of that region to exceed their critical limits W and U, respectively, where W and U are obtained using Algorithm 6.3. This is because both SSD and GSWM indexes can capture the variation of topographic values from the normal mean. Thus, if $(T_4 > W)$ and $(T_5 > U)$, then we identify the fault type of the q^{th} anomaly region as $T_6 =$ "variance change". However, when an anomaly region is observed with an autocorrelation change, the GSWM index of that region is only expected to exceed its critical limit. This is because the GSWM index is only effective in capturing changes in the spatial autocorrelation of surface structure. Therefore, if $(T_4 < W)$ and $(T_5 > U)$, we identify the fault type of the q^{th} anomaly region as $T_6 =$ "autocorrelation change".

6.4 Performance Analysis

We use the Gaussian random model described in Section 2.3 to generate topographic surfaces. We also consider the same fault scenarios described earlier in Section 3.3.1.

6.4.1 Effectiveness of the Proposed NSHT and NSCCL Algorithms

In this section, we initially present the effectiveness of the proposed NSHT algorithm in representing changes in surface characteristics. Figure 6.4 shows examples of the selected regions of three generated surfaces and their corresponding binary results obtained by the NSHT algorithm. Note that we integrate the negative and positive binary results (i.e., $B = B_1 + B_2$) to better illustrate the change in each surface. In general, we observe that the NSHT algorithm effectively captures both valley (blue color) and peak (red color) characteristics due to the separate binarization of negative and positive topographic values. We also notice that the surface characteristics of the normal surface, which are represented by "1" (or white areas) in the obtained binary results, tend to be more random and distant from each other. However, the characteristics of the binary result obtained under the variance change are appeared to be distinguished from the ones of the normal surface such a way that they become more clustered and closer to each other. This is due to the fact that the variance change yields more topographic values to exceed the normal surface hard threshold (i.e., $h_1 = 1.28\text{E-}06$). Moreover, the surface characteristics of the binary results obtained under the autocorrelation change are shown to be different from the ones of the normal surface as they become more clustered with their neighbors. This is because of the change in the spatial relationships among topographic values.

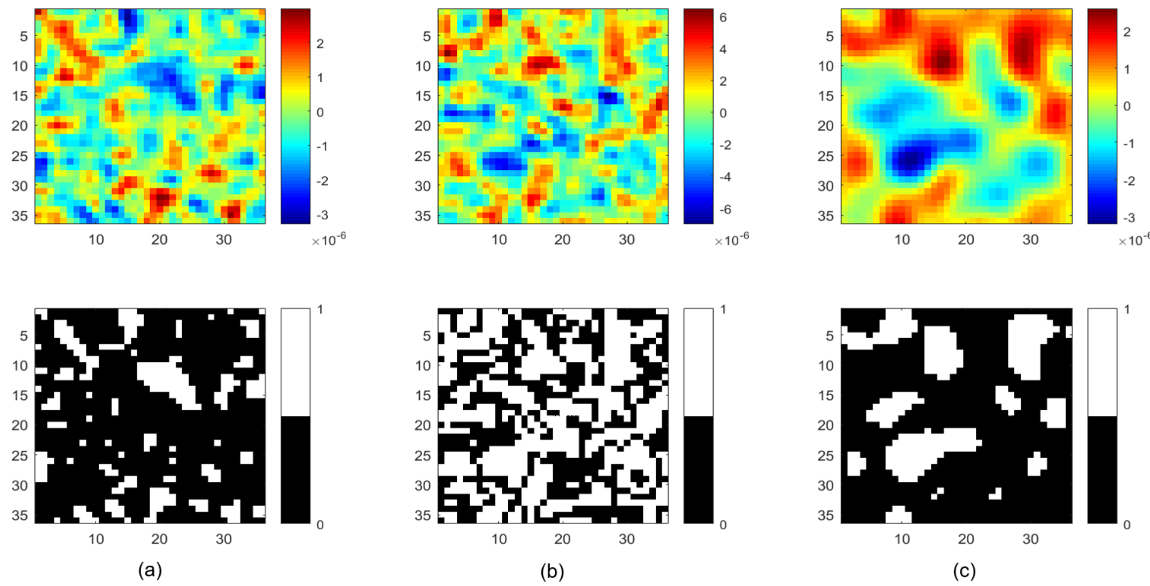


Figure 6.4 Three topographic images and their corresponding binary results obtained by the NSHT algorithm from (a) normal surface, and two anomaly surfaces with (b) variance change, and (c) autocorrelation change.

We also present the effectiveness of the proposed NSCCL algorithm in identifying the topographic regions with suspicious characteristics. After we obtain the binary results of the three generated surfaces presented in Figure 6.4, we apply the NSCCL algorithm to identify the regions with suspicious characteristics as shown in Figure 6.5. We notice that the NSCCL algorithm effectively assigns each connected region that is associated with a region size that exceeded the threshold of the normal region size (i.e., $h_2 = 19$) to a distinct label (or color). In particular, Figure 6.5 (a) shows only two suspicious regions identified from the normal surface. However, we notice an increase in the number of suspicious regions identified from the anomaly surfaces where 14 and 5 suspicious regions identified from the surfaces with the variance and autocorrelation changes as shown in Figure 6.5 (b and c), respectively.

