Jump Regression, Image Processing and Quality Control

Peihua Qiu

Department of Biostatistics, University of Florida

2004 Mowry Road, Gainesville, FL 32610

Abstract

Images have been widely used in manufacturing applications for monitoring production processes, partly because they are often convenient and economic to acquire by different types of imaging devices. Medical imaging techniques, such as CT, PET, X-ray, ultrasound, magnetic resonance imaging (MRI), and functional MRI, have become a basic medical diagnosis tool nowadays. Satellite images are also commonly used for monitoring the changes of the earth's surface. In all these applications, image comparison and monitoring are the common and fundamentally important statistical problems that should be addressed properly. In computer science, applied mathematics, statistics and some other disciplines, there have been many image processing methods proposed. In this paper, I will discuss (i) a powerful statistical tool, called jump regression analysis (JRA), for modeling and analyzing images and other types of data with jumps and other singularities involved, (ii) some image processing problems and methods that are potentially useful for image comparison and monitoring, and (iii) some of my personal perspectives about image comparison and monitoring.

Key Words: Discontinuities; Edges; Features; Image comparison; Image monitoring; Jumps; Process monitoring; Statistical process control.

1 Introduction

In manufacturing industries, images have been widely used for quality control purposes, partly because they are relatively convenient and economic to acquire. See Figure 1 for a demonstration. Applications include stress and strain analysis of products (Patterson and Wang 1991), anomaly detection of rolling processes (Jin et al. 2004, Yan et al. 2017), inspection of composite material fabrication (Sohn et al. 2004), quality control in liquid crystal display manufacturing (Jiang et al. 2005), structural health monitoring (Balageas et al. 2006), and so forth. Medical imaging techniques, such as CT, PET, SPECT, X-ray, ultrasound, magnetic resonance imaging (MRI), and functional MRI, have already become one of the basic medical diagnosis tools (Kahn et al. 2007, Parker et al. 2011). Satellite images have been widely used for monitoring changes in the earth's surface (http://landsat.gsfc.nasa.gov). They have become a basic tool in studying agriculture, forestry and range resources, land use and mapping, geology, hydrology, coastal resources, and environment. In many of these imaging applications, comparison of two or more images of a same object (or similar objects) is a basic task for information fusion, structure/target localization, difference detection, and other purposes. In some applications, image data are obtained in the form of data streams, in the sense that new images are acquired from a longitudinal process sequentially over time (e.g., fMRI images, or images acquired from a rolling process). In such applications, one fundamental task is to monitor the image sequence to see whether the underlying longitudinal process changes significantly over time. This is the image monitoring problem. Therefore, both the image comparison and image monitoring problems are critically important in many applications.

(a) (b) (c)

Figure 1: (a) An imaging system for checking steel surface and detecting cracks and defects. (b) A camera. (c) An observed image of steel surface.

To process and analyze the data in a single image, there are already many methods proposed in the computer science, applied mathematics, statistics and some other disciplines. These methods serve different purposes and belong to different areas in image processing (cf., Gonzalez and Woods 1992, Qiu 2005), including edge detection, image denoising, image segmentation, image deblurring, image registration, and many others. These methods would definitely be helpful for image comparison and monitoring, although they cannot solve the image comparison and monitoring problems directly. To describe an image, there are several different tools. One powerful tool in the statistics literature is jump regression analysis (JRA), which is for estimating curves and surfaces with jumps and other singularities. JRA is relevant to statistical process control (SPC), not only because it can handle image data effectively so that it is helpful for image comparison and monitoring, but also because it has a great potential for solving some other SPC problems, since shifts (or drifts) are the major concern in most SPC problems and JRA methodologies can detect such singularities effectively. In this paper, we will first introduce the JRA approach and its connection to the image processing problem, then discuss some image processing problems and certain commonly used methods for solving these problems that are potentially useful for image comparison and monitoring, and finally provide some personal perspectives about image comparison and monitoring.

The rest of the article is organized as follows. Some JRA models and major JRA problems and methods are introduced in Section 2. Certain image processing problems and the related representative methods are discussed in Section 3. Our personal perspectives about image comparison and monitoring are provided in Section 4. Several remarks conclude the article in Section 5.

2 Jump Regression Analysis

Nonparametric regression analysis provides statistical tools for estimating regression curves or surfaces from noisy data. Conventional nonparametric regression procedures, such as the local kernel smoothers and splines, are appropriate for estimating continuous regression functions only (Fan and Gijbels 1996, Wahba 1990). In practice, however, the underlying regression function could have jumps or other singularities. For instance, the small diamonds in Figure 2 denote the sea-level pressures observed by a Bombay weather station in India during 1921–1992. We can notice that there is a jump around the year 1960, which was confirmed by us in Qiu and Yandell (1998). If a conventional local kernel smoother is used for estimating the underlying regression function, then the estimated function looks like the dashed curve in the plot. Obviously, the jump, which could be an important data structure that reveals certain interesting atmospheric phenomena, is blurred by the smoother. JRA is specifically for handling cases with jumps or other singularities in the underlying regression function (Qiu 2005). The regression function estimated by a JRA method is shown by the solid curve in Figure 2. It can be seen that the jump structure is preserved well by this method.

One-dimensional (1-D) JRA: A 1-D JRA model has the form

$$y_i = f(x_i) + \varepsilon_i, \quad \text{for } i = 1, 2, \ldots, n, \tag{1}$$

where $\{y_i,\ i = 1, 2, \ldots, n\}$ are observations of a response variable y, $\{x_i,\ i = 1, 2, \ldots, n\}$ are design points, f is an unknown regression function, and $\{\varepsilon_i,\ i = 1, 2, \ldots, n\}$ are random errors. For

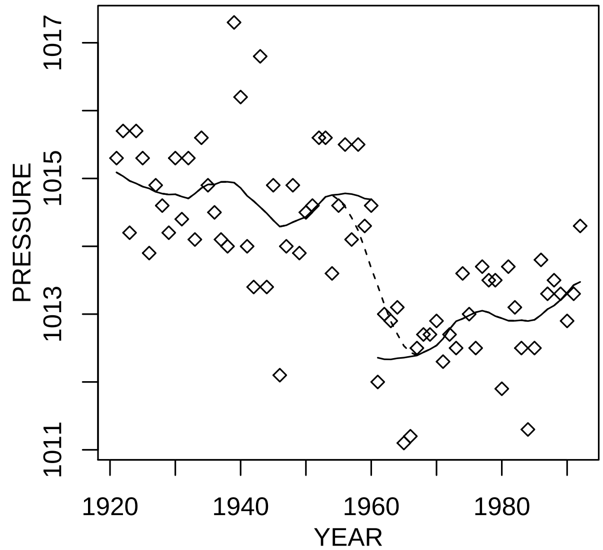


Figure 2: The small diamonds denote the sea-level pressures observed by a Bombay weather station in India during 1921–1992. The solid curves denote the estimated regression function with a detected jump around the year 1960 preserved. The dashed curve denotes the conventional local linear kernel estimator of the regression function that blurs the jump around the year 1960.

simplicity, we assume that the design interval is $[0, 1]$. In (1), f is assumed to have the expression

$$f(x) = g(x) + \sum_{j=1}^{p} d_j I(x > s_j), \text{ for } x \in [0, 1], \tag{2}$$

where g is a continuity part of f (i.e., $g(x)$ is a continuous function of x in $[0, 1]$), p is the number of jump points, $\{s_j,\ j = 1, 2, \ldots, p\}$ are the jump positions, and $\{d_j,\ j = 1, 2, \ldots, p\}$ are the jump sizes. The major goal of JRA is to estimate the unknown quantities in (2) from the observed data. More specifically, there are two major tasks. One is to estimate the quantities in the jump part of f (i.e., $\sum_{j=1}^{p} d_j I(x > s_j)$), which is called *jump detection*. The other is to estimate the entire regression function f with the possible jumps preserved, which is called *jump-preserving curve estimation*. Because jumps are often important data structures, both the first and the second tasks are important in applications.

