Towards the Utilization of Machine Vision Systems as an Integral Component of Industrial Quality Monitoring Systems

Fadel Mounir Megahed

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Master of Science
In
Industrial and Systems Engineering

Jaime A. Camelio, Chair

William H. Woodall, Member

Joel A. Nachlas, Member

December 15th, 2009

Blacksburg, Virginia

Keywords: Fault-Detection, Process Monitoring, Quality Monitoring, Machine Vision Systems, and Digital Image Processing

Towards The Utilization of Machine Vision Systems (MVS) as an Integral Component of Industrial Quality Monitoring Systems

Fadel Mounir Megahed

ABSTRACT

Recent research discussed the development of image processing tools as a part of the quality control framework in manufacturing environments. This research could be divided into two image-based fault detection approaches: 1) MVS; and 2) MVS and control charts. Despite the intensive research in both groups, there is a disconnect between research and the actual needs on the shop-floor. This disconnect is mainly attributed to the following:

- The literature for the first category has mainly focused on improving fault detection accuracy through the use of special setups without considering its impact on the manufacturing process. Therefore, many of these methods have not been utilized by industry, and these tools lack the capability of using images already present on the shop floor.
- The studies presented on the second category have been mainly developed in isolation. In addition, most of these studies have focused more on introducing the concept of utilizing control charts on image data rather than tackling specific industry problems.

In this thesis, these limitations are investigated and are disseminated to the research community through two different journal papers. In the first paper, it was shown that a face-recognition tool could be successfully used to detect faults in real-time in stamped processes, where the changes in image lighting conditions and part location were allowed to emulate actual manufacturing environments. On the other hand, the second paper reviewed the literature on image-based control charts and suggested recommendations for future research.

DEDICATION

To my grandma and Montasser, in heaven

And my mum, my guiding light on earth

ACKNOWLEDGMENT

First of all, I would like to express my deepest thanks to Dr. Jaime Camelio, my advisor and my committee chair. This thesis would not have been possible if it were not for your continuous support, trust and encouragement.

I am also indebted to Prof. William H. Woodall who taught me a lot about quality control and process monitoring. I also thank him for serving on my committee and for his constant support and guidance in writing the review paper.

I would also like to Dr. Joel Nachlas for serving on my committee. I also thank him for being there for me from the first day in Virginia Tech through his role as the advisor of the “Operations Research” Option in ISE.

I am also extremely grateful for Hannah Swiger, Kim Ooms and the rest of the ISE staff who made me feel home here in Virginia Tech.

Also, I would like to thank my colleagues and friends in the iMAS lab; Anna Schuh, Jeremy Rickli, and Lee Wells who have helped me a lot during my graduate research and taught me the trades of being a graduate researcher.

I owe my deepest gratitude to the faculty of the Mechanical Engineering Department in the American University of Cairo, where I did my undergraduate studies. Without you, I would have never had the knowledge to ever-succeed on this level.

I would like to thank my best friends in Cairo for being there for me every day in my life and more importantly for being there for my family when I could not. Amal, Dina, Fatma, Gamil, Mokhtar, Nourhan, Omar, Sherif and Yehia; no words are enough to show my gratitude for you.

Finally, I would like to thank my closest friends in the US, especially: Andrew Henry, Kammy Mann, Na Mi and Supreet Dangore. Without you, I could have never made it, and I could have never been able to withstand being all these many thousand miles away from my family, friends and my beloved country. I am indebted to your kindness, love and endless support.

TABLE OF CONTENTS

LIST OF FIGURES

PREFACE

This thesis is submitted in partial fulfillment of the requirements for the Master of Science in Industrial and Systems Engineering. It involves the research work carried out, mostly by the author, starting late December 2008 through early December 2009 in the iMAS lab, a part of the Center for High Performance Manufacturing in the ISE department in Virginia Tech. My advisor during this period is: Dr. Jaime Camelio, Director of the iMAS Lab. The thesis has been written solely by the author; some of the ideas, as in any research, are based on the research of others, and I have done my best to provide references to these sources.

In November 2008, Dr. Jaime Camelio presented his research to the ISE graduate seminar. Since, at the time, I was a fresh masters student, and I was also exposed to the fundamentals of digital image processing research at my 2007 summer internship in Institut für Textiltechnik (ITA), I was both impressed and intrigued by the opportunity of joining his research team to work on a continuation of a research project where a digital image processing tool was being developed for industrial fault detection purposes. During the first couple of months, I was assisted by Lee Wells who introduced me to what was done by the lab, which was documented in a conference paper that compared between two different image processing tools (this paper is cited in this report).

With the permission of Dr. Camelio, I built on that project to develop a tool, which could be used to detect faults in real-time in simulated manufacturing environments with better detection rates than the previous tools. I have written the manuscript containing the research under the supervision of Dr. Camelio and in May 2009, we have submitted to the *Journal of Intelligent Manufacturing* (Dr. Camelio is 2nd author). We got the “accept with minor revisions” decision a few months later and we are currently waiting for the final acceptance.

In the spring semester of 2009, I had the idea of utilizing control charts combined with digital image processing to detect faults and trends in manufacturing processes. Prof. Woodall suggested reviewing the current literature as a starting point. This process has lead to the second paper (to be submitted to the *Journal of Quality Technology*) as we have found out that the developed techniques were developed in isolation. In the second manuscript here, I am the sole author, but by the time of submission Prof. Woodall would be putting a lot of his expertise to further develop this paper. Prof. Woodall is a renowned professor of Statistics and his expertise in statistical quality control is recognized world-wide and therefore, he would definitely add a lot to this review paper. Also, Dr. Camelio would be another co-author on the final paper, where his expertise in intelligent manufacturing systems would be extremely important.

Writing this thesis has been extremely challenging; however, in the process my knowledge about statistical process monitoring and digital image processing has increased a lot. The multi-disciplinary focus of the thesis has lead to the incorporation of many statistical, computer science and industrial engineering concepts.

1 Introduction

1.1 Motivation:

Research has been recently presented discussing the development of image processing techniques to detect product faults as an addition to existent fault detection methods in the Statistical Process Control (SPC) framework. This research could be divided into two categories: 1) image-based fault detection (inspection) methods that fully utilize digital image processing tools; and 2) image-based fault detection methods that combine digital image processing tools and control charts. Despite the intensive research in both groups, there is a gap between these research studies and the actual needs on the shop-floor. This gap is mainly attributed to the following two limitations:

- The literature for the first category has mainly focused on improving fault detection accuracy through the use of special setups without considering its impact on the manufacturing process. Therefore, not only many of these methods have not been utilized by industry, but also these tools lack the capability of using images already present on the shop floor.
- The studies presented on the second category have been mainly developed in isolation. In addition, most of these studies have focused more on introducing the concept of utilizing control charts on image data rather than tackling specific industry problems.

These two limitations hamper the utilization of Machine Vision Systems (MVS) to a greater extent in the quality monitoring framework and consequently, the current state of many quality monitoring systems within industry is mainly limited to the traditional dimensional data. This thesis sets the groundwork necessary for integrating different measurement technologies such as coordinate measuring machines, optical measuring machine and 3-D laser scanners in one robust quality framework, which optimizes the knowledge generated from these data sources to detect and diagnose out-of-control conditions so that a quick effective process recovery could be reached.

1.2 Thesis Layout:

This thesis is compromised of four major chapters. The second chapter presents the proposed image-based fault detection method that detects quality faults in real-time using a linear subspace method (Fisher's Linear Discrimnant). Chapter three provides a survey on the added benefits from combining control charts with image processing tools to detect product defects as well as specific industry-to industry recommendations. Chapter four is the concluding chapter for this work.

It should be noted that chapter two is an accepted paper in the *Journal of Intelligent Manufacturing* and it is currently under final review. Chapter three represents the first draft of a paper to be submitted to the *Journal of Quality Technology*.

2 Real-Time Fault Detection in Manufacturing Environments Using Face Recognition Tools

2.1 Abstract

New image processing techniques as well digital image capture equipment provide an opportunity for fast detection and diagnosis of quality problems in manufacturing environments compared with traditional dimensional measurement techniques. This paper proposes a new use of image processing to detect in real-time quality faults using images traditionally obtained to guide manufacturing processes. The proposed method utilizes face recognition tools to eliminate the need of specific feature detection on determining out-of-specification parts. The focus of the proposed methodology is on computational efficiency to ensure that the algorithm runs in real time in high volume manufacturing environments. The algorithm is trained with previously classified images. New images are then classified into two groups, healthy and unhealthy. This paper proposes a method that combines Discrete Cosine Transform (DCT) with Fisher's Linear Discriminant Analysis (FLD) to detect faults, such as cracks, directly from aluminum stamped parts.

2.2 Introduction

Research has been recently presented discussing the development of image processing techniques to detect product and process faults. Despite the intensive research (summarized in the Background section), there remains two major limitations, which are the driving force behind that study:

- There have not been any attempts to present the fault detection methods for digital images from a quality monitoring perspective. Therefore, it is difficult for a practitioner to integrate fault detection methods based on digital images with traditional dimensional statistical quality control methods.

- Literature shows that previous work has mainly focused on improving fault detection accuracy, but very little work has been presented on computational efficiency. Instead, most of the work has concentrated on increasing the fault detection accuracy through using special setups that increases the likelihood of detecting required faults. Consequently, many of the developed tools lack the capability of using images that are already present in the shop floor, which are used for other purposes—rather than fault detection—such as part location in robotic systems.

This paper investigates the feasibility of applying image processing techniques for the purpose of fault detection in manufactured products in real-time such that known faults (such as: surface cracks, wrinkles and splits) as well as unknown faults are efficiently and effectively detected. There are four major contributions in this study: 1) the use of image-based face recognition tools (linear subspace methods) rather than trying to measure directly some particular features in the image such as valleys and contours as shown in the Background Section; 2) the use of image-based face recognition tools, allows for detecting unknown faults and therefore, overcoming the limitations found in some of the current techniques; 3) the ability to utilize images that are currently used in the shop floor for other purposes without the need of any modifications; 4) the focus on computational efficiency to ensure that the algorithm runs in real time in high volume manufacturing environments. The proposed methodology is presented—with the practitioner in mind—and then, tested in a pilot study to detect faults in aluminum stamped parts.

2.3 Background

This section reviews the state of the art in image processing techniques used for quality purposes. The section is divided into three subsections: the first one provides a brief explanation of some of the image processing tools that are used in this paper; the second provides a brief survey of applying face recognition techniques in fault detection; and the third subsection provides specifics of the techniques developed for sheet metal work.

2.3.1 Related Image Processing Background

As discussed in the Introduction, the proposed technique uses face recognition tools to detect product faults. This section presents a brief discussion of the most important face detection techniques, for more detailed surveys see (Hjelmas and Low 2001; Yang et al. 2002; Zhao et al. 2003).

Numerous classifications have been presented for categorizing face detection techniques. However, most of these classifications seem to follow the categories proposed by Hjelmas and Low (2001) and Yang et al. (2002). Hjelmas and Low's classification closely relates to the work done, in the literature, in fault detection; therefore, it will be adopted in this paper.

Hjelmas and Low (2001) classified face detection techniques into two major categories, feature-based approaches and image-based approaches. The feature based approaches are divided into three subcategories: low-level analysis (naïve), feature analysis, and active shape models (advanced). The low-level analysis uses segmentation concepts see (Gonzalez and Woods 2008) to detect facial features/faces through numerous ways, which include edge detection methods (Canny 1986; Sirohey 1993; Zobel et al. 2000) and gray-levels (Graf et al., 1995). On the other hand both feature analysis and active shape models have been developed to overcome the problems in the low-level analysis techniques, mainly the ambiguity of the generated features, through the knowledge of the face geometry. According to Yang et al. (2002) and Abdallah (2007), these techniques are mainly for face localization rather than face detection and therefore, will not be discussed further. The image-based approaches are also divided into three categories neural networks (Viennet and Soulie 1995; Rowley et al. 1998a and 1998b), statistical methods—for a basic method based on maximum likelihood face detection see (Colmenarez and Huang 1996)—and linear subspace methods (Hjelmas and Low 2001).

Linear subspace methods are techniques that apply concepts from multivariate statistics to represent the face, which is a subspace of the image space. These techniques are model-based methods, which require a training algorithm as no prior knowledge of face/facial features is needed (Yang et al. 2002). They include principal component analysis (PCA) and linear discriminant analysis (LDA). PCA was first used in face detection/recognition by Sirovich and Kirby (1987) and then by Turk and Pentland (1991). In general, both approaches make use of PCA to reduce the dimensions of the data by eigenvector decomposition where images are presented as a matrix. LDA-based methods, on the other hand, are usually recognized to be superior as they treat in-class variance and between-class variance separately (Lotlikar and Kolthari 2000), unlike PCA that depends on the overall variance matrix (Hjelmas and Low 2001; Er et al. 2005). Thus, LDA-based methods are more insensitive to variations in lighting conditions and facial expressions.

In applications that require high speed face decomposition, LDA is preceded by a data reduction stage. The data reduction can be carried out by wavelets (Zhu et al. 2000), PCA—as in any of the Fisher linear discriminant (FLD) based approaches see (Lotlikar and Kolthari 2000; Lu et al. 2003)—and Discrete Cosine Transforms (DCT) (Er et al. 2005). Er et al. (2005) recommend applying DCT followed by FLD for the following reasons: DCT is implemented by fast algorithms and make the data independent, thus, eliminating the need for re-training the system when a new class of faces is added to the set; and on the other hand, FLD allows for keeping more DCT coefficients and thus, the most discriminating features could be extracted at high speed.

2.3.2 Related Fault Detection Background

The purpose of this section is to provide some insights on how face detection techniques are applied in detecting faults (such as cracks, splits and abnormalities) in a number of different fields. These fields could be categorized into three major groups: 1) Medical Applications, where feature-based approaches (Liu et al. 1996; Hafez and Abdel Azeem 2002; Li et al. 2004) and neural networks (Joo et al. 2004; Mircic and Jorgovanovic 2006; Liang et al. 2007) are utilized to detect and diagnose abnormalities in different areas of the body. 2) Transportation and Structures, where edge detection, feature analysis, neural networks, and PCA-based approaches are the most widely-implemented methods. They are used in detecting cracks and surface anomalies in bridges (Abdel-Qader et al. 2003; Abdel-Qader et al. 2006; Lee et al. 2008), pavements (Ayenu-Prah and Attoh-Okine 2007) and roads (Bursanescu et al. 2000). 3) Manufacturing environments.

In a manufacturing environment, face detection techniques are used for two different functions: process fault detection and product quality. Process fault detection is when the measurement systems in an industrial plant are monitored for detecting faulty performance and diagnosing root-causes for process failures. It is outside the scope of this paper as the neural networks used in detecting such faults, in many cases, do not involve digital images, as in (Chumakov 2008; Huang and Kovacevic 2009), and therefore, will not be discussed further. On the other hand, product quality assurance, via digital image processing, was originally performed through feature-based approaches; one example of those approaches is Ho et al. (1990) where the Sobel edge detection technique is utilized to detect cracks in ferromagnetic materials steel plates. One constraint in this approach is the use of magnetic particles to increase the accuracy of their algorithm, and thus, it cannot be easily generalized. Now, most of the quality assurance/fault detection techniques, applied to products, focus on utilizing neural networks/expert systems see (Chang et al. 2009; Chen et al. 2009; Selek et al. 2009).

These approaches utilize the learning ability of neural networks to continuously improve the fault identification/detection process. However, none of the aforementioned techniques, including (Selek et al. 2009), reported the computational speed of their developed algorithms, which might be an obstacle in applying those methodologies in high volume production environments. Furthermore, it is clear that the linear subspace algorithms developed for face detection have not been fully utilized in fault detection techniques in manufacturing environments despite its success in face detection (Er et al. 2005) and crack detection in bridges (Abdel-Qader et al. 2006).

2.3.3 Related Sheet Metal Fault Detection Techniques

Since the proposed model is tested on stamped aluminum sheets; it is important to describe the techniques developed for detecting faults in sheet metal parts. Shima et al. (1989) developed an algorithm to detect stamping defects in the surfaces of electronic parts. Their method utilized digital image processing techniques such as point matching techniques, position alignment methods, and gray level analysis for detecting defects in stamping. The limitation of that approach is that it focuses on character stamping defects and thus, its application in identifying other possible faults, such as cracks and splits, is doubtful. More recently, Fuente et al. (2003) presented vision system for detecting defects in large stamped parts. Their method combined statistical techniques, Markov

Random Field models see (Li, 2001), and feature based approaches to detect splits and cracks. One of the advantages of this paper is the ability of their algorithm to successfully detect splits, which are greater than 0.2 mm. However, their algorithm focuses only on detecting splits and cracks, and thus, other possible defects (such as inclusions) may not be detected. Garcia (2005) provided a more complete model combining image processing techniques and fuzzy logic to detect cracks and wrinkles in stamped parts in real-time with good precision. However, an obvious limitation in Garcia's approach is that it requires changes in the work station, which cannot always be attained on a shop floor.

Our proposed technique fills in some of the gaps found in the literature, while testing the feasibility of applying linear subspace models in fault detection in manufactured parts, as it: 1) ensures that there are no modifications required on the shop floor as it utilizes images that are currently used for part location in robotic systems, 2) focuses on computational efficiency so that it can run in real-time and 3) is able to detect unknown faults.

2.4 Methodology

2.4.1 Overview of the Proposed Technique:

A The proposed methodology utilizes linear subspace methods (FLD-based methods) along with standard digital image processing techniques to ensure that an effective and efficient measurement system (via digital images) is obtained. This method follows a two phase quality approach: *Phase A* and *B*. During *Phase A*, a set of images representing healthy and unhealthy (faulty) parts—including the representation of all known faults—are collected. These images are then processed, analyzed and then, tested on a historical set of images to ensure that the algorithm is ready to be used in monitoring future production. During *Phase B*, real-time images of the parts are compared to *Phase A* data to determine if the observed part is healthy or not. A schematic of the proposed methodology is shown in Figure 1.

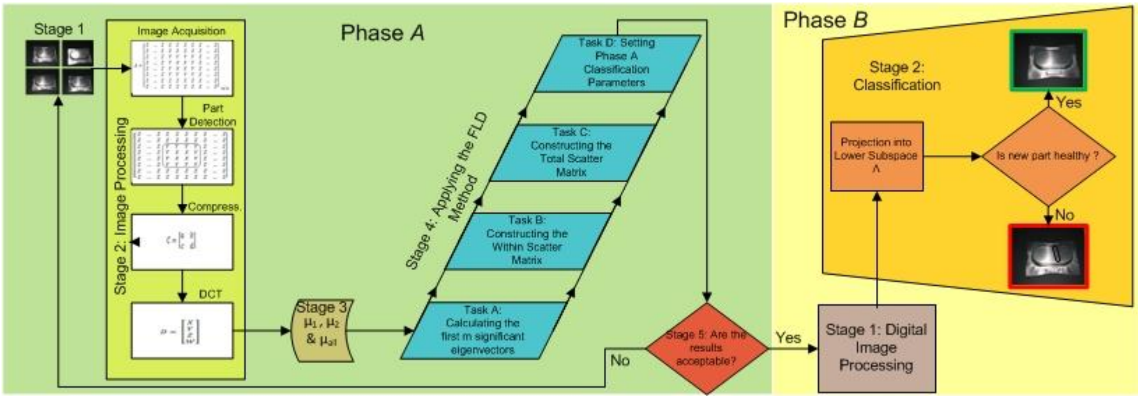


Figure 1 Schematic of the Proposed Methodology

The image-based quality monitoring approach proposed in this paper is significantly different than the traditional quality monitoring approach. The difference is mainly attributed to the objective of the two Phases in each of the approaches. In the traditional quality approach, Phase I is when historical data is used to ensure that the process is brought into a state of statistical control, usually through an iterative process, which primarily targets building an understanding of the process to be able to improve it (Woodall 2000). On the other hand, in our proposed approach, as well as any image-based approach, the set of data (images) constituting Phase A is different, as the target is to be able to fully represent the healthy and unhealthy sets of data based on an understanding of product requirements and process capability to minimize the probability of a false classification by the developed algorithms. In other words, it is a process, where the developers teach the computer what is a healthy and a faulty product. Thus, it is a step preceding the traditional Phase I, as the traditional Phase I assumes that the in-place measurement systems—such as dimensional measurements from Coordinate Measurement Machines (CMMs)—are capable of measuring the critical-to-quality characteristics (Montgomery 2008), which is not the case in an image based measurement system.