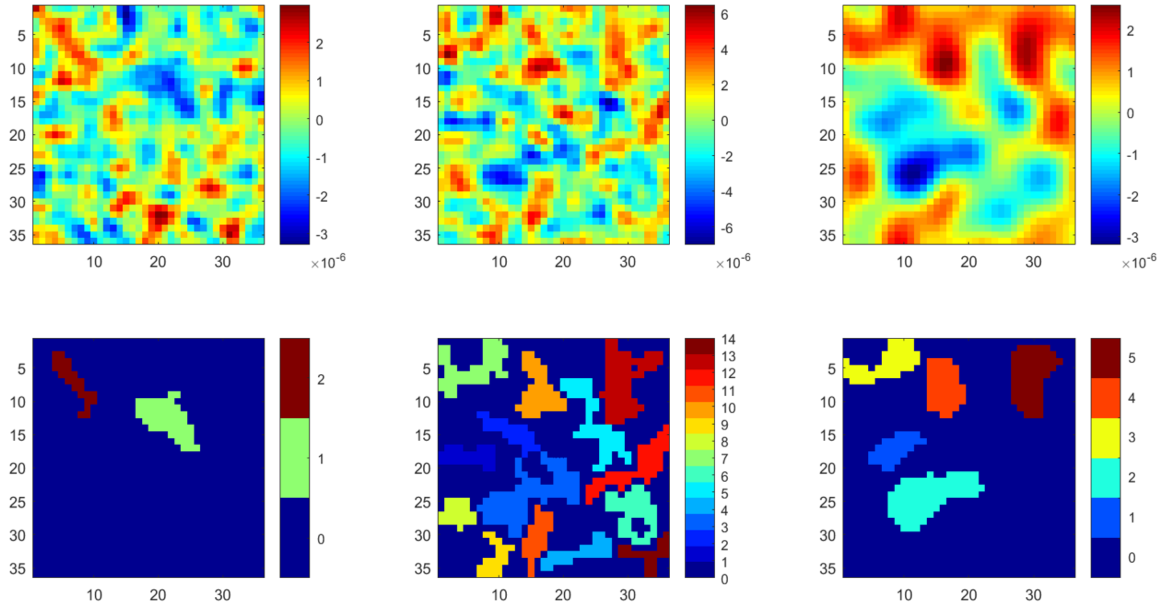


Figure 6.5 Illustrations of the regions with suspicious characteristics identified by the NSCCL algorithm from (a) a normal surface, and two anomaly surfaces with (b) variance change, and (c) autocorrelation change.

6.4.2 Performance of the Proposed Approach

This section assesses the performance of the proposed monitoring approach with the existing approaches discussed earlier in Section 2.4. Note that we use the performance measure stated earlier in Equation (3.15). Particularly, Table 6.1 presents the power of detection P of the proposed and the existing monitoring approaches under different fault sizes, areas, and types. We observe that the power spectral density feature S_{psd} yields low detection performance of local faults since S_{psd} is calculated using a global decomposition filter (the FFT filter), which is likely to remove the local fault information. The watershed feature S_w also results in low power of detection because S_w is heavily dependent on the watershed algorithm that suffers from the over-segmentation issue, which can cause low segmentation performance of local faults. Next, the Fiedler connectivity feature S_c shows improved detection performance of the variance change due to the quantification of the pairwise distance between all nodes of the graph network. However, S_c yields poor performance in detecting the autocorrelation change because S_c does not capture the spatial autocorrelation among topographic pixels.

We also show the performance of Moran's index I, which results in a lower power of detection for both variance and autocorrelation changes due to the limitations of this index discussed earlier in Section 6.2. Next, the one-sample Anderson–Darling approach S_{ad} results in a high detection accuracy for the variance change because S_{ad} captures the deviation of the CDF of the observed residuals from the CDF of the normal residuals. However, S_{ad} yields a low detection performance for the autocorrelation change because the normal surface characteristics are likely to obscure the characteristics of local faults.

In contrast, the proposed NSHT algorithm is shown to be effective in representing the variance and autocorrelation changes in a way that they are likely to cause a change in the structure of the binarized topographic values, and the proposed NSCCL algorithm is shown to be accurate in utilizing the binarized values to identify the regions of with suspicious structures. In addition, the proposed MGSWM index A is effective in characterizing and monitoring changes in surface characteristics by calculating the spatial autocorrelation level of the region with the most suspicious structure. As a result, A yields outstanding detection performance for both variance and autocorrelation changes under all specified fault scenarios and subsequently outperforms the other monitoring approaches as shown in Table 6.1.

Table 6.1 Power of detection P of the proposed and the existing monitoring approaches for anomaly surfaces generated by the Gaussian random model

Fault type	Fault region	Fault size		Monitoring approaches					
		σ	θ	S_{psd}	S_w	S_c	I	S_{ad}	A
Variance	2.5%	1.5	1	0.238	0.123	0.409	0.149	0.568	**0.714**
		1.75	1	0.296	0.129	0.677	0.160	0.928	**0.993**
	3.0%	1.5	1	0.240	0.130	0.401	0.158	0.645	**0.794**
		1.75	1	0.278	0.130	0.568	0.166	0.963	**0.998**
Autocorrelation	2.5%	1	2	0.102	0.262	0.206	0.205	0.263	**0.760**
		1	2.25	0.104	0.339	0.217	0.255	0.263	**0.876**
	3.0%	1	2	0.105	0.331	0.222	0.234	0.298	**0.813**
		1	2.25	0.108	0.411	0.234	0.288	0.307	**0.917**

We also assess and compare the diagnostic performance of the proposed and the existing monitoring approaches. Particularly, after we detect an anomaly surface, we calculate the following three diagnostic performance measures: (1) precision; expressed as the percentage of the identified anomalies that are true anomalies (this may include true and untrue anomalies), (2) recall; expressed as the percentage of the true anomalies that are correctly identified, and (3) F1-score; expressed as the harmonic mean of the precision and recall measures (Janssens *et al.*, 2015). These measures are respectively given in Equations (6.8-6.10)

$$\text{Precision} = \frac{1}{V}\sum_{j=1}^{V} P_j = \frac{1}{V}\sum_{j=1}^{V}\frac{TN_j}{TN_j + FN_j}, \tag{6.8}$$

$$\text{Recall} = \frac{1}{V}\sum_{j=1}^{V} R_j = \frac{1}{V}\sum_{j=1}^{V} \frac{TN_j}{TN_j + FP_j}, \tag{6.9}$$

$$\text{F1-score } = \frac{1}{V}\sum_{j=1}^{V} 2 \times \frac{P_j \times R_j}{P_j + R_j}, \tag{6.10}$$

where V is the size of simulation replicates ($V = 100$ replicates), (P_j, R_j) are the precision and recall measures in the j^{th} simulation replicate, TN_j is the number of true negative alarms, i.e., the number of anomaly pixels that are correctly identified as anomaly pixels, FN_j is the number of false negative alarms, i.e., the number of normal pixels that are incorrectly identified as anomaly pixels, and FP_j is the number of false positive alarms, i.e., the number of anomaly pixels that are incorrectly identified as normal pixels.