In the literature, there have been many jump detectors proposed (e.g., Joo and Qiu 2009, Loader 1996, Müller 1992, Qiu 1991, 1994, Qiu et al. 1991, Qiu and Yandell 1998, Wang 1995, Wu and Chu 1993). The jump detection criteria in most of these papers have the form

$$M_n(x) = \frac{\sum_{i=1}^n w_1(x_i - x, h_n)Y_i}{\sum_{i=1}^n w_1(x_i - x, h_n)} - \frac{\sum_{i=1}^n w_2(x_i - x, h_n)Y_i}{\sum_{i=1}^n w_2(x_i - x, h_n)}, \text{ for } x \in [h_n/2, 1 - h_n/2], \tag{3}$$

where $\{w_1(x_i - x, h_n)\}$ are weights for design points in the right neighborhood $(x, x+h_n/2)$ of x, and $\{w_2(x_i - x, h_n)\}$ are weights for design points in the corresponding left neighborhood $(x - h_n/2, x)$. Then, $M_n(x)$ is a difference of two *one-sided* local weighted averages. Intuitively, $|M_n(x)|$ would be large if x is a jump point, and small otherwise. So, if we know that there is only one jump point (i.e., $p = 1$) in $[0, 1]$, then the jump point s_1 can be estimated by the maximizer of $|M_n(x)|$ over $x \in [h_n/2, 1 - h_n/2]$, denoted as $\widehat{s}_1$, and d_1 can be estimated by $M_n(\widehat{s}_1)$. In cases when $p > 1$ and p is known, the jump positions $\{s_j, j = 1, 2, \ldots, p\}$ and the jump magnitudes $\{d_j, j = 1, 2, \ldots, p\}$ can be estimated in a similar way (cf., Zhang et al. 2009). In different papers mentioned above, the weights $\{w_1(x_i - x, h_n)\}$ and $\{w_2(x_i - x, h_n)\}$ in (3) are defined differently. In some papers, they are defined based on local constant or linear kernel smoothing (e.g., Qiu 1991, Loader 1996). In some others, they are defined based on local least squares estimation, wavelets, or other smoothing approaches (e.g., Qiu and Yandell 1998, Wang 1995). The cases when the number of jumps is unknown were recently discussed in Xia and Qiu (2015), where a *jump information criterion* was proposed for estimating the number of jumps.

To estimate the regression function $f(x)$ with possible jumps preserved, there are two possible approaches proposed in the literature. In the first approach, the jump part of $f(x)$, i.e., $\sum_{j=1}^p d_j I(x > s_j)$, is first estimated based on jump detection. Then, the estimated jump part $\sum_{j=1}^{\widehat{p}} \widehat{d}_j I(x > \widehat{s}_j)$ is deducted from the observations. Namely, we define

$$y_i^* = y_i - \sum_{j=1}^{\widehat{p}} \widehat{d}_j I(x_i > \widehat{s}_j), \quad \text{for } i = 1, 2, \ldots, n.$$

Finally, the continuity part $g(x)$ of $f(x)$ can be estimated as usual from the modified data $\{y_i^*, i = 1, 2, \ldots, n\}$. See Joo and Qiu (2009) for a detailed discussion. In the second approach, the jumps do not need to be detected explicitly. Instead, the estimator of $f(x)$ can be obtained from a single formula. Interested readers can see papers such as Qiu (2003) and Gijbels et al. (2007).

As a side note, besides jumps in the regression function $f(x)$ itself, JRA can also handle jumps in the first-order or higher-order derivatives of $f(x)$, although jumps in the 3rd-order or higher-order derivatives are usually difficult to interpret in practice. For related discussions, see Zhang et al. (2009) and Joo and Qiu (2009).

Two-dimensional (2-D) JRA: In 2-D cases, the regression model becomes

$$Z_i = f(x_i, y_i) + \varepsilon_i, \; i = 1, 2, \ldots, n, \tag{4}$$

where n is the sample size, $\{(x_i, y_i), i = 1, 2, \ldots, n\}$ are the design points in the design space, f is the 2-D regression function, $\{Z_i, i = 1, 2, \ldots, n\}$ are n observations of the response variable Z, and $\{\varepsilon_i, i = 1, 2, \ldots, n\}$ are random errors. For simplicity, we assume that the design space is the unit square $[0, 1] \times [0, 1]$. In such cases, jump positions of f are curves in the design space, which are called *jump location curves (JLCs)*. Again, due to the importance of jumps, 2-D JRA mainly focuses on estimating JLCs and on estimating f with the jumps preserved, which are often referred to as *2-D jump detection* and *jump-preserving surface estimation* in the literature (cf., Qiu 2005, Chapters 4 and 5).

Early 2-D jump detection methods are based on the assumptions that the number of JLCs is known and the JLCs are smooth (Hall and Rau 2000, Hall et al. 2001, Korostelev and Tsybakov 1993, Müller and Song 1994, O'Sullivan and Qian 1994, Qiu 1997, Wang 1998). These methods are usually the generalized versions of their 1-D counterparts, based on estimation of certain first-order directional derivatives of f. As pointed out in Qiu (2002), the methods listed above treat JLCs as curves and try to estimate them by curves. Because of the global nature of curves in the sense that any two points on a given curve are connected through other points on the same curve, estimation of curves by curves is often challenging and it requires restrictive assumptions, such as the one that there is only one JLC. Of course, in reality, such assumptions are hardly valid. To overcome that difficulty, Qiu and Yandell (1997) started to describe the JLCs as *a pointset* in the design space, and suggested estimating the JLCs by another pointset in the same design space. Since points in a pointset need not form curves, the treatment of the JLCs as a pointset makes estimation of arbitrary JLCs possible. Different 2-D jump detectors have been proposed based on that idea in papers, such as Kang and Qiu (2014), Qiu (2002), Qiu and Yandell (1997), and Sun and Qiu (2007). Hall et al. (2008) suggested a tracking method for estimating arbitrary JLCs.

To estimate the regression surface $f(x, y)$ in model (4), two types of jump-preserving surface estimation methods have been proposed in the literature, as in 1-D cases. The first type of methods estimates $f(x, y)$ after jumps are detected (cf., Qiu 1998). Around the detected jumps, the surface estimator at a given point is often defined by a weighted average of the observations whose design points are located on the same side of the estimated JLC as the given point in a neighborhood of the point. Thus, potential jumps are preserved in the estimated surface. The second type of methods estimates the surface without detecting the jumps explicitly. Instead, they usually accommodate some information about jumps in the observed data directly in surface estimation. See papers such as Gijbels et al. (2006), Qiu (2004, 2009), and Qiu and Mukherjee (2010).

Connections between JRA and SPC: In manufacturing industries, we usually monitor a set of p quality variables of a production process over time. If we are interested in monitoring the mean vector of the quality variables, then we expect that the mean vector remains unchanged when the process is in-control (IC). If the mean vector has a shift, then we should stop the production process as soon as possible. To detect a shift, a control chart is often used. See, for instance, Hawkins and Olwell (1998), Montgomery (2012) and Qiu (2014). Figure 3 presents a typical univariate dataset (plot (a)) and a conventional CUSUM chart (plot (b)). The CUSUM chart gives a signal of process mean shift at $n = 35$.

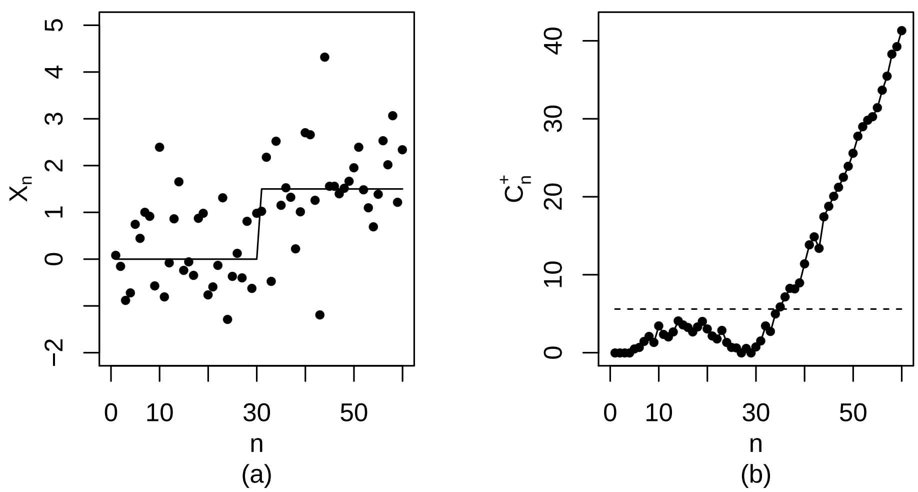


Figure 3: (a) Dot points denote observations from a production process with its mean function denoted by the solid line. A mean shift occurs at $n = 31$. (b) A conventional CUSUM chart with its control limit presented by the horizontal dashed line.

By comparing the traditional SPC problem and the JRA problem, we can see that they are connected in the sense that jump detection is central in both of them. Of course, these two problems are also substantially different. Their major differences can be summarizes as follows. (i) In the traditional SPC problem, the mean function is often a step function and it is a constant when the process is IC, while the mean function could be an arbitrary continuous curve between two consecutive jump points in the JRA problem. Because of these features of the two problems, *all* history data should be used in the SPC problem while the jumps are usually detected based on *local* information in the JRA problem. In this sense, the traditional SPC problem is more related to the piecewise linear regression modeling (e.g., Hinkley 1971, Worsley 1983), which could be regarded as a special case of JRA. As a matter of fact, this connection was explored when we suggested a control chart based on change-point detection in Hawkins et al. (2003). (ii) In Phase II SPC, we are usually interested in detecting the first shift because the quality of the products has deteriorized after the first shift, while there could be multiple jumps in a JRA model. Of course, there could be multiple shifts in Phase I SPC. But, another major concern in Phase I SPC, besides shift detection, is outlier detection that is usually not a major concern in JRA. (iii) One fundamental difference between the Phase II SPC problem and the JRA problem is that the former is a sequential monitoring problem in the sense that the sample size keeps increasing, while the latter has a fixed sample size. Because of this and other differences mentioned above, strategies to construct jump/shift detection schemes and evaluate their performance are different in the two problems.

