Statistical modeling and monitoring of image data in the presence of temporal and spatial correlations

Dariush Eslami[a], Hamidreza Izadbakhsh[a,*], Orod Ahmadi[a] and Marzieh Zarinbal[b]

[a]*Department of Industrial Engineering, Faculty of Engineering, Kharazmi University, Tehran, Iran*
[b]*Iranian Research Institute for Information Science and Technology (IranDoc), Tehran, Iran*

ABSTRACT

Owing to the increasing sensitivity of the processes and the inadequacy of the methods based on human inspection, use of product images in statistical process control has been considered by some researchers. In this paper, the regression-based approach was developed to monitor image data under two-scale analysis, large-scale and small-scale. Because geometric profiles created from images have complex and non-linear behaviour, wavelet transformation was used to extract the main features (regression coefficients) under large-scale analysis. Observations in mentioned profiles could be temporally or spatially correlated as well. To monitor the small-scale components which could be expressed by correlation in error terms, one parametric and one non-parametric methods were developed. After extracting features for both scales, some appropriate test statistics were computed. Then, monitoring the process was performed by plotting these test statistics on corresponding control charts. Performance of the proposed method was evaluated in terms of run-length and change-point measures. Simulation and industrial case studies were also performed to evaluate the proposed method's performance in detecting different shifts. The results indicated the proper performance of the proposed method in monitoring industrial processes to detect out-of-control conditions and identify the source of variability.

1. Introduction

Recently, use of non-contact image-based data-collection techniques has received increasing attention in a statistical framework to evaluate process stability and capability to produce high quality products with low variation. In just a few seconds, each image can provide a huge amount of information at a relatively low cost. Consequently, image data plays a significant role in industrial production. In this framework, big data streams and new data formats present a challenge to practitioners. These data structures require the use of analytical methods to characterize and synthesize the relevant information content. The content of an image stream needs to be summarized into a format that makes statistical monitoring easier to implement. To this aim, Wang and Tsung [1] presented a seminal method to monitor image data through profile monitoring technique. The main idea of Wang and Tsung [1], later extended by Wells et al. [2] to high-density point clouds, consists in translating large image data into linear profiles through the use of Q-Q plots. Then, traditional profile monitoring is applied to check the stability of Q-Q plot parameters. Similar to the work of Wang and Tsung [1], Menafoglio et al. [3] presented a profile monitoring procedure to determine the stability of random features in image data over time. However, they attempted to generalize the seminal idea suggested by Wang and Tsung [1] through monitoring functional shape of probability density function rather than Q-Q plot parameters.

Generally, the use of profile monitoring approach to monitor image data has increased in industrial production for both in-line process monitoring and post-process quality inspections. However, there are not enough researches in which the processes are studied after out-of-control conditions are detected [4]. When a process is identified as out-of-control, it is necessary to bring the process to in-control condition. To achieve this purpose, Megahed et al. [5] applied profile monitoring based on image data to monitor the process. Their method can provide valuable diagnostic information for users and artisans to detect shifts and bring the process back to in-control conditions. In their method, animated masks with different sizes move on the image, and the average values of intensity of pixels inside each window are recorded as a feature in a vector. Thus, image information summarizes in a vector of mean values called profile, which is monitored using a control chart over time. The main idea of Megahed et al. [5], later extended by some

researchers. He et al. [6] developed the above-mentioned method for detecting multiple defects simultaneously. Koosha, Noorossana and Megahed [7] proposed a novel wavelet-based approach to monitor statistically gray-scale image data. In their approach, each image is transformed to a number of 1-dimensional (1D) profiles. Wavelet transformation is then applied to extract the main features of the profiles. Then, a generalized likelihood ratio (GLR) chart is used to monitor extracted features (wavelet coefficients) over time. Menafoglio et al. [3] presented a profile monitoring method which is based on a simplistical variant of the functional principal components analysis (FPCA). FPCA allows the preservation of the limited nature of the data due to Bayes space geometry. The proposed method summarizes the random occurrences of the observed features by using their probability density functions (PDFs). Zuo, He and Zhang [8] presented a image monitoring method which combines the region growing algorithm with the EWMA-based statistic. In their method, quality characteristic is either characterized by uniformity or a specific pattern. Huang et al. [9] proposed a statistical monitoring method based on the surface image data of fused deposition modeling (FDM) parts with consistent layers and a specified period. The probability of alarm in a specified period (PASP) and the cumulative PASP are presented to compute the control limits and investigate the monitoring performance. A novel method for spatiotemporal monitoring processes with image data has been presented by Amirkhani and Amiri [10]. In their method, a P-value–based control chart and Dunnett's test are applied to detect the out-of-control condition of process. For this task, the intensity matrix of the nominal and captured images is divided into specific regions. Next, the equality of the means of the regions is examined to identify region(s) with remarkable differences. Eslami et al. [11] developed a statistical method based on multi-channel profiles to monitor RGB color image data. In their method, a Multi-channel Functional Principal Component Analysis (MFPCA) is utilized to achieve a number of extracted features that can effectively express process variations. Then, an Exponentially Weighted Moving Average (EWMA) control chart is constructed based on these features to investigate the ability of the proposed method in detecting shifts and change-point.

To the best of our knowledge, none of the proposed image-based monitoring techniques enables one to recognize the source of assignable causes when out-of-control conditions are detected, and take advantage of them for the purpose of SPC. Here, we propose a novel method to cover this item that can accelerates the recovery process. To do this, it is first necessary to study the structure of the profile created from the image.

The nature of the profile created from the image is similar to geometric profiles which express the functional relationship between the quality characteristic (intensity) and the physical location of the product's points (pixels) [7, 5]. Geometric profiles have several properties that must be taken into account for modeling the profile created from the images. Important properties of geometric profiles include large amount of data, complex and non-linear behavior, and spatial/temporal correlation [12].

Due to the first two properties, it is necessary to use dimension reduction techniques such as non-linear regression and non-parametric regression to monitor these types of profiles. Non-linear regression method loses its performance as the number of parameters increases in complex processes. This problem becomes more pronounced when there are sudden jumps in profile [13, 14]. Therefore, non-parametric regression methods are more streamlined because of their ability in approximating complicated functions. Many of non-parametric regression methods have been suggested for profile monitoring, some of these include the use of spline estimators, component analysis, wavelet transformation, and simple data-driven metrics [15, 16, 17].

On the other hand, in geometric profiles, measurements are taken at adjacent points in time and close locations. The results of one measurement may be related to the results of a previous measurement if the measurements within a profile are taken at very short intervals. This dependence over time is often referred to as temporal correlation. The measurements within a profile may also be autocorrelated due to the fact that they are on the same physical space. This correlation due to physical distance is often referred to as spatial correlation [16, 18, 19, 14].

According to the properties of geometric profiles created from product's images, two sources of variability in the quality characteristic of interest can be considered in practice. One source is a result of variations in manufacturing process such as part-to-part variation, fixture or tooling tolerance, and/or process operation condition variations. The second one is mostly due to measurement errors and environmental disturbances. In the literature of geometric profiles, monitoring of variations in manufacturing process is performed by considering regression coefficients as large-scale components. Also, monitoring of variations in measurement and machinery conditions is done by considering small-scale components expressed by correlation in error terms [20, 21]. Hence, under out-of-control conditions, identifying the scale in which the change occurred is important for variation reduction and corrective actions. It can also help users to make suitable decisions and quickly bring the process back to in-control condition.

Briefly, the motivations for the present study can be summarized as follows:

- Accelerating the recovery process by identifying sources of variability in the quality characteristic of interest

- Using non-parametric regression methods in approximating complicated profiles in large-scale monitoring

- Modeling temporal and spatial correlations in error terms

- Presenting a novel non-parametric method with no assumption about the structure of correlation in error terms in small-scale monitoring

The main aim of this paper is to present a regression-based approach under two-scale analysis for monitoring processes that produce image data. Images are considered as geometric profiles in which measurements of the same characteristic, i.e. light intensity, are taken at adjacent points. Under large-scale analysis, wavelet transformation is applied to extract the main features of images which are called large-scale components. These features should be studied over time. For this task, extracted components are monitored using GLR chart. The following step is to monitor the small-scale components which can be expressed as correlations in error terms. Observations could be temporally or spatially correlated. To monitor correlation in error terms, one parametric and one non-parametric methods are developed in this study. In the parametric method, correlation in error terms is modelled by first-order spatial correlation model. Estimation of this model's parameters, including spatial correlation coefficient and error term variance, is carried out with Ordinary Least Squares (OLS) and Generalized Least Squares (GLS) methods, respectively. Based on these estimates, small-scale components are monitored using OMNIBUS control chart. In the non-parametric method, no assumption is made about the structure of correlation in error terms. To extract useful information about the nature of correlation in error terms, FPCA is used in this step. Using FPCA one could reduce data dimension while keeping useful information. Extracted features can be effectively used to characterize process variations. After extracting these features, EWMA chart is constructed based on principal components (PCs) scores. Simulation and industrial case studies are performed to investigate the proposed method's performance to detect different shifts.

The remainder of this paper is organized as follows: In section 2, basic principles of wavelet transformation, FPCA and spatial correlation are presented, respectively. In section 3, the proposed method is discussed. Section 4 has been devoted to simulation studies that have been conducted to evaluate the performance of the proposed method. In section 5, tile manufacturing industry is considered as a real case study to analyze the proposed method. Finally, section 6 concludes the paper with several remarks.

2. Basic principles

2.1. Wavelet transformation

As mentioned in section 1, as manufacturing products become more complex, their profiles could be complicated, consisting of a lot of sharp changes with many abrupt jumps. In these cases, non-parametric regression could be utilized as a effective method to model the process and variation explanation [7]. In non-parametric regression, the objective is to obtain a function with the fewest coefficients while retaining a good fit to the observations [14]. In this field, wavelet transformation is an important and widespread method.

Wavelets are some basic functions that have great ability to approximate highly complex functions with irregular fluctuations. Another advantage of wavelet transformation is the ability to simultaneously analyze data in the spatial and frequency domains. This feature greatly helps to detect the time and location of changes accurately [14].

Wavelets are basic functions that are applied to set up an orthonormal basis. Using these basic functions, squared integrable functions could be expressed. Each orthonormal basis is created by father and mother wavelets which are shown by $\phi(x)$ and $\psi(x)$, respectively. Several alternatives are available for these two functions such as "Haar", "Daubechies", and "Mexican hat". Based on the case under consideration, these functions may differ in their properties and performance [22]. In the proposed method, the basis function named "Haar" is used due to computational simplicity.