In Table 6.2, we present the precision, recall, and F1-score measures for the proposed and the existing approaches. Particularly, we generate anomaly Gaussian random surfaces under two fault types, two fault sizes, and one fault area (i.e., 2.5%). It is clear that the proposed approach, denoted as A, has a much better diagnostic performance than the other approaches (i.e., S_w and S_{ad}). This difference is more apparent in the autocorrelation change where S_w and S_{ad} fail to capture the spatial structure of surface anomalies. In particular, for the variance change with the fault size 1.5 μm, the F1-score, which is based on precision and recall measures, of S_w and S_{ad} are 0.609 and 0.293, respectively, whereas this measure is 0.825 for A. In addition, for the autocorrelation change with the fault size 2 μm, the F1-score of A is 0.822, which is much better than the F1-scores of S_w and S_{ad}, which are obtained as 0.390 and 0.273, respectively. This is because the spatial structure of surface anomalies is effectively considered in the proposed approach through the GSWM index.

Table 6.2 Diagnostic performance results of the proposed and the existing monitoring approaches for anomaly surfaces generated by the Gaussian random model

Fault type	Fault size		Precision			Recall			F1		
	σ	θ	S_w	S_{ad}	A	S_w	S_{ad}	A	S_w	S_{ad}	A
Variance	1.5	1	0.625	0.743	**0.843**	0.614	0.185	**0.817**	0.609	0.293	**0.825**
	1.75	1	0.632	0.766	**0.866**	0.839	0.312	**0.928**	0.719	0.439	**0.893**
Autocorrelation	1	2	0.633	0.711	**0.817**	0.046	0.173	**0.845**	0.390	0.273	**0.822**
	1	2.25	0.773	0.697	**0.809**	0.021	0.167	**0.885**	0.268	0.264	**0.838**

Figure 6.6 shows examples of normal and two anomaly surfaces, along with the corresponding locations of anomaly regions extracted by the proposed approach. In Figure 6.6, we observe that no normal region is incorrectly identified as anomalies for all three surfaces, whereas 2 and 9 anomaly regions are correctly identified as anomalies from the anomaly surfaces in Figure 6.6 (b and c), respectively. Moreover, in Table 6.3, we extract additional fault information from each identified anomaly region in Figure 6.6 (b and c), including the fault size T_2, the fault magnitude T_4 , and the fault autocorrelation T_5. We also identify the fault type T_6 of the anomaly regions obtained from Figure 6.6 (b) as "variance change" since T_4 and T_5 of these regions exceeded their critical limits (i.e., $W = 5.56E-07$ and $U = 87.56$). However, we identify the fault type T_6 of the anomaly regions obtained from Figure 6.6 (c) as "autocorrelation change" because only T_5 of these regions exceeded its critical limit (i.e., $U = 87.56$).

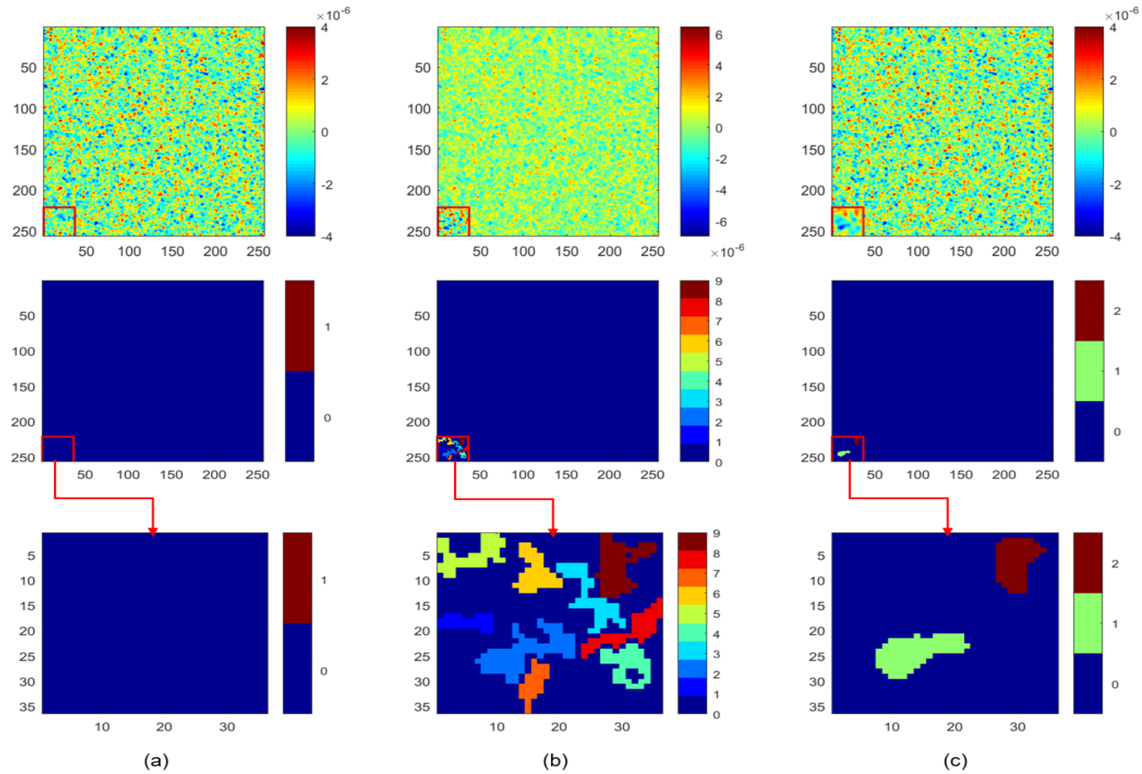


Figure 6.6 Examples of thee three generated surfaces (first row) and their corresponding anomaly regions (second and third rows) identified by the proposed approach from (a) a normal surface, and two anomaly surfaces with (b) variance change, and (c) autocorrelation change.