In this big data era, SPC has more and more applications. Many such applications are beyond the traditional manufacturing industries. For instance, in public health and medical studies, we are interested in monitoring the incidence rates of certain diseases (e.g., infectious diseases, cancers, and chronic diseases) over time. In environmental studies, we are interested in monitoring the air quality and other environmental conditions longitudinally. In cybersecurity, we are interested in monitoring internet traffic or occurrence of certain cyber attacks. In all these applications, the IC process distribution could change over time, the number of jumps/shifts could be more than one in a specific time period, and the process cannot be stopped after a shift is detected. In these sequential monitoring problems, the traditional SPC methods need to be modified or extended. To this end, JRA would be helpful because the jumps and other singularities should be properly accommodated when we model all available data by the current time point for detecting trend shifts in future observations. In cases when we are interested in monitoring images, because JRA is a powerful tool for analyzing image data, it is especially important, which will be discussed in the next two sections. Therefore, there is a great potential to develop new SPC methodologies based on the ideas and existing research in JRA, to meet the great demand in monitoring processes with complicated data structures.

As a side note, there has been some recent research in modifying traditional SPC charts to accommodate varying IC process distribution for applications of disease early detection and disease surveillance. See references such as Fricker (2013), Li and Qiu (2016), Qiu and Xiang (2014, 2015), and Zhang et al. (2015, 2016).

3 Some Image Processing Topics and Methods

Images are widely used nowadays in various applications. They have become one of the most common data formats in practice. Not so long ago, we could only acquire 2-D images. Today, because of the rapid progress in image acquiring devices, 3-D images are routinely used in certain applications (e.g., MRI images in medical studies). Figure 4 shows a 3-D MRI brain image and its three 2-D slices from three different directions. In this section, we introduce some basic concepts, descriptions, and methods about images and image processing.

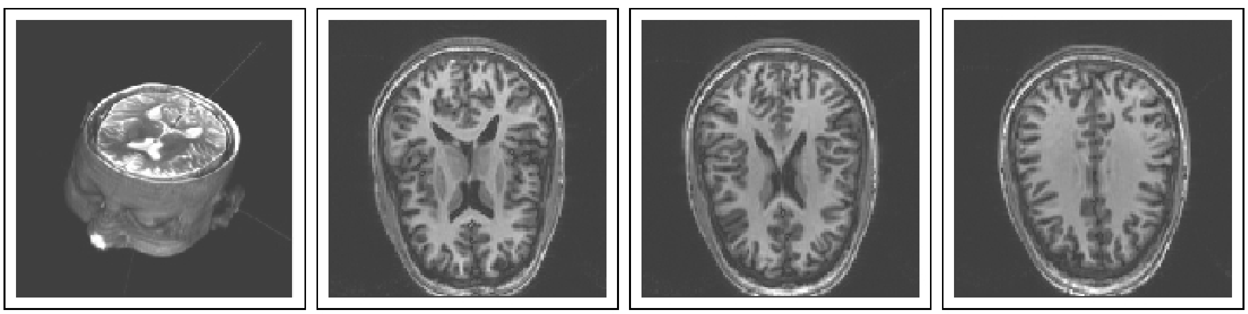


Figure 4: A 3-D image and its three slices along three different directions.

Images and their descriptions: For simplicity of presentation, our discussion will focus mainly on 2-D images, and many statements can be extended to 3-D cases without much difficulty. A typical 2-D gray-scale image can be described by the following 2-D JRA model:

$$Z_{ij} = f(x_i, y_j) + \varepsilon_{ij},\ i = 1, 2, \ldots, n_1;\ j = 1, 2, \ldots, n_2, \tag{5}$$

where (x_i, y_j) is the (i, j)th pixel, $f(x_i, y_j)$ is the true image intensity level at (x_i, y_j), ε_{ij} is the pointwise noise, and Z_{ij} is the observed image intensity level at (x_i, y_j). Model (5) is similar to model (4). Their only difference is that the design points (i.e., the pixels) in model (5) are regularly spaced in n_1 rows and n_2 columns while the design points in model (4) are located more arbitrarily in the design space. In that sense, model (4) is more general. But, model (5) is appropriate for describing images because pixels of a typical image are regularly spaced. From the images shown in Figure 4, it can be seen that the image intensity function $f(x, y)$ has jumps at various different places, including the outlines of image objects. In the image processing literature, positions at which $f(x, y)$ has jumps are called *step edges*, and positions at which its first-order derivatives have jumps are called *roof edges*. So, 2-D JRA can provide a powerful statistical tool for edge detection and many other tasks in image processing (cf., Qiu 2005).

In the image processing literature, people traditionally describe an image as a *Markov random field (MRF)*. In that framework, observed image intensities of an image are assumed to have the Markov property that the observed intensity at a given pixel depends only on the observed intensities in a neighborhood of the given pixel. It has been proved that the random field has a Gibbs distribution in that case (e.g., Besag 1974). If the true image intensities are assumed to have a prior distribution that is also a Gibbs distribution, then the posterior distribution of the true image would be a Gibbs distribution too and the true image can be estimated by the *maximum a posteriori (MAP)* estimator (cf., Geman and Geman 1984). Many image processing methods were developed in that framework.

In applied mathematics and computer sciences, people also use *diffusion equations* as a tool to handle images. In that framework, the process of noise removal from an observed image is regarded as a diffusion process described by a partial differential equation, and the observed image is regarded as the initial state of this process (e.g., Perona and Malik 1990). Many different methods have been proposed for edge-preserving image denoising and other tasks in image processing (e.g., Barash 2002, Sapiro 2001).

Among the three tools described above (i.e., JRA, MRF, and diffusion equations), JRA is the most flexible one for describing and handling images, although people often assume that the noise in model (5) is spatially independent. Based on the mechanism of image acquisition, people think that noise is a pointwise contamination and the spatial independence assumption is roughly true in many image applications. If spatial contamination is involved, then the corresponding problem is mainly handled by image deblurring methods (see related discussion in the next part). However, if spatial correlated noise is a concern in an application, then some statistical research in handling correlated data can be adopted here (e.g., Opsomer et al. 2001). The MRF framework is quite general and it has been used widely in the image processing literature for handling various problems. However, it is often difficult to check whether a given image has the Markov property in practice, and the performance of the methods based on MRF may not be robust to the Markovian assumption either. The diffusion equation approach is also quite versatile for solving many different image processing problems, although statisticians usually avoid using partial differential equations because derivatives are sensitive to noise.

Some image processing problems and methods: Image processing is a broad area, containing many different research problems. In this part, we only discuss some basic image processing problems with which we have some research experience and they might be related to SPC of images. For a more complete description about image processing, read books such as Gonzalez and Woods (2008).

Edge detection and edge-preserving image denoising are two fundamental problems in image processing. Edge detection is mainly for detecting edges accurately, while edge-preserving image denoising is for removing the noise and recovering the true image with possible edges preserved. Obviously, these problems are essentially the same as jump detection and jump-preserving surface estimation in 2-D JRA. See Qiu (2007) for a discussion about the connections and differences between the two areas. For edge detection, most existing methods are based on the facts that can be demonstrated by Figure 5. Plot (a) in the figure shows a cross-sectional 1-D profile of an image intensity surface around a slightly blurred step edge curve, and plots (b) and (c) show its first-order and second-order derivatives, respectively. A slightly blurred step edge is considered in plot (a) to make the profile differentiable. In practice, the jumps in image intensity around an edge curve could be sharp, and the directional derivatives of the image intensity function can be replaced by directional differences that are computed from local pixels. Based on plot (b), the absolute value of the first-order derivative of the image intensity function at the edge position would be large if the first-order derivative is computed along a direction that is perpendicular to the edge curve. Plot (c) shows that the second-order derivative of the image intensity function at the edge position would have the following *zero-crossing* property: (i) the second-order derivative equals zero at the edge position, and (ii) its sign changes at two different sides of the edge curve. It has been well demonstrated in the literature that edge detectors based on the first-order derivatives are more effective in detecting edges and less sensitive to noise, compared to the edge detectors based on the second-order derivatives. However, the detected edges by the former are thicker in the sense that they are located in a wider band of the true edge curves, and they may not form closed curves. In comparison, the detected edges by the latter are thinner and they tend to form closed curves (e.g., Canny 1986, Joo and Qiu 2009, Kang and Qiu 2014, Marr and Hildreth 1980, Qiu and Bhandarkar 1996, Sun and Qiu 2007, Torre and Poggio 1986).