In addition, the traditional Phase II approach is built on the assumption that the process is stable as the most significant sources of variation have been removed in Phase I. Based on this assumption, Phase II focuses on process monitoring to detect shifts in the process. Depending on the size of the shift required to be detected, the appropriate method—

usually a Phase II control chart such as CUSUM (Page 1954) and MEWMA (Lowry et al. 1992)—is selected. In the image-based Phase B approach, the main focus is to check on the capability of the developed method to accurately detect the state of the part (if it is healthy or not). In other words, the image-based Phase B approach could be seen as a naïve process monitoring/measurement system calibration step, where the focus is on whether the part is healthy or not. Thus, the proposed two phase image-based quality approach sets the groundwork needed for applying traditional quality monitoring techniques to digital images. The methodology developed for these two phases is described in the subsections below.

2.4.2 Phase A: Measurement System Capability

Before discussing the work done in Phase A, it should be noted that the human capability of easily distinguishing between healthy and unhealthy parts is a complex problem for a machine/computer. One of the most important problems, yet often neglected in the literature, is creating an understanding of the way that a machine "sees" and how to overcome its limitations to ensure that the developed image-based processing systems deliver continuous and consistent results that are better than the average performance of humans over time (Demant et al. 1999). There are five major stages that are pivotal in ensuring that the developed image-based system meets the aforementioned criteria.

Stage 1: Image Selection

The image selection process should be developed such that all permissible variations of the product as well as all types of known faults are taken into account in the Phase A sample images. The representation of all permissible variations of the product is relatively easy, where images are taken for products that are in both ends of the specification limits set by the product design team. The mapping of historical faults is a more complex process, which could involve using one or more of the Statistical Process Control problem-solving tools, especially histograms and check sheets (Banks 1989). Also, it is recommended that the acquired images are shot in different environmental conditions (lighting conditions and minor changes in part positioning) to ensure that the system is intact to changes in environmental conditions.

The number of images (samples) that are needed for Phase A is a function of the inherent variability in the process, the required accuracy level, the degree of variation in environmental conditions, and the speed required for the real-time classification in Phase B. The traditional Phase I sample size of 25-40 samples should also be sufficient for an excellent representation of healthy and unhealthy parts for most processes. After these images are captured and stored, Stage 1 is complete and the subsequent digital image processing stage begins.

Stage 2: Digital Image Processing

This stage encompasses four consecutive tasks, which are described below:

Task A: Image Acquisition

This is the starting step of all digital image processing codes. It is where the images are fed into the software environment to be processed. Basically, this task is about transforming the JPEG (could be other formats as well) image into the following matrix:

$$A = \begin{bmatrix} f(1,1) & f(1,2) & \cdots & f(1,n) \\ f(2,1) & f(2,2) & \cdots & f(2,n) \\ \vdots & \vdots & \vdots & \vdots \\ f(m,1) & f(m,2) & \cdots & f(m,n) \end{bmatrix} \tag{1}$$

It should be noted that (x, y) are the spatial co-ordinates of the image and the function f returns the intensity of that particular pixel. The value of the function f is characterized by the amount of illumination incident and reflected by the object in the image (Gonzalez and Woods 2008). The values of *m* and *n* depend on the resolution of the camera used, where *m* and *n* increases as the resolution (number of megapixels) increases.

The captured images are always assumed to be in a RGB (colored) format. Colored images take more time to be processed and therefore, it is customary to rescale the image (matrix) into gray scale. This is achieved by scaling the intensity values such that it takes values between 0 (black) and 255 (white). All image processing software have standard functions to achieve this transformation. This transformation marks the end of Task A.

Task B: Part Detection

In order to be able to run Phase B in real-time, most of the unnecessary background in the image must be removed. This background increases the size of the transformed gray-scale Matrix A, without providing additional information, and thus, the computational time needed to manipulate that matrix increases. There are two standardized approaches for part detection: 1) a face recognition approach, as discussed in the background, and 2) an automated crop approach, which utilizes knowledge of the process to define the area of the image to be cropped.

The automated crop approach is preferred as it increases the computational efficiency when applied in Phase B; the same digital image processing techniques have to be applied in Phase B to ensure consistency. The automated crop approach utilizes the knowledge of the process to determine the maximum amount of variability in part location and thus, define a window that will always encompass the part. For example, assume that the transformed gray-scale matrix is a 6 by 6 matrix—Matrix A' in Eq.2—and that the part's nominal location is in the center of that matrix. The part's dimensions are 2 by 2 pixels. Then, the window could be set to form the new image Matrix B, which will always encompass the part as shown in Eq.2. This reduction in matrix size provides a further boost to the computational efficiency of the proposed system.

$$A' = \begin{bmatrix} Z & Z & Z & Z & Z & Z \\ Z & Y & Y & Y & Y & Z \\ Z & Y & X & X & Y & Z \\ Z & Y & X & X & Y & Z \\ Z & Y & Y & Y & Y & Z \\ Z & Z & Z & Z & Z & Z \end{bmatrix} \rightarrow B = \begin{bmatrix} Y & Y & Y & Y \\ Y & X & X & Y \\ Y & X & X & Y \\ Y & Y & Y & Y \end{bmatrix} \quad (2)$$

* X: Nominal set of Pixels for Part Location; Y: Possible set of pixels for Part Peripheries & Z: Impossible set of pixel for Part Peripheries

Task C: Compression

The Matrix B, output of task B, does not always lead to major improvements in computational efficiency. That is because one of the main motivations behind the proposed system is to make use of images that are currently in use in production for other purposes. Thus, the aim is avoid introducing changes in the facility to maximize the performance of the proposed algorithm. On the contrary, the backbone of this paper is based on developing a system that would detect faults in real-time based on images that are already utilized in the production facility for other purposes. Therefore, the proposed algorithm is developed to be robust for different camera conditions, as long as the part is fully captured in the photo. Consequently, the algorithm should work for the cases when the zoom level is too high to the extent there is no room for the unwanted background pixels Z, which are shown in task B. In this case (B = A), which means that matrix B is a 3264 × 2448 sized-matrix in a standard 8.0 Megapixel camera, a much larger matrix if the camera's resolution is better. The processing operations performed on such a large matrix—in the subsequent tasks and stages—would definitely prevent real-time processing of the newly produced parts in Phase B.

To solve this problem (based on B = A), the proposed algorithm utilizes one of the most known concepts in the field of digital image processing, which is inter-pixel redundancy. It basically states that in an image, neighboring pixels are highly correlated and thus, any pixel value could be guessed based on its neighbors. Therefore, the image (or its matrix) could be compressed (loosing data) without the loss of any information and with a substantial improvement in computational efficiency (Gonzalez et al. 2004). For example, a compression ratio of 4, would lead to a 200 % improvement in the computational efficiency as well as a new transformed matrix C.

$$B = \begin{bmatrix} a & b & e & f & \cdots \\ c & d & g & h & \cdots \\ m & n & q & r & \cdots \\ o & p & s & t & \cdots \\ \vdots & \vdots & \vdots & \vdots & \cdots \end{bmatrix} \xrightarrow{Yields} C = \begin{bmatrix} \frac{a+b+c+d}{4} & \frac{e+f+g+h}{4} & \cdots \\ \frac{m+n+o+p}{4} & \frac{q+r+s+t}{4} & \cdots \\ \vdots & \vdots & \vdots \end{bmatrix} \tag{3}$$

It should be noted that different characters have been used to represent Matrix B elements in each of Task B and C. In Task B, the characters reflected location; however, in Task C, the character indicates the pixel value. Matrix C could, also, be utilized in the case ($B \neq A$) to provide a further improvement in the system's computational efficiency. The construction of Matrix C marks the end of Task C, which is going to be followed by a further reduction stage in Task D.

Task D: Discrete Cosine Transform (DCT)

The previous reduction tasks, Tasks B and C, aim at providing assistance to the two major reduction stages performed in this algorithm, mainly DCT and FLD. Experimental work in face recognition suggests that DCT is a very efficient method that can be utilized to reduce image redundancy as only a subset of the transform coefficients is needed to represent the face. High recognition rates were achieved by Pan et al. (2000) using just 0.19 % of transform components. Thus, this method is very appealing in enabling the proposed system to achieve the aforementioned objectives of real-time detection.

In general, the DCT could be performed in two ways: a 2-D matrix transformation and a 1-D vector transformation. In this paper, the 1-D vector transformation is followed based on the recommendations in Er et al. (2005). In order to be able to apply the 1-D vector transformation, the Matrix C must be, first, transformed into a vector. This transformation is carried out by a "Zigzag Scanning Approach" (McAndrew 2004; Er et al. 2005). The Zigzag Scanning approach is illustrated by the following example. Assuming that Matrix C is a 4 X 4 matrix, then the vector D1 is obtained as follows to maintain the continuity of the data in the vector.

$$C = \begin{bmatrix} c_{11} & c_{12} & c_{13} & c_{14} \\ c_{21} & c_{22} & c_{23} & c_{24} \\ c_{31} & c_{32} & c_{33} & c_{34} \\ c_{41} & c_{42} & c_{43} & c_{44} \end{bmatrix} \rightarrow \begin{bmatrix} c_{11} \\ c_{12} \\ c_{21} \\ c_{31} \\ c_{22} \\ c_{13} \\ c_{14} \\ c_{23} \\ c_{32} \\ c_{41} \\ c_{42} \\ c_{33} \\ c_{24} \\ c_{34} \\ c_{43} \\ c_{44} \end{bmatrix} \tag{4}$$

After the vector D1 is obtained, the vector is further transformed by the DCT. DCT transforms vector D1, which in general is sized (1 × xy), into an equal length vector D2, as shown in Eq. (5).

$$D1 = \begin{bmatrix} c_{11} \\ c_{12} \\ c_{21} \\ \vdots \\ c_{xy} \end{bmatrix} \xrightarrow{yields} D2 = \begin{bmatrix} d_1 \\ \vdots \\ d_k \\ \vdots \\ d_{xy} \end{bmatrix} \xrightarrow{where} d_k = w_k \sum_{i=1}^{xy} \cos \frac{(2i-1)(k-1)\pi}{xy},\ k = 1,2...xy \tag{5}$$

$$where\ w_k = \begin{cases} \dfrac{1}{\sqrt{xy}}, k = 1 \\ \sqrt{\dfrac{2}{xy}}, k \neq 1 \end{cases}$$

It should be noted that the first DCT coefficient (d_1) represents the mean value of the gray scale index of all pixels in the picture, corresponding to the overall brightness of the picture. Utilizing this coefficient, it becomes possible to remove the variation from image to image caused by different lighting conditions. Consequently, the algorithm becomes robust to changes in lighting conditions. The robustness to changes in lightening conditions is a key component in the developed algorithm. It enables the algorithm to be used for shop-floor images without the need for adjustments.

On the other hand, dimensional reduction is achieved through, only, preserving the significant DCT coefficients. The threshold that determines the DCT coefficients, which are kept, varies from one process to another. However, in most processes, the percentages of the DCT coefficients that should be maintained may not exceed ten percent (Pan et al. 2000; Zhao and Camelio 2008). The preserved coefficients are then stored in Vector D. The constructing of Vector D denotes the end of Task D and Stage 2.

Stage 3: Phase A—Parameter Setting

The operations carried out in Stages 1 and 2 are repeated—healthy first and then faulty—on every image in the Phase A training set. After the end of the digital image processing stage, vector D is stored in a column of Matrix M. Therefore, after processing each image in the Phase A training set, the Matrix M should contain the maintained DCT coefficients for each image. According to the processing order, it is known that the first ‘u’ columns in the matrix represent the cluster of healthy parts and the remaining ‘v’ columns in the matrix denote the cluster of the faulty parts. Accordingly, the vector μ_1 representing the average coefficients vector for healthy parts could be calculated through averaging the first ‘u’ columns of the Matrix M. Similarly, the average coefficients vector for faulty parts μ_2 could be calculated. The global average Vector μ_{all} is then calculated through the mean of μ_1 and μ_2. Then, the Vector μ_{all} is then used to standardize the Vectors μ_1 and μ_2 as shown in Eq. (6).

$$\mu_i = \mu_i - \mu_{all} \ , i = 1, 2 \tag{6}$$

Stage 4: Applying the FLD-Based Method

FLD is a well-known dimensional reduction technique in face detection. It divides the image covariance matrix into between-class and within-class variance. The objective of FLD is to maximize the between-class variance, while preserving the within class variance. Thus, the strength of FLD is in minimizing the effect of environmental conditions on the algorithm's accuracy and consequently be used for shop-floor images.

FLD, also, leads to the extraction of the most discriminating features and projects the discriminating features vector from the truncated DCT domain into an optimal subspace Λ (Er. et al. 2005). It should be noted that the FLD-based method used in this algorithm was preferred over the mainstream Fisher Face approach (Belhumeur et al. 1997; Zhao et al. 1999) as it was found to be more computationally efficient. The reasons behind the computational inefficiency of the Fisher Face were provided by Zhao and Camelio (2008). The following four tasks are needed to construct the FLD-based method:

Task A: Calculation of the First *m* Significant Eigenvectors of the Between Class Scatter Matrix

The output of Stage 3, vectors μ_1 and μ_2, is used to calculate the between class scatter matrix (S_B) as shown in Eq. (7).

$$S_B = \phi_B \phi_B^T \xrightarrow{where} \phi_B = \begin{bmatrix} \mu_1(1) & \mu_2(1) \\ \mu_1(2) & \mu_2(2) \\ \vdots & \vdots \\ \mu_1(N) & \mu 2(N) \end{bmatrix} \begin{matrix} \mu_i(j)\, is\, the\; j^{th}\; component\, of\; \mu_i, i=1,2 \\ N\, is\, the\, length\, of\; the\, vector\, \mu_i, i=1,2 \end{matrix} \tag{7}$$

Then, the first *m* significant eigenvectors (V) could be calculated as shown in Eq.(8).

$$V = \phi_B E_m \xrightarrow{where} E_m = [e_1, ..., e_m]\, is\, the\, vector\, of\; the\, nonzero\, eigenvalues\, of\; S_B \tag{8}$$

Task B: Constructing the Within Scatter Matrix

Then, the within class scatter matrix could calculated in a similar way to the between class scatter matrix;

$$S_w = \phi_w \phi_w^T;$$

$$\phi_w = \left[\begin{array}{cccc:ccc} M_1(1)-\mu_1(1) & M_2(1)-\mu_1(1) & \cdots & M_u(1)-\mu_1(1) & M_{u+1}(1)-\mu_2(1) & \cdots & M_{u+v}(1)-\mu_2(1) \\ M_1(2)-\mu_1(2) & M_2(2)-\mu_1(2) & \cdots & M_u(2)-\mu_1(2) & M_{u+1}(2)-\mu_2(2) & \cdots & M_{u+v}(2)-\mu_2(2) \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ M_1(N)-\mu_1(N) & M_2(N)-\mu_1(N) & \cdots & M_u(N)-\mu_1(N) & M_{u+1}(N)-\mu_2(N) & \cdots & M_{u+v}(N)-\mu_2(N) \end{array}\right] \tag{9}$$

* u: No. of healthy images; v: No. of faulty images; M: output matrix of Stage 3

Task C: Constructing the Total Scatter Matrix

Based on Task A, the Matrix Δ_B, shown below in Eq. (10), is a diagonal matrix.

$$\Delta_B = V^T S_B V \tag{10}$$

Since, the total scatter matrix (S_T) could be represented as the summation of the between class matrix (S_B) and the within class matrix (S_W)—Eq. (11a); the total scatter matrix will be a diagonal matrix as shown in Eq. (11b).

$$\begin{aligned} &\because S_T = S_B + S_W && (a) \\ &\because S_B \; is\, a\, diagonal matrix && (b) \\ &\therefore S_T = U^T S_T U && (c) \\ &where: U = V\Delta_B^{-0.5} \end{aligned} \tag{11}$$

Task D: Setting Phase A Classification Parameters

There are three parameters, which form the classification criteria: 1) lower dimensional subspace, Λ, 2) average projected vector for healthy cluster, Ω_H, 3) average projected vector for unhealthy cluster, Ω_F. These parameters are calculated as shown in Eqs. (12a, 12b, and 12c), respectively.

$$\begin{aligned} &\Lambda = U \times Eigenvalues\, of\; S_T \times (Eigenvectors\, of\; S_T)^{-0.5} && (a) \\ &\Omega_H = \Lambda^T \times (\Lambda^T M_H) && (b) \\ &\Omega_F = \Lambda^T \times (\Lambda^T M_F) && (c) \\ &where \begin{array}{l} M_H is\, the\; first\, u\, columns\, in\, the\, Matrix\, M \\ M_F is\, the\, last\, v\, columns\, in\, the\, Matrix\, M \end{array} \end{aligned} \tag{12}$$

These three parameters are stored in the system to be used in classifying real-time images in Phase B into either the healthy or faulty clusters. The system can be viewed, at this point of time, as a hypothesis, which needs to be tested to ensure that the performance of the system meets the requirements needed by the users. The testing stage is carried out in Stage 5.

Stage 5: Hypothesis Testing

This stage could be viewed as an iterative hypothesis testing stage, where the system is tested on a set of historical images to ensure that the performance of the algorithm meets user expectations. In the previous four stages, lots of assumptions have been utilized to aid the user in developing a measurement system, which is capable of detecting faults without the need of introducing any modifications in the shop-floor. These assumptions do not guarantee optimal results for every process, camera resolution, and type of faults to be detected; however, they provide a good starting point from which the user can begin changing the parameters—Image Selection, compression and the number of eigenvectors retained—to reach the required performance. This expected performance should be evaluated based on false alarms rate, processing speed, and accuracy. Therefore, this stage tests whether the algorithm parameters lead to results which maximizes the accuracy for a given processing speed and a false alarm rate threshold. More details about this stage are given later in this paper.

2.4.3 Phase B: Real-Time Fault Detection

In Phase A, the system (algorithm) has been trained to recognize the difference between a healthy part (image) and a non-healthy one. This process encompassed several stages that started with the image selection stage to the hypothesis testing stage, where the performance of the system is deemed to be up-to-standards. In order to be able to utilize the three classification parameters stored in Stage 4 of Phase A, the acquired image of the newly produced part must be processed in a similar fashion to Phase A. Accordingly, Phase B encompasses the following two stages:

Stage 1: Digital Image Processing

The method used in this stage is the same as the one utilized in the digital image processing stage in Phase A. Therefore, the four tasks—image acquisition, part detection, compression and DCT—carried out in Phase A will be performed in Phase B using the exact parameters' values. The use of the same parameters' values ensures the final Vector D' for the newly acquired image is of the same size as the Vector D of the Phase A images so that they can be compared.

Stage 2: Classification

The Vector D' of the newly acquired image will be projected to the lower dimensional subspace Λ, as shown in Eq. (13). Then, the Euclidean distance classification measure, predominately used in clustering applications in image processing will be utilized to determine if the newly produced product is in class **C** as shown in Eq. (14) (Gonzalez et al. 2004).

$$D'_{sub} = \Lambda^T D' \tag{13}$$

$$C = \arg_c \min \left| D'_{sub} - \Omega_c \xrightarrow{where} c = H, F \right| \tag{14}$$

Based on Eq. (14), the newly produced image is categorized either as healthy or faulty. Accordingly, a system has been developed to detect if a part is faulty or not in real-time without requiring any changes in the production process. This system uses different statistical and digital image processing concepts to ensure that the system and algorithm are robust against changes in the environmental conditions and, therefore, could be utilized on the shop-floor. Also, the proposed system utilizes linear subspace methods (FLD-based method) and reduction techniques (window selection, compression, and DCT) to ensure the computational efficiency of the algorithm so that it can be used in real-time fault detection.

2.5 Application of the Developed Measuring System in a Pilot Study

The proposed algorithm was tested in a pilot study to detect faults in aluminum stamped parts. The description of the aluminum parts, experimental setup, chosen values for algorithm parameters, and the performance metrics measured in the pilot study are described in the subsections below.