Table 6.3 Summary of the fault information extracted from anomaly surfaces shown in Figure 6.6

Surface index	Region	Fault information			
		T_2	T_4	T_5	T_6
b	1	26	**9.83E-07**	**142.63**	Variance change
	2	83	**1.17E-06**	**400.10**	
	3	50	**1.26E-06**	**184.10**	
	4	43	**5.83E-07**	**93.81**	
	5	49	**1.21E-06**	**232.94**	
	6	42	**1.26E-06**	**208.81**	
	7	28	**8.95E-07**	**99.13**	
	8	45	**9.53E-07**	**161.51**	
	9	65	**7.70E-07**	**212.39**	
c	1	78	5.20E-07	**155.42**	Autocorrelation change
	2	60	3.31E-07	**95.83**	

Finally, we evaluate the average computation time for calculating the proposed and the existing monitoring approaches. Particularly, we calculate the average and standard deviation of the computation time for each approach based on 1000 samples as presented in Table 6.4. Referring to the proposed approach as A, its computation time is the total time that the approach takes from Stage 1 to Stage 4. Due to the model complexity of A, the average computation time of A is larger than the other approaches such as S_{psd} and S_c. However, it is still efficient and smaller than the other complex approaches, such as S_w, I, and S_{ad}. In particular, A needs an average of 0.34 seconds to detect an anomaly surface, while S_w, I, and S_{ad} need an average of 0.874, 0.772, and 5.363 seconds, respectively. This makes the proposed approach suitable for the online monitoring of topographic images.

Table 6.4 Average computation time of the proposed and the existing monitoring approaches

Monitoring approaches	S_{psd}	S_w	S_c	I	S_{ad}	A
Time in seconds (standard deviation)	0.001 (7.54E-05)	0.874 (0.024)	0.012 (0.002)	0.772 (0.001)	5.575 (0.067)	**0.340 (0.009)**

6.5 A Case Study of Semiconductor Wafer Surfaces

In this section, we examine the effectiveness of the proposed approach by considering the case study of semiconductor wafer presented earlier in Section 3.4. In Table 6.5, we present the performance comparison results in terms of the power of detection of anomaly wafers for the proposed and the existing monitoring approaches. We observe that the existing approaches result in low power of detection for both scratch and ridge faults because they lack the quantification of local changes in topographic surfaces. The proposed approach, in contrast, outperforms the existing approaches under all specified fault scenarios because of the accurate identification, characterization, and monitoring of local changes in topographic surfaces. In addition, Figure 6.7 shows the effectiveness of the proposed approach in identifying faults in anomaly wafers. In particular, the proposed approach correctly identified 26 scratches and 20 ridges from the anomaly wafers with an accuracy of 81.25% and 100%, respectively, as shown in Figure 6.7 (b and c).

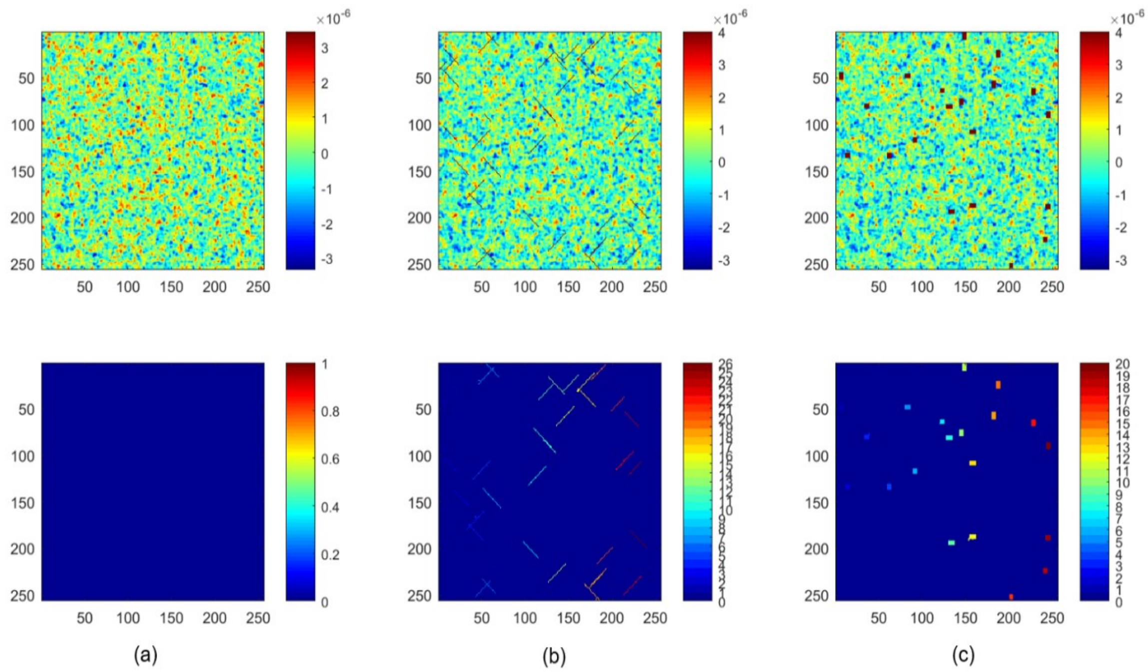


Figure 6.7 Examples of the topographic images of (a) normal surface, and two anomaly wafers with (b) scratch faults, and (c) ridge faults, along with the anomaly regions identified by the proposed approach.

