The Pennsylvania State University
The Graduate School
College of Engineering

# DYNAMIC NETWORK MODELING AND ANALYSIS OF LARGE-SCALE INTERNET OF THINGS WITH MANUFACTURING AND HEALTHCARE APPLICATIONS

A Dissertation in
Industrial and Manufacturing Engineering
by
Chen Kan

© 2018 Chen Kan

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

May 2018

The dissertation of Chen Kan was reviewed and approved* by the following:

Hui Yang

Harold and Inge Marcus Career Associate Professor of Industrial and Manufacturing Engineering

Dissertation Co-Advisor, Co-chair of Committee

Sounder Kumara

Allen E. Pearce and Allen M. Pearce Professor of Industrial and Manufacturing Engineering

Dissertation Co-Advisor, Co-chair of Committee

Jingjing Li

William and Wendy Korb Early Career Professor (Associate Professor) of Industrial and Manufacturing Engineering

Lin Lin

Assistant Professor of Statistics

Janis Terpenny

Peter and Angela Dal Pezzo Department Head and Professor of Industrial and Manufacturing Engineering

*Signatures are on file in the Graduate School

# Abstract

With the advancement of sensing and information technology, sensor networks and imaging devices have been increasingly used in manufacturing and healthcare to improve information visibility and enhance operational quality and integrity. Furthermore, the Internet of Things (IoT) technology has been integrated for the continuous monitoring of machine and patient conditions. As such, large amounts of data are generated with high dimensionality and complex structures. This provides an unprecedented opportunity to realize smart automated systems such as smart manufacturing and connected health care. However, it also raises new challenges in data analysis and decision making. Realizing the full potential of the data-rich environment calls for the development of new methodologies for data-driven information processing, modeling, and optimization.

This dissertation develops new methods and tools that enable 1) better handling of large amounts of multi-channel signals and imaging data generated from advanced sensing systems in manufacturing and healthcare settings, 2) effective extraction of information pertinent to system dynamics from the complex data, and 3) efficient use of acquired knowledge for performance optimization and system improvement. The accomplishments include:

1) *Fusion and analysis of multi-channel signals*. In Chapter 2, a spatiotemporal warping approach was developed to characterize the dissimilarity among 3-lead functional recordings. A network was then optimally constructed based on the dissimilarity and network features were extracted for the identification of different types of diseases.

2) *Statistical process control based on time-varying images*. In Chapter 3, a stream of time-varying images were represented as a dynamic network. Then, community structure of the network was characterized and community statistics were extracted from time to time. Finally, a new control charting approach was developed for in-situ monitoring of manufacturing processes.

3) *Monitoring and Control of large-scale Industrial Internet of Things*. In Chapter 4, a stochastic learning approach was developed to construct a large-scale dynamic network from IIoT machines. A parallel computing scheme was further proposed to harness

multiple computation resources to increase the efficiency. The constructed large-scale network can be used to visually and analytically monitor machine condition, inspect product quality, and optimize manufacturing processes.

4) *Modeling and Analysis of large-scale Internet of Health Things*. In Chapter 5, the large-scale network model was extended to the Internet of Health Things. A new IoT technology specific to the heart disease, namely, the Internet of Hearts, was proposed. Dynamic network modeling and parallel computing schemes were designed and developed for multi-level cardiac monitoring: patient-to-patient variations in the population level and beat-to-beat dynamics in the individual patient level. Control charting schemes were further proposed to harness network features for change detection in the cardiac processes.

# Contents

# List of Figures

# List of Tables

# Acknowledgements

I would like to express my deepest gratitude to Dr. Hui Yang, my advisor, for his continuous support, guidance, and care throughout my Ph.D. study. I will always remember his instructions and help on how to design experiments, how to write papers, and how to give presentations. It is Dr. Hui Yang who let me make up my mind to pursue an academic career. Also, I owe my sincere gratitude to my co-advisor, Dr. Soundar Kumara, for this insightful guidance and encouragement. Dr. Kumara's rich knowledge, rigorous thinking, and professionalism have an immense impact on my current and future research. Both Dr. Yang and Dr. Kumara are my role models. Through their guidance and encouragement, I have grown both personally and professionally.

I would like to thank my other committee members, Dr. Jingjing Li and Dr. Lynn Lin for their support on my research. No matter how busy they are, they always find time to meet and give me valuable suggestions. They provide knowledge from different domains and their valuable comments help me a lot to improve the quality of my dissertation. My deep gratitude also goes to professors and scholars who helped me and/or worked collaboratively with me in the past few years, including Dr. Fabio Leonelli, Dr. Changqing Cheng, Dr. Kay-Pong Yip, Dr. Prohalada Rao, Dr. Tapas Das, Dr. Satish Bukkapatnam, Dr. Robert Voigt, and Dr. Janis Terpenny. Special thanks to Dr. Jose Zayas-Castro for his encouragement and support during my Ph.D. study.

I greatly appreciate the opportunity to work in Dr. Hui Yang's Complex Systems Monitoring, Modeling and Analysis Laboratory. I would like to thank my fellow students and colleagues in the lab, including but not limited to Dr. Dongping Du, Dr. Gang Liu, Dr. Yun Chen, Ms. Bing Yao, Mr. Cheng-Bang Chen, Mr. Farhad Imani, Ms. Shengli Pei, and Ms. Rui Zhu. I would also like to thank administrative staffs in the IME department Ms. Lisa Fuoss, Mr. James Wyland, Ms. Laurette Gonet for their generous help. In addition, I gratefully acknowledge the financial support from the National Science Foundation (CMMI-1454012, CMMI-1266331, CMMI-1619648, IIP-1447289, and IOS-1146882) and other funding resources.

Most importantly, I wish to thank my beloved parents for their endless love and unconditional support. Their encouragement and trust always help me get through difficulties in my life.

# Chapter 1: Introduction

## 1.1. Motivation

The modern industries have increasingly invested in advanced sensing technology to cope with the ever-increasing complexity of systems, improve the visibility of information, and enhance operational quality and integrity. 1) Instead of single-channel sensing, distributed sensor networks have been more and more used in manufacturing and healthcare to facilitate the all-round collection of space-time information. For example, 12-lead electrocardiogram (ECG) system deploys 10 electrodes to provide multi-directional views of cardiac electrical activity. Multiple sensors (e.g., cutting force, vibration, and acoustic emission) are installed for *in situ* quality assurance of ultra-precision machining (UPM). 2) Furthermore, imaging technologies are increasingly adopted for online and high-resolution monitoring of process dynamics. High dimensional functional images are collected to assess the quality of manufactured products or study the pathology of diseases. For example, microscopic images have been used to inspect the surface finish of UPM samples. Cardiac optical mapping captures emitted fluorescent lights from the heart and produces time-varying 2-D and 3-D images for electrophysiological analysis. 3) Recently, the Internet of Things (IoT) has been hailed as a revolution in automation science and information technology. The IoT deploys a multitude of sensors, computing units and physical objects in an Internet-like infrastructure. The cyber-physical system of IoT enables continuous condition monitoring and timely anomaly detection in both manufacturing and healthcare settings.

As a result, advanced sensing brings rich data, which provides an unprecedented opportunity to realize smart automated systems such as smart manufacturing and smart health care. However, advanced sensing produces massive data with complex structures, which poses ***significant challenges***:

1) Distributed sensor networks capture spatiotemporal signals, which show high-level of nonlinear and nonstationary patterns with the presence of extraneous noises. This hampers the performance of traditional approaches with linearity and stationarity assumptions.

2) Imaging technology generates 2-D, 3-D, and higher-dimensional functional images. Most of traditional statistical process control (SPC) approaches are designed for the monitoring of 1-D

profiles and they are limited in their ability to handle high-dimensional images for process monitoring and fault diagnosis.

3) IoT connects large amounts of machines in digital manufacturing, as well as human subjects in smart and connected health. This gives rise to large amounts of data with complex and multifarious structures. Traditional approaches are limited in their ability to effectively extract pertinent knowledge about system dynamics in the large-scale IoT contexts.

Therefore, realizing the full potential of the data-rich environment requires fundamentally new methodologies for data-driven information processing, modeling, and optimization. There is a pressing need to develop advanced methodologies and associated tools that will enable and assist in 1) the handling of massive, complex data generated from advanced sensing systems in manufacturing and healthcare settings; 2) the extraction of pertinent information about system dynamics; and 3) the exploitation of acquired knowledge for decision making and performance optimization.

## 1.2. State-of-the-Art

With sensor networks and imaging technology, high-dimensional data (such as multi-channel signals and 2-D images) are increasingly available in manufacturing and healthcare. In the past decade, various methods have been developed in the research community of Industrial Engineering to extract useful information from the data for process monitoring and anomaly detection. The typical idea is to reduce the high-dimensional data as representative features in the low-dimensional space. To achieve this, one way is to represent the data by different models, and further extract model parameters as features. See, for example, linear models of Kang and Albin [1] , Kim *et al.* [2], and Zou *et al.* [3], and mixed-effect models of Apley [4] and Zhou *et al*. [5]. Also, some works apply dimensionality reduction techniques. For example, the principal component analysis used by Paynabar *et al*. [6] and independent component analysis by Ding *et al*. [7]. Furthermore, some papers transform the original data into other domains for feature extraction. For example, wavelet decomposition by Yang [8] and Zhou *et al.* [9], and fractal representation by Ruschin-Rimini *et al*. [10]. Methods such as variable selection (e.g., Liu and Yang [11] and Sun *et al*. [12]) can then be applied to further reduce the size of the feature set. In addition, other approaches (e.g., recurrence analysis of Chen and Yang [13], and Kan *et al.* [14]) are also developed in the literature to extract features from sensing signals. Although these works

made great contributions to the research in system informatics and quality engineering, they are designed for the monitoring of 1-D signals and are limited in their ability to handle high-dimensional images.

In the literature, pioneering works have been done to extract useful information from 2-D and 3-D profiles for process monitoring and control. Instead of directly using the imaging data, earlier efforts are made to obtain representative sample points from 2-D profiles. For example, Jin *et al*. [15] and Zhang *et al*. [16] collect a reduced set of thickness measurements to characterize the 2-D geometric shape of a wafer. Some works studied the spectral bands of hyperspectral images. For example, approaches were developed by Du *et al*. [17] and Wilcox *et al*. [18] to optimally select a subset of spectral bands from hyperspectral images for quality control. Some papers directly handle images or a sequence of images collected in the manufacturing process. For example, Yan *et al*. [19] proposed a low-rank tensor decomposition to extract features from images for process monitoring and anomaly detection. Megahed *et al*. [20] segmented the image into multiple regions of interests and developed a generalized likelihood ratio chart to monitor the color shift in those small regions. Yao *et al*. [21] proposed a 2-D multifractal analysis to characterize the surface integrity of products based on CT scan images. However, most of these studies focus on fault detection from snapshot images in discrete-part manufacturing. They are limited in their ability to monitor the time-varying image profiles that are highly nonlinear and nonstationary.

Recently, there is an increasing interest in bringing IoT into manufacturing and healthcare to increase the smartness level of the system. Some works have been done in the design of IoT-based cloud manufacturing. For example, see the four-layer cloud manufacturing system developed by Tao *et al* [22], real-time plant monitoring system by Georgakopoulos [23], and IoT-based manufacturing resources perception and access by Tao *et al.* [24]. Some research focus on the development of cyber-physical manufacturing systems. For example, the UML-based framework (i.e., UML4IoT) proposed by Thramboulidis and Christoulakis [25] and feature-based manufacturing system by Adamson *et al*. [26]. Some works leverage IoT technology for energy efficiency management. For example, Qin *et al*. [27] implemented IoT in the optimization of energy consumption in the additive manufacturing. Tan *et al*. [28] used IoT for the real-time monitoring of energy efficiency on manufacturing shop floors. Furthermore, IoT technology has been applied in manufacturing operations management. Xu and Chen [29] developed an IoT-based dynamic production scheduling framework for just-in-time manufacturing. Ding *et al*. [30] developed an

approach for optimal allocation of sensors in a multi-station assembly process. In addition, some research have been conducted to achieve IoT-enabled connected healthcare. For example, Pasluosta *et al*. [31] presented an IoT-enabled solution for the diagnosis and treatment of Parkinson's disease. Al-Taee *et al*. [32] developed a mobile health platform to incorporate humanoid robots for the management of diabetes in children. Islam *et al*. [33] presented a comprehensive survey on IoT-enabled healthcare applications. However, most of these works focus on the conceptual design of IoT infrastructures or hardware development. Little has been done to leverage sensing data from a large-scale IoT to develop new methods and tools for process monitoring, diagnosis, and improvement.

## 1.3. Research Objectives

The ***long-term goal of my research*** is to develop innovative sensor-based methodologies for the modeling, monitoring, and optimization of large-scale complex systems. Specifically, the ***objectives*** of this dissertation include:

1) Proposing a network model to optimally quantify the spatiotemporal dissimilarity among multi-channel signals for the anomaly detection.
2) Developing a dynamic network methodology to represent, model and control of time-varying image profiles for the in-situ process monitoring and quality control.
3) Developing a large-scale network model to efficiently handle big data in the Industrial Internet of Things for machine information processing, condition monitoring, and fault diagnosis.
4) Extending the large-scale network model for the multi-level modeling and analysis of the Internet of Health Things.

## 1.4. Organization of the Dissertation

This dissertation is organized based multiple manuscripts. Each of chapters 2, 3, 4, and 5 is written as a research paper, which has been published or under review. The outline of this dissertation is illustrated in Fig. 1-1.

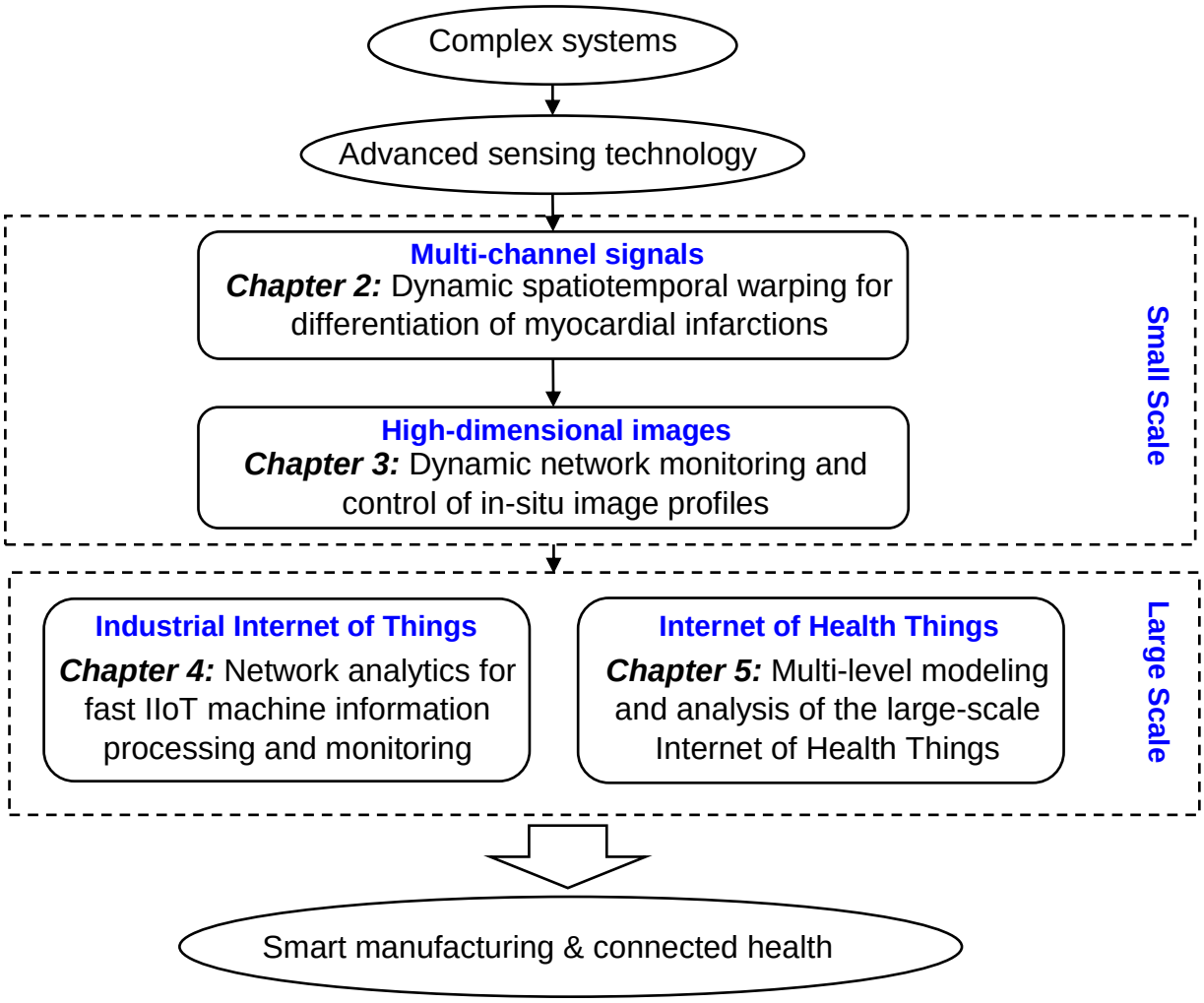



Figure 1-1. Outline and flow of this dissertation.

In Chapter 2, we address the challenge of multi-channel signals. A spatiotemporal warping approach is developed to quantify the dissimilarity of disease-altered patterns in the 3-lead vectorcardiogram. Each functional recording is then optimally embedded as a node in a high-dimensional network that preserves the dissimilarity matrix. Classification models are constructed based on the network, which yields a high accuracy for hierarchical differentiation of different types of myocardial infarctions.

In Chapter 3, the challenge of high-dimensional images is addressed by a new dynamic network scheme. Specifically, we first represent each image as a network. As such, a dynamic network is obtained from the stream of time-varying images. Then, we characterize the community structures of the dynamic network. Finally, a new control charting approach is developed to detect the change point of the underlying dynamics based on the community statistics. The developed

dynamic network scheme shows its effectiveness on the monitoring of both UPM and biomanufacturing processes.

In Chapter 4, we develop a large-scale network model to address the challenge of Industrial Internet of Things (IIoT). We measure the dissimilarity among machine signatures and further represent each machine signature as a node in the network. A stochastic learning approach is developed to reduce the computational complexity of network embedding. Also, a parallel computing scheme is developed to increase the computational efficiency for the embedding of large amounts of IIoT-enabled machines. The developed approach shows a strong potential for predictive maintenance and optimal machine scheduling in the context of IIoT.

In Chapter 5, we further extend the large-scale network model to address the challenge of the Internet of Health Things. In this study, we propose a new IoT technology, namely the Internet of Hearts (IoH), to empower the smart and connected heart health. We designed and developed a new approach of dynamic network modeling and parallel computing for multi-level cardiac monitoring: patient-to-patient variations in the population level and beat-to-beat dynamics in the individual patient level. After the construction of the large-scale dynamic network, control charts are proposed to harness network features for change detection of cardiac processes. The dynamic network reveals important information pertinent to patients' cardiac dynamics, which shows a strong potential for real-time cardiac monitoring and disease management.

In the end, Chapter 6 concludes the dissertation and summarizes the contributions. Future research directions are also discussed in this chapter.

# Chapter 2: Dynamic Spatiotemporal Warping for the Detection and Location of Myocardial Infarctions[1]


## Abstract

Myocardial infarction (MI), also known as heart attack, is the leading cause of death – about 452,000 per year – in US. It often occurs due to the occlusion of coronary arteries, thereby leading to the insufficient blood and oxygen supply that damage cardiac muscle cells. Because the blood vessels are all over the heart, MI can happen at different spatial locations (e.g., anterior and inferior portions) of the heart. The spatial location of diseases causes the variable excitation and propagation of cardiac electrical activities in space and time. Most of previous studies focused on the relationships between disease and time-domain biomarkers (e.g., QT interval, ST elevation/depression, heart rate) from 12-lead ECG signals. Few, if any, previous approaches have investigated how the spatial location of diseases will alter cardiac vectorcardiogram (VCG) signals in both space and time. This paper presents a novel warping approach to quantify the dissimilarity of disease-altered patterns in 3-lead spatiotemporal VCG signals. The hypothesis testing shows there are significant spatiotemporal differences between healthy controls (HC), MI-anterior, MI-anterior-septal, MI-anterior-lateral, MI-inferior, and MI-inferior-lateral. Further, we optimize the embedding of each functional recording as a feature vector in the high-dimensional space that preserves the dissimilarity distance matrix. This novel spatial embedding approach facilitates the construction of classification models and yields an accuracy of 94.7% for separating MIs and HCs and an accuracy of 96.5% for anterior-related MIs and inferior-related MIs.


## 2.1. Introduction

Myocardial infarction (MI), commonly known as a heart attack, is the leading cause of death in US. It is reported that nearly 452,000 Americans die from MI each year and almost every 34 seconds, there will be a heart attack in the US. MI occurs due to the occlusion of the coronary arteries. This results in the insufficient blood and oxygen supply that damages the cardiac muscle

[1] This chapter has been published in IEEE International Conference on Automation Science and Engineering (CASE) 2012

cells and trigger the disease-induced degradation process. The accurate detection of myocardial infarctions is critical for the timely medical interventions and the improvement of quality of life.

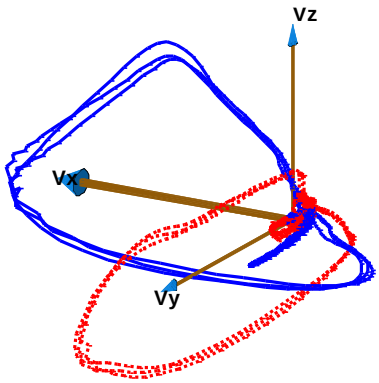


Figure 2-1. Cardiac electrical activities of HC (blue/solid) and MI subjects (red/dashed).

It is well known that cardiac electrical activities are spatio-temporally excited and propagated, i.e., initially excited at the sinoatrial (SA) node, conducted in both atria, then relayed through the atrioventricular (AV) node to further propagate through bundle of His and Purkinje fibers toward ventricular depolarization and repolarization. Because MI takes place in different spatial locations of the heart, e.g., anterior, inferior, space-time cardiac electrical activities are often perturbed by the location of lesions in the heart. As shown in Fig. 2-1, the MI and healthy control (HC) subjects show vastly dissimilar cardiac electrical activities in space and time. The blue/solid trajectories represent the vectorcardiogram (VCG) signals from a HC subject, and the red/dash trajectories are from a MI subject. Our previous studies demonstrated there are statistically significant differences in spatiotemporal paths of cardiac electrical activities between healthy and diseased subjects [8, 34, 35]. It may be noted that we developed the approaches of multiscale recurrence analysis [35] and spatial octant analysis [36] for the characterization and quantification of disease-altered cardiac electrical dynamics in space and time. However, few progresses have been made previously to quantitatively measure spatiotemporal distances of VCG signals between healthy and diseased subjects, and further utilize these distance measures to identify the locations of myocardial infarctions.

Most of previous studies focused on the relationships between diseases and time-domain biomarkers (e.g., QT interval, ST elevation or depression, heart rate) from ECG signals. As shown in Fig. 2-2, the 1-lead ECG is obtained by measuring the potential difference between 2 electrodes that are placed on the body surface (e.g., left foot, left arm and right arm). The ECG trace is often temporally segmented into P wave, QRS complex, and T wave (see Fig. 2-2). Each segment is closely associated with the specific physical activities of heart components [37]. For example, atrial depolarization (and systole) is represented by the P wave, ventricular depolarization (and

systole) is represented by the QRS complex, and ventricular repolarization (and diastole) is represented by the T wave.

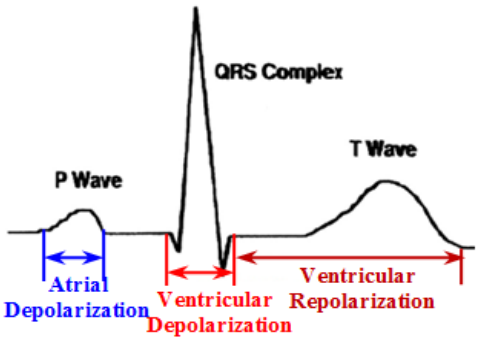


Figure 2-2. A time-domain ECG trace with P wave, QRS complex, and T wave.

However, time-domain projections of space-time cardiac electrical activities will diminish important spatial information of cardiac pathological behaviors. As such, these medical decisions that are made can be significantly influenced by such an information loss, especially when the lesion is in its initial stage. Therefore, multiple lead ECG systems, for e.g., 12-lead ECG and 3-lead VCG, are designed to provide multi-directional views of space-time electrical activities. The 12-lead ECG is widely used because medical doctors are accustomed to using them in the clinical practice. It has thus proven its value, time tested, and considered as the Gold Standard. However, much of that information is redundant and even in that, only a small fraction of the data is used by the physicians. As shown in Fig. 2-1, 3-lead VCG signals monitor cardiac electrical activities along three orthogonal X, Y, Z planes of the body, namely, frontal, transverse, and sagittal. The VCG vector loops contain 3 dimensional recurring and near-periodic patterns of space-time cardiac electrical dynamics. Dower et al. [38] and our previous study [39] showed that 3-lead VCG can be linearly transformed to 12-lead ECG without a significant loss of clinically useful information. Thus, 3-lead VCG surmounts not only the information loss in 1-lead ECG but also the redundant information in 12-lead ECG.

This paper presents a novel warping approach to quantify the dissimilarity of disease-altered patterns in 3-lead spatiotemporal VCG signals. We have addressed three research challenges in this presented investigation as follows:

1) The 3-lead VCG is a type of functional data that captures cardiac electrical activities in both space and time. Two VCG signals from different subjects may be misaligned due to the discrete sampling and phase shift. A novel warping method is developed to measure the distances between two misaligned functional signals.

2) However, the dissimilarity matrix, obtained from MI and HC subjects, can only be used for characterization but not directly for classification. A new embedding method is designed to

extract feature vectors from such a dissimilarity matrix, and these feature vectors will preserve the warping distances between functional recordings.

3) Furthermore, we quantitatively associate the disease properties (e.g., the spatial locations of myocardial infarctions) with the feature vectors that represent disease-altered spatiotemporal patterns underlying the 3-lead VCG signals.

The reminder of this paper is organized as follows: Section 2.2 introduces the methodology used in this research. Section 2.3 presents the material and experimental results, and Section 2.4 includes the discussion and conclusions arising out of this investigation.

## 2.2. Research Methodology

The research methodology is schematized as shown in Fig. 2-3. First, 3-lead VCG signals are utilized to represent the cardiac electrical activities in both space and time. Then, dynamic spatio-temporal warping algorithm is developed to measure dissimilarities between the 3-D VCG signals and a matrix of warping distances is obtained. Furthermore, spatial embedding algorithm is designed to extract feature vectors from the dissimilarity matrix. The functional recordings are optimally embedded as points in high-dimensional space and the warping distances among them are preserved. Finally, the supervised classification models of self-organizing map are developed to associate feature vectors with disease properties and identify the spatial locations of different MI subjects.

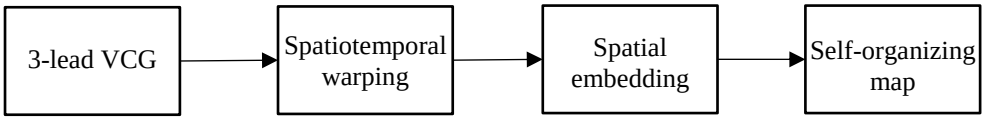



Figure 2-3. Flow diagram of the research methodology.

### *A. Dynamic Spatiotemporal Warping*

As shown in Fig. 2-1, spatiotemporal dissimilarity is found between the functional recordings of 3-lead VCG of MI and HC subjects. Quantification of such dissimilarity will provide a great opportunity for the identification of cardiovascular diseases. However, it is always challenging to measure the spatiotemporal dissimilarity between 2 functional signals in both space and time. Due to phase shift and discrete sampling, two VCG signals can be misaligned, e.g., both signals show a typical pattern and yet there are variations in shape, amplitude and phase between them. A new method needs to be developed for measuring the time-normalized spatial dissimilarity between 3-

D functional signals. Dynamic time warping is widely used in pattern recognition to measure the similarity between two time series. Previous researches have shown that dynamic time warping methods efficiently handle the misalignment in time series by searching the optimal matching path [40].

However, it may be noted that previous approaches are limited to the warping of 1-dimensional time series. The dissimilarity between the two time series is characterized mainly in time domain. As discussed above, cardiac electrical signals are excited and propagated spatio-temporally. The disease-altered signal patterns are shown in both space and time. Thus, as opposed to conventional 1-dimensional time warping methods, a dynamic spatiotemporal warping algorithm is developed in this study to measure the dissimilarities between space-time functional recordings.

Given two 3-D VCG signals $\overrightarrow{\boldsymbol{v_1}}(t)$ and $\overrightarrow{\boldsymbol{v_2}}(t)$, the time-normalized spatial distance between $\overrightarrow{\boldsymbol{v_1}}(t)$ and $\overrightarrow{\boldsymbol{v_2}}(t)$ is calculated as $\sum_{(t_i,t_j)\in p}\left\|\overrightarrow{\boldsymbol{v_1}}(t_i)-\overrightarrow{\boldsymbol{v_2}}(t_j)\right\|$ by alignment $p$. The warping path $p(i,j)$ connects (1, 1) and (*M*, *N*) in a 2-dimensional square lattice as well as satisfying constraints such as monotonicity condition and step size condition. To find the optimal path, an exhaustive search of alignment path is intractable and computationally expensive. However, this problem is solved efficiently using dynamic programming (DP) methods. The DP algorithm is started at the initial condition: $g(1,1)=d(1,1)=\|\overrightarrow{\boldsymbol{v_1}}(t_1)-\overrightarrow{\boldsymbol{v_2}}(t_1)\|$ and is searching forward as:

$$g(i,j)=min\begin{pmatrix} g(i,j-1)\ +\ d(i,j) \\ g(i-1,j-1)\ +\ d(i,j) \\ g(i-1,j)\ +\ d(i,j) \end{pmatrix}$$

Finally, the time-normalized spatial distance is calculated as:

$$\Delta(\overrightarrow{\boldsymbol{v_1}},\overrightarrow{\boldsymbol{v_2}})=\frac{g(\mathrm{M},\mathrm{N})}{\mathrm{M}+\mathrm{N}}$$

which represents the spatio-temporal dissimilarity between two multi-dimensional functional recordings. Therefore, disease-altered characteristics of 3-lead VCG signals are obtained in the warping matrix $\Delta$.

As shown in Fig. 2-4 (a), the warping matrix is symmetric with on-diagonal squares representing the in-group dissimilarities and off-diagonal squares representing the dissimilarities across subject groups. In this present investigation, subject groups include 309 MIs and 79 HCs. Within the 309 MI recordings, there are 47 anterior (AN), 77 anterior-septal (ANSP), 43 anterior-

lateral (ANLA), 89 inferior (IN) and 53 inferior-lateral (INLA). Then, the warping matrix is mapped into a color scale as shown in Fig. 2-4 (b). The small distances are in blue colors with gradual transition to the large distances in red colors. It may be noted that in-group distances are smaller than between-group distances. For example, the on-diagonal squares of HC-HC, AN-AN, ANSP-ANSP, ANLA-ANLA, IN-IN and INLA-INLA are mainly in blue colors, while the off-diagonal squares display more red colors. However, some off-diagonal squares are also in blue colors, e.g., AN-ANSP, IN-INLA. This is because the locations of diseases are spatially close, and the disease-altered patterns of 3-lead VCG are less dissimilar.

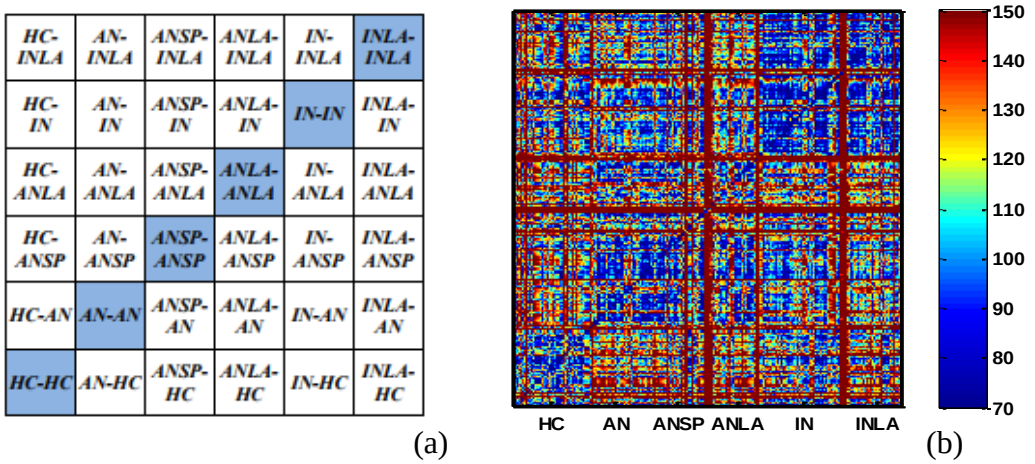



Figure 2-4. (a) Labels between each pair of groups in the warping distance matrix; (b) Color mapping plot of the warping distance matrix.

In addition, we have statistically tested the significance of spatiotemporal distances for different MI locations using 3 hypothesis tests, i.e., two-sample unpaired *t*-test, F-test and Kolmogorov-Smirnov (KS) test. As shown in Table 2-1, "HC-HC" denotes the distances within HC groups, and "HC-AN" denotes the distances between HC and MI anterior (AN) groups. The nomenclature is the same for all other groups.

Table 2-1. Two Sample t-Test for Testing the Mean of Spatiotemporal Distances at Different MI Locations

| Group | *HC-HC vs. HC-AN* | *HC-HC vs. HC-ANSP* | *HC-HC vs. HC-ANLA* | *HC-HC vs. HC-IN* | *HC-HC vs. HC-INLA* |
|---|---|---|---|---|---|
| $t_0$ | -3.55 | -4.79 | -4.40 | -2.58 | -2.74 |
| t crit. | 1.98 | 1.98 | 1.98 | 1.97 | 1.98 |
| **Group** | ***AN-AN vs. AN-ANSP*** | ***AN-AN vs. AN-ANLA*** | ***AN-AN vs. AN-IN*** | ***AN-AN vs. AN-INLA*** | ***ANSP-ANSP vs. ANSP-ANLA*** |
| $t_0$ | -1.96 | -1.27 | -2.37 | -1.90 | -0.01 |
| t crit. | 1.98 | 1.99 | 1.98 | 1.98 | 1.98 |
| **Group** | ***ANSP-ANSP vs. ANSP-IN*** | ***ANSP-ANSP vs. ANSP-INLA*** | ***ANLA-ANLA vs. ANLA-IN*** | ***ANLA-ANLA vs. ANLA-INLA*** | ***IN-IN vs. IN-INLA*** |
| $t_0$ | -0.54 | 0.02 | -1.62 | -1.11 | 0.35 |
| t crit. | 1.97 | 1.98 | 1.98 | 1.99 | 1.98 |

First, the central tendency of spatiotemporal distances is tested for different MI location with 2-sample *t*-test. Let $\mu_i$ and $\mu_j$denote the means of $i^{th}$ and $j^{th}$ groups, the hypotheses of *t*-test are:

$$H_0: \mu_i = \mu_j \ \& \ H_1: \mu_i \neq \mu_j$$

For each pair of groups, the *t*-test statistic is compared with critical value given by $t_{\alpha/2,\, n_1+n_2-2}$, where $n_1$and $n_2$ are the number of independent observations in the corresponding group and $\alpha$ = 0.05. If $|t_0|>t_{\alpha/2,\, n_1+n_2-2}$, the null hypothesis $H_0$ will be rejected and the mean values of two groups are declared to be different at significant level of 0.05. Statistically significant groups are marked as red colors in Table 1-1. It may be noted that the distances within HC group are significantly different from MI groups (i.e., AN, ANSP, ANLA, IN and INLA) in terms of mean values. It is also found that MI-anterior (AN) and MI-inferior (IN) are statistically significant in the mean values of spatiotemporal distances.

Second, the dispersion of spatiotemporal distances is tested for different MI location with 2-sample F-test. Let $\sigma_i^2$ and $\sigma_j^2$ denote the variances of $i^{th}$ and $j^{th}$ groups, the hypotheses of F-test are:

$$H_0: \sigma_i^2 = \sigma_j^2 \ \& \ H_1: \sigma_i^2 \neq \sigma_j^2$$

The test statistics are compared with critical values $F_{\alpha/2,\, n_1-1,\, n_2-1}$ and $F_{1-\alpha/2,\, n_1-1,\, n_2-1}$. Here, $n_1$and $n_2$ are the number of independent observations in corresponding groups and the significant level $\alpha = 0.05$. If $F_0 > F_{\alpha/2,\, n_1-1,\, n_2-1}$ or $F_0 < F_{1-\alpha/2,\, n_1-1,\, n_2-1}$, the null hypothesis $H_0$ will be rejected at significant level of 0.05 and there are significant differences in the variances of spatiotemporal distances for two groups. As shown in Table 2-2, the statistics and critical values

of significant groups are marked in red colors. The variances within the HC group are significantly different from MI-anterior-septal (ANSP) and MI-inferior (IN). In addition, the variance of spatiotemporal distances within MI-anterior (AN) is different from the distances between MI-anterior (AN) and MI-anterior-septal (ANSP). The variances of MI-inferior-lateral (INLA) are significantly different from another two groups, i.e., MI-anterior-lateral (ANLA) and MI-inferior (IN).