2.5.1 Description of the Product Under Study

The proposed image-based fault detection system is applied to detect faults in stamped aluminum sheet metal pans. These pans are produced using a high volume production stamping process. The nominal dimensions of the pans are shown in Figure 2. The proposed measurement system was tested on a set of one hundred and twenty images, which were divided equally into quality (healthy) products and faulty products. The faulty products are divided into three fault categories: splits, cracks, and inclusions. The width of the cracks and the splits ranged from 1.2-1.8 mms with a length ranging from 1.8-4 cms. An example of a quality product as well as each fault category is shown in Figure 3. It should be noted that the wave-formed surface variations—seen in Figure 3—are not considered a fault in this study and its effect is neutralized during the first stage of Phase A in this experiment, where healthy and unhealthy pans are selected for the fault-detection algorithm.

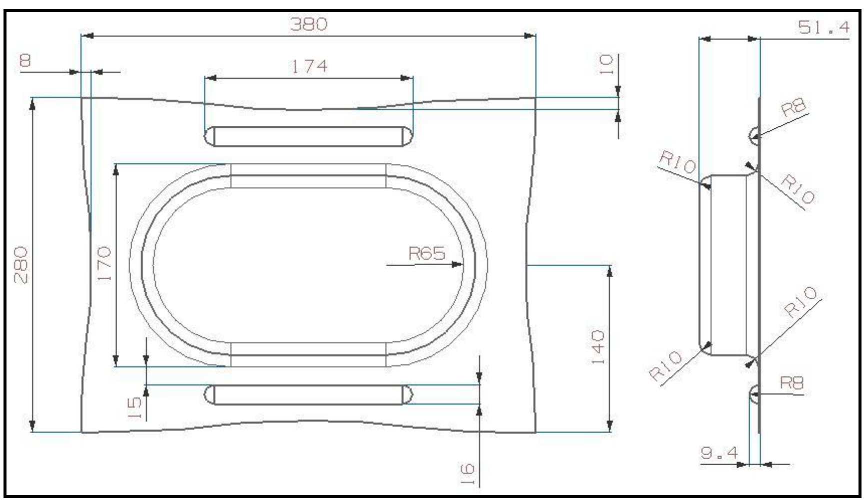


Figure 2 Draft Showing the Nominal Dimensions of the Pan (in millimeters)

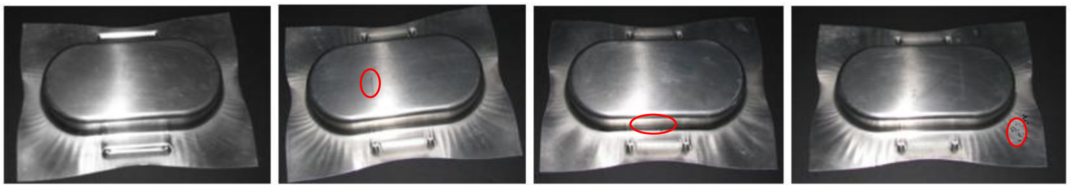


Figure 3 Representation of healthy pans and faulty (a crack, split and inclusions) pans respectively

2.5.2 Description of the Experimental Setup

The images were captured using a CANON SX 100 IS PowerShot 8.0 Mega-Pixels camera. These images were then processed using MATLAB 7.6 software, running on an Intel Core 2 Duo 4.0 GB desktop. It should be noted that the camera was mounted on an approximately (3ft*4ft*3ft) aluminum frame, shown in Figure 4. Unfortunately, the proposed techniques could not be tested in a real-manufacturing shop-floor; however, different environmental conditions—mainly, part location and lightening conditions—were altered during the image capturing process to emulate some of the difficulties expected in a real-manufacturing environment. The part location—of each pan during the image capturing process—was changed by rotating the pans around its central axis up to $\pm 15^{o}$, and moving the pan up to ± 0.6 cm, from its nominal location, in both vertical and horizontal directions. On the other hand, the lightening conditions were also varied between three lightening schemes. These changes attempt to artificially induce variations in the process, which would be equivalent to those occurring in the-shop-floor.

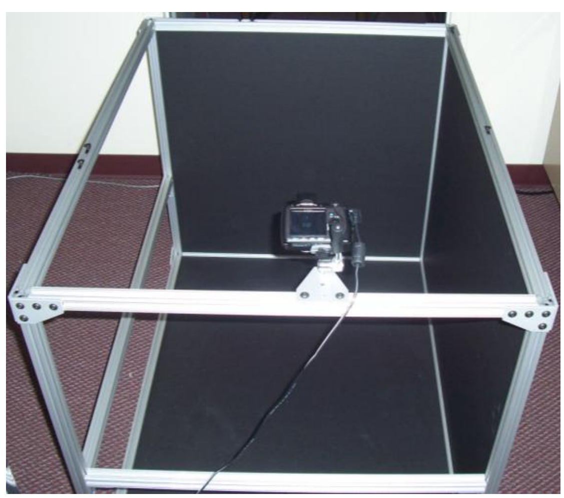


Figure 4 The Experimental Setup

2.5.3 Description of the Algorithm's Parameter Setup

This subsection shows how the proposed algorithm is applied in the case study to explain the choice of the parameters' values utilized in the image selection and digital image processing stages.

Image Selection Stage

During the training stage, the challenge is how to select images to fully represent the current knowledge of the process. In this pilot study, there are three categories of faults in the product, which are cracks, splits, and inclusions. These images are not optimal, as the conditions at which the images were captured, are not standardized. This problem can only be remedied in the image selection process, where the images selected for Phase A should represent both the healthy and faulty clusters in nominal as well as extreme conditions. Therefore, 9-15 images (3 faults × 3-5 variations of lightening/position) are required to represent the different variations of the faults and around 3-5 should be sufficient to represent variations within the healthy parts. It is customary to use equal number of images in training the system/computer to detect each cluster and therefore, 9-15 training images would be used for each set—healthy and faulty images. Thus, out of the 120 images, 20 (10×2) images would be used in Stage 1 of Phase A. The remaining images are used to verify that the code is effectively working in Stage 5. Since this is a pilot study, it is reasonable to assume that the Phase *B* results would be consistent with Stage 5 results.

Digital Image Processing Stage

After describing the process behind setting the number of required images, the approach followed in choosing appropriate values for the algorithm's digital image processing variables would be discussed. In the methodology section, the digital image processing stage is divided into four tasks. Each of these tasks needed input variable(s) to be processed. Some of these values were outputs from the previous task and the remaining values needed to be predefined by the user. In the following paragraphs, the techniques used in setting those variables, predefined by the user, in the baseline case study are shown.

The first task in the digital image processing stage is the image acquisition task, where the captured images are set into the computer. The variables in that step are the total number of Phase A images and the number of images in each class (healthy and faulty clusters). The selection criteria for these variables were discussed in the previous subsection. This led to the decision to use 20 images as Phase A data, divided equally between the two clusters.

The second task was to utilize the automatic cropping system to compress the image to allow for real-time detection in Phase B. Based on the process knowledge; the crop window was calculated to be 1200 ×1800 pixels. Therefore, the unwanted background data was removed leading to significant improvements in the computational efficiency of the code.

The third and fourth tasks induced further reduction to provide a further boost for real-time processing capabilities. A compression ratio of 4 was selected for the compression task and *m* coefficients of the DCT were maintained. By setting the values for the compression ratios and the DCT coefficients the parameter initialization for the developed system is concluded.

2.5.4 Description of the Expected Implementation Problems

In this pilot study, as well as for any image-based fault detection system, there are three variables that must be controlled in order to be able to benefit from such system. These three variables are false alarm rates, processing speed and fault detection accuracy. It is pivotal that the false alarm rates are minimized, the processing speed to be compatible with the manufacturing process cycle time and the fault detection accuracy to be maximized. If such conditions are reached, then the user will benefit from the fault detection system. A detailed discussion of overcoming expected problems for these three variables is provided in the following section.

Phase A: Implementation and Discussion

The proposed algorithm is meant to be utilized in different shop-floors with images that are currently available to detect product faults. These faults vary from one facility to another and therefore, any fault detection technique based on digital images would not necessarily be optimal for this facility. In this section, the tools needed to overcome traditional implementation problems (out-of-control-conditions for the variables aforementioned in the previous section) for digital image based systems will be explained to ensure the proposed algorithm optimally performs. It should be noted that in the literature there is no reference to the magnitude of these problems and therefore, the discussion is based on the most significant problems that could be encountered in implementing the proposed algorithm.

Problem 1: High False Alarm Percentage

Significance:

As in any fault detection scheme, an elevated false alarm rate drastically decreases the confidence of the facility's personnel in the accuracy of the system. Therefore, it is of prime importance to maintain a very low false alarm rate to ensure that the most alarms correspond to an actual fault in the product. There is usually a trade-off between false alarms and detection; however, in our application it is important to minimize false alarms—even if that decreases fault detection percentages—as detecting faults is a bonus we are trying to add to the system.

Causes:

An obvious reason for a high false alarm rate is the overlapping between the healthy and faulty clusters. There are two root-causes, shown in Figure 5, for such a problem. First, the inability of the system to differentiate between the two clusters as the current measurement system settings are incapable of detecting such a difference, as shown in Figure 6. An extreme example, is trying to detect a micro-fracture with a camera. Second, the system is capable; however, the in-class variation within one of the (both) clusters is so huge that the two clusters intersect, which could be seen in Figure 7.

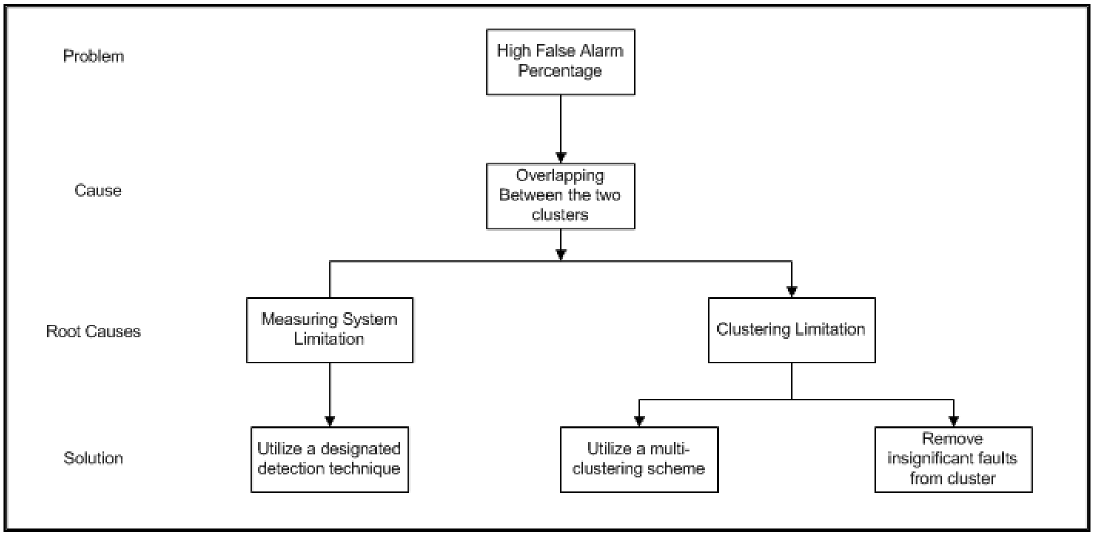


Figure 5 Root Causes for High False Alarms

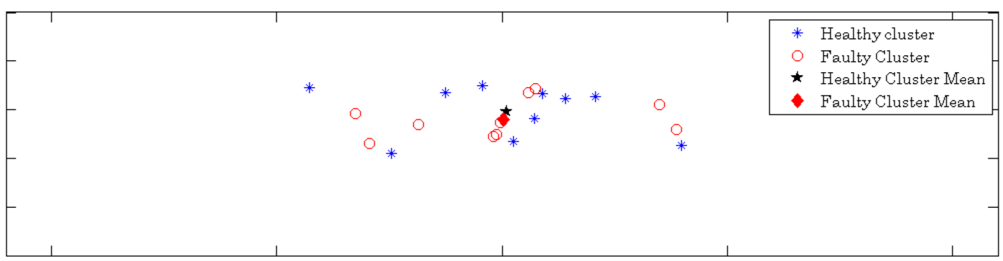


Figure 6 Representation of Cluster Interference due to System Inability

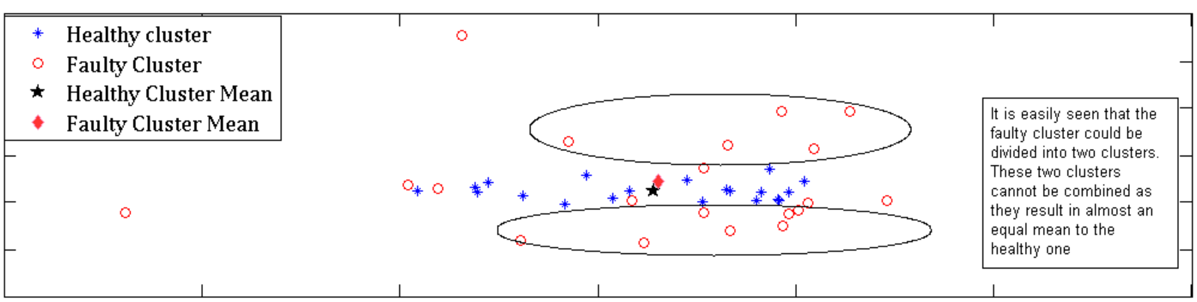


Figure 7 Representation of Cluster Interference due to Clustering Limitation

General Troubleshooting Techniques:

Based on running the actual program, the user should be able to identify the root-cause behind the overlapping between the two clusters. If the root cause is the inability of the system to differentiate between the two classes, then in-class and within class variation between classes would be tight as shown in Figure 6. Unlike, when the system is capable (Figure 7), where the faulty class could be divided into a group(s) whose projection is above the healthy cluster and a group(s) whose projection is below. Therefore, it is good practice to visually represent the clusters, in case of high false detection in order to understand the root-causes behind the problem.

How False Alarm Rates are Overcome in Implementation

Upon testing the code in Stage 5, both the false alarms and the missed-alarms (not detecting faults) were high. Upon performing a root-cause analysis, the system was found to encompass both features of incapability and clustering as shown in Figure 8.

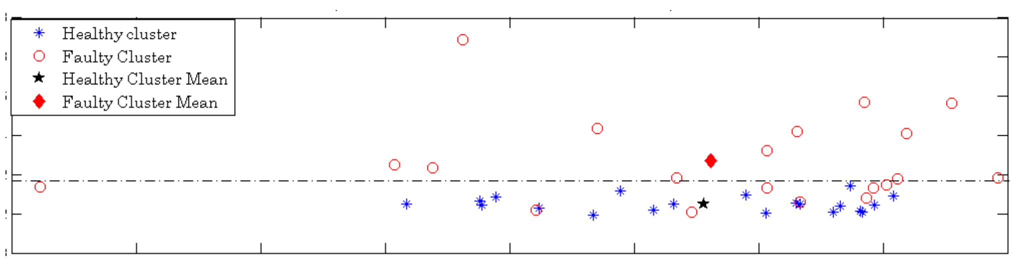


Figure 8 Encountering High False Alarms Problem

In Figure 8, the clustering effect is apparent that the faulty parts are divided into two groups, which are separated by the dotted line. It is also obvious that the lower cluster is in fact inseparable from the healthy cluster and therefore, the system is incapable of separating the lower cluster from the healthy one. Upon tracking the data for the lower cluster, it was found that the lower cluster represents the inclusions fault category. Therefore, it is concluded that based on the current conditions, the system cannot be used to detect inclusions, which makes sense as the camera is around 3 feet away from the pan and the image capturing conditions are not optimal. Thus, by removing the inclusions from the measurement step (to be detected by specialized algorithms in the quality monitoring phase), the system's performance will be greatly improved. Actually, it was improved to the extent that the false alarms rate went down to 2 percent and the faults detection rate was 66 percent.

Problem 2: Relatively Slow Processing Speed

Significance:

In traditional statistical quality control (SQC), 100 % inspection schemes have been (almost) abandoned. This is due to the inspection process is time consuming—even with the emergence of the Coordinate Measurement Machines (CMMs)—and therefore, is costly. However, in fault detection based on digital images, 100 % is useful as: 1) false alarm rates and missed-alarm rates are significantly higher than in traditional fault detection techniques; 2) in the case of the proposed algorithm, the images are captured—anyways—and therefore, utilizing these images is consistent with concurrent quality initiatives that are based on waste reduction; 3) even for traditional SQC, in the case of an alarm, there has been an increasing trend to undergo variable sampling times in response to these alarms (Reynolds 1996). Therefore, a real-time detection is beneficial for any organization as they will be reducing waste and in the process increasing the efficiency through decreasing the response time for any process drifts (causing product faults).

Causes:

There are two evident reasons for a relatively low processing speed, which are limited data reduction and inefficient programming, as shown in Figure 9. These two causes are, according to the authors' experience, the most common reasons for a relatively low processing speed when the system is capable.

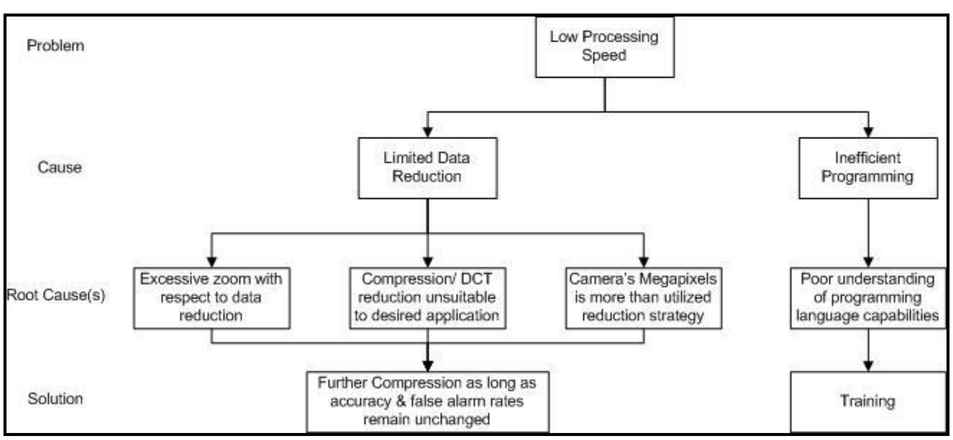


Figure 9 Low Processing Speed Diagnostics Chart

General Trouble Shooting Techniques:

The user should always question how to write the code in the most efficient method, as inefficient programming could be computationally intensive. As a result, inefficient programming could prevent any digital image processing algorithm from working up to speed even if the parameters are properly selected. Therefore, the user is strongly recommended to be fully acquainted with the programming language that will be used in order to ensure that the user could write the code to be as efficient as possible. For example, in MATLAB, vectorization is much more efficient than the use of loops or sequence of nested loops. McAndrew (2004) provided an example, where the time elapsed to calculate the sine values for all integers between 1 and 1 million radians was 1.35 seconds, which is significantly less than 27.5 seconds through using a ‘for’ loop. These seemingly trivial details can lead to a significant decrease in processing time, as shown by McAndrew’s example, and thus, can be more effective than data reduction techniques especially because the algorithm’s accuracy—in terms of false and missed alarms—is not jeopardized. However, the inefficient programming problems are outside the scope of this paper and thus, will not be discussed further.

The limited data reduction, on the other hand, is a quantifiable problem and thus, through understanding how each reduction stage contributes to the overall processing time of the algorithm and the accuracy of the results, the users will be able to make the best decision for their application. The following table provides a summary of the contribution of each reduction stage.