Table 6.5 Power of detection of the proposed and the existing monitoring approaches using anomaly wafers

Fault type	Fault area	Monitoring approaches				
		S_{psd}	S_w	S_c	I	A
Scratches	1.5%	0.087	0.191	0.141	0.303	**0.769**
	2%	0.092	0.212	0.143	0.349	**0.866**
	2.5%	0.094	0.272	0.148	0.445	**0.963**
Ridges	2%	0.094	0.109	0.516	0.550	**0.765**
	2.5%	0.094	0.114	0.620	0.660	**0.855**
	3%	0.098	0.129	0.697	0.737	**0.915**

6.6 Conclusions

In this chapter, we propose the generalized spatially weighted autocorrelation approach for accurate monitoring and detection of different patterns of local surface faults. We propose two algorithms to assign spatial weights to the important characteristics of new surfaces based on information learned from normal surfaces. The NSHT algorithm enhances the surface representation by binarizing the important surface characteristics, and the NSCCL algorithm exploits the obtained binary results to identify and assign spatial weights to the topographic regions with suspicious characteristics. The developed GSWM index subsequently utilizes the assigned weights to characterize the spatial autocorrelation structure of each identified region. After an anomaly surface is detected based on its maximum GSWM index, different fault information is extracted for fault diagnosis.

The proposed approach shows its robustness and efficiency in monitoring and diagnosing various faults that emerge in different types of topographic images and subsequently outperforms the existing monitoring approaches. Specifically, the NSHT algorithm is shown to be effective in representing changes in surface characteristics through binarization. Since the variance and autocorrelation changes are likely to cause a change in the structure of the obtained binary results, the NSCCL algorithm is proven to be accurate in identifying the regions with suspicious structures. Different changes in surface characteristics, including both variance and autocorrelation, are accurately described and detected by monitoring the GSWM index of the most suspicious regions. Since the proposed approach only characterizes and monitors the regions with suspicious characteristics, this makes the approach efficient to implement in the online monitoring environment.

CHAPTER 7

CONCLUSION AND FUTURE RESEARCH

In this chapter, we present the summary and conclusions of this dissertation in Section 7.1 and discuss future research in Section 7.2.

7.1 Summary and Conclusions

The proposed monitoring approaches bridge the gap between the existing literature and the need for effective online monitoring of 3D topographic surfaces. In particular, the proposed approaches show superior performance compared to the existing approaches in characterizing and assessing the complex characteristics of topographic surfaces. We briefly summarize the four proposed monitoring approaches as follows:

1. We introduce the multi-level spatial randomness approach for monitoring global changes in 3D topographic surfaces. The approach enhances the representation of topographic values by slicing the 3D surface topography into different levels in reference to the characteristics of normal surfaces through the developed multi-level surface thresholding algorithm. We effectively quantify the spatial and random properties of topographic values at each surface level using the proposed spatial randomness (SR) profile. We also develop a monitoring statistic based on the functional principal component analysis for effectively detecting different forms of global changes in topographic surfaces based on their SR profiles.
2. We present the multi-label separation-deviation surface model for detecting local variance changes in 3D topographic surfaces. The approach improves the

representation of topographic defects through the developed multi-label separation-deviation surface (MSS) model, which labels the important surface characteristics and smoothes out the noisy ones. We develop two features for monitoring changes in surface characteristics. The MSS feature is introduced for capturing deviations within the label assignments, and the generalized spatial randomness feature is derived for quantifying deviations between the label assignments. These two features are integrated into a single monitoring statistic to detect local variance changes in topographic surfaces.

3. We develop a spatially weighted graph theory-based approach for accurate monitoring of local autocorrelation changes in 3D topographic surfaces. The approach improves the representation of surface characteristics by segmenting the observed surface pixels into predefined clusters through the developed in-control multi-region surface segmentation algorithm. We also propose the maximum local spatial randomness feature, which accurately describes the local and spatial characteristics among topographic values. After representing the surface as a spatially weighted graph network, we monitor its connectivity through the developed spatial graph connectivity to detect changes in topographic structures.
4. We present the generalized spatially weighted autocorrelation approach for fault detection and diagnosis in topographic surfaces. We propose two algorithms to identify and assign spatial weights to the regions with suspicious characteristics. The normal surface "hard" thresholding algorithm initially enhances the representation of surface characteristics through binarization and followed by the normal surface connected-component labeling algorithm, which utilizes the

obtained binary results to identify and assign spatial weights to the suspicious topographic regions. We also develop the generalized spatially weighted Moran (GSWM) index, which exploits the assigned weights to characterize and monitor local changes in the spatial autocorrelation structure of each identified region. After an anomaly surface is detected based on its maximum GSWM index, we extract different fault information for fault diagnoses.

7.2 Future Research

The developed approaches for monitoring surfaces assume an independent monitoring scheme, i.e., the past monitored surfaces do not influence the monitoring of new surfaces. This may result in a lower assessment of small process shifts that may occur during manufacturing. Therefore, effective monitoring of small process shifts can be accomplished by using the characteristics of topographic surfaces from previous images to improve the current monitored image analogous to the Exponentially Weighted Moving Average (EWMA) and cumulative sum (CUSUM) in SPC, which are both sensitive in detecting small process shifts (Qiu, 2013).

Multistage processes are used in many manufacturing processes, such as automotive manufacturing processes, oil refining processes, fabrication processes, and batch manufacturing processes (Kim *et al.*, 2019, Sangahn, 2019). Monitoring and controlling multistage processes are of vital importance to detect shifts in such processes. However, variation propagation from one stage to another is one of the most challenges in multistage processes, which is known as the cascade property, i.e., the output of upstream stages is likely to influence the input of downstream stages (Kim *et al.*, 2019). Examples of variations that may occur in any process stage include operator errors, defective raw materials, or jammed machines (Montgomery, 2007). Existing multistage monitoring approaches are limited to univariate or multivariate data (Kim *et al.*, 2019, Sangahn, 2019). However, multistage monitoring approaches based on topographic data are sparse. Therefore, the proposed approaches can be extended from a single stage process to multistage processes where the topographic images among stages can be correlated to determine potential shifts in the process parameters and identify “faulty” stages.