The $\left\{ \phi_{j_o k}(x), \psi_{jk}(x); j \geq j_o; k \in Z \right\}$ for $j_o \geq 0$ form an orthonormal basis where $\phi_{jk}(x) = 2^{j/2}\phi(2^j x - k)$ and $\psi_{jk}(x) = 2^{j/2}\psi(2^j x - k)$ are the translation and dilation of ϕ and ψ, respectively [22]. Each squared integrable $f(x)$ can
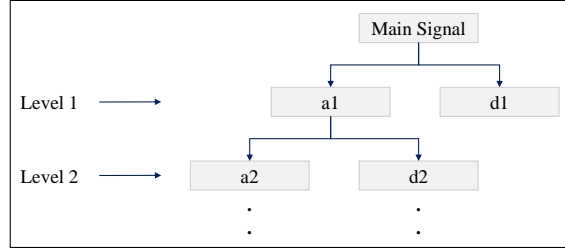
Figure 1: The signal decomposition by wavelet [23].

be represented by Equation (1) in terms of the father and mother functions [23].

$$f(x) = \sum_{k=0}^{2^{j_o}-1} c_{j_o,k}\, \phi_{j_o,k}(x) \;+\; \sum_{j=j_o}^{J-1} \sum_{k=0}^{2^{j_o}-1} d_{jk}\, \psi_{jk}(x) \tag{1}$$

In Equation (1), $c_{j_o,k}$ is the inner product of the $f(x)$ and the father function, $\left\langle f, \phi_{j_o,k} \right\rangle$, and d_{jk} is the inner product of the $f(x)$ and the mother function, $\left\langle f, \psi_{j_o,k} \right\rangle$. These coefficients are named wavelet coefficients. The parameter j demonstrates the level of decomposition.

The analysis of functions in both frequency and spatial domains is one of the beneficial properties of wavelet transformation. For this task, the function decomposes into some basic functions which are generated by father and mother wavelets. The decomposition of function is done by passing it into a low-pass and a high-pass filters. Approximation coefficients, $c_{j_o,k}$, are obtained by passing the function into the low-pass filter. Then, these coefficients contain the low-frequency components of a function or main signal. Passing the function into high-pass filter results in detail coefficients, d_{jk}. These coefficients illustrate high-frequency components of the function [14]. Figure 1 indicates the decomposition scheme based on wavelets.

From Figure 1 it is found that, the signal is divided into an approximate part ($a1$) and a detail part ($d1$) for the first level of decomposition. For the next level of decomposition, $a1$ is divided into a new approximate part ($a2$) and a new detail part ($d2$). This process was continued to a desired level of decomposition. At this level, the approximation part contains the main features of the main signal. The approximate part will resemble the main signal and has a large number of coefficients at lower decomposition levels. Therefore, there is a direct relationship between the number of coefficients and the similarity of the approximate part with the main signal. The choice of the decomposition level depends on two factors: computational time and the least error area to be discovered. The computational time must be equal to or less than the process cycle time so that the monitoring process is implemented with 100% online inspection. The decomposition level should be small enough to make the algorithm more sensitive to detect smaller shifts. On the other hand, the number of coefficients and computational time will decrease at higher decomposition levels. Therefore, the decomposition level must be chosen in accordance with the limits of accuracy and computational time [14].

2.2. Functional principal components analysis

Principal component analysis (PCA) is a statistical method for investigating the dominant modes of variation in multivariate data. PCA finds directions in the observations space along which the data have the greatest variability using an eigenvalue decomposition of the covariance matrix. The analysis generates a loading vector or weight vector for every principal component, which indicates the direction of variability associated with that component.

Based on the functional context, each principal component is characterized by a principal component weight function or eigenfunction $u(t)$ defined on the same range of t as the functional data.

According to Deregowski and Krzysko [24], suppose a sample of the process $X(t) \in L_2([0,T])$, where $L_2([0,T])$ is the Hilbert space of square integrable functions on the interval $[0,T]$, equipped with the scalar product.

$$\langle u, v \rangle = \int u(s)v(a)\, ds \tag{2}$$

Remark. All integrals are taken over the interval $[0,T]$.

Moreover, suppose that $E X(t) = 0$ and assume the existence of the variance $Var_X(t) = E[X(t)^2]$ and covariance

$Cov_X(s,t) = E[X(s)X(t)]$ functions of $X(t)$; $s,t \in [0,T]$.

In FPCA, we want to find orthonormal weight functions $u_1, u_2, \ldots$, such that the variance of a linear combination (U_r) is maximal.

The weight functions satisfy:

$$\|u_r\|^2 = \int u_r^2(t)\,dt = 1,$$

$$\langle u_r, u_{\acute{r}} \rangle = \int u_r(t)u_{\acute{r}}(t)dt = 0, \; r \neq \acute{r}.$$

The linear combination is:

$$U_r = \langle u_r, X \rangle = \int u_r(t)X(t)\,dt, \tag{3}$$

and the desired weight functions solve:

$$\arg\max_{\langle u_r, u_{\acute{r}} \rangle = \delta_{\acute{r}r}, \acute{r} \leq r} = Var\langle u_r, X \rangle, \tag{4}$$

or equivalently:

$$\arg\max_{\langle u_r, u_{\acute{r}} \rangle = \delta_{\acute{r}r}, \acute{r} \leq r} = \iint u_r(s)Cov_X(s,t)u_r(t)\,ds\,dt, \tag{5}$$

where $\delta_{\acute{r}r}$ is the Kronecker delta.

The solution is obtained by solving the Fredholm functional eigenequation

$$\int Cov_X(s,t)u(t)\,dt = \lambda\,u(s). \tag{6}$$

The eigenfunctions $u_1(s), u_2(s), \ldots$, sorted with respect to the corresponding eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq 0$ solve the FPCA problem in Equation (4). The following relationship holds between eigenvalues and eigenfunctions:
$\lambda_r = Var(U_r) = Var[\int u_r(t)X(t)\,dt] = \iint u_r(s)Cov_X(s,t)u_r(t)\,ds\,dt$

In practice, the covariance function $Cov_X(s,t)$, is unknown and should be estimated from the functional data set. For the functional sample, the estimator of the covariance function $Cov_X(s,t)$, has the form:

$$\hat{Cov}_X(s,t) = \frac{1}{N}\sum_{i=1}^{N}X_i(s)X_i(t) \tag{7}$$

By considering u_a as the ath principal component (PC), the score of the ath PC of functional data $i = 1, 2, ..., N$ is

$$S_{i,a} = \langle X_i(t), u_a(t) \rangle = \int_0^1 X_i(t)\,u_a(t)\,dt \qquad a = 1, 2, ..., d \tag{8}$$

The score vector for i^{th} functional data is $S_i = [S_{i,1}, S_{i,2}, ..., S_{i,d}]'$, where d is the number of PCs selected to provide a certain amount of the variation.

2.3. Correlation in data

In reality situations, oftentimes, the measurements are correlated to one another because of a wide variety of reasons. For instance, the results of one measurement may be related to the results of a previous measurement if the measurements within a profile are taken at very short intervals. This dependence over time is often referred to as temporal correlation. Note that when we refer to correlation, we are referring to the correlation in the errors or data values themselves. In other words, the correlation in the parameter estimators that are used to represent the profile with fewer values than the original data itself are not referring. To illustrate one way in which correlation can exist in a profile, consider the following model for a profile [17].

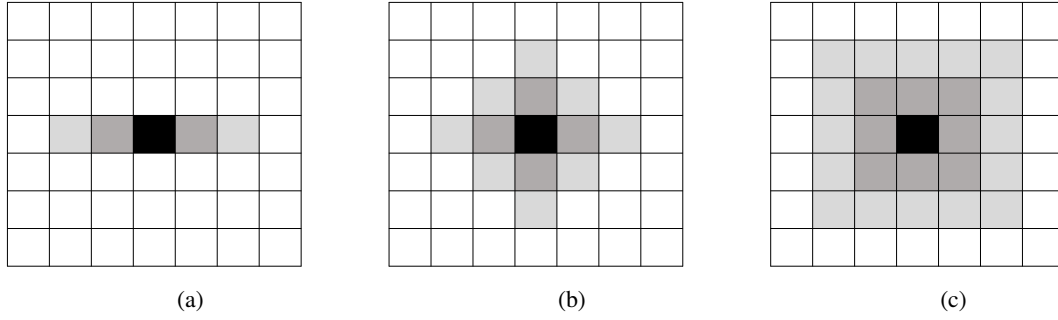$$y_i = \beta_o \mathbf{1} + \beta_1 x_i + \varepsilon_i, \tag{9}$$

Figure 2: The grid view of spatial patterns: (a) Linear, (b) Rook, (c) Queen

where y_i and x_i are $n \times 1$ vectors contain the responses and predictor values, respectively. If there is no correlation within measurements in a profile (or it is negligible), then it is assumed that $\varepsilon_i \sim N(0, \sigma^2 I)$, i.e., the error terms are independent and identically distributed (i.i.d.) observations from a Gaussian distribution. On the other hand, if we want to allow for correlation, then we can relax the limitation on the error terms to characterize the correlation. In fact, we allow $\varepsilon_i \sim N(0, R)$, where R is some positive definite covariance matrix:

$$
R = \sigma^2
\begin{bmatrix}
1 & \rho & \rho^2 & . & . & . & \rho^n \\
\rho & 1 & & & & & . \\
\rho^2 & & . & & & & . \\
. & & & . & & & . \\
. & & & & . & & \rho^n \\
. & & & & & 1 & \rho \\
\rho^n & . & . & . & \rho^2 & \rho & 1
\end{bmatrix}
$$

In this case, error terms can be considered as a stochastic process and off-diagonal elements of the covariance matrix that are not zero, allowing us to characterize the correlation within measurements in a profile [25, 26].

The measurements within a profile may also be autocorrelated due to the fact that they are on the same physical space. Thus, besides time-based correlation, there is a location-based correlation in measurements within a profile called spatial correlation [12]. Spatial correlation implies how one point affects another. This effect is caused by the adjacency of two points (having a common border) or by the effect of the other point being within a certain distance (without having a common border).

The spatial contiguity matrix is used to model spatial correlation. The spatial contiguity matrix $\mathbf{W} = \left[w_{p,q} \right]$; $p, q = 1, 2, ..., N$ is $N \times N$ matrix, where $w_{p,q}$ is one if p is adjacent to q and otherwise it is zero.

Spatial correlation models are defined based on different patterns which include linear, rook, and queen. In these models, the two spatial units are adjacent if they have a common border and/ or vertex. In the linear model $w_{p,q}$ is equal to one if $q = p + 1$ or $q = p - 1$. In the rook model, two spatial units are adjacent if they have a common border, and in the queen model, two spatial units are adjacent if they have a common border or vertex [27]. Figure 2 shows the grid view of these patterns.