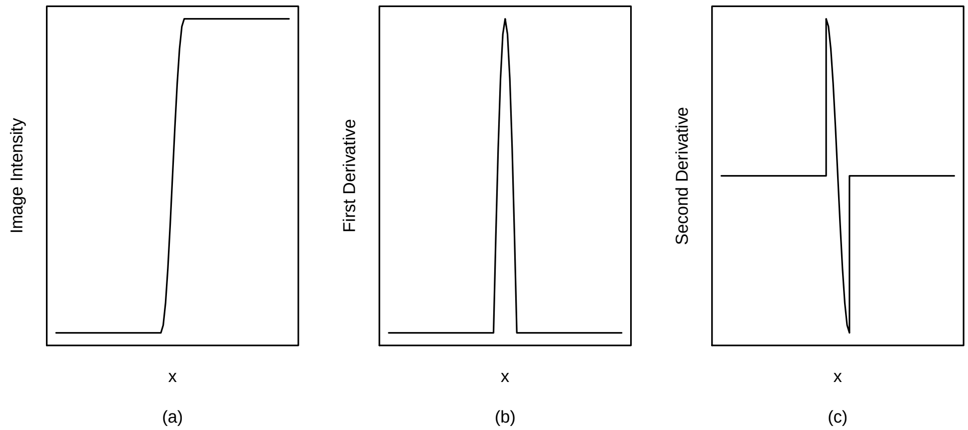


Figure 5: (a) A cross-sectional 1-D profile of an image intensity surface around a slightly blurred step edge curve. (b) The first-order derivative of the profile. (c) The second-order derivative of the profile.

In some applications, we need to segment one or more specific objects in an image to study their properties (e.g., size, shape, etc.). This is the *image segmentation* problem (cf., Gonzalez and Woods 2008). It should be pointed out that image segmentation is different from edge detection, although the two problems are related. First, edge detection is usually for an entire image, instead of for some specific objects in the image. So, the detected edges may not describe the outlines of the objects under consideration well. Second, the outlines of the image objects under consideration are most probably edges of the image. But, part of the outlines may not appear to be edges in the observed image due to lighting conditions, contrast in image intensity between the objects and their background, and so forth. Therefore, the detected edges are helpful for segmenting the specific objects of interest, but attention is needed to address the challenges of image segmentation due to the differences between the two problems. As a side note, effective image segmentation methods depend heavily on the nature of a specific application. That is because a successful image segmentation depends on many factors, including the number and shape of the objects to segment, complexity of their background, contrast in image intensity between the objects and their background, and so forth. All such factors differ in different applications. For segmenting foreground from background in microarray images, see our papers Qiu and Sun (2007, 2009) and the references cited therein.

For edge-preserving image denoising, besides the JRA methods discussed in Section 2 for jump-preserving surface estimation, there are many methods proposed under the framework of MRF or diffusion equations. To remove noise and preserve edges under the MRF framework, usually a line process is considered together with an observed image (or called intensity process). By a line process, it is assumed that there is a line element between two horizontally or vertically neighboring pixels, and the line element equals 1 if there is an edge between the two pixels and 0 otherwise. Then, the line process is estimated together with the true image using MAP or similar approaches (e.g., Besag 1986, Fessler et al. 2000, Geman and Geman 1984). Under the diffusion equation framework, a quantity called diffusivity controls the degree of smoothing at a given position in an image. To preserve edges when denoising an image, it should be chosen small at edge positions and relatively large at non-edge positions. The edge positions can be indicated roughly by the gradient of the observed image. For a more detailed discussion, see papers such as Perona and Malik (1990), Keeling and Stollberger (2002), Weickert (1998), and Mrázek et al. (2003). In the literature, there are some alternative methods for edge-preserving image denoising. For instance, image denoising based on wavelet transformation is also a popular approach, although selection of proper father and mother wavelets is often challenging to properly accommodate edges (e.g., Chang et al. 2000, Mrázek et al. 2003). Another alternative denoising method is based on adaptive weights smoothing. See a sequence of papers by Joerg Polzehl and his co-authors (e.g., Polzehl and Spokoiny 2000).

In model (5), we only consider pointwise image contamination (i.e., noise). In practice, another common image contamination is spatial blur, caused by a relative move between the camera and the scene to be pictured and other reasons. A blurred image of a bird is shown in the left panel of Figure 6. A research area called *image deblurring* tries to recover the true image from the observed but blurred one. Mathematically, the image deblurring problem can be formulated as follows:

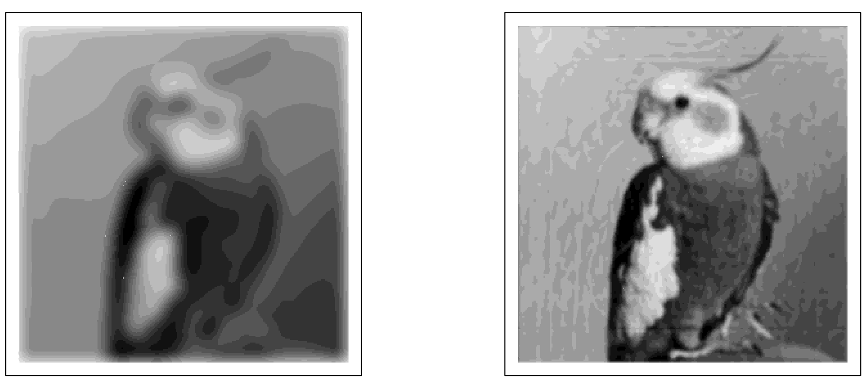
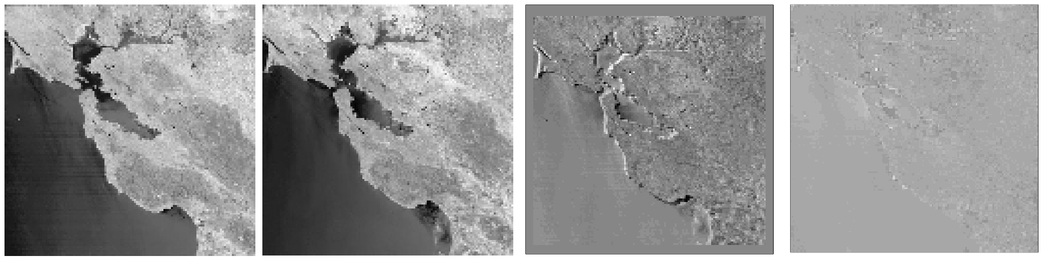


Figure 6: A blurred image and its deblurred version.

$$Z(x, y) = H\{f\}(x, y) + \varepsilon(x, y), \tag{6}$$

where $H\{f\}(x,y) = \int\int_{R^2} h(u,v)f(x-u,y-v)\, dudv$ is the convolution between a *point spread function (psf)* h and the true image intensity function f, and $\varepsilon(x,y)$ is the pointwise noise. The psf h describes how f is spatially contaminated (i.e., blurred) in the imaging process. Image deblurring is for estimating $f(x,y)$ from $Z(x,y)$. This problem is "ill-posed" in the sense that there could be multiple sets of h and f corresponding to the same Z, even when there is no noise in Z. So, in the early literature, h is often assumed to be known, and the main research effort is on estimating $f(x,y)$ from $Z(x,y)$ in such cases, using different inverse filtering algorithms (cf., Carraso 1999, Figueiredo and Nowak 2003, Skilling 1989). That task is also challenging because the inverse filtering is often numerically unstable because of the noise. In practice, it may not be realistic to assume h to be known. For instance, satellite images are often blurred because of wind, atmospheric turbulence, aberrations of the optical system, and other various reasons, and it is difficult to completely specify h in such cases. More recent research in this area is for estimating $f(x,y)$ when h is not completely specified. This problem is called *blind image deblurring*. One type of blind image deblurring methods assumes that h has a parametric form, and its parameters are estimated together with the true image in image deblurring (e.g., Carasso 2001, Joshi and Chaudhuri 2005, Katsaggelos and Lay 1990, Rajagopalan and Chaudhuri 1999). Of course, validity of such parametric forms for h is not guaranteed in practice. To overcome this difficulty, Hall and Qiu (2007a) suggested estimating h from an observed but blurred test image whose true structure is known. Then, the estimated h can be used for deblurring other observed images taken by the same imaging device (Hall and Qiu 2007b, Qiu 2008). A more flexible method without using any test images was recently proposed in Qiu and Kang (2015).

In different applications, we often want to compare two or more images of a same scene (e.g., a patient's brain) that are taken at different times. Because the relative positions between the camera and the scene cannot be exactly the same at different times in practice, the images should be geometrically aligned well before we can make a reliable comparison. As a demonstration, plots (a) and (b) in Figure 7 show two satellite images taken at the San Francisco bay area in 1990 and 1999, respectively. Their pixelwise difference is shown in plot (c). It can be seen that there is a clear pattern in this difference image, mainly because the two original images are not geometrically aligned well. The problem to geometrically align the images is called *image registration* in image processing. It is a fundamental task in many imaging applications. After the two images shown in Figure 7(a)-(b) are properly registered, their pixelwise difference is shown in Figure 7(d). Obviously, the pattern is much lighter now. Most existing image registration methods assume that the



(a) (b) (c) (d)

Figure 7: (a)-(b) Two satellite images taken at the San Francisco bay area in 1990 and 1999, respectively. (c) Their difference without image registration. (d) Their difference after image registration.

mapping transformation has a parametric form or satisfies certain regularity conditions (e.g., it is a smooth function with the first-order or higher order derivatives). They often estimate the mapping transformation globally by solving a global minimization/maximization problem or by using a global smoothing technique. Such global smoothing methods usually cannot preserve singularities (e.g., discontinuities) and other features of the mapping transformation well. Further, the ill-posed nature of the image registration problem, i.e., the mapping transformation is not well defined at certain places (e.g., at places where the true image intensity surfaces are straight), is undisclosed by such methods. Recently, we suggested handling the image registration problem locally, by first studying local properties of a mapping transformation. To this end, we suggested the concept of non-degenerate pixels. A local smoothing method for estimating the mapping transformation was proposed accordingly in Xing and Qiu (2011). Because of the flexibility of local smoothing, our method did not require many regularity conditions on the mapping transformation. In a recent paper (Qiu and Xing 2013a), this topic was further studied. Several concepts, including the 2-D degenerate pixels, 2-D partial degenerate pixels, 1-D degenerate pixels, and 1-D partial degenerate pixels, were proposed for describing the local properties of the mapping transformation. The relationship among these concepts and the statistical properties of the estimated transformation were also studied. In Qiu and Xing (2013b), we further demonstrated that non-degenerate pixels were ideal features for image registration, compared to some other commonly used features.