Table 2-2. Two Sample F-Test for Testing the Variance of Spatiotemporal Distances at Different MI Locations

| **Group** | ***HC-HC vs. HC-AN*** | ***HC-HC vs. HC-ANSP*** | ***HC-HC vs. HC-ANLA*** | ***HC-HC vs. HC-IN*** | ***HC-HC vs. HC-INLA*** |
|---|---|---|---|---|---|
| **$F_0$** | **1.17** | **2.10** | **1.22** | **1.57** | **1.06** |
| **F crit.** | **1.71** | **1.56** | **1.74** | **1.53** | **1.67** |
| **Group** | ***AN-AN vs. AN-ANSP*** | ***AN-AN vs. AN-ANLA*** | ***AN-AN vs. AN-IN*** | ***AN-AN vs. AN-INLA*** | ***ANSP-ANSP vs. ANSP-ANLA*** |
| **$F_0$** | **2.02** | **1.11** | **1.21** | **1.63** | **1.33** |
| **F crit.** | **1.65** | **1.82** | **1.63** | **1.75** | **1.75** |
| **Group** | ***ANSP-ANSP vs. ANSP-IN*** | ***ANSP-ANSP vs. ANSP-INLA*** | ***ANLA-ANLA vs. ANLA-IN*** | ***ANLA-ANLA vs. ANLA-INLA*** | ***IN-IN vs. IN-INLA*** |
| **$F_0$** | **1.14** | **1.45** | **1.11** | **1.91** | **1.71** |
| **F crit.** | **1.54** | **1.67** | **1.65** | **1.77** | **1.65** |

Third, two-sample KS test is utilized to test the differences in cumulative distribution function (CDF) for MIs at various locations. Let $F_i(x)$and $F_j(x)$ denote cumulative distribution functions of the ith and jth groups respectively and the hypotheses of KS test are given as:

$$H_0: F_i(x) = F_j(x) \;\&\; H_1: F_i(x) \neq F_j(x)$$

The test statistics (KS stat.) and critical values (KS crit.) are shown in the Table 2-3. The KS test statistic is compared with the corresponding critical value given by $c_\alpha\sqrt{\frac{n_1+n_2}{n_1 n_2}}$, where n1 and n2 are the number of independent observations in corresponding groups, the significant level is $\alpha = 0.05$, and $c_\alpha$is approximated as 1.36. If the KS statistic is greater than the critical value, the null hypothesis $H_0$ will be rejected and the cumulative distribution functions of two groups are declared to be different at significant level of 0.05. As shown in Table 2-3, four significant groups are highlighted in red colors. It may be noted that, the cumulative distribution function of HC group is significantly different with MI-anterior (AN), MI-anterior-lateral (ANLA), MI-anterior-septal (ANSP) and MI-inferior-lateral (INLA).

Table 2-3. Two Sample KS-Test for Testing the CDF of Spatiotemporal Distances at Different MI Locations

| **Group** | ***HC-HC vs. HC-AN*** | ***HC-HC vs. HC-ANSP*** | ***HC-HC vs. HC-ANLA*** | ***HC-HC vs. HC-IN*** | ***HC-HC vs. HC-INLA*** |
|---|---|---|---|---|---|
| **KS stat.** | **0.35** | **0.34** | **0.39** | **0.17** | **0.26** |
| **KS crit.** | **0.25** | **0.21** | **0.25** | **0.21** | **0.24** |
| **Group** | ***AN-AN vs. AN-ANSP*** | ***AN-AN vs. AN-ANLA*** | ***AN-AN vs. AN-IN*** | ***AN-AN vs. AN-INLA*** | ***ANSP-ANSP vs. ANSP-ANLA*** |
| **KS stat.** | **0.11** | **0.09** | **0.18** | **0.19** | **0.06** |
| **KS crit.** | **0.25** | **0.28** | **0.24** | **0.27** | **0.25** |
| **Group** | ***ANSP-ANSP vs. ANSP-IN*** | ***ANSP-ANSP vs. ANSP-INLA*** | ***ANLA-ANLA vs. ANLA-IN*** | ***ANLA-ANLA vs. ANLA-INLA*** | ***IN-IN vs. IN-INLA*** |
| **KS stat.** | **0.11** | **0.12** | **0.16** | **0.20** | **0.05** |
| **KS crit.** | **0.21** | **0.24** | **0.25** | **0.27** | **0.23** |

The results of three hypothesis tests show that the spatiotemporal distances within the HC group are significantly different from myocardial infarctions at different locations. Furthermore, anterior myocardial infarctions (AN) show significant differences in terms of means and variances from inferior myocardial infarctions (IN). It may also be noted that there are significant differences within the group of anterior (AN) or inferior (IN) subjects, e.g., MI-inferior (IN) is different from MI-inferior-lateral (INLA) in the variances of spatiotemporal distances.

### *B. Spatial Embedding*

The results of statistical tests show significant spatiotemporal differences among groups. However, it may be noted that the warping matrix itself can hardly be used as features for classification models. In addition, the measure of Euclidean distance is not directional and can mix the distances that are equal in magnitudes but along different spatial directions. A novel method needs to be developed for transforming the warping matrix into feature vectors that preserve the warping distances among functional recordings. The spatial embedding method embeds the functional recordings as the points in a high-dimensional space. These points can be used as features vectors that recover not only the distance matrix but also directional differences between functional recordings.

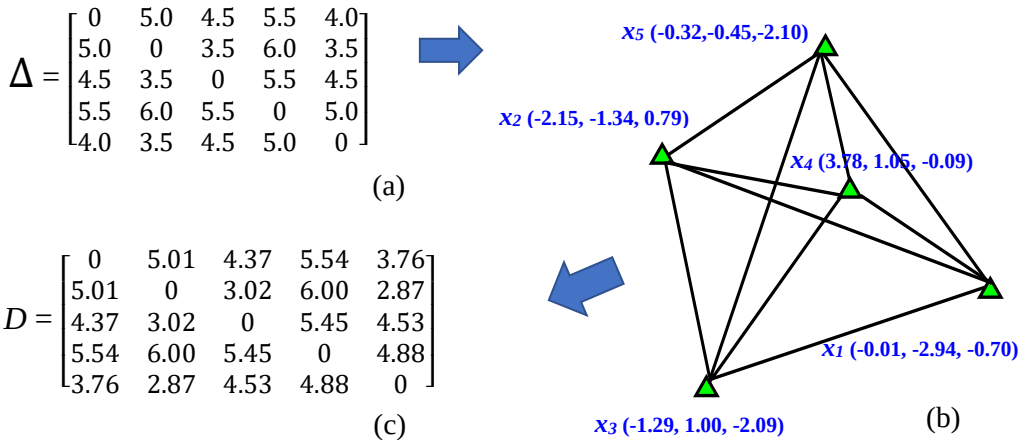



Figure 2-5. Spatial embedding algorithm; (a) original distance matrix Δ; (b) network and nodes; (c) Reconstructed distance matrix D.

This is similar to a network problem, i.e., how to reconstruct the nodes' location in a high-dimensional space if the node-to-node distance matrix is known. As shown in Fig. 2-5, a network comprises a number of nodes that are connected by edges. Each node stands for an individual component in the system and the edges show the relationship (e.g., distances or causal relationships) between nodes. As given in Fig. 2-5 (a), assume the distance matrix Δ for 5 nodes is known. If the network is reconstructed in the 3-dimensional space, this is analogues to optimally identify the coordinate vector $\boldsymbol{x}_i = (x_{i1}, x_{i2}, x_{i3}),\ i = 1, 2, \ldots, 5$ for 5 nodes that can preserve the distance matrix Δ. As shown in Fig. 2-5 (b), all the nodes and their connections preserve the dissimilarities matrix Δ. The matrix *D* is the pairwise distances between reconstructed nodes in the 3-dimensional space (see Fig. 2-5 (c)). It may be noted that we are bridging from functional signals to the distance matrix to feature vectors (nodes in the network) with the use of spatial embedding method. The feature vectors should approximately preserve the distance matrix Δ between functional signals.

Let $\delta_{ij}$ denotes the dissimilarity between $i^{th}$ and $j^{th}$ functional recordings in $n \times n$ warping matrix Δ, $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ denote the $i^{th}$ and $j^{th}$ feature vectors in a high-dimensional space. Then, the optimization of spatial embedding algorithm is formulated as:

$$min \sum_{i<j} (\|\boldsymbol{x}_i - \boldsymbol{x}_j\| - \delta_{ij})\,;\ i, j\ \in [1, n]$$

where $\|\cdot\|$ is the Euclidean norm. This optimization problem is solved as follows: Given an $n \times n$ distance (dissimilarity) matrix Δ, the squared matrix $\Delta^{(2)}$ is obtained. Each element in $\Delta^{(2)}$ is

$(\delta_{ij})^2$, i.e., the squares of $\delta_{ij}$ in the matrix Δ. Secondly, a Gram matrix $B$ is constructed as: $B = -\frac{1}{2}H\Delta^{(2)}H$, where the centering matrix H = $I - n^{-1}\mathbf{1}\mathbf{1}^T$ and 1 is a column vector with $n$ ones. The matrix B is decomposed as: $B = V\Lambda V^T$, where $V$ = [$\boldsymbol{v_1}$, $\boldsymbol{v_2}$, ..., $\boldsymbol{v_n}$] is a matrix of eigenvectors and $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ is a diagonal matrix of eigenvalues. Then, the Gram matrix $B$ is rewritten as: $B = V\sqrt{\Lambda}\sqrt{\Lambda}V^T$. Therefore, feature vectors will be obtained as: $X = V\sqrt{\Lambda}$.

The spatial embedding algorithm embeds each functional recording as a point in the $n$-dimensional space ($n$ = 2, 3, 4...). Fig. 2-6 (a) and (b) illustrated the 2-D and 3-D embedding. The blue circles represent HCs and red triangles represent MIs. It may be noted that HC subjects are away from MI subjects and there are two distinct clusters in Fig. 2-6 (a) and (b).

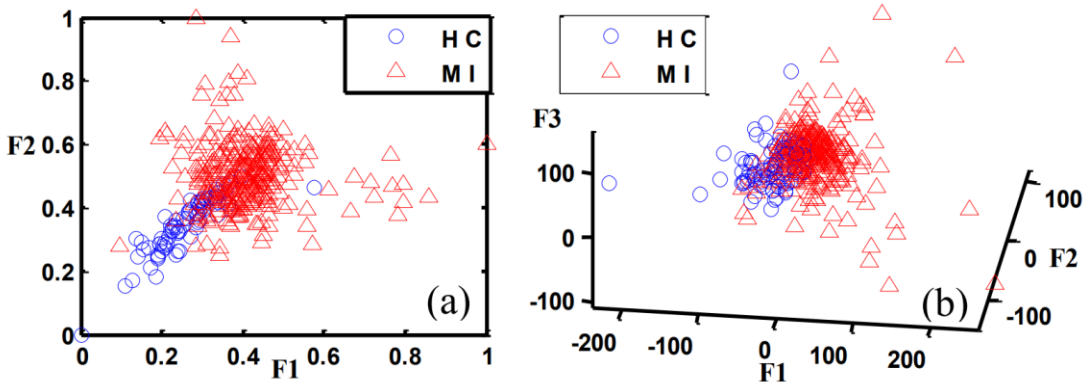



Figure 2-6. (a) 2-D and (b) 3-D scatter plots of spatial embedding vectors.

## *C. Self-organizing Map*

Self-organizing map (SOM) is utilized to investigate how self-organized patterns of warping features are correlated with the spatial locations of myocardial infarctions. The SOM represents the high-dimensional data in a low dimensional map, preserves the topological relationahip that cannot be visualized otherwise, and organizes the data automatically according to inherent structures [41]. The supervised SOM integrates the label information as an additional element in the feature vector during the training phase [42]. In this investigation, the SOM is utilized as a classification model to identify MI at different locations based on the feature vectors obtained with the method of spatial embedding. For more information on supervised and unsupervised SOM, see, for example, [41, 42].

## 2.3. Materials and Experimental Results

In this investigation, 3-lead VCG recordings from 388 subjects (309 MIs and 79 HCs) available in the PTB Database of PhysioNet [10] are analyzed to evaluate the performance of developed research methodologies. The VCG signals were digitized at 1 kHz sampling rate with a 16-bit resolution over a range of ±16.384 mV. Within the 309 MI recordings, there are 47 anterior (AN), 77 anterior-septal (ANSP), 43 anterior-lateral (ANLA), 89 inferior (IN) and 53 inferior-lateral (INLA). To optimize the design of research methodology, we have conducted experiments to address the following questions:

1) *The influence of embedding dimension*: As the embedding dimension increases, the warping dissimilarities between functional recordings are better preserved. However, a higher embedding dimension will increase the model complexity and introduce "curse of dimensionality" problem. A lower embedding dimension may not fully preserve the distance relationships of functional recordings. In this study, a cost function is defined to represent the performance of each embedding dimension. Experiments are conducted to investigate how the performance will vary as the embedding dimension increases.

2) *The influence of neuron map size*: The classification performance of SOM can be influenced by the map size. The larger the map size, the better performance can be achieved. However, a large number of neurons will result in "overfitting" problems. In this study, experiments are conducted for selecting optimal neuron map size.

3) *Hierarchical classification*: As shown in Fig. 2-7, the structure of hierarchical classification consists of 4 levels: (1) *level 1* – HC vs. MI subjects. (2) *level 2* – AN-all vs. IN-all. The group of AN-all includes MI-anterior, MI-anterior-septal and MI-anterior-lateral. The group of IN-all includes MI-inferior and MI-inferior-lateral. (3) *level 3* – within AN-all, MI-anterior-septal and MI-anterior-lateral are combined as AN* to be separated from MI-anterior. Within IN-all, MI-inferior and MI-inferior-lateral are classified. (4) *level 4* – MI-anterior-septal vs. MI-anterior-lateral.

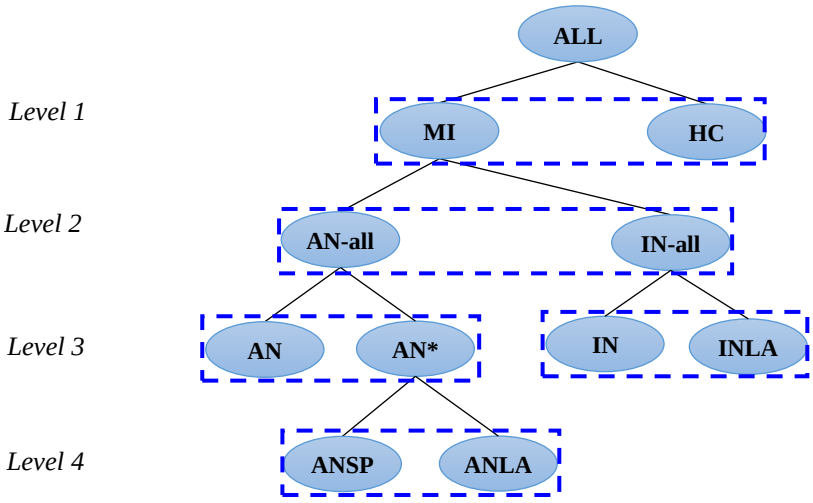



Figure 2-7. Structure of hierarchical classification

In addition, K-fold cross validation (i.e., K = 4 in this study) is used to estimate the performance of classification models and prevent the "overfitting" problem. Three statistics, i.e., correct rate (CR), sensitivity (SEN) and specificity (SPEC), are used as model performance metrics.

*A. Influence of embedding dimension*

The *stress* metric is used to optimally select the embedding dimension. The *stress* is calculated as $\sqrt{\sum\sum(\delta_{ij} - d_{ij})^2)/\sum\sum {d_{ij}}^2}$, where $d_{ij}$is the distance between feature vectors $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$, and $\delta_{ij}$ is the dissimilarity matrix $\Delta$ measured from functional recordings. The embedding dimension is varied from 1 to 16 and the stress values are as shown in Fig. 2-8. The decreasing trend of stress values becomes stable when the embedding dimension reaches 8, 9 and higher. For 8-D embedding, the stress value is around 0.058. The stress value is 0.0393 for the 10-D embedding. As shown in Fig. 2-8, the stress is 0.0337 when the embedding dimension is increased to 12. Thus, we chose the embedding dimension to be 10 in this present study.

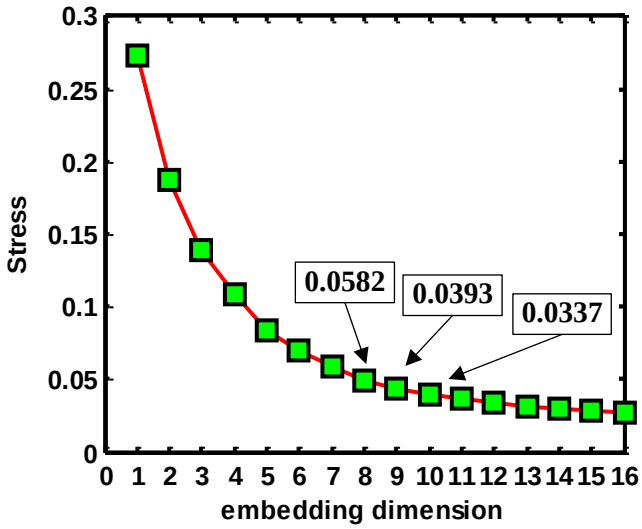



Figure 2-8. Stress values for the embedding dimensions from 1 to 16

Table 2-4. The Variations of Correct rates with Different SOM Map Sizes

| **Correct Rate (%)** | | **SOM Map Width** | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| **SOM Map Length** | 3 | 89.5 | 91.3 | 91.6 | 92.0 | 92.9 | 92.7 | 92.4 | 93.5 |
| | 4 | 91.2 | 92.0 | 92.8 | 92.6 | 93.2 | 93.8 | 93.4 | 93.8 |
| | 5 | 91.4 | 92.8 | 93.2 | 93.2 | 93.9 | 94.2 | 94.9 | 95.1 |
| | 6 | 92.1 | 93.0 | 93.2 | 93.8 | 94.6 | 94.8 | 95.4 | 95.8 |
| | 7 | 92.8 | 93.3 | 93.4 | 94.7 | 95.0 | 96.0 | 95.8 | 96.8 |
| | 8 | 92.5 | 93.7 | 93.9 | 94.1 | 95.3 | 95.8 | 96.3 | 97.0 |
| | 9 | 92.7 | 93.9 | 94.3 | 95.2 | 95.4 | 95.9 | 96.6 | 96.9 |
| | 10 | 93.2 | 94.1 | 94.7 | 95.7 | 95.9 | 96.4 | 96.6 | 96.5 |

*B. Influence of neuron map size*

As shown in Table 2-4, the correct rates are obtained when the length and width of SOM map are increased from 3 to 10. It may be noted that correct rates yield an increasing trend when the map length and width are getting bigger, and all the SOMs show >89% performances. In addition, the *corrected version of Akaike information criterion* (*AICc*) [108] is calculated to select an optimal SOM model as:

$$AICc = 2k + Nln(QE) + \frac{2k(k+1)}{N-k-1}$$

where $k$ is the number of neurons in the SOM model, and $N$ is the number of recordings. The quantization error (QE) is defined as: $QE = \frac{1}{N}\sum_{n=1}^{N}\left\|\boldsymbol{x}_n - \boldsymbol{w}_{c(n)}\right\|$, where $\boldsymbol{x}_n$is the input feature vector and $\boldsymbol{w}_{c(n)}$is the weighting vector of the best matching neuron for the input feature vector

$\boldsymbol{x}_n$. It may be noted that $2k$ represents the penalty of model complexity, and $Nln(QE)$ represents model performance. The term of $\frac{2k(k+1)}{N-k-1}$ is added to adjust the penalty of model complexity (i.e., $2k$) on the basis of sample size $N$ [108]. In other words, a balance between model performance and complexity will be achieved by minimizing the AICc values. As shown in Fig. 2-9, the response surface of BIC has higher amplitudes at both small and large map sizes. An optimal SOM map size of (7, 6) (i.e., red dot in Fig. 2-9) is selected for further classification and visualization of SOM structure.

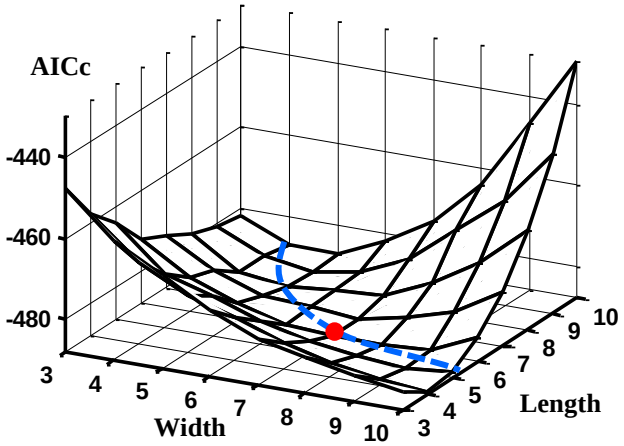



Figure 2-9. Response surface of AICc for different SOM map sizes.

## *C. Hierarchical Classification*

Table 2-5. Hierarchical Classification Results

| Level | Group | CR (%) | SEN (%) | SPEC (%) |
|---|---|---|---|---|
| *Level 1* | HC vs. MI | 94.7 | 93.7 | 98.7 |
| *Level 2* | AN-all vs. IN-all | 96.5 | 97.0 | 96.1 |
| *Level 3* | AN vs. AN* | 89.5 | 88.1 | 93.1 |
| | IN vs. INLA | 91.7 | 93.4 | 90.7 |
| *Level 4* | ANSP vs. ANLA | 93.1 | 93.0 | 93.1 |

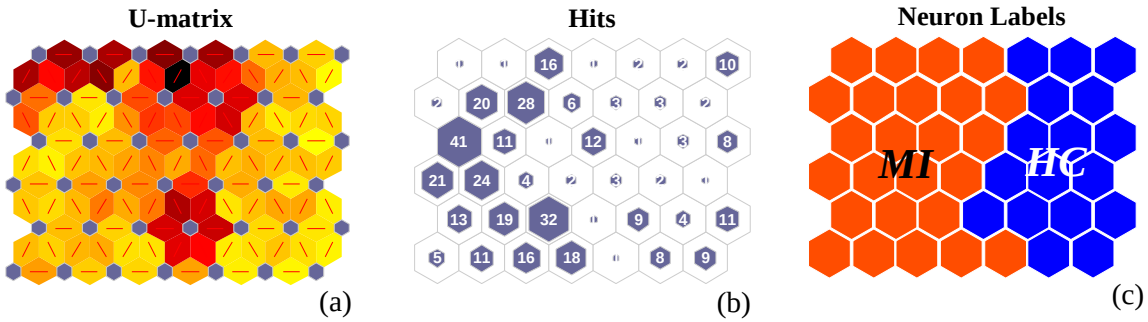



Figure 2-10. (a) SOM U-matrix for the neuron distances; (b) SOM sample hits plot; (c) SOM Neuron labels of HC (blue) and MI (red)

With the selected embedding dimension 10 and SOM map size (7, 6), the results are shown in Table 2-5 for the hierarchical classification (see Fig. 2-7). At level 1 (HC and MI), we achieved a correct rate of 94.7%, a sensitivity of 93.7% and a specificity of 98.7%. As shown in Fig. 2-10 (a), the distances between neurons in the SOM are represented in a U-matrix. It may be noted that the distances in left and right portions are in light yellow colors and separated by the darker colors in the middle. The U-matrix indicates that features vectors are organized into 2 clusters according to the self-organizing SOM structure. Fig. 2-10 (b) and (c) shows the plots of SOM sample hits and neuron labels. It may be noted that the neurons in the left portion of SOM map (colored as light red) represent MI and the right portion (colored as deep blue) represent HC. The two groups are separated effectively in the SOM map. At level 2, i.e., all anterior infarctions vs. all inferior infarctions, we achieved 96.5%, 97% and 96.1% for CR, SEN and SPEC respectively. At level 3, MI-anterior-septal and MI-anterior-lateral are combined and compared with MI-anterior. 89.5%, 88.1% and 93.1% are achieved for CR, SEN and SPEC respectively. In addition, MI-inferior and MI-inferior-lateral are classified with the performances of 91.7%, 93.4% and 90.7% for CR, SEN and SPEC respectively. Finally, at level 4, 93.1%, 93.0% and 93.1% for CR, SEN and SPEC respectively are reached for the classification of MI-anterior-septal and MI-anterior-lateral.

## 2.4. Conclusion and Discussions

Few, if any, previous approaches have characterized and quantified the patterns of spatio-temporal cardiac electrical activities that are altered by myocardial infarctions at different locations. As commonly known, MI can take place in different portions of the heart, i.e., anterior, anterior-lateral, anterior-septal, inferior and inferior–lateral. Space-time cardiac electrical activities are often perturbed by the spatial location of lesions. In this present study, we

characterized the disease-altered cardiac electrical activities in the 3-lead VCG. However, the VCG signals are often misaligned due to discrete sampling and phase shift. A novel method of dynamic spatio-temporal warping is designed to overcome the misalignment problem and measure the spatiotemporal distances between VCG signals. The hypothesis tests show that there are significant spatio-temporal differences between the subject groups in the dissimilarity matrix. Furthermore, functional recordings are optimally embedded as points in a high-dimensional space that preserve the warping distance matrix. Finally, the proposed multi-class classification models yield superior classification results (>94.7%) in the first 2 levels. At levels 3 and 4, because MIs are spatially close to each other, disease-altered cardiac electrical activities are similar, thereby achieving a relatively low accuracy (>89.5%). The developed methodologies facilitate the understanding of disease-altered spatio-temporal patterns in the complex cardiac systems.

# Chapter 3: Dynamic Network Modeling and Control of in-situ Image Profiles from Ultraprecision Machining and Biomanufacturing Processes[2]


## Abstract

In modern industries, advanced imaging technology has been more and more invested to cope with the ever-increasing complexity of systems, improve the visibility of information, and enhance operational quality and integrity. As a result, large amounts of imaging data are readily available. This presents great challenges on the state-of-the-art practices in process monitoring and quality control. Conventional statistical process control (SPC) focuses on key characteristics of the product or process, and is rather limited to handle complex structures of high-dimensional imaging data. New SPC methods and tools are urgently needed to extract useful information from in-situ image profiles for process monitoring and quality control. In this study, we developed a novel dynamic network scheme to represent, model and control of time-varying image profiles. Potts model Hamiltonian approach is introduced to characterize community patterns and organizational behaviors in the dynamic network. Further, new statistics are extracted from network communities to characterize and quantify dynamic structures of image profiles. Finally, we design and develop a new control chart, namely network generalized likelihood ratio chart (NGLR), to detect the change point of the underlying dynamics of complex processes. The proposed methodology is implemented and evaluated for real-world applications in ultra-precision machining and biomanufacturing processes. Experimental results show that the proposed approach effectively characterize and monitor the variations in complex structures of time-varying image data. The new dynamic network SPC method is shown to have strong potentials for general applications in a diverse set of domains with in-situ imaging data.


[2] This Chapter has been published in the Journal of Quality and Reliability Engineering International

## 3.1. Introduction

With the rapid advancement of technology, modern industries are shifting from sparse and episodic measurement to online and high-resolution sensing such as laser-scanning video microscopy. As such, high-dimensional image data become readily available. This provides an unprecedented opportunity to improve the process controllability, address the ever-increasing complexity of systems, and achieve competitive advantages in the global market. As shown in Fig. 3-1a, ultra-precision machining (UPM) is equipped with air-bearing spindles and diamond tools to produce optical surface finishes. The 2D and 3D images are often collected to inspect the quality of surface finishes, which is closely pertinent to machine setup and the operation of cutting tools [43]. Fig. 3-1b shows time-varying 3D (4D) images of cardiac electrical waves captured through optimal mapping techniques. Such dynamic images are critical to the study of the pathology of heart diseases. While advanced sensing improves the level of intelligence and automation in manufacturing and healthcare systems, dynamic image data present great challenges on existing methodologies of statistical process control (SPC).

(a) (b)

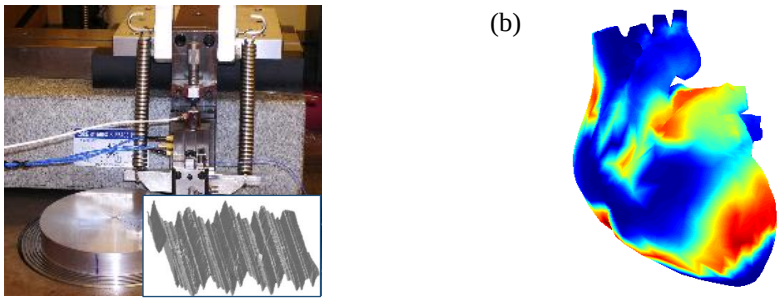


Figure 3-1. (a) Ultra-Precision Machining; (b) whole-heart optical mapping.

Traditional SPC methods and tools primarily focus on key characteristics of the product and the conformance to specifications. In order to address data complexity, there has been significant development of SPC methods in the past two decades. For example, a number of control schemes were designed to monitor discrete data points, linear profiles, and nonlinear waveforms. In addition, image data are increasingly utilized for quality inspection in various fields, including nanomanufacturing [44], semiconductor manufacturing [16] and food industry [17]. However, most previous works reported in the literature focused on the detection of stationary faults from snap-shot images (e.g., wafer heat map and hyperspectral image of poultry carcasses). In other words, existing image-based SPC methods are mainly developed to inspect whether pixel

intensities are uniform within a region of interest or whether it conforms to a nominal pattern. In-situ sensing, however, generates a stream of time-varying images and dynamic faults/defects that are nonlinear and nonstationary. As such, new SPC methods and tools are urgently needed to extract useful information from in-situ image profiles for process monitoring and quality control.

This paper presents a new dynamic network scheme that is designed and developed to represent, model, and monitor real-time image data. It may be noted that dynamic network models have intrinsic advantages to handling complex structures of image data: (1) *Nodes of the network can readily represent the data with a high dimensionality*; (2) *Edges of the network enables the perseveration of complex structures and patterns in the dynamic image*. In the proposed approach, each image is firstly represented as a network. As such, a dynamic network is obtained from the stream of time-varying images. Second, we introduce an approach of Potts model Hamiltonian to characterize community structures of the dynamic network. When the Hamiltonian function is minimized, strongly connected network nodes are clustered into the same community. Third, we design and develop a new control chart, namely *network generalized likelihood ratio chart* (NGLR), to detect the change point of the underlying dynamics of complex processes. Case studies on UPM and biomanufacturing processes show that the developed NGLR approach effectively characterizes and monitors the variations in complex structures of time-varying image data.

The remainder of this paper is organized as follows: Section 3.2 introduces the research background. Section 3.3 presents the developed dynamic network methodology. Case studies are demonstrated in Section 3.4. Section 3.5 concludes and discusses this present investigation.

## 3.2. Research Background

In the past decade, the SPC methodology has been advanced from monitoring individual quality features to linear profiles, nonlinear waveforms, and image data. Recently, the wide application of in-situ imaging technology brings the proliferation of image data. This, in turn, fuels increasing interests to develop new image-based SPC methods and tools for process monitoring and quality control.

### *A. Profile Monitoring*

At the early stage, linear profiles were widely used to represent the relationship between control parameters and process characteristics for quality control. For example, Kang and Albin [1]

delineated the performance of a mass flow controller as a linear function. Parameters such as the slope and intercept of linear models are monitored by a multivariate Hotelling $T^2$ chart. Zhou *et al*. [5] proposed a linear mixed-effects model to investigate fault diagnosability. A general framework of diagnosability analysis was developed for the purpose of fault diagnosis in multi-stage manufacturing process. Mahmoud and Woodall [45] developed a linear model to analyze Phase I behaviors of the data gathered in calibration applications. Control charts based on slope and intercept were constructed for the root cause diagnosis. Zou *et al*. [3] designed a change-point model with likelihood ratio test to detect shifts in the slope, intercept and standard deviation of the linear function, which was used to characterize linear profiles. In addition, Mahmoud *et al*. [46] used a segmented regression approach to characterize linear profile data. The constancy of regression parameters was evaluated by a likelihood ratio test model for change-point detection.

However, linear functions were not adequate to model the data when more complex data structures were presented, e.g., nonlinear waveforms. As such, many process monitoring schemes were developed to model and analyze nonlinear profiles. For example, Koh and Shi *et al.* [47] integrated engineering knowledge with Haar transformation for tonnage signal analysis and fault detection in stamping processes. Jin and Shi [48] compressed tonnage signals in a "feature-preserving" away, which are further decomposed to obtain signals from individual station in transfer or progressive die processes. Bukkapatnam *et al*. [49] and Kamarthi *et al*. [50] characterized acoustic emission, cutting force and vibration signals for chatter control, flank wear estimation and fault diagnosis in lathe-cutting processes. Ding *et al*. [7] analyzed the nonlinear profile in Phase I for dimensionality reduction and data separation. Nonlinear profiles were transformed into lower dimensions to reveal pertinent quality information. Kim *et al*. [51] proposed a principal curve method for change-point detection and fault diagnosis with multi-channel sensing signals collected from forging process. Zhou *et al*. [9] monitored cycle-based waveform signals by Haar wavelet decomposition. The location of process shift as well as its magnitude was detected by the developed wavelet scheme. Jeong and Lu [52] further developed an adaptive thresholding approach to optimally extracting monitoring statistics from wavelet coefficients of functional profiles. Wavelet coefficients that significantly deviate from target values were selected to construct the $T^2$ chart, which is shown to be more sensitive to small changes in the process. Apley [4] provided an in-depth discussion of nonparametric profile monitoring with mixed-effect models and emphasized on the importance of both false alarms and root cause analysis. Paynabar and Jin [6] developed a mixed-

effect model and used wavelet decomposition to identify within- and between-profile variations for process monitoring. Chen and Yang [53] proposed a heterogeneous recurrence control chart for online monitoring in nonlinear dynamic processes.

### *B. Image Profiles*

However, existing SPC methods and tools focus on key product characteristics, linear and nonlinear profiles instead of image profiles that are nonlinear and nonstationary. In the past few years, the increasing availability of time-varying image profiles has attracted growing interests on image-based SPC. For example, Du *et al.* [17] analyzed hyperspectral images of poultry carcasses. A spectral band selection approach was developed to extract features for the detection of skin tumors. Wilcox *et al.* [18] also investigated how to optimally select a subset of spectral bands from hyperspectral images for the automatic monitoring of animal feed quality. Megahed *et al.* [20] compared the manufactured tiles with a specific pattern and designed a generalized likelihood ratio chart to detect the non-conforming tiles. Park *et al.* [44] investigated microscopic images of nanoparticle dynamics. Morphology of nanoparticles was characterized by a multistage procedure and then semi-automatically classify them into homogeneous groups. Yan *et al.* proposed to integrate low-rank tensor decomposition with multivariate control charts for image-based process monitoring [19]. In addition, Zhang *et al*. [16] measured the variations of wafer thickness from image profiles using an additive Gaussian process model.

However, most previous studies focus on fault detection from snap-shot images in discrete-part manufacturing. Very little has been done to model the stream of time-varying image profiles that are highly nonlinear and nonstationary. It may be noted that faults or defects are stationary in snap-shot images, but nonstationary in the in-situ stream of image data. There is an urgent need to design and develop new SPC methods and tools that can extract useful information from in-situ image profiles as well as exploit the acquired information for process monitoring and quality control.

## 3.3. Research Methodology

As shown in Fig. 3-2, this paper presents a novel dynamic network scheme to represent and model time-varying image streams. Notably, we introduce an approach of Potts model Hamiltonian to detect community patterns and characterize organizational behaviors in the dynamic network. Finally, a new network generalized likelihood ratio chart is developed to characterize variations in sub-graph structures, which, in turn, facilitates change-point detection in image streams.

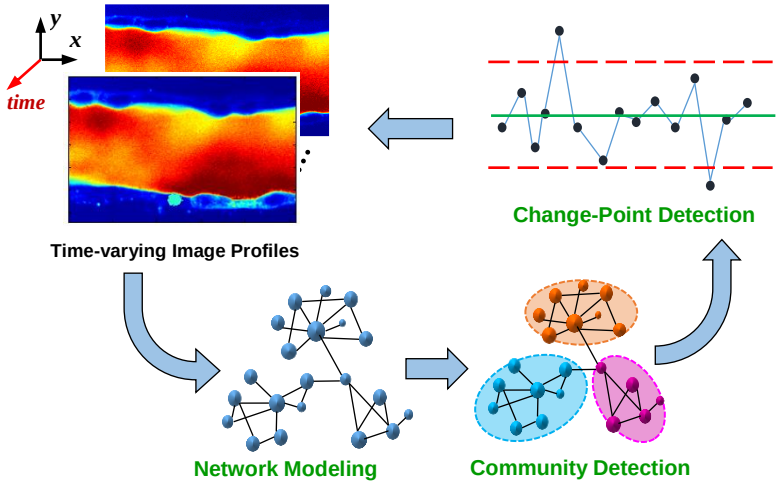



Figure 3-2. Flow chart of the research methodology.

### *A. Network Representation of Image Profiles*

In the literature, both unweighted and weighted networks are commonly used to represent and model image profiles. The unweighted network assigns a binary variable (i.e., 0 and 1) to describe the connectivity between each pair of nodes. In contrast, weighted networks not only contain the information of connectivity as in unweighted networks, but also characterize how close each pair of nodes are connected by assigning a weight. A weighted network can be represented as $G = (V, E, W)$, where $V$, $E$, and $W$ denote the set of network nodes, edges and weights, respectively. In the present study, we extend further propose a new regularized formulation that uses the radial basis functions to calculate the weight $W_{ij}$. Let $\boldsymbol{x}_i$ denotes the intensity of the pixel $i$ and $\boldsymbol{s}_i$ represents its spatial location. The exponential kernel function of $W_{ij}$ is formulated based on the intensity difference of pixels $i$ and $j$ as well as the spatial distance between them:

$$W_{i,j} = \exp\left(\frac{\alpha \left\|\boldsymbol{x}_i - \boldsymbol{x}_j\right\| + (1-\alpha)\left\|\boldsymbol{s}_i - \boldsymbol{s}_j\right\|}{2 \times l^2}\right) \tag{3.1}$$

where $l$ denotes the bandwidth of the exponential kernel. The regularization parameter $\alpha \in [0,1]$ balances between intensity differences $\left\|\boldsymbol{x}_i - \boldsymbol{x}_j\right\|$ and spatial distances $\left\|\boldsymbol{s}_i - \boldsymbol{s}_j\right\|$. When α is larger, the weight $W_{ij}$ primarily focuses on intensity differences among pixels. Nodes *i* and *j* will be assigned a smaller weight if their corresponding pixels share similar intensity values. When α is smaller, the weight $W_{ij}$ focuses more on distances between spatial locations of pixels. Hence, $W_{ij}$ is smaller if corresponding pixels of nodes *i* and *j* are located closer to each other. Notably, Eq. (3.1) can be readily used to cast different types of images (e.g., gray-scale and RGB) into networks. In gray-scale image, the intensity of pixel $x_i$ is denoted by a scalar within the range [0, 255]. In RGB image, the intensity of pixel $\boldsymbol{x}_i$ is represented by a vector with 3 elements of primary colors, i.e., red, green, and blue.