In the following, the spatial contiguity matrix representations of the above patterns are given. Spatial contiguity matrices for the linear, rook, and queen models for 9 spatial units in a grid structure (3×3) are as follows:

$$
\mathbf{W}_{Linear} =
\begin{bmatrix}
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0
\end{bmatrix}
\quad
\mathbf{W}_{Rook} =
\begin{bmatrix}
0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0
\end{bmatrix}
\quad
\mathbf{W}_{Queen} =
\begin{bmatrix}
0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 \\
0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 1 \\
0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0
\end{bmatrix}
\tag{10}
$$

$$
\text{(a)} \qquad\qquad \text{(b)} \qquad\qquad \text{(c)}
$$

Matrices (*a*) to (*c*) in Equation (10) are the first-order spatial contiguity matrices. The z-order spatial contiguity matrix ($\mathbf{W}^z$) is the exponentiation of the spatial contiguity matrix [27].
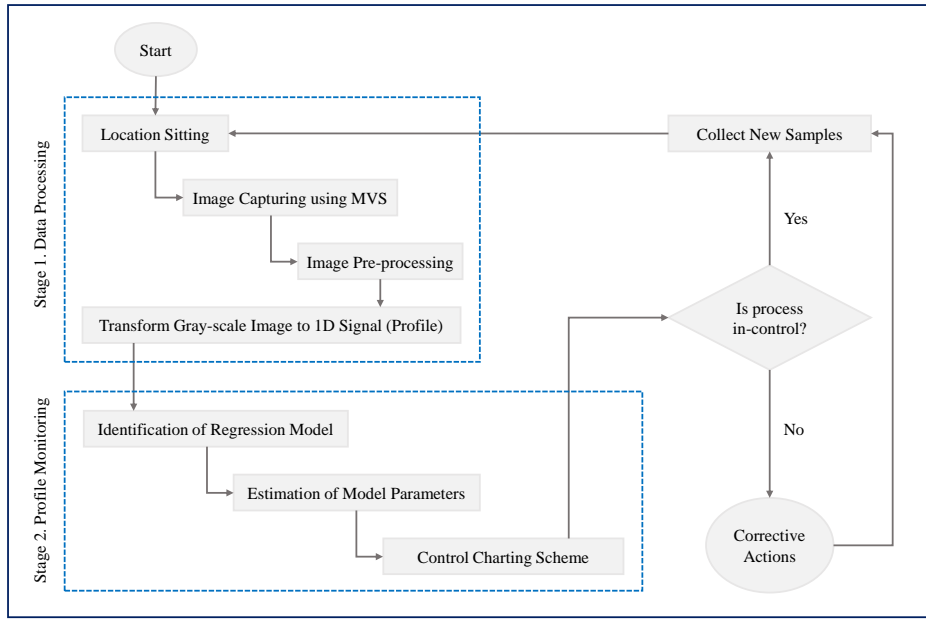
Figure 3: The proposed procedure.

3. The proposed method

As mentioned in section 1, it is necessary to study the structure of the profile created from the image to accelerate the recovery process. The nature of the profile created from the image is similar to geometric profiles in which measurements of the same characteristic, i.e., light intensity, are taken at adjacent points. Geometric profiles have properties such as a large volume of observations, complex behavior and spatial/temporal correlation that must be taken into account when modeling these profiles. After modeling, components of two scales should be monitored to identify sources of variability in the quality characteristic of interest. This task can help users to make proper decisions and take corrective actions. In this study, the proposed method is performed in two stages: data processing and profile monitoring. In the first stage, after adjusting the product to its proper location, the image capturing is done, and the pre-processing steps are performed using a machine vision system (MVS). Then, the final image of the product is converted into a two-dimensional profile. In the profile monitoring stage, the model of the process is determined based on the desired properties. After that, two-scale components of the model should be estimated. Next, process monitoring is performed through a control charting scheme.

Briefly, the assumptions for the present study can be summarized as follows:

- The profile created from the image is similar to geometric profiles which express the functional relationship between the quality characteristic (intensity) and the physical location of the product's points (pixels).

- The profiles created from image data have complex and non-linear behavior, so non-parametric regression technique (wavelet transformation) is applied to monitor these types of profiles in large-scale monitoring.

- Since measurements are taken at adjacent points in time and close locations, temporal and spatial correlations in error terms are considered in small-scale modeling and monitoring.

- Monitoring of image data is done under two scales, small-scale and large-scale. Under out-of-control conditions, identifying the scale in which the change occurred is important for variation reduction and corrective actions.

The flowchart of the proposed procedure is shown in Figure 3.
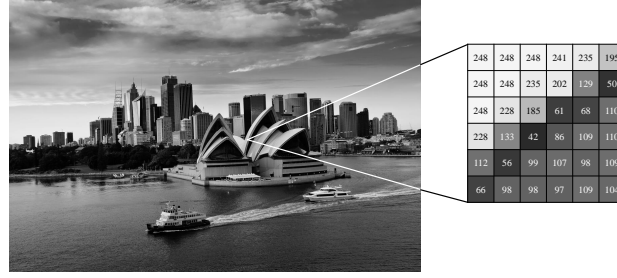The stages of the proposed method are described as follows:

Figure 4: A gray-scale image and the pixel values in a 6×6 neighborhood [29].
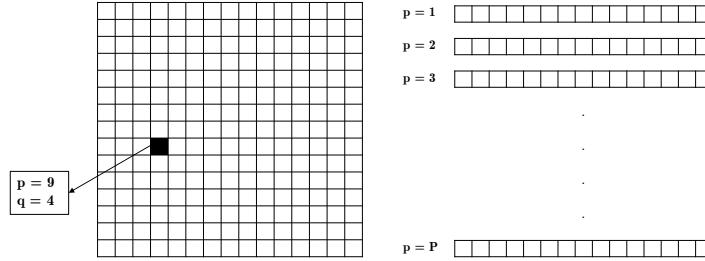

Figure 5: The matrix of image [7].

3.1. Data processing

At this stage, the data are prepared for process monitoring. The input data is in the form of an image that is available after processing as a profile. For data preparation, the following steps are considered:

3.1.1. Image acquisition

The first step of any vision system is the image acquisition step. In this step, the digital image is captured by one or several image sensors such as range sensors, tomography devices, ultra-sonic cameras, etc. In this study, it is assumed that the product images of an industrial process are taken by an MVS as successive and online.

3.1.2. Image pre-processing

This step consists of several pre-processing steps that performing some of them may be useless depending on new MVS devices and industrial cases. For more information on image pre-processing, see [28]. According to the capturing device and the process in this study, image pre-processing consists of converting color image to gray-scale image, removing unnecessary parts from the image, contrast improvement, and removing image background.

3.1.3. Transforming gray-scale image to 1D signal (Profile)

The digital image represents a two-dimensional image using a finite number of points, usually referred to as pixels. One or more numerical values represent each pixel. In gray-scale images, a single value represents the intensity of the pixel in the range 0 to 255, as shown in Figure 4.

The zigzag approach, sweep line algorithm, and linear approach are some of the methods used to convert a 2D matrix to 1D vectors [30, 7]. As a result of its simplicity and short computational time, the latter is adopted in the proposed method [7]. The matrix shown in Figure 5 is related to an image that has been digitized. This matrix is a $P \times Q$ matrix with P rows and Q columns.

To convert the 2D matrix into a 1D vector (signal), the rows of the matrix can be separated and aligned to form a vector of size $PQ \times 1$. After preparing the signal, a diagram can be plotted based on the intensity and the index of pixels. For instance, Figure 6b shows a diagram obtained from the first row of the intensity matrix of a tile image (Figure 6a).
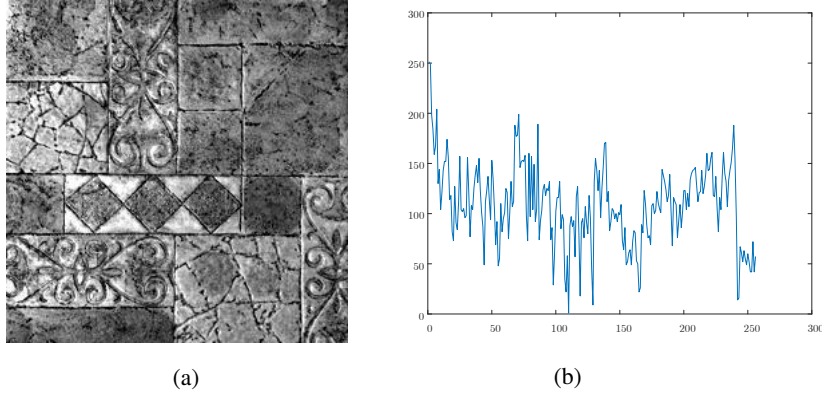
| (a) | (b) |

Figure 6: (a) The image of tile, (b) First row profile of intensity matrix [7].

The diagram obtained from Figure 6b can be considered as a geometric profile based on the concept stated in section 1. The general framework of geometrical profiles includes applications in which different measurements of a single quality characteristic such as thickness are taken at different locations.

3.2. Profile monitoring

The profiles obtained from the previous stage should be monitored to check for out-of-control conditions. Generally, profile monitoring consists of three main steps: identification of regression model, estimation of model parameters and designing control charts. In this study, source of assignable cause is also determined in addition to the above steps. Each of these steps is evaluated under two scales, large-scale and small-scale. Under large-scale evaluation, regression coefficients are extracted as main features of image and monitoring of variations in manufacturing process is performed. Under small-scale evaluation, monitoring of variations in measurements and machinery condition is done by considering components that is expressed through correlation in error terms.

3.2.1. Large-scale monitoring

This subsection describes the steps for large-scale monitoring:

Step 1: Identification of Regression Model. The first step in profile monitoring is to identify the process model. For this purpose, the appropriate model should be determined based on the desired properties. As discussed in section 1, geometric profiles (profiles created from images) have several properties that must be taken into account. Two of the most important properties of geometric profiles are high data volume, complex and non-linear behavior. A high volume of observations in geometric profiles makes the estimated variance-covariance matrix of observations singular which cannot be used in multivariate control charts. As a result, there is a need for dimension reduction techniques. One of these techniques is the regression approach whereby the estimators of regression coefficients can be considered as random variables for monitoring the profiles rather than dealing with large data. On the other hand, because of complex and non-linear behavior of observations in each profile, it is necessary to use techniques such as non-linear regression or non-parametric regression for modeling. As the complexity of the process intensifies, the number of parameters (regression coefficients) increases which, in turn, reduces the power of control charts to detect shifts. Therefore, to identify the model for the geometric profile, non-parametric regression methods are considered.

According to the above-mentioned two properties, the non-parametric regression approach is used to model the profile created by the image. Recently, the non-parametric regression approach using wavelet transformation has become more popular than other non-parametric methods for modeling profiles [15, 16]. Using wavelets profiles could be modelled efficiently without losing important features of the data. Wavelet coefficients represent a profile in the wavelet domain. Profile information is compressed into a few coefficients, which can be interpreted and extracted in terms of scale-frequencies and locations of features within the profile.

In this case, a general form of the profile can be considered as Equation (11):

$$\mathbf{y}_i = \mathbf{X}\,\boldsymbol{\beta}_i + \boldsymbol{v}_i \qquad i = 1, 2, ..., m \tag{11}$$

so that,

$$
\mathbf{X} = \begin{bmatrix}
\phi_{j_o,1}(x_1) & \phi_{j_o,2}(x_1) & . & . & . & \phi_{j_o,h}(x_1) \\
\phi_{j_o,1}(x_2) & \phi_{j_o,2}(x_2) & & & & \phi_{j_o,h}(x_2) \\
. & . & . & & & . \\
. & . & . & & & . \\
. & . & . & . & & . \\
\phi_{j_o,1}(x_n) & \phi_{j_o,2}(x_n) & . & . & . & \phi_{j_o,h}(x_n)
\end{bmatrix}
$$

where, x_j, $j = 1, 2, ..., n$ and $\phi_{j_o,k}(\cdot)$, $k = 1, 2, ..., h$ are independent variables and wavelet functions, respectively. In Equation (11), observation $Y_{i,j}$ indicates the value of quality characteristic in location j of profile i where $j = 1, 2, ..., n$; $i = 1, 2, ..., m$ and $\mathbf{y}_i$ is a $n \times 1$ vector includes observations for the ith profile. $\boldsymbol{\beta}_i$ is the $h \times 1$ vector of regression coefficients. The wavelet approximation coefficients are regarded as the regression coefficients. These coefficients are obtained through wavelet transformation by passing the profile into the low-pass filter. $\boldsymbol{v}_i$ is the $n \times 1$ vector of error terms. Further explanation of the covariance structure of $\boldsymbol{v}_i$ will be discussed in the next subsection.