In contrast to the traditional monitoring approaches, multimode monitoring approaches are designed for processes with multiple operating modes, which is known as multimode processes (Yu and Qin, 2008, Jin and Liu, 2013, Turkoz, 2018). In such processes, each operating mode can be characterized by a single baseline model, and subsequently multiple baseline models can be integrated for effective monitoring of the quality of overall system behavior. In the proposed monitoring approaches, we assume that the process has only a single operating mode under normal process behavior. Therefore, extending the proposed approaches to multimode processes is crucial to effectively detect anomaly surfaces in such processes.

Appendix A. Proof of Lemma 4.1

We show the proof of lemma 4.1 for obtaining the optimal weights of the GSR feature discussed in Section 4.2.2. For the given n_l and n_r, where n_l is the number of pixels with the label l and $n_r = \sum_{j=1|j\neq l}^{k} n_j$ is the number of the pixels without the label l, the first and the second moments of J_{ll} are, respectively, given as follows (Cliff and Ord, 1981)

$$\mathrm{E}(J_{ll}) = cn_l^{(2)} / n^{(2)}, \text{ and } \mathrm{E}(J_{ll}^2) = s_0^2 \, [\frac{n_l^{(4)}}{n^{(4)}}] + s_1 \, [\frac{n_l^{(2)}}{n^{(2)}} - \frac{2n_l^{(3)}}{n^{(3)}} + \frac{n_l^{(4)}}{n^{(4)}}] + s_2 \, [\frac{n_l^{(3)}}{n^{(3)}} - \frac{n_l^{(4)}}{n^{(4)}}],$$

where $c = \sum_{l=1}^{k} \sum_{j=1|j\leq l}^{k} J_{lj}$ is the sum of JC statistics and $n^{(k)} = \prod_{i=1}^{k} (n-i+1)$ is a positive integer constant. We refer to Cliff and Ord (1981) to obtain s_0, s_1, and s_2. Therefore, we derive the expectation and variance of S_l, $l = 1,2,...,k$, as follows

$$\begin{aligned}
&\mathrm{E}(S_l) = d_l \mathrm{E}(J_{ll}) + d_r \mathrm{E}(J_{rr}) = \frac{c}{n^{(2)}} (d_l n_l^{(2)} + d_r n_r^{(2)}), \text{ and} \\
&\mathrm{Var}(\mathrm{S}_l) = \mathrm{E}(\mathrm{S}_l^2) - \mathrm{E}(S_l)^2 \\
&\qquad = \mathrm{s}_0^2 \begin{bmatrix} d_l^2 \left(\dfrac{n_l^{(4)}}{n^{(4)}} - \left\{ \dfrac{n_l^{(2)}}{n^{(2)}} \right\}^2 \right) + d_r^2 \left(\dfrac{n_r^{(4)}}{n^{(4)}} - \left\{ \dfrac{n_r^{(2)}}{n^{(2)}} \right\}^2 \right) \\ -2d_l d_r \left(\dfrac{n_l^{(2)} n_r^{(2)}}{n^{(4)}} - \dfrac{n_l^{(2)} n_r^{(2)}}{n^{(2)} n^{(2)}} \right) \end{bmatrix}
\end{aligned}$$

$$+\,\mathrm{s}_1\begin{bmatrix} d_l^2\left(\dfrac{n_l^{(2)}}{n^{(2)}}-2\dfrac{n_l^{(3)}}{n^{(3)}}+\dfrac{n_l^{(4)}}{n^{(4)}}\right)+d_r^2\left(\dfrac{n_r^{(2)}}{n^{(2)}}-2\dfrac{n_r^{(3)}}{n^{(3)}}+\dfrac{n_r^{(4)}}{n^{(4)}}\right)+ \\ 2d_l d_r\dfrac{n_l^{(2)}n_r^{(2)}}{n^{(4)}} \end{bmatrix}$$

$$+\,\mathrm{s}_2\begin{bmatrix} d_l^2\left(\dfrac{n_l^{(3)}}{n^{(3)}}-\dfrac{n_l^{(4)}}{n^{(4)}}\right)+d_r^2\left(\dfrac{n_r^{(3)}}{n^{(3)}}-\dfrac{n_r^{(4)}}{n^{(4)}}\right) \\ -2d_l d_r\dfrac{n_l^{(2)}n_r^{(2)}}{n^{(4)}} \end{bmatrix}.$$

As $n \to \infty$, it can be assumed that $n_l^{(k)} / n^{(k)} \approx \beta_l^k$, $n_r^{(k)} / n^{(k)} \approx \beta_r^k$. Therefore, the variance of S_l can be written as

$$\mathrm{Var}(\mathrm{S}_l) = (s_2 - 4s_0^2 / n)\beta_l\beta_r(d_l\beta_l - d_r\beta_r)^2 + \mathrm{s}_1\ \beta_l^2\beta_r^2\ (d_l + d_r)^2.$$

Thus, after minimizing $\mathrm{Var}(\mathrm{S}_l)$ subjected to $d_l + d_r = 1$, the optimum solution is obtained as $(d_l^*, d_r^*) = (\beta_r, \beta_l)$.

References

ALI, S. H. 2012. Advanced nanomeasuring techniques for surface characterization. *ISRN Optics,* 2012**,** 1-23.

ALQAHTANI, M. A., JEONG, M. K. and ELSAYED, E. A. 2020a. Generalized spatially weighted autocorrelation approach for monitoring faults in 3D topographic surfaces with application to wafer surface monitoring. Working paper.

ALQAHTANI, M. A., JEONG, M. K. and ELSAYED, E. A. 2020b. Multi-label separation-deviation surface model for detecting spatial defects in topographic surfaces. *IEEE Transactions on Industrial Informatics*. Accepted.

ALQAHTANI, M. A., JEONG, M. K. and ELSAYED, E. A. 2020c. Multilevel spatial randomness approach for monitoring changes in 3D topographic surfaces. *International Journal of Production Research,* 58**,** 5545-5558.

ALQAHTANI, M. A., JEONG, M. K. and ELSAYED, E. A. 2020d. Spatially weighted graph theory-based approach for monitoring faults in 3D topographic surfaces. *International Journal of Production Research*. Accepted.