### *B. Network Community Modeling and Characterization*

To this end, image profiles are represented as networks, with nodes correspond to pixels and weights preserve the relationship among pixels. Next, it is critical to detect sub-graph structures from the network, i.e., network communities. A community is defined as a cluster of tightly connected network nodes. The internal connectivity of network nodes within a community is higher than their external connectivity to the remainder of the network. As the network preserves the relationship among image pixels, each community correspondingly represents a group of pixels that form a specific pattern in the image. Detection of network communities provides a great opportunity to uncover organizational structures of the network for the study of interacting objects and the relations between them. In the literature, a number of approaches were reported for the detection of network communities. For example, Newman [54] calculated the network modularity to quantify the connectivity among network nodes. By maximizing the modularity, densely connected nodes tend to be clustered into the same community. Furthermore, Mankad and Michailidis [55] applied the non-negative matrix factorization to characterize community structures from a sequence of networks. By adding sparse and smooth constraints, the algorithm is scalable to large networks and accommodates sudden changes in the network topology. In addition, prior research extended the network community detection to image segmentation [56].

However, most previous studies focused on networks (e.g., social network [54] and biological network [57]) and did not specifically consider the stream of imaging data. Little work has been done on the community detection of dynamic image streams for the purpose of process quality control.

In this study, we introduce the Potts model Hamiltonian [56] to detect communities in the dynamic network for the modeling and analysis of time-varying image profiles. First, a "Potts type" variable $g_i$ is assigned to each network node to label the index of its community. That is, $g_i = 1$ when node $i$ belongs to the 1st community and $g_i \in [1, K]$ for the network with $K$ communities. As such, the network can be characterized by the state of "Potts type" variables of $N$ nodes, i.e., $\{g_i\}_{i=1,\dots,N}$. The network energy is then characterized by the Potts model Hamiltonian function as:

$$\mathcal{H}(\{g_i\}_{i=1}^{N}) = \frac{1}{2}\sum_{k=1}^{K}\sum_{i,j\in C_k} (W_{ij} - \bar{W})[\Theta(\bar{W} - W_{ij}) + \gamma\Theta(W_{ij} - \bar{W})]\delta(g_i, g_j) \quad (3.2)$$

Here, $i, j \in C_k$ indicate nodes $i$ and $j$ belong to the same community $C_k$. $\Theta(\cdot)$ is the Heaviside function, whose value is 0 for negative argument and 1 for positive argument. $\bar{W}$ denotes the average of all weights and $\delta(g_i, g_j)$ is the Kronecker delta with $\delta(g_i, g_j) = 1$ if $g_i = g_j$. Notably, assigning node $i$ and $j$ into the same community will increase the network Hamiltonian when the weight $W_{ij}$ is small (i.e., $W_{ij} > \bar{W}$). Hence, a regularization parameter $\gamma$ is introduced to penalize the increase of network Hamiltonian. A larger $\gamma$ will result in a more significant increase of network community when node $i$ and $j$ with a large weight $W_{ij}$ are assigned into the same community.

It may be noted from Eq. (3.2) that if nodes $i$ and $j$ share a small weight (i.e., $W_{ij} < \bar{W}$), then $\Theta(W_{ij} - \bar{W}) = 0$. Assigning nodes $i$ and $j$ into the same community (i.e., having $g_i = g_j$) will decrease Hamiltonian $\mathcal{H}$ by $\Delta\mathcal{H} = (W_{ij} - \bar{W})$. However, if node $i$ and $j$ are associated with a large weight (i.e., $W_{ij} > \bar{W}$), then $\Theta(W_{ij} - \bar{W}) = 1$. A penalized increase of $\Delta\mathcal{H} = \gamma(W_{ij} - \bar{W})$ will incur if node $i$ and $j$ are assigned to the same community (i.e., $g_i = g_j$). Therefore, minimizing the Potts model Hamiltonian, i.e., $\min_{g_i,g_j}\mathcal{H}(\{g_i\}_{i=1}^{N}), s.t.\ 1 \le g_i \le K$, will reveal the underlying structure of network communities. The optimization of Potts model Hamiltonian

aggregates densely connected nodes by assigning the same value to their Potts type variables, whereas separates loosely related nodes into different communities.

However, the objective function of Potts model Hamiltonian is NP-hard to be minimized exactly. Thus, a heuristic algorithm is introduced to extract network communities that approximately optimizes the Potts model Hamiltonian in Eq. (3.2). As shown in Table 3-1, the order of nodes are first randomly permuted and each node is considered as an individual community. Further, we assign each node to its neighboring communities and calculate the change of Hamiltonian $\Delta\mathcal{H}$. The assignment that decreases $\mathcal{H}$ to the largest extent will be retained. Here, a neighboring community of one node is the community that contains nodes with stronger connections (e.g., smaller weights) to that node. For example, all nodes $j$ that have weights $W_{ij}$ less than a user-defined threshold are considered to be neighbors of node $i$ and communities that nodes $j$ belong to are neighboring communities of node $i$. Assigning node $i$ to its neighboring communities will reduce the computation time when the number of nodes $N$ is large. In addition, we iterate through all the $N$ nodes to achieve the convergence of Hamiltonian $\mathcal{H}$. It is noteworthy that the Potts model Hamiltonian is an unsupervised approach and the number of communities $K$ does not need to be fixed beforehand.

Table 3-1. The Potts Model Hamiltonian Algorithm for Network Community Detection

| | |
|---|---|
| 1: | $G$ - network, $\mathcal{H}$ - network Hamiltonian, $g_i$ – community label |
| 2: | **repeat** |
| 3: | Random permutation of the order of nodes |
| 4: | Treat each node as a community |
| 5: | **while** $\mathcal{H}$ does not converge |
| 6: | **for all** node *i* in *G* **do** |
| | Assign node *i* to its neighboring communities |
| 7: | Calculate $\Delta\mathcal{H}$ for each assignment |
| | Choose the assignment which maximally decreases $\mathcal{H}$ |
| 8: | **end for** |
| 9: | **end while** |
| 10: | **If** $\mathcal{H}$ is smaller than that obtained from last replication |
| 11: | Keep the newly obtained $\{g_i\}_{i=1,\ldots,N}$ |
| 12: | **else** |
| 13: | Terminate |
| 14: | **end if** |
| 15: | **until** min($\mathcal{H}$) is achieved |

Fig. 3-3 shows community detection results on benchmark images using Potts model Hamiltonian. Original images from the Berkeley BSDS500 dataset [58] are shown in Fig. 3-3a. Notably, a smaller number of communities separates the main object from the background, e.g., the plane and building are depicted by two communities in Fig. 3-3b. On the contrary, a larger number of communities delineates detailed information of the original image. For example, ten communities not only characterize detailed patterns on the plane, but also depict the morphology of clouds (see the top figure in Fig. 3-3c).

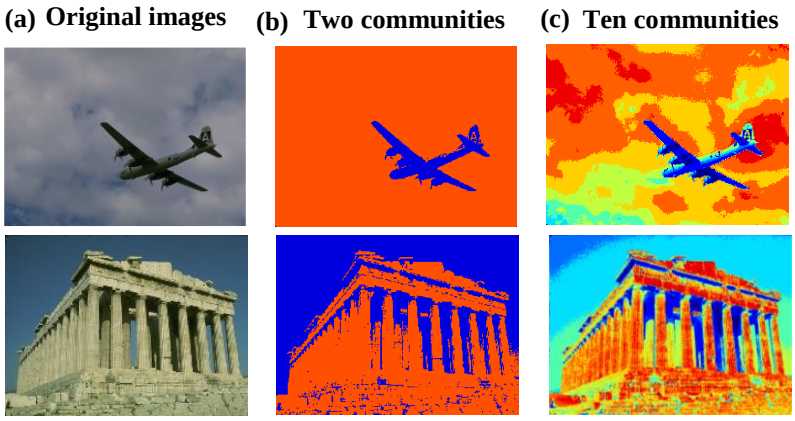



Figure 3-3. Examples of community detection using the Hamiltonian algorithm. (a) original images; (b) two communities; (c) ten communities.

## C. *Multivariate Monitoring of Network Statistics*

As the Hamiltonian approach favors strongly connected nodes into one community, image pixels within each community are homogeneous. Therefore, pixel intensities within the $k^{th}$ community of the image profile $I^{(i)}$ approximately follows a normal distribution $N\left(\mu_k^{(i)}, {\sigma_k^{(i)}}^2\right)$ with mean $\mu_k^{(i)}$ and variance ${\sigma_k^{(i)}}^2$. Further, we extracted a Q-variate vector of community statistics for the $i^{th}$ image profile as:

$$\boldsymbol{y}^{(i)} = \left(\bar{x}_1^{(i)}, \ldots, \bar{x}_K^{(i)}, s_1^{(i)}, \ldots, s_K^{(i)}, n_1^{(i)}, \ldots, n_K^{(i)}\right)^T \tag{3.3}$$

where $i = 1, 2, \ldots, m$ is the time index of $m$ image profiles, $\bar{x}_k^{(i)}$ is the sample mean of pixel intensities, $s_k^{(i)}$ is the sample standard deviation, and $n_k^{(i)}$ is the number of pixels in the $k^{th}$ community. The dimensionality of the community feature vector is $Q = 3 \times K$.

*Proposition 1*: If the process is in control, $\boldsymbol{y}^{(i)}$ is approximately a multivariate normal distribution as $\boldsymbol{y}^{(i)} \sim MVN(\boldsymbol{\mu}_0, \boldsymbol{\Sigma})$, where $\boldsymbol{\mu}_0$ and $\boldsymbol{\Sigma}$ are in-control mean and covariance matrix of feature vectors.

The proof of *Proposition 1* is given in Appendix A.

The variations of community statistics distinctly mark the emergent properties of anomalies in the stream of images. If process shift occurs at the $\tau^{th}$ image, the process mean is shifted to $\boldsymbol{\mu}_1$ and $\boldsymbol{y}^{(i)} \sim MVN(\boldsymbol{\mu}_1, \boldsymbol{\Sigma})$. As a result, the hypothesis of change-point detection is formulated as:

$$\begin{cases} H_0: E[\boldsymbol{y}^{(i)}]_{i=1}^{\tau} = E[\boldsymbol{y}^{(i)}]_{i=\tau+1}^{m} = \boldsymbol{\mu}_0 \\ H_1: E[\boldsymbol{y}^{(i)}]_{i=1}^{\tau} \neq E[\boldsymbol{y}^{(i)}]_{i=\tau+1}^{m} = \boldsymbol{\mu}_1 \end{cases} \tag{3.4}$$

Furthermore, we propose a NGLR chart to monitor the multivariate vector of network community statistics. For a stream of *m* image profiles, we have the data sample of ($\boldsymbol{y}^{(1)}, \boldsymbol{y}^{(2)}$, …, $\boldsymbol{y}^{(i)}$, …, $\boldsymbol{y}^{(m)}$). Under the null hypothesis $H_0$, all *m* feature vectors are from in-control distribution and the likelihood function is:

$$L(\boldsymbol{\mu}_0) = \prod_{i=1}^{m} \frac{1}{(2\pi)^{\frac{Q}{2}} |\boldsymbol{\Sigma}|^{\frac{1}{2}}} \exp\left( -\frac{1}{2} \left(\boldsymbol{y}^{(i)} - \boldsymbol{\mu}_0\right)^T \boldsymbol{\Sigma}^{-1} \left(\boldsymbol{y}^{(i)} - \boldsymbol{\mu}_0\right) \right) \tag{3.5}$$

Under the alternative hypothesis $H_1$, a process shift occurs at the $\tau^{th}$ image. The likelihood function becomes:

$$\begin{aligned} L(\boldsymbol{\mu}_0, \boldsymbol{\mu}_1) = \left( \frac{1}{(2\pi)^{\frac{Q}{2}} |\boldsymbol{\Sigma}|^{\frac{1}{2}}} \right)^m \times \prod_{i=1}^{\tau} \exp\left( -\frac{1}{2} \left(\boldsymbol{y}^{(i)} - \boldsymbol{\mu}_0\right)^T \boldsymbol{\Sigma}^{-1} \left(\boldsymbol{y}^{(i)} - \boldsymbol{\mu}_0\right) \right) \\ \times \prod_{i=\tau+1}^{m} \exp\left( -\frac{1}{2} \left(\boldsymbol{y}^{(i)} - \boldsymbol{\mu}_1\right)^T \boldsymbol{\Sigma}^{-1} \left(\boldsymbol{y}^{(i)} - \boldsymbol{\mu}_1\right) \right) \end{aligned} \tag{3.6}$$

The process mean after the shift at the $\tau^{th}$ image profile is estimated as:

$$\widehat{\boldsymbol{\mu}}_{m,\tau,\boldsymbol{y}} = \frac{1}{(m-\tau)} \sum_{i=\tau+1}^{m} \boldsymbol{y}^{(i)} \tag{3.7}$$

The maximum log likelihood ratio statistic is defined as:

$$R_{m,\boldsymbol{y}} = log\frac{L(\boldsymbol{\mu}_0, \boldsymbol{\mu}_1)}{L(\boldsymbol{\mu}_0)} = log\frac{\max_{0\le\tau<m} L\left(\boldsymbol{\mu}_0, \widehat{\boldsymbol{\mu}}_{m,\tau,\boldsymbol{y}}\right)}{L(\boldsymbol{\mu}_0)} \tag{3.8}$$

Similar to the results of Wang and Reynolds [59], the multivariate NGLR statistic at the $m^{th}$ image is computed as:

$$R_{m,\boldsymbol{y}} = \max_{0\le\tau<m} \frac{m-\tau}{2}\left(\widehat{\boldsymbol{\mu}}_{m,\tau,\boldsymbol{y}} - \boldsymbol{\mu}_0\right)^T \boldsymbol{\Sigma}^{-1}(\widehat{\boldsymbol{\mu}}_{m,\tau,\boldsymbol{y}} - \boldsymbol{\mu}_0) \tag{3.9}$$

If population mean $\boldsymbol{\mu}_0$ and covariance $\boldsymbol{\Sigma}$ are replaced with the sample mean $\overline{\boldsymbol{y}}$ and sample covariance $\boldsymbol{S}$ estimated from the in-control data, then the NGLR statistic in Eq. (3.9) becomes:

$$R_{m,\boldsymbol{y}} = \max_{0\le\tau<m} \frac{m-\tau}{2}\left(\widehat{\boldsymbol{\mu}}_{m,\tau,\boldsymbol{y}} - \overline{\boldsymbol{y}}\right)^T \boldsymbol{S}^{-1}(\widehat{\boldsymbol{\mu}}_{m,\tau,\boldsymbol{y}} - \overline{\boldsymbol{y}}) \tag{3.10}$$

However, the computation of NGLR statistic will be expensive if we take account of all the historical image profiles (i.e., $m$ is large). To facilitate the online monitoring, it is common to consider images within a sliding window, $m - w \le \tau < m$, as opposed to $0 \le \tau < m$. Therefore, the computation of online NGLR statistic becomes:

$$R_{m,\boldsymbol{y}} = \max_{m-w\le\tau<m} \frac{m-\tau}{2}\left(\widehat{\boldsymbol{\mu}}_{m,\tau,\boldsymbol{y}} - \overline{\boldsymbol{y}}\right)^T \boldsymbol{S}^{-1}(\widehat{\boldsymbol{\mu}}_{m,\tau,\boldsymbol{y}} - \overline{\boldsymbol{y}}) \tag{3.11}$$

As shown in the top portion of Fig. 3-4, each solid bar represents a $Q$-variate vector $\boldsymbol{y}^{(i)}$ and the square box is a sliding window of size $w$ that contains feature vectors from $\boldsymbol{y}^{(m-w)}$ to $\boldsymbol{y}^{(m-1)}$. According to Eq. (3.7), $\widehat{\boldsymbol{\mu}}_{m,\tau,\boldsymbol{y}}$ is process mean calculated from $\boldsymbol{y}^{(\tau+1)}$ to $\boldsymbol{y}^{(m)}$. Eq. (3.11) shows that the computation is repeated in the sliding window for $\tau$ values from $m - w$ to $m$. Finally, $R_{m,\boldsymbol{y}}$ is the maximal statistic obtained by varying $\tau$ values within the current window.

However, there are technical difficulties in the computation of NGLR statistics $R_{m,\boldsymbol{y}}$. Note that Eq. (3.11) involves the inversion of sample covariance matrix $\boldsymbol{S}$. If two communities share the same distribution in pixel intensities, then $\boldsymbol{S}$ will be singular that makes the NGLR computation impractical. For example, two communities in Fig. 3-3 may be farther apart but correspond to the same object (i.e., the cloud). As such, these two communities have the same distribution of pixel intensities. The $Q$-variate vector $\boldsymbol{y}^{(i)}$ will have features that are very close to each other for these

two communities, thereby leading to a singular covariance matrix. On the other hand, the feature vector $\boldsymbol{y}^{(i)}$ will pose the "curse of dimensionality" issue for online monitoring if an image is segmented into a large number of communities (i.e., $K$ is large). It is necessary to decompose the high-dimensional feature vector with complex correlation structure into a set of uncorrelated principal components.

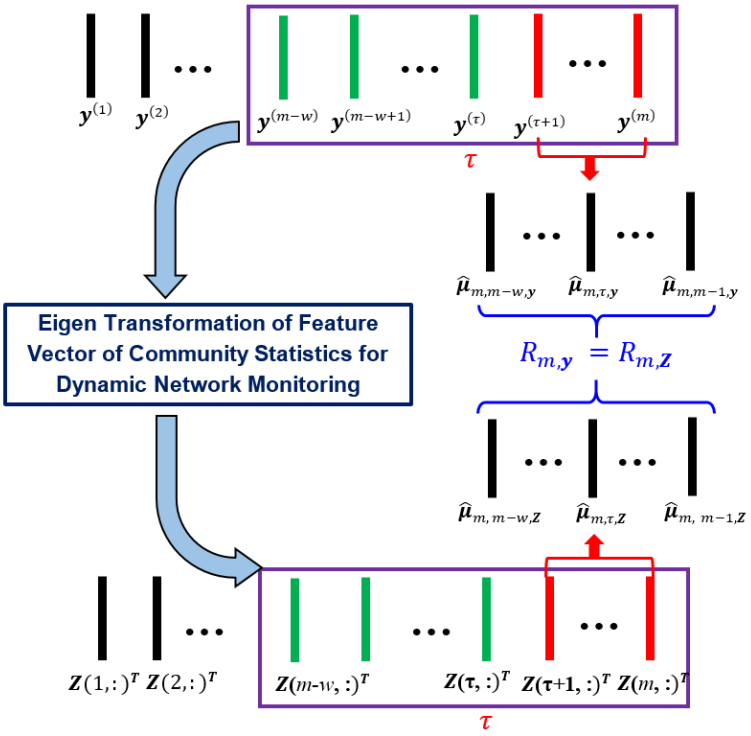



Figure 3-4. Multivariate network GLR statistics in the eigenspace.

To address the aforementioned challenges, we propose to further transform the $Q$-variate vectors $\boldsymbol{y}^{(i)}$, $i = 1,2,\dots,m$ to uncorrelated principal components $\mathbf{Z}(i,:)^T$, $i = 1,2,\dots,m$. Then, the projected NGLR statistics $R_{m,\mathbf{Z}}$ in a sliding window is calculated from $\mathbf{Z}(m-w,:)^T$ to $\mathbf{Z}(m-1,:)^T$ (see Fig. 3-4).

*Proposition 2*: In the eigenspace, multivariate NGLR statistic $R_{m,\boldsymbol{y}}$ in Eq. (3.11) is equivalently computed as:

$$R_{m,\mathbf{Z}} = \max_{m-w\le\tau<m} \frac{m-\tau}{2} \left(\widehat{\boldsymbol{\mu}}_{m,\tau,\mathbf{Z}}\right)^T \boldsymbol{S}_{\boldsymbol{z}}{}^{-1} \left(\widehat{\boldsymbol{\mu}}_{m,\tau,\mathbf{Z}}\right) \tag{3.12}$$

where

$$\widehat{\boldsymbol{\mu}}_{m,\tau,\mathbf{Z}} = \frac{1}{(m-\tau)} \sum_{i=\tau+1}^{m} \mathbf{Z}(i,:)^T \tag{3.13}$$

and $\mathbf{Z}(i,:)^T$, $i = 1,2,\ldots,m$ are principal components obtained by eigen transformation of feature vector $\boldsymbol{y}^{(i)}$.

The proof of *Proposition 2* in Appendix B.

Furthermore, a truncated $m$-by-$q$ score matrix $\tilde{\mathbf{Z}}$ can be obtained by considering only the first $d$ eigen values and eigenvectors ($q < Q$), which explains the majority of variations in the feature matrix **Y**. Hence, the NGLR statistic $R_{m,\tilde{\mathbf{Z}}}$ in the reduced dimension *d* becomes:

$$\widehat{\boldsymbol{\mu}}_{m,\tau,\tilde{\mathbf{Z}}} = \frac{1}{(m-\tau)} \sum_{i=\tau+1}^{m} \mathbf{Z}(i,1:q)^T \tag{3.14}$$

The NGLR control chart signals when the statistic $R_{m,\mathbf{Z}}$ falls beyond the control limits. In order to achieve a desired in-control performance, the control limit is obtained by simulating large number of in-control profiles and then estimate the distribution of NGLR statistics. Average run length (ALR) is also used to measure in-control performance (type I error) of control limits. Case studies to evaluate performances of our developed NGLR algorithm will be described in Section 3.4.

## 3.4. Experiments and Results

In this present paper, we implement and evaluate the proposed methodology for real-world applications in both biomanufacturing process and UPM. The laser-scanning video microscopy captures in-vivo images of living cells during the synthesis of biopharmaceutical products. Also, advanced surface metrology brings the increasing availability of UPM images, whose surface characteristics are pertinent to setup and conditions of cutting tools. Our experiments and results will be detailed in the following sections.

### *A. Case Study in the Biomanufacturing Process*

The microscopic image of a living cell is shown in Fig. 3-5. Cytoplasm is shown in red and yellow, and extracellular fluids are in dark blue. A vesicle is inserted into the cell, which carries a specific type of transgene to change the cell phenotype. As such, genetically engineered cells are

produced. In synthetic biology, the cell works as a "factory" to produce biopharmaceutical products for therapeutics. Biologists developed molecular approaches that regulate genes to synthesize cells with specific properties. For example, Chinese Hamster Ovary cells are engineered to resemble a human-like ion channels for scientific experiments. Notably, our previous studies used the genetically engineered Chinese Hamster Ovary cells to investigate whether and how reduced glycosylation alter the kinetics of human potassium ion channels [60, 61].

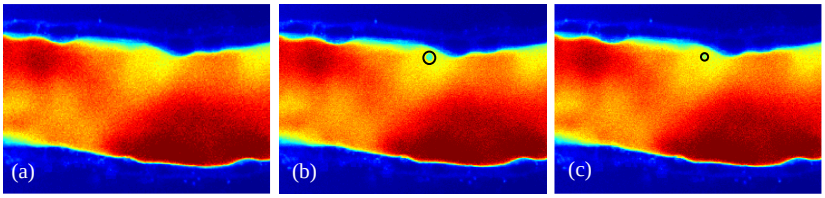



Figure 3-5. (a) Real-world image of a cell in the biomanufacturing process; (b) Out-of-control image with a vesicle (size r = 6); (c) Out-of-control image with a vesicle (size r = 4).

The quality of genetically engineered cells has a direct impact on the success and replicability of biomedical experiments. A high-resolution video microscope is commonly utilized to monitor and control the biomanufacturing process for quality assurance in the cellular level. As opposed to static images (e.g., wafer profiles in semiconductor manufacturing), video microscopy brings a stream of dynamic images that describe time-varying activities of living cells. Dynamic imaging profiles pose significant challenges in the modeling and analysis of biomanufacturing process. Little work has been done to develop SPC methods for monitoring high-dimensional dynamic imaging profiles. In this present study, we collected a stream of 100 images from in-control processes. Each image is with the size of 160 × 256 pixels. In addition, 50 out-of-control imaging profiles with inserted vesicles were collected. In each experimental scenario, the vesicle keeps the same shift (intensity and size). Table 3-2 shows the total of 30 scenarios. Together, these 30 scenarios capture various types of dynamic properties of vesicles. These images serve as the basis for computer experiments to evaluate the performance of proposed dynamic network methodology.

*1) Dynamic network for adaptive community detection*

We implement the developed Hamiltonian approach for community detection of 100 in-control images. Parameters $\gamma$ and $\overline{W}$ in Eq. (3.2) are fixed at 3 and 12 so as to achieve 15 communities for each image. As shown in Fig. 3-6, each image is segmented into 15 adaptive communities based on the patterns and sub-graph structures of network models. As a result, image pixels are self-organized

into 15 adaptive region of interests (ROIs), each is with a different color. As aforementioned, we assumed that pixel intensities within the same community follow a normal distribution with mean $\mu_k$ and variance $\sigma_k^2$. Here, normal probability plot is employed to graphically validate this assumption. Fig. 3-6b shows the normal probability plot for the pixel intensities of the fifteenth community $C_{15}$ in Fig. 3-6a. The approximately straight line indicates the validity of normal assumption.

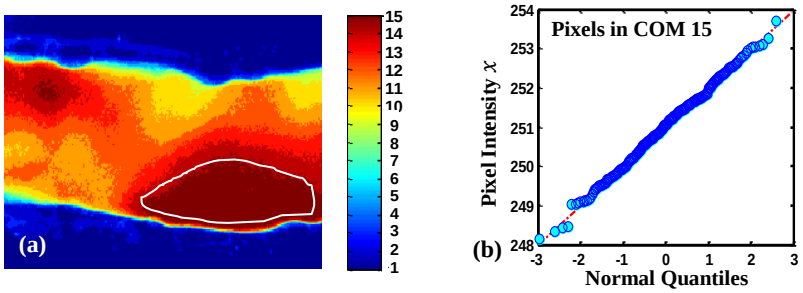



Figure 3-6. (a) Community detection result of a cellular image (15 communities), (b) Normal probability plot of pixel intensities in the 15th community.

Further, we extracted a *Q*-variate vector of community statistics from the $i^{th}$ image profile, $\boldsymbol{y}^{(i)} = \left(\bar{x}_1^{(i)}, \ldots, \bar{x}_K^{(i)}, s_1^{(i)}, \ldots, s_K^{(i)}, n_1^{(i)}, \ldots, n_K^{(i)}\right)^T$, where $i = 1, 2, \ldots, m$ is the time index in the stream of *m* image profiles. For the in-control process, we assume that $\boldsymbol{y}^{(i)}$ follows the multivariate normal distribution. Therefore, we validate the assumption by the normal probability plotting of the image-to-image mean intensity $\bar{x}_k^{(i)}$, standard deviation $s_k^{(i)}$, number of pixels $n_k^{(i)}$ in each community $C_k$. As shown from Fig. 3-7a to 3-7c, sorted values of image-to-image mean, standard deviation and number of pixels are plotted against normal quantiles, respectively. It may be noted that each normal probability plot follows an approximately straight line, which validates normal assumptions of image-to-image community statistics.

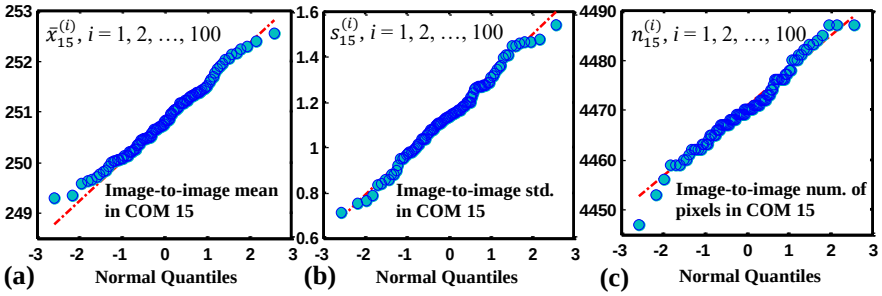



Figure 3-7. Normal probability plots of image-to-image (a) mean intensities (b) standard deviation of intensities (c) number of pixels in the 15th community for a stream of dynamic images.

*2) Performance comparison with traditional square-shaped Region of Interests (ROI) approaches*

It may be noted that the square-shaped ROI is commonly used to inspect the uniformity within an image or compare to a specific pattern. In order to detect the change point of process mean shift, the assumption is that in-control distribution of image-to-image mean within the square ROI is approximately normal. However, this assumption is not always valid for time-varying imaging profiles. For example, Fig. 3-8a shows a square-shaped ROI in the same image as Fig. 3-6a. Pixels within the square ROI are not homogeneous and their intensities are not normally distributed. Fig. 3-8b shows the normal probability plot for pixel intensities of the ROI in Fig. 3-8a. The normal assumption is invalid because the data distribution does not follow a straight line in the plot.

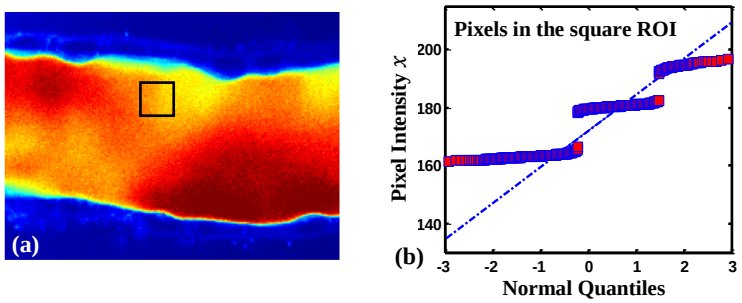



Figure 3-8. (a) Square-shaped ROI in the cellular image, (b) Normal probability plot of pixel intensities within the square ROI.

If we zoom into the square ROI in Fig. 3-8a, Fig. 3-9a shows the intensity values of pixels vary significantly from image to image. We have further extracted the same *Q*-variate vector $\boldsymbol{y}^{(i)}$ from the square ROI (see Fig. 3-8a), where $i = 1, 2, \ldots, m$ is the time index in the stream of $m$ image profiles. Fig. 3-9b to 3-9d show normal probability plotting of the image-to-image mean intensity $\bar{x}_k^{(i)}$, standard deviation $s_k^{(i)}$, number of pixels $n_k^{(i)}$ within the square ROI. It may be noted that normal probability plots of $\bar{x}_k^{(i)}$ and $s_k^{(i)}$ are highly skewed, which indicate the invalidity of normal assumption. However, the normal probability plot of $n_k^{(i)}$ shows a strictly horizontal line, because the number of pixels is constant when the size of square ROI is fixed. It is worth mentioning that each community in Fig. 3-6a, however, does have a changing number of pixels because of adaptive ROIs.

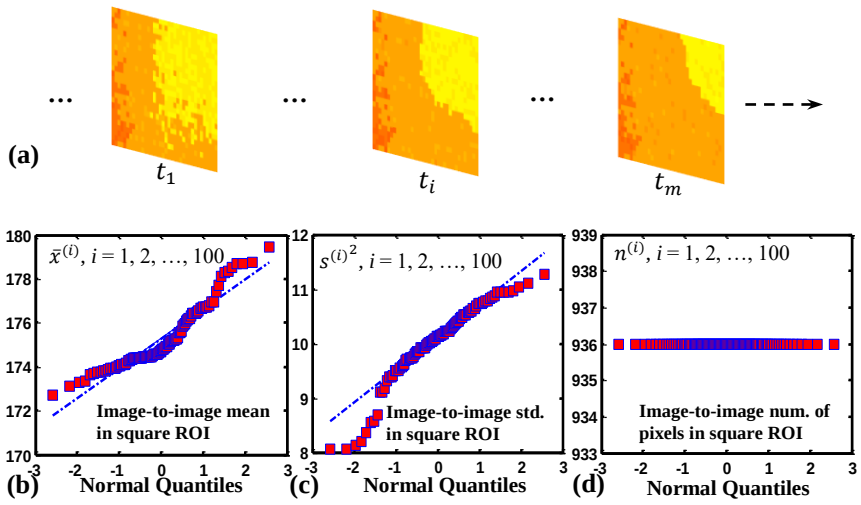



Figure 3-9. (a) Pixel intensity variations in the square ROI with respect to time, and normal probability plots of image-to-image (b) mean intensities (c) standard deviation of intensities (d) number of pixels within the square ROI for a stream of dynamic image profiles.

Therefore, square ROIs are not well suited for process monitoring of time-varying images because the normal assumption within the square ROI is violated. Fig. 3-10 shows the control chart obtained with the monitoring algorithm using the square ROI. Note that GLR chart signals out of control even when the process is in control. If the images are static and the monitoring objective is to inspect the uniformity or comparison to a specific pattern, then pixel intensities within each square ROI vary slightly from one image to another. Hence, image-to-image mean intensity is approximately normally distributed. When a large number of pixels are contained in the square ROI, the normality assumption is guaranteed by the central limit theorem. However, time-varying dynamic images bring significant challenges for monitoring algorithms that use square ROIs. Dynamic network models provide a new means to characterize and detect adaptive ROIs in the images for process monitoring.

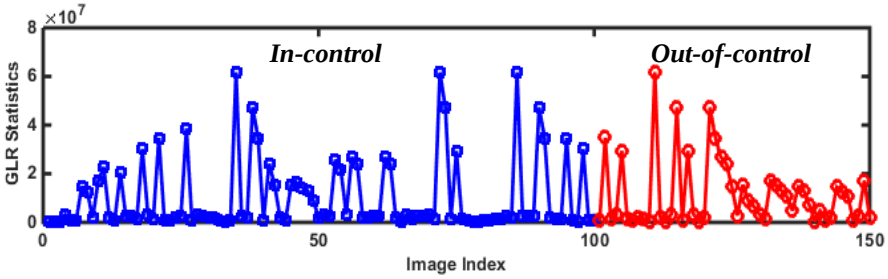



Figure 3-10. The performance of GLR control chart with square-shaped ROIs for the monitoring of biomanufacturing process.

*3) Performance comparison between NGLR and $T^2$ charts*

Furthermore, we conducted computer experiments to compare the performance of network GLR and Hotelling $T^2$ charts for biomanufacturing process monitoring. It is worth mentioning that either network GLR or $T^2$ chart is calculated from the Q-variate vector of network community statistics. $T^2$ statistics are calculated based on features extracted from network modeling and characterization of image data. Note that the first 100 images are in-control profiles of cell dynamics. A vesicle with transgenes is inserted into the cell after the 100th image ($\tau$ = 100). We varied the intensity shift ($\Delta x$) and the radius ($r$) of inserted vesicles to evaluate the performances of the developed network GLR algorithms. In total, there are 30 scenarios including 3 levels of radius and 10 levels of intensity shifts. Average run length is used as a metric to evaluate the performance of monitoring schemes. ARL is the average period after which the process-monitoring scheme first signals the out-of-control image profile. In this study, we focused on Phase II out-of-control ARLs (i.e., $ARL_1$). Notably, $ARL_1$s are calculated by randomly replicating each scenario for 100 times.

For the $i^{th}$ image profile, the Hotelling $T^2$ statistic is computed using the $Q$-variate vector of network community statistics $\boldsymbol{y}^{(i)}$ as:

$$T^2(i) = (\boldsymbol{y}^{(i)} - \bar{\boldsymbol{y}})^T \boldsymbol{S}^{-1} (\boldsymbol{y}^{(i)} - \bar{\boldsymbol{y}}) \tag{3.15}$$

In order to address the singular issue of the covariance matrix $\boldsymbol{S}$, we computed the $T^2$ statistic in Eq. (3.15) using the set of principal components as in our previous study [53]:

$$T^2(i) = \left(\boldsymbol{y}^{(i)} - \bar{\boldsymbol{y}}\right)^T \boldsymbol{S}^{-1} \left(\boldsymbol{y}^{(i)} - \bar{\boldsymbol{y}}\right) = \sum_{r=1}^{Q} \frac{\mathbf{Z}(i,r)^2}{\lambda_r^2} \tag{3.16}$$

In addition, we considered the first $q$ principal components ($q < Q$) to obtain the Hotelling $T^2$ statistic in the reduced dimension $q$ as:

$$T^2(i) = \sum_{r=1}^{q} \frac{\mathbf{Z}(i,r)^2}{\lambda_r^2} \tag{3.17}$$

This reduced-dimension $T^2$ statistic will be further used in the next section to benchmark the performance of network GLR chart in the reduced dimension. Notably, the upper control limit (*UCL*) of Hotelling $T^2$control chart is estimated as:

$$UCL_{T^2} = \frac{q(M+1)(M-1)}{M^2 - Mq} F_{\alpha,q,M-q} \tag{3.18}$$

where $q$ is the number of features (i.e., the dimensionality of $\boldsymbol{y}^{(i)}$), $M$ is the number of imaging profiles, $F_{\alpha,q,M-q}$ is the upper $100\alpha\%$ of critical point of $F$ distribution with $q$ and $M-q$ degree of freedoms. The significance level $\alpha$ is set as 0.05.

Table 3-2 shows the results of performance comparison between NGLR and Hotelling $T^2$ charts for 30 scenarios of varying size ($r$) and intensity shift ($\Delta x$) of the vesicle. Experiments were replicated 100 times for each scenario. When the intensity shift is large, $ARL_1$s of both charts reach 1.00. In other words, one sample is adequate on average for the monitoring scheme to first signal the vesicle. For instance, both charts have $ARL_1$ values of 1.00 when $|\Delta x|$ = 10, which means the vesicle is immediately detected when it emerges. When the intensity shift is decreased to $|\Delta x|$ = 7 and 8, the $ARL_1$ begin to increase. Notably, $ARL_1$ is bigger for the scenarios with smaller radius of vesicles. For example, both charts yield $ARL_1$s of 1.00 when $\Delta x$ = 8 and $r$ = 8, while the $ARL_1$ of NGLR chart is increased to 1.63 and the $ARL_1$ of $T^2$ chart reaches 4.88 when $\Delta x$ = 8 and $r$ = 4. When the intensity shift is very small (i.e., $|\Delta x|$ = 5 and 6), NGLR charts signal the out-of-control faster than $T^2$ charts. For example, NGLR chart yields the $ARL_1$ of 3.38 for $\Delta x = 5, r = 4$ and 3.71 for $\Delta x = -5, r = 4$, while $T^2$ chart are 16.69 (for $\Delta x = 5$) and 14.96 (for $\Delta x = -5$).