In Equation (11), the deterministic term $\mathbf{X}\boldsymbol{\beta}_i$ represents a systematic pattern of the ith profile. In complicated profiles, functions such as Fourier, B-spline, or wavelet basis can be applied as a model to represent the columns of $\mathbf{X}$. When the process is in-control, it is assumed that $\boldsymbol{\beta}_i = \boldsymbol{\beta}_o$ where $\boldsymbol{\beta}_o$ is an in-control parameter which must be estimated using sample data.

After determination of regression model, the process must be monitored. There are two phases in process monitoring, namely phase I and phase II. In phase I, the control chart is designed based on a set of in-control sample data. For this purpose, the parameters of the process are estimated. Then control limits are obtained based on these estimated parameters. In phase II, the process is monitored by comparing the test statistics with the control limits obtained in phase I. In the following, these two phases are considered for the proposed method.

Step 2: Phase I - Estimation of Model Parameters. Parameters of the non-parametric regression model including regression coefficients ($\boldsymbol{\beta}_i$) must be estimated from phase I data set to construct a control chart. For this purpose, wavelet transformation is applied to estimate the regression coefficients (as discussed in Section 2.1). As a result, estimated coefficients $\hat{\boldsymbol{\beta}}_i = (\hat{\beta}_{i,1}, \hat{\beta}_{i,2}, ..., \hat{\beta}_{i,h})'$; $i = 1, 2, ..., m$ for the entire image space are obtained. The number of coefficients for each profile is obtained in terms of the decomposition level, h. In this study, $h = 4$ is considered, which is appropriate according to [7]. Also, it is supposed that the process is in-control in phase I. Thus, sample's mean vector of $\hat{\boldsymbol{\beta}}_i$; $i = 1, 2, ..., m$ may be used as estimators of $\boldsymbol{\beta}_o$.

Step 3: Phase II - Control Charting Scheme. In this step, several suitable test statistics are determined in terms of parameters estimated from phase II data set. Next, the process is monitored by plotting these test statistics on corresponding control charts. In this subsection, a multivariate control chart should be utilized to monitor the process. As discussed in Section 1, the performance of traditional multivariate control charts such as T^2-Hoteling or multivariate exponentially weighted moving average (MEWMA) decreases as the number of parameters increases. In monitoring image data, the number of main features (approximate wavelet coefficients) reaches from hundreds to thousands in accordance with the level of decomposition. Therefore, a control chart should be used to monitor a huge number of coefficients over time without dramatically decreasing its performance. GLR chart is one of these charts which is based on the following statistics [31]:

$$
R_{i,l} = \begin{cases}
\displaystyle\max_{0 \le \tau < i, k} \frac{(i-\tau)}{2\hat{\sigma}_k^2}(\hat{\beta}_{1,\tau,i}(k) - \beta_{o,k})^2 & i = 1, 2, ..., l \\
\displaystyle\max_{i-l \le \tau < i, k} \frac{(i-\tau)}{2\hat{\sigma}_k^2}(\hat{\beta}_{1,\tau,i}(k) - \beta_{o,k})^2 & i = l+1, l+2, ...
\end{cases}
\tag{12}
$$

where, i is the image index and τ is the potential change-point. $\hat{\beta}_{1,\tau,i}(k)$ is equal to $(i-\tau)^{-1} \sum_{t=\tau+1}^{i} \hat{\beta}_{t,k}$, where $\hat{\beta}_{t,k}$ is a kth component of vector $\hat{\boldsymbol{\beta}}$ for t^{th} sample. $\beta_{o,k}$ is a kth component of vector $\boldsymbol{\beta}_o$ for the nominal image (in-control image), and $\hat{\sigma}_k^2$ is equal to $S^2(\hat{\beta}_k)$, where $\hat{\beta}_k$ is k^{th} estimated profile coefficient from phase I data and $S^2(\cdot)$ is the sample variance operator. The parameter l is considered as a time window due to computational time restriction. That is, each image is compared only with its l predecessors. Therefore, this control chart approach can estimate change-point that is important diagnostic information for users. In this study, $l = 5$ is considered.

To summarize the above discussion, GLR control chart can give diagnostic information including out-of-control conditions and change-point estimation, simultaneously. This information help practitioners to retrieve the process to in-control condition.

3.2.2. Small-scale monitoring

As stated in subsection 2.3, sometimes the measurements are correlated with each other. This correlation can be caused by collecting the measurements within a profile at very small time\location intervals. The dependence over time is known as temporal correlation. Alternatively, the dependence thanks to physical distance is often known as spatial correlation. Note that the existence of temporal correlation is very obvious because of the simultaneous measurement of data through imaging. In this study, the consideration of spatial correlation (if exists) in the process and its effect on process monitoring has been discussed.

In this study, one parametric method and one non-parametric method are developed to monitor correlation in error terms. In the parametric method, correlation in error terms is modelled by first-order spatial correlation model. Alternatively, no assumption is made about the structure of correlation between error terms in the non-parametric method. In the following, the steps for parametric and non-parametric methods are described:

- **Parametric method**

    Step 1: Identification of regression model. Another important property of geometric profiles is correlation in measurements that should be considered in the model identification. In the parametric method, correlation in error terms is modelled by z-order spatial correlation model. As a result, the model in Equation (11) for a profile can be considered as Equation (13):

$$\mathbf{y}_i = \mathbf{X}\boldsymbol{\beta}_i + \boldsymbol{v}_i \qquad i = 1, 2, ..., m$$
$$\boldsymbol{v}_i = \sum_{z=1}^{\Omega} \rho_i^{(z)} \mathbf{W}^{(z)} \boldsymbol{v}_i + \varepsilon_i \tag{13}$$
$$\varepsilon_i \sim N(0, \sigma_\varepsilon^2 I)$$

where, $\boldsymbol{v}_i$ is the $n \times 1$ vector of error terms with z-order spatial correlation, $\mathbf{W}^{(z)}$ is the z-order spatial contiguity matrix, $\rho_i^{(z)}$ is the z-order correlation coefficient and ε_i is $N(0, \sigma_\varepsilon^2 I)$ in which σ_ε^2 is error terms variance. In this study, $z = 1$ is considered, which is appropriate according to Noorossana and Nikoo [12] and Anselin [32]. When the process is in-control, it is assumed that $\rho_i = \rho_o$; $i = 1, 2, ..., m$ where ρ_o is in-control parameters. Equation (13) is the simplest spatial regression model called spatial error model (SEM) [27]. The designed model can respond to the challenges of geometric profiles created from images.

In practical situations, it is necessary to test whether or not spatial correlation exists between error terms in a profile. In the spatial regression literature, three test statistics including Wald (WL), Likelihood Ratio (LR) and Lagrange Multiplier (LM) are presented to investigate the existence of a spatial correlation between error terms [27]. These test statistics are used to test the following hypothesis:

$$H_o : \quad \rho = 0$$
$$H_1 : \quad \rho \neq 0 \tag{14}$$

The WL test statistic is more commonly used to test existence of first-order spatial correlation and is calculated as software outputs. This test statistic could be computed as follows [27]:

$$WL = \hat{\rho}_{oMLE}^2 \left[t_2 + t_3 - \frac{1}{n} t_1^2 \right]$$

$$s.t.$$

$$t_1 = tr\left(\mathbf{W}.* \mathbf{B}^{-1} \right)$$
$$t_2 = tr\left[\left(\mathbf{W}\mathbf{B}^{-1} \right)^2 \right] \tag{15}$$
$$t_3 = tr\left[\left(\mathbf{W}\mathbf{B}^{-1} \right)' \left(\mathbf{W}\mathbf{B}^{-1} \right) \right]$$
$$\mathbf{B} = \left(\mathbf{I}_n - \hat{\rho}_{oMLE}\mathbf{W} \right)$$

where, $\hat{\rho}_{\circ MLE}$ is the maximum likelihood estimator of $\rho_\circ$ which is the first-order spatial correlation coefficient, $(.*)$ is the element-by-element matrix multiplication and tr stands for trace operator. It could be shown that under $H_\circ$, WL is distributed according to chi-squared distribution with one degree of freedom, χ_1^2. If WL > 6.635, $H_\circ$ is rejected and it is concluded that there is a spatial correlation between the error terms. It should be noted that this test must be done in phase I. In this phase, in-control sample data are used to check the existence of a spatial correlation between error terms.

Similar to large-scale monitoring, the process in this scale must be monitored through two phases, phase I and phase II. In the following, these two phases are considered for the proposed method.

Step 2: Phase I - Estimation of model parameters. Parameters of the spatial correlation model include spatial correlation coefficient (ρ_i), and error terms variance (σ_ϵ^2) that must be estimated from phase I data set to construct a control chart. The spatial correlation coefficients, (ρ_i; $i = 1, 2, ..., m$), and error term variance (σ_ϵ^2) are estimated using OLS and GLS estimators, respectively. The estimation procedure of these parameters has been described below:

The estimated error term variance is obtained through the Equations (16) and (17) [33]:

$$\mathbf{b}_{i_{EGLS}} = \left[\mathbf{X}_i^{*'} \mathbf{X}_i^* \right]^{-1} \mathbf{X}_i^{*'} \mathbf{y}_i^* \tag{16}$$

$$\hat{\sigma}_{i_{EGLS}}^2 = \frac{\left(\mathbf{y}_i^* - \mathbf{X}_i^* \mathbf{b}_{i_{EGLS}} \right)' \left(\mathbf{y}_i^* - \mathbf{X}_i^* \mathbf{b}_{i_{EGLS}} \right)}{n} \tag{17}$$

where, $\mathbf{X}_i^* = \left(\mathbf{I} - \hat{\rho}_i \mathbf{W} \right) \mathbf{X}$, $\mathbf{y}_i^* = \left(\mathbf{I} - \hat{\rho}_i \mathbf{W} \right) \mathbf{y}$, $\hat{\rho}_i$; $i = 1, 2, ..., m$ are the least squares estimators of the spatial correlation coefficient for the ith sample in phase I, $\hat{\sigma}_{i_{EGLS}}^2$; $i = 1, 2, ..., m$ are the estimated error term variance for the ith sample in phase I and n is the number of observations per samples. In this paper, it is supposed that the process is in-control in phase I. Thus, sample mean of $\hat{\rho}_i$; $i = 1, 2, ..., m$ and $\hat{\sigma}_{i_{EGLS}}^2$; $i = 1, 2, ..., m$ may be used as estimators of $\rho_\circ$ and σ_ϵ^2, respectively.

In this study, MATLAB spatial econometric toolbox [27] is used to estimate parameters of spatial correlation model include spatial correlation coefficient (ρ_i), and error term variance (σ_ϵ^2)[12].