ASPINWALL, D., SOO, S., BERRISFORD, A. and WALDER, G. 2008. Workpiece surface roughness and integrity after WEDM of Ti–6Al–4V and Inconel 718 using minimum damage generator technology. *CIRP Annals-Manufacturing Technology,* 57**,** 187-190.

BEUCHER, S. and MEYER, F. 1993. The morphological approach to segmentation: the watershed transformation. *Mathematical Morphology in Image Processing,* 34**,** 433-481.

BHUSHAN, B. 2000. *Modern tribology handbook, two volume set*, CRC press.

BLUNT, L. and JIANG, X. 2003. *Advanced techniques for assessment surface topography: development of a basis for 3D surface texture standards " surfstand"*, Elsevier.

BOYKOV, Y., VEKSLER, O. and ZABIH, R. 2001. Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 23**,** 1222-1239.

BROWN, L. and BLUNT, L. 2008. Surface metrology for the automotive industry. *Inaugural Automotive Researchers' Conference.* University of Huddersfield.

BRUNE, D., HELLBORG, R., WHITLOW, H. J. and HUNDERI, O. 2008. *Surface characterization*, John Wiley & Sons.

BUI, A. T. and APLEY, D. W. 2018a. A monitoring and diagnostic approach for stochastic textured surfaces. *Technometrics,* 60**,** 1-13.

BUI, A. T. and APLEY, D. W. 2018b. Monitoring for changes in the nature of stochastic textured surfaces. *Journal of Quality Technology,* 50**,** 363-378.

CLIFF, A. D. and ORD, J. K. 1981. *Spatial processes: models & applications*, Pion.

FEBRERO, M., GALEANO, P. and GONZÁLEZ-MANTEIGA, W. 2008. Outlier detection in functional data by depth measures, with application to identify abnormal NOx levels. *Environmetrics,* 19**,** 331-345.

GAETANO, R., MASI, G., POGGI, G., VERDOLIVA, L. and SCARPA, G. 2015. Marker-controlled watershed-based segmentation of multiresolution remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing,* 53**,** 2987-3004.

GARCIA, N. and STOLL, E. 1984. Monte Carlo calculation for electromagnetic-wave scattering from random rough surfaces. *Physical Review Letters,* 52**,** 1798.

HANSEN, C. and THYREGOD, P. 1998. Use of wafer maps in integrated circuit manufacturing. *Microelectronics Reliability,* 38**,** 1155-1164.

HE, L., REN, X., GAO, Q., ZHAO, X., YAO, B. and CHAO, Y. 2017. The connected-component labeling problem: a review of state-of-the-art algorithms. *Pattern Recognition,* 70**,** 25-43.

HOCHBAUM, D. S. 2011. An efficient and effective tool for image segmentation, total variations and regularization. *International Conference on Scale Space and Variational Methods in Computer Vision.* Springer.

JANSSENS, O., SCHULZ, R., SLAVKOVIKJ, V., STOCKMAN, K., LOCCUFIER, M., VAN DE WALLE, R. and VAN HOECKE, S. 2015. Thermal image based fault diagnosis for rotating machinery. *Infrared Physics & Technology,* 73**,** 78-87.

JEONG, Y. S., KIM, S. J. and JEONG, M. K. 2008. Automatic identification of defect patterns in semiconductor wafer maps using spatial correlogram and dynamic time warping. *IEEE Transactions on Semiconductor Manufacturing,* 21**,** 625-637.

JIANG, X., SCOTT, P. J., WHITEHOUSE, D. and BLUNT, L. 2007. Paradigm shifts in surface metrology. Part II. The current shift. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences,* 463**,** 2071-2099.

JIANG, X. J. and WHITEHOUSE, D. J. 2012. Technological shifts in surface metrology. *CIRP Annals-Manufacturing Technology,* 61**,** 815-836.

JIN, R. and LIU, K. 2013. Multimode variation modeling and process monitoring for serial-parallel multistage manufacturing processes. *Iie Transactions,* 45**,** 617-629.

JOLIC, K., NAGARAJAH, C. and THOMPSON, W. 1994. Non-contact, optically based measurement of surface roughness of ceramics. *Measurement Science and Technology,* 5**,** 671.

JOMAA, W., SONGMENE, V. and BOCHER, P. 2014. Surface finish and residual stresses induced by orthogonal dry machining of AA7075-T651. *Materials,* 7**,** 1603-1624.

KANAFI, M. M. and TUONONEN, A. J. 2017. Top topography surface roughness power spectrum for pavement friction evaluation. *Tribology International,* 107**,** 240-249.

KIM, S., KIM, J., JEONG, M.-K., AL-KHALIFA, K., HAMOUDA, A. and ELSAYED, E. 2019. Monitoring and control of beta-distributed multistage production processes. *Quality Technology & Quantitative Management,* 16**,** 1-18.

KORKUT, I. and DONERTAS, M. 2007. The influence of feed rate and cutting speed on the cutting forces, surface roughness and tool–chip contact length during face milling. *Materials & Design,* 28**,** 308-312.

KROLCZYK, G., KROLCZYK, J., MARUDA, R., LEGUTKO, S. and TOMASZEWSKI, M. 2016. Metrological changes in surface morphology of high-strength steels in manufacturing processes. *Measurement,* 88**,** 176-185.

LIANG, X., LIU, Z., YAO, G., WANG, B. and REN, X. 2019. Investigation of surface topography and its deterioration resulting from tool wear evolution when dry turning of titanium alloy Ti-6Al-4V. *Tribology International,* 135**,** 130-142.

LIAO, P.-S., CHEN, T.-S. and CHUNG, P.-C. 2001. A fast algorithm for multilevel thresholding. *Journal of Information Science and Engineering,* 17**,** 713-727.

MONTGOMERY, D. C. 2007. *Introduction to statistical quality control*, John Wiley & Sons.

MORAN, P. A. 1950. Notes on continuous stochastic phenomena. *Biometrika,* 37**,** 17-23.