Table 3-2. Performance Comparison Between Network GLR and Hotelling $T^2$ Charts

| Δ | $r$ | $ARL_1$ | | Distribution of $\Delta_\tau = \hat{\tau} - \tau$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | $\Delta_\tau = 0$ | | $0 < \Delta_\tau \leq 3$ | | $3 < \Delta_\tau \leq 10$ | | $\Delta_\tau > 10$ | |
| | | *NGLR* | $T^2$ | *NGLR* | $T^2$ | *NGLR* | $T^2$ | *NGLR* | $T^2$ | *NGLR* | $T^2$ |
| -10 | 4 | 1.00 | 1.00 | 100 | 100 | 0 | 0 | 0 | 0 | 0 | 0 |
| | 6 | 1.00 | 1.00 | 100 | 100 | 0 | 0 | 0 | 0 | 0 | 0 |
| | 8 | 1.00 | 1.00 | 100 | 100 | 0 | 0 | 0 | 0 | 0 | 0 |
| -8 | 4 | 1.24 | 1.90 | 78 | 51 | 22 | 39 | 0 | 10 | 0 | 0 |
| | 6 | 1.01 | 1.05 | 99 | 95 | 1 | 5 | 0 | 0 | 0 | 0 |
| | 8 | 1.00 | 1.00 | 100 | 100 | 0 | 0 | 0 | 0 | 0 | 0 |
| -7 | 4 | 2.18 | 4.44 | 37 | 22 | 49 | 30 | 14 | 40 | 0 | 8 |
| | 6 | 1.10 | 1.60 | 90 | 61 | 10 | 34 | 0 | 5 | 0 | 0 |
| | 8 | 1.00 | 1.02 | 100 | 98 | 0 | 2 | 0 | 0 | 0 | 0 |
| -6 | 4 | 2.56 | 13.30 | 23 | 7 | 56 | 9 | 21 | 32 | 0 | 51 |
| | 6 | 1.41 | 3.44 | 61 | 28 | 39 | 31 | 0 | 41 | 0 | 0 |
| | 8 | 1.06 | 1.23 | 94 | 82 | 6 | 18 | 0 | 0 | 0 | 0 |
| -5 | 4 | 3.71 | 14.96 | 18 | 7 | 39 | 10 | 40 | 29 | 3 | 46 |
| | 6 | 1.98 | 7.37 | 34 | 14 | 60 | 22 | 6 | 40 | 0 | 24 |
| | 8 | 1.30 | 2.36 | 71 | 42 | 29 | 39 | 0 | 17 | 0 | 2 |
| 5 | 4 | 3.38 | 16.69 | 24 | 4 | 41 | 9 | 33 | 25 | 2 | 55 |
| | 6 | 1.82 | 6.00 | 43 | 23 | 54 | 19 | 3 | 42 | 0 | 16 |
| | 8 | 1.25 | 2.04 | 76 | 48 | 24 | 37 | 0 | 15 | 0 | 0 |
| 6 | 4 | 3.01 | 11.11 | 21 | 11 | 46 | 14 | 33 | 33 | 0 | 41 |
| | 6 | 1.47 | 3.24 | 62 | 33 | 38 | 36 | 0 | 30 | 0 | 1 |
| | 8 | 1.03 | 1.15 | 97 | 86 | 3 | 14 | 0 | 0 | 0 | 0 |
| 7 | 4 | 2.18 | 7.23 | 35 | 9 | 53 | 31 | 12 | 33 | 0 | 27 |
| | 6 | 1.13 | 1.51 | 87 | 67 | 13 | 29 | 0 | 4 | 0 | 0 |
| | 8 | 1.00 | 1.00 | 100 | 100 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | 4 | 1.63 | 4.88 | 53 | 19 | 44 | 34 | 3 | 36 | 0 | 11 |
| | 6 | 1.03 | 1.17 | 97 | 85 | 3 | 15 | 0 | 0 | 0 | 0 |
| | 8 | 1.00 | 1.00 | 100 | 100 | 0 | 0 | 0 | 0 | 0 | 0 |
| 10 | 4 | 1.05 | 1.11 | 96 | 90 | 4 | 10 | 0 | 0 | 0 | 0 |
| | 6 | 1.00 | 1.00 | 100 | 100 | 0 | 0 | 0 | 0 | 0 | 0 |
| | 8 | 1.00 | 1.00 | 100 | 100 | 0 | 0 | 0 | 0 | 0 | 0 |

Table 3-2 also shows the distribution of $\Delta_\tau$ for 100 replications, which is the time delay between signaled change point ($\hat{\tau}$) and actual change point ($\tau$). Smaller delay $\Delta_\tau$ indicates better performances in the change-point detection. When the intensity shift is large (i.e.,$|\Delta x|$ = 10), both NGLR and $T^2$ charts yield the time delay $\Delta_\tau = 0$ for 100 replications and only slight deviation to $\Delta_\tau \in (0, 3]$ when $\Delta x = 10$ and $r = 4$. In other words, both charts immediately signal the out-of-control sample. When the intensity shift is decreased to $|\Delta x|$ = 7 and 8, most of replications reside in the intervals $\Delta_\tau \in [0, 3]$ and (3, 10]. Nonetheless, NGLR charts have more replications with shorter time delays $\Delta_\tau$ and thereby perform better than $T^2$ chart. The $T^2$ chart also has $\Delta_\tau > 10$ for a fraction of replications (e.g., 27 replications for $\Delta x = 7, r = 4$). When the intensity shift is very small (i.e., $|\Delta x|$ = 5 or 6), 100 replications spread over the intervals $\Delta_\tau \in [0, 3]$, (3, 10] and [10, +∞]. Table 3-2 shows that NGLR charts have more replications with shorter time delays $\Delta_\tau$, thereby performing better than $T^2$ chart. Notably, time delay $\Delta_\tau$ is bigger for the scenarios with smaller radius of vesicles.

It should be noted that both NGLR and $T^2$ chart are established based on the *Q*-variate vector of network community statistics. Table 3-2 shows that both charts perform well in the change-point detection, but NGLR chart signals the out-of-control quicker than $T^2$ chart. This is because NGLR chart not only considers the present imaging profile, but also tracks historical imaging profiles within a sliding window to identify the maximal change in the likelihood ratio (also see Eq. (3.11) and Eq. (3.15)). However, computational complexity of NGLR is slightly higher than that of $T^2$ chart due to the computation of likelihood ratio in the sliding window. In practice, the tradeoff between performance and computational complexity needs to be considered in the selection of control chart for the purpose of process monitoring.

*4) Dynamic network GLR chart in the reduced dimension*

Furthermore, we conducted experiments to investigate the performance of dynamic network GLR charts in the reduced dimension. Notably, Table 3-2 is obtained with the complete set of community statistics (45 features = 15 communities × 3 statistics in each community). The next step is to evaluate the performance of NGLR and $T^2$ charts in the reduced dimension. Monitoring schemes with a low-dimensional set of features tend to be more efficient because of the simplified model form. Therefore, we reduced the dimensionality of *Q*-variate vector of network community statistics by keeping the first $q$ principal components with relatively large eigenvalues that explain the majority of data variance. The Pareto chart in Fig. 3-11 shows the percentage of variance explained by principal components (PC) that are sorted with respect to the descending eigenvalues. The first PC with the largest eigenvalue explains 20% of total variance. It should be noted that Fig. 3-11 does not have a few predominant PCs (e.g., < 5) that explain more than 90% of total variance. A total of 30 principal components are necessary to achieve the performance to explain 95% of total data variance.

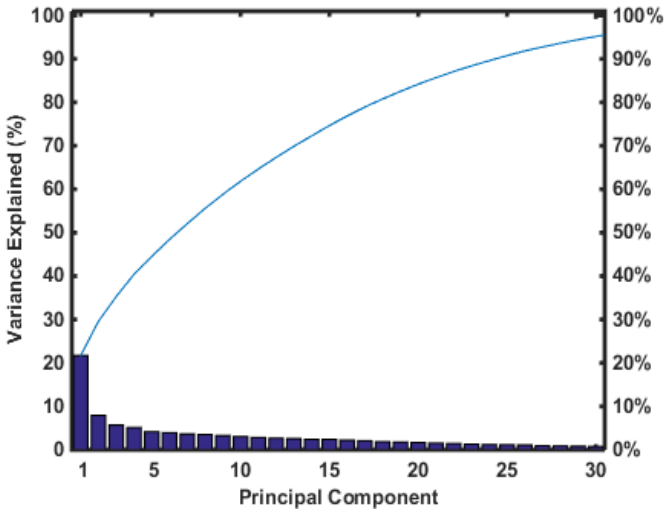



Figure 3-11 Pareto chart of the percentage of variance explained by principal components of the Q-variate vector of network community statistics.

As shown in Table 3-3, we considered the reduced dimension $q$ from 40, 35, 30, to 10, which explains 99%, 97%, 95% and 60% of data variance, respectively. The performances of NGLR and $T^2$ monitoring schemes were evaluated for four different $q$ values under 18 scenarios of varying size ($r$) and intensity shift ($\Delta x$). First of all, $ARL_1$s of both charts increase when the dimension $q$ gets smaller. This is more evident when the intensity shift ($\Delta x$) and radius ($r$)is very small (i.e., $|\Delta x| = 5$ and $r = 4$). But it may be noted that NGLR generally yields better performance than $T^2$ chart. For example, NGLR chart yields the $ARL_1$ of 4.41 and $T^2$ chart is 18.85 when $\Delta = -5$, $r = 4$ and $q = 40$. When the dimensionality $q$ is further reduced to 35, the $ARL_1$ of NGLR chart is increased to 6.23 and $T^2$ chart is 20.42. In this present study, $ARL_1$ is marked as N/A if out-of-control is not signaled for more than 70 (out of 100) replications. When the dimensionality $q$ is reduced to 10, only 60% of data variance is explained. Both charts signal the out-of-control when the intensity shift is large (i.e., $|\Delta x|$=8 or 10), but NGLR chart detects the change point quicker than $T^2$ chart. When the intensity shift is small (i.e., $|\Delta x|$=5), both charts encounter difficulty to detect the out-of-control. Notably, most $ARL_1$s of $T^2$ chart are N/A when $|\Delta x| = 5$. It should be noted that $T^2$ chart is more sensitive to the reduced dimensionality than NGLR chart. In general. NGLR chart performs better than Hotelling $T^2$ chart in the reduced dimension.

Table 3-3. Performance Comparison of Multivariate Network GLR and $T^2$ Charts in the Reduced Dimension

| $\Delta x$ | $r$ | $ARL_1$ ($q$ = 40) Variance 99% | | $ARL_1$ ($q$ = 35) Variance 97% | | $ARL_1$ ($q$ = 30) Variance 95% | | $ARL_1$ ($q$ = 10) Variance 60% | |
|---|---|---|---|---|---|---|---|---|---|
| | | *NGLR* | $T^2$ | *NGLR* | $T^2$ | *NGLR* | $T^2$ | *NGLR* | $T^2$ |
| -10 | 4 | 1.00 | 1.00 | 1.00 | 1.00 | 1.02 | 1.00 | 1.82 | 2.31 |
| | 6 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | 8 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.01 | 1.00 |
| -8 | 4 | 1.43 | 1.69 | 1.55 | 1.98 | 1.59 | 2.12 | 4.28 | 5.86 |
| | 6 | 1.01 | 1.09 | 1.05 | 1.19 | 1.08 | 1.11 | 2.04 | 2.28 |
| | 8 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.80 | 2.38 |
| -5 | 4 | 4.41 | 18.85 | 6.23 | 20.42 | 9.34 | N/A | N/A | N/A |
| | 6 | 2.38 | 7.38 | 2.69 | 18.11 | 3.19 | 22.38 | 23.96 | N/A |
| | 8 | 1.36 | 2.37 | 1.56 | 5.04 | 1.83 | 5.46 | 9.22 | N/A |
| 5 | 4 | 4.13 | 17.28 | 6.94 | 26.33 | 10.45 | N/A | N/A | N/A |
| | 6 | 2.08 | 6.65 | 2.99 | 20.68 | 3.40 | 19.97 | 23.31 | N/A |
| | 8 | 1.31 | 2.47 | 1.77 | 6.79 | 1.79 | 5.83 | 9.23 | N/A |
| 8 | 4 | 1.57 | 5.58 | 2.39 | 12.46 | 2.72 | 14.13 | 18.49 | N/A |
| | 6 | 1.04 | 1.18 | 1.15 | 1.71 | 1.11 | 1.51 | 4.33 | 15.92 |
| | 8 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 2.59 | 6.25 |
| 10 | 4 | 1.02 | 1.08 | 1.07 | 1.13 | 1.09 | 1.06 | 4.10 | 15.66 |
| | 6 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.18 | 1.09 |
| | 8 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.01 | 1.02 |

## *B. Case Study in Ultra-Precision Machining*

This case study is aimed at evaluating the performances of dynamic network algorithms for quality inspection of imaging profiles from the UPM process (see Fig. 3-1a). It is generally known that UPM produces mirror finish surfaces, which are used for a variety of engineering applications, including sensors, laser scanners, computer hard drives, medical equipment and optical elements in aerospace industries [62] – [63]. The UPM process employs a single crystal diamond tool to remove metal materials at the depth of cut in the micrometer range, i.e., 1~50 $\mu m$, which can achieve surface finish ($R_a$) in the nanometer range. In the industry practice, real-world uncertainty factors such as microstructure heterogeneity, environmental changes and process drifts often cause the variations of surface finishes in UPM machining (see Fig. 3-12). As shown in Fig. 3-1a, the UPM machine is equipped with sensors, an aerostatic spindle bearing (model Block-Head® 4R)

and air-slide tool carriage. In this present study, image profiles were acquired using the confocal optical microscope MicroXAM® from the finished aluminum alloy workpieces ($\Phi 8cm \times 3cm$). Monitoring and control of surface morphology variations is critical to ensuring the quality of surface finishes, improving the manufacturing performance and increasing competitive advantage in the global market.

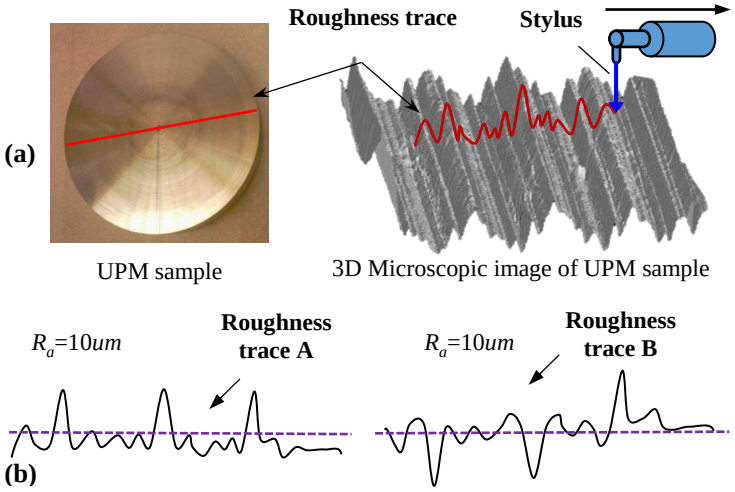



Figure 3-12. (a) Linear roughness measurement (roughness trace is marked in red); (b) Two distinct roughness traces (black/solid) with same $R_a$ values.

Traditionally, the surface finish of UPM process is measured using a diamond stylus profilometer. Although effective to a certain level, the profilometer measures only a line across the surface that is not sufficient to characterize the surface quality. As shown in Fig. 3-12a, the travel path of the stylus is recorded as a roughness trace that serves as input data for the calculation of surface roughness parameters. One commonly used parameter is $R_a$, which is the arithmetic average of absolute distances from each point of the roughness trace to the mean line. However, the 1-dimensional roughness trace is limited in its ability to represent and characterize the 3-dimensional surface. Quality of unexplored regions in the 2D surface remains unknown. In addition, $R_a$ averages out information and cannot adequately describe the surface characteristics. For example, two surface morphologies may be different but have the same $R_a$ value (also see Fig. 3-12b). As advanced metrology brings the increasing availability of imaging profiles, there is an urgent need to develop new methods and tools to characterize surface variations in the UPM process.

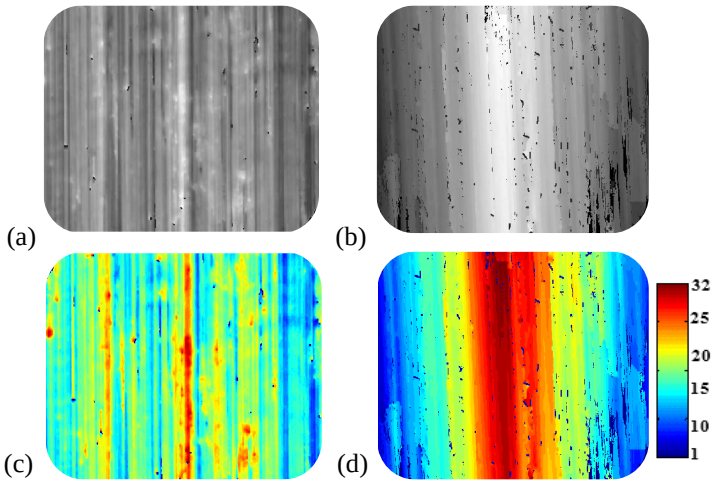



Figure 3-13. Surface images from UPM process, (a) a smooth surface with surface roughness $R_a$ = 43nm; (b) a rough surface with $R_a$ = 585 nm; (c) community detection results of the smooth surface in (a); (d) community detection results of the rough surface in (b).

In this present study, we propose to utilize the distribution of community statistics (i.e., extracted from network models of imaging profiles) to monitor and control the UPM surface finishes. Fig. 3-13a shows a smooth surface ($R_a$ = 43 nm), while Fig. 3-13b shows a rough surface ($R_a$ = 585 nm). In total, 32 communities are detected for both surfaces. Fig. 3-13c and Fig. 3-13d demonstrate obtained community labels, which are represented with 32 different colors of UPM images. Notably, communities with small mean intensity values are in blue colors with a gradual transition to communities with large mean intensity in red colors. The distribution of communities is significantly different between smooth and rough surfaces.

Furthermore, new statistics are extracted from each community. Fig. 3-14a shows community statistics of the smooth surface and Fig. 3-14b shows those of the rough surface. As shown in the figures, pixel amplitudes of the smooth surface are approximately in the same range as the rough surface (i.e., from 0 to 200). Nonetheless, a noteworthy difference can be identified in the distribution of the percentage of pixels among communities (see histograms in Fig. 3-14a and 3-14b for smooth and rough surfaces, respectively). The 32 communities are sorted by ascending values of mean intensity in each community. The percentage of pixels for one community is calculated as the ratio between the number of pixels in that community and the total number of pixels in the image profile.

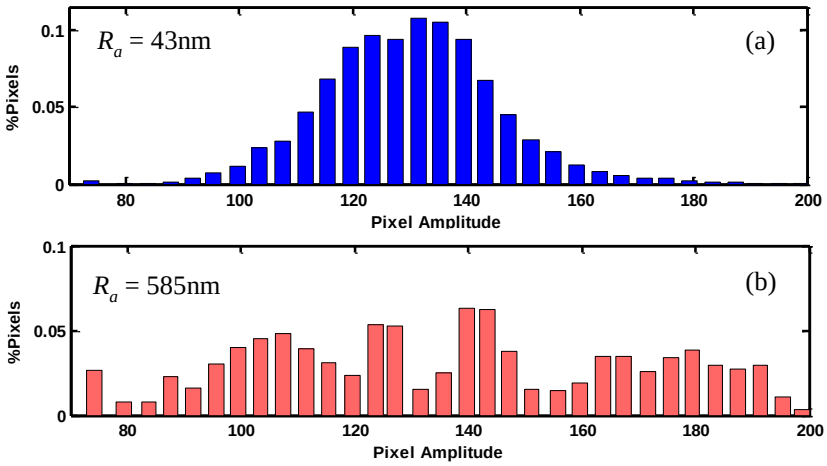



Figure 3-14. Distribution of community statistics – the percentage of pixels vs. average pixel amplitude in each community for (a) Smooth surface with $R_a$ = 43nm; (b) Rough surface with $R_a$ = 585nm.

As shown in Fig. 3-14a, most of pixels from the image of smooth surface belong to communities with amplitudes from 110 to 150. It is approximately a bell shape and most of pixels are centralized to a mean value. As such, the surface is smooth with a small variation of pixel intensities. In contrast, rough surface is with intensity values spreading from 0 to 200 (Fig. 3-14b), which results in a large variation of pixels. As a result, 32 communities do not approximately yield a normal distribution of mean intensity values. Therefore, the proposed network model effectively reveals hidden features from the image that pertinent to the quality of surface finishes.

It is worth mentioning that the extracted $Q$-variate vector $\boldsymbol{y}^{(i)} = \left(\bar{x}_1^{(i)}, \ldots, \bar{x}_K^{(i)}, s_1^{(i)^2}, \ldots, s_K^{(i)^2}, n_1^{(i)}, \ldots, n_K^{(i)}\right)^T$ consists of sample mean, variance and number of pixels in each network community of the image, thereby providing an effective characterization of finished surface profiles. We further extracted community statistics from imaging profiles of 80 UPM aluminum alloy samples. The dimensionality of Q-variate vector is 96, i.e., 32 communities × 3 statistics in each community. In order to avert the issue of "curse of dimensionality", we transformed these features into linearly uncorrelated PCs. The first 10 PCs were found to retain 97% of the data variance, and then used for the computation of GLR statistics in the reduced dimension. Fig. 3-15 shows the scatter plot of GLR statistics and $R_a$ values for 80 UPM imaging profiles. The $R_a$ value is captured using the MapVue software available in the MicroXAM® microscope. The process shift (e.g., rough surfaces with large $R_a$) is mainly resulted from the

change in machine parameters and/or conditions. In our experiments, the same shift occurs over the time in UPM surfaces until the chart detects the shift.

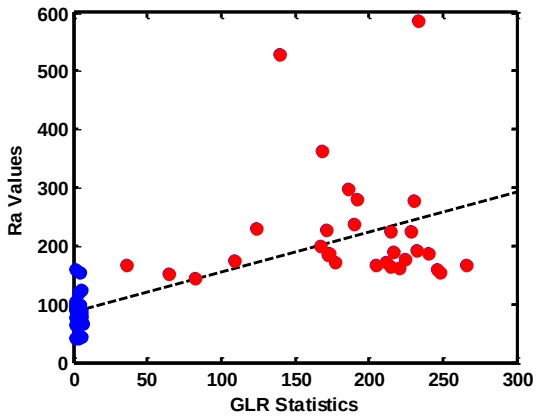



Figure 3-15. Scatter plot of $R_a$ values and GLR statistics from 80 images.

As shown in Fig. 3-15, network GLR statistics are highly correlated with $R_a$ values. The correlation coefficient is about 0.69. Notably, smooth surfaces with small GLR statistics and low $R_a$ values (blue dots) are concentrated in the lower left corner. On the contrary, rough surfaces with large GLR statistics and high $R_a$values (red dots) are distributed in the middle and upper right portion of the figure. It may also be noted that GLR statistics are well separated into two groups along the x-axis, while $R_a$ values have overlaps on the y-axis). The variation of GLR statistics among rough surfaces are higher than smooth surfaces (see the distribution of red dots in Fig. 3-15). This is due to the fact that community statistics of smooth surfaces are homogeneous and close to the normal distribution as shown in Fig. 3-14a. However, community statistics of rough surfaces vary significantly in their distributions, which are not always similar to Fig. 3-14b. Fig. 3-16 shows two distributions of community statistics from 2 rough UPM finishes. Fig. 3-16a is highly skewed and has a long right tail, but Fig. 3-16b is bimodal. As such, GLR statistics differs between Fig. 3-16a and Fig. 3-16b.

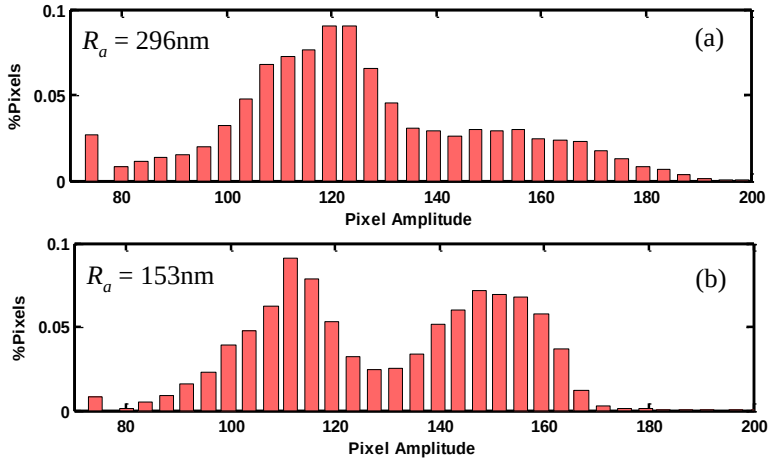



Figure 3-16. Distributions of community statistics – the percentage of pixels vs. average pixel amplitude in each community for two rough surfaces.

## 3.5. Discussion and Conclusions

In the past few years, advanced imaging technology is increasingly invested to cope with system complexity and increase information visibility. As such, large amounts of high-dimensional images are generated, which pose great challenges on existing practices in process monitoring and quality control. Most of existing methods are designed for monitoring 1D linear and nonlinear profiles, and are limited in their ability to handle complex structures of imaging profiles. Although some image-based SPC methods were developed, they mainly focused on fault detection in snap-shot images. Few, if any previous research has been conducted to extract useful information from time-varying images for process control. In this study, we developed a new dynamic network scheme to represent, model and control in-situ image profiles. We first cast each image as a weighted network. As such, a dynamic network is obtained from the time-varying image stream. Then, a new approach of Potts model Hamiltonian is introduced to detect communities from the dynamic network. New statistics are extracted from network communities to characterize the spatiotemporal patterns of image profiles. Finally, we design a new NGLR to effectively detect the change from the underlying process.

The new dynamic network scheme is implemented and evaluated for real-world applications in biomanufacturing and ultra-precision machining processes. Experimental results show that the proposed approach effectively characterizes variations in complex structures of high-dimensional images and can be used an enabling tool for online monitoring of time-varying image profiles. The dynamic network scheme shows strong potentials to be applied to diverse industries with in-situ

imaging data. While existing statistical methods mainly focus on the modeling and analysis of 2D images, dynamic network has advantageous features to effectively represent high-dimensional imaging profiles (i.e., > 2D images), Network nodes are flexible to represent pixels in 3D, 4D and multi-dimensional images. Network edges effectively characterize spatiotemporal patterns in the stream of images. In the future, we will further investigate and develop network methods and tools for quality control of complex spatial patterns in 3D/4D image streams.

# Chapter 4: Parallel Computing and Network Analytics for Fast Industrial Internet-of-Things (IIoT) Machine Information Processing and Condition Monitoring


## Abstract

Rapid advancement in sensing, communication, and mobile technologies brings a new wave of Industrial Internet of Things (IIoT) technology. IIoT integrates a large number of sensors for smart and connected monitoring of machine conditions. Sensor observations contain rich information on operational signatures of machines, thereby providing a great opportunity for machine condition monitoring and control in the context of large-scale IIoT. However, realizing the full potential of IIoT depends to a great extent on the development of new methodologies using big data analytics. Existing approaches are limited in their ability to effectively extract pertinent knowledge about manufacturing operations from the large volume of IIoT data of networked machines. This paper presents new parallel algorithms for large-scale IIoT machine information processing, network modeling, condition monitoring, and fault diagnosis. First, we introduce a dynamic time warping algorithm to characterize and quantify the dissimilarity of machine signatures (e.g., power profiles during operations). Second, we develop a stochastic network embedding algorithm to construct a large-scale network of IIoT machines, in which the dissimilarity between machine signatures is preserved in the network node-to-node distance. When the machine condition varies, the location of the corresponding network node changes accordingly. As such, node locations will provide diagnostic information about machine conditions. However, the network embedding algorithm is computationally expensive in the presence of large amounts of IIoT enabled machines. Therefore, we further develop a parallel algorithm that harnesses the power of multiple processors for efficient network modeling of large-scale IIoT enabled machines. Experimental results show that the new parallel-computing algorithm efficiently and effectively characterizes the variations of machine signatures for network modeling and monitoring. This new approach shows strong potentials for optimal machine scheduling and maintenance in the context of large-scale IIoT.

## 4.1. Introduction

The manufacturing sector has a large footprint in the US economy and produced the gross output of $5940.3 billions in 2015, which accounted for 18.92% of total American gross domestic product. To achieve competitive advantages and promote sustainable infrastructure, Internet of Things (IoT) technology has been increasingly invested to transfer capital and labor-intensive processes into smart manufacturing systems with high-level reliability and safety. In particular, rapid growth of Industrial IoT is driving the next generation of manufacturing technology. The IIoT technology enables the deployment of a multitude of wireless sensors, computing units and physical objects (e.g., machines and tools) in a network that helps "things" to communicate and exchange data. Sensors collect data pertinent to the condition of machines and tools and transmit them to computing units for in-situ analysis and process control. Such a connected cyber-physical system provides an unprecedented opportunity to realize intelligent manufacturing systems that leverage sensor signals and advanced analytics for decision making, e.g., predictive maintenance, machine scheduling, and energy consumption optimization.

In the IIoT, heterogeneous types of data are continuously collected, e.g., vibration, force, acoustic, and power consumption. Together, these data provide real-time machine signatures that contain rich information pertinent to variations in the manufacturing process. Fig. 4-1. shows one type of machine signature – an power profile – that characterizes the power consumption of successive operations in the turning process. For example, the spindle is first accelerated, which is followed by movements of the turret. The cutting phase is with the longest duration of the entire cycle and constantly consumes ~4KW power. Notably, the morphology of the power profile (e.g., amplitude and duration) can vary significantly due to different machine settings (feed rate and cutting depth), coolant use, type of materials, etc. [64].

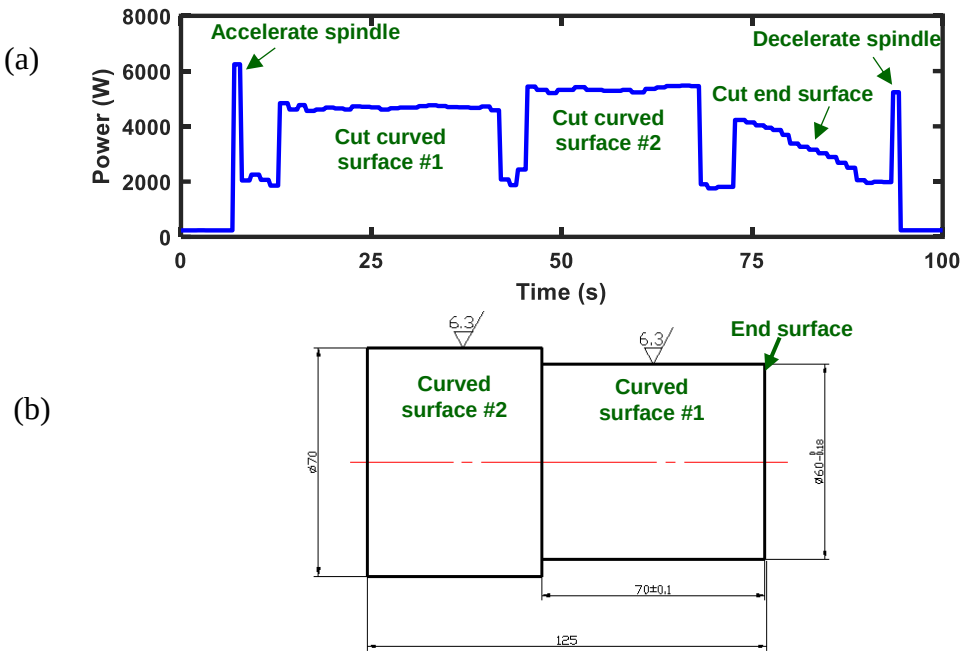



Figure 4-1. (a) The power profile of a turning machine; (b) The machined part (JIS S45C).

Also, discrete-part manufacturing generates cycle-based signatures in continuous operations. As shown in Fig. 4-2, the cycle-based signature is collected during repetitive operation cycles of the manufacturing process using advanced data acquisition systems. Cycle-based signatures are very common in various manufacturing processes, e.g., forming force in the stamping process and insertion force in the engine assembly process [9], [7]. Because the machine condition is continuously monitored in the IIoT settings, signatures with tens of thousands of operation cycles can be collected from one machine. As shown in Fig. 4-2, successive cycles of the signature demonstrate patterns that are similar to each other but with variations. For example, each cycle contains movements of the turret and the cutting phase (similar to the profile shown in Fig. 4-1). However, the cutting phase of cycle 3 consumes higher power and has a larger variation compared with cycles 1 and 2. This indicates a process shift, which may be due to extraneous factors (e.g., the onset of chatter or the start of tool wear). Hence, cycle-to-cycle variations of machine signatures provide a wealth of information for manufacturing process monitoring and control. This will enable engineers to make proactive decisions about process adjustment and machine maintenance, which improves the quality of products and reduces the re-work rate. Therefore, there is a need to characterize and quantify cycle-to-cycle variations in signatures from a single machine for the development of customized monitoring schemes.

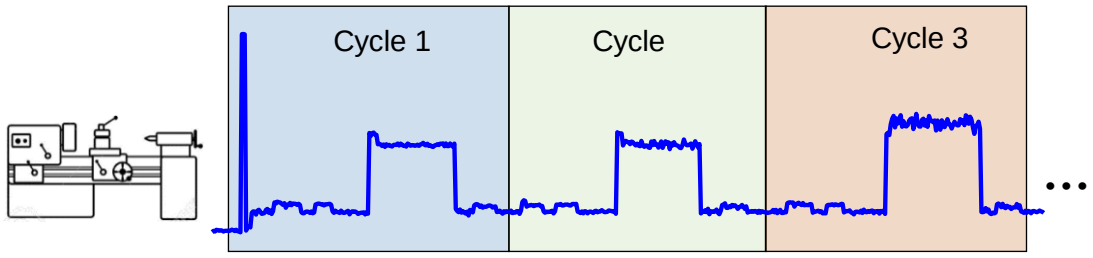



Figure 4-2. Cycle-to-cycle variations of power profiles in a single machine.

Furthermore, the IIoT enables the internet-like connection of a large number of machines. Note that machine signatures may vary due to a number of factors such as products, machine types, procedures, and anomalies. For example, networked machines in Fig. 4-3 demonstrate distinct morphology of power consumptions in the cutting phase. Some of them have normal patterns (e.g., M2 and M3), whereas some are associated with larger variations (e.g., M1 and M6) and elevated patterns (e.g., M4 and M5). Quantifying machine-to-machine dissimilarities and developing an IIoT-based machine-to-machine network provides a great opportunity for 1) Condition monitoring and quality control: Machines with similar conditions can be grouped into the same cluster. The structure of a machine-to-machine network not only provides useful information on the machine status and utilization statistics but also offers a good opportunity for profile-based machine categorization, product categorization, and online quality control. 2) Planning and scheduling: The structure of a machine-to-machine network is dynamically varying because machine profiles are changing over time. Such a dynamic network can further help optimize maintenance decision, manufacturing planning, and scheduling. For example, we can proactively assign its workload to other (normal) machines and schedule maintenance, when a machine is moving towards the cluster of “machine failure” in the network. 3) The large-scale smart manufacturing: For a large-scale manufacturing system, advanced sensing increases the information visibility and helps cope with the high-level complexity in the system. IIoT provides an opportunity to realize the machine-to-machine network for smart manufacturing. For example, machines communicate with each other to report status and exchange information for optimal planning and scheduling. This will significantly help create value from data, optimize factory operations, reduce maintenance costs and equipment downtime.

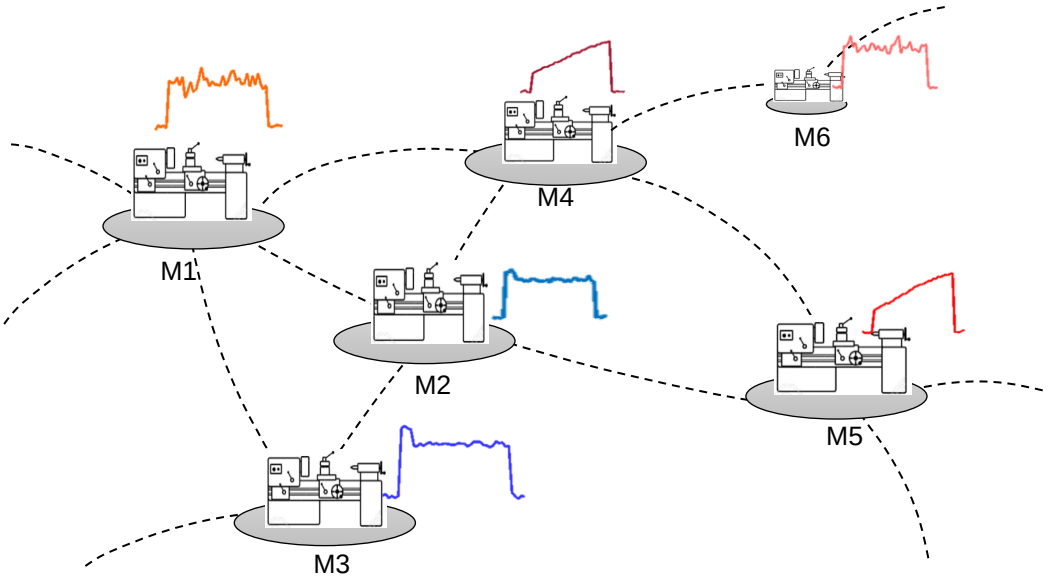



Figure 4-3. A machine-to-machine network in the IIoT.

In the IIoT context, there is an urgent need to characterize and quantify cycle-to-cycle variations customized to one machine and machine-to-machine dissimilarities of the population for process monitoring and control. However, two major gaps remain in current practice. First, there is a lack of effective tools to characterize dissimilarities between machine signatures. It may be noted that cycles of machine signatures are similar to each other but with some variations. Such variations can be induced by the shift of process-related factors, including tool wear, incorrect machine settings, and misuse of coolant. Some other factors, nevertheless, may also result in variations of machine signatures, e.g., measurement noises and cycle misalignment due to discrete sampling or phase shift. Hence, there is a need to discern various factors and extract salient features that are sensitive to operational, sustainable states of machines.

Second, existing algorithms are limited in their ability to handle large volumes of real-time data. It is worth mentioning that real-time condition monitoring generates big data for a single machine or a network of machines. Big data generated from the large-scale IIoT not only provides an unprecedented opportunity for optimal manufacturing decision making, but also poses significant challenges for analysis, management, and storage [65]. Realizing the full potential of the IIoT depends to a great extent on big data analytics for smart manufacturing.