Step 3: Phase II - Control Charting Scheme. In this step, several suitable test statistics are determined in terms of parameters estimated from phase II data set. Next, process monitoring is done by plotting these test statistics on corresponding control charts. In this subsection, OMNIBUS control chart proposed by Zhang, Zou and Wang [34] is used to monitor small-scale parameters. In OMNIBUS control chart, it is possible to monitor the mean and variance of a process in one control chart, simultaneously. Note that monitoring of mean and variance of a process simultaneously is simpler and more efficient than using two control charts to monitor these parameters separately [34]. OMNIBUS control chart presented by Zhang, Zou and Wang [34] is augmented to detect change-point in small-scale components of the model in Equation (18). The test statistic of the proposed control chart is defined as below:

$$ELR_{i,l} = \begin{cases} \displaystyle\min_{0 \leq \tau < i} \frac{1}{(i-\tau)}(u_i^2 + v_i - \ln(v_i)) & i = 1, 2, ..., l \\ \displaystyle\min_{i-l \leq \tau < i} \frac{1}{(i-\tau)}(u_i^2 + v_i - \ln(v_i)) & i = l+1, l+2, ... \end{cases} \tag{18}$$

s.t.

$$u_i = \lambda \hat{\rho}_i + (1-\lambda)u_{i-1}$$
$$v_i = \lambda \hat{\sigma}_{\epsilon_i}^2 + (1-\lambda)v_{i-1}$$

where, i is the image index and τ is the potential change-point. $u_o = 0$, $v_o = 1$ and $0 < \lambda < 1$ is a smoothing parameter. $\hat{\rho}_i$ and $\hat{\sigma}_{\epsilon_i}^2$ represent estimated spatial autocorrelation and estimated error term variance of the ith sample, respectively. The parameter l is considered as a time window similar to GLR control chart.

It should be noted that the control limits for large-scale and small-scale statistics are chosen to achieve a predetermined in-control average run length (ARL$_o$) in phase I. In this study, the values of control limits are estimated by simulation. Through try and error approach, the in-control data is generated. Then, the values of control limits will be determined so that the ARL$_o$ is set to 200.

For a set of phase II data, the proposed statistic in Equation (18) i.e. ELR_i is calculated for each image and located on the corresponding control chart to access the stability of process. When the value of statistics exceeds control limit, the control chart warns.

- **Non-parametric method**

Step 1: Identification of regression model. Unlike the parametric method, no assumption is made about the structure of correlation in error terms in the non-parametric method. This property in the non-parametric method causes a great amount of flexibility in the structure of the covariance matrix of error terms to enable for a variety types of autocorrelation [26]. The non-parametric method is also applied to decrease the number of parameters required to model the data with high flexibility more than the parametric one.

Another non-parametric method for modeling profiles as functional data is through the use of component analysis. FPCA projects the functional data into a new spatial space where dimensions are sorted based on the amount of variability they explain. This method uses an eigenvalue decomposition of the covariance matrix of the data to find directions in the observations space along which the data have the highest variability.

In the following, a guideline is provided on how to apply FPCA for the monitoring process:

Step 2: Phase I - Estimation of Model Parameters. To describe the guideline, consider the model in Equation (11) for a profile. Suppose that measurements within a profile are correlated. In this case, error terms can be considered as a stochastic process. As a result, residuals related to error terms should be monitored. To calculate residuals, first, it is necessary to estimate regression terms. After that, residuals are obtained as follows:

$$\xi_i = \mathbf{y}_i - \hat{f}_o \tag{19}$$

where, $\xi_i = (\xi_{i,1}, \xi_{i,2}, ..., \xi_{i,n})$ is a $n \times 1$ vector contains residuals values with $E\left[\xi_i\right] = 0$, $\Sigma_\xi = E\left[\xi_i \xi_i'\right]$ and $Cov_\xi(s,t) = E\left[\xi_{i,\frac{n}{T} \times s} \xi_{i,\frac{n}{T} \times t}\right]$ in which $T = 1$. $\mathbf{y}_i$ is a $n \times 1$ vector contains observations for the ith profile, and $\hat{f}_o = \mathbf{X}\boldsymbol{\beta}_o$.

The vector ξ_i has a large number of elements, which makes it difficult to monitor directly. The FPCA is applied to reduce its dimensionality and extract a few main features for small-scale monitoring. Similar to Standard PCA, which is based on the covariance matrix, the FPCA is based on the covariance function of ξ_i.

In practice, the covariance function $Cov_\xi(s,t)$, is unknown and must be estimated from phase I dataset of size (m_o). For this purpose, the estimator of the covariance function $Cov_\xi(s,t)$, has the form:

$$\hat{Cov}_\xi(s,t) = \frac{1}{m_o} \sum_{i=1}^{m_o} \xi_{i,\frac{n}{T} \times s} \; \xi_{i,\frac{n}{T} \times t} \tag{20}$$

As stated in Subsection 2.2, in FPCA we want to find orthonormal weight functions $u_1, u_2, ...,$ such that the variance of the linear transformation (U_r) is maximal. The solution is obtained by solving the Fredholm functional eigenequation

$$\int Cov_\xi(s,t)\, u(t)\, dt = \lambda\, u(s) \tag{21}$$

Next, the eigenfunctions $u_1, u_2, ...,$ are sorted with respect to the corresponding eigenvalues $\lambda_1 \geq \lambda_2 \geq ... \geq 0$. Let λ_a be the ath largest eigenvalue and u_a the corresponding eigenfunction. By considering u_a as the ath principal

component (PC), the score of the ath PC of profile i is

$$s_{i,a} = \langle \xi_i(t), u_a(t) \rangle = \int_0^1 \xi_i(t)\, u_a(t)\, dt \qquad a = 1, 2, ..., d \tag{22}$$

Where, $\xi_i(t)$ is the scaler valued continues in time stochastic process of residuals. Since, residuals are observed at discrete time instances Equation (22) must be approximated as follows:

$$s_{i,a} = \langle \xi_i(t), u_a(t) \rangle = \frac{1}{n} \sum_{j=1}^n \xi_{i,j}\, u_a(j \times \frac{T}{n}) \qquad a = 1, 2, ..., d \tag{23}$$

The score vector for i^{th} profile is $s_i = \left[s_{i,1}, s_{i,2}, ..., s_{i,d} \right]'$, where d is the number of PCs selected to provide a certain amount of the variation.

Step 3: Phase II - Control Charting Scheme. In this step, a test statistic is designed in terms of PC scores for control charting scheme. For this task, an EWMA sequence based on $s_{i,a}$ is defined as:

$$\eta_{i,a} = (1 - w)\eta_{i-1,a} + w\, s_{i,a} \qquad a = 1, 2, ..., d \tag{24}$$

where, the initial vector $\eta_{o,a}$ is usually taken to be zero. $0 < w < 1$ is the smoothing constant. The suggested control chart warns if

$$Q_{i,l} = \begin{cases} \max\limits_{0 \leq \tau < i} \frac{(i-\tau)(2-w)}{w} \sum_{1 \leq a < d} \eta_{i,a}^T \Sigma_a^{-1} \eta_{i,a} & i = 1, 2, ..., l \\[2ex] \max\limits_{i-l \leq \tau < i} \frac{(i-\tau)(2-w)}{w} \sum_{1 \leq a < d} \eta_{i,a}^T \Sigma_a^{-1} \eta_{i,a} & i = l+1, l+2, ... \end{cases} \tag{25}$$

where, Σ_a is covariance of the score vectors, i is the image index and τ is the potential change-point. The control limit is determined to achieve a predetermined (ARL$_o$). Notice that the weighted average in Equation (24) indicates that the recent observations contain more information for detecting the shift and as a result they have higher weights [35]. Thus, this chart is referred to as the principal components-based EWMA chart, or PCEWMA.

3.2.3. Identifying sources of variability

In geometric profiles, the quality characteristic has two sources of variability, firstly, the variations in manufacturing process and secondly, the variations in material properties and machinery conditions [20]. As stated in section 1, monitoring of changes in process is performed by considering large-scale components related to regression function. Also, monitoring of changes in materials and machinery conditions is done by considering small-scale components related to error term [21, 20]. Therefore, under out-of-control conditions, identifying the scale in which the change occurred is important for corrective actions. For this purpose, several suitable test statistics are determined using profile data. Next, these test statistics are located on control charts. In monitoring process, if either large-scale or small-scale statistics exceed control limits, the control charts give out-of-control signals. In this case, the source of the assignable cause can be determined.

4. Simulation study

The simulation study is expressed according to stages described in Section 3. In the first step, a $2664 \times 3264 \times 3$ RGB color image from a tile is captured by the MVS. In the next step, the image pre-processing is done including eliminating image background, enhancing contrast and converting the RGB color image to a gray-scale image. The resulted image is considered as the nominal image (Figure 7).

The next step is to generate images as phase I data set by adding random error to the nominal image. For creating random error, first, a matrix of white noises with dimensions same as nominal image is generated. Then, first-order autoregressive filter is applied to this matrix. By doing this, random error is actually obtained through a linear process with autoregressive parameter ϕ ($|\phi| < 1$). Adding correlated random errors to image creates spatial correlation
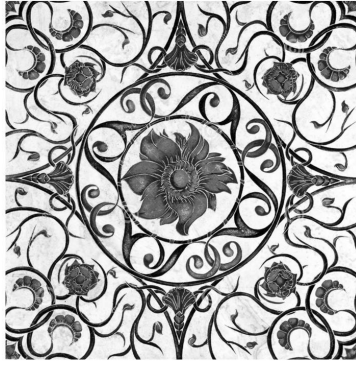
Figure 7: The nominal image after pre-processing.

between image pixels. In the simulation study, white noise is defined as a zero-mean Gaussian distribution with a standard deviation equal to 0.01. Signal-to-noise ratio (SNR) is computed to obtain the correct amount for white noise standard deviation. More information on SNR can be found at Koosha et al. [7]. Also, the in-control parameter of the linear process, ϕ is considered 0.4.

After generating 1000 in-control images (phase I dataset), the control limits of the control charts are estimated by $10,000$ simulation runs. For this task, two groups of control charts must be designed for each class of the proposed method. For each group, two control charts are considered to monitor the process simultaneously: GLR\OMNIBUS, GLR\PCEWMA. Control charts are designed to obtain a predetermined (ARL_o) in phase I. In this study, the values of control limits are estimated by simulation. Through try and error approach, the in-control data is generated. Then, the values of control limits will be determined so that the ARL_o is set to 200. After the simulation, the control limits for GLR, OMNIBUS and PCEWMA control charts are found to be 12.24, 1.67 and 2.54, respectively.

In phase II, two types of assignable causes are considered to examine the performance of the proposed method. In first type, ϕ is shifted as small-scale shift while in second type, shift in regression function is generated as large-scale shift. Thus, it is desired to detect these two different types of shifts. For this purpose, three scenarios are defined based on ϕ values: $\phi = 0, 0.4, 0.8$. It should be noted that under scenario 1, only large-scale shift exists. While, in other scenarios both large-scale and small-scale shifts exist. Each scenario is studied under 11 shift magnitudes (Δ). Fault region is considered squared shaped of size 10×10. Also, to investigate the preciseness of change-point estimation in the simulation study, τ sets equal to 10, which means that in each repetition, a shift occurs from the eleventh image.