Table 1 Effects of Reduction on Algorithm’s Performance

		Baseline	Best Exp. Results	Compression (4 X, No zoom)			DCT Reduction (4 X, No zoom)					
		4X Red.	1.5X Z., 8X Comp.	0X	8X	16X	10%	20%	40%	60%	80%	95%
Acc.	False Alarms	2 %	0%	0%	0%	0%	2%	2%	2%	2%	2%	2%
	Missed Alarms	34 %	12%	42%	46%	46%	38%	38%	38%	38%	38%	40%
Time (sec)		0.96	1.06	1.95	0.51	0.40	0.94	0.93	0.92	0.92	0.90	0.89

Based on the table, it is clear that the best results are when the camera is focusing on the pan (1.5 X Z.) and the resulting image is compressed 8 times. In that case, over a 50 healthy image span, there are not any false alarms and the fault detection rate is 88 %. As mentioned in the introduction, it is not our purpose, to change the settings of the manufacturer; however, these numbers reflect how the code performs in optimal camera zoom. It should also be noted that the 88 percent fault detection is a very solid result as the lighting conditions as well as the pan positions are changing to reflect expected variations when the code is implemented in the shop-floor. It could also be inferred from the table that the most significant processing time reduction is obtained through the image compression stage. This aligns with the expected results as the data is reduced in an early stage of the code implementation and thus, has a greater impact than when the data is reduced in the later stages of the code.

Problem 3: Fault Detection Inadequacies

Significance:

For the purpose of the proposed algorithm, this is the least significant problem. Any percentage of fault detection is directly linked to system improvement and thus, even a mediocre 10 percent fault detection improves the system by 10 percent with no additional cost for implementation. Thus, this algorithm could be thought of as an excellent six sigma initiative, which improves productivity through waste reduction without adding any additional costs to the manufacturing process. In the following subsections, an overview of the techniques that could be used to improve the fault detection performance of the algorithm—without increasing the false alarm rates and processing speed—will be presented to allow the user to effectively and efficiently utilize the proposed system.

Causes and Troubleshooting:

In general most of the fault detection accuracy problems are caused by the clustering problems discussed in the high false alarm rates section. Therefore, if the user ensures that the false alarm rate is in control, then the accuracy would automatically improve.

Minimizing false alarms rate usually ensures that the code is performing near its best—for the current image acquisition settings (camera and environmental conditions).

Since the motivation of this paper is to utilize current capabilities to detect product faults; the environmental conditions are assumed to be unchangeable. However, the camera settings can usually be changed without affecting the functionality of the robot arms and thus, if the zoom can be changed the between clusters variation will significantly increase leading to better fault detection accuracy. Thus, at this stage, the three variables (false alarm rates, computational speed and fault detection accuracy) are within control and the system is ready to be utilized in detecting real-time faults.

2.6 Results and Discussion

In this section, the detailed results for the using the algorithm to detect quality faults for the pans are provided. As a reminder, in the previous sections, the code was to be run on a set of hundred images (50 healthy and 50 faulty).

2.6.1 Validation of Pilot Study Settings

To emulate manufacturing requirements, two functional constraints have been added to the system: the false alarm rate should be less than or equal to two percent and the processing time of the algorithm should not exceed one second. Since the algorithm was only tested in a pilot study; the following three assumptions are made: 1) in the manufacturing floor, any changes in the lightening conditions would be captured within the two different lightening conditions interval used in the image acquisition stage; 2) the 1.5X zoom used in this stage does not adversely affect the functionality of the robot arm; 3) the performance of the code will be the same for Phase A and Phase B. These three assumptions are justified as in a manufacturing plant, the variations in the lightening conditions are smaller than the variations applied to the image acquisition process and thus, the system is expected to maintain its performance. As for the 1.5 X zoom, it is a reasonable camera setting given that the camera is mounted 3 feet from the part and thus, it lies within the expected setting range for part location purposes. The final assumption will always remain valid if the data collection (image acquisition) encompasses the product's history—similar to Phase I measurement data in a traditional SPC. Therefore, there is no statistical or functional reason, which prevents the code from maintaining compatible results with the results achieved throughout this experiment.

2.6.2 Step by Step Results

The achieved results are really encouraging; a summary of the results is provided in Figure 10. This summary is compromised of two parts. In the first part, the dimensions of the most notable parameters in the algorithm are provided in order to provide future users with a baseline for how the code works. In the second part, a graph representing the projection of the 100 images in the reduced space is presented to show the missed alarms and the formation of the healthy and faulty images.

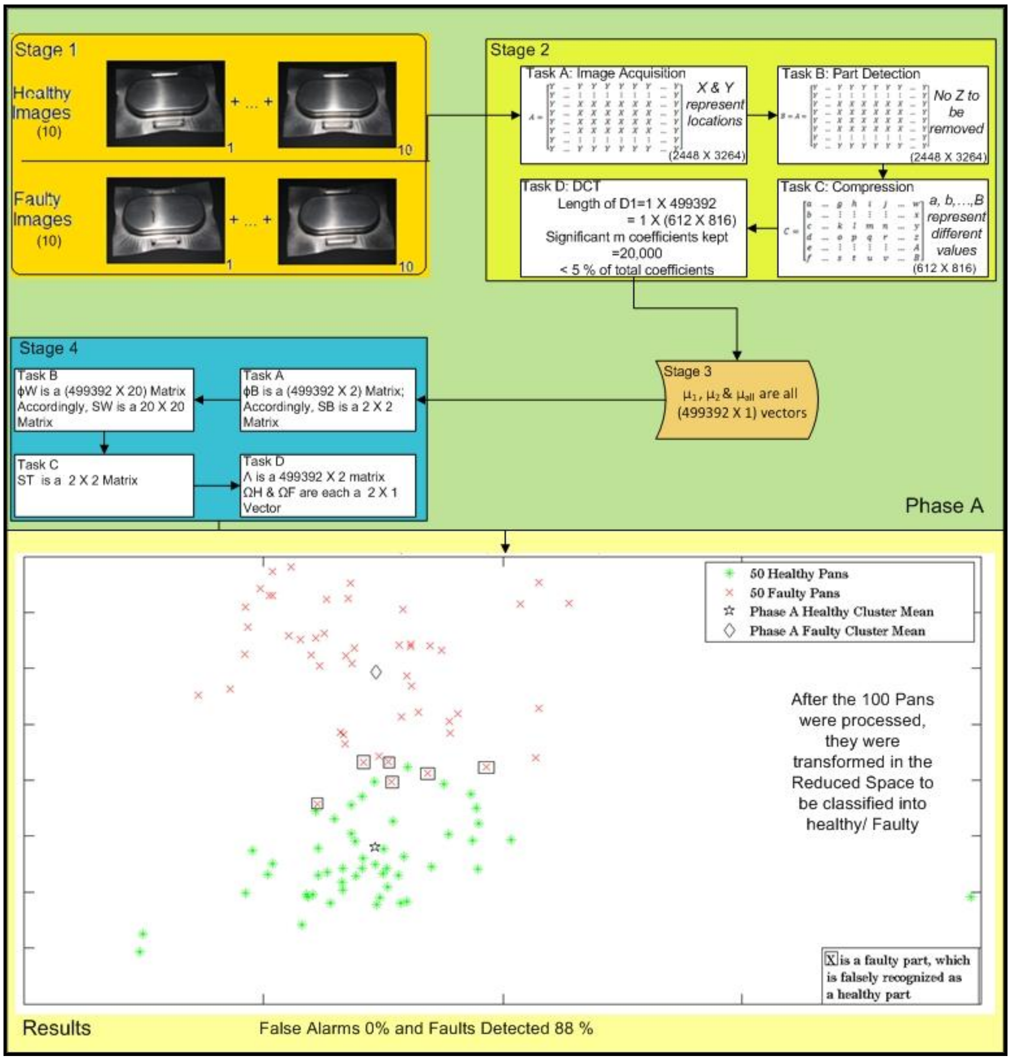


Figure 10 Results Summary

2.7 Conclusions

This paper presents a new methodology for fault diagnosis quality problems using image recognition techniques. The proposed method does not consider any particular feature detection technique but image classification based on a holistic approach. The proposed method combines two image data dimension reduction techniques. First, every image is filtered using a Discrete Cosine Transformation. Then, a projection sub-space is constructed using labeled training images. Each new image that requires classification is filtered and then projected onto the sub-space. The Euclidean distance between the projected image and the previously defined training classes is calculated. The image is classified according to the shortest distance to a class. It can be concluded that for aluminum stamped parts crack detection a DCT combined with a FLD projection method is computational efficient to detect faulty parts in real-time (under 1s). It is also shown that the proposed techniques provide a feasible solution for fault detection in a manufacturing environment, especially in cases where images are collected for other purposes such as robot localization. The strength of this algorithm was shown in its very low false alarm rate accompanied with relatively high fault detection accuracy; up to 88 % of the faults were detected in one experiment. Therefore, the proposed algorithm is a reliable tool that complements image processing methods currently used in industry. Due to its real-time detection capabilities, and its application to existing images on the shop-floor, this proposed algorithm is able to increase fault detection and response speed. Thus, the algorithm creates the possibility for significant cost-reduction in the manufacturing environment.

2.8 References

Abdallah, A (2007). "Investigation of new techniques of face detection", (Virginia Polytechnic Institute and State University).

Abdel-Qader, I., Pashaie-Rad, S., Abudayyeh, O., and Yehia, S. (2006). "PCA-Based algorithm for unsupervised bridge crack detection", *Advances in Engineering Software,* 37 (12), 771-778.

Abdel-Qader, L., Abudayyeh, O., and Kelly, M. E. (2003). "Analysis of edge-detection techniques for crack identification in bridges", *Journal of Computing in Civil Engineering,* 17 (4), 255-263.

Ayenu-Prah, A. and Attoh-Okine, N. (2007). "Exploring pavement crack evaluation with bidimensional empirical mode decomposition - art. no. 65760Q", *Independent*

Component Analyses, Wavelets, Unsupervised Nano-Biomimetic Sensors, and Neural Networks V, 6576, Q5760-Q5760.

Banks, Jerry (1989). *Principles of quality control* (New York: Wiley) xxx, 634 p.

Belhumeur, P. N., Hespanha, J. P., and Kriegman, D. J. (1997). "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection", *Ieee Transactions on Pattern Analysis and Machine Intelligence,* 19 (7), 711-720.

Bursanescu, L., Bursanescu, M., Hamdi, M., Lardigue, A., and Paiement, D. (2000). "Three-dimensional infrared laser vision system for road surface features analysis", *Romopto 2000: Sixth Conference on Optics,* 4430, 801-808.

Canny, J. (1986). "A Computational Approach to Edge-Detection", *Ieee Transactions on Pattern Analysis and Machine Intelligence*, 8 (6), 679-698.

Chang, C. Y., Li, C. H., Chang, J. W., and Jeng, M. D. (2009). "An unsupervised neural network approach for automatic semiconductor wafer defect inspection", *Expert Systems with Applications,* 36 (1), 950-958.

Chen, L. F., Su, C. T., and Chen, M. H. (2009). "A Neural-Network Approach for Defect Recognition in TFT-LCD Photolithography Process", *Ieee Transactions on Electronics Packaging Manufacturing,* 32 (1), 1-8.

Chumakov, R. (2008). "An artificial neural network for fault detection in the assembly of thread-forming screws", *Journal of Intelligent Manufacturing,* 19 (3), 327-333.

Colmenarez, A. J. and Huang, T. S. (1996). "Maximum likelihood face detection", *Proceedings of the Second International Conference on Automatic Face and Gesture Recognition*, 307-311.

Demant, Christian, Streicher-Abel, Bernd, and Waszkewitz, Peter (1999). *Industrial image processing : visual quality control in manufacturing* (Berlin ; New York: Springer) 353 p.

Er, M. J., Chen, W. L., and Wu, S. Q. (2005). "High-speed face recognition based on discrete cosine transform and RBF neural networks", *Ieee Transactions on Neural Networks,* 16 (3), 679-691.

Fuente, E., Trespademe, F. M. , and Gayubo, F. (2003). "Detection of small splits in car-body manufacturing", *IASTED International Conference* (Rhodes, Greece), 354-359.

Garcia, C. (2005). "Artificial intelligence applied to automatic supervision, diagnosis and control in sheet metal stamping processes", *Journal of Materials Processing Technology*, 164, 1351-1357.

Gonzalez, Rafael C. (2004). *Digital Image processing using MATLAB*, eds Richard E. Woods and Steven L. Eddins (xiv, 609 p. :; Upper Saddle River, N.J. :: Pearson Prentice Hall).

Gonzalez, Rafael C. and Woods, Richard E. (2008). *Digital image processing* (3rd edn.; Upper Saddle River, N.J.: Prentice Hall) xxii, 954 p.

Graf, H. P., Chen, T., Petajan, E., and Cosatto, E. (1995). "Locating Faces and Facial Parts", *Proceedings of Third IEEE International Conference on Computer Vision and Pattern Recognition*, 41-46.

Hafez, Mohammed and Azeem, Sherif Abdel (2002). "Using adaptive edge technique for detecting microaneurysms in fluorescein angiograms of the ocular fundus", *11th IEEE Mediterranean Electronical Conference (Melecon 2002), May 7, 2002 - May 9, 2002* (Proceedings of the Mediterranean Electrotechnical Conference -

MELECON, 1; Cairo, Egypt: Institute of Electrical and Electronics Engineers Inc.), 479-483.
Hjelmas, E. and Low, B. K. (2001). "Face detection: A survey", *Computer Vision and Image Understanding,* 83 (3), 236-274.
Ho, S. K., White, R. M., and Lucas, J. (1990). "A vision system for automated crack detection in welds", *Measurement Science & Technology,* 1 (Copyright 1990, IEE), 287-294.
Huang, Wei and Kovacevic, Radovan (2009). "A neural network and multiple regression method for the characterization of the depth of weld penetration in laser welding based on acoustic signatures", *Journal of Intelligent Manufacturing*.
Joo, Segyeong, Moon, Woo Kyung, and Kim, Hee Chan (2004). "Computer-aided diagnosis of solid breast nodules on ultrasound with digital image processing and artificial neural network", *Conference Proceedings. 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 1-5 Sept. 2004* (Conference Proceedings. 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE Cat. No.04CH37558), Vol.2; Piscataway, NJ, USA: IEEE), 1397-1400.
Lee, J. H., Lee, J. M., Park, J. N. , and Moon, Y.S. (2008). "Efficient algorithm for automatic detection of a crack on a concrete bridge", *23rd International Technical Conference on Circuits/ Systems, Computers and Communications*, 1213-1216.
Li, K. Y., Xiang, Y., Yang, X. D., and Hu, J. M. (2004). "Extracting Pathologic patterns from NIR breast images with digital image processing techniques", *Medical Imaging and Augmented Reality, Proceedings,* 3150, 62-69.
Li, S. Z. (2001). *Markov random field modeling in image analysis* (Computer science workbench; Tokyo: Springer) xix, 323 p.
Liang, Tan Kok, Tanaka, Toshiyuki, Nakamura, Hidetoshi, and Ishizaka, Akitoshi (2007). "A neural network based computer-aided diagnosis of emphysema using CT lung images", *SICE(Society of Instrument and Control Engineers)Annual Conference, SICE 2007, September 17, 2007 - September 20, 2007* (Proceedings of the SICE Annual Conference; Takamatsu, Japan: Society of Instrument and Control Engineers (SICE)), 703-709.
Liu, Z. Q., Austin, T., Thomas, C. D. L., and Clement, J. G. (1996). "Bone feature analysis using image processing techniques", *Computers in Biology and Medicine,* 26 (1), 65-65.
Lotlikar, Rohit and Kothari, Ravi (2000). "Fractional-step dimensionality reduction", *Ieee Transactions on Pattern Analysis and Machine Intelligence,* 22 (6), 623-627.
Lowry, C. A., Woodall, W. H., Champ, C. W., and Rigdon, S. E. (1992). "A Multivariate Exponentially Weighted Moving Average Control Chart", *Technometrics,* 34 (1), 46-53.
Lu, Juwei, Plataniotis, Kostantinos N., and Venetsanopoulos, Anastasios N. (2003). "Face recognition using LDA-based algorithms", *Ieee Transactions on Neural Networks,* 14 (1), 195-200.
McAndrew, Alasdair (2004). *An introduction to digital image processing with MATLAB* (1st edn.; Belmont, CA: Brooks/Cole) p.
Mircic, S. and Jorgovanovic, N. (2006). "Application of neural network for automatic classification of leukocytes", *2006 8th Seminar on Neural Network Applications*

in Electrical Engineering, 25-27 Sept. 2006 (2006 8th Seminar on Neural Network Applications in Electrical Engineering (IEEE Cat No. 06EX1402C); Piscataway, NJ, USA: IEEE), 4 pp.

Page, E. S. (1954). "Continuous inspection scheme", *Biometrika*, 41, 100-115.

Pan, Zhengjun, Rust, Alistair G., and Bolouri, Hamid (2000). "Image redundancy reduction for neural network classification using Discrete Cosine Transforms", *International Joint Conference on Neural Networks (IJCNN'2000), July 24, 2000 - July 27, 2000* (Proceedings of the International Joint Conference on Neural Networks, 3; Como, Italy: IEEE), 149-154.

Reynolds, M. R. (1996). "Shewhart and EWMA variable sampling interval control charts with sampling at fixed times", *Journal of Quality Technology*, 28 (2), 199-212.

Rowley, H. A., Baluja, S., and Kanade, T. (1998). "Rotation invariant neural network-based face detection", *Proceedings 1998 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 23-25 June 1998* (Proceedings. 1998 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No.98CB36231); Los Alamitos, CA, USA: IEEE Comput. Soc), 38-44.

Rowley, Henry A., Baluja, Shumeet, and Kanade, Takeo (1996). "Neural network-based face detection", *Proceedings of the 1996 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June 18, 1996 - June 20, 1996* (Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition; San Francisco, CA, USA: IEEE), 203-208.

Selek, M., Sahin, O. S., and Kahramanli, S. (2009). "Using artificial neural networks for real-time observation of the endurance state of a steel specimen under loading", *Expert Systems with Applications*, 36 (4), 7400-7408.

Shima, Y., Kashioka, S., and Yasue, T. (1989). "Consideration on automatic defect detection algorithm for stamped patterns in electronic parts", *Systems and Computers in Japan*, 20 (Copyright 1990, IEE), 48-58.

Sirohey, S.A (1993). "Human Face Segmentation and Identification", (University of Maryland).

Sirovich, L. and Kirby, M. (1987). "Low-dimensional procedure for the characterization of human faces", *Journal of the Optical Society of America A (Optics and Image Science)*, 4 (Copyright 1987, IEE), 519-524.

Turk, Matthew and Pentland, Alex (1991). "Eigenfaces for Recognition", *Journal of Cognitive Neuroscience*, 3 (1), 71-86.

Viennet, V and Soulie, F. F (1998). *Connectionist methods for human face processing, in Face Recognition: From Theory to Application* (Berlin/ New-York: Springer-Verlag).

Woodall, W. H. (2000). "Controversies and contradictions in statistical process control", *Journal of Quality Technology*, 32 (Copyright 2000, IEE), 341-350.

Yang, Ming-Hsuan, Kriegman, David J., and Ahuja, Narendra (2002). "Detecting faces in images: A survey", *Ieee Transactions on Pattern Analysis and Machine Intelligence*, 24 (1), 34-58.

Yu, H. L., MacGregor, J. F., Haarsma, G., and Bourg, W. (2003). "Digital imaging for online monitoring and control of industrial snack food processes", *Industrial & Engineering Chemistry Research*, 42 (13), 3036-3044.

Zhao, Q. and Camelio, J. (2008). "Quality monitoring and fault detection on stamped parts using DCA and LDA image recognition techniques", *International Conference on Manufacturing Science and Engineering, MSEC* (Illinois, USA.).

Zhao, W., Chellappa, R., and Phillips, P. J. (1999). "Subspace linear discriminant analysis for face recognition", (College Park: Center for Automation Research, University of Maryland).

Zhao, W., Chellappa, R., Phillips, P. J., and Rosenfeld, A. (2003). "Face recognition: A literature survey", *Acm Computing Surveys,* 35 (4), 399-458.

Zobel, M., Gebhard, A., Paulus, D., Denzler, J., and Niemann, H. (2000). "Robust facial feature localization by coupled features", *Proceedings of the Fourth International Conference on Automatic Face and Gesture Recognition, 28-30 March 2000* (Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580); Los Alamitos, CA, USA: IEEE Comput. Soc), 2-7.

3 Image-Based Control Charts: A Review with Discussion

3.1 Abstract

The advent of new technologies often results in the ability to collect more data about the process/product being monitored. The availability of higher dimensional-more frequent data raises the need for novel, innovative, and specialized quality monitoring tools that effectively utilize the data to better understand and control the production process.