In this investigation, a fast parallel-computing algorithm is developed to leverage the big data in the large-scale IIoT for machine condition monitoring and fault diagnosis. IIoT analytical

methods and tools will enable and assist (i) the handling of massive data communicated by the IIoT; (ii) the extraction of pertinent knowledge about machine dynamics; (iii) the exploitation of acquired knowledge for data-driven analysis and performance optimization of machines. Specifically, a dynamic time warping approach is introduced to measure dissimilarities among machine signatures. Then, machines are optimally embedded into a network in the high-dimensional space (e.g., 2D and 3D). Each node corresponds to one machine, and Euclidean distances between network nodes preserve the dissimilarity of machine signatures. As such, the location of network nodes reveals diagnostic features pertinent to machine conditions. However, the ability to optimize a large network is hampered in traditional algorithms, due to their high computational complexity. Thus, we further introduce a stochastic network embedding approach to handle the large-scale machine network. Parallel-computing algorithms are also developed to speed up the stochastic embedding approach. Experimental results show significant reduction of running time for a large-scale IIoT network. Our contributions in the present investigation are highlighted as follows:

1) We introduce a stochastic embedding approach (rather than traditional deterministic algorithms) to significantly improve the computational efficiency of network embedding and optimize the embedding of each machine in the network.

2) We develop a fast parallel algorithm (rather than traditional serial computing) to enable the embedding of large-scale machine networks in the context of IIoT.

The remainder of this paper is organized as follows: Section 4.2 introduces the research background in condition monitoring, IIoT, and parallel computing. Section 4.3 presents the methodology of stochastic network embedding and parallel computing algorithms developed in this research. Case studies and experimental results are shown in Section 4.4. Section 4.5 discusses and concludes this research.

## 4.2. Research Background

### *A. Advanced Sensing and Condition Monitoring*

Rapid advances in sensing technology have prompted the proliferation of advanced sensing in manufacturing. 1) Instead of single-channel sensors, multi-channel sensors are more and more adopted for the machine condition monitoring. By distributing sensors at different positions, multi-channel sensing provides a comprehensive view of the manufacturing system. For example, four

strain gage sensors are commonly installed in a forging machine. The sensing signals provide rich information pertinent to the product quality [51, 66, 67]. 2) Instead of a single type of sensor, heterogeneous sensors are used to monitor the manufacturing process from different perspectives. For example, ultraprecision machining is equipped with multiple types of sensors, including cutting force, vibration and acoustic emission [68]. Data fusion of heterogeneous sensors provides a great opportunity for quality assurance of manufactured products [69, 70].

Notably, realizing the full potential of advanced sensing for manufacturing process control hinges on the modeling and analysis of collected machine signatures. In the literature, mechanical signatures (such as vibration and force signals) have been widely used for machine condition monitoring and quality control. For example, Zhou, Sun and Shi [9] implemented Haar wavelet transform to decompose tonnage force signals for the multi-scale characterization of variations in the stamping process. Statistical control chart was then used to detect the process shift with wavelet coefficients. Kan, Cheng and Yang [14] investigated heterogeneous recurrence behaviors in vibration signals of ultraprecision machining. Recurrence quantifiers were used as features for in-situ monitoring of the surface integrity of manufactured wafers. In addition, Shao et al. [71] monitored the ultrasonic metal welding with acoustic signals collected by watt meter and microphone. A cross-validation approach was developed to select representative features from acoustic signals to evaluate the quality of joints between battery tabs.

Recently, electrical signatures are increasingly used for the characterization of machine conditions. For example, Seera et al. [72] analyzed the stator-current signals for the fault detection and diagnosis of induction motors. A hybrid fuzzy min-max neural network model was developed to extract pertinent features about various motor faults, which were then differentiated by a decision tree model. Emec, Kruger and Seliger [73] developed an online approach for the analysis of electrical intake of machine tools for predictive maintenance of milling and drilling machines. The morphology of the power profile was compared with a selected reference pattern to identify the malfunction of machine tools. It may be noted that the acquisition of electrical signatures does not require sensors installed in machine tools or drive trains. As such, it allows for high flexibility of condition monitoring on a variety of machines and facilitates remote control. Also, electrical power data has high signal-to-noise ratio. As opposed to mechanical signatures, it is less influenced by external perturbations.

However, most previous approaches focus on the analysis of representative cycles from a single machine. Very little has been done to characterize cycle-to-cycle and machine-to-machine variations for machine condition monitoring in the context of large-scale IIoT. It may be noted that continuous monitoring of machines as collaborative agents in a network will facilitate real-time loading and maintenance scheduling, and thereby improving the productivity. Thus, there is a need to design and develop new approaches that effectively leverage continuously collected IIoT data for in-situ process monitoring and control.

### *B. Industrial Internet of Things*

In the past few years, the manufacturing industry is hurtling into the era of IIoT [74]. By connecting wireless sensors and mobile computing units, IIoT is triggering a paradigm shift in manufacturing, and it will be a major trend for the global industry. Nowadays, a massive number of machines and tools, from oil-drilling rigs to steel mills, are equipped with sensors to collect and send data. Data analytics help to detect the operational issues and process anomalies in the manufacturing system. Parts are automatically replenished, and machine tools are replaced before failure occurs. As such, the IIoT significantly reduces the risk of unplanned outages and improves the production throughput. It is estimated that IIoT-enabled predictive maintenance saves up to 12% of over scheduled repairs, reduces 30% maintenance costs, and eliminates breakdowns up to 70% [75].

In addition, a large-scale machine network can be established under the cyber-physical system of IIoT. Machines are represented as nodes in the network, and the machine-to-machine communication is realized through network arcs. Manufacturers can access the machines' operational dynamics anywhere anytime via the machine network and deliver remote, even overseas maintenance services. This will help reduce machine downtime and avoid shutdowns of factories. Nowadays, manufacturers are identifying new growth opportunities by adding intelligent products and digital services that communicate with each other and with people globally.

However, existing studies of the IIoT mainly focus on sensor development, wireless communication and information architecture [76][77]. Few analytics algorithms have been developed to leverage data collected in the IIoT for machine condition monitoring. There is a need to develop efficient algorithms to characterize collected machine signatures and quantify salient

features that are sensitive to operational, environmental, and sustainable states of manufacturing processes.

### *C. Serial Computing vs. Parallel Computing*

Traditionally, analytics algorithms are implemented in the serial-computing framework. As shown in Fig. 4-4a, serial computing decomposes the overall task (e.g., condition assessment of a large group of machines) into a discrete series of sub-tasks (e.g. assess the condition of individual machines). Then, sub-tasks are executed sequentially on a single processor. Although serial computing is efficient for analyzing small datasets, the computational complexity may increase exponentially with the growth of sample size, which results in prohibitive computation time. This, in turn, significantly hampers the use of serial computing for big data analytics in large-scale IIoT. Therefore, researchers are seeking a way to scale up the algorithm and utilize more computing resources to complete a task collaboratively.

Rapid advancements in information technology have catalyzed the incredibly fast growth in computing power. Recently, parallel computing has been employed to harness multiple processing units to solve problems. As opposed to serial processing, parallel computing is achieved by pipelining the overall task into multiple processors. As shown in Fig. 4-4b, each processor executes its part of the computation and works simultaneously with other processors. As such, the computing time can be critically reduced. Nowadays, the availability of multi-core CPUs, cell processors (e.g., GPUs) and cloud computing technology make parallel computing easily achievable by deploying off-the-shelf strategies such as multi-threaded and single-instruction-multiple-data. Parallel computing has been widely used in various research areas, including genetics, geology, medical imaging, and graph mining.

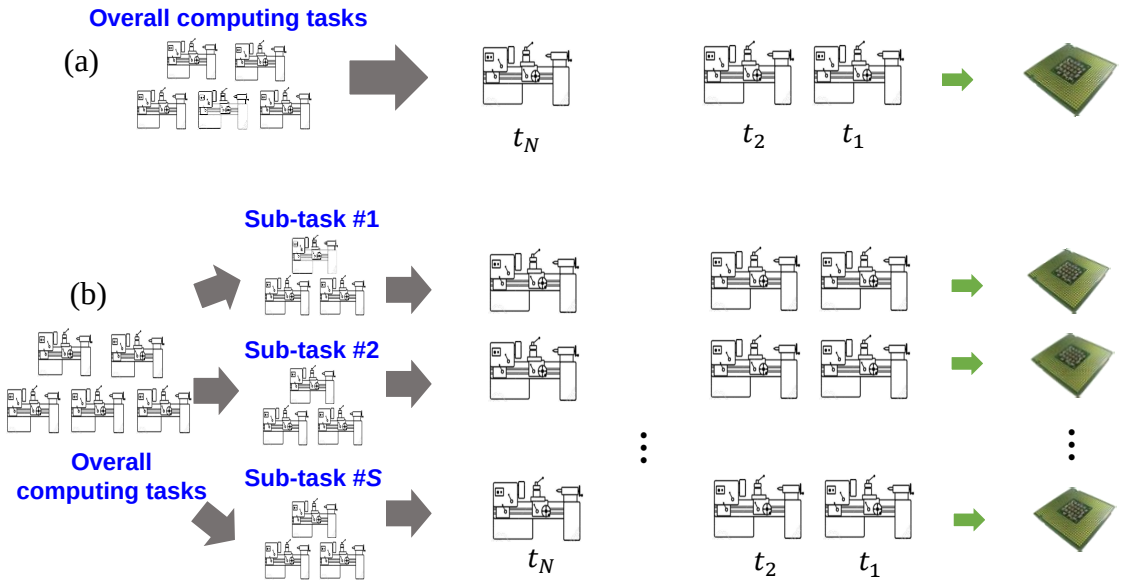



Figure 4-4. Machine-network information processing: (a) Serial computing on a single processor vs. (b) Parallel computing on multiple processors.

Parallel computing has been increasingly investigated in manufacturing. As every agent is capable of sensing and communicating with others, large-scale IIoT systems lead to the accumulation of big data, which poses a great challenge for analytics. To address this challenge, researchers attempted to harness the power of a large number of computing resources (e.g., multi-core computers) and parallelize the execution of computing tasks. For example, Raghavan and Waghmare [78] developed a distributed and parallel computing framework to integrate diverse computing resources for manufacturing applications, e.g., machine scheduling. The work-stealing scheduler and the tree-structured model allowed workers (i.e., computing units) to join and leave the framework dynamically. Mourtzis et al. [79] utilized the cloud technology to optimize the process planning. By real-time monitoring of machines in distributed shop floor, it dynamically assigned jobs to available machines with suitable tools. Tao et al. [80] developed a parallel computing algorithm to achieve highly efficient service decomposition and optimal selection. The proposed parallel algorithm was able to handle large numbers of constraints and it was shown to outperform traditional serial algorithms.

However, most of these studies focus on strategical problems in manufacturing, e.g., job scheduling and service optimization. In other words, little has been done to develop parallel algorithms to improve the efficiency and speed in the modeling and analysis of large amounts of machine signatures. As we discussed in the introduction section, quantifying dissimilarities among

machine signatures is of great importance for machine condition monitoring. Hence, there is a need to develop new parallel algorithms to analyze machine signatures for condition monitoring and performance improvements in the large-scale IIoT network of machines.

## 4.3. Research Methodology

In this study, a fast parallel algorithm is developed for condition monitoring of large-scale machine networks. As shown in Fig. 4-5, a dynamic warping approach is firstly introduced to measure pattern dissimilarities between machine signatures. Then, network embedding algorithms are developed to optimally represent each machine signature as a network node in the high-dimensional space. The distance between network nodes preserves the dissimilarity of corresponding machine signatures (detailed in Section 3.2). The location of network nodes reveals important information pertinent to dynamics of machine conditions, which can be further used for decision making in the manufacturing process.

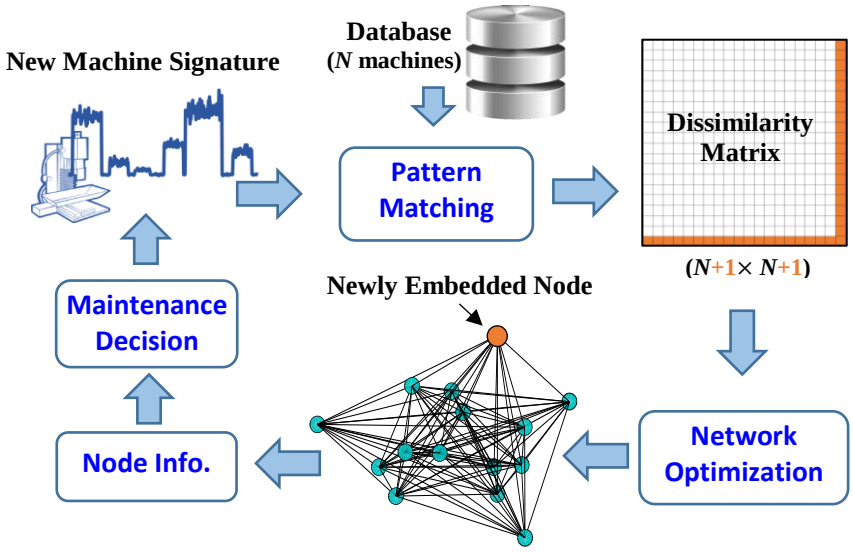



Figure 4-5. The proposed methodology for condition monitoring of IIoT machines.

### *A. Pattern Matching and Signature Warping*

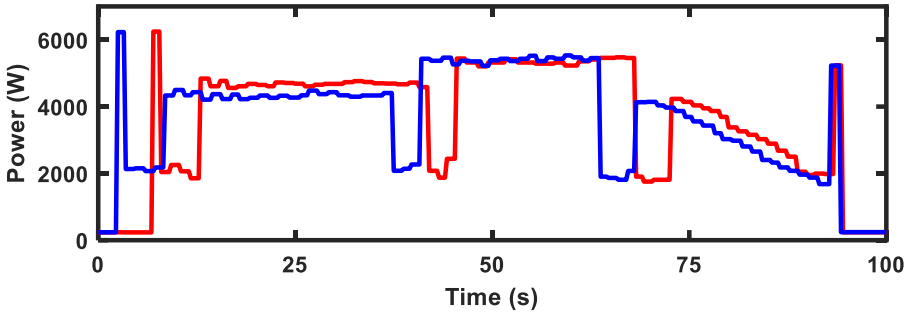



Figure 4-6. Power consumptions of two turning processes.

As shown in Fig. 4-6, dissimilarities are found between two energy profiles turning processes. The red profile consumes higher energy in the cutting phase, which indicates the potential tool wear. Quantification of such dissimilarities will provide a great opportunity for condition monitoring of machines and tools. Traditional methods focus on the comparison with reference signatures (i.e., “standard” or “normal” ones). We propose to perform a pairwise comparison of machine signatures in either cycle-to-cycle or machine-to-machine ways. Note that a dissimilarity matrix (consists of the dissimilarity between each pair of signatures) is obtained from pairwise comparison, rather than only a column in reference comparison.

However, two signatures can be misaligned due to discrete sampling and phase shift [40, 81]. For example, the blue and red signatures show a typical pattern, but there are variations in shape, amplitude, and phase (see Fig. 4-6). Directly taking the Euclidean distances between two misaligned signatures (e.g., Fig. 4-7a) will result in a large dissimilarity, even for phases that are alike. Therefore, it is imperative to optimally align the profiles in terms of spindle and cutting activities in different phases. For example, we should compare the cutting phases between two signatures, as opposed to an incorrect comparison between the cutting of end surface (in red) and spindle deceleration (in blue).

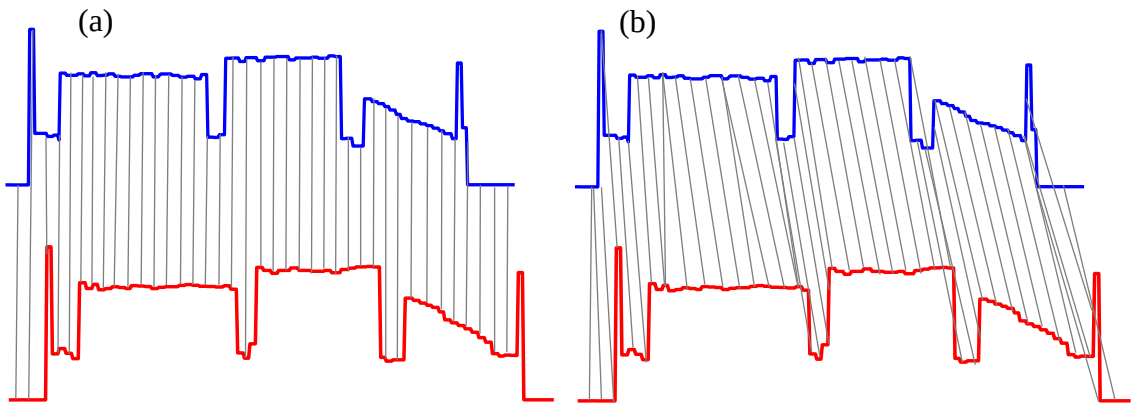



Figure 4-7. (a) Direct pattern matching of misaligned signatures; (b) Optimal pattern matching of aligned signatures by the dynamic time warping.

In this study, a dynamic time warping approach is introduced to optimally align two power signatures for the characterization of dissimilarity. Given two signatures $\overrightarrow{\boldsymbol{p}_1}(t)$, $t = 1, 2, \ldots, n_1$ and $\overrightarrow{\boldsymbol{p}_2}(t)$, $t = 1, 2, \ldots, n_2$, the dissimilarity between $\overrightarrow{\boldsymbol{p}_1}(t)$ and $\overrightarrow{\boldsymbol{p}_2}(t)$ is then measured as $\sum_{(t_i,t_j)\in\zeta}\left\|\overrightarrow{\boldsymbol{p}_1}(t_i) - \overrightarrow{\boldsymbol{p}_2}(t_j)\right\|$. To find the optimal warping path $\zeta$, a dynamic programming algorithm iteratively searches as:

$$W(i,j) = min\begin{pmatrix} W(i,j-1) + w(i,j) \\ W(i-1,j-1) + w(i,j) \\ W(i-1,j) + w(i,j) \end{pmatrix} \tag{4.1}$$

with the initial condition $W(1,1) = w(1,1) = \|\overrightarrow{\boldsymbol{p}_1}(t_1) - \overrightarrow{\boldsymbol{p}_2}(t_1)\|$ and a window size constraint $|i-j| < r$. The normalized dissimilarity between $\overrightarrow{\boldsymbol{p}_1}(t)$ and $\overrightarrow{\boldsymbol{p}_2}(t)$ are obtained as $\Delta(\overrightarrow{\boldsymbol{p}_1}, \overrightarrow{\boldsymbol{p}_2}) = W(N_1, N_2)/(n_1 + n_2)$. In this way, corresponding phases of machine signatures are optimally aligned for the measurement of pattern dissimilarities (see Fig. 4-7b).

Notably, the morphology of machine signatures is closely related to process dynamics. Thus, we are concerned about the morphology dissimilarity of signatures in this study. The proposed approach is not a "one-size-fit-all" solution, but profile alignment is an imperative step to measure signature dissimilarity. However, the misalignment information can be extracted to some extent from the length of each profile if they are of primary interests. If we do not use the warping approach and directly measure the difference between machine signatures, this will contaminate useful information and cannot yield meaningful results in most cases (see Fig. 4-7). On the

contrary, dynamic time warping optimally aligns two signatures and yields meaningful results by comparing the morphology of corresponding segments.

### *B. Network Embedding and Predictive Modeling of Machine Signatures*

Furthermore, each machine signature is optimally embedded as a network node in the high-dimensional space. The distance between nodes preserves the dissimilarity of two corresponding machine signatures. Fig. 4-8 illustrates the network embedding of six machine signatures. A dissimilarity matrix provides pertinent information about variations of machine signatures. By optimizing the location of nodes, node-to-node distances preserve the signature-to-signature dissimilarities in the matrix of Fig. 4-8a. For example, dissimilarities between machine 1 and others are preserved as Euclidean distances between node 1 and others (highlighted in red in Fig. 4-8a and 4-8b).

(a)

| | M1 | M2 | M3 | M4 | M5 | M6 |
|---|---|---|---|---|---|---|
| M1 | 0.00 | 2.55 | 2.63 | 1.40 | 2.59 | 3.79 |
| M2 | 2.55 | 0.00 | 1.93 | 1.40 | 1.45 | 1.96 |
| M3 | 2.63 | 1.93 | 0.00 | 2.45 | 1.80 | 3.10 |
| M4 | 1.40 | 1.40 | 2.45 | 0.00 | 2.03 | 2.84 |
| M5 | 2.59 | 1.45 | 1.80 | 2.03 | 0.00 | 1.52 |
| M6 | 3.79 | 1.96 | 3.10 | 2.84 | 1.52 | 0.00 |

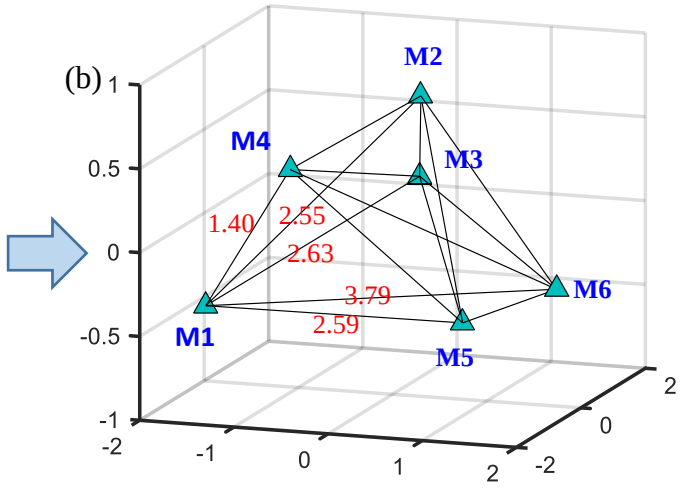



Figure 4-8. (a) Dissimilarity matrix of six machine signatures; (b) An embedded network with node-to-node distances preserving the dissimilarity matrix in (a).

Such a network embedding approach is usually achieved through classic multidimensional scaling. Let $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ denote the location of $i^{th}$ and $j^{th}$ nodes in the network and $\delta_{ij}$ is the dissimilarity between $i^{th}$ and $j^{th}$ machine signatures in the warping matrix $\Delta$. Then, the cost function of network embedding can be formulated as:

$$F = \sum_{i,j} (\|\boldsymbol{x}_i - \boldsymbol{x}_j\| - \delta_{ij})^2 \tag{4.2}$$

To optimize the location of network nodes, a Gram matrix $B$ is reconstructed from dissimilarity matrix $\Delta$ as $B = -\frac{1}{2}H\Delta^{(2)}H$, where H = $I - N^{-1}\mathbf{1}\mathbf{1}^T$ is a centering matrix. The element $b_{ij}$ in matrix $B$ is obtained as:

$$b_{ij} = -\frac{1}{2}[\delta_{ij}^2 - \frac{1}{N}\sum_{k=1}^{N}\delta_{ik}^2 - \frac{1}{N}\sum_{k=1}^{N}\delta_{kj}^2 + \frac{1}{N^2}\sum_{g=1}^{N}\sum_{h=1}^{N}\delta_{gh}^2] \tag{4.3}$$

The Gram matrix $B$ can be defined as the scalar product $B = XX^T$, which is further written as $B = V\Lambda V^T = V\sqrt{\Lambda}\sqrt{\Lambda}V^T$ by eigen-decomposition [81]. Here, $V$ is a matrix of eigenvectors and $\Lambda$ is a diagonal matrix of eigenvalues. Hence, optimal coordinates of network nodes are obtained as: $X = V\sqrt{\Lambda}$.

As a result, each machine signature is represented as a node in the network. Euclidean distances between nodes preserve warping distances, thereby characterizing signature-to-signature variations in the manufacturing process. This, in turn, allows the use of network information (e.g., coordinates of nodes) as features for condition monitoring and predictive modeling. 1) For the monitoring of cycle-to-cycle variations of a single machine, each node is embedded from one operation cycle of machine signature. If embedded nodes are located apart from the cluster of normal condition, maintenance service needs to be scheduled to prevent machine failure. 2) For the analysis of a machine network, each node represents a machine or the signature (e.g., power profiles, features, or patterns) of the machine. Classification models can be constructed to predict the machine condition based on coordinates of the corresponding node. If the machine condition is more likely to be abnormal, engineers can re-assign its jobs to other machines and schedule the maintenance. In our previous studies, various predictive modeling approaches were developed, including linear and quadratic discriminant models [8], neural networks [81], self-organizing networks [82], and particle filtering [83]. Practitioners can select an optimal predictive modeling approach based on the complexity of data and requirements of processing speed.

### *C. New Algorithms for Large-scale Network Modeling*

#### *1) Stochastic network embedding – serial computing algorithm*

However, classic multidimensional scaling is limited in its ability to handle large volumes of data. In other words, the computation time will be significantly increased when large numbers of

machines are involved in the network (e.g., $N$ >5,000). This is because the computational complexity of classical multidimensional scaling is $O(N^3)$, due to the centering operation and eigen-decomposition. In fact, most of traditional algorithms for network embedding are limited in the ability to handle big data due to high level of computational complexity. In addition to classic multidimensional scaling, another example is scaling by majorizing of a complicated function (SMACOF) [84]. It may be noted that SMACOF requires Guttman transform, which consists of a multiplication between an $N \times N$ matrix and an $N \times L$ matrix ($L$ is the dimensionality of the network space). As such, it is difficult to use traditional algorithms for the optimization of a large-scale machine network.

Notably, big data poses two significant challenges for optimization. On one hand, the dimensionality of the data is high. As such, second order optimization is sometimes prohibitive. Thus, first order methods are preferred, e.g., gradient-based method. On the other hand, large volumes of data make it impractical for algorithms to scan the entire dataset many times. Hence, a stochastic approach is better suited to handle big data, because only one data sample is considered at every iteration. Moreover, it has been shown that the convergence of stochastic approaches for the optimization of convex functions is guaranteed in most cases. Therefore, a stochastic embedding approach is introduced to extend the integrated framework in Section 3.1 to large-scale machine networks.

Motivated by [85], the objective function of network optimization in Eq. (4.2) is redefined as:

$$F = \frac{1}{2}\sum_{i}\sum_{j\neq i}\left(\left\|\boldsymbol{x}_i - \boldsymbol{x}_j\right\| - \delta_{ij}\right)^2 \Psi\left(\left\|\boldsymbol{x}_i - \boldsymbol{x}_j\right\|, \lambda\right) \tag{4.4}$$

where $\Psi(\cdot)$ is a bounded and monotonically decreasing function to favor the local topology in the network [85]. One example of $\Psi(\cdot)$ is a step function:

$$\Psi\left(\left\|\boldsymbol{x}_i - \boldsymbol{x}_j\right\|, \lambda\right) = \begin{cases} 1 & \text{if } \left\|\boldsymbol{x}_i - \boldsymbol{x}_j\right\| \leq \lambda \\ 0 & \text{if } \left\|\boldsymbol{x}_i - \boldsymbol{x}_j\right\| > \lambda \end{cases} \tag{4.5}$$

and $\lambda$ is defines the neighborhood radius. Minimization of Eq. (4.4) with respect to $\boldsymbol{x}_j$'s identifies the optimal location of network nodes that preserve the signature-to-signature dissimilarity measures. Notably, the decrease of $F$ values can be considered as a summation of partial decrease generated by adjusting the location of each network node $i$, i.e., $F = \frac{1}{2}\sum_i F_i$. The optimization of

Eq. (4.4) works as follows: in each iteration, a node $i$ is randomly selected and its location (e.g., $\boldsymbol{x}_i$) is fixed. Then, all other nodes $\boldsymbol{x}_j (j \neq i)$ are updated based on the following rule:

$$\boldsymbol{x}_j \leftarrow \boldsymbol{x}_j - \eta_t \frac{\partial F_i}{\partial \boldsymbol{x}_j} \tag{4.6}$$

where the learning rate $\eta_t$ is monotonically deceasing over time. Different forms of $\eta_t$ can be used and we adopt an exponential function:

$$\eta_t = \eta_0 \left(\frac{\eta_T}{\eta_0}\right)^t \tag{4.7}$$

where constants $\eta_0$ and $\eta_T$ specify the initial and final scales of the learning rate and the ratio of $\eta_T$ and $\eta_0$ determines the decaying speed. Furthermore, a similar form of exponential function can be applied to shrink the radius of neighborhood $\lambda$ in the learning process, which generates a monotonically decreasing function $\lambda_t = \lambda_0 \left(\frac{\lambda_T}{\lambda_0}\right)^t$.

*Proposition*: *The updating rule of stochastic network embedding is given as*

$$\begin{cases} \boldsymbol{x}_j \leftarrow \boldsymbol{x}_j - \eta_t \dfrac{\left\|\boldsymbol{x}_i - \boldsymbol{x}_j\right\| - \delta_{ij}}{\left\|\boldsymbol{x}_i - \boldsymbol{x}_j\right\|} \left(\boldsymbol{x}_j - \boldsymbol{x}_i\right) & \text{if } \left\|\boldsymbol{x}_i - \boldsymbol{x}_j\right\| \leq \lambda_t \\ \boldsymbol{x}_j \text{ is not updated} & \text{if } \left\|\boldsymbol{x}_i - \boldsymbol{x}_j\right\| > \lambda_t \end{cases} \tag{4.8}$$

*Proof*: Introduce an auxiliary parameter: $D_{ij} = \left\|\boldsymbol{x}_i - \boldsymbol{x}_j\right\| = \sqrt{\sum_{h=1}^{p} \left(x_{ih} - x_{jh}\right)^2}$. According to Eq. (4.4), $F_i$ can be written as $\frac{1}{2}\sum_j \left(D_{ij} - \delta_{ij}\right)^2 \Psi\left(D_{ij}, \lambda\right)$. Furthermore, $\frac{\partial \Psi}{\partial \boldsymbol{x}_j} = 1$, if $D_{ij} \leq \lambda_t$. Otherwise, $\frac{\partial \Psi}{\partial \boldsymbol{x}_j} = 0$.

According to the chain rule:

$$\frac{\partial F_i}{\partial \boldsymbol{x}_j} = \frac{\partial F_i}{\partial D_{ij}} \cdot \frac{\partial D_{ij}}{\partial \boldsymbol{x}_j}$$

$$\frac{\partial F_i}{\partial D_{ij}} = \frac{1}{2} \frac{\partial \left(\sum_j \left(D_{ij} - \delta_{ij}\right)^2\right)}{\partial D_{ij}} = \frac{1}{2} \frac{\partial \left(\sum_j \left({D_{ij}}^2 - 2D_{ij}\delta_{ij} + {\delta_{ij}}^2\right)\right)}{\partial D_{ij}} = D_{ij} - \delta_{ij}$$

Also we have:

$$\frac{\partial D_{ij}}{\partial x_{jh}} = \frac{\partial \sqrt{\sum_{h=1}^{p}\left(x_{ih}-x_{jh}\right)^2}}{\partial x_{jh}} = \frac{1}{2\sqrt{\sum_{h=1}^{p}\left(x_{ih}-x_{jh}\right)^2}} \cdot \frac{\partial\left(\sum_{h=1}^{p}\left(x_{ih}-x_{jh}\right)^2\right)}{\partial x_{jh}}$$

$$= \frac{1}{2D_{ij}}\, 2\left(x_{jh}-x_{ih}\right)$$

$$= \frac{x_{jh}-x_{ih}}{D_{ij}}$$

Thus,

$$\frac{\partial D_{ij}}{\partial \boldsymbol{x}_j} = \left(\frac{\partial D_{ij}}{\partial x_{j1}}, \frac{\partial D_{ij}}{\partial x_{j2}}, \ldots, \frac{\partial D_{ij}}{\partial x_{jh}}, \ldots\right) = \left(\frac{x_{j1}-x_{i1}}{D_{ij}}, \frac{x_{j2}-x_{i2}}{D_{ij}}, \ldots, \frac{x_{jh}-x_{ih}}{D_{ij}}, \ldots\right) = \frac{\boldsymbol{x}_j-\boldsymbol{x}_i}{D_{ij}}$$

Therefore:

$$\frac{\partial F_i}{\partial \boldsymbol{x}_j} = \frac{\left(\left\|\boldsymbol{x}_i-\boldsymbol{x}_j\right\|-\delta_{ij}\right)\left(\boldsymbol{x}_j-\boldsymbol{x}_i\right)}{\left\|\boldsymbol{x}_i-\boldsymbol{x}_j\right\|}$$

and substitute it into Eq. (4.6), we have Eq. (4.8). ■

This approach iterates until the maximum number of learning epoch is reached. Notably, the update rule in Eq. (4.8) can be done by matrix operation and vectorization. In other words, the location of all other nodes (i.e., $\boldsymbol{x}_j (j \neq i)$) are simultaneously updated instead of one by one, which accelerates the update process. The stochastic embedding approach is summarized in Algorithm 4-1.

**Algorithm 4-1** Stochastic network embedding – serial computing

**Input:** dissimilarity matrix $\Delta$ $(\delta_{ij})$
**Output:** coordinates of nodes $\boldsymbol{x}_i, i \in \{1, 2, \ldots, N\}$ in the high-dimensional network
1: $T$ – number of epochs
2: initialize $\boldsymbol{x}_i$, learning rate $\{\eta_t\}_{t=1}^{T}$ and neighborhood radius $\{\lambda_t\}_{t=1}^{T}$
3: **for** $t$ =1, 2, …, $T$ **do**
4: create index sequence $\vartheta$ by random permutation of $\{1, 2, \ldots, N\}$
5: **for** $n$ = 1 to $N$ **do**
6: draw $i \in \vartheta$ , fix the $\boldsymbol{x}_i$
7: update using Eq. (4.8)
8: **end for**
9: **end for**

*2) Stochastic network embedding – new parallel computing algorithm*

Algorithm 1 processes one node per iteration and the computational complexity in each iteration is low. However, it takes a long time for Algorithm 1 to find the optimal solution when a large number of machines or production cycles of a machine are involved. Thus, it is imperative to parallelize the algorithm and utilize more computing resources collaboratively. However, the

inherently sequential nature of Algorithm 1 poses significant challenges for parallelization. An algorithm can be parallelized if the computing task can be decomposed and results from individual processors (or computers) are independent of each other. Notably, fixing the location of nodes $a$ or $b$ and updating the location of other nodes will result in different locations of $c$ especially when $a$ and $b$ are far from each other. Thus, Algorithm 1 is difficult to be parallelized.

To address this challenge, one practical solution is to combine stochastic embedding with batch learning, which results in the mini-batch stochastic embedding. Instead of considering one sample at a time, the mini-batch stochastic embedding considers multiple samples (i.e., a mini-batch) simultaneously at each iteration [86]. Assume the total number of data $N$ can be partitioned into $M$ mini-batches, each of the size $m$. Given a randomly selected mini-batch $\gamma_t \subset \{1, 2, \dots, N\}$, the objective function on the mini-batch can be written as $F_{\gamma_t} = \frac{1}{|\gamma_t|}\sum_{i\in\gamma_t} F_i$, where $|\gamma_t| = m$ is the size of a batch. At each iteration, the mini-batch stochastic embedding accumulates sub-gradients with respect to $m$ samples in the batch and calculate the average gradients as:

$$\nabla F_{\gamma_t} = \frac{1}{|\gamma_t|}\sum_{i\in\gamma_t} \nabla F_i \tag{4.9}$$

where $\nabla F_i = \frac{\partial F_i}{\partial \boldsymbol{x}_j}$. As such, each mini-batch generates a single average gradient. At the end of each batch, mini-batch stochastic embedding feeds $\nabla F_{\gamma_t}$ to update the location of all other nodes $\boldsymbol{x}_j$ ($j \neq i$) as:

$$\boldsymbol{x}_j \leftarrow \boldsymbol{x}_j - \eta_t \nabla F_{\gamma_t} \tag{4.10}$$

Therefore, the update rule of Eq. (4.8) can be rewritten as:

$$\begin{cases} \boldsymbol{x}_j \leftarrow \boldsymbol{x}_j - \eta_t \dfrac{1}{|\gamma_t|}\displaystyle\sum_{i\in\gamma_t} \frac{\|\boldsymbol{x}_i - \boldsymbol{x}_j\| - \delta_{ij}}{\|\boldsymbol{x}_i - \boldsymbol{x}_j\|}\left(\boldsymbol{x}_j - \boldsymbol{x}_i\right) & \text{if } \|\boldsymbol{x}_i - \boldsymbol{x}_j\| \le \lambda_t \\ \boldsymbol{x}_j \text{ is not updated} & \text{if } \|\boldsymbol{x}_i - \boldsymbol{x}_j\| > \lambda_t \end{cases} \tag{4.11}$$

It may be noted that Eq. (4.11) is equivalent to Eq. (4.8) when the size of batch $m$ = 1. As the update is executed less frequently, the mini-batch stochastic embedding provides computational benefits.

Further, each mini-batch can be further considered as a combination of multiple subsets, i.e., $\gamma_t = \gamma_{t,1} \cup \gamma_{t,2} \cup, \ldots, \cup \gamma_{t,S}$. Thus, the calculation of sub-gradients in Eq. (4.9) can be further represented as:

$$\nabla F_{\gamma_t} = \frac{1}{|\gamma_t|}\sum_{i\in\gamma_t} \nabla F_i = \frac{1}{|\gamma_t|}\left[\sum_{i\in\gamma_{t,1}} \nabla F_i + \sum_{i\in\gamma_{t,2}} \nabla F_i + \cdots + \sum_{i\in\gamma_{t,S}} \nabla F_i\right] \tag{4.12}$$

As a result, the mini-batch stochastic embedding can be readily implemented in a parallel manner. That is, the mini- batch $\gamma_t$ can be divided and the computation of average gradient can be assigned to $S$ individual processors. Each processor $s$ will compute $\sum_{i\in\gamma_{t,s}} \nabla F_i$ and results from $S$ processors are combined using Eq. (4.12) to return the average gradient with respect to all $m$ samples in the mini-batch. The parallel stochastic network embedding algorithm is summarized in Algorithm 4-2.