The results of the simulation are presented in Table 1. The ARL and SDRL are evaluation metrics for fault detection capability, which indicate average and standard deviation of run lengths, respectively. The absolute deviation between the estimated and actual change-points ($\theta = |\hat{\tau} - \tau|$) is regarded as the metric for examining the preciseness of change-point estimation by the proposed methods. $E(\theta)$, $STD(\theta)$ and $MED(\theta)$ represent mean, standard deviation and median of θ's.

Insert Table 1 about here.

For comparative analysis, Figure 8 is drawn based on different values of shift magnitude for two classes of the proposed methods. As shown in Table 1, as the shift magnitude increases (from +1 to +10), the ARL values decrease under all scenarios. The downward trend of ARL values in small shifts ($\Delta \leq 4$) is more intense than in larger shifts ($\Delta > 4$). The ARL values also decrease with increasing degree of correlation. It is worth noting that the performance of both two classes based on ARL converges in larger shifts under three scenarios. The analysis mentioned for the ARL performance metric also applies to the SDRL performance metric with bigger convergence intensity under three scenarios. The estimation accuracy of change-point is evaluated through three metrics using mean, standard deviation and median of $\theta = |\hat{\tau} - \tau|$. As can be seen from Table 1, the values for the metrics are drastically decreased for smaller shifts ($\Delta \leq 3$) under three scenarios. In other words, the estimation accuracy increases sharply. It is also inferred that the proposed method is better in estimating the change-point in larger shifts. This trend is observed under three scenarios.

In summary, the performance of the control charting groups for monitoring is almost the same and reasonable, however GLR\OMNIBUS charting group performed slightly better to monitor process.
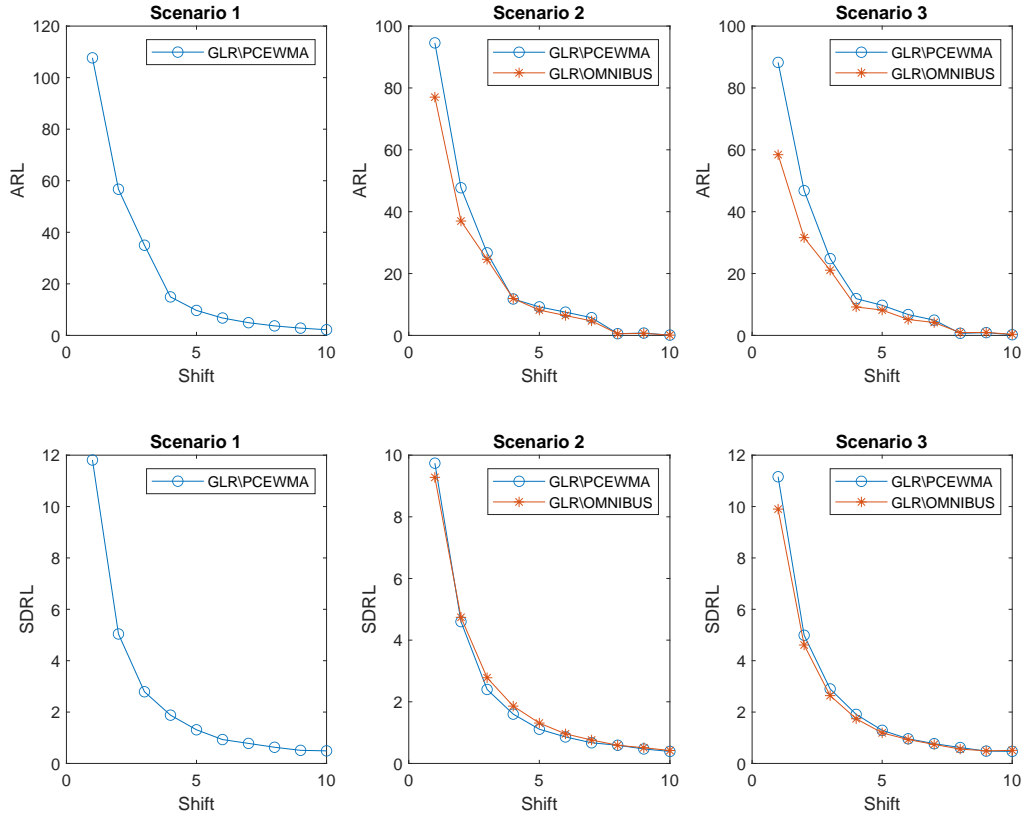
Figure 8: Performance of the control charting groups of proposed approach under three scenarios in terms of ARL and SDRL.

Under the first scenario, the parametric method loses its efficiency in modeling and monitoring process due to the lack of spatial correlation. However, the non-parametric method does not face this problem due to considering no assumptions in correlation modeling and can be used to monitor the process. In the second and third scenarios, the parametric method has a relatively better performance than the non-parametric method due to better spatial correlation modeling. This performance is also better by increasing the value of ϕ. It can be inferred that in the presence of spatial correlation, the parametric modeling improves process monitoring. On the other hand, if the presence of spatial correlation is not detectable or does not exist, the use of nonparametric method is recommended. Note that the degree of correlation is determined based on Wald test.

A competitive method presented by Koosha et al. [7] is considered to compare consequences and confirm the proper performance of the proposed method. By comparing the results, it can be concluded that the proposed method in this study has a better ability to detect shifts due to consideration of spatial correlation and simultaneous use of two control charts. This difference in performance, especially in detecting small shifts, is more visible.

To evaluate the efficiency of the proposed method in detecting out-of-control conditions, an indicator is defined as probability of true detection. This indicator can be computed as follows:

$$Probability\ of\ True\ Detection = \frac{Number\ of\ out\text{-}of\text{-}control\ conditions\ detected\ in\ each\ scale}{Total\ number\ of\ out\text{-}of\text{-}control\ conditions\ detected} \tag{26}$$

To compute the indicator in Equation (26), two type of deviations are considered: Large-scale deviation (shift in intensities of image) and Small-scale deviation (shift in parameter of linear process, ϕ). In phase II, image data is generated by considering each type of deviations separately. Firstly, a case is examined in which large-scale deviation has occurred ($\Delta\ :\ 0\ \rightarrow\ +2$). After 1000 simulations, number of out-of-control conditions detected related to

each scale and total number of out-of-control conditions detected are counted. Then, probability of true detection is computed based on Equation (26). This procedure is also performed for a case where a small-scale deviation occurs ($\phi : 0.4 \rightarrow 0.6$). The results of the simulation are presented in Table 2.

Insert Table 2 about here.

According to results obtained in Table 2, it can be seen that GLR\OMNIBUS charting group has slightly better performance than the second group to discover out-of-control statue. The proposed methods also have better true detection capabilities in cases where Large-scale deviation occurs. Briefly, it can be concluded that the proposed methods have an appropriate efficiency in discovering out-of-control statues.

5. Case study

In this section, a real case study is analyzed to confirm the ability of the proposed methods in an actual industrial operation. For this purpose, the tile production process is analysed. It should be noted that the case study considered in this manuscript is parallel to that considered by Koosha et al. [7] and Megahed et al. [5] to make the results comparable.

The first step is to make sure that the captured images are appropriate for statistical monitoring and are minimally affected by the alteration in the environmental conditions such as shift in camera positioning, part positioning, and lighting conditions. The camera (*Samsung Galaxy Camera 16 MP EK-GC120VRAMC4, 21 x OpticalZoom, 23mm Wide Zoom Lens, Android(TM) 4.1, Jelly Bean*) mounted on an $3ft \times 4ft \times 3ft$ aluminium frame to fix in the duration of the experiment. Next, the tiles were stabilized using a simple jig so that each tile was located in the same location (see Figure 9a). Finally, the images were captured in a light-controlled environment. Because of the room's lack of external illumination, there was no temporal variation in lighting conditions in this environment.

In the image acquisition step, 160 successive images were captured from a tile production line. Some image pre-processing methods such as converting red, green, blue (RGB) images to gray-scale, contrast enhancement and removing image background are applied to the captured images for use in the monitoring process (see Figure 9b). Next, the images are divided into two parts for use in phase I and phase II. For this task, the first 40 images are considered as phase I data (in-control images). This number of images is sufficient to estimate the process parameters in phase I [5]. However, the user can generate many in-control images in accordance with these 40 images to achieve greater accuracy in parameter estimation. Based on phase I data, the control limits for GLR, OMNIBUS and PCEWMA control charts are determined 5.8365, −0.6451 and 1.7221, respectively.

The other 120 images are regarded as two sets of phase II data with equal numbers. It is supposed that the process is in-control for the first 30 images of phase II data in each set. After that, two types of deviations are applied in process separately. For first set, a shift in intensities of image with shift magnitude, +5 and Fault region, squared shaped of size 10×10 is added (Figure 9c). For second one, a shift in parameter of linear process ϕ is added to images by 0.2 (Figure 9d).

For first set of phase II data, the proposed statistics in Equation (12), R_i, Equation (18), ELR_i and Equation (25), Q_i are calculated for each image (see Table 3) and located on the corresponding control charts to investigate the out-of-control conditions. When the value of statistics exceeds control limits, the control charts warn (see Figure 10a). This procedure is also performed for secondary set and the proposed statistics are calculated for each image (see Table 4). Subsequently, these statistics are also plotted on the corresponding control charts (see Figure 10b).

As can be seen in Figure 10a, GLR control chart warns the out-of-control condition and detected the shift on the 32nd image (first image after exerted the shift). While, other control charts related to small-scale do not warn. Thus, it can be concluded that the source of variability is related to the manufacturing process rather than the materials and environmental conditions of machinery operations. Industrial users can take corrective actions using this information. This performance indicates the proper efficiency of the proposed methods. The competitive method related to Megahed et al. [5] detected this shift on the 34th image, and the results achieved by Koosha et al. [7] indicate a similar performance.

Insert Table 3 about here.
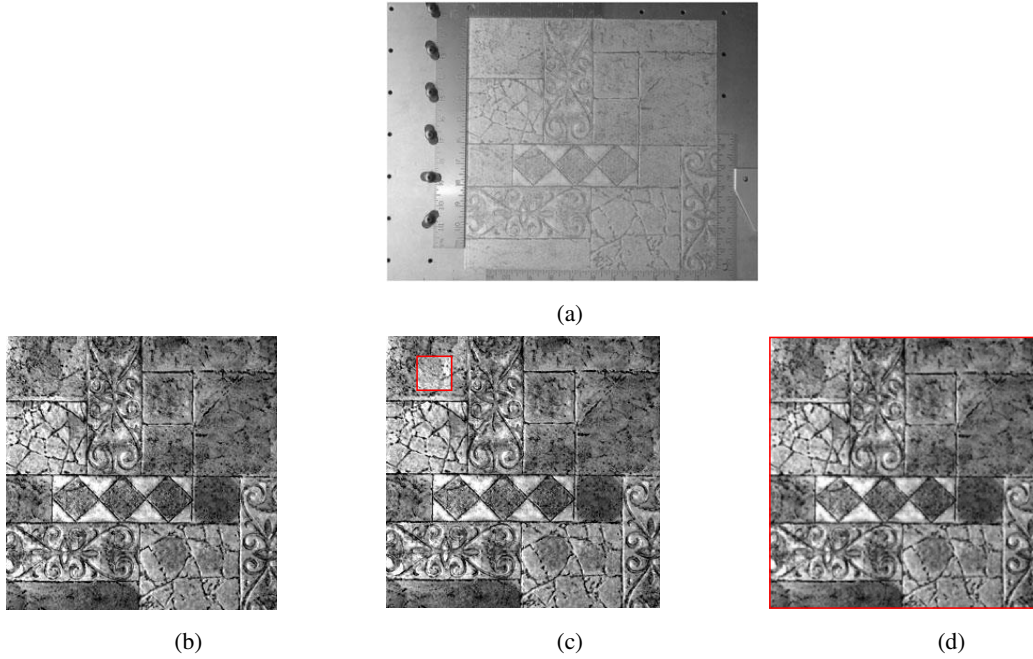
Insert Table 4 about here.