Machine vision systems are one of the ever-increasingly utilized technologies in providing data that not only explains the traditional dimensional parameters, but also provide information on product geometry, surface defects, and surface finish. The integration of these new parameters into the SQC framework provides an opportunity for decreasing defect rates, and increasing process efficient and profitability through detecting faults, which were not monitored by traditional control charts. In this paper, a thorough review of image-based control charts is provided, along with ideas for future work, to set the groundwork needed for incorporating dimensional and surface data into one robust quality framework.

3.2 Introduction

Continuous improvement has played an important role in manufacturers' quest to constantly increase their production efficiency through the reduction of process/product variability. Continuous improvement is not an exclusive feature to manufacturing operations as it also shaped the process of quality control; specifically, key statistical process control (SPC) tools control charts and measurement tools have been continuously enhanced to allow manufacturers to achieve better results on the shop-floor. For a long period, the basic X-bar and R charts (Shewhart charts) for variables have been recognized as the standard optimal control charts. Deming (1993) echoed this sentiment, "The Shewhart control charts do a good job under a wide range of conditions. No one has yet wrought improvement". However, sixteen years later, industry disapproves this sentiment as explained by the various control charts that are currently utilized on the shop-floor. These control charts could be categorized into two major categories: 1) Univariate control charts, which have been proven to surpass the Shewhart's charts in detecting small process trends such as CUSUM and EWMA charts; and 2) Specialized control charts, which have demonstrated better representation of complex manufacturing operations such as MEWMA—proposed by Lowry et al. (1992)—and profile monitoring techniques (for thorough reviews consult Kim et al. (2003), Woodall et al. (2004) and Woodall (2007)).

With the advent of new technologies, more data is collected about the process/product to be monitored, which raises the need for innovative and specialized quality control techniques that effectively utilize the data to understand more about the process/product. Woodall et al. (2004) showed that manufacturing operations are better characterized in terms of a profile—a relationship between a response variable and one or more explanatory variables—instead of the traditional assumption that a single univariate quality characteristic or a general multivariate distribution could adequately represent the quality of process/product. On the other hand, Feng and Kapur (2008) have showed that the current advancements in sensing and computer technologies have lead to reduced time between samples, and as a result 100 % sampling became a reality as manifested by the development of the Coordinate Measurement Machines (CMMs) for dimensional data. The consequences of this (reduced time between sampling) is the strong correlation between successive samples, which violates the assumptions of the standard univariate control charts and leads to the incorporation of statistical approaches to remove the correlation before monitoring the data via control charts (see e.g. Alwan and Roberts (1988)).

Machine vision technologies have been gaining popularity over the last two decades. They are increasingly employed in different inspection (quality assurance) applications as will be explained in the Background section. The strength of these systems is in their ability to provide data that not only explain traditional dimensional parameters, but can also provide information about product geometry, surface defects and surface finish in manufacturing; tumors in medicine; and crack growth in civil engineering. The successful utilization of this data provides immense opportunities for further advancements in SPC as the monitoring and control of industrial products could be extended to measuring the aesthetical aspects of the product, in addition, to the traditional manufacturing/design requirements. This addition of aesthetical aspects could lead to significant improvements in the SPC framework of many companies, especially if the application of the machine vision systems (MVS) is extended beyond inspection.

In this paper, we aim at providing a thorough review of image-based control charting techniques (input is a single image—Multi-Image based control charting is not covered by this review), which are indispensible in extending the application of MVS beyond inspection. These control charts cover the application spectrum of MVS (manufacturing, medical diagnosis and civil engineering), which follows the current trend of extensive SPC applications outside manufacturing (e.g. Woodall (2006)). There is a need behind this review paper, despite the novelty of the application, as the reviewed literature (to be discussed later in the paper) showed that each of the control charting techniques for image-data was developed in isolation. This lack of communication between the different researchers in the different applications hinders the progression and the development of control charting techniques for the most effective utilization of this (MVS) technology. We hope that this compilation, as well as the presented discussion, will build the foundation needed for future work that would consequently fill the gaps in the current image-based control charting techniques. In addition, this will allow for the effective implementation of MVS in SPC for the different industrial and medical applications.

It is assumed that the reader is familiar with the construction and use of control charts (for detailed introductions see Wheeler and Chambers (1992), Woodall and Adams (1998), or Montgomery (2009)). On the other hand, it is assumed that the reader is not familiar with the issues involved with image data and therefore, an overview is presented in the following section.

3.3 Background

This section is divided into four subsections. In the first subsection, an overview of machine/computer vision systems (MVS) is presented, where the differences between computer vision and human vision is highlighted. Then, in the following two subsections, the basic/relevant fundamentals of 2-D and 3-D images are provided. Finally, some of the relevant inspection/fault detection works are underlined as they set the tone for the literature reviewed in the subsequent sections.

3.3.1 Machine/Computer Vision Systems: An Overview

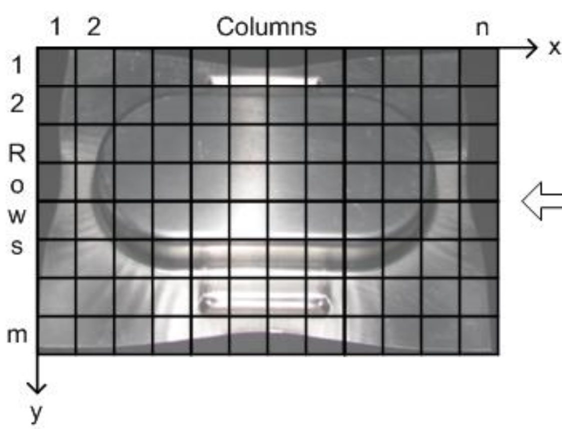
There are many different perspectives to describe machine vision systems; however, in terms of quality control, they are considered data acquisition devices that are capable of playing a critical role in the future of manufacturing and quality management. Zuech (2000) builds upon this perspective by stating that the data from machine vision systems is the foundation for Computer Integrated Manufacturing (CIM), where the vast amount of manufacturing data originating from numerous assembly line sensors is integrated and monitored beyond traditional part inspection and introduced into new fields such as product design matching, process planning, and production processes. These areas can be viewed as novel explanatory variables that assist in the formulation of a better quality model (profile) that can be used to monitor the process/product.

In spite of the machine vision systems' increasingly important role in manufacturing, they remain extremely complex as they combine two different fields of computer science. The first field deals with image acquisition/ processing, where an image is formed to mimic the object scene, through the use of an image capturing device—such as a charge-coupled device (CCD)/camera or an x-ray (image capturing devices cover all the range of the electromagnetic spectrum)—and then, the image is enhanced and objects are extracted from the image through the digital image processing activities (explained in more details in the next two subsections). The second, and final, field is image analysis and understanding, which is a branch of artificial intelligence (AI) where the cognitive functions normally associated with human vision are emulated and automatically performed via a computer (Gonzalez and Woods (2007)).

In spite of the similarities in notation between human and machine vision, there are significant differences between them. Zuech (2000) stated that current machine vision systems are primitive when compared to the eye-brain capacity as the current MVS are prone to variations in lighting conditions, reflection, minor changes in texture among other tasks that the human eye is immune to. Despite these limitations, machine vision systems are increasingly used in manufacturing as they are much more superior to the human eye in: 1) Monitoring high production rates; 2) Performing multiple simultaneous tasks to different objects; 3) Immunity to fatigue (see Zuech (2000) and Gonzalez and Woods (2007) for detailed discussions). The aforementioned advantages position MVS as a much more attractive option when compared to visual inspection, which explains their widespread utilization and the consequent need for SPC monitoring techniques.

3.3.2 2-D Image Fundamentals

A 2-D image is often represented as a 2-D function, f(x, y), where x and y are the spatial coordinates on the image and the value of f at any pair of (x, y) is called the intensity. In general x, y and f could be continuous; however, in this paper, digital images (ones that could be stored and read by a computer) are the focus and thus x, y and f are all finite discrete integers. A property of digital images is that they are composed of a finite number of elements (pairs of (x, y)), which are commonly referred to as pixels. The number of pixels in an image (usually noted as m*n where m and n are the number of pixels in the vertical and horizontal directions respectively) is directly proportional to the resolution of the image capturing device. Accordingly, an image "A" could also be represented as a matrix as shown in Figure 11.



$$\Longleftrightarrow \quad A = \begin{bmatrix} f(1,1) & f(1,2) & \cdots & f(1,n) \\ f(2,1) & f(2,2) & \cdots & f(2,n) \\ \vdots & \vdots & \vdots & \vdots \\ f(m,1) & f(m,2) & \cdots & f(m,n) \end{bmatrix}$$

Figure 11 Matrix Representation of Images

Since digital images could either be black and white, grayscale or colored, the intensity should reflect those different possibilities. In the case of a black and white (binary) image, f(x, y) is only capable of taking a value of 0 (black) or 1 (white). On the other hand, if the image is grayscale, then f(x, y) takes any integer value between 0 (black) and 255 (white). Finally, when the image is colored, f(x, y) is not an element, but a vector of (red, green and blue) individual components and thus, many of the techniques used for binary and grayscale images are easily extended to colored images by simply processing the three component images individually (Gonzalez et al. (2004)).

The next logical step after classifying the image visualization type is how information is extracted from the image. Snyder and Qi (2004) showed that 2-D image information is represented in six different ways, two of which are of relevance: 1) the matrix representation (shown in Figure 11) and 2) the spatial frequency representation. Digital image processing is the field of computer science that is used to extract information from images, especially from a matrix format. Gonzalez and Woods (2007) have categorized digital image activities into three levels: 1) Low-level processes—these are basic operations of noise reduction and contrast enhancement, where the inputs and outputs are images; 2) Mid-level processes—processes where the inputs are images and the outputs are attributes extracted from images such as edges, parts and contours; and 3) High-level operations—processes that aim at understanding and "making sense" of the data. It should be noted that the high-level operations can also be applied to the spatial frequency data—image data transformed to the frequency domain (e.g. see Jahne (2005)—and that typically involves the use of statistical methods to show that the data follows a certain distribution, or to cluster the data into different categories (examples are provided in the fault detection subsection).

There are a number of internal and external factors that significantly affect the performance of any 2-D digital image processing technique; the most relevant three factors are discussed. The first factor is an inherent property of any image, where neighboring pixels are usually extremely correlated. This problem is referred to, in the digital image processing field, as interpixel redundancy and is the prime motivation behind image compression techniques, which aim at maximizing the information gained while minimizing the data stored. The two remaining factors, which are changes in part location and lighting conditions, represent how the environment/external conditions can affect the accuracy of machine vision systems. These two factors are of high relevance to manufacturing and civil engineering applications for fault detection scenarios. Additional introductory material on 2-D images can be found in Nalwa (1993), Jahne (2005), Gonzalez and Woods (2007) and Pratt (2007).

3.3.3 3-D Image Fundamentals

The term 3-D image seems contradicting as an image is always thought of as a 2-D object. However, as explained in Nikolaidis and Pitas (2001), a more appropriate term for 3-D images are 3-D data sets (matrices) that reflect different 3-D measurements on the part represented by the image. Toriwaki and Yoshida (2009) provided some examples of 3-D images in the medical field: 1) CT (Computer Tomography) scans; 2) MRI (Magnetic Resonance Imaging) and 3) 3-D ultrasound imaging. As for manufacturing and civil engineering, 3-D laser scanner images, illustrated in Figure 12, are the most common, with the exception of the textile industry where CT images are more common. It could be inferred from these examples that 3-D images can either represent the details of a body/part, as in CT and MRI, or represent a group of surfaces of the body as in 3-D laser scanners, which do not provide any information on the internal structure of the part.

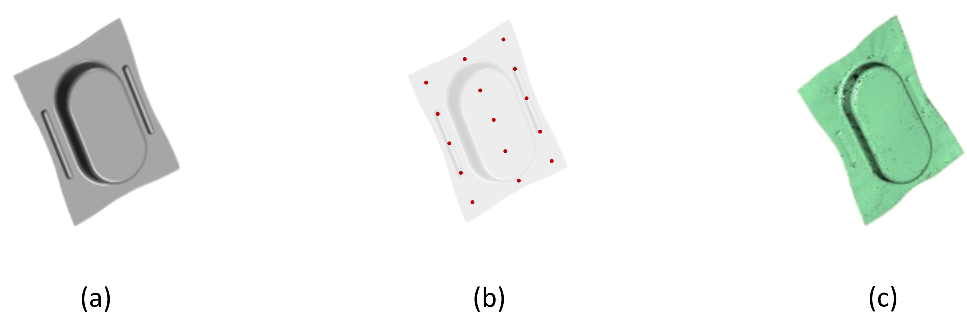


Figure 12 a) Represents a CAD model of a metal formed production part, (b) a CMM measurement system only measures a set of specified points, (c) an actual laser scan that captures all product characteristics

The concepts applied to 2-D images can also be extended, with some differences, to 3-D images. Therefore, a 3-D image can be represented as a 3-D function, f (x, y, z), where f is a characteristic value at a point (x, y, z). Toriwaki and Yoshida (2009) provide a detailed description of the major differences between 2-D and 3-D images; the most relevant two are discussed: 1) As opposed to 2-D images, the meaning of the value function f is different according to the measurement technology. For example, in an x-ray CT image, f is the absorption factor to the X-ray of the part as opposed to in MRI, where f is the strength of magnetic resonance at the part. 2) The building block of 3-D images is called a voxel (see Figure 13 a), which is a small cubic cell that divides the 3-D space. According to the definition of the voxel, the 3-D space is represented as a cubic lattice and therefore, the neighborhood concept becomes more complex. For any given voxel, there are 50 other voxels which are neighbors to it; 6 of them (referred to as 6-neighborhood, shown in Figure 13 b) only differ by ± 1 in one of the coordinates, 18 (shown in Figure 13 c) of them differ by ± 1 in two of the coordinates and the remaining 26 (shown in Figure 13 d) differ by ± 1 in three coordinates. Therefore, neighborhood relationships are weaker than in 2-D images.

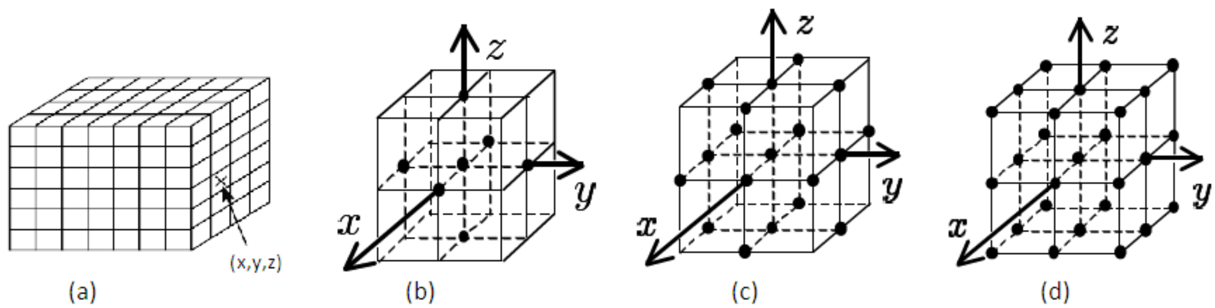


Figure 13: a) 3-D building block (voxel) b) 6-Neighborhood Representation c) 18-Neighborhood Representation d) 26-Neighborhood Representation (Toriwaki and Yoshida (2009))

3-D image information, as for 2-D images, is extracted and analyzed using 3-D image processing techniques, which are similar to the corresponding 2-D techniques. Nikolaidis and Pitas (2001) provide two different methods for processing 3-D images; first of which is to apply a series of 2-D processing operations on consecutive 2-D slices of the 3-D image; however, they do not recommend this approach as it does not take into consideration the spatial relationships and correlations in the 3-D data. Therefore, they recommend using true 3-D image processing techniques. These techniques are categorized into three levels defined by Gonzalez and Woods (2007), where the operations initially focus on image enhancement activities and end with image understanding activities; however, they only differ from the corresponding 2-D activities in terms of how the algorithms work (e.g. consult Nikolaidis and Pitas (2000), Jan (2006) and Toriwaki and Yoshida (2009)).

As in 2-D images, there are issues which effect the performance of 3-D digital image processing systems. There are three main issues in 3-D digital image processing (Nikolaidis and Pitas (2001) and Toriwaki and Yoshida (2009)): 1) Image Compression—as 3-D images involve much more data than 2-D images; 2) Challenges in the digitization of the images by the computer due to the complexity of 3-D digital geometry; and 3) Limitation of the Human capabilities in viewing the internal structure of 3-D digital images, which can provide major problems in medical diagnosis. Since the focus of this paper is on showing image-based control chart applications, image compression is the most significant problem that faces automated systems and the two other limitations are mentioned for the sake of completion.

3.3.4 Image-Based Inspection/ Fault Detection: A very brief overview

Fault detection methods can be categorized in a number of different ways; for example, by digital image processing techniques, application areas, image types (2-D or 3-D), or image data representations (matrix or spatial frequency data). In this section, the classification by Megahed and Camelio (2009)—based on the face detection classification developed by Hjelmas and Low (2001)—is adopted as it provides a generic framework that focuses on the key digital image processing techniques used in inspection/fault detection technique rather than—less SPC relevant categorizations—image types and image data representations.

Megahed and Camelio (2009) classified fault detection techniques into two major groups, feature-based approaches and image-based approaches. The feature based approaches explicitly utilize product knowledge, where the historical product faults (such as: different fault geometries and textures) are exploited at different steps to assist in accurately detecting the expected different faults. They (feature-based methods) use the segmentation technique, which is a digital image processing technique where the image is divided into its constituent subcomponents to detect the different features within the image, such as: part boundary, holes and cracks.

In almost all feature-based methods, the segmentation is followed by a representation and description stage. In this (representation and description) stage, the output of the segmentation stage (raw pixel data) is represented either as edge data or as region data. The edge representation (consult Ziou and Tabbone (1998) for a detailed review) is suitable for applications, where the prime interest is to use the MVS as an optical dimensional measurement device (e.g. see Couweleers et al. (2003) and Deng et al. (2008)). On the other hand, region data representation is usually applied when the focus is on detecting internal properties such as texture—an important characteristic in diagnosing changes in tumor cells in medical images (Lyu et al. 2009), and detecting surface defects in textile manufacturing and integrated circuit (IC) manufacturing (e.g. refer to Ngan and Pang (2009)). Gonzalez and Woods (2007) noted that, in some applications, these two representations complement each other especially when the edge representation is used as a data reduction technique.

The description stage, which follows the representation stage, is used to extract attributes that are utilized in differentiating a class of objects from another; for example, differentiating cracked objects from healthy ones. This stage is usually followed by the final recognition/detection stage, which usually utilizes priory knowledge about the part to detect the faults in the part. The utilization of this priory knowledge (known as model-based approaches) could be a simple operation—such as pointing out an area of an image where the information of interest is known to be found (see e.g. Region of Interest in Rafajlowicz et al. (2008))—or could be a complex method, which for instance lists all known defects in an industrial inspection problem (further examples are discussed in Gonzalez and Woods (2007)).

As for the image-based approaches for fault detection, they directly classify different faults into fault groups/ bins using training algorithms without the need for feature derivation and analysis. Unlike the feature-based approaches, these relatively new techniques incorporate product/process knowledge only implicitly into the detection algorithm through mapping and training schemes. They (image-based approaches) are divided into three categories: 1) neural networks—nonlinear mathematical/ statistical models intensively used, in image processing applications, in pattern recognition. It is extensively used in medical diagnosis, e.g. screening for cervical cancer (Cenci et al. (2000) or in breast tumors (Campos et al. (2005) and Huang et al. (2008)), and in detecting surface cracks in bridges (e.g. Kabir et al. (2008)), and in many industrial inspection applications (most recent examples include Chang et al. (2009), Chen et al. (2009) and Selek et al. (2009)); 2) statistical methods—are methods that are based on maximum likelihood operations to detect faults (a textile manufacturing application was addressed by Campbell et al. (1999)); 3) Linear subspace methods—techniques that apply concepts from multivariate statistics to represent the part, which is a subspace of the image space. These methods utilize three widely-used approaches: principal component analysis (PCA), Fisher-based linear discriminant methods (FLD) and wavelets transform (WT). Linear subspace methods are widely utilized in all three disciplines as in (Bharati and MacGregor (2001), Gurcan et al. (2002), Yu and MacGregor (2004), Abdel-Qader et al. (2006) and Megahed and Camelio (2009)).