**Algorithm 4-2** Stochastic network embedding – parallel computing

```
initialize as Algorithm 4-1 step 1 and 2
randomly partition data into M mini-batch γ_1, γ_2, …, γ_M
start parallel pool with S processors
for t =1, 2, …, T do
    randomly choose a mini-batch γ_t
    partition γ_t into subsets γ_t,1, γ_t,2,…, γ_t,S, each with size m
    for n = 1 to M do {in parallel}
        processor # s gets partition γ_t,s
        solve the sub-problem ∇F_{i∈γ_t,s} on γ_t,s by Algorithm 1
    end for
    average over the batch size m and update x_j
end for
close parallel pool
```

## 4.4. Experiments and Results

In the present investigation, we evaluate the proposed methodology in the development of a large-scale machine-to-machine network, as well as a cycle-to-cycle network for the monitoring of machine conditions. First, we compare the performance of our proposed parallel algorithm over traditional approaches for handling a large machine network. Second, the proposed methodology is implemented for the development of a cycle-to-cycle network customized to one machine. Power profiles are simulated and used in experimental studies to characterize dynamics of the underlying machining process. Our experiments and results are detailed in the following sections.

*A. Optimization of a Large-scale Machine-to-Machine Network*

We first evaluate the efficiency of stochastic network embedding over classic multidimensional scaling using a simulated large-scale machine network. Here, both algorithms are implemented using serial computing. Then, the proposed parallel-computing algorithm is compared with serial computing for network embedding.

*1) Stochastic network embedding vs. classic multidimensional scaling*

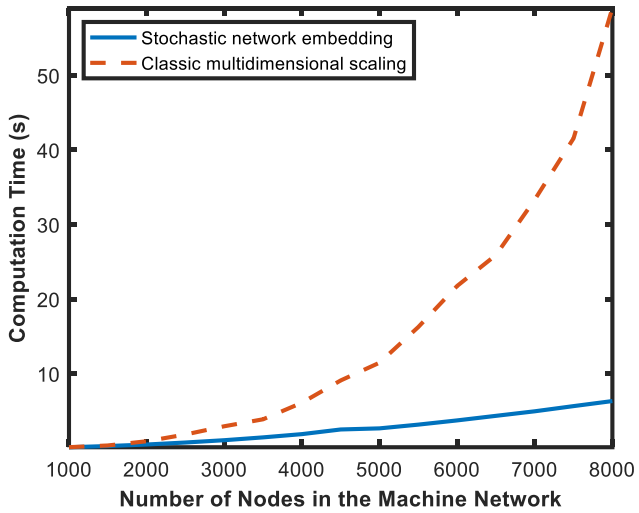



Figure 4-9. The computation time of stochastic network embedding and classic multidimensional scaling in a single computing unit.

As shown in Fig. 4-9, the stochastic approach is more efficient than classic multidimensional scaling, especially when a large number of machines are involved in the network. Notably, the computation time of the stochastic network embedding increases slowly with more machines in the network. However, the computation time of classic multidimensional scaling increases significantly with respect to the size of the network., due to its $O(N^3)$ computational complexity. It may be noted that the stochastic network embedding is 10-fold faster than classic multidimensional scaling to embed a network with 8000 nodes.

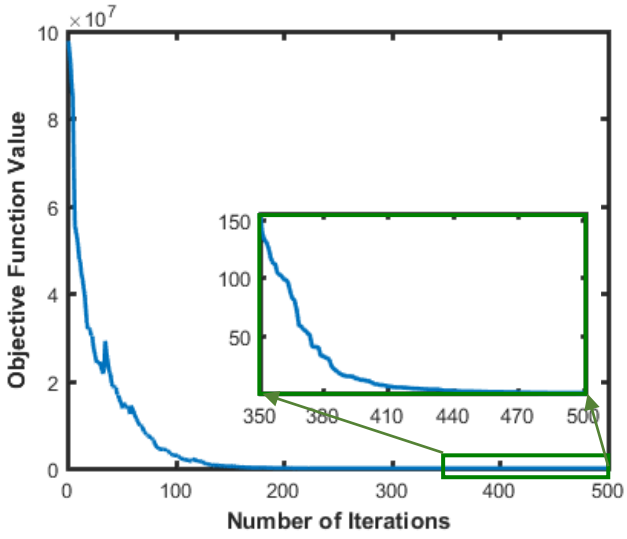



Figure 4-10. The convergence of stochastic network embedding.

As shown in Algorithm 4-1, larger number of epochs will enhance the convergence, but result in a longer computation time. As shown in Fig. 4-10, the value of objective function quickly drops with the increase of iterations (i.e., each visit of a node) and quickly converges. This is because all nodes $\boldsymbol{x}_j$ $(j \neq i)$ are updated when the algorithm visits node $\boldsymbol{x}_i$, rather than only update $\boldsymbol{x}_i$ at each iteration (see Algorithm 4-1).

*2) Parallel vs. serial computing – stochastic network embedding*

Further, the performance of proposed parallel algorithm (Algorithm 4-2) is compared with serial computing (Algorithm 4-1). We increase the size of the simulated machine network from 18,000 to 180,000 and record the computation time of both algorithms for each machine network. The entire set of machines is divided and assigned to 12 CPU cores. As shown in Fig. 4-11, both algorithms achieve similar performance when less than 20000 machines are included in the network. However, the computation time of serial computing increases significantly when a large number of machines are included. On the contrary, the computation time of the parallel framework increases much slower than serial computing. It may be noted that the gap between two curves increases dramatically with the number of machines. When 70000 machines are embedded, the serial approach is 400s slower than the parallel computing. When the number of machines reaches 180,000, serial computing is >2500s slower than the parallel framework.

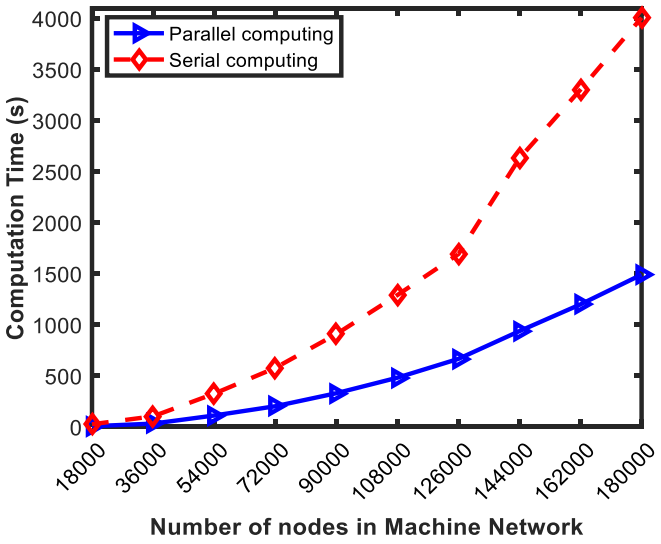



Figure 4-11. Performance comparison of parallel (12 cores) and serial computing.

*B. Optimization of a Cycle-to-Cycle Network Customized to One Machine*

The proposed methodology shows great potentials to be used for continuous monitoring of cycle-to-cycle variations of a single machine, e.g., a Hass turning center. As shown in Fig. 4-12, raw materials (e.g., metal bars) are automatically fed into the turning center by the Haas Bar Feeder. The turning center processes raw materials according to pre-programmed specifications, and one sample is produced in each turning cycle. Notably, the turning center is performing the same machining job for each discrete sample repeatedly. As such, power profiles will be near-periodically varying from one sample to the next one. The variations of power profiles can be potentially used for quality monitoring and control of the machining process.

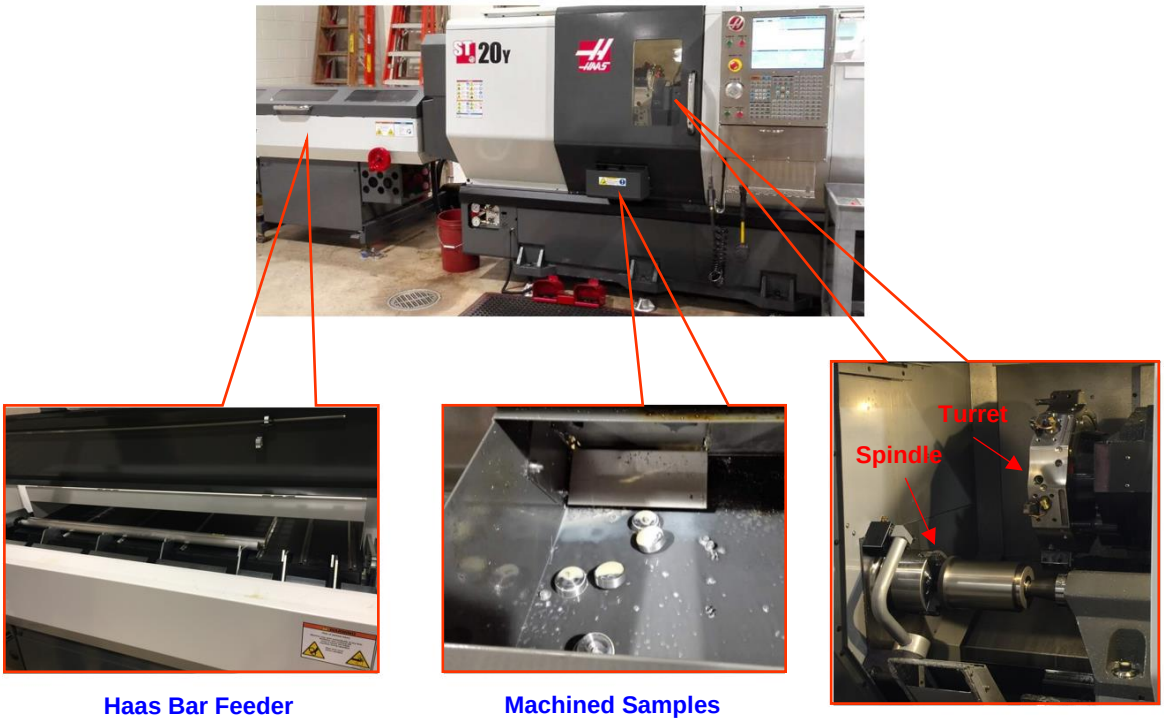



Figure 4-12. A Haas ST-20Y turning center at the Pennsylvania State University.

Here, we focus on the monitoring of different conditions of the cutting tool. It may be noted that a worn cutting tool (see Fig. 4-13) will result in a high surface roughness of the machined samples, which will further impact the quality of products and increase the re-work rate. Hence, it is imperative to monitor the condition of cutting tool and replace it in the early stage of failure. Notably, tool wear will result in an elevated power consumption in the cutting phase: If the condition of a cutter degrades during the cutting, the corresponding power consumption will significantly increase (red curve in Fig. 4-14). This happens when the hardness of raw material is high, and the failure of cutting tool occurs during cutting. In addition, we considered a common problem encountered in the manufacturing process, i.e., chatter. Chatter is the self-excited vibration that will impact the surface finish of parts and even the life of spindle. The onset of chatter usually results in a larger variation in the power consumption (blue curve in Fig. 4-14).


Figure 4-13. A worn (left) and a new (right) coated tungsten carbide cutting tools.

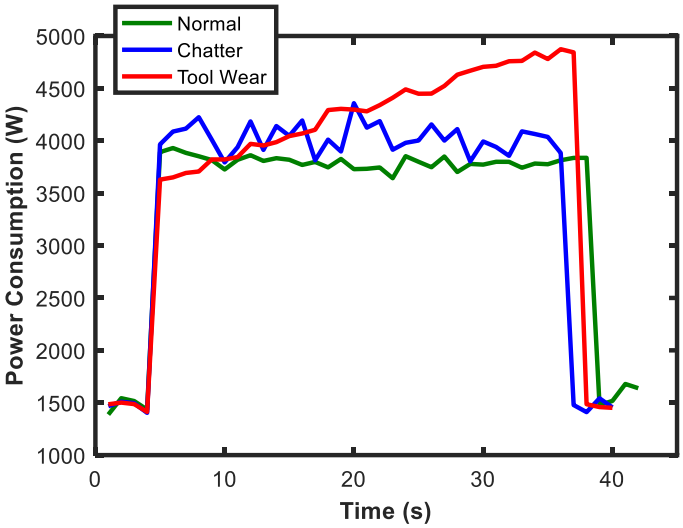



Figure 4-14. Morphology of power profiles in the cutting phase for three conditions in the turning process.

Hence, we seek to investigate cycle-to-cycle variations of machine signatures using the power profile for each discrete part produced in the turning process. A total of 6000 power profiles of the turning center are simulated under each of the three conditions. Note that 3000 are normal cycles, 1000 cycles are with the onset of chatter, 2000 are with tool wear. First, dissimilarities among the power profiles are measured using the dynamic time warping approach. Then, the developed parallel-computing approach for network embedding is used to optimally project the 6000 cycles into a 3-D network. Notably, three distinct clusters of nodes are shown in Fig. 4-15. Nodes corresponding to normal condition (green) and the onset of chatter (blue) are concentrated densely, while those associated with tool wear are distributed sparsely. This is because the morphology of power consumption in normal cycles are similar to each other, while the severity of tool wear can be vastly different depending on the cutting speed, fed rate and types of materials. As each node corresponds to one operation cycle, the distribution of network nodes reveals the variation of machine conditions in the turning process. Based on the network structure, predictive maintenance can be scheduled. For example, when a new cycle is projected closely to the red cluster, i.e., tool wear, it is imperative to order a new cutting tool for replacement.

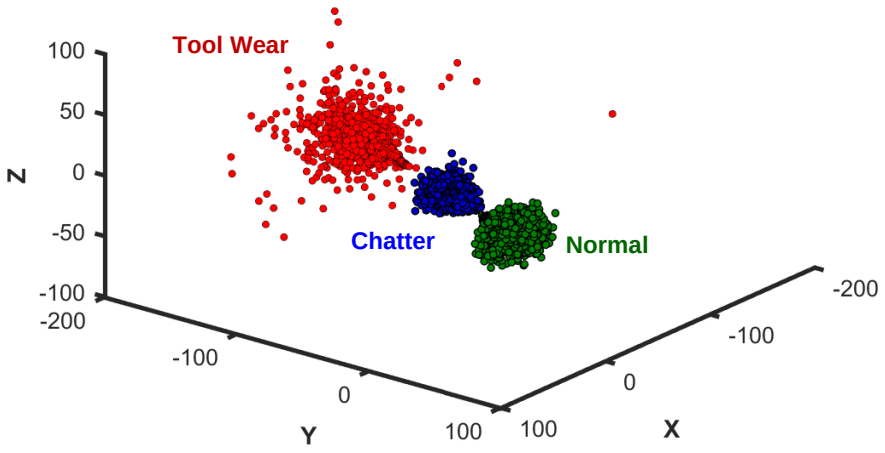



Figure 4-15. Visualization of a 3-D network (nodes) embedded from 6000 turning cycles with 3 different conditions (Normal, Chatter and Tool wear).

## 4.5. Conclusion and Discussions

Recent advances in sensing and communication technology open up the new era of IIoT. The IIoT enables smart and connected monitoring of machine conditions and generates large amounts of data. Analyzing these data, known as machine signatures, provides a great opportunity for machine condition monitoring and control. However, realizing the full potential of IIoT depends, to a great extent on the development of new methodologies for big data analytics. Existing approaches are limited in their ability to effectively extract pertinent knowledge about manufacturing operations from the large volume of IIoT data of networked machines. There is a pressing need to develop an effective and efficient algorithm to leverage big data in the IIoT for machine condition monitoring and fault diagnosis.

In this study, a new parallel algorithm is developed for large-scale IIoT machine information processing, network modeling, and condition monitoring. First, dissimilarities between machine signatures are characterized and quantified by the dynamic time warping algorithm. Second, a stochastic approach is introduced to optimally embed machine signatures into a network that preserves the warping distance. The location of network nodes reveals underlying machine conditions and can be used as features for predictive modeling and process diagnosis. The stochastic network embedding, nevertheless, is computationally expensive when large numbers of machines are involved in the IIoT network. Therefore, a parallel algorithm of stochastic network embedding is further developed to harness the power of multiple processors for efficient network

modeling of a large-scale IIoT machines. It may be noted that the proposed approach provides a great opportunity for:

a) *Dimensionality reduction*: The network approach represents each power profile as a node in the network based on the pairwise dissimilarity measures, which greatly reduced the dimensionality of data and thereby identifies the "best data" to represent the machine condition.
b) *Visual analytics*: The proposed approach provides a network representation of large amounts of machines and further groups machines with similar conditions into homogeneous clusters. This enables visual analytics of the data for condition monitoring of machines. For example, we can pinpoint each profile from a machine in the network clusters for monitoring or classification purposes.
c) *High-performance computing*: As a vertical step to improve computational efficiency, we developed the stochastic network approach and parallelized the algorithm into multiple processors to reduce computational complexity. In this way, the "best data" can be efficiently identified for machine condition monitoring and performance improvement.

Notably, most traditional methods focus on the conformance to reference signatures (i.e., "standard" or "normal" ones). However, our method performs a pairwise comparison of signatures. It may be noted that a dissimilarity matrix (consists of the dissimilarity between each pair of signatures) is obtained from the pairwise comparison, rather than only a column in reference comparison. For traditional reference comparison, the workload of computation is low, and it is easy to implement. The difference against the reference can be directly used as an indicator to determine if the current signature is normal or not. However, there is a need to empirically and/or statistically establish a "normal" signature from a historical record of signatures. However, even under the "normal" condition, the morphology of machine signatures may vary. Thus, it is difficult to use one signature to represent the "normal" condition. On the contrary, there is no need for our pairwise comparison to establish a "normal" signature. Instead, we develop a data-driven network approach that leverages the pairwise dissimilarity information to automatically group large amounts of signatures into homogeneous clusters. As such, the proposed network approach provides a better representation of information in the data and further offers a great opportunity for visual analytics of machine conditions. However, the workload of pairwise computation is large. Also, it is challenging to extract information from the resulted dissimilarity matrix. In fact,

this motivates the development of an efficient and effective approach (i.e., fast parallel stochastic network embedding) to fully utilize the pairwise comparison results for machine condition monitoring.

This paper presents a pilot study on the development of cycle-to-cycle network for the purpose of quality monitoring. In other words, one machine repeatedly manufactures the same type of product in large quantities (*the high volume, low mix scenario*). Also, the cycle-to-cycle network can be established from a large number of profiles pertinent to different types of parts. For example, if there are two different kinds of parts, then power profiles from the same part tend to have a higher level of similarity than those from a different part. This will lead to another application of product classification. That is, network models leverage pairwise dissimilarity measures to group profiles from the same part into a homogeneous cluster. Node location in the network reveals useful information on the type of parts in production. Network visualization will provide the categorization of parts, the evaluation of energy consumption, and further help production planning. For *the low volume high mix scenario*, network models are not for quality monitoring but rather can be used for potential applications to classify part or process characteristics (e.g., types of materials, machining procedures, and specific tools used). For example, the cycle-to-cycle network on the same machine can be used for product categorization and energy monitoring. The machine-to-machine network can help extract useful information on machine utilization, power usage, and condition monitoring, which will help further optimize factory operations, reduce equipment downtime and maintenance costs.

To the best of our knowledge and review, little has been done to develop new methodologies of parallel computing and network analytics for quality monitoring and control of large-scale IIoT machines. The proposed methodology has a strong potential to promote existing manufacturing practices towards globalized production optimization and management. Furthermore, it provides a great opportunity to optimize the energy consumption and realize green and sustainable manufacturing. In addition, the parallel-computing approach for network optimization can be potentially applicable to not only manufacturing but also other research areas such as disease diagnosis and prognosis, infection prevention and control, supply chain optimization and service networks.

# Chapter 5: Parallel Computing for Multi-level Modeling and Analysis of the Large-scale Internet of Health Things


## Abstract

Heart disease remains the leading cause of death and disability, thereby causing high healthcare costs. Optimal management of heart health hinges on the detection of early signs of disease progression. In this study, we proposed a new IoT technology, namely, the Internet of Hearts (IoH), to empower smart and connected heart health. However, real-time monitoring of a large number of patients generates big data. Early diagnosis calls for the development of a new methodology to efficiently handle big data and extract pertinent information about the disease-altered cardiac activity. Therefore, we designed and developed a new approach of dynamic network modeling and parallel computing for multi-level cardiac monitoring: patient-to-patient variations in the population level and beat-to-beat dynamics in the individual patient level. Specifically, dissimilarities among heartbeats or patients are firstly characterized. Then, a stochastic learning approach is developed to optimize the embedding of heartbeats or patients into a dynamic network that preserves the dissimilarity matrix. Furthermore, we developed a parallel computing algorithm of the stochastic network embedding to improve computational efficiency and effectiveness. Finally, we proposed the network control charts that harness network features for online monitoring and change detection of cardiac processes. Experimental results show the developed parallel computing and network modeling methods have strong potentials for realizing smart connected healthcare in large-scale IoH contexts.

## 5.1. Introduction

According to American Heart Association, cardiac diseases remain the No. 1 killer worldwide. Over 30 percent of global deaths are attributable to cardiac diseases. In the US, cardiac diseases claimed approximately 1 in 7 deaths each year [87]. Considering costs of medications, services, and lost productivity, the annual cost from cardiac diseases amounts to over $300 billion in the US [87]. Optimal heart health management calls for early detection of disease signs and timely treatment. For example, the blockage of coronary arteries will reduce oxygen and nutrients supply to the myocardium and result in ischemia. In the early stage, effects of ischemia on heart muscle cells are reversible. However, the cells die and trigger heart attack when the episode of ischemia lasts for a longer time. It is estimated that a 30-minute delay in treatment will increase the risk of 1-year mortality by 7.5% [88]. Hence, advanced cardiac monitoring and early-stage diagnosis are critical for the prevention and control of cardiac diseases.

More importantly, two-third of deaths caused by cardiac diseases occur in out-of-hospital settings [89]. There is an urgent need to develop smart and connected systems for anytime, anywhere cardiac monitoring and delivery of life-saving therapies. In the past decade, advances in wearable sensing and mobile technology make it possible and affordable to remotely collect patients' cardiac data. Caregivers are able to gain access to patients' complete information to match the right patient with the right service at the right time. Meanwhile, patients including virtually every age, gender, and race segments are increasingly motivated to participate in their own care [89]. This significantly diminishes care disparities in the population. In the past few years, the Internet of Things (IoT) has been hailed as a revolution in the healthcare industry. The IoT deploys a multitude of wireless sensors, mobile communication units, and physical objects (e.g., patients, cardiologists and medical devices) in an Internet-like infrastructure [65]. Such a cyber-physical system provides an unprecedented opportunity for revolutionizing the cardiac care. For example, IoT-enabled smart sensors provide more integrated solutions for data collection and device-to-analytics data flow. This, in turn, enables an information-intensive environment for disease diagnosis and prognosis. Also, ubiquitous communication networks in the IoT facilitates anywhere, anytime data exchange and care delivery between patients and cardiologists. This helps to boost patient engagement and satisfaction.

In this study, we propose a new IoT-based cardiac monitoring system, namely, the Internet of Hearts (IoH) to integrate wearable sensing, wireless communication with cloud computing for continuous cardiac monitoring and early disease diagnosis. The IoH enables 24/7 data collection from hundreds of thousands of patients. Collected data are transmitted seamlessly to the cloud, where tremendous computing power is utilized for data processing and storage. Cardiologists are able to access patients' information anytime and anywhere and deliver on-demand health services in a timely manner. The IoH-based cardiac care essentially integrates patients and cardiologists into a physical network, which enables seamless information exchange, early diagnosis, and timely response to medical emergencies. However, large-scale IoH leads to the accumulation of big data. Continuous monitoring of an individual patient generates a large volume of data when performed in hours, days, months and years. Also, the number of patients in the IoH is growing. The goal of IoH is to embody patients all across the world to empower smart telehealth and preventive medicine. Such big data not only provides a wealth of opportunities to promote patient-centered, personalized cardiac care but also poses significant challenges for analytics and management. Conventional cardiac care system lacks the ability to handle the big data and extract information pertinent to the underlying cardiac disorders. Realizing the full potential of IoH depends, however, to a great extent on big data analytics for optimal decision making and system control.

Therefore, a new dynamic network model is developed for big data analytics in the large contexts of IoH. Specifically, we first measure the similarity among the sequence of cardiac diastoles and systoles. This can be used to characterize 1) beat-to-beat dynamics of a single patient or 2) stochastic behaviors of cardiac activity in the patient-to-patient scale. Then, cardiac signals are optimally embedded into a high-dimensional network. The distance between network nodes preserves the dissimilarity of corresponding signals. As such, the location of network nodes reveals diagnostic features pertinent to patients' cardiac conditions and the change of network structure reveals the stochastic cardiac dynamics. However, the computation time will be prohibitive when large amounts of cycles or patients are involved. Hence, a parallel computing scheme is developed to scale up the network embedding and harness the computing power of multiple processors simultaneously. Notably, such a dynamic network model resembles as a virtual counterpart of the physical IoH network in reality. The physical network provides data to the virtual network, and results of data analysis from the virtual network are fed back to the physical network for medical decision making. Our contributions in the present investigation are highlighted as follows:

1) We propose the IoH, a new IoT-based cardiac care system, which integrates IoT sensing, wireless communication, big data analytics, and medical decision making to realize smart and connected cardiac health management.

2) We design a new approach of dynamic network modeling for multi-level cardiac monitoring: i) a beat-to-beat network to create personalized baselines and reveal progressive cardiac variations of a specific patient; ii) a patient-to-patient network to delineate different types of cardiac conditions among various patients in the IoH for population-level monitoring.

3) We developed a parallelized stochastic learning approach to harness the power of multiple processors for fast network optimization. This enables the inclusion of a large patient population in the IoH for real-time monitoring and disease management.

The remainder of this paper is organized as follows: Section 5.2 presents the details of system design of the IoH. Section 5.3 will present the methodology of proposed dynamic network models. Section 5.4 includes simulation and real-world studies for the evaluation and validation of proposed methods. Section 5.5 concludes this study.

## 5.2. Research Background

### *A. Cardiac Monitoring and Pattern Recognition*

Optimal management of heart health depends to a great extent on the effective cardiac monitoring and early recognition of disease patterns. For over 100 years, the electrocardiogram (ECG) system has been used for cardiac monitoring. One-lead ECG captures a one-dimensional view of temporal cardiac dynamics, and it has been available in home-based devices, sports use, and emergency care. Multi-lead ECG systems, on the other hand, provide multi-directional views of space-time cardiac activity. Among which, the 12-lead ECG system is considered as the Gold Standard and has proven its value for the diagnosis of acute cardiac events. For example, Elmberg *et al*. [90] measured QRS prolongation in 12-lead ECGs to quantify the severity of ischemia. Meo *et al*. [91] characterized the variability of f-wave amplitude from 12-lead ECGs to predict the catheter ablation outcome of atrial fibrillation. Perlman *et al*. [92] extracted features from QRS complex in 12-lead ECGs and developed a tree scheme for the classification of supraventricular tachycardia. Cheng *et al.* [93] proposed a heterogeneous recurrence model to characterize heart rate variability from ECG signals for the identification of obstructive sleep apnea. Yang *et al*. [37] designed customized wavelet functions to extract fiducial patterns of ECG signals for the

identification of atrial fibrillations. Chen and Yang [94] developed multiscale recurrence analysis with a self-organizing map for the quality control of 12-lead ECG signals.

In the past decade, the use 3-lead vectorcardiogram (VCG) has regained interest in cardiac monitoring and diagnosis. Instead of measuring scalar amplitudes (i.e., ECG), VCG observes cardiac activity as a vector in three orthogonal planes of the body (i.e., frontal, transverse, and sagittal). Yang and Leonelli [95] developed a self-organizing network for the characterization of pattern dissimilarities among QRS loops in VCG signals. Man *et al*. [96] analyzed features from VCG (e.g., QRS-T angles) to improve the diagnostic accuracy for left ventricular hypertrophy. In our previous works, a representation approach was developed for the dynamic display of VCG trajectories and characterization of spatiotemporal pathological patterns [34]. Also, recurrence [35] and graphical models [81] were proposed for the detection and differentiation of myocardial infarctions using VCG signals.

Due in large part to advances in sensing and mobile technology, cardiac care has been transformed from reactive and hospital-centered to proactive, preventive and patient-centered. In particular, connected cardiac health has gained increasing interests to improve patient outcomes and reduce costs. As opposed to conventional cardiac care, highly-connected care ecosystem consists of doctors, patients, databases and other entities. Doctors are able to gain rapid access to all relevant evidence pertaining to the patient and bring them to the point of care anywhere and anytime. Also, connected cardiac health enables individuals to participate in their own care. The smart connected healthcare provides a data-rich environment for efficient medical decision making, enables personalized patient-centered care, and diminishes care disparities.

### *B. The Internet of Health Things*

One important enabling technology for connected cardiac health is the IoT. The IoT integrates a multitude of sensors into a large-scale information network for continuous monitoring of interconnected “Things”. In healthcare, “Things” may include patients, caregivers, sensors, medical devices, medications, data, computers, and virtual infrastructures such as cloud computing. Notably, patients are of greater importance vis-à-vis other physical objects in the internet of health things system. Therefore, a multitude of IoT-enabled smart sensors are deployed for patient-centered monitoring and transmit data to the cloud. Data in the IoT system not only include historical data (e.g., electronic health records (EHR)) collected from a large number of

patients but also online data from the real-time monitoring. Realizing the full potential of internet of health things system calls for the development of new methodologies for data-driven information processing and condition monitoring in the large-scale IoT contexts.

In the literature, there are increasing interests to bring the IoT to the context of healthcare. For example, Pasluosta *et al*. [31] presented an IoT-enabled solution for the diagnosis and treatment of Parkinson's disease. The authors discussed the combination of wearable devices and the IoT technology for patient monitoring and data collection. Also, they highlighted the challenges in transforming IoT data to knowledge and discussed the potential of integrating artificial intelligence for decision support. Al-Taee *et al.* [32] developed a mobile health platform to incorporate humanoid robots for the management of diabetes in children. Based on the IoT technology, the robot connected to medical sensors and were also linked with a web-centric disease management hub. Such a platform helped collect patients' data and extracted simple features such as blood glucose level, and improved the engagement of patients with caregivers. Yasin *et al*. [97] designed a hardware solution with high energy efficiency and security for the IoT-based monitoring of ventricular arrhythmia. The design reduced power consumption of hardware by incorporating algorithmic enhancement that required less memory and clock speed. Also, a logic locking technique was implemented to protect the security-critical components in the hardware. Yang *et al.* [98] developed an IoT platform to improve in-home healthcare services. The critical components included an intelligent medicine box, an RFID-based pharmaceutical packaging, and a flexible and wearable sensing device. The platform enabled the fusion of IoT devices with in-home healthcare services for an improved user experience and service efficiency. More IoT-enabled healthcare applications can be found in a comprehensive survey provided by Islam *et al*. [33].

Although these pioneering studies present various IoT-enabled solutions to improve health management, most of them focus on the conceptual design of IoT infrastructures or hardware development. Very little has been done to leverage sensing data from a large-scale internet of health things to develop new methods and tools for patient condition monitoring, and disease diagnostics and prognostics. Our previous studies [84, 99] preliminarily discussed the potential of leveraging IoT sensing for the monitoring of cardiac conditions. In this paper, we propose the new IoH system to integrate network modeling and parallel computing for multi-level cardiac monitoring in large-scale IoT contexts.

## 5.3. The Internet of Hearts

### *A. IoH Architecture*

Figure 5-1. An Illustration of the Large-scale Internet of Hearts.

As shown in Fig. 5-1, the IoH embodies a number of networked components, including ECG monitors, mobile devices, computing units, databases, patients, caregivers, and hospitals. Both inpatients and outpatients are monitored 24/7 and data are seamlessly transmitted to the cloud server via wireless networks. Analytics models are running on the cloud server to uncover hidden cardiac patterns and detect early signs of cardiac events. Cardiologists are able to access patients' data, review analytical results and communicate with patients and other cardiologists anytime and anywhere. In addition, emergency centers and hospitals are integral components of the IoH. If a patient's condition is highly risky, medical intervention will be delivered in a timely manner.

### *B. IoT Sensing and the Physical Network of Patients*

The IoH provides ubiquitous cardiac care solutions by connecting patients and caregivers using the IoT technology. Besides traditional ECG monitors that routinely used in hospitals, wearable devices are becoming increasingly available to monitor cardiac conditions of outpatients. Besides simple measures (such as blood pressure and heart rate), high-resolution ECG data can be collected using wearable devices. However, the low durable value remains an important challenge for wearables. It may be noted that only 80% people continue using their wearable devices regularly

after 3 months and less than 50% remain faithful to their wearables after 1.5 years [100]. Hence, unobtrusive and energy-efficient sensing devices are needed to increase the durable values of wearable ECG devices.

In our previous research [65], a coin-sized board was designed that integrated 3-lead ECG sensing with a wireless communication module. The thin two-layer PCB can be easily implanted into the clothes, and collected data can be transmitted to the smartphone via Bluetooth. Recently, advances in materials science make it possible to develop soft sensors that directly mounted on the skin. For example, Jang *et al*. [101] developed a thin electrophysiological sensor with stretchable properties that can fold, twist, and wrap around the complex geometry of the skin. This will significantly improve the durable value of wearable ECG devices and make it possible for continuous monitoring of patients' cardiac conditions.

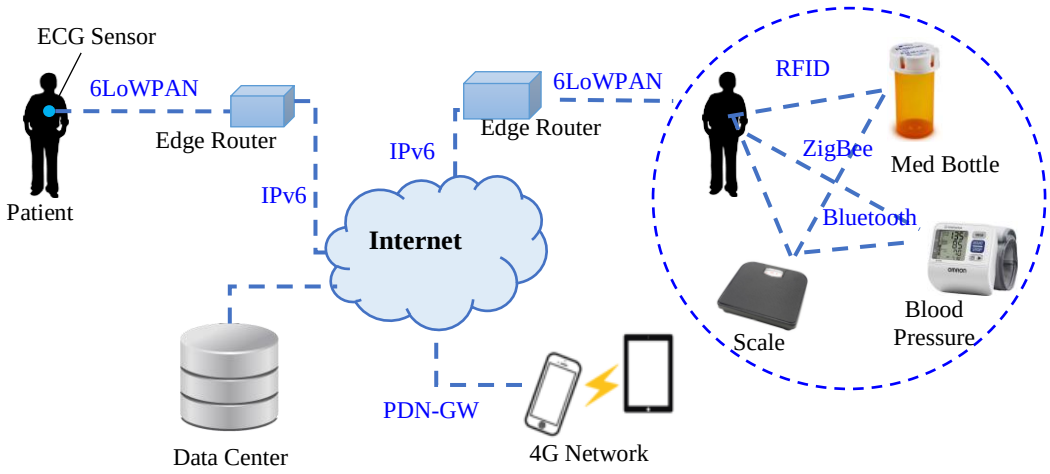



Figure 5-2. Wireless communication protocols in the IoH.

The IoH can adopt an IP-based network to allow one-to-one and one-to-many communications between physical subjects (e.g., patients and caregivers). Each subject is assigned with a unique IP address, and the emerging IPv6-based low-power wireless personal area network (6LoWPAN) technology is used for communication and data exchange. Unlike traditional 32-bit addresses in IPv4, the 128-bit IPv6 supports a larger address space ($2^{128}$) that is capable of handling every IoH device in the world. Also, 6LoWPAN provides an effective solution to integrate individual LoWPANs (with various protocols like ZigBee, EnOcean, and SNAP) into the standard Internet-level network. The large packets from the Internet are stripped down and handed over to LoWPANs, whereas small LoWPAN packets are reformed and passed back to the standard

internet-level network. As such, the IoH is additive, which enables to incorporate small-scale cardiac sensing networks into the large-scale Internet-of-Things system. In addition, the Internet enables the connection with the 4G mobile network, which extends the communication between LoWPANs to a broader scale. Fig. 5-2 shows the wireless communication protocols in the IoH.

### *C. The Virtual Network of Patients*

The P2P network in the IoH communicates large volumes of data containing rich information pertinent to patients' cardiac dynamics. The availability of big data not only provides a great opportunity for patients' condition monitoring but also calls for the development of new information-processing technology. Early detection of cardiovascular diseases hinges on effective and efficient analysis of cardiovascular big data to extract useful information from individual patients as well as the large-scale patient population. However, it poses significant challenges to direct analyze the physical patient network in the real world due to its high complexity and heterogeneity.

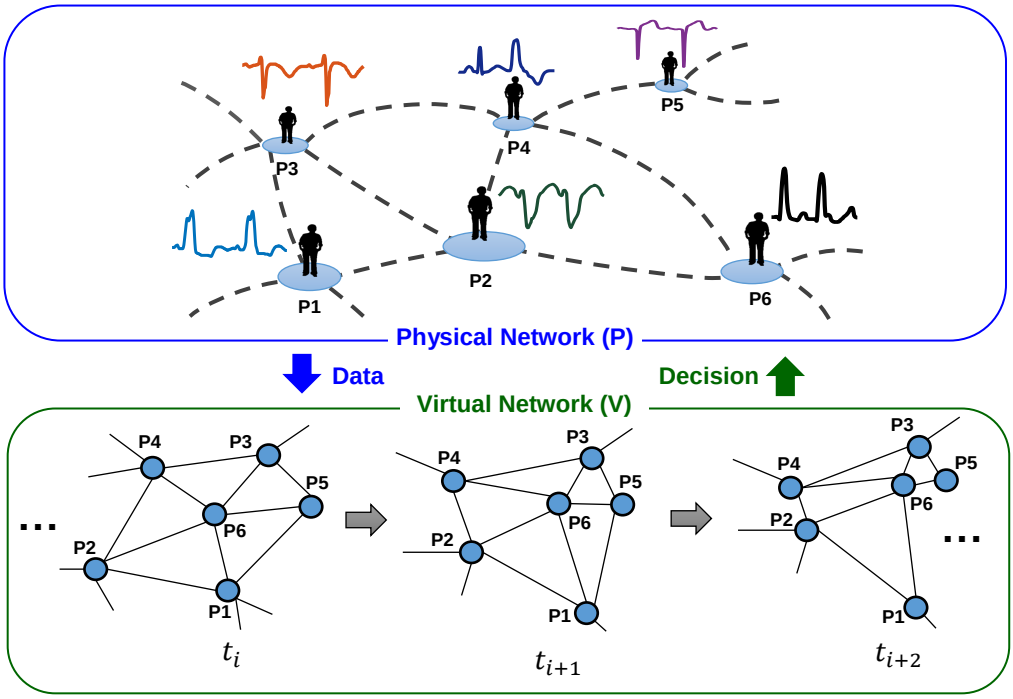



Figure 5-3. Physical network to virtual network (P2V) communication in the IoH.