Figure 9: The case study: (a) Original tile Image, (b) After pre-processing, (c) Large-scale deviation, (d) Small-scale deviation.

The calculated statistics from secondary set are plotted on corresponding control charts. As can be seen in Figure 10b, control charts related to small-scale warn out-of-control condition. The difference is that OMNIBUS control chart warned faster. This performance also indicates the appropriate efficiency of the proposed methods in detecting the type of deviation. It should be noted that the competitive methods lack the ability to determine the source of variability. Therefore, determining the source of variability and providing information to users in order to return the process to in-control condition is one of the advantages of the proposed methods.

After detecting the out-of-control condition, the proposed methods estimated change-point. As the result shows, the accuracy of the proposed methods in change-point estimation is appropriate and correctly estimates the real change-point. Also, both competitive methods are able to estimate the change-point accurately on the 31st image.

6. Conclusion

In most industrial cases, the use of an image in statistical quality control has been taken into account to detect defective or separate conforming products from non-conforming ones. There were rarely statistical and process control viewpoints in this field. In this paper, a regression-based methods under two-scale analysis were presented for monitoring processes that produce image data. Under large-scale analysis, wavelet transformation was applied to geometric profile created from images to extract the main features. Then, the extracted components were monitored using GLR chart. On the next scale, small-scale components were monitored, which could be described by correlation in error terms. One parametric and one non-parametric methods were developed in this study to monitor correlation in error terms. In the parametric method, correlation in error terms was modelled by first-order spatial correlation model. After that, small-scale components were monitored using OMNIBUS control chart. In the non-parametric method, no assumption was made about the structure of correlation in error terms. To extract useful information about the nature of correlation in error terms, FPCA was used. After extracting some features for characterizing process variations, an EWMA chart was constructed based on principal components PC scores. Simulation and industrial case studies were implemented to examine the performance of the proposed methods in detecting different shifts. The proposed methods were capable of estimating a change-point in addition to detecting out-of-control conditions. Simulation studies indicated the appropriate performance of the proposed methods in detecting process deviations and estimating the change-point. The results also proved the appropriate performance of the proposed method in
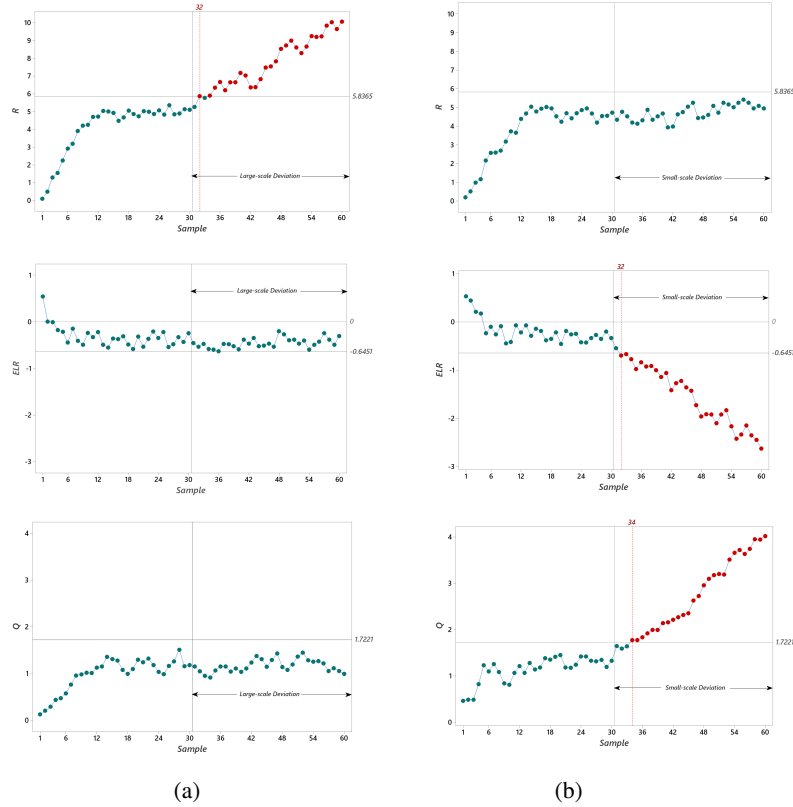
Figure 10: Control chart scheme in phase II under (a) Large-scale deviation, and (b) Small-scale deviation.

manipulating industrial cases to discover out-of-control conditions and identify the source of variability. The use of other basic functions of wavelet, spatial correlation modes, and other control charts is one of the issues that can be considered in future research.

References

1. Wang, K. and Tsung, F. "Using profile monitoring techniques for a data-rich environment with huge sample size", *Quality and reliability engineering international*, 21(7):677–688 (2005).
2. Wells, L. J., Megahed, F. M., Niziolek, C. B., et al. "Statistical process monitoring approach for high density point clouds", *Journal of Intelligent Manufacturing*, 24(6):1267–1279 (2013).
3. Menafoglio, A., Grasso, M., Secchi, P., et al. "Profile monitoring of probability density functions via simplicial functional pca with application to image data", *Technometrics*, 60(4):497–510 (2018).
4. Eslami, D., Izadbakhsh, H., Ahmadi, O., et al. "Spatial-nonparametric regression: an approach for monitoring image data", *Communications in Statistics-Theory and Methods*, pages 1–24 (2021a).
5. Megahed, F. M., Wells, L. J., Camelio, J. A., et al. "A spatiotemporal method for the monitoring of image data", *Quality and Reliability Engineering International*, 28(8):967–980 (2012).
6. He, Z., Zuo, L., Zhang, M., et al. "An image-based multivariate generalized likelihood ratio control chart for detecting and diagnosing multiple faults in manufactured products", *International Journal of Production Research*, 54(6):1771–1784 (2016).
7. Koosha, M., Noorossana, R., and Megahed, F. "Statistical process monitoring via image data using wavelets", *Quality and Reliability Engineering International*, 33(8):2059–2073 (2017).
8. Zuo, L., He, Z., and Zhang, M. "An ewma and region growing based control chart for monitoring image data", *Quality Technology & Quantitative Management*, 17(4):470–485 (2020).
9. Huang, T., Wang, S., Yang, S., et al. "Statistical process monitoring in a specified period for the image data of fused deposition modeling parts with consistent layers", *Journal of Intelligent Manufacturing*, pages 1–16 (2020).
10. Amirkhani, F. and Amiri, A. "A novel framework for spatiotemporal monitoring and post-signal diagnosis of processes with image data", *Quality and Reliability Engineering International*, 36(2):705–735 (2020).
11. Eslami, D., Izadbakhsh, H., Ahmadi, O., et al. "Statistical monitoring of image data using multi-channel functional principal component analysis", *Communications in Statistics-Theory and Methods*, pages 1–18 (2021b).

12. Noorossana, R. and Nikoo, M. "Phase ii monitoring of geometric profiles", *Communications in Statistics-Simulation and Computation*, 44(4):1036–1049 (2015).

13. Čížek, P. and Sadıkoğlu, S. "Robust nonparametric regression: A review", *Wiley Interdisciplinary Reviews: Computational Statistics*, 12(3):e1492 (2020).

14. Nikoo, M. and Noorossana, R. "Phase ii monitoring of nonlinear profile variance using wavelet", *Quality and Reliability Engineering International*, 29(7):1081–1089 (2013).

15. Cohen, A. and Atoui, M. A. "On wavelet-based statistical process monitoring", *Transactions of the Institute of Measurement and Control*, 44(3):525-538 (2020).

16. Lim, M. and Bae, S. J. "Spatial monitoring of wafer map defect data based on 2d wavelet spectrum analysis", *Applied Sciences*, 9(24):5518 (2019).

17. Noorossana, R., Saghaei, A., and Amiri, A. "Statistical analysis of profile monitoring", *John Wiley & Sons*, volume 865 (2011).

18. Cheng, T. C. and Yang, S. F. "Monitoring profile based on a linear regression model with correlated errors", *Quality Technology & Quantitative Management*, 15(3):393–412 (2018).

19. Colosimo, B. M. "Modeling and monitoring methods for spatial and image data", *Quality Engineering*, 30(1):94–111 (2018).

20. Colosimo, B. M., Mammarella, F., and Petro, S. "Quality control of manufactured surfaces", *In Frontiers in Statistical Quality Control*, pages 55–70. Springer (2010).

21. Paynabar, K. and Jin, J. "Characterization of non-linear profiles variations using mixed-effect models and wavelets", *Iie transactions*, 43(4):275–290 (2011).

22. Vidakovic, B. "Statistical modeling by wavelets", *John Wiley & Sons*, volume 503 (2009).

23. Daubechies, I. "Ten lectures on wavelets", *CBMS-NSF series*, vol 61 (philadelphia: Siam) (1992).

24. Deregowski, K. and Krzysko, M. "Principal components analysis for functional data", *In Colloquium Biometricum*, volume 41 (2011).

25. Soleimani, P., Noorossana, R., and Amiri, A. "Simple linear profiles monitoring in the presence of within profile autocorrelation", *Computers & Industrial Engineering*, 57(3):1015–1021 (2009).

26. Littell, R. C., Milliken, G. A., Stroup, W. W., et al. "Sas system for mixed models" (1996).

27. LeSage, J. and Pace, R. "Introduction to spatial econometrics", *Chapman & hall/crc* (2009).

28. Qiu, P. "Image processing and jump regression analysis", *John Wiley & Sons*, volume 599 (2005).

29. Marques, O. "Practical image and video processing using MATLAB", *John Wiley & Sons* (2011).

30. Wei, G., Zhou, Y., Gao, X., et al. "Field-aligned quadrangulation for image vectorization", *In Computer Graphics Forum*, volume 38, pages 171–180. Wiley Online Library (2019).

31. Reynolds Jr, M. R. and Lou, J. "An evaluation of a glr control chart for monitoring the process mean", *Journal of quality technology*, 42(3):287–310 (2010).

32. Anselin, L. "Spatial econometrics: methods and models", *Springer Science & Business Media*, volume 4 (2013).

33. Greene, W. H. "The econometric approach to efficiency analysis", *The measurement of productive efficiency and productivity growth*, 1(1):92–250 (2008).

34. Zhang, J., Zou, C., and Wang, Z. "A control chart based on likelihood ratio test for monitoring process mean and variability", *Quality and Reliability Engineering International*, 26(1):63–73 (2010).