It should be noted that the neural networks and the linear subspace methods are the most implemented techniques in fault/abnormality detection, nowadays, as they are seen to be more capable in representing more complex situations, where the number of different faults that could occur is significantly high. This is also justified by the fact that feature-based approaches require an intensive knowledge of possible faults, to be incorporated within the system, for achieving acceptable results.

3.4 Classification of Image-Based Control Charts

As in the Fault Detection subsection, image-based control charts could be classified according to different criteria: control chart technique, data structure, application and preliminary image processing methods. Upon reviewing the relevant literature, we found that the classification by application is the most natural as the performance targets and quality characteristics vary significantly from one field to the other and consequently, the required functionality/detection technique will change. As expected, the broad and specific application areas in which control charts are applied is dictated by the areas in which image processing is heavily used. Therefore, image-based control charts could be classified according to two major domains: 1) manufacturing and 2) medicine; these two domains constitute two of the three major domains, where image processing is applied (manufacturing, construction and medicine).

The application of control charts for image-data is a complicated problem as the target is not to only obtain specific average run lengths, for a definition refer to (Montgomery (2009)), but also to account for the traditional challenges in machine vision. Surprisingly, Rafajlowicz and Steland (2009) represent the first attempt to set a generic framework that could be followed, in any image-based control chart application, to tackle some of the challenges in constructing control charts for image data—mainly interpixel redundancy, background noise and their effect on the expected shift. By utilizing images as data-sources for control charts, the expected shift between the in-control and out-of-control conditions is very small. This observation would lead to high out-of-control ARLs, which would lead to a lot of waste due to the slow detection of the out-of-control condition by the control chart. Accordingly, they suggested a “moving average kind of a p-chart”, which would outperform the traditional p-chart for smaller shifts. Due to the absence of a universal framework that tackles the inherent challenges in image processing, most of the literature focused on the industry specific problems without taking into account the effect of the challenges on their control chart performance as seen in the following sections.

3.5 Image-Based Charts for Manufacturing Applications

Control charts for image data have been utilized in a number of different fields, which include: textiles, LCD monitors (phone displays) and electronic components manufacturing. The control charts for each of these fields are described in one of the subsections below.

3.5.1 Image-Based Charts for Manufacturing Applications: Textiles and Fabrics Manufacturing

Challenges and SPC Targets:

Rafajlowicz et al (2008) coined the monitoring of fabrics production to be the classic example of applying quality monitoring using images. This is justified by the fact that out-of-spec products in the textile industry cannot be remanufactured and thus, the real-time detection of any out-of-spec products could result in significant process yield and profitability improvements. Anagnostopoulos et al. (2001) showed that the main difficulty with computer-based fabrics quality monitoring is the huge diversity in the type of fabrics and their defects. They have provided a detailed list of fifteen different textile faults and their causes, which are great assets for both the fault detection and diagnosis stages of textile quality monitoring (see Chiang et al (2001) for a detailed description of the different procedures in quality monitoring).

Another often missed challenge in quality monitoring of textiles using MVS is the presence of many different classifications for the quality of the finished product. In a traditional manufacturing environment, the finished product is often seen as conforming or non-conforming. On the other hand, the textile industry utilizes a more detailed classification where the finished product is classified into a number of different grades of quality through some sort of a Demerit system (see Montgomery (2009)). The grading is often based on calculating the numbers of major and minor defects as point values per square meter (1973). Therefore, any framework for the quality monitoring of textiles manufacturing should be able to detect shifts based on the fifteen different fault types, and signal only if that resulted in a change of grade. This goal has yet to be achieved and the available literature for quality monitoring has focused more on detecting a subset of the fifteen defects.

Available Literature:

Horst and Negin (1992) were the first researchers to highlight the advantages of control charts in web applications (paper, textiles and plastic films). They suggested that the utilization of Shewhart charts and histograms when combined with machine vision could lead to significant productivity improvements. It could be inferred that this paper was more of a dissemination approach as the authors did not specify what quality characteristics should be represented on the control charts.

Tunak and Linka (2008) provided a robust technique for detecting the occurrence and location of woven defects, on plain weave structure, through utilizing the observation that any woven image is periodic in nature and therefore, can be considered to have directional texture. Based on this observation, they utilized second-order texture statistics (more specifically the Gray Level Co-occurrence Matrix, GLCM, first introduced by Haralick et al. (1973)) as it allows for maintaining both the brightness and spatial information of the parts. Even though Cartensen (1992) showed that fifteen features could be calculated using the standardized GLCM, the authors decided to retain the most significant features based on classification and regression tree techniques, CART (for detailed introduction on CART see Breiman (1984)).

CART lead to the extraction of a set of five significant features from the matrix: energy, correlation, homogeneity, cluster shade and cluster prominence. Since the presence of a defect over texture causes regular structure changes and consequently, statistical changes; therefore, the authors used multivariate control charts—Hotelling T^2 charts—as a tool to integrate multiple texture features and judge the existence of defects as shown in Figure 14. The Hotelling T^2 charts involved ten quality characteristics, each of the five features in the warp and weft directions.

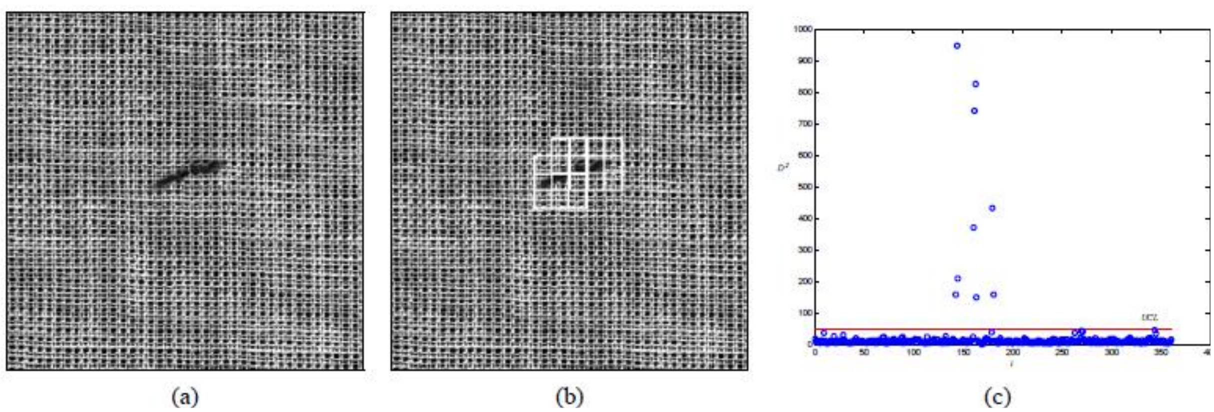


Figure 14 (a) Foreign body defect, (b) result of applied algorithm, (c) plotting the quality characteristics on the T^2 control chart Tunak and Linka (2008)

Tunak et al. (2009) built on the defect detection capability on plain weave structures by relaxing some of the assumptions and adding more features to the SPC algorithm. Mainly, their new algorithm can account for the misplacement of the yarn on the image i.e. possible rotation of the image. In addition, it can be used not only to detect defects associated with the change of weaving density of weft yarns, but also to monitor weaving density in direction of length of fabric; fabrics could also have different pattern (plain, twill, satin). In order to accommodate for these performance targets, the authors have followed a totally different approach than Tunak and Linka (2008).

After the image is captured, Tunak et al. (2009) performed contrast enhancement to increase the difference between the woven fabric and the background. As mentioned earlier in the background, that 2-D image information could be represented either as a matrix representation or as spatial frequency representation. In Tunak and Linka (2008), the matrix representation was used. On the other hand, in Tunak et al. (2009), the spectral (frequency) approach is utilized through transforming the matrix representation to a frequency representation via a 2-D Discrete Fourier Transformation (2-D DFT), which was found to be also suited in representing the directionality of periodic patterns. Through some manipulations for the parameters involving the (2-D DFT) and its inverse, the resulting images can be made to contain only warp or weft set of yarns. Therefore, the restored images can be used for automatic assessing of weaving density. The authors then used control charts (X-bar chart) on the weaving density as a tool to find sites of potential defects. A sample of their results is shown in Figure 15.

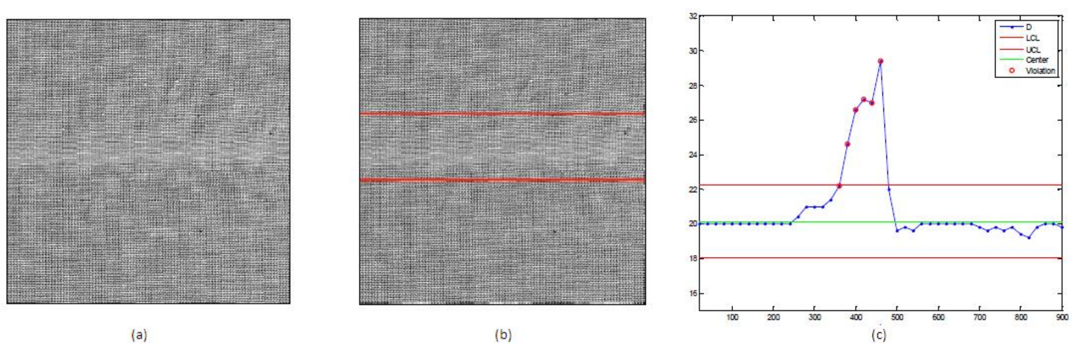


Figure 15 (a) Image of woven fabric in a plain weave with defect weft stripe, (b) defective region, (c) control charts for weaving density of weft yarns (Tunak et al. (2009)).

Future Research Ideas: Image-Based Control Charting Techniques for Textiles

Over the last two decades, there has been intensive research over machine vision applications for the textile industry. More work is needed; however, to investigate the issues related to the statistical basis of image-based control charts and to fill in the gaps of research that has already been done. A summary of specific research ideas and unanswered research questions is provided below in addition to some general research ideas, which are provided at the end of the document.

(1) The textile industry is by far the most active area of application for MVS and therefore, it is really surprising to find out that there is no standard image datasets to allow the researchers to benchmark their techniques. Therefore, comparisons between the different digital image processing techniques—the preliminary step for control charts—are often neglected. We suggest that the textile industry stakeholders (researchers, labor and management) should work on constructing a large dataset for which benchmarking for different types of textiles could be done. This would follow in the footsteps of face recognition, where several datasets have been successfully created for that purpose (see Gross (2005)).

(2) The methods devised by Tunak and Linka (2008) and Tunak et al. (2009) were tested on a sample subset of all possible error types, sizes and locations. Further studies are needed to show their performance with respect to each of these three parameters. In addition, the Average Run length should be recorded for such parameters to see whether the false alarm ratios and the number of samples needed to detect a certain fault given specific values for the previously mentioned parameters is acceptable or not.

(3) Related to the previous research question is the effect of these parameters on the grade of the textile. It might be more interesting for manufacturers to know whether a certain fault would cause a change in the grade or not and therefore, the development of control charts for such a purpose should be explored.

(4) The marginal distribution should be thoroughly investigated to make useful window-size recommendations. This is analogous to the utilization of the marginal distribution in the traditional control charts in deciding a sample size value.

(5) The used control charts (Shewhart and Hotelling T^2 charts) are all known to be good for moderate to large shifts in the quality characteristic. This is contradictory with the suggestions of Rafajlowicz and Steland (2009), where they demonstrated that shifts are usually small when the data-source is an image. Therefore, the use of the EWMA and the CuSUM charts for univariate quality characteristics and the Multivariate EWMA for multivariate quality characteristics should be investigated.

3.5.2 Image-Based Charts for Manufacturing Applications: Liquid Crystal Display (LCD) Manufacturing

Challenges and SPC Targets:

Due to its high quality and space efficiency, TFT-LCD (Thin Film Transistor-Liquid Crystal Display) has become the mainstream display device in many electronic products (Yamazakia et al. (1996)). In many of these products, the aesthetical display of the monitor is the key quality-characteristic demanded by the customer. Accordingly, 100% visual defect inspection is common practice in the manufacturing process (Jiang et al. (2004)). Since the human visual inspection is slow, subjective, and costly, it is often considered the bottleneck (Jiang et al. (2004)) in the LCD manufacturing process. Accordingly, there is an increasing trend to utilize machine vision to overcome the aforementioned limitations (Sokolov and Treskunov (1992), Saitoh (1999), Lin et al. (2001), Jiang et al. (2004), Jiang et al. (2005), Wang and Tsung (2005) and Lin and Chiu (2006)).

Defects on TFT panels are classified into two categories: macro and micro defects (Nakashima (1994)). Lu and Tsai (2005) provided examples for each of the two defect categories. They have defined the appearance of macro defects as regions of high contrast with irregular shapes and sizes such as: MURA (unevenness in the TFT panels), SIMI (stains on TFT panels) and ZURE (misalignment of TFT panels). On the other hand, micro defects include pinholes, fingerprints, particles and scratches, which are generally very small and hard to detect using human inspectors. Accordingly, an effective machine vision system should be able to detect both defect categories. This goal has yet to be achieved and the available literature for image-based control charts has initially focused more on detecting the macro defects as explained in the following paragraphs. It should be noted; however, that detecting macro, especially MURA, faults is not an easy task as they often have no clear contour or contrast. (Taniguchi et al. (2006))

Available Literature:

Wang and Tsung (2005) used profile monitoring techniques to detect changes in a Q-Q plot, which reflected the relationship between a current sample and a baseline conforming sample. The authors used their method, in a case study, to detect macro-defects in mobile-phone LCD panels. It could be inferred that this technique serves only the fault detection (inspection) function of quality monitoring as it does not provide any information on the type, size or location of the fault.

Jiang et al. (2005) provided a method that is capable of overcoming the limitations in Wang and Tsung (2005) through the detection of the type, size and location of MURA defects (> 5 cm^2) in high-grade TFT-LCD panels. They have used luminance measurement equipment to collect the needed data for analysis. Their analysis consisted of two phases: 1) Fault Detection—ANOVA was used to determine the existence of MURA defects and 2) Fault Diagnosis—EWMA control charts were used to determine the location and size of the defects as shown in Figure 16.

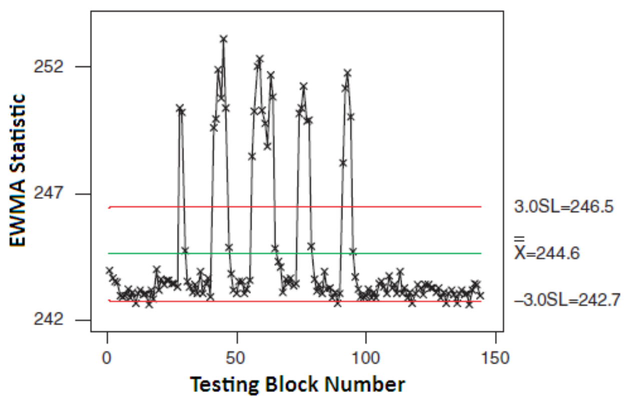


Figure 16 EWMA for MURA defects (Jiang et al. (2005)

It should be noted that there are two main advantages for the Jiang et al. (2005) method: 1) it utilizes equipment that is readily available in the plant; and 2) it is capable of identifying the defect type, location and size, which could be utilized in process improvement. On the other hand, their procedure has three major limitations: 1) the number of blocks would vary significantly if the manufacturer is producing a large variety of LCD panel sizes, which can make the analysis more complicated; 2) the detection of a defective area crossing two of the testing blocks would be difficult; and 3) there is no guidelines for determining the value of the EWMA parameter λ, as the authors determined it by trial and error.

Lin and Chiu (2006) suggested the utilization of Hotelling T^2 chart, among other methods, to overcome the limitations in Jiang et al. (2005). They used the multivariate Hotelling T^2 statistic is to integrate the different coordinates of the color models and computer science algorithmic to determine the regions of small color variation representing the MURA defects. The Lin and Chiu (2006) method raises a concern as the utilization of color images, as the data-source, increases the computational time significantly and therefore, it is not clear whether this method can be used to detect/diagnose faults in real-time. The real-time detection is a really important tool as it allows for 100 % sampling, which is heavily emphasized in the LCD industry according to Jiang et al. (2005). Therefore, the adoption of this method (in its current form) in an actual manufacturing environment is questionable.

Future Research Ideas: Image-Based Control Charting Techniques for LCD monitors

The need for MVS in the quality monitoring stage of LCD is well-documented as it not only immune to human fatigue, but also removes a bottle-neck by speeding up the process. Accordingly, we expect more research in the coming years that will build on the foundations set by the current studies and allow for the detection/diagnosis of more than one fault type. As in the textile section, a list of specific research ideas and unanswered research questions is provided below, which supplements the more general recommendations at the end of this paper.

(1) As in the textile section, the LCD quality monitoring research could significantly improve by the introduction of research datasets. This would not only affect the digital image processing side of the research, but would also improve the quality monitoring component through the introduction of more errors of different sizes, locations and types. In addition, it would allow the researchers to test their techniques on different LCD monitor sizes (a limitation in Jiang et al. (2005)). We envision the datasets to not only have macro and micro defects, but also LCD monitors of significantly different sizes—from cell phone displays to 60’’ TV monitors. This would allow the researchers to evaluate whether their developed

techniques is specific to a certain display size or is it more robust and could be applied industry-wide.

(2) The assumption in the Jiang et al. (2005) paper of having the defects lie inside the testing blocks could be relaxed by optimizing the size of the block and the sampling strategy (number and location of measurement points within the block).

(3) Related to the previous research question what are the effects of the size of the block and the sampling strategy on the OC (Operating Characteristic) curves and the ARLs of the control charts? Also, the question of whether these two parameters could be designed to uphold the concept of rationale sub-grouping without the need for any transformation is another interesting topic that should be explored.

(4) The Lin and Chiu (2006) method could be modified using some digital image processing techniques to ensure that it is up to speed with typical LCD production rates, which could be seen as an alternative approach to building upon the Jiang et al. (2005) method. The successful utilization of an industry-wide dataset would allow for a true evaluation of the two methods.

(5) The detection of micro defects using image-based control charting techniques has not yet been investigated and the successful detection of these defects would lead to the development of a total quality monitoring tool that could be used to detect any visual defect in a LCD panel. This tool would be of great practical use if these error types could be linked with some actual root-cause on the production floor using some sort of an expert system. Thus, the quality monitoring cycle would be closed from the detection of a product fault to process recovery.

(6) Demerit systems could be used to evaluate the grade of a certain product. A Demerit-based control chart could be an introductory step for advancements in the life cycle analysis (see e.g. Finkbeiner et al. (2006)) of these products, where a Demerit score could be linked to a certain product's life expectancy.

3.5.3 Image-Based Charts for Manufacturing Applications: Electronic Component Manufacturing

Challenges and SPC Targets:

For the last two decades, the electronics industry has been active in applying automated machine vision systems to various manufacturing processes, which include, as reported by Zuech (1992); printed circuit boards (PCBs), integrated circuit (IC) manufacturing, photo-masks, etc. Moganti et al. (1996) provided a thorough explanation for the importance of machine vision systems to the electronics industry. According to their list, the following targets should be met by any MVS/SPC approach:

(1) 100 % inspection is an industry requirement to ensure the very high quality levels set by industry. Therefore, any image-based quality monitoring technique should be run in real-time corresponding to the production rate of the process being monitored.
(2) There is a lot of variability between one process and another. Therefore, it is often the case that the developed quality monitoring systems can only work for effectively for one process, which makes the development of a generalized quality monitoring approach extremely difficult. Efforts should focus more on customizing the quality monitoring system to a certain production function.
(3) Related to the previous performance target, both the false and missed alarms (Type I and Type II) errors should be minimized for any image-based quality monitoring system. Consequently, the ever-increasing costs of scrap could be minimized.

The literature presented in the following subsection would be depicted based on the three aforementioned performance targets. This would also serve as a tool in evaluating the practicality of the presented research.