OTSU, N. 1979. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics,* 9**,** 62-66.

PALANI, S. and NATARAJAN, U. 2011. Prediction of surface roughness in CNC end milling by machine vision system using artificial neural network based on 2D Fourier transform. *The International Journal of Advanced Manufacturing Technology,* 54**,** 1033-1042.

PATIL, D. D. and DEORE, S. G. 2013. Medical image segmentation: a review. *International Journal of Computer Science and Mobile Computing,* 2**,** 22-27.

PAYNABAR, K., ZOU, C. and QIU, P. 2016. A change-point approach for phase-I analysis in multivariate profile monitoring and diagnosis. *Technometrics,* 58**,** 191-204.

QIU, P. 2013. *Introduction to statistical process control*, Chapman and Hall/CRC.

RAMESH, S., KARUNAMOORTHY, L. and PALANIKUMAR, K. 2008. Surface roughness analysis in machining of titanium alloy. *Materials and Manufacturing Processes,* 23**,** 174-181.

RAMSAY, J. and SILVERMAN, B. W. 2005. *Functional data analysis,* New York, Springer.

RAO, P., BUKKAPATNAM, S., KONG, Z., BEYCA, O., CASE, K. and KOMANDURI, R. 2015a. Quantification of ultraprecision surface morphology using an algebraic graph theoretic approach. *Procedia Manufacturing,* 1**,** 12-26.

RAO, P. K., BEYCA, O. F., KONG, Z., BUKKAPATNAM, S. T., CASE, K. E. and KOMANDURI, R. 2015b. A graph-theoretic approach for quantification of surface morphology variation and its application to chemical mechanical planarization process. *IIE Transactions,* 47**,** 1088-1111.

RENGIER, F., MEHNDIRATTA, A., VON TENGG-KOBLIGK, H., ZECHMANN, C. M., UNTERHINNINGHOFEN, R., KAUCZOR, H.-U. and GIESEL, F. L. 2010. 3D printing based on imaging data: review of medical applications. *International Journal of Computer Assisted Radiology and Surgery,* 5**,** 335-341.

SANGAHN, K. 2019. Variable selection-based SPC procedures for high-dimensional multistage processes. *Journal of Systems Engineering and Electronics,* 30**,** 144-153.

SCHMÄHLING, J. 2006. *Statistical characterization of technical surface microstructure.* University of Heidelberg.

SCHOLES, S. and UNSWORTH, A. 2009. Wear studies on the likely performance of CFR-PEEK/CoCrMo for use as artificial joint bearing materials. *Journal of Materials Science: Materials in Medicine,* 20**,** 163.

SILVELA, J. and PORTILLO, J. 2001. Breadth-first search and its application to image processing problems. *IEEE Transactions on Image Processing,* 10**,** 1194-1199.

SÖDERFJÄLL, M., HERBST, H. M., LARSSON, R. and ALMQVIST, A. 2017. Influence on friction from piston ring design, cylinder liner roughness and lubricant properties. *Tribology International,* 116**,** 272-284.

ŠT, O. and BENEŠ, B. 2011. *Connected component labeling in CUDA*, Elsevier.

TAAM, W. and HAMADA, M. 1993. Detecting spatial effects from factorial experiments: An application from integrated-circuit manufacturing. *Technometrics,* 35**,** 149-160.

TOOTOONI, M. S., LIU, C., ROBERSON, D., DONOVAN, R., RAO, P. K., KONG, Z. J. and BUKKAPATNAM, S. T. 2016. Online non-contact surface finish measurement in machining using graph theory-based image analysis. *Journal of Manufacturing Systems,* 41**,** 266-276.

TOWNSEND, A., SENIN, N., BLUNT, L., LEACH, R. and TAYLOR, J. 2016. Surface texture metrology for metal additive manufacturing: a review. *Precision Engineering,* 46**,** 34-47.

TURKOZ, M. 2018. *Distribution-free fault identification and anomaly detection in high-dimensional data.* Rutgers University-School of Graduate Studies.

VALA, M. H. J. and BAXI, A. 2013. A review on Otsu image segmentation algorithm. *International Journal of Advanced Research in Computer Engineering & Technology,* 2**,** 387-389.

WANG, Y. H. 2010. Tutorial: image segmentation. National Taiwan University, Taipei.

WYCISK, E., SOLBACH, A., SIDDIQUE, S., HERZOG, D., WALTHER, F. and EMMELMANN, C. 2014. Effects of defects in laser additive manufactured Ti-6Al-4V on fatigue properties. *Physics Procedia,* 56**,** 371-378.

XIAO, S., XIE, F., BLUNT, L., SCOTT, P. and JIANG, X. 2006. Feature extraction for structured surface based on surface networks and edge detection. *Materials Science in Semiconductor Processing,* 9**,** 210-214.

YANG, S., LI, W. and CHEN, H. 2018. *Surface finishing theory and new technology,* New York, Springer.

YANG, W., ZHANG, X., LIANG, C. and LIAO, Q. 2011. Evaluating surface roughness of castings using K-means clustering and watershed transform. *International Conference on Multimedia Technology.* IEEE.

YU, G., ZOU, C. and WANG, Z. 2012. Outlier detection in functional observations with applications to profile monitoring. *Technometrics,* 54**,** 308-318.

YU, J. and QIN, S. J. 2008. Multimode process monitoring with Bayesian inference-based finite Gaussian mixture models. *AIChE Journal,* 54**,** 1811-1829.

ZHANG, L., WANG, K. and CHEN, N. 2016. Monitoring wafers' geometric quality using an additive Gaussian process model. *IIE Transactions,* 48**,** 1-15.

ZHANG, P., CUI, T. and LI, Q. 2017. Effect of surface roughness on thermal contact resistance of aluminium alloy. *Applied Thermal Engineering,* 121**,** 992-998.

ZHANG, S., ZHOU, K., DING, H., GUO, J., LIU, Q. and WANG, W. 2018. Effects of grinding passes and direction on material removal behaviours in the rail grinding process. *Materials,* 11**,** 2293.