35. Ren, H., Chen, N., and Wang, Z. "Phase-ii monitoring in multichannel profile observations", *Journal of Quality Technology*, 51(4):338–352 (2019).

Dariush Eslami completed his PhD program in Industrial Engineering at the Kharazmi University of Tehran, Iran in 2021. He received his MS from faculty of Industrial Engineering at Iran University of Science and Technology in 2015. He received his BS in Industrial Engineering from Noshirvani University of Technology in 2013. His research interests are Statistical Process Control, Process Optimization, and Data Analysis.

Hamidreza Izadbakhsh is an Assistant Professor of Industrial Engineering at Kharazmi University. His research interests include System Modeling, Simulation, Quality Control, and Data Science. He has a PhD in Industrial Engineering from Iran University of Science and Technology. He received his MSc degree in Industrial Engineering from University of Tehran . He has completed several innovative projects in industry. He is the author and co-author of books and many papers and advisor of many PhD and master theses.

Orod Ahmadi is a Ph.D. graduate in Industrial Engineering from K. N. Toosi University of Technology in Tehran, Iran, in 2015. He is presently an Assistant Professor at Kharazmi University. His research interests include Statistical Quality Control, Applied Multivariate Statistics, Robust Statistics, Data Analysis and Time Series Analysis.

Marzieh Zarinbal is an Assistant Professor of Information Systems (IS) Research Group at Iranian Research Institute for Information Science and Technology (Irandoc). Her research interests include System Modeling, Intelligent Systems, Fuzzy Logic, Machine Learning, Image Processing, and Gamification. She has a PhD in Industrial Engineering from Amirkabir University of Technology (Tehran Polytechnic). She received her MSc degree in Industrial Engineering from Amirkabir University of Technology and holds a BSc degree in Industrial Engineering from Alzahra University. She is the author and co-author of some books and many papers and advisor of some PhD and master theses.

Table 1
Simulation results.

Δ	Scenario 1: $\phi = 0$					Scenario 2: $\phi = 0.4$					Scenario 3: $\phi = 0.8$				
	ARL	SDRL	E(θ)	std(θ)	med(θ)	ARL	SDRL	E(θ)	std(θ)	med(θ)	ARL	SDRL	E(θ)	std(θ)	med(θ)
							Group 1: GLR\OMNIBUS								
+10	–	–	–	–	–	0.08	0.42	0.00	0.47	0.00	0.35	0.51	0.00	0.13	0.00
+9	–	–	–	–	–	0.68	0.51	0.00	0.81	0.00	0.96	0.49	0.00	0.51	0.00
+8	–	–	–	–	–	0.53	0.59	0.00	0.57	0.00	0.91	0.56	0.00	0.02	0.00
+7	–	–	–	–	–	4.69	0.76	0.00	0.66	0.00	4.16	0.74	0.00	0.31	0.00
+6	–	–	–	–	–	6.43	0.96	0.00	0.36	0.00	5.15	0.93	-0.04	0.53	0.00
+5	–	–	–	–	–	8.18	1.31	0.00	0.74	0.00	8.18	1.19	-0.03	0.08	0.00
+4	–	–	–	–	–	11.92	1.86	0.00	1.18	0.00	9.20	1.74	-0.15	1.38	0.00
+3	–	–	–	–	–	24.61	2.78	0.00	1.82	0.00	21.08	2.64	0.75	1.97	0.00
+2	–	–	–	–	–	36.97	4.74	0.01	2.00	1.00	31.60	4.61	5.13	2.99	2.00
+1	–	–	–	–	–	77.03	9.28	0.90	5.63	1.00	58.42	9.90	11.59	4.79	3.00
							Group 2: GLR\PCEWMA								
+10	2.20	0.49	-0.01	0.96	0.00	0.09	0.39	0.00	0.14	0.00	0.22	0.47	0.00	0.82	0.00
+9	2.82	0.51	-0.04	0.75	0.00	0.75	0.47	0.00	0.22	0.00	0.89	0.48	0.00	0.40	0.00
+8	3.69	0.63	-0.34	0.92	0.00	0.56	0.59	0.00	0.16	0.00	0.67	0.62	0.00	0.78	0.00
+7	4.93	0.78	-0.08	0.50	1.00	5.73	0.67	0.00	0.27	0.00	4.93	0.77	0.00	0.97	0.00
+6	6.74	0.93	-0.12	0.63	1.00	7.52	0.86	0.00	0.97	0.00	6.74	0.96	-0.02	0.36	0.00
+5	9.69	1.31	0.49	1.26	2.00	9.22	1.11	0.00	0.05	0.00	9.73	1.29	-0.01	0.74	1.00
+4	14.93	1.88	3.81	1.16	1.00	11.77	1.60	-0.07	1.60	0.00	11.88	1.91	-0.07	1.18	0.00
+3	34.98	2.79	12.34	2.57	3.00	26.69	2.40	-0.11	1.44	1.00	24.82	2.90	0.18	1.82	1.00
+2	56.71	5.04	34.27	3.04	13.00	47.73	4.60	0.07	2.10	0.00	46.80	4.99	2.47	2.00	3.00
+1	107.67	11.81	108.67	6.89	34.00	94.57	9.74	1.08	6.27	2.00	88.26	11.16	12.49	5.63	4.00

Table 2
Simulation results (Probability of True Detection).

Control Charting Group	Large-scale Deviation	Small-scale Deviation
GLR\PCEWMA	%90.1	%83.6
GLR\OMNIBUS	%90.2	%84.3

Table 3
The calculated statistics under first type of deviation ($\lambda = 0.5$).

i	R_i	ELR_i	Q_i	i	R_i	ELR_i	Q_i	i	R_i	ELR_i	Q_i
1	0.1043	0.5426	0.1321	21	5.0228	-0.5328	1.2471	41	7.0232	-0.3833	1.1131
2	0.5093	0.0022	0.2089	22	4.9935	-0.3668	1.3227	42	6.3667	-0.4660	1.2361
3	1.2966	-0.0077	0.2941	23	4.8664	-0.2103	1.1835	43	6.3796	-0.3500	1.3803
4	1.5617	-0.1750	0.4424	24	5.0657	-0.3458	1.0375	44	6.8165	-0.5207	1.3078
5	2.2498	-0.2160	0.4794	25	4.8363	-0.2182	0.9880	45	7.4799	-0.5083	1.1459
6	2.9254	-0.4448	0.5776	26	5.3617	-0.5435	1.1676	46	7.5440	-0.4661	1.2904
7	3.1861	-0.1460	0.7673	27	4.8492	-0.4790	1.2640	47	7.8237	-0.5357	1.4299
8	3.9056	-0.4046	0.9625	28	4.9047	-0.3267	1.5124	48	8.5347	-0.2031	1.1400
9	4.2103	-0.4914	0.9867	29	5.1374	-0.4301	1.1614	49	8.7232	-0.2659	1.0813
10	4.2576	-0.2365	1.0201	30	5.1077	-0.2449	1.1820	50	8.9940	-0.3944	1.1989
11	4.7054	-0.3289	1.0134	31	5.2694	-0.4548	1.1559	51	8.6078	-0.3854	1.3638
12	4.7187	-0.2197	1.1282	32	5.8655	-0.5330	1.0514	52	8.2910	-0.4697	1.4515
13	5.0344	-0.4913	1.1536	33	5.7798	-0.4738	0.9519	53	8.6582	-0.4019	1.2860
14	5.0037	-0.5513	1.3575	34	5.8995	-0.5809	0.9148	54	9.2478	-0.5937	1.2557
15	4.9352	-0.3603	1.3128	35	6.3417	-0.5973	1.0694	55	9.1893	-0.4926	1.2698
16	4.4845	-0.3713	1.2831	36	6.6639	-0.6322	1.1514	56	9.2213	-0.4279	1.2184
17	4.6689	-0.3119	1.0792	37	6.2089	-0.4732	1.1543	57	9.8265	-0.2459	1.0591
18	5.0627	-0.4889	0.9955	38	6.6445	-0.4806	1.0427	58	10.0215	-0.3843	1.1143
19	4.8664	-0.5847	1.1013	39	6.6517	-0.5203	1.1103	59	9.6464	-0.4934	1.0598
20	4.7336	-0.3177	1.2972	40	7.1735	-0.5877	1.0402	60	10.0495	-0.3063	0.9984

Table 4
The calculated statistics under second type of deviation ($\lambda = 0.5$).

i	R_i	ELR_i	Q_i	i	R_i	ELR_i	Q_i	i	R_i	ELR_i	Q_i
1	0.2017	0.5312	0.4688	21	4.6973	-0.1861	1.2036	41	3.9847	-1.0575	2.1444
2	0.5192	0.4401	0.4899	22	4.4207	-0.2611	1.2014	42	4.6356	-1.4136	2.2043
3	0.9970	0.2056	0.4944	23	4.6993	-0.2485	1.2643	43	4.7622	-1.2706	2.2667
4	1.1702	0.1721	0.8279	24	4.8625	-0.4214	1.4203	44	5.0522	-1.2203	2.3117
5	2.1686	-0.2377	1.3124	25	4.9621	-0.4260	1.4235	45	5.2608	-1.3572	2.3701
6	2.5858	-0.1007	1.1040	26	4.6832	-0.3345	1.3409	46	4.4343	-1.4269	2.4368
7	2.6020	-0.2591	1.2014	27	4.1926	-0.2676	1.3176	47	4.4643	-1.7253	2.7106
8	2.7085	-0.0898	1.1008	28	4.5477	-0.3527	1.3441	48	4.6007	-1.9574	2.9099
9	3.1899	-0.4434	0.8714	29	4.5637	-0.1978	1.2100	49	5.0834	-1.9096	3.9654
10	3.7303	-0.4155	0.9043	30	4.7341	-0.3348	1.3276	50	4.7213	-1.9176	3.1763
11	3.6512	-0.0706	1.0164	31	4.3523	-0.5471	1.6481	51	5.2602	-2.0998	3.1971
12	4.3913	-0.2151	1.1743	32	4.7704	-0.6951	1.5803	52	5.1634	-1.9190	3.1896
13	4.6769	-0.0726	1.1024	33	4.5357	-0.6648	1.4726	53	5.0082	-1.8310	3.5102
14	5.0387	-0.2860	1.2792	34	4.2045	-0.7741	1.7154	54	5.2605	-2.1596	3.6533
15	4.7886	-0.1424	1.1603	35	4.1327	-0.9752	1.7310	55	5.4274	-2.4194	3.7152
16	4.9400	-0.1874	1.1777	36	4.3166	-0.8373	1.8002	56	5.2566	-2.3334	3.6147
17	5.0279	-0.3840	1.3958	37	4.8736	-0.9218	1.9523	57	4.9517	-2.1439	3.7630
18	4.9510	-0.3538	1.3701	38	4.3503	-0.9097	1.9946	58	5.0867	-2.3484	3.9500
19	4.5340	-0.2190	1.4120	39	4.5255	-0.9976	1.9973	59	4.9542	-2.4436	3.9689
20	4.2374	-0.4567	1.4789	40	4.6866	-1.1428	2.0954	60	3.9847	-2.6197	4.1048