Available Literature:

In this subsection, three image-based control charting techniques are presented; the first one is on monitoring the wafer (IC) manufacturing process and the other two are on the monitoring semiconductor manufacturing processes using a combined MVS/control charting approach.

Tong et al. (2005) used a MVS approach that utilizes the Hotelling T^2 control charts to monitor the wafer production process. As depicted in Jeong et al. (2008), a wafer is the building block of semiconductor manufacturing, where several hundred ICs are simultaneously fabricated on one wafer (Fenner et al. (2005)). After the IC fabrication process is completed, each chip is then classified as either functional or defective, where a visual display (wafer map) is then used to show the locations of the nonconforming IC chips on the wafer (Jeong et al. (2008)). Hansen et al. (1997) showed that defective chips often occur in clusters or display systematic patterns, which could be used to make some inferences about the manufacturing process conditions (Cunnigham and McKinnon (1998)). The defect clustering effects are important as they can reflect the causes of the defects and therefore, clustering indices (CIs) have been developed to accurately represent the clustering phenomenon as in Jun et al. (1999). One key aspect of CIs is that they are independent of the chip area and they don't require any distributional assumptions.

It might be intuitive to use the C chart for monitoring the defects in wafer processes; however, the defect count often does not obey a Poisson distribution and therefore, the corresponding traditional C-chart have high false alarm rates (Friedman and Albin (1991) and Hsieh et al. (2007)). Accordingly, a number of different approaches have been suggested to tackle this problem by suggesting a potential distribution for the defect count as in (Friedman and Albin (1991) and Hsieh et al. (2007)); however, none of these suggested a total quality monitoring framework that includes both the actual measurement of defects and the defect detection component.

The method proposed by Tong et al. (2005) has the potential to be applied in practice due to their utilization of the number of defects and the CI as the two quality characteristics, which are being monitored by the Hotelling T^2 control chart as shown in Figure 17. In the case of a signal in the T^2 control chart, the T^2 statistic is decomposed based on the method by Mason et al. (1997), to identify whether the cause of the signal whether it is too many defects, the clustering of the defects, or both as shown in Figure 18.

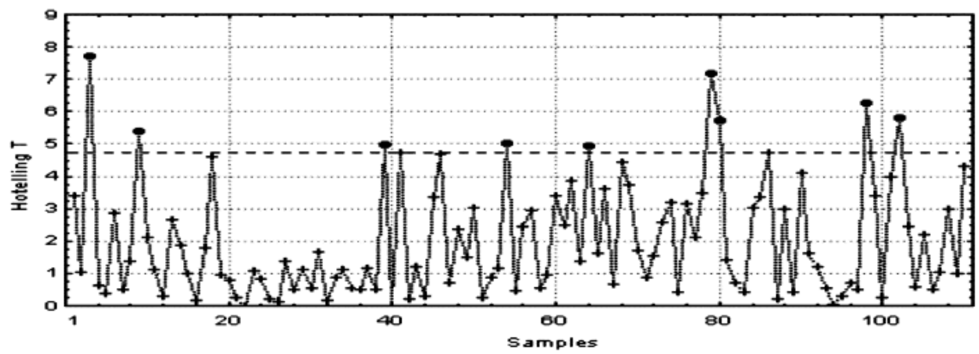


Figure 17 Hotelling T^2 control charts (Tong et al. (2005))

Wafer	T^2	T_1^2	$T_{2.1}^2$	T_2^2	$T_{1.2}^2$	Source of out of control
3	7.6997	**5.2329**	2.3642	0.2115	7.4882	Number of defects
9	5.3977	1.1333	**4.1557**	0.5557	**4.8420**	Interactions
39	4.9953	**4.1494**	0.7998	0.5684	4.4269	Number of defects
54	5.0393	**5.0218**	0.0125	2.1369	2.9025	Number of defects
64	4.9371	**3.5265**	1.3489	0.2031	4.7340	Number of defects
79	7.1585	**4.7309**	2.3291	0.1527	7.0058	Number of defects
80	5.7293	1.9179	**3.7015**	0.1952	**5.5342**	Interactions
98	6.2831	0.5096	5.7072	**4.9464**	1.3367	Defect clustering
102	5.7998	1.0584	**4.6240**	0.7245	**5.0753**	Interactions
Critical point	$UCL = \frac{m+1}{m} F_{(0.1,1,109)} = (111/110) \times 2.752 = 2.777$					

Figure 18 Explanation of the Cause of Signals in the Hotelling T^2 Control Chart (Tong et al. (2005))

Lin (2007) used wavelets and multivariate statistical approaches, including the Hotelling T^2 control charts, to detect ripple defects in electronic components (SBL chips of ceramic capacitor). The wavelet characteristics were used to describe the surface texture properties and then, the author utilized a Hotelling T^2 control chart to judge the existence of a defect based on integrating the different texture properties (quality characteristics). In a later paper, Lin et al. (2008) conducted a comparison between the capabilities of a wavelet-Hotelling T^2 control chart approach and a wavelet-PCA based approach in detecting surface defects in light-emitting diode (LED) chips. Their results showed that the wavelet-PCA based approach was more effective.

Future Research Ideas: Image-Based Control Chart Techniques for Electronic Components Manufacturing

Based on the discussion in the previous two subsections, it could be seen that more work is needed if these control charting techniques are to meet their potential. A summary of interesting and potentially useful research ideas is provided below. It should be noted that some of these research ideas were suggested in Woodall (1997) and are now found to be of relevance to electronic components manufacturing.

(1) Based on the paper by Moganti et al. (1996), it is surprising to find that none of the papers (discussed in the previous subsection) reported real-time analysis capabilities, which would allow for 100 % sampling. It is not clear whether this disconnect is actually a limitation or that the authors just did not state the obvious that the algorithms have to be running in real-time to allow for the 100% sampling requirement in industry.
(2) What is the effect of combining attribute data (number of defects) and the variables data (CI) on the performance of the Hotelling T^2 control chart? Would the utilization of a variable control chart and an attribute control chart separately lead to better performance?
(3) Most of the reported research has been in detecting step shifts in the underlying parameters. The question of how effective is these attribute control charts in detecting linear and other type of trends in the underlying parameter(s) seem of interest? It is well known that semiconductors are manufactured in highly automated facilities with near dust free conditions; however, defects still occur. The effectiveness of control charts to account for an increasing number of defects arising from unexpected increase in the dust content seem to be of relevance to the semiconductor industry.
(4) With the high production rates and increased product variability, some "legitimate batch-to-batch variability" (see Heimann (1996)) could be expected. Could these methods be adopted for the construction of image-based control charts?
(5) How can multivariate CUSUM and EWMA charts be used with the reported quality characteristics for the electronics industry?
(6) The design of any of the aforementioned control charts should be considered from an economic view as it seems that there is no justification for setting up the control chart limits so that a certain percentage of Type I and Type II errors are randomly selected.
(7) Can the relationship between the quality characteristics and the output be represented in terms of a profile? If so, how can profile monitoring techniques be utilized in detecting

instances of non-conforming and how to identify the reasons for non-conforming conditions?

3.5.4 Image-Based Charts for Manufacturing Applications: Other Industrial Applications

Challenges and SPC Targets:

The introduction of new technologies on the shop-floor has allowed industrial image-based MVS quality monitoring systems capable of tackling more complex problems. Malamas et al. (2003) noted that most of the problems tackled by these systems are related to one of the following quality characteristics: 1) dimensional quality; 2) surface quality; 3) structural (assembly) quality; and 4) operational quality. They have also showed that the degree of flexibility of these systems can be described by the following degrees of freedom (DoF): a) size; b) pose; c) shape; d) color; e) texture; and f) illumination. Therefore, it could be inferred that the end goal is to have an image-based quality monitoring system that is capable of not only detecting faults in more than one of the aforementioned quality characteristics, but also diagnosing the root-causes for these faults, and allowing for different variations in the manufacturing setting parameters (i.e. having more DoF). It should be noted that the literature presented below focuses on measuring either the dimensional quality, surface quality or both and there have not been any reference to structural quality and operational quality (except in Tan et al. (1996) and Liu and MacGregor (2006), where they suggested the utilization of feedback controllers).

Available Literature:

The measurement of the dimensional quality of manufactured components is one of the classical motivational examples for the construction of control charts. Therefore, it was surprising to see that there are only two image-based control chart methods for that purpose. In an exploratory paper, Tan et al. (1996) used a MVS to sample and measure the quality characteristics (length, width and area) of extruded food products as a part of total statistical process control system, where the X-bar control chart was used to monitor the state of statistical control. Corrective actions were then determined using an engineering process control (EPC) scheme, mainly a proportional and integral (PI) control scheme, to minimize the product size variations arising from material inconsistencies in terms of the moisture content. In another application, Lyu and Chen (2009) used digital image processing techniques to measure the diameters of concentric circles and then used one of three multivariate control charts, Chi-Square, Hotelling T^2 and MEWMA, to check for out of control conditions. It is not clear; however, their definition of out of control conditions as in their case study they have only used 35 standard concentric circles of diameters 2 and 3.5 cm.

On the other hand, Jiang et al. (1998) used a digital image processing system to inspect oil seals for oversize and Individuals-Moving Range (I-MR) charts to detect surface defects on the oil seals. Their reported measurement accuracy is really good (0.4 of a pixel); however, it is undermined by the poor performance of their control charts (a 17.31 false alarms %). Cheng (2003) tried to combine subjective experts rating with actual shape measurements obtained by machine vision to make decisions about product's shape quality based on fuzzy control charts. On the other hand, Armingol et al. (2003) used an image-based control chart (I-MR) technique to detect surface bumps and scratches in manufactured products; however, the quality characteristic plotted on the control chart is not clear.

Liu and MacGregor (2006) provided a more robust method for measuring the appearance and aesthetics of manufactured products. The strength of their approach is on the consistent/ quantitative estimation of the continuous variation in the visual appearance rather than trying to classify these variations into discrete classes. Therefore, SPC and EPC approaches could be used for surveillance and feedback control functions. The surveillance (monitoring) of the quality of the surface was done using both Hotelling T^2 control charts and the Squared Prediction Error (SPE), see Figure 19, as they are both known to complementary statistics (refer to Kresta and MacGregor (1991) and MacGregor and Kourti (1995)).

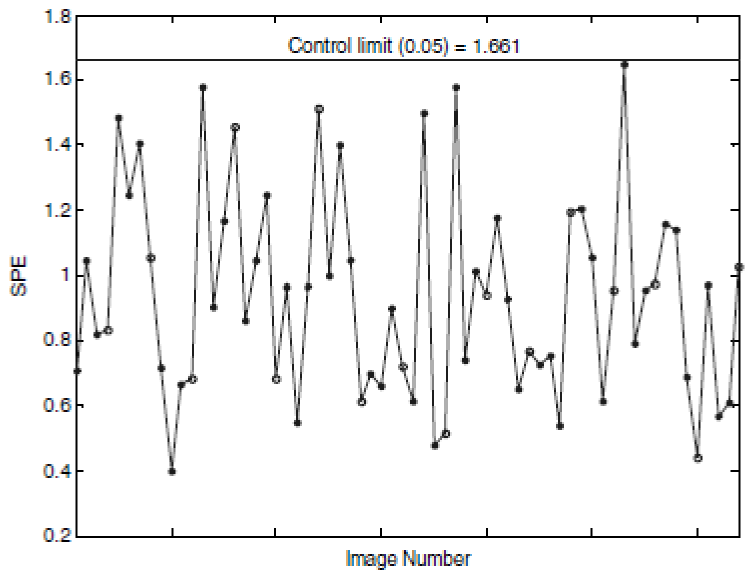


Figure 19 SPE Plot used in monitoring the aesthetics of Countertops (Liu and MacGregor (2006))

Colosimo et al. (2007) dealt with a more complex dataset in their attempt to measure the surface quality of metallic surfaces, where they extended the application of profile monitoring techniques to three dimensional surfaces. Their proposed approach combines a regression model having spatial correlated noise with univariate and multivariate control charts. They have successfully tested their approach on metallic surfaces produced by turning.

It should be noted that there are several papers that discuss the utilization of machine vision to detect trends on control charts (e.g. Pham and Wani (1997) and Hassan et al. (2003)); however, they are not going to be discussed further as they are outside the scope of this review paper due to the fact that the inputs for the control charts are not based on image-data.

Future Research Ideas: Image-Based Control Chart Techniques for Other Industrial Applications

Based on the literature described above, it is clear that more work is needed before quality monitoring systems based on MVS and image-based control charts could be widely adopted in industry. Very little work has focused on any statistical considerations in the design of control charts and with the exception of Liu and MacGregor (2006) and Colosimo et al. (2007), there is no thorough explanations for basic concepts such as the choice of a specific control charts. In addition to filling these obvious research gaps, a summary of research ideas and unanswered research questions is provided below.

(1) Based on Malamas et al. (2003), there are two major research ideas that should be explored. On the digital image processing side, how can an effective machine vision system be developed so that it integrates more dimensional, surface, and/or structural quality characteristics with their operational implications? This could be done through the development of some novel expert systems and neural networks that allow for the measurement and detection of these different features. On the quality control side, how can a control chart be developed for these different quality characteristics? More importantly, how can a signal be tracked down to a specific quality characteristic (noting that data is a combination of attribute and variables data in a huge sample size problem)?

(2) Related to the previous questions, how can these be constructed taking into account the effect of changes in the DoF on the performance of the machine vision system? In manufacturing settings, changes in the pose (part location) and the lighting conditions are inevitable so how can the machine vision system be robust against these changes? If not, how can the developed control charts account for these changes and therefore, minimizing false alarms without a significant increase in Type II errors?

(3) Building on points 1 and 2, how can trends be detected quickly for such systems? The work of MacGregor and Kourti (1995), Bharati and MacGregor (1998), Kourti (2005), Kourti (2006) and Pereira et al (2009) seem to be of relevance for this topic. In addition some economic analysis of the control charts should be considered in determining the control chart parameters and their consequent effect on Type I errors, Type II errors and ARLs.

(4) In the digital image processing literature, there has been some recent work describing the utilization of 3-D laser scanners in providing one-to-one comparisons between the as-built product and the original CAD model (see e.g. Varady et al. (1997), Griffin et al. (1998), Thompson et al. (1999) and Mohaghegh et al. (2007)). The question now is in

spite of the significance of this breakthrough, how can the SPC framework utilize this information in improving the manufacturing process and reducing product variability?

(5) Related to the previous research question, how can variable sampling plans be devised for two different technologies within the SPC framework? It is well known that the current capabilities of 3-D Laser scanners does not allow for real-time analysis for products. On the other hand, traditional machine vision systems have a real-time ability for detecting faults. The question now is how to integrate a high precision/accuracy and low frequency measurement device with one with lower degrees of precision and accuracy, but significantly higher frequency in one SPC framework/control chart? Should this question only be framed as an extension to variable sampling intervals (VSI) control charts? (e.g. of VSI and VSI control charts see Cui and Reynolds (1988), Reynolds et al. (1988), Chengalur et al. (1989), Sethuraman et al. (1994) and Reynolds and Stoumbos (2001)).

3.6 Image-Based Charts for Medical Applications

Challenges and SPC Targets:

Radiologists and clinicians routinely use a number of imaging techniques in their daily diagnostic decisions on a huge scale. Accordingly, a vital aspect of the diagnosis is the correct and fast interpretation of the collected data referring to the physician's knowledge of "typical, healthy and pathological anatomy and physiology of examined organs and structures, completed by experience and cognitive intuition" (Przelaskowski (2008)). In addition, the process of radiological interpretation, the understanding of medical image contents, involves a number of sequential activities (see Przelaskowski (2008)), which starts by image-based detection of disease and ends with following the response to therapy—another task that is heavily dependent on medical imaging. Therefore, from a quality monitoring perspective, it is indispensable to ensure that these image capturing devices are functioning properly. On the other hand, future challenges are identified in (2000).

Available Literature:

Simmons et al. (1999) developed a quality assurance protocol that can be used in evaluating the fMRI (functional Magnetic Resonance Imaging) system performance. In addition, it allows for a fast system recovery through taking corrective actions whenever necessary. The evaluation of the fMRI system performance is achieved through monitoring the quality characteristics of the fMRI, mainly signal to noise ratio (SNR) and signal to ghost ratio (SGR), through two separate Xbar charts with run rules. The control charts allow for the detection of trends reflecting deterioration in the performance of the fMRI machine. To the best of our knowledge, the research done by Simmons et al. (1999) represents the only study where image-based control charts, where explicitly used in the medical field.

Future Research Ideas: Image-Based Control Chart Techniques for Medical Applications

It is definitely obvious that based on one study, not a lot could be inferred on what needs to be done in the field. However, some recommendations could be made based on the current research in the field of medical image processing and by showing the gaps that exists in the Simmons el al. (1999) study. Recommendations based on these two criteria are described below.

(1) Based on the Simmons et al. (1999) study, the statistical reasoning behind their control charts is questionable to say the least. There are a number of reasons for that evaluation: a) it is well known within the SPC community that the utilization of a multivariate control chart causes fewer false alarms than a set of separate control charts; b) the statistical basis for the choice of the run rules is not clear and therefore, the expected type I and type II error rates are unknown to the developers of the chart; c) based on the two previous comments, a high false alarm rate is expected, which is a note that they have reported; d) the independent variable that they use in the control charts is a sample per day, which seems contradictory to the requirement of the quick detection of trends; and e) a phase II multivariate control chart such as the multivariate EWMA should be investigated. These remarks reflect important gaps that need to be tackled in the Simmons et al. (1999) study.

(2) In the Przelaskowski (2008) study, it was shown that the current trend in medical images is leading towards an increased role of machine vision in supporting the physicians' decisions in the diagnosis. This situation is equivalent to the proposed research area of fast detection of trends using image-based quality monitoring systems in industrial applications. This analogy is also supported by the increasing trend of maintain electronic medical records, which provides an opportunity for multivariate image analysis using SPC. Therefore, the work of MacGregor and Kourti (1995), Bharati and MacGregor (1998), Kourti (2005), Kourti (2006) and Pereira et al (2009) could be benchmarks for the development of such capabilities.

3.7 Conclusions and Recommendations:

We view image-based control charts are as promising area of research within statistical process control. It expands the current body of SPC knowledge to include the capability of monitoring dimensional data, product geometry, surface defects and surface finish concurrently in real-time. The incorporation of these quality characteristics, especially non-dimensional parameters, to the SPC framework could be really beneficial as they represent often neglected quality characteristics, which could not be quantitatively measured. The quantitative measurement of these parameters would lead to better process control through the monitoring of a more representative set of variation sources.

In addition to recommending ideas for future research in each of the image-based control charts application, we believe that these three research questions and ideas are pivotal for future developments in image-based control charts:

(1) The modeling of image-based control charts as a spatial-temporal problem seems natural and therefore, the question of how image-data could be manipulated to allow for this representation seems of great value.

(2) Woodall et al. (2004) discussed the importance of addressing the issue of common cause profile-to-profile variation. We believe that image-based control charts could also be represented as an extension to profile monitoring as reflected in Woodall (2007). Therefore, addressing the effect of part location change, lighting condition change and measurement error due to image segmentation is important by similarity.

(3) What is the effect of integrating attributes and variables data in one control chart on Type I and Type II errors? It seems that for most of the image-based control charts this would be the end goal so studying the effect of integrating these two sets of data in a multivariate control chart's performance is a key issue, which has been overlooked in the research studies.

3.8 References:

(1973). *Experience with a Demerit System for Cloth Quality Grading* (Manchester: Shirley Institute).

(2000). "Future needs for medical imaging in HC: Medical Imaging Technology Roadmap", <http://www.ic.gc.ca/eic/site/mitr-crtim.nsf/eng/h_hm00038.html>.

Abdel-Qader, I., Pashaie-Rad, S., Abudayyeh, O., and Yehia, S. (2006). "PCA-Based algorithm for unsupervised bridge crack detection", *Advances in Engineering Software,* 37 (12), 771-778.

Alwan, L. C. and Roberts, H. V. (1988). "Time-Series Modeling for Statistical Process-Control", *Journal of Business & Economic Statistics,* 6 (1), 87-95.