Therefore, we construct a virtual network from the physical patient network for data-driven information processing (see Fig. 5-3). Data in the physical network (e.g., an ensemble ECG of one patient or an ECG beat) are represented as nodes in the virtual network. Time-evolving cardiac

conditions are then characterized by structural variations of the virtual network. Analysis performed in the dynamic virtual network reveals hidden patterns in patients' electrical activity, which facilitates early medical decision making and care delivery in the physical network. Notably, the virtual network essentially serves as a dual for the primal physical network, which facilitates the efficient handling of cardiac big data and effective information processing and decision making. The physical to virtual (P2V, and vice versa) interaction forms a feedback control system to leverage big data to improve the cardiac care service.

### *D. Multi-level Networks Modeling and Analysis of Beat-to-beat (B2B) Dynamics and Patient-to-patient (P2P) variations.*

Notably, cardiac electrical activity is stochastic, which poses significant challenges for information processing and decision making. The stochastic behavior of the cardiac electrical activity consists of two aspects: dynamics within a patient and variations between patients. On the one hand, beat-to-beat cardiac activity within a patient demonstrates temporal dynamics. As shown in Fig. 5-4a, the amplitude of the $4^{th}$ beat of the ECG signal is smaller than the first three, so as the $8^{th}$ beat. Furthermore, the $6^{th}$ beat shows a significant S wave and an elevated T wave. Moreover, apparent variability can be identified even among those beats that look similar, for example, beat #1, #2, #3, #9, and #10. The stochastic behavior of cardiac activity for an individual patient is critical to the identification of arrhythmic events. Taking consideration of historical variabilities in cardiac activity is conducive to the delivery of personalized treatment planning. On the other hand, there exist variations between patients. As shown in Fig. 5-4b, the heart rate differs significantly among patients. Patient P1, P3, and P6 have only two ECG beats, but the others have 2.5–3 beats within 2 seconds. More importantly, the morphology of these ECG signals shows significant dissimilarities. Patient P3 shows inverted T wave (i.e., T wave is pointing downward instead of upward). P3 has an abnormal wave before the onsite of Q wave, and the R peak of P6 is notched. Notably, between-patient stochastic behaviors are closely pertinent to the disease-altered cardiac patterns. The detection and differentiation of cardiac diseases hinge on the effective characterization of both beat-to-beat dynamics within a patient and patient-to-patient variations of the patient population.

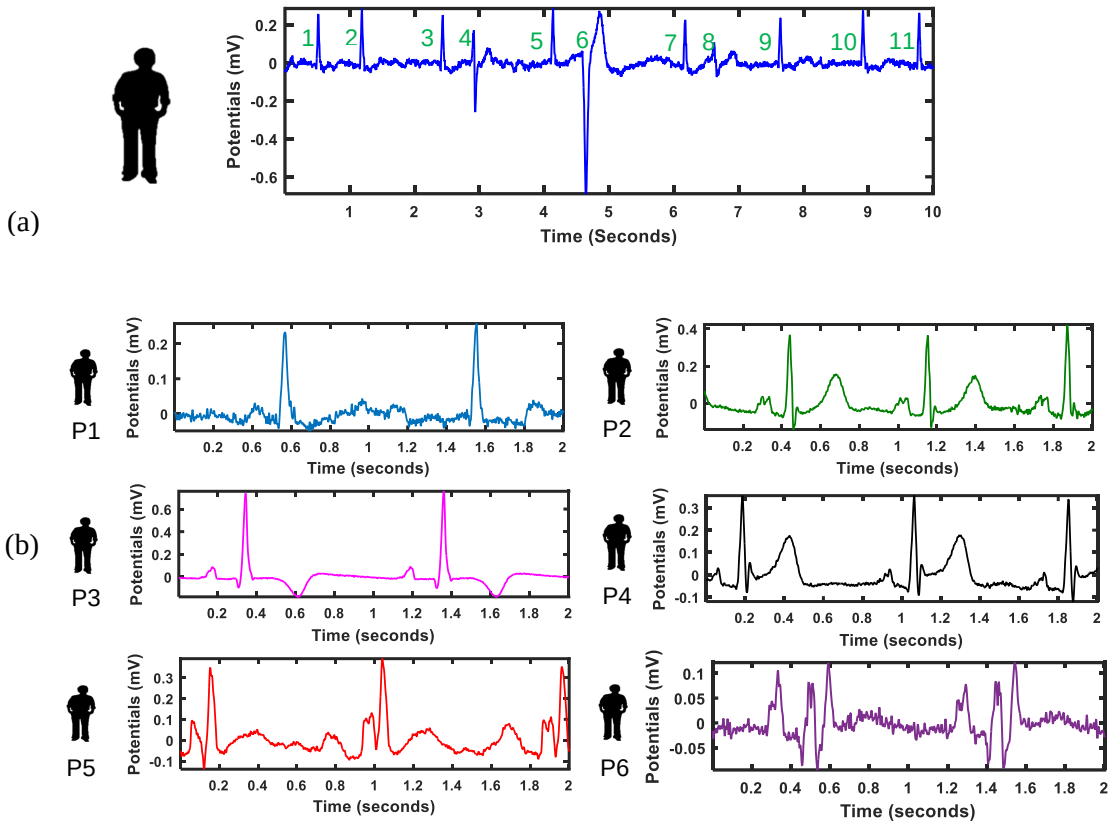



Figure 5-4. (a) ECG dynamics within a patient (cycle-to-cycle) and (b) ECG variations among multiple patients (patient-to-patient).

Therefore, the virtual network of IoH is multi-level, which consists of a patient-to-patient network and a beat-to-beat network (see Fig. 5-5). In the patient-to-patient network, an ensemble ECG/VCG beat from one patient is embedded as a network node, which represents the aggregated information of that patient. The patient-to-patient network characterizes the disease-altered cardiac patterns among large numbers of patients. This provides a great opportunity for the accurate detection and differentiation of cardiovascular diseases at the early stage. The beat-to-beat network, on the other hand, embeds each ECG/VCG beat as a network node, which represents the detailed information of a single patient. The beat-to-beat network depicts the dynamics among successive beats of one patient, which enables the delivery of precise and personalized therapies based on the long-term cardiac monitoring.

However, the IoH connects large numbers of patients and generates big data. It remains significant challenges to optimize the virtual network in the large-scale IoH contexts. Realizing the smart and connected cardiac care using IoH depends on the capability of big data analytics. Existing approaches (e.g., [5–15] discussed in Section II A) are limited in their ability to scale up

and handle the large-scale IoH data. Little has been done to effectively and efficiently characterize cardiac dynamics and detect early disease patterns from large numbers of patients. In this research, we will address this challenge by introducing a new network analytics scheme in the large-scale IoH contexts.

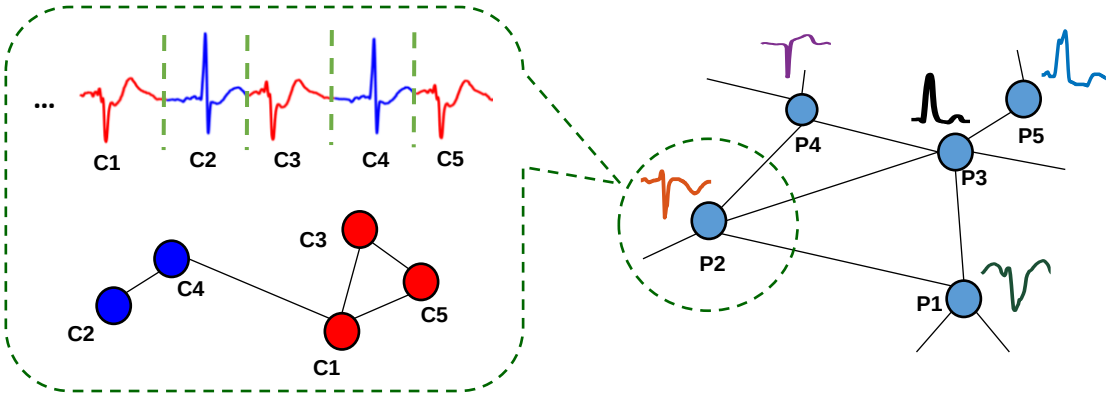



Figure 5-5. Cycle-to-cycle network (left) and patient-to-patient network (right) in the IoH.

## 5.4. Network Analytics of the Internet of Hearts

In the IoH, a virtual network is optimized to handle cardiac big data generated from the physical patient network. As shown in Fig. 5-6, when a new cycle is presented, it is first compared against historical recordings of the patient (for a beat-to-beat network), or other patients (for the patient-to-patient network) in databases. To this end, a symmetric dissimilarity matrix is generated, which is then optimally embedded as a large-scale network. Each signal is represented as a network node and the distance between nodes, as designed, preserves the dissimilarity of corresponding cardiac signals. The network structure reveals important information pertinent to patients' cardiac dynamics. Features extracted from the network are used to for automated diagnosis and prognosis of patients' cardiac conditions.

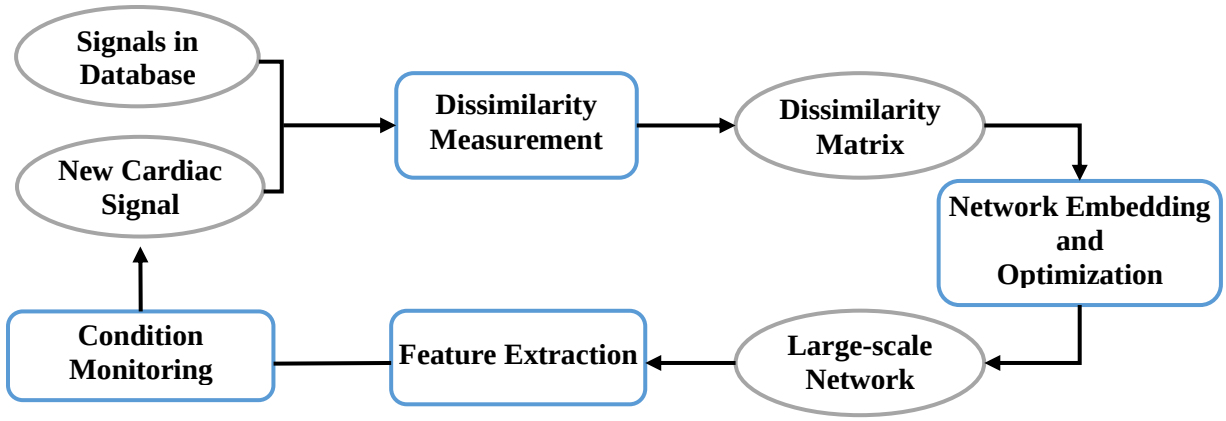



Figure 5-6. Flowchart of proposed big data analytics scheme in the virtual network of IoH.

*A. Dissimilarity Measure*

As discussed above, disease-altered cardiac electrical activity leads to variations in cardiac signals. Quantification of such variations will provide a great opportunity for the identification of cardiac diseases. It is worth mentioning that there are various ways to measure the pattern dissimilarity such as correlation, mutual information, and dynamic spatiotemporal warping. These measures can also be computed in a variety of signal representations such as time domain, frequency domain, state space domain, and time-frequency domain. Note that the Pearson's correlation coefficient quantifies linear interdependence between two signals. Mutual information is derived from information theory to determine how similar the joint distribution of two signals with respect to the product of marginal distributions. Given two 3-D VCG signals $\overrightarrow{\boldsymbol{v}_1}$ and $\overrightarrow{\boldsymbol{v}_2}$, the mutual information is computed as:

$$MI(\overrightarrow{\boldsymbol{v}_1}, \overrightarrow{\boldsymbol{v}_2}) = \sum_{\overrightarrow{\boldsymbol{v}_1}} \sum_{\overrightarrow{\boldsymbol{v}_2}} P(\overrightarrow{\boldsymbol{v}_1}, \overrightarrow{\boldsymbol{v}_2}) \log \left( \frac{P(\overrightarrow{\boldsymbol{v}_1}, \overrightarrow{\boldsymbol{v}_2})}{P(\overrightarrow{\boldsymbol{v}_1}) P(\overrightarrow{\boldsymbol{v}_2})} \right) \tag{5.1}$$

In addition, distance measurements (e.g., Euclidean and Mahalanobis) are also widely used to quantify signal similarity. An intuitive way is to directly measure the distance. However, two VCG signals can be misaligned, e.g., both signals show a typical pattern and yet there are variations in shape, amplitude, and phase between them. As such, a large distance may be resulted even between similar signal patterns due to the misalignment. In the literature, R-peak based matching has been used to improve the solution [102]. That is, R peaks from two ECGs are first aligned and then the Euclidean distance is calculated. In this way, ventricular depolarization (QRS complex) of two subjects are compared together, whereas the atrial depolarization (P wave) and ventricular

repolarization (T wave) may still be misaligned. Thus, there is a need to optimally align two ECG signals to compare the electrical activity of corresponding heart chambers.

Dynamic warping is a well-known approach to nonlinearly warp two signals for optimal alignment [40]. In our previous study [81], we extended time domain warping into spatiotemporal warping to address misalignment between multi-lead ECG signals. For example, given two VCG signals $\overrightarrow{\boldsymbol{v}_1}$ and $\overrightarrow{\boldsymbol{v}_2}$, with length $V_1$ and $V_2$, respectively. The spatiotemporal warping can be defined as: $\min_{W}\left[\sum_{(t_i,t_j)\in W}\left\|\overrightarrow{\boldsymbol{v}_1}(t_i)-\overrightarrow{\boldsymbol{v}_2}(t_j)\right\|\right]$, where $W$ is the warping path. $t_i$ is an index selected from $\overrightarrow{\boldsymbol{v}_1}$ and $t_j$ is from $\overrightarrow{\boldsymbol{v}_2}$. To find the warping path, a distance matrix $D$ is computed with each element $D(i,j)=\|\overrightarrow{\boldsymbol{v}_1}(i)-\overrightarrow{\boldsymbol{v}_2}(j)\|$, $i=1,2,\dots,V_1$ and $j=1,2,\dots,V_2$. The warping path $W$ connects $D(1,1)$ and $D(V_1,V_2)$, and recursively computed using dynamic programming approach:

$$G(i,j)=D(i,j)+\min\begin{cases}G(i,j-1)\\G(i-1,j-1)\\G(i-1,j)\end{cases} \tag{5.2}$$

The spatiotemporal dissimilarity between two multidimensional signals $\overrightarrow{\boldsymbol{v}_1}$ and $\overrightarrow{\boldsymbol{v}_2}$ (i.e., the warping distance) can then be calculated as: $\delta_{\overrightarrow{\boldsymbol{v}_1},\overrightarrow{\boldsymbol{v}_2}}=\frac{G(V_1,V_2)}{V_1+V_2}$. In this way, dynamic warping optimally aligns two signals to ensure a fair comparison. In other words, atrial depolarization (P wave), ventricular depolarization (QRS complex) and repolarization (T wave) from one beat is compared with the corresponding one from the other beat, as opposed to the incorrect comparison between atrial and ventricular patterns that generates misleading results.

*B. Data-driven Modeling and Optimization of a Large-scale Dynamic Network*

Dynamic warping between $N$ VCG signals generates a $N\times N$ warping matrix $\Delta$, with each element $\delta_{ij}$ characterizes the spatiotemporal dissimilarity between VCGs $\overrightarrow{\boldsymbol{v}_i}$ and $\overrightarrow{\boldsymbol{v}_j}$. However, the warping matrix itself is difficult to be used as features for the identification of disease properties in predictive modeling and monitoring. In addition, the measure of Euclidean distance is not directional and can mix the distances that are equal in magnitudes but along different spatial directions [81]. There is a need to transfer the warping matrix into feature vectors that not only preserve the warping matrix but also recover directional differences among space-time signals.

One way to realize such transformation is network embedding, which represents the signals as network nodes in a high-dimensional space. The Euclidean distance between each pair of nodes

preserves the warping distance of corresponding space-time signals. A commonly used network embedding approach is classic multidimensional scaling (MDS). Let $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ denote the location of $i^{th}$ and $j^{th}$ nodes in the network and $\delta_{ij}$ is the dissimilarity between $i^{th}$ and $j^{th}$ cardiac signals in the warping matrix $\Delta$. Then, the cost function of network embedding can be formulated as:

$$f = \sum_{i,j} (\|\boldsymbol{x}_i - \boldsymbol{x}_j\| - \delta_{ij})^2 \tag{5.3}$$

To optimize the location of network nodes, a Gram matrix $B$ is reconstructed from dissimilarity matrix $\Delta$ as $B = -\frac{1}{2} H \Delta^{(2)} H$, where H = $I - N^{-1}\mathbf{1}\mathbf{1}^T$ is a centering matrix. The element $b_{ij}$ in matrix B is obtained as:

$$b_{ij} = -\frac{1}{2}\left[\delta_{ij}^2 - \frac{1}{N}\sum_{k=1}^{N}\delta_{ik}^2 - \frac{1}{N}\sum_{k=1}^{N}\delta_{kj}^2 + \frac{1}{N^2}\sum_{g=1}^{N}\sum_{h=1}^{N}\delta_{gh}^2\right] \tag{5.4}$$

The Gram matrix $B$ can be defined as the scalar product $B = XX^T$, which is further written as $B = V\Lambda V^T = V\sqrt{\Lambda}\sqrt{\Lambda}V^T$ by eigen-decomposition. Here, $V$ is a matrix of eigenvectors and $\Lambda$ is a diagonal matrix of eigenvalues. Hence, optimal coordinates of network nodes are obtained as: $X = V\sqrt{\Lambda}$.

In the past decade, stochastic learning approaches have been increasingly adopted in analytics methods to solve optimization problems that involve big data. Notably, big data poses two significant challenges for optimization. On one hand, the dimensionality of the data is high. As such, second-order optimization is sometimes prohibitive. Thus, first order methods are preferred, e.g., gradient-based method. On the other hand, large volumes of data make it impractical for algorithms to go through the entire dataset many times. Hence, a stochastic approach is more suited to handle big data, because only one data sample is considered at every iteration. Moreover, it has been shown that the convergence of stochastic approaches for the optimization of convex and concave functions is guaranteed in most cases.

Motivated by [85], we rewrite the cost function of Eq. (5.3) as:

$$f = \frac{1}{2}\sum_{i}\sum_{j\neq i}\left(\left\|\boldsymbol{x}_i^t - \boldsymbol{x}_j^t\right\| - \delta_{ij}\right)^2 \Psi\left(\left\|\boldsymbol{x}_i^t - \boldsymbol{x}_j^t\right\|, \lambda\right) \tag{5.5}$$

where $\Psi(\cdot)$ is a bounded and monotonically decreasing function to favor the local topology in the network and $\lambda$ is defines the neighborhood radius. One example of $\Psi(\cdot)$ is a step function:

$$\Psi\left(\left\|\boldsymbol{x}_i^t - \boldsymbol{x}_j^t\right\|, \lambda\right) = \begin{cases} 1 & \text{if } \left\|\boldsymbol{x}_i^t - \boldsymbol{x}_j^t\right\| \leq \lambda \\ 0 & \text{if } \left\|\boldsymbol{x}_i^t - \boldsymbol{x}_j^t\right\| > \lambda \end{cases} \tag{5.6}$$

Minimization of Eq. (5.5) with respect to $\boldsymbol{x}_j$'s achieves the optimal location of the network nodes. Notably, the decrease in the value of $F$ can be considered as a summation of partial decrease resulted from each network node $i$, i.e., $f = \frac{1}{2}\sum_i f_i$. The optimization of Eq. (5.5) works as follows [85]: in each iteration, a node $i$ is randomly selected and its location (e.g., $\boldsymbol{x}_i$) is fixed. Then, all other nodes $\boldsymbol{x}_j (j \neq i)$ are updated based on the following rule:

$$\boldsymbol{x}_j^{t+1} = \boldsymbol{x}_j^t - \eta_t \frac{\partial f_i}{\partial \boldsymbol{x}_j^t} \tag{5.7}$$

where the learning rate $\eta_t$ is a monotonically deceasing function. Different forms of $\eta_t$ can be used to decay the learning rate along the learning process and we adopt an exponential function: $\eta_t = \eta_0 \left(\frac{\eta_T}{\eta_0}\right)^t$, where constants $\eta_0$ and $\eta_T$ specify the initial and final scales of the learning rate and the ratio of $\eta_T$ and $\eta_0$determines the decaying speed.

In the update rule of Eq. (5.7), one node is pinned and all $N-1$ nodes need to be updated at each iteration. This is different from conventional stochastic optimization rules, which update only one subject at each iteration. Advantages of updating $N-1$ nodes include 1) easy to escape from local minimum, 2) robust to initial node layout, and 3) fast convergence rate. However, this may introduce a larger variation compared with conventional update. Here, the update of $\boldsymbol{x}_j$ only depends on its relationship (i.e., the gradient) with respect to the pinned node $\boldsymbol{x}_i$. On the contrary, the update of $\boldsymbol{x}_j$ is controlled by its relationship with all other nodes in the conventional rule. As such, the convergence of the update rule of Eq. (5.7) can be influenced by the selection of model parameters (e.g., $\eta_t$) and requires fine tunes of these parameters, which makes it not practical in

the clinical practice. Therefore, a new cost function is introduced to regularize the update of $\boldsymbol{x}_j$ as:

$$f = \frac{1}{2}\left[\sum_i \sum_{j \neq i} \left(\left\|\boldsymbol{x}_i^t - \boldsymbol{x}_j^t\right\| - \delta_{ij}\right)^2 + \vartheta_t \left\|\boldsymbol{x}_j^t - \boldsymbol{x}_j^{t-1}\right\|_2^2\right] \Psi\left(\left\|\boldsymbol{x}_i^t - \boldsymbol{x}_j^t\right\|, \lambda\right) \tag{5.8}$$

where the second term $\vartheta_t \left\|\boldsymbol{x}_j^t - \boldsymbol{x}_j^{t-1}\right\|_2^2$ regularizes the update of $\boldsymbol{x}_j$ between successive iterations. Notably, the new cost function contains both spaital regularization term ($f_S = \sum_i \sum_{j \neq i} \left(\left\|\boldsymbol{x}_i^t - \boldsymbol{x}_j^t\right\| - \delta_{ij}\right)^2$) and temporal regularization term ($f_T = \vartheta_t \left\|\boldsymbol{x}_j^t - \boldsymbol{x}_j^{t-1}\right\|_2^2$). The update rule of stochastic network embedding is then given as:

$$\begin{cases} \boldsymbol{x}_j^{t+1} = \boldsymbol{x}_j^t - \eta_t \left(\nabla f_s + \vartheta_t \left(\boldsymbol{x}_j^t - \boldsymbol{x}_j^{t-1}\right)\right) & \text{if } \left\|\boldsymbol{x}_i^t - \boldsymbol{x}_j^t\right\| \leq \lambda_t \\ \boldsymbol{x}_j^{t+1} = \boldsymbol{x}_j^t & \text{if } \left\|\boldsymbol{x}_i^t - \boldsymbol{x}_j^t\right\| > \lambda_t \end{cases} \tag{5.9}$$

where $\nabla f_s = \frac{\left\|x_i^t - x_j^t\right\| - \delta_{ij}}{\left\|x_i^t - x_j^t\right\|} \left(\boldsymbol{x}_j^t - \boldsymbol{x}_i^t\right)$. This approach iterates until the maximum number of learning epoch is reached. The stochastic embedding approach is summarized in Algorithm 5-1.

**Algorithm 5-1** Stochastic network embedding – serial computing

**Input:** dissimilarity matrix $\Delta$ $(\delta_{ij})$
**Output:** coordinates of nodes $\boldsymbol{x}_i, i \in \{1, 2, \ldots, N\}$ in the high-dimensional network
1: $T$ – number of epoch
2: initialize $\boldsymbol{x}_i$, learning rate $\{\eta_t\}_{t=1}^T$ ,neighborhood radius $\{\lambda_t\}_{t=1}^T$, and regularization parameter $\{\vartheta_t\}_{t=1}^T$
3: **for** $t$ =1, 2, …, $T$ **do**
4: create index sequence $\varkappa$ by random permutation of $\{1, 2, \ldots, N\}$
5: **for** $n$ = 1 to $N$ **do**
6: randomly draw $i \in \varkappa$ , Fix the $\boldsymbol{x}_i^t$
7: update using Eq. (5.9)
8: **end for**
9: **end for**

Compared with MDS, the computational complexity of Algorithm 5-1 is reduced (i.e., $O(N)$ instead of $O(N^2)$). However, it is still expensive when large numbers of nodes are involved in the network. Thus, it is imperative to parallelize the algorithm and utilize more computing resources collaboratively. However, the inherently sequential nature of Algorithm 5-1 poses significant challenges for parallelization. An algorithm can be parallelized if the computing task can be

divided and results from individual processors (or computers) are independent of each other. Notably, fixing the location of nodes $a$ or $b$ and updating the location of other nodes will result in different locations of $c$ especially when $a$ and $b$ are far from each other. Thus, Algorithm 5-1 is difficult to be parallelized.

To address this challenge, we combine stochastic embedding with batch learning, which results in the mini-batch stochastic embedding. Instead of considering one sample at a time, the mini-batch algorithm considers multiple samples (i.e., a mini-batch) simultaneously at each iteration [86]. Assume the total number of data $N$ can be partitioned into $M$ mini-batches, each of the size $m$. Given a randomly selected mini-batch $\gamma$, the objective function on the mini-batch can be written as $f_\gamma = \frac{1}{|\gamma|}\sum_{i\in\gamma} f_i$, where $|\gamma|$ is the size of a batch. At each iteration, the mini-batch stochastic embedding accumulates sub-gradients with respect to $|\gamma|$ samples in the batch and calculate the average gradients as:

$$\nabla f_\gamma\left(\boldsymbol{x}_j^t\right) = \frac{1}{|\gamma|}\sum_{i\in\gamma} \nabla f_i\left(\boldsymbol{x}_j^t\right) \tag{5.10}$$

where $\nabla f_i\left(\boldsymbol{x}_j^t\right) = \frac{\partial f_i}{\partial \boldsymbol{x}_j^t}$. As such, each mini-batch generates a single average gradient. At the end of each batch, mini-batch stochastic embedding feeds $\nabla f_\gamma\left(\boldsymbol{x}_j^t\right)$ to update the location of all other nodes $\boldsymbol{x}_j^t$ $(j \neq i)$ as:

$$\boldsymbol{x}_j^t = \boldsymbol{x}_j^t - \eta_t \nabla f_\gamma\left(\boldsymbol{x}_j^t\right) \tag{5.11}$$

As the update rule is used less frequently, the mini-batch stochastic embedding provides computational benefits.

More importantly, each mini-batch $\gamma$ can be further considered as a combination of multiple subsets, i.e., $\gamma = \gamma_1 \cup \gamma_2 \cup, \ldots, \cup \gamma_S$. Thus, the calculation of sub-gradients in Eq. (5.10) can be represented as:

$$\nabla f_{\gamma}(\boldsymbol{x}_j^t) = \frac{1}{|\gamma|}\sum_{i\in\gamma}\nabla f_i(\boldsymbol{x}_j^t) = \frac{1}{|\gamma|}\left[\sum_{i\in\gamma_1}\nabla f_i(\boldsymbol{x}_j^t) + \sum_{i\in\gamma_2}\nabla f_i(\boldsymbol{x}_j^t) + \cdots + \sum_{i\in\gamma_S}\nabla f_i(\boldsymbol{x}_j^t)\right] \quad (5.12)$$

To this end, the mini-batch stochastic embedding can be readily implemented in a parallel manner. That is, the mini- batch $\gamma$ can be divided and the computation of average gradient can be assigned to $S$ individual processors. Each processor s will compute $\sum_{i\in\gamma_s}\nabla f_i(\boldsymbol{x}_j^t)$ and results from $S$ processors are combined using Eq. (5.12) to return the average gradient with respect to all samples in the mini-batch. The parallelized stochastic network embedding algorithm is summarized in Algorithm 5-2.

**Algorithm 5-2** Stochastic network embedding – parallel computing

1: initialize as Algorithm 5-1 step 1 and 2
2: randomly partition data into $M$ mini-batches $\gamma_1, \gamma_2, \ldots, \gamma_M$
3: start parallel pool with $S$ processors
4: **for** $t$ =1, 2, …, $T$ **do**
5: randomly choose a mini-batch $\gamma_t$
6: partition $\gamma_t$ into subsets $\gamma_{t,1}, \gamma_{t,2},\ldots, \gamma_{t,S}$, each with size $m$
7: **for** $n$ = 1 to $M$ **do {in parallel}**
8: processor $n$ gets partition $\gamma_{t,n}$
9: solve the sub-problem $\nabla f_{i\in\gamma_{t,n}}(\boldsymbol{x}_j^t)$ on $\gamma_{t,n}$ by Algorithm 5-1
10: **end for**
11: average over the batch size $m$ and update $\boldsymbol{x}_j^t$
12: **end for**
13: close parallel pool

*C. Network-based Monitoring of Cardiac Processes*

To this end, cardiovascular big data generated from the physical patient network are handled by the virtual IoH network. Coordinates of each node in the virtual network are optimized to represent the cardiac condition of the corresponding patient. Patients with similar conditions are located closely, thereby forming multiple communities in the network. Distances among nodes with the same community are small, whereas large distances exist among different groups. Network features (such as coordinates of nodes and community structures) can then be extracted for the diagnosis and prognosis of cardiovascular diseases.

In the literature, various predictive modeling approaches are available to quantitatively associate network features to patients' cardiac conditions. For example, k-nearest neighbors, logistic regression, and neural networks. Practitioners can select the most suitable approach based on the complexity of data and requirements of processing speed. In this study, we introduce another domain of methods, namely, statistical control charts. The most widely used control chart is the Shewhart chart. Two types of Shewhart charts, i.e., X-bar chart and R-chart, are constructed to monitor the mean and variance of process data in both manufacturing and healthcare. With the advancement of sensing technology, complex and higher dimensional process data are obtained. To address the data complexity, multivariate control charts were more and more used in the process control. The most widely used multivariate chart is the Hotelling $T^2$ chart, which is an extension of Shewhart chart for the monitoring of the mean vector of the process. The $T^2$ statistic for the $i^{\text{th}}$ patient is computed as:

$$T^2(i) = (\boldsymbol{x}_i - \overline{\boldsymbol{x}})^T \boldsymbol{S}^{-1}(\boldsymbol{x}_i - \overline{\boldsymbol{x}}) \tag{5.13}$$

Here, $\overline{\boldsymbol{x}}$ is an averaged location of network nodes and $\boldsymbol{S}$ is the covariance matrix. However, $\boldsymbol{S}$ can be singular when the dimensionality of node coordinate $\boldsymbol{x}$ is high. Therefore, we computed the $T^2$ statistic in Eq. (5.14) using the set of principal components as in our previous study [103]:

$$T^2(i) = (\boldsymbol{x}_i - \overline{\boldsymbol{x}})^T \boldsymbol{S}^{-1}(\boldsymbol{x}_i - \overline{\boldsymbol{x}}) = \sum_{r=1}^{Q} \frac{\mathbf{Z}(i,r)^2}{\lambda_r^2} \tag{5.14}$$

The upper control limit (UCL) of Hotelling $T^2$control chart is estimated as:

$$UCL_{T^2} = \frac{q(N+1)(N-1)}{N^2 - Nq} F_{\alpha,q,M-q} \tag{5.15}$$

where $q$ is the dimensionality of $\boldsymbol{x}_i$. $F_{\alpha,q,M-q}$ is the upper $100\alpha\%$ of critical point of F distribution with $q$ and $N-q$ degree of freedoms. The significance level $\alpha$ is set as 0.05.

## 5.5. Materials and Experimental Design

In this investigation, we evaluate and validate the developed methodology using both simulation studies and real-world applications. As shown in Fig. 5-7, we evaluate four factor groups in the simulation study as follows:

1) *ECG segment*: Different types of cardiovascular diseases will impact the morphology of different ECG segments. For example, pulmonary hypertension will result in the right atrial enlargement, which is reflected as enlarged P waves. Also, diseases occur in ventricles (e.g., myocardial infarctions) will impact the activity of ventricular depolarization and repolarization, which lead to abnormal morphologies in QRS and T waves, respectively. Notably, the length of QRS is the smallest, whereas P and T waves are larger. Segments with smaller lengths pose greater challenges for the detection. It is important to test the proposed algorithm on shifted ECG segments with different lengths.

2) *Mean shift magnitude within an ECG segment*: Disease-altered ECG morphology also reveals the severity of the disease. When the disease is in the early stage, the morphology of an ECG segment is changed slightly from the normal one. However, the shift magnitude is large when the disease lasts for a long time. Notably, small shift magnitudes pose greater challenges for the automated detection. It is important to implement the proposed algorithm on varying shift magnitudes of ECG segments to evaluate the ability of early disease pattern recognition.

3) *Dissimilarity measure*: Effective measurement of the dissimilarity among ECG signals is a prerequisite for achieving high detection power. The signal dissimilarity can be characterized from various perspective, including morphological dissimilarity, signal correlation, and the mutual dependence of the amount of information held in both signals. It is important to evaluate the effectiveness of different measures and identify the most suitable one for a particular type of disease.

4) *Statistical control chart*: Early diagnosis of cardiac diseases calls for not only the accurate measure of signal dissimilarities but also effective detection and prediction tools. It is important to evaluate various monitoring schemes under different scenarios (e.g., length of shifted ECG segments and shift magnitudes) to identify the one that most sensitive to small and early-stage abnormalities in the ECG signals.

In real-world applications, a continuously recorded ECG signals from the PhysioNet Long-term ST database [104, 105] are used to evaluate the performance of the developed methodology. In the database, signals were digitized at 250 Hz sampling rate with a 12-bit resolution over a range of $\pm$10 mV. Beat-by-beat annotations are also provided, which labels each beat to be normal or a certain type of disease. Two types of beats are identified in the continuous ECG records: beats

with normal rhythm and beats with premature ventricular contraction. It may be noted that we first preprocess the ECG with an FFT band-pass filter (1~120 Hz), which removes low-frequency artifacts (e.g., baseline wander) and high-frequency noises. Individual beats are then extracted for similarity measurement. Details of signal preprocessing and beat extraction are available in our previous studies [8, 81].

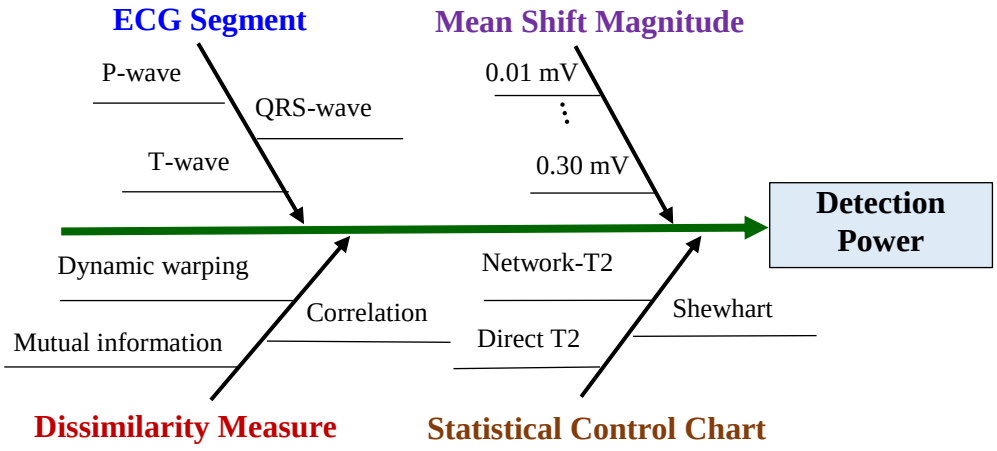



Figure 5-7. Experimental design of the simulation study.

## 5.6. Experimental Results

### *A. Simulation Studies*

In simulation studies, we first evaluate the proposed parallel computing algorithm over traditional serial algorithm for computational efficiency of handling large numbers of patients. Second, the proposed methodology is implemented for the detection of disease-altered patterns in cardiac activity. Following the experimental design in Section 4, the detection power of the proposed methodology is evaluated by considering the following four factors: 1) ECG segment, 2) mean shift magnitude within an ECG segment, 3) dissimilarity measure and 4) control chart.

*1) Evaluation of the computational efficiency*

The computational efficiency of dynamic warping algorithm and stochastic network embedding are evaluated with both parallel and serial computing schemes. As shown in Fig. 5-8a and 9b, the parallel computing scheme consumes much less time than serial computing, especially when a larger number of patients are involved in the IoH network. As shown in Fig. 5-8a, the computation time of parallel computing increases slowly with the increasing number of patients, whereas computation time of serial computing boosts significantly. Notably, we stop the

computation after 20,000 seconds for dynamic warping. In other words, the serial computing will take > 5.5 hours to calculate the pairwise warping distances for 1300 and 1500 patients. As shown in Fig. 5-8b, parallel network embedding (Algorithm 5-2) is compared with the serial computing version (Algorithm 5-1). We increase the size of the simulated network from 36,000 to 180,000 and record the computation time of both algorithms. Notably, the computation time of serial computing increases significantly when a large number of patients are included, whereas computation time of the parallel computing increases much slower. It may be noted that the performance gap between two algorithms increases dramatically with the increasing size of network. When 72000 patients are embedded, the serial approach is 400s slower than the parallel computing. When the number of patients reaches 180,000, serial computing is >2500s slower than the parallel framework.

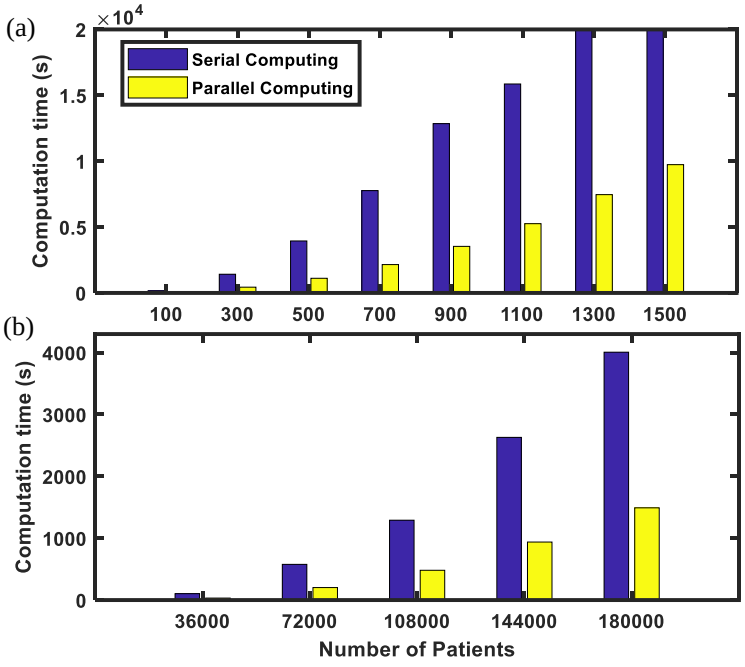



Figure 5-8. Computation time using parallel and serial computing for (a) dynamic warping and (b) stochastic network embedding.