As pointed out at the beginning of this section, 3-D images get increasingly popular in various applications, including medical studies using MRI and functional MRI (fMRI) images. Generally speaking, the structure of a typical 3-D image is substantially more complicated than that of a typical 2-D image. For instance, edge locations are surfaces in 3-D cases, which would be much more challenging to handle, compared to edge curves in 2-D cases. So, in many cases, we cannot generalize 2-D methods for solving the 3-D image processing problems. We proposed a 3-D method for image denoising in Qiu and Mukherjee (2012), and discussed 3-D image registration in Song and Qiu (2017a,b).

4 Statistical Process Control of Images, Its Challenges and Some Research Problems

As discussed in Section 1, comparison and monitoring of images acquired from a longitudinal process have broad applications. However, the complex image structure and the sequential nature of image data streams pose significant analytic challenges on such tasks. The major challenges include at least the following ones. 1) Images have complicated structures, including edges and other singularities in the image intensity surface, and the spatial nature of the image data. Also, image streams are sequential, in the sense that their sample sizes keep increasing over time. Sequential data are usually challenging to analyze properly. 2) A typical 3-D MRI image has millions of voxels. If we are concerned about image monitoring, then hundreds or even thousands of images need to be handled. Therefore, data volume is big, and this requires our methods to be computationally efficient. 3) There are some new features in the monitoring of image data, compared to the traditional SPC problems. For instance, the process IC distribution may change over time due to seasonality and other reasons. Therefore, when we develop methodologies for image comparison and monitoring, these challenges should be handled properly. This and some other related topics are discussed in this section.

Image pre-processing: Observed images acquired from a longitudinal process often contain different types of contamination, including geometric misalignment among images, pointwise noise in individual images, spatial blur, and some others. Figure 7 shows that when geometric misalignment among images exists, proper image registration is critically important for reliable image comparison and monitoring. In the image processing literature, proper processing of the observed images to make the subsequent image analysis more reliable is called *image pre-processing* (cf., Sonka et al. 1993, Chapter 4), which is common in different applications of image processing and analysis.

In industrial applications, the geometric misalignment among different images might be mainly due to relative movement between a camera and products to be pictured. In such cases, the following *rigid-body transformation* should be appropriate:

$$\begin{aligned} T_1(x,y) &= x\cos(\phi) + y\sin(\phi) + \Delta x, \\ T_2(x,y) &= -x\sin(\phi) + y\cos(\phi) + \Delta y, \end{aligned} \tag{7}$$

where ϕ is a rotation parameter, Δx and Δy are the translation parameters in the $x-$ and $y-$axes, respectively, and $\mathbf{T}(x,y) = (T_1(x,y), T_2(x,y))$ defines the geometric transformation to map the point (x,y) in one image to $\mathbf{T}(x,y)$ in another image. By a rigid-body transformation, the Euclidean distance between any two points on an image will not change after the transformation. Then, the major goal of image registration is to estimate the parameters $(\phi, \Delta x, \Delta y)$ from the two observed images. Estimation of $(\phi, \Delta x, \Delta y)$ and the properties of the estimators have been discussed recently in Feng and Qiu (2017).

Besides image registration, image denoising and image debluring might also be important for comparison of different images acquired from a longitudinal process and for monitoring of image streams as well. That is because pointwise noise and spatial blur are common in observed images and it might make image comparison and monitoring more effective if they can be removed to a certain degree in the image pre-processing stage. To this end, the edge-preserving image denoising methods and the image deblurring methods discussed in Section 3 and the related 2-D JRA methods discussed in Section 2 should be helpful.

There are many open research problems to address related to image pre-processing for image comparison and monitoring. For instance, when we monitor a sequence of images, what is the appropriate way to formulate the image registration problem? More specifically, should we always register the observed image at the current time point with the first image in the sequence, or with a certain average of the images at all previous time points? For image denoising and deblurring, should we accommodate them in image comparison and monitoring, or make these image pre-processings separately before image comparison and monitoring? Intuitively, a rough registration of images (e.g., using the simple rigid-body transformation (7)) can improve the results in image comparison and monitoring quite dramatically, and any further fine tuning in image registration might not generate a meaningful improvement, which needs to be confirmed in future research.

How to compare and monitor images? In the literature, there is a limited discussion about image comparison, mainly by researchers in computer vision and graphics (e.g., Davis et al. 1997). A major existing tool for comparing two images is to use a quantitative similarity or dissimilarity measure (Freire et al. 2002). Commonly used measures include the mean squared difference between two images, Pearson's correlation coefficient of the observed image intensities of two images, entropy of the difference between two images, and so forth (Qiu and Nguyen 2008). However, these measures alone cannot tell us whether two related images are significantly different or not, especially in cases when image misalignment is present. They do not take into account the complicated image structure, such as edges, either. Because observed images almost always contain noise and other contamination (e.g., spatial blur), image comparison is a statistical problem and it should be addressed in the framework of hypothesis testing or other statistical inferences. This problem has not been well discussed yet in the statistical literature, and requires much future research effort.

Because images have been widely used in manufacturing industries in recent years, mainly for quality control purposes, there is some existing research on sequential monitoring of image data streams in the chemical and industrial engineering literature (e.g., Megahed et al. 2011, Prats-Montalban 2014, Yan et al. 2015). Most existing methods for image monitoring proceed in two steps. They first extract some features from each image, using methods such as the principal component analysis (PCA), and then monitor the extracted features by a conventional control chart (e.g., Duchesne et al. 2012, Lin et al. 2008). Some other methods focus on certain pre-specified regions, called regions of interest (ROIs), in individual images, and then monitor the images by a control chart constructed based on a summary statistic (e.g., the mean image intensity) of the ROIs (e.g., Jiang et al. 2011, Megahed et al. 2012). These methods make a good start in our research effort to solve the image monitoring problem. But, they have some limitations for us to address in our future research as well. For instance, The first type of methods mentioned above ignores the spatial data structure of images completely, while the second type considers the spatial data structure partially in the pre-specified ROIs only. All these methods do not take into account the edges and other complicated structure in the image intensity surfaces. They usually do not properly accommodate geometric misalignment among different images either.

To properly compare images or monitor an image data stream, we first need to decide what specific features or parts of the images that we should compare or monitor. In some applications, we have a clear idea about what to compare or monitor. For instance, in the steel surface monitoring problem demonstrated in Figure 1, we are usually concerned about the cracks and other specific defectives (Yan et al. 2017). In such cases, some edge detection and image segmentation methods can be applied for figuring out the size/shape/location or other summary statistics of these defectives for image comparison and monitoring. To this end, jump detection methods in 2-D JRA should be extremely useful. In many other applications, we may not have a clear idea about what to compare or monitor. In such cases, the following possible approaches can be considered, as in the image registration literature (e.g., Klein et al. 2009, Qiu and Xing 2013a,b, Zitova and Flusser 2003): (i) feature-based approach, and (ii) intensity-based approach. By a feature-based approach, we first extract some specific features from the observed images and then compare or monitor them for image comparison and monitoring purposes. Commonly used features in the image processing include landmarks or control points that can be selected manually or automatically by a computer (Wu et al. 2006), edge lines or curves that are often detected by gradient-based methods (Hsieh et al. 1997), regions, centroids or templates that are usually determined by ways of thresholding and segmentation (Saeed 1998), and the so-called degenerate pixels discussed in Qiu and Xing (2013a,b). In SPC of images, the features should be able to measure the quality of products well. To this end, the features defined in the image processing literature may not be the most appropriate ones to use, and new features might need to be defined for a specific application. This topic requires much future research effort.

It should be pointed out that there are many challenges in developing and applying a feature-based method. For instance, we need to decide which feature type or types we should use, how many features should be extracted, how to extract the features, and so forth. All these decisions would involve many subjective choices. To avoid these challenges in using a feature-based method, we can consider using intensity-based approaches. By an intensity-based approach, we use all observed intensities in the related images for image comparison and monitoring. However, such approaches would have their own challenges. For instance, because of the relatively large variation of image intensities around the edges, the edge structure of the related images would play a dominating role in image comparison and monitoring if all observed image intensities are treated equally. To overcome this difficulty, we recently suggest dividing an image into two parts that consist of edge pixels and non-edge pixels, and then compare two images using their edge pixels, or non-edge pixels, or both (Feng and Qiu 2017). In intensity-based image comparison and monitoring, the jump-preserving surface estimation methods discussed in Section 2 and the edge-preserving image denoising methods discussed in Section 3 should be relevant.