Anagnostopoulos, C., Vergados, D., Kayafas, E., Loumos, V., and Stassinopoulos, G. (2001). "A computer vision approach for textile quality control", *The Journal of Visualization and Computer Animation,* 12 (1), 31-44.

Bharati, M. H. and MacGregor, J. F. (1998). "Multivariate Image Analysis for Real-Time Process Monitoring and Control", *Industrial & Engineering Chemistry Research,* 37 (12), 4715-4724.

Bharati, M. H. and MacGregor, J. F. (2001). "Texture analysis of images using Principal Component Analysis", *Process Imaging for Automatic Control,* 4188, 27-37.

Breiman, L. (1984). *Classification and regression trees*, ed. Leo Breiman (x, 358 p. :; Belmont, Calif. :: Wadsworth International Group).

Campbell, J. G., Fraley, C., Stanford, D., Murtagh, F., and Raftery, A. E. (1999). "Model-based methods for textile fault detection", *International Journal of Imaging Systems and Technology,* 10 (4), 339-346.

Campos, L. F. A., Silva, A. C., and Barros, A. K. (2005). "Diagnosis of breast cancer in digital mammograms using independent component analysis and neural networks", *Progress in Pattern Recognition, Image Analysis and Applications, Proceedings,* 3773, 460-469.

Carstensen, J. (1992). "Description and Simulation of Visual Texture", (Technical University of Denmark).

Cenci, M., Nagar, C., and Vecchione, A. (2000). "PAPNET-assisted primary screening of conventional cervical smears", *Anticancer Research,* 20 (5C), 3887-3889.

Chang, C. Y., Li, C. H., Chang, J. W., and Jeng, M. D. (2009). "An unsupervised neural network approach for automatic semiconductor wafer defect inspection", *Expert Systems with Applications,* 36 (1), 950-958.

Chen, L. F., Su, C. T., and Chen, M. H. (2009). "A Neural-Network Approach for Defect Recognition in TFT-LCD Photolithography Process", *Ieee Transactions on Electronics Packaging Manufacturing,* 32 (1), 1-8.

Cheng, C-B (2003). "Fuzzy process control based on fuzzy regression and possibility measures", *Fuzzy Information Processing Society, 2003. NAFIPS 2003. 22nd International Conference of the North American*, 127-131.

Chengalur, I. N., Arnold, J. C., and Reynolds, M. R. (1989). "Variable Sampling Intervals for Multiparameter Shewhart Charts", *Communications in Statistics-Theory and Methods,* 18 (5), 1769-1792.

Chiang, L.H. (2001). *Fault detection and diagnosis in industrial systems*, eds R. D. Braatz and E. Russell (xiv, 279 p. :; London ; New York :: Springer).

Colosimo, B. M., Mammarella, F., and Petro, S. (2007). "Quality Control of Manufactured Surfaces", *International Workshop on Intelligent Statistical Quality Control* (Tsinghua University, Beijing).

Couweleers, F., Skoitheim, O., Schulerud, H., and Kaspersen, K. (2003). "In-line geometric fault detection in car parts based on structured light projection and image processing", *Optical Measurement Systems for Industrial Inspection Iii,* 5144, 697-706.

Cui, R. Q. and Reynolds, M. R. (1988). "Xbar-Charts with Runs Rules and Variable Sampling Intervals", *Communications in Statistics-Simulation and Computation,* 17 (3), 1073-1093.

Cunningham, S. P. and MacKinnon, S. (1998). "Statistical methods for visual defect metrology", *Semiconductor Manufacturing, IEEE Transactions on,* 11 (1), 48-53.

Deming, W. E. (1993). *The new economics for industry, government, education* (Cambridge, MA: Massachusetts Institute of Technology, Center for Advanced Engineering Study) xii, 240 p.

Deng, S., Liu, T., Xiao, Z., and Li, X. (2008). "An automatic detection system for flatness of integrated circuit pins", in Y. Sheng, Y. Wang, and L. Zeng (eds.), (1 edn., 7156; Beijing, China: SPIE).

Feng, Q. and Kapur, K. C. (2008). "Quality Engineering: Control, Design and Optimization", 171-186.

Fenner, J. S., Jeong, M. K., and Lu, J-C. (2005). "Optimal automatic control of multistage production processes", *Semiconductor Manufacturing, IEEE Transactions on,* 18 (1), 94-103.

Finkbeiner, M., Inaba, A., Tan, R. B. H., Christiansen, K., and Kluppel, H. J. (2006). "The new international standards for life cycle assessment: ISO 14040 and ISO 14044", *International Journal of Life Cycle Assessment,* 11 (2), 80-85.

Friedman, D. J. and Albin, S. L. (1991). "Clustered defects in IC fabrication: Impact on process control charts", *Semiconductor Manufacturing, IEEE Transactions on,* 4 (1), 36-42.

Gonzalez, R. C., Woods, R. E., and Eddins, S. L. (2004). *Digital Image processing using MATLAB* (xiv, 609 p. :; Upper Saddle River, N.J. :: Pearson Prentice Hall).

Gonzalez, R. C. and Woods, R. E. (2007). *Digital image processing* (3rd ed. edn., xxii, 954 p. :; Harlow :: Prentice Hall).

Griffin, A., McMillin, S., and Knox, C. (1998). "Practical examples, using scan data for reverse engineering", *Technical paper PE98-122* (Dearborn, MI: Society of Manufacturing Engineers).

Gross, R. (2005). "Face Databases", in S. Li and A. Jain (eds.), *Handbook of Face Recognition* (New York: Springer), 22.

Gurcan, M. N., Sahiner, B., Petrick, N., Chan, H. P., Kazerooni, E. A., Cascade, P. N., and Hadjiiski, L. (2002). "Lung nodule detection on thoracic computed tomography images: Preliminary evaluation of a computer-aided diagnosis system", *Medical Physics,* 29 (11), 2552-2558.

Hansen, M. H., Nair, V. N., and Friedman, D. J. (1997). "Monitoring wafer map data from integrated circuit fabrication processes for spatially clustered defects", *Technometrics,* 39 (3), 241-253.

Haralick, R. M., Shanmugam, K., and Dinstein, I. (1973). "Textural Features for Image Classification", *Systems, Man and Cybernetics, IEEE Transactions on,* 3 (6), 610-621.

Hassan, A., Baksh, M. S. N., Shaharoun, A. M., and Jamaluddin, H. (2003). "Improved SPC chart pattern recognition using statistical features", *International Journal of Production Research,* 41 (7), 1587-1603.

Heimann, P. A. (1996). "Attributes control charts with large sample sizes", *Journal of Quality Technology,* 28 (4), 451-459.

Hjelmas, E. and Low, B. K. (2001). "Face detection: A survey", *Computer Vision and Image Understanding,* 83 (3), 236-274.

Horst, R. L. and Negin, M. (1992). "Vision system for high-resolution dimensional measurements and on-line SPC: web process application", *Industry Applications, IEEE Transactions on,* 28 (4), 993-997.

Huang, S. F., Chen, Y. C., and Moon, W. K. (2008). "Neural network analysis applied to tumor segmentation on 3D breast ultrasound images", *2008 Ieee International Symposium on Biomedical Imaging: From Nano to Macro, Vols 1-4*, 1303-1306.

Jähne, B. (2005). *Digital image processing* (6th rev. and ext. ed. edn., xiii, 607 p. :; Berlin ; New York :: Springer).

Jan, J. (2006). *Medical image processing, reconstruction and restoration concepts and methods* (xxiii, 710 p. :; Boca Raton, FL :: Taylor & Francis).

Jeong, Y-S., Kim, S-J., and Jeong, M. K. (2008). "Automatic Identification of Defect Patterns in Semiconductor Wafer Maps Using Spatial Correlogram and Dynamic Time Warping", *Semiconductor Manufacturing, IEEE Transactions on,* 21 (4), 625-637.

Jiang, B. C. and Jiang, S. J. (1998). "Machine vision based inspection of oil seals", *Journal of Manufacturing Systems,* 17 (3), 159-166.

Jiang, B. C., 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.

Jiang, B.C., Wang, C.C. , Tsai, D. M., and Lu, C.J. (2004). "LCD Surface Defect Inspection using Machine Vision", *Fifth Asia Pacific Industrial Engineering and Management Systems Conference* (Australia).

Jun, C-H., Hong, Y., Kim, S. Y., Park, K-S., and Park, H. (1999). "A simulation-based semiconductor chip yield model incorporating a new defect cluster index", *Microelectronics Reliability,* 39 (4), 451-456.

Kabir, S., Rivard, P., and Ballivy, G. (2008). "Neural-network-based damage classification of bridge infrastructure using texture analysis", *Canadian Journal of Civil Engineering,* 35 (3), 258-267.

Kim, K., Mahmoud, M. A., and Woodall, W. H. (2003). "On the monitoring of linear profiles", *Journal of Quality Technology,* 35 (3), 317-328.

Kourti, T. (2005). "Application of latent variable methods to process control and multivariate statistical process control in industry", *International Journal of Adaptive Control and Signal Processing,* 19 (4), 213-246.

Kourti, T. (2006). "The Process Analytical Technology initiative and multivariate process analysis, monitoring and control", *Analytical and Bioanalytical Chemistry,* 384 (5), 1043-1048.

Kresta, J. V., Macgregor, J. F., and Marlin, T. E. (1991). "Multivariate Statistical Monitoring of Process Operating Performance", *Canadian Journal of Chemical Engineering,* 69 (1), 35-47.

Lin, C-S., Wu, W-Z., Lay, Y-L., and Chang, M-W. (2001). "A digital image-based measurement system for a LCD backlight module", *Optics & Laser Technology,* 33 (7), 499-505.

Lin, H-D. and Chiu, S. (2006). "Computer-Aided Vision System for MURA-Type Defect Inspection in Liquid Crystal Displays", in Chang, L. W. and Lie, W. N. (ed.), *Advances in Image and Video* Technology, 442-452. Berlin: Springer.

Liu, J. and MacGregor, John (2006). "Estimation and monitoring of product aesthetics: application to manufacturing of “engineered stone” countertops", *Machine Vision and Applications,* 16 (6), 374-383.

Lowry, C. A., Woodall, W. H., Champ, C. W., and Rigdon, S. E. (1992). "A Multivariate Exponentially Weighted Moving Average Control Chart", *Technometrics,* 34 (1), 46-53.

Lu, C-J. and Tsai, D-M. (2005). "Automatic defect inspection for LCDs using singular value decomposition", *The International Journal of Advanced Manufacturing Technology,* 25 (1), 53-61.

Lyu, J. and Chen, M. (2009). "Automated visual inspection expert system for multivariate statistical process control chart", *Expert Systems with Applications,* 36 (3, Part 1), 5113-5118.

MacGregor, J. F. and Kourti, T. (1995). "Statistical process control of multivariate processes", *Control Engineering Practice,* 3 (3), 403-414.

Malamas, E. N., Petrakis, E. G., Zervakis, M., Petit, L., and Legat, J-D (2003). "A survey on industrial vision systems, applications and tools", *Image and Vision Computing,* 21 (2), 171-188.

Mason, R. L., Tracy, N. D. , and Young, J. C. (1997). "A practical approach for interpreting multivariate T2 control chart signals", *Journal of Quality Technology,* 29 (4), 396.

Moganti, M., Ercal, F., Dagli, C. H., and Tsunekawa, S. (1996). "Automatic PCB Inspection Algorithms: A Survey", *Computer Vision and Image Understanding,* 63 (2), 287-313.

Mohaghegh, K., Sadeghi, M. H., and Abdulla, A. (2007). "Reverse engineering of turbine blades based on design intent", *International Journal of Advanced Manufacturing Technology,* 32 (9-10), 1009-1020.

Montgomery, D. C. (2009). *Introduction to statistical quality control* (6th edn.; Hoboken, N.J.: Wiley) xiv, 734 p.

Nakashima, K. (1994). "Hybrid Inspection System for Lcd Color Filter Panels", *10th Anniversary, Imtc/94 - Advanced Technologies in I & M, Conference Proceedings, Vols 1-3 and Supplement*, 689-692.

Nalwa, V. S. (1993). *A guided tour of computer vision* (xii, 361 p. :; Reading, Mass. :: Addison-Wesley).

Ngan, H. Y. T. and Pang, G. K. H. (2009). "Regularity Analysis for Patterned Texture Inspection", *Ieee Transactions on Automation Science and Engineering,* 6 (1), 131-144.

Nikolaidis, N. (2001). *3-D image processing algorithms*, ed. I. Pitas (x, 176 p. :; New York :: John Wiley).

Pereira, A. C., Reis, M. S., and Saraiva, P. M. (2009). "Quality Control of Food Products using Image Analysis and Multivariate Statistical Tools", *Industrial & Engineering Chemistry Research,* 48 (2), 988-998.

Pham, D. T. and Wani, M. A. (1997). "Feature-based control chart pattern recognition", *International Journal of Production Research,* 35 (7), 1875-1890.

Pratt, W. K. (2007). *Digital image processing : PIKS Scientific inside* (4th ed., Newly updated and rev. ed. edn., xix, 782 p., [4] p. of plates :; Hoboken, N.J. :: Wiley-Interscience).

Przelaskowski, A. (2008). "Computer-Aided Diagnosis: From Image Understanding to Integrated Assistance", in Pietka, E. and Kawa, J. (ed.), *Information Technologies in Biomedicine*, 44-54. Berlin : Springer.

RafajŁOwicz, E., Wnuk, M., and RafajŁOwicz, W. (2008). "Local Detection of Defects From Image Sequences", *International Journal of Applied Mathematics & Computer Science,* 18 (4), 581-592.

RafajŁOwicz, E. and Steland, A. (2009). "A binary control chart to detect small jumps", *Statistics,* 43, 295-311.

Reynolds, M. R., Amin, R. W., Arnold, J. C., and Nachlas, J. A. (1988). "Xbar Charts with Variable Sampling Intervals", *Technometrics,* 30 (2), 181-192.

Reynolds, M. R. and Stoumbos, Z. G. (2001). "Monitoring the process mean and variance using individual observations and variable sampling intervals", *Journal of Quality Technology,* 33 (2), 181-205.

Saitoh, F. (1999). "Boundary extraction of brightness unevenness on LCD display using genetic algorithm based on perceptive grouping factors", *Image Processing, 1999. ICIP 99. Proceedings. 1999 International Conference on* (2), 308-312 vol.302.

Selek, M., Sahin, O. S., and Kahramanli, S. (2009). "Using artificial neural networks for real-time observation of the endurance state of a steel specimen under loading", *Expert Systems with Applications,* 36 (4), 7400-7408.

Sethuraman, S., Reynolds, M. R., and Arnold, J. C. (1994). "Process control strategies using variable sampling intervals", *American Statistical Association 1994 Proceedings of the Section on Quality and Productivity*, 109-114.

Simmons, A., Moore, E., and Williams, S.C.R (1999). "Quality control for functional magnetic resonance imaging using automated data analysis and Shewhart charting", *Magnetic Resonance in Medicine,* 41 (6), 1274-1278.

Sokolov, S. M. and Treskunov, A. S. (1992). "Automatic vision system for final test of liquid crystal displays", *Robotics and Automation, 1992. Proceedings., 1992 IEEE International Conference on*, 1578-1582 vol.1572.

Tan, J., Chang, Z., and Hsieh, F. (1996). "Implementation of an Automated Real-Time Statistical Process Controller", *Journal of Food Process Engineering,* 19 (1), 49-61.

Taniguchi, K., Ueta, K., and Tatsumi, S. (2006). "A mura detection method", *Pattern Recognition,* 39 (6), 1044-1052.

Thompson, W. B., Owen, J. C., St Germain, H. J. D., Stark, S. R., and Henderson, T. C. (1999). "Feature-based reverse engineering of mechanical parts", *Ieee Transactions on Robotics and Automation,* 15 (1), 57-66.

Tong, L-I., Wang, C-H., and Huang, C-L. (2005). "Monitoring defects in IC fabrication using a Hotelling T2 control chart", *Semiconductor Manufacturing, IEEE Transactions on,* 18 (1), 140-147.

Toriwaki, J. (2009). *Fundamentals of three-dimensional digital image processing*, ed. Hiro Yoshida (xv, 269 p. :; Dordrecht ; London ; New York :: Springer).

Tunak, M. and Linka, A. (2008). "Directional Defects in Fabrics", *Research Journal of Textile and Apparel,* 12 (2), 10.

Varady, T., Martin, R. R., and Cox, J. (1997). "Reverse engineering of geometric models - An introduction", *Computer-Aided Design,* 29 (4), 255-268.

Wang, K. B. and Tsung, F. (2005). "Using profile monitoring techniques for a data-rich environment with huge sample size", *Quality and Reliability Engineering International,* 21 (7), 677-688.

Wheeler, D. J. and Chambers, D. S. (1992). *Understanding statistical process control* (2nd edn.; Knoxville, Tenn.: SPC Press) xx, 406 p.

Woodall, W H. and Adams, B. M. (1998). "Statistical Process Control", in Harrison M. Wadsworth (ed.), *Handbook of statistical methods for engineers and scientists* (2nd edn.; New York: McGraw-Hill), 1 v. in various pagings.

Woodall, W. H. (1997). "Control charts based on attribute data: Bibliography and review", *Journal of Quality Technology,* 29 (2), 172-183.

Woodall, W. H., Spitzner, D. J., Montgomery, D. C., and Gupta, S. (2004). "Using control charts to monitor process and product quality profiles", *Journal of Quality Technology,* 36 (3), 309-320.

Woodall, W. H. (2006). "The use of control charts in health-care and public-health surveillance", *Journal of Quality Technology,* 38 (2), 89-104.

Woodall, W. H. (2007). "Current research on profile monitoring", *Produção,* 17, 420-425.

Yamazaki, T., Kawakami, H., and Hori, H. (1996). *Color TFT liquid crystal displays* (Mountain View, CA: Semiconductor Equipment and Materials International).

Yu, H. L., MacGregor, J. F., Haarsma, G., and Bourg, W. (2003). "Digital imaging for online monitoring and control of industrial snack food processes", *Industrial & Engineering Chemistry Research,* 42 (13), 3036-3044.

Ziou, D. and Tabbone, S. (1998). "Edge Detection Techniques-An Overview", *Pattern Recognition & Image Analysis,* 8 (4), 537-559.

Zuech, N. (1992). "Introductory thoughts on machine vision/ AOI applications in the electronic industry", *NEPCON* (2), 443-444.

Zuech, N. (2000). *Understanding and applying machine vision*. New York : Marcel Dekker.

4 Conclusions and Future Work

4.1 Thesis Contributions:

The main contributions of this thesis affect two fields: a) industrial image processing; and b) Statistical Process Control as shown below:

(1) For the industrial image processing, image-based face recognition tools (FLD based methods) were utilized—for the first time in detecting faults in manufactured products—rather than trying to measure directly some particular features in the image such as valleys and contours

(2) The use of image-based face recognition tools, allows for detecting unknown faults and therefore, overcoming the limitations found in some of the current industrial image-processing techniques

(3) The proposed industrial image processing technique allows for the ability to utilize images that are currently used in the shop floor for other purposes, such as part location in robotic systems, without the need for any modifications on the shop floor;

(4) Finally, unlike many image-based fault detection tools, the proposed technique stresses on the importance of computational efficiency to ensure that the algorithm runs in real time in high volume manufacturing environments.

(5) On the SPC front, this thesis represents one of the pioneering work that highlights the benefits from incorporating machine vision systems into the current SPC framework.

(6) In addition, it sets the groundwork needed for more research in image-based control charts through a detailed review and analysis of the current literature, which would be published in the Journal of Quality Technology.

4.2 Future Work

In addition to the future work highlighted in chapter 3, the huge momentum generated from the intensive research in the field of industrial image processing should focus more on reflecting the needs of the industry. One example is to use expert systems coupled with image processing tools not only to detect faults, but also diagnose the root-causes of these faults. Another research idea would utilize the high dimensional data to gain insights on the effect of these faults on the manufacturing process i.e. if these faults would result in incineration, selling the final product as a lower grade or some minor adjustment on the shop floor for selling it as a high grade-product. In that essence, digital image processing would be closing the process monitoring cycle (fault detection → process recovery) and provide the production personnel with additional data on the magnitude of these faults.