*2) Performance Evaluation*

Further, we evaluate the statistical performance of the proposed methodology for the detection of disease patterns. Here, Gaussian noise is added to ECG cycles extracted from a normal subject. Following the experimental design in Section IV, three segments of ECG signals are shifted (i.e., ST, P, and QRS) to simulate different types of diseases. For example, an elevated ST segment

corresponds to acute myocardial infarction. Also, a larger P wave indicates the potential atrial enlargement. When the patient is associated with massive pericardial effusion, the voltage of QRS will be lower, which is shown as a reduced QRS complex in the ECG signals. The severity of disease can be represented by the shift magnitude of cardiac signals. In other words, a tiny shift indicates the disease is in the early stage, whereas a large shift is associated with well-developed disease conditions and calls for immediate medical intervention. Notably, ST is the longest segment in the ECG signal. The P segment is shorter and QRS is with the smallest length.

In this study, we construct statistical control charts for the monitoring of patients' cardiac conditions. Similar to the SPC monitoring system of Zhou, Sun and Shi [9], three types of control charts are constructed, including 1) a $T^2$chart based on the network features (i.e., coordinates of embedded network nodes), 2) a $T^2$chart directly using the original ECG signals, and 3) a Shewhart chart monitoring the mean change of the ECG signals. In total, 150 new ECG cycles are generated. The first 50 cycles are considered as normal (i.e., in-control) signals, whereas mean shifts with different magnitudes are added to the rest 100 cycles to generate disease-altered (i.e., out-of-control) signals. The detection power is calculated as the number of out-of-control signals over 100. Notably, dissimilarities among ECG signals are measured by the spatiotemporal warping approach.

As shown in Fig. 5-9, the network $T^2$chart performs better than the Shewhart chart and the $T^2$chart of original signals under all three scenarios. Specifically, the detection power of the network $T^2$ chart (blue triangles) quickly reaches one at very small shift magnitude (0.03 mV) when the length of shifted segment is larger (shifted ST segments in Fig. 5-9a). When shifted segments are shorter, network $T^2$ chart is also sensitive to small disease-induced changes and the detection power reaches one at the shift magnitude of 0.08 mV and 0.12 mV for P (Fig. 5-9b) and QRS segments (Fig. 5-9c), respectively. The detection power of the $T^2$chart of original signals (orange squares) increases slower with the increase of shift magnitude compared with the network $T^2$chart. When the length of shifted segment is short (shifted QRS segments in Fig. 5-9c), the detection power is only 0.4 when the mean shift magnitude is 0.15 mV and it reaches one when the shift magnitude is over 0.22 mV. Notably, the detection power of Shewhart chart is comparable to the $T^2$chart of original signals when the shift magnitude is large (e.g., ST segments in Fig. 5-

9a). However, no out-of-control signals can be detected by the Shewhart chart when shift magnitudes are small (see Fig. 5-9b and 5-9c).

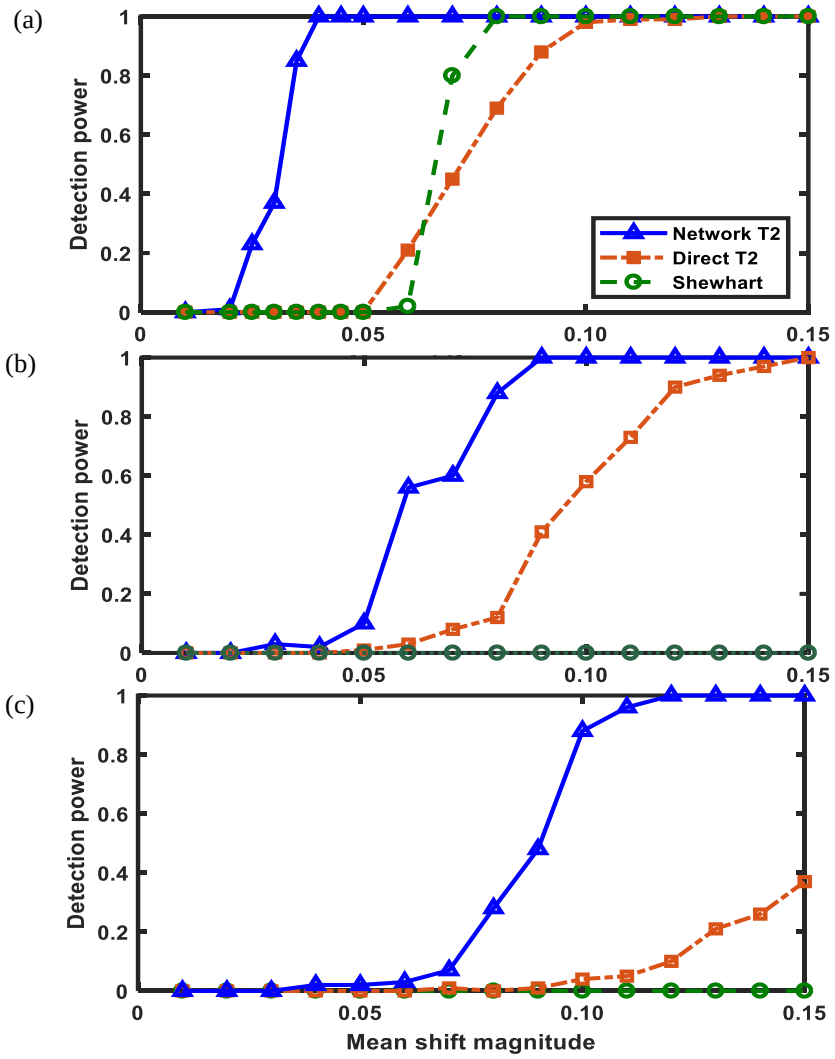



Figure 5-9. Performance comparison between three types of control charts under different shift magnitudes for (a) shifted ST segments, (b) shifted P segments and (c) shifted QRS segments.

In addition, performances of three similarity measures (i.e., dynamic warping, mutual information, and correlation) are also compared. Here, we assume the disease-induced changes occur in ST segments and the developed network $T^2$ chart is implemented to calculate the detection power of all three measures. As shown in Fig. 5-10, it may be noted that dynamic warping performs better than the mutual information. For example, the network $T^2$ chart claims a 100% detection with dynamic warping when the shift magnitude reaches 0.04. However, the detection power

reaches one with mutual information when the shift magnitude is over 0.07. Notably, the control chart cannot yield a detection based on correlation even if the shift magnitude is relatively large.

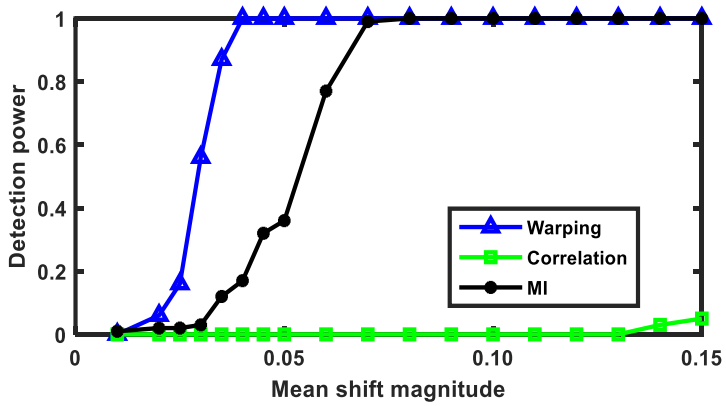



Figure 5-10. Performance comparison between three types of similarity measures (dynamic warping, correlation and mutual information) under different shift magnitudes.

## *B. Real-world Studies*

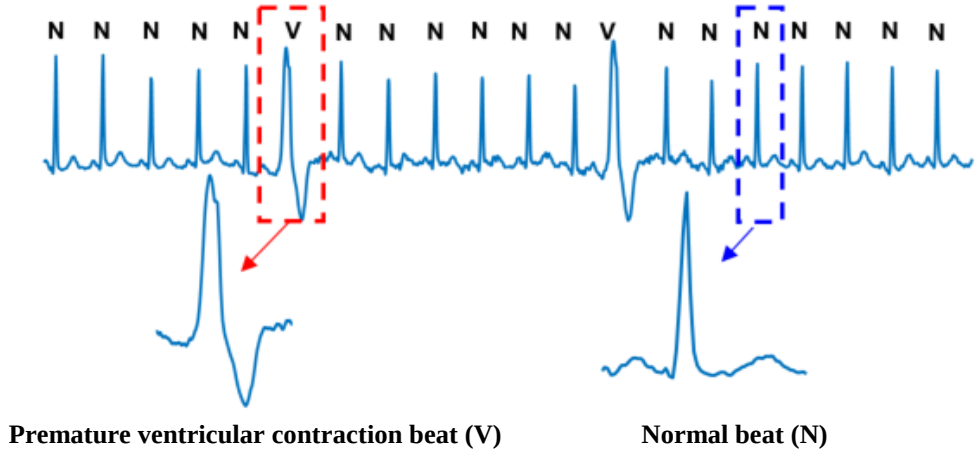



Figure 5-11. A segment of real-world ECG signals containing normal beats (N) and premature ventricular contraction beats (V).

A segment of real-world ECG signals is shown in Fig. 5-11. Notably, the morphology of ECG beats varies with respect to time: 1) the normal beats (with annotation N) show similar patterns, but there are variations among them, and 2) the premature ventricular contraction beat (with annotation V) appears once in a while. The premature ventricular contraction is caused by an ectopic cardiac pacemaker in the ventricle, which usually manifests as premature and bizarrely shaped QRS complex with longer duration. Potential causes of premature ventricular contractions include heart attack, congestive heart failure, and diseases of heart valves. It is important to monitor

the onset and progress of premature ventricular contractions for the early diagnosis of fatal cardiac diseases.

In this study, we illustrate the proposed methodology for the monitoring of beat-to-beat variations of one cardiac patient. As shown in Fig. 5-12, blue nodes represent normal beats and red nodes correspond to beats with premature ventricular contractions. Notably, the embedded network structure is changing when beats are increasingly embedded into the network. For example, the first 15 beats contain three premature ventricular contractions, which are located apart from the main group of normal beats (red dots in Fig. 5-12a). Fig. 5-12b shows the embedded network when first 25 beats are included in the network. Five beats associated with premature ventricular contraction are located closer, and they together are away from the normal cluster. When 50 beats are embedded (see Fig. 5-12c), 13 premature ventricular beats are identified, which forms a distinct group that located away from the normal beats.

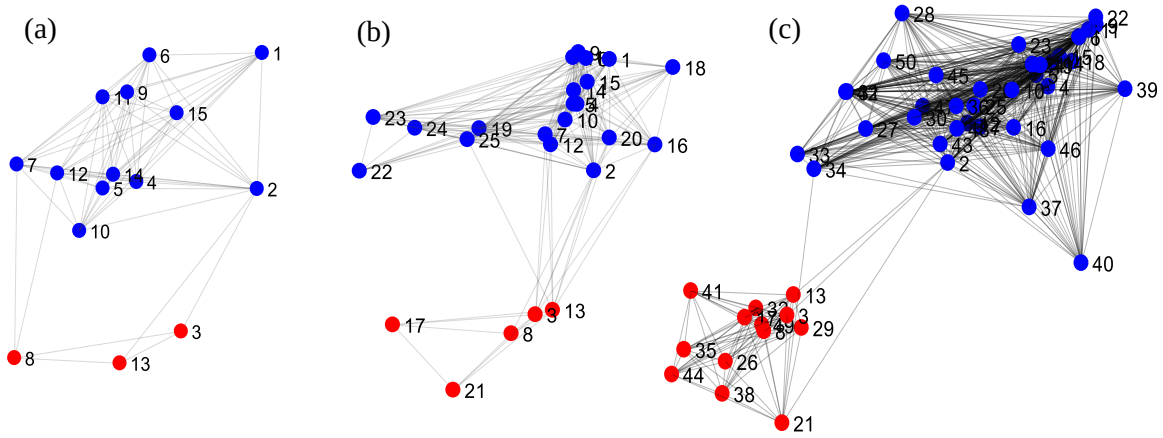



Figure 5-12. The constructed beat-to-beat networks for (a) 15 beats; (b) 25 beats; and (c) 50 beats from the real-world ECG signals.

An embedded network of 2000 successive ECG beats is shown in Fig. 5-13. Based on the embedded network, cardiologists are able to visually inspect the progress of the disease and provide timely medical intervention. For example, the heart function of the patient deteriorates when new beats are increasingly embedded near the disease group. Cardiologists need to evaluate the severity of the patient's condition and provide medical intervention in a timely manner to reduce the risk of the disease. More importantly, the dynamic network provides a great opportunity for cardiologists to identify the trend of the change in the patient's cardiac condition and provide personalized therapies. For example, the patient always shows 1-2 premature ventricular

contractions every minute. This may be due to a constant effect in the functionality of the patient's ventricle. This situation can be considered as normal and no action is needed. However, urgent medical intervention needs to be given when the number of disease-altered beats is suddenly increased. This indicates a significant change in the patients' ventricle condition and it may be associated with the fatal myocardial infarctions.

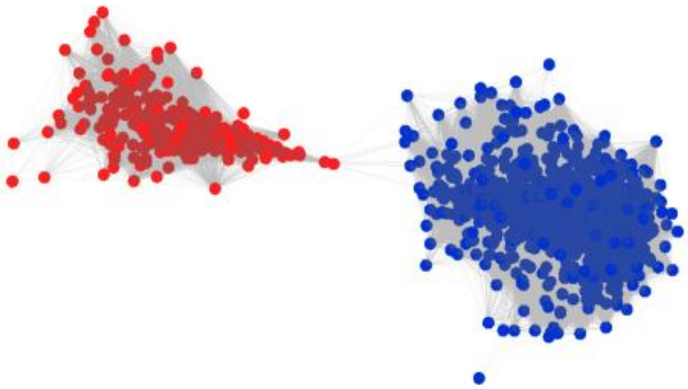


Figure 5-13. The constructed beat-to-beat network with blue nodes representing normal beats and read nodes representing premature ventricular contractions.

## 5.7. Discussion and Conclusions

Heart disease is the leading cause of death in the world. Optimal health management hinges on the development of advanced cardiac monitoring system for the early detection of disease progression and timely delivery of life-saving therapies. In this study, we propose the new IoT technology, namely the Internet of Hearts to enable and assist in the inclusion of a large number of patients in a networked structure for condition monitoring and health management. However, continuous patient monitoring leads to the accumulation of big data. Traditional approaches are limited in their ability to handle big data and extract pertinent information about patients' cardiac conditions. This is an urgent need to develop new analytical methods that can help realize the full potential of big data from the IoH system for effective and efficient process monitoring and optimal decision making.

Therefore, we develop new parallel computing algorithms for sensor-based modeling and analysis of the large-scale network of patients. First, we propose to measure dissimilarities among time-varying heartbeats from a specific patient for personalized monitoring, as well as dissimilarities among representative heartbeats from various patients for population monitoring. This, in turn, is conducive to study disease progression dynamics in the levels of individual patients

and the population. It is worth mentioning that there are various ways to measure the pattern dissimilarity such as correlation, mutual information, and dynamic spatiotemporal warping. These dissimilarity measures can also be computed in a variety of signal representations such as time domain, frequency domain, state space domain, or time-frequency domain. Second, we transform the obtained dissimilarity matrix into the B2B and P2P networks. Indeed, ECG patterns can be matched between subjects to identify communities, or benchmarked between heartbeats for personalized health monitoring. For the personalized B2B network, each beat is represented as a network node and node attributes are ECG patterns within the heartbeat. For the P2P network, each node represents an individual and node attributes are the salient ECG pattern from this individual.

However, such network modeling is computationally expensive due to the high volume of data. Traditional serial computing schemes are limited in their ability for efficient network representation and real-time analytics in the IoH setting. In this investigation, we developed new parallel computing algorithms to reduce the computational complexity of network modeling by distributing the network optimization into multiple processors for fast processing. As such, this is well suited for modeling and optimization of the large-scale network of patients in the IoH contexts. Finally, we leverage network features to develop new statistical control charts for real-time cardiac monitoring.

Notably, multi-level network modeling is a critical component in the IoH that will help clinicians identify communities that share a similarity in cardiac activity, study disease mechanisms in each community, and pinpoint a patient in one of the communities. The P2P network reveals the aggregated information of patients. For example, community detection in the P2P network will help delineate different types of cardiac conditions. The number of nodes in each community discloses how many patients are associated with each cardiac condition. The distribution of nodes in each community shows whether large or small variations exist among patients in that community. The spatial location of a node (i.e., a patient) helps identify the cardiac condition. For example, if one node is embedded close to the normal group, it will have a high probability to be normal. However, if nodes are increasingly embedded close to the diseased group, a closer attention should be paid to them. Furthermore, the personalized B2B network over time reveals the progressive variations of cardiac conditions for a specific individual. In this investigation, multi-level network analytics are shown to have strong potentials for a better

understanding of disease mechanisms in various populations and the identification of high-risk individuals.

# Chapter 6: Conclusions and Future Research

In recent years, there has been growing interest to investigate increasingly complex systems, e.g., from conventional machining (~mm) to ultra-precision machining (~nm) and from organs to cells and genes in the medicine. To address the system complexity, advanced sensing technology has been used more and more to increase the information visibility and system controllability. In particular, sensor networks and imaging technology have been more and more used to capture high-resolution space-time information of the system. Furthermore, the Internet of Things deploys a multitude of sensors for the continuous monitoring of machine and patient conditions. As a result, the structure of sensing data has become increasingly ***complex***:

1) Data from individual sensors are with higher dimensionality, e.g., from single-channel to multi-channel signals and from 1-dimensional profiles to 2- or 3- dimensional imaging profiles.
2) Data are collected from larger-scale systems, e.g., from a single machine to a large-scale manufacturing system and from a single patient to the patient population.

In the literature, limited works have been done in industrial engineering to cope with the ever-increasing complexity of data for system modeling, process monitoring, anomaly detection, and performance improvement. The methodologies developed in this dissertation are interdisciplinary, which integrates techniques from engineering domain, statistics, and medicine. ***Contributions*** of this dissertation are summarized as follows:

1) In Chapter 2, a new network approach is developed for the differentiation of disease-altered patterns in 3-lead VCG signals. i) A new warping algorithm is designed to quantify the spatiotemporal dissimilarity among VCG recordings. ii) A network embedding approach is introduced to embed the dissimilarity matrix into a network, which preserves warping distance among functional recordings. iii) Disease properties are quantitatively associated with network features that represent disease-altered spatiotemporal patterns underlying 3-lead VCG signals. The developed approach facilitates the understanding of disease-altered spatiotemporal patterns in cardiac electrical activity and can be potentially applicable for the identification of not only myocardial infarctions but also other cardiac disorders.

2) In Chapter 3, a new dynamic network scheme is designed and developed to represent, model, and monitor real-time image data. i) The steam of time-varying images is represented as a dynamic network. ii) Community structures of the dynamic network are characterized. iii) A new control chart, namely, the network-generalized likelihood ratio chart is developed to detect the change point of the underlying dynamics of complex processes. The developed network scheme has advantageous features to effectively represent not only 2D images but also higher-dimensional profiles. Thus, the new dynamic network SPC method is shown to have strong potentials for general applications in diverse domains with in-situ imaging data.

3) In Chapter 4, a new large-scale network approach is developed for IIoT machine information processing, condition monitoring, and fault diagnosis. i) We introduce a stochastic learning approach to significantly improve the computational efficiency of network embedding and optimize the embedding of each machine in the network. ii) The stochastic embedding is further parallelized to enable a fast optimization of large-scale machine networks in the contexts of IIoT. The proposed methodology has a strong potential to promote existing manufacturing practices towards globalized production optimization and management. Furthermore, it provides a great opportunity to optimize the energy consumption and realize green and sustainable manufacturing. In addition, the parallel-computing approach for network optimization can be potentially applicable to not only manufacturing but also other research areas such as disease diagnosis and prognosis, infection prevention and control, supply chain optimization and service networks.

4) In Chapter 5, a new approach of dynamic network modeling is developed for multi-level cardiac monitoring to realize the Internet of Health Things. i) The IoH, a new IoT-based cardiac care system is proposed, which integrates IoT sensing, wireless communication, big data analytics, and medical decision making to realize smart and connected cardiac health management. ii) A beat-to-beat network is developed to create personalized baselines and reveal progressive cardiac variations of a specific patient. iii) A patient-to-patient network is constructed to delineate different types of cardiac conditions among various patients in the IoH for population-level monitoring. iv) A parallelized stochastic learning approach is developed to harness the power of multiple processors for fast network optimization. This enables the inclusion of a large patient population in the IoH for real-

time monitoring and disease management. The developed scheme of multi-level network analytics is shown to have strong potentials for a better understanding of disease mechanisms in various populations and the identification of high-risk individuals.

Based on the developed methodologies in this dissertation, there are some research topics that deserve further investigations, including:

- Dynamic network monitoring of time-varying image profiles
  - Network community statistics: In Chapter 3, we extract mean and standard deviation of pixel intensities, and the number of pixels in each community as features and fed into the NGLR chart for change point detection. Future studies will be conducted to extract other types of community statistics. For example, each community is a sub-network and network topological metrics such as density and clustering coefficient can be extracted.
  - Sensitivity analysis and performance comparison: Other methods also have potentials for online image-based process monitoring, e.g., wavelet analysis. Future studies can be done to compare the network-based approach and other methods. Also, sensitivity analysis will be conducted to evaluate the influence of model parameters (e.g., number of communities and the weight $\alpha$) on the performance of change-point detection.

- Industrial Internet of Things
  - Heterogeneous sensor fusion: In Chapter 4, we focus on the analysis of energy profiles in the IIoT. Instead of one type of sensor, heterogeneous sensors are more and more used for collaborative monitoring of advanced manufacturing processes. How to extract useful information from different types of sensors remains a challenge. In the future, IIoT-based heterogeneous sensor fusion is worth investigating.
  - Low-volume and high-mix scenario: In Chapter 4, we assume the same type of products. In other words, we focus on a high-volume and low-mix scenario. Future works may extend our algorithm to the low-volume and high-mix scenario. That is, different types of machines are monitored simultaneously in the IIoT. This calls for the development of new algorithms that incorporate the variations among heterogeneous machines or products.

- Internet of Health Things
  - Beyond morphological features: In Chapter 5, we characterize morphological features from cardiac signals. Future studies will integrate other types of features (e.g., frequency domain features) into the IoH virtual network for disease diagnosis. Furthermore, the data-rich environment of IoH not only provides cardiac signals but also other patient data, e.g., blood pressure, respiratory rate, and other vital signs. These data will also be considered in the IoH in the future.
  - Advanced parallel computing techniques: In Chapter 4 and 5, the focus is to parallelize the stochastic network embedding, which is with a sequential nature and difficult to be directly parallelized. Future studies will continue developing advanced parallel computing techniques. For example, studies will be done to develop approximate and probabilistic algorithms to compute sufficiently optimal solutions while using bounded time, space, communication, and energy; and design new programming logic and language mechanisms to support existing and new computational models.

# Appendix

## Appendix A – Proof of Proposition 1 in Chapter 3

*Proposition 1*: If the process is in control, $\boldsymbol{y}^{(i)}$ is approximately a multivariate normal distribution as $\boldsymbol{y}^{(i)} \sim MVN(\boldsymbol{\mu}_0, \boldsymbol{\Sigma})$, where $\boldsymbol{\mu}_0$ and $\boldsymbol{\Sigma}$ are in-control mean and covariance matrix of feature vectors.

*Proof*: Let's focus on the $k^{th}$ community and thus dropping the subscript $k$. As the Hamiltonian approach favors strongly connected nodes into one community, image pixels within each community are homogeneous. In other words, pixel intensities $x_1, x_2, \ldots, x_n$ within one community approximately follow a normal distribution $N(\mu, \sigma^2)$ (see Fig. 3-6).

1) By the central limit theorem, $x_1, x_2, \ldots, x_n$ be random samples from a distribution with finite mean $\mu$ and variance $\sigma^2$, the sample mean $\bar{x}$ follows an approximate normal distribution when $n$ is large. That is, the sample mean $\bar{x}$ of one community from image to image can be approximated by a normal distribution as $\bar{x} \sim N(\mu, \frac{\sigma^2}{n})$.

2) Notably, $\sum_{i=1}^{n} \left(\frac{x_i - \mu}{\sigma}\right)^2 \sim \chi^2_{(n)}$. If we replace $\mu$ with $\bar{x}$, then we have:

$$\sum_{i=1}^{n} \left(\frac{x_i - \bar{x}}{\sigma}\right)^2 \sim \chi^2_{(n-1)}$$

Because $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$, we have:

$$\sum_{i=1}^{n} \left(\frac{x_i - \bar{x}}{\sigma}\right)^2 = \frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{(n-1)}$$

When $n$ is large, the chi-square distribution asymptotically converges to a normal distribution based on the central limit theorem [106]. In practice, $n >$50 the distribution is sufficiently close to a normal distribution for the difference to be ignored [107]. In our case, the number of pixels in each community is much greater than 50.

Further, let $w = \frac{(n-1)s^2}{\sigma^2}$ and $h(w)$ denotes the probability density function. The expectation of $\frac{\sqrt{(n-1)}s}{\sigma}$ is then given as:

$$E\left(\frac{\sqrt{(n-1)}s}{\sigma}\right) = \int_0^\infty \sqrt{w}h(w)dw = \int_0^\infty \sqrt{w}\frac{1}{2^{(n-1)/2}\Gamma((n-1)/2)} w^{\frac{(n-1)}{2}-1} e^{-w/2} dw$$

$$= \sqrt{2}\frac{\Gamma(n/2)}{\Gamma((n-1)/2)}$$

Hence,

$$E(s) = \sqrt{\frac{2}{n-1}}\frac{\Gamma(n/2)}{\Gamma((n-1)/2)}\sigma$$

In addition,

$$Var(s) = E(s^2) - \big(E(s)\big)^2 = \left(1 - \left(\sqrt{\frac{2}{n-1}}\frac{\Gamma(n/2)}{\Gamma((n-1)/2)}\right)^2\right)\sigma^2$$

Based on the normality assumption, the $s$ chart is commonly used to monitor process variance as follows:

$$UCL = c_4\sigma + 3\sqrt{(1-{c_4}^2)}\sigma$$

$$CL = c_4\sigma$$

$$LCL = c_4\sigma - 3\sqrt{(1-{c_4}^2)}\sigma$$

where

$$c_4 = \sqrt{\frac{2}{n-1}}\frac{\Gamma(n/2)}{\Gamma((n-1)/2)}$$

If the process is in control, the number of pixels $n$ within one community randomly fluctuates around a mean value. Thus, distribution of $n$ can be approximated by a normal distribution. The normal probability plot in Fig. 3-7 shows that the number of pixels of one community from image to image is approximately normally distributed.■

## Appendix B – Proof of Proposition 2 in Chapter 3

*Proposition 2*: In the eigenspace, multivariate NGLR statistic $R_{m,\boldsymbol{y}}$ in Eq. (3-11) is equivalently computed as:

$$R_{m,\mathbf{Z}} = \max_{m-w\le\tau<m} \frac{m-\tau}{2} \left(\widehat{\boldsymbol{\mu}}_{m,\tau,\mathbf{Z}}\right)^T \boldsymbol{S_z}^{-1} (\widehat{\boldsymbol{\mu}}_{m,\tau,\mathbf{Z}})$$

where:

$$\widehat{\boldsymbol{\mu}}_{m,\tau,\mathbf{Z}} = \frac{1}{(m-\tau)} \sum_{i=\tau+1}^{m} \mathbf{Z}(i,:)^T$$

and $\mathbf{Z}(i,:)^T$, $i = 1,2,\dots,m$ are principal components obtained by eigen transformation of feature vector $\boldsymbol{y}^{(i)}$.

*Proof*: First, the mean is subtracted from the feature matrix, i.e., $\mathbf{Y}^* = [\boldsymbol{y}^{(1)} - \overline{\boldsymbol{y}},\ \boldsymbol{y}^{(2)} - \overline{\boldsymbol{y}}, \dots, \boldsymbol{y}^{(m)} - \overline{\boldsymbol{y}}]^T$. The $m \times Q$ matrix $\mathbf{Y}^*$ is then factorized into three matrices (**U**, **Ω** and **V)** by singular value decomposition:

$$\mathbf{Y}^* = \mathbf{U\Omega V}^T$$

where **U** and **V** are $m \times m$ and $Q \times Q$ orthogonal matrices, **Ω** is a $m \times Q$ diagonal matrix with diagonal entries $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_Q \ge 0$ (i.e., singular values of $\mathbf{Y}^*$). Each vector in **V** is an eigenvector that represents a principal component direction of $\mathbf{Y}^*$. Principal components **Z** ($m \times Q$) are computed by projecting $\mathbf{Y}^*$ onto the eigenvector **V** as:

$$\mathbf{Z} = \mathbf{Y}^*\mathbf{V} = \mathbf{U\Omega V}^T\mathbf{V} = \mathbf{U\Omega}$$

Note that the first principal component $\mathbf{Z}(1,:) = \mathbf{Y}^*\boldsymbol{v_1}$ has the largest sample variance among all principal components. The feature covariance matrix $\boldsymbol{S}$ becomes:

$$\boldsymbol{S} = \frac{1}{m-1} \sum_{i=1}^{m} \left(\boldsymbol{y}^{(i)} - \overline{\boldsymbol{y}}\right)\left(\boldsymbol{y}^{(i)} - \overline{\boldsymbol{y}}\right)^T$$

$$= \frac{1}{m-1} \mathbf{V}\mathbf{Z}^T\mathbf{Z}\mathbf{V}^T = \mathbf{V}\boldsymbol{S_z}\mathbf{V}^T$$

where $\boldsymbol{S_z}$ denotes the covariance matrix of **Z**. By substituting Eq. (3-7) into (3-11), we have:

$$\left(\widehat{\boldsymbol{\mu}}_{m,\tau,\boldsymbol{y}} - \overline{\boldsymbol{y}}\right)^T \boldsymbol{S}^{-1} (\widehat{\boldsymbol{\mu}}_{m,\tau,\boldsymbol{y}} - \overline{\boldsymbol{y}})$$

$$= \left( \frac{1}{(m-\tau)} \sum_{i=\tau+1}^{m} \boldsymbol{y}^{(i)} - \bar{\boldsymbol{y}} \right)^T \boldsymbol{S}^{-1} \left( \frac{1}{(m-\tau)} \sum_{i=\tau+1}^{m} \boldsymbol{y}^{(i)} - \bar{\boldsymbol{y}} \right)$$

$$= \left( \frac{\sum_{i=\tau+1}^{m} (\boldsymbol{y}^{(i)} - \bar{\boldsymbol{y}})}{(m-\tau)} \right)^T \boldsymbol{S}^{-1} \left( \frac{\sum_{i=\tau+1}^{m} (\boldsymbol{y}^{(i)} - \bar{\boldsymbol{y}})}{(m-\tau)} \right)$$

$$= \left( \frac{\sum_{i=\tau+1}^{m} \mathbf{V}\,\mathbf{Z}(i,:)^T}{(m-\tau)} \right)^T \boldsymbol{S}^{-1} \left( \frac{\sum_{i=\tau+1}^{m} \mathbf{V}\mathbf{Z}(i,:)^T}{(m-\tau)} \right)$$

$$= \left( \mathbf{V}\,\hat{\boldsymbol{\mu}}_{m,\tau,\mathbf{Z}} \right)^T \boldsymbol{S}^{-1} \left( \mathbf{V}\,\hat{\boldsymbol{\mu}}_{m,\tau,\mathbf{Z}} \right)$$

$$= \left( \mathbf{V}\,\hat{\boldsymbol{\mu}}_{m,\tau,\mathbf{Z}} \right)^T (\mathbf{V}\boldsymbol{S}_{\boldsymbol{z}}\mathbf{V}^T)^{-1} \left( \mathbf{V}\,\hat{\boldsymbol{\mu}}_{m,\tau,\mathbf{Z}} \right)$$

Hence, the multivariate NGLR statistic $R_{m,\boldsymbol{y}}$ is equivalent to $R_{m,\mathbf{Z}}$ if all the $Q$ dimensions of principal components $\mathbf{Z}$ are used in Eq. (3-12):

$$R_{m,\boldsymbol{y}} = \max_{\max(0,m-w)\le\tau<m} \frac{m-\tau}{2} \left( \hat{\boldsymbol{\mu}}_{m,\tau,\boldsymbol{y}} - \bar{\boldsymbol{y}} \right)^T \boldsymbol{S}^{-1} \left( \hat{\boldsymbol{\mu}}_{m,\tau,\boldsymbol{y}} - \bar{\boldsymbol{y}} \right)$$

$$= \max_{\max(0,m-w)\le\tau<m} \frac{m-\tau}{2} \left( \mathbf{V}\,\hat{\boldsymbol{\mu}}_{m,\tau,\mathbf{Z}} \right)^T (\mathbf{V}\boldsymbol{S}_{\boldsymbol{z}}\mathbf{V}^T)^{-1} \left( \mathbf{V}\,\hat{\boldsymbol{\mu}}_{m,\tau,\mathbf{Z}} \right)$$

$$= \max_{\max(0,m-w)\le\tau<m} \frac{m-\tau}{2} \left( \hat{\boldsymbol{\mu}}_{m,\tau,\mathbf{Z}} \right)^T \boldsymbol{S}_{\boldsymbol{z}}{}^{-1} \left( \hat{\boldsymbol{\mu}}_{m,\tau,\mathbf{Z}} \right)$$

$$= R_{m,\mathbf{Z}} \quad \blacksquare$$

# Appendix C – Copyright Permissions

## Copyright permission for Chapter 2



**Title:** Dynamic spatiotemporal warping for the detection and location of myocardial infarctions

**Conference Proceedings:** Automation Science and Engineering (CASE), 2012 IEEE International Conference on

**Author:** Chen Kan

**Publisher:** IEEE

**Date:** Aug. 2012

Copyright © 2012, IEEE


### Thesis / Dissertation Reuse

**The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:**

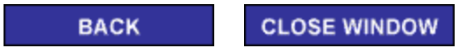
*Requirements to be followed when using any portion (e.g., figure, graph, table, or textual material) of an IEEE copyrighted paper in a thesis:*

1) In the case of textual material (e.g., using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © 2011 IEEE.
2) In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.
3) If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author's approval.

*Requirements to be followed when using an entire IEEE copyrighted paper in a thesis:*

1) The following IEEE copyright/ credit notice should be placed prominently in the references: © [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication]
2) Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis on-line.
3) In placing the thesis on the author's university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity's name goes here]'s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.

If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.




Copyright © 2018 Copyright Clearance Center, Inc. All Rights Reserved. Privacy statement. Terms and Conditions.
Comments? We would like to hear from you. E-mail us at customercare@copyright.com

Copyright permission for Chapter 3

This Agreement between Dr. CHEN KAN ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

| | |
|---|---|
| License Number | 4256731055841 |
| License date | Dec 26, 2017 |
| Licensed Content Publisher | John Wiley and Sons |
| Licensed Content Publication | Quality & Reliability Engineering International |
| Licensed Content Title | Dynamic network monitoring and control of in situ image profiles from ultraprecision machining and biomanufacturing processes |
| Licensed Content Author | Chen Kan,Hui Yang |
| Licensed Content Date | May 19, 2017 |
| Licensed Content Pages | 20 |
| Type of use | Dissertation/Thesis |
| Requestor type | Author of this Wiley article |
| Format | Electronic |
| Portion | Full article |
| Will you be translating? | No |
| Title of your thesis / dissertation | Dynamic network modeling and analysis of large-scale internet of things with manufacturing and healthcare applications |
| Expected completion date | Dec 2017 |
| Expected size (number of pages) | 160 |
| Requestor Location | Dr. CHEN KAN<br>503 Waupelani Dr.<br><br>COLLEGE, PA 16801<br>United States<br>Attn: Dr. CHEN KAN |
| Publisher Tax ID | EU826007151 |
| Billing Type | Invoice |
| Billing Address | Dr. CHEN KAN<br>503 Waupelani Dr.<br><br>COLLEGE, PA 16801<br>United States<br>Attn: Dr. CHEN KAN |
| Total | 0.00 USD |
| Terms and Conditions | |

## TERMS AND CONDITIONS

This copyrighted material is owned by or exclusively licensed to John Wiley & Sons, Inc. or one of its group companies (each a"Wiley Company") or handled on behalf of a society with which a Wiley Company has exclusive publishing rights in relation to a particular work (collectively "WILEY"). By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the billing and payment terms and conditions established by the Copyright Clearance Center Inc., ("CCC's Billing and Payment terms and conditions"), at the time that you opened your RightsLink account (these are available at any time at http://myaccount.copyright.com).

**Terms and Conditions**

- The materials you have requested permission to reproduce or reuse (the "Wiley Materials") are protected by copyright.

- You are hereby granted a personal, non-exclusive, non-sub licensable (on a stand-alone basis), non-transferable, worldwide, limited license to reproduce the Wiley Materials for the purpose specified in the licensing process. This license, **and any CONTENT (PDF or image file) purchased as part of your order,** is for a one-time use only and limited to any maximum distribution number specified in the license. The first instance of republication or reuse granted by this license must be completed within two years of the date of the grant of this license (although copies prepared before the end date may be distributed thereafter). The Wiley Materials shall not be used in any other manner or for any other purpose, beyond what is granted in the license. Permission is granted subject to an appropriate acknowledgement given to the author, title of the material/book/journal and the publisher. You shall also duplicate the copyright notice that appears in the Wiley publication in your use of the Wiley Material. Permission is also granted on the understanding that nowhere in the text is a previously published source acknowledged for all or part of this Wiley Material. Any third party content is expressly excluded from this permission.

- With respect to the Wiley Materials, all rights are reserved. Except as expressly granted by the terms of the license, no part of the Wiley Materials may be copied, modified, adapted (except for minor reformatting required by the new Publication), translated, reproduced