Some new issues with image comparison and monitoring: As mentioned at the beginning of Section 4, image data often have a big data volume and a typical 3-D image has millions of voxels. So, computation would be a big issue for online image monitoring. To this end, recursive formulas for computing the charting statistic would be important, as in other sequential monitoring problems (e.g., Hawkins et al. 2003, Li and Qiu 2016). Another possible approach is to apply the dimension reduction or variable selection methods to the SPC problem (e.g., Capizzi and Masarotto 2011, Wang and Jiang 2009, Zou and Qiu 2009). However, the spatial information of images, including their edges and other structures, should be properly accommodated, which may not be straightforward. In our opinion, recent research in fast computing should also be considered in the SPC community. For instance, it has been shown that distributed parallel computing can reduce the computing time dramatically for 3-D image registration (Song and Qiu 2017c). Such fast computing platforms should be helpful for online image monitoring.

For some image monitoring problems, the IC distribution of an image stream might change over time. For instance, the sequence of satellite images of a specific region (cf., Figure 7 for satellite images of the San Francisco bay area) may reveal seasonality and other longitudinal variation that could be part of the IC process. In such cases, the time varying IC distribution needs to be properly accommodated in process monitoring. See Qiu and Xiang (2014, 2015) for a related discussion in the dynamic screening system context. In this regard, there are many challenges to address, including proper estimation of the time varying IC distribution, accommodation of seasonality and other covariates, accommodation of spatial and/or temporal correlation in image data, and so forth.

SPC of images and profile monitoring: In SPC, profile monitoring receives much attention from researchers in recent years (Qiu 2014, Chapter 10). By profile monitoring, we are mainly concerned about the relationship between two or more variables. Early research in this area focuses on univariate linear or parametric profile monitoring (e.g., Jin and Shi 1999, Kang and Albin 2000, Zou et al. 2007). In such cases, the profile monitoring problem can be described by the following model:

$$Z_{ij} = f(x_{ij}, \boldsymbol{\theta}_i) + \varepsilon_{ij}, \qquad \text{for } j = 1, 2, \ldots, n_i, \;\; i = 1, 2, \ldots, \tag{8}$$

where Z is a response variable, x is a predictor, $f(x, \boldsymbol{\theta})$ is a known parametric function with parameter vector $\boldsymbol{\theta}$, (x_{ij}, Z_{ij}) is the jth observation of the ith product to monitor, and ε_{ij} is the error term. In this profile monitoring problem, we mainly monitor the sequence $\{\boldsymbol{\theta}_i, i = 1, 2, \ldots\}$ (after $\{\boldsymbol{\theta}_i\}$ are estimated properly) to make sure that it is stable when i increases. In practice, the parametric model (8) may not describe the true relationship between x and Z well. So, recent research in this area focuses more on nonparametric profile modeling. A number of nonparametric profile monitoring approaches have been proposed in the literature (e.g., Qiu et al. 2010, Qiu and Zou 2010, Zou et al. 2008, 2009). In some applications, there are multiple response variables involved (i.e., Z in (8) is a vector). This multiple profile monitoring problem has been discussed recently by a number of papers (e.g., Paynabar et al. 2013, 2016).

From the description about image processing in Section 3, we can see that images can be regarded as profiles. For simplicity, let us focus on 2-D images here, and 3-D images can be discussed in a similar way. A sequence of images can be described by the following model:

$$Z_{ijk} = f_k(x_{ik}, y_{jk}) + \varepsilon_{ijk},\ i = 1, 2, \ldots, n_1;\ j = 1, 2, \ldots, n_2;\ k = 1, 2, \ldots, \tag{9}$$

where (x_{ik}, y_{jk}) is the (i, j)th pixel of the kth image, $f_k(x, y)$ is the true image intensity function of the kth image, $\{Z_{ijk}\}$ is the kth observed image, and ε_{ijk} is the pointwise noise. So, $\{f_k(x, y), k = 1, 2, \ldots\}$ can be regarded as 2-D profiles, and monitoring of the image sequence is equivalent to monitoring of the 2-D profile sequence. By comparing (9) with the univariate profile monitoring problems (cf., (8)) and the multiple profile monitoring problem discussed in the previous paragraph, we can see that there are two location variables x and y in (9) and only one predictor x in the latter problems. Because of this difference, the two types of profile monitoring problems are substantially different in several aspects, including the following ones. First, in the univariate profile monitoring problems and the multiple profile monitoring problem discussed, the profiles are curves, and they are surfaces in the image monitoring problem (9). Surfaces usually have much more complicated structure than that of curves. Second, edges and other singularities are important parts of images. But, in the existing research on profile monitoring, profiles are usually assumed to be continuous functions. Third, in some image processing applications (e.g., applications involving MRI images), noise is often non-additive, due to the specific image acquisition techniques (cf., Mukherjee and Qiu 2013). Also, observed images could be blurred, or contain other types of contaminations, as discussed in Section 3. The existing profile monitoring research usually does not handle such issues. For these reasons, we may not be able to solve the image monitoring problem properly by simply using the existing methods in the profile monitoring literature. Instead, new methods should be developed. To this end, the JRA and image processing methods discussed in the previous two sections should be helpful.

In the literature, there are some recent discussions about surface monitoring, motivated by different applications (e.g., Colosimo et al. 2014, del Castillo et al. 2014, Wang et al., 2014, Xia et al. 2008, Zang and Qiu 2017). These surface monitoring problems are relevant to image monitoring, because images can be regarded as surfaces of the image intensity functions and proper surface registration also need to be addressed properly in the surface monitoring problems. However, one fundamental difference between the two types of problems is that the image intensity surfaces usually have jumps and other singularities (e.g., step and roof edges), but the surfaces discussed in the surface monitoring literature are often assumed to be continuous. Because of this difference, we cannot apply the existing surface monitoring methods to the image monitoring problem directly, although some aspects of the former (e.g., registration and denoising) might be useful for solving the image monitoring problem.

5 Concluding Remarks

In this big data era, SPC has found more and more applications in engineering, environmental research, medicine, public health, and other disciplines and areas (Qiu 2016). Many such applications involve image data streams. Because images usually have a big data volume and a complicated data structure, proper comparison and monitoring of images are challenging. To address different challenges with image data, existing research in image processing and jump regression analysis should be helpful. In this paper, we have briefly introduced some basic research problems and methods in these two areas and discussed how the related methods can be potentially used for solving the image comparison and monitoring problems. It should be pointed out that the image comparison and monitoring problems are mostly open, and they need much future research effort.

Acknowledgments: The author appreciates the organizing committee of the 2017 Stu Hunter Research Conference for providing an opportunity to present this work in the conference. The guest-co-editor, Professor Giovanna Capizzi, and two referees provided many constructive comments and suggestions in the review process, which improved the quality of the paper greatly. This research is supported in part by an NSF grant.

References

Balageas, D., Fritzen, C.P., and Güemes, A. (2006), *Structural Health Monitoring*, London, UK: ISTE Ltd.

Barash, D. (2002), "A fundamental relationship between bilateral filtering, adaptive smoothing, and the nonlinear diffusion equation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **24**, 844–847.

Besag, J. (1974), "Spatial interaction and the statistical analysis of lattice systems (with discussions)," *Journal of the Royal Statistical Society (Series B)*, **36**, 192–236.

Canny, J. (1986), "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **8**, 679–698.

Capizzi, G., and Masarotto, G. (2011), "A least angle regression control chart for multidimensional data," *Technometrics*, **53**, 285–296.

Carasso, A.S. (2001), "Direct blind deconvolution," *SIAM Journal on Applied Mathematics*, **61**, 1980–2007.

Chang, G.S., Yu, B., and Vetterli, M. (2000), "Spatially adaptive wavelet thresholding with context modeling for image denoising," *IEEE Transactions on Image Processing*, **9**, 1522–1531.

Colosimo, B.M., Cicorella, P., Pacella, M., and Blaco, M. (2014), "From profile to surface monitoring: SPC for cylindrical surfaces via Gaussian processes," *Journal of Quality Technology*, **46**, 95–113.

Davis, M.H., Khotanzad, A., Flamig, D.P., and Harms, S.E. (1997), "A physics-based coordinate transformation for 3-D image matching," *IEEE Transactions on Medical Imaging*, **16**, 317–328.

del Castillo, E., Colosimo, B.M., and Tajbakhsh, S. (2014), "Geodesic Gaussian processes for the reconstruction of a free-form surface," *Technometrics*, **57**, 87–99.

Duchesne, C., Liu, J., and MacGregor, J. (2012) "Multivariate image analysis in the process industries: A review," *Chemometrics and Intelligent Laboratory Systems,* **117**, 116–128.

Fan, J., and Gijbels, I. (1996), *Local Polynomial Modelling and Its Applications,* New York: Chapman and Hall.

Feng, L., and Qiu, P. (2017), "Difference detection between two images for image monitoring," *Technometrics*, in press.

Fessler, J.A., Erdogan, H., and Wu, W.B. (2000), "Exact distribution of edgepreserving MAP estimators for linear signal models with Gaussian measurement noise," *IEEE Transactions on Image Processing*, **9**, 1049–1055.

Figueiredo, M.A.T. and Nowak, R.D. (2003), "An EM algorithm for waveletbased image restoration", *IEEE Transactions on Image Processing,* **12**, 906–916.

Freire, L., Roche, A., and Mangin, J.-F. (2002), "What is the best similarity measure for motion correction in fMRI times series?" *IEEE Transactions on Medical Imaging*, **21**, 470–484.

Fricker, R.D. (2013), *Introduction to Statistical Methods for Biosurveillance: With an Emphasis on Syndromic Surveillance*, New York: Cambridge University Press.

Geman, S., and Geman, D. (1984), "Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **6**, 721–741.

Gijbels, I., Lambert, A., and Qiu, P. (2006), "Edge-preserving image denoising and estimation of discontinuous surfaces," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* **28**, 1075–1087.

Gijbels, I., Lambert, A., and Qiu, P. (2007), "Jump-preserving regression and smoothing using local linear fitting: a compromise," *Annals of the Institute of Statistical Mathematics*, **59**, 235–272.

Gonzalez, R.C., and Woods, R.E. (2008), *Digital Image Processing (3rd edition)*, Prentice Hall.

Hall, P., Peng, L., and Rau, C. (2001), "Local likelihood tracking of fault lines and boundaries," *Journal of the Royal Statistical Society (Series B)*, **63**, 569–582.

Hall, P., and Qiu, P. (2007a), "Nonparametric estimation of a point spread function in multivariate problems," *Annals of Statistics*, **35**, 1512–1534.

Hall, P., and Qiu, P. (2007b), "Blind deconvolution and deblurring in image analysis," *Statistica Sinica*, **17**, 1483–1509.

Hall, P., Qiu, P., and Rau, C. (2008), "Edge, corners and vertex estimation for images and regression surfaces," *Scandinavian Journal of Statistics*, **35**, 1–17.

Hall, P., and Rau, C. (2000), "Tracking a smooth fault line in a response surface," *Annals of Statistics*, **28**, 713–733.

Hawkins, D.M., and Olwell, D.H. (1998), *Cumulative Sum Charts and Charting for Quality Improvement*, New York: Springer-Verlag.

Hawkins, D.M., Qiu, P., and Kang, C.W. (2003), "The changepoint model for statistical process control," *Journal of Quality Technology*, **35**, 355–366.

Hinkley, D.V. (1971), "Inference in two-phase regression," *Journal of the American Statistical Association*, **66**, 736–743.

Hsieh, J.W., Liao, H.Y.M., Fan, K.C., Ko, M.T., and Hung, Y.P. (1997), "Image registration using a new edge-based approach," *Computer Vision and Image Understanding*, **67**, 112–130.

Jiang, W., Han, S.W., Tsui, K.L., and Woodall, W.H. (2011), "Spatiotemporal surveillance methods in the presence of spatial correlation," *Statistics in Medicine*, **30**, 569–583.

Jiang, B., Wang, C.C., and Liu, H.C. (2005). "Liquid crystal display surface uniformity defect inspection using analysis of variance and exponentially weighted moving average techniques," *International Journal of Production Research*, **43**, 67–80.

Jin, J., and Shi, J. (1999), "Feature-preserving data compression of stamping tonnage information using wavelets," *Technometrics*, **41**, 327–339.

Jin, N., Zhou, S., and Chang, T.S. (2004), "Identification of impacting factors of surface defects in hot rolling processes using multi-level regression analysis," *Transactions of the North American Manufacturing Research Institute of SME*, **32**, 557–564.

Joo, J., and Qiu, P. (2009), "Jump detection in a regression curve and its derivative," *Technometrics*, **51**, 289–305.

Joshi, M.V. and Chaudhuri, S. (2005), "Joint blind restoration and surface recovery in photometric stereo," *Journal of the Optical Society of America, Series A,* **22**, 1066–1076.

Kahn, C.E., Carrino, J.A., Flynn, M.J., Peck, D.J., and Horii, S.C. (2007), "DICOM and radiology: past, present, and future," *Journal of the American College of Radiology*, **4**, 652–657.

Kang, L., and Albin, S.L. (2000), "On-line monitoring when the process yields a linear profile," *Journal of Quality Technology*, **32**, 418–426.

Kang, Y., and Qiu, P. (2014), "Jump detection in blurred regression surfaces," *Technometrics*, **56**, 539–550.

Katsaggelos, A.K. and Lay, K.-T. (1990), "Image identification and image restoration based on the expectation-maximization algorithm," *Optical Engineering*, **29**, 436–445.

Keeling, S.L., and Stollberger, R. (2002), "Nonlinear anisotropic diffusion filters for multiscale edge enhancement," *Inverse Problems*, **18**, 175–190.

Klein, A., Andersson, J., Ardekani, B.A. et al. (2009), "Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration," *NeuroImage*, **46**, 786–802.

Korostelev, A.P., and Tsybakov, A.B. (1993), *Minimax Theory of Image Reconstruction*, Lecture Notes in Statistics, **82**, New York: Springer-Verlag.

Li, J., and Qiu, P. (2016), "Nonparametric dynamic screening system for monitoring correlated longitudinal data," *IIE Transactions*, **48**, 772–786.

Loader, C.R. (1996), "Change point estimation using nonparametric regression," *Annals of Statistics*, **24**, 1667–1678.

Marr, D., and Hildreth, E. (1980), "Theory of edge detection," *Proceedings of the Royal Society in London*, **B207**, 187–217.

Megahed, F.M., Woodall, W.H., and Camelio, J. A. (2011), "A review and perspective on control charting with image data," *Journal of Quality Technology*, **43**, 83–98.

Megahed, F.M., Wells, L.J., Camelio, J.A., and Woodall, W.H. (2012), "A spatiotemporal method for the monitoring of image data," *Quality Reliability and Engineering International*, **28**, 967–980.

Montgomery, D.C. (2012), *Introduction to Statistical Quality Control,* New York: John Wiley & Sons.

Mrázek, P., Weickert, J., and Steidl, G. (2003), "Correspondences between wavelet shrinkage and nonlinear diffusion," in *Scale-Space 2003* (edited by Griffin, L.D., and Lillholm, M.), 101–116.

Mukherjee, P.S., and Qiu, P. (2013), "Efficient bias correction for MRI image denoising," *Statistics in Medicine*, **32**, 2079–2096.

Müller, H.G. (1992), "Change-points in nonparametric regression analysis," *Annals of Statistics*, **20**, 737–761.

Müller, H.G., and Song, K.S. (1994), "Maximin estimation of multidimensional boundaries," *Journal of Multivariate Analysis*, **50**, 265–281.

Opsomer, J., Wang, Y., and Yang, Y. (2001), "Nonparametric regressin with correlated errors," *Statistical Science*, **16**, 134–153.

OSullivan, F., and Qian, M. (1994), "A regularized contrast statistic for object boundary estimation - implementation and statistical evaluation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **16**, 561–570.

Parker K.J., Doyley, M.M., and Rubens, D.J. (2011), "Imaging the elastic properties of tissue: the 20 year perspective," *Physics in Medicine and Biology,* **56**, 5359-5360.

Patterson, E., and Wang, Z. (1991), "Towards full field automated photoelastic analysis of complex components," *Strain*, **27**, 49–53.

Paynabar, K., Jin, J., and Pacella, M. (2013), "Analysis of multichannel nonlinear profiles using uncorrelated multilinear principal component analysis with applications in fault detection and diagnosis," *IIE Transactions*, **45**, 1235–1247.

Paynabar, K., Qiu, P., and Zou, C. (2016), "A change point approach for phase-I analysis in multivariate profile monitoring and diagnosis," *Technometrics*, **58**, 191–204.

Perona, P., and Malik, J. (1990), "Scale space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **12**, 629–639.

Polzehl, J., and Spokoiny, V.G. (2000), "Adaptive weights smoothing with applications to image restoration," *Journal of the Royal Statistical Society (Series B)*, **62**, 335–354.

Prats-Montalban, J.M., and Ferrer, A. (2014), "Statistical process control based on Multivariate Image Analysis:A new proposal for monitoring and defect detection," *Computers and Chemical Engineering,* **71**, 501–511.

Qiu, P. (1991), "Estimation of a kind of jump regression functions," *Journal of Systems Science and Complexity*, **4**, 1–13.

Qiu, P. (1994), "Estimation of the number of jumps of the jump regression functions," *Communications in Statistics - Theory and Methods*, **23**, 2141–2155.

Qiu, P. (1997), "Nonparametric estimation of jump surface," *Sankhya (Series A)*, **59**, 268–294.

Qiu, P. (1998), "Discontinuous regression surfaces fitting," *Annals of Statistics*, **26**, 2218–2245.

Qiu, P. (2002), "A nonparametric procedure to detect jumps in regression surfaces," *Journal of Computational and Graphical Statistics*, **11**, 799–822.

Qiu, P. (2003), "A jump-preserving curve fitting procedure based on local piecewise-linear kernel estimation," *Journal of Nonparametric Statistics*, **15**, 437–453.

Qiu, P. (2004), "The local piecewisely linear kernel smoothing procedure for fitting jump regression surfaces," *Technometrics*, **46**, 87–98.

Qiu, P. (2005), *Image Processing and Jump Regression Analysis*, New York: John Wiley & Sons.

Qiu, P. (2007), "Jump surface estimation, edge detection, and image restoration," *Journal of the American Statistical Association*, **102**, 745–756.

Qiu, P. (2008), "A nonparametric procedure for blind image deblurring," *Computational Statistics and Data Analysis*, **52**, 4828–4841.

Qiu, P. (2009), "Jump-preserving surface reconstruction from noisy data," *Annals of the Institute of Statistical Mathematics*, **61**, 715–751.

Qiu, P. (2014), *Introduction to Statistical Process Control,* Boca Raton, FL: Chapman Hall/CRC.

Qiu, P. (2017), "Statistical process control charts as a tool for analyzing big data," In *Big and Complex Data Analysis: Statistical Methodologies and Applications* (Ejaz Ahmed ed.), 123–138, New York: Springer-Verlag.

Qiu, P., Asano, Chi., and Li, X. (1991), "Estimation of jump regression functions," *Bulletin of Informatics and Cybernetics*, **24**, 197–212.

Qiu, P., and Bhandarkar, S.M. (1996), "An edge detection technique using local smoothing and statistical hypothesis testing," *Pattern Recognition Letters*, **17**, 849–872.

Qiu, P., and Kang, Y. (2015), "Blind image deblurring using jump regression analysis," *Statistica Sinica*, **25**, 879–899.

Qiu, P., and Mukherjee, P.S. (2010), "Edge structure preserving image denoising," *Signal Processing*, **90**, 2851–2862.

Qiu, P., and Mukherjee, P.S. (2012), "Edge structure preserving 3-D image denoising by local surface approximation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **34**, 1457–1468.

Qiu, P., and Nguyen, T. (2008), "On image registration in magnetic resonance imaging," *IEEE Proceedings of the 2008 International Conference on BioMedical Engineering and Informatics*, 753–757.

Qiu, P., and Sun, J. (2007), "Local smoothing image segmentation for spotted microarray images," *Journal of the American Statistical Association*, **102**, 1129–1144.

Qiu, P., and Sun, J. (2009), "Using conventional edge detectors and post-smoothing for segmentation of spotted microarray images," *Journal of Computational and Graphical Statistics*, **18**, 147–164.

Qiu, P., and Xiang, D. (2014), "Univariate dynamic screening system: an approach for identifying individuals with irregular longitudinal behavior," *Technometrics*, **56**, 248–260.

Qiu, P., and Xiang, D. (2015), "Surveillance of cardiovascular diseases using a multivariate dynamic screening system," *Statistics in Medicine*, **34**, 2204–2221.

Qiu, P., and Xing, C. (2013a), "On nonparametric image registration," *Technometrics*, **55**, 174–188.

Qiu, P., and Xing, C. (2013b), "Feature Based Image Registration Using Non-Degenerate Pixels," *Signal Processing*, **93**, 706–720.

Qiu, P., and Yandell, B. (1997), "Jump detection in regression surfaces," *Journal of Computational and Graphical Statistics*, **6**, 332–354.

Qiu, P., and Yandell, B. (1998), "A local polynomial jump detection algorithm in nonparametric regression," *Technometrics*, **40**, 141–152.

Qiu, P., and Zou, C. (2010), "Control chart for monitoring nonparametric profiles with arbitrary design," *Statistica Sinica*, **20**, 1655–1682.

Qiu, P., Zou, C., and Wang, Z. (2010), "Nonparametric profile monitoring by mixed effects modeling (with discussions)," *Technometrics*, **52**, 265–293.

Rajagopalan, A.N., and Chaudhuri, S. (1999), "MRF model-based identification of shift-variant point spread function for a class of imaging systems," *Signal Processing*, **76**, 285–299.

Saeed, N. (1998), "Magnetic resonance image segmentation using pattern recognition, and applied to image registration and quantitation," *NMR in Biomedicine*, **11**, 157–167.

Sapiro, G. (2001), *Geometric Partial Differential Equations and Image Analysis,* New York: Cambridge University Press.

Skilling, J. (1989, eds), *Maximum Entropy and Bayesian Methods*, Norwell, MA: Kluwer Academic.

Sohn, H., Park, G., Wait, J. R., Limback, N. P., and Farrar, C. R. (2004), "Wavelet-based active sensing for delamination detection in composite structures," *Smart Materials and Structures*, **13**, 153.

Song, H., and Qiu, P. (2017a), "A parametric intensity-based 3D image registration method for magnetic resonance imaging," *Signal, Image and Video Processing*, **11**, 455–462.

Song, H., and Qiu, P. (2017b), "Intensity-based 3D local image registration," *Pattern Recognition Letters*, **94**, 15–21.

Song, H., and Qiu, P. (2017c), "High speed 3D image registration using distributed parallel computing," under review.

Sonka, M., Hlavac, V., and Boyle, R. (1993), *Image Processing, Analysis and Machine Vision,* New York: Springer.

Sun, J., and Qiu, P. (2007), "Jump detection in regression surfaces using both first-order and second-order derivatives," *Journal of Computational and Graphical Statistics*, **16**, 289–311.

Torre, V., and Poggio, T. (1986), "On edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **8**, 147–163.

Wahba, G. (1990), *Spline Models for Observational Data,* Philadelphia: SIAM.

Wang, Y. (1995), "Jump and sharp cusp detection by wavelets," *Biometrika*, **82**, 385–397.

Wang, Y. (1998), "Change curve estimation via wavelets," *Journal of the American Statistical Association*, **93**, 163–172.

Wang, K., and Jiang, W. (2009), "High-dimensional process monitoring and fault isolation via variable selection," *Journal of Quality Technology*, **41**, 247–258.

Wang, A., Wang, K., and Tsung, F. (2014), "Statistical surface monitoring by spatial-structure modeling," *Journal of Quality Technology*, **46**, 359–376.

Weickert, J. (1998), *Anisotropic Diffusion in Imaging Processing*, Stuttgart, Germany: Teubner.

Worsley, K.J. (1983), "Testing for a two-phase multiple regression," *Technometrics*, **25**, 35–42.

Wu, J.S., and Chu, C.K. (1993), "Kernel type estimators of jump points and values of a regression function," *Annals of Statistics*, **21**, 1545–1566.

Wu, G., Qi, F., and Shen, D. (2006), "Learning-based deformable registration of MR brain images," *IEEE Transactions on Medical Imaging*, **25**, 1145–1157.

Xia, H., Ding, Y., and Wang, J. (2008), "Gaussian process method for form error assessment using coordinate measurements," *IIE Transactions*, **40**, 931–946.

Xia, Z., and Qiu, P. (2015), "Jump information criterion for statistical inference in estimating discontinuous curves," *Biometrika*, **102**, 397–408.

Xing, C., and Qiu, P. (2011), "Intensity based image registration by nonparametric local smoothing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **33**, 2081–2092.

Yan, H., Paynabar, K., and Shi, J. (2015), "Image-based process monitoring using low-rank tensor decomposition," *IEEE Transactions on Automation Science and Engineering,* **12**, 216–227.

Yan, H., Paynabar, K., and Shi, J. (2017), "Anomaly detection in images with smooth background via smooth-sparse decomposition," *Technometrics*, **59**, 102–114.

Zang, Y., and Qiu, P. (2017), "Phase I monitoring of spatial surface data from 3D printing," *Technometrics*, in press.

Zhang, J., Kang, Y., Yang, Y., and Qiu, P. (2015), "Statistical monitoring of the hand, foot, and mouth disease in China," *Biometrics*, **71**, 841–850.

Zhang, J., Qiu, P., and Chen, X. (2016), "Statistical monitoring-based alarming systems in modeling the AIDS epidemic in the US, 1985-2011," *Current HIV Research*, **14**, 130–137.

Zitova, B., and Flusser, J. (2003), "Image registration methods: a survey," *Image and Vision Computing*, **21**, 977–1000.

Zou, C., and Qiu, P. (2009), "Multivariate statistical process control using LASSO," *Journal of American Statistical Association*, **104**, 1586–1596.

Zou, C., Qiu, P., and Hawkins, D. (2009), "Nonparametric control chart for monitoring profiles using change point formulation and adaptive smoothing," *Statistica Sinica*, **19**, 1337–1357.

Zou, C., Tsung, F., and Wang, Z. (2007), "Monitoring general linear profiles using multivariate EWMA schemes," *Technometrics*, **49**, 395–408.

Zou, C., Tsung, F., and Wang, Z. (2008), "Monitoring profiles based on nonparametric regression methods," *Technometrics*, **50**, 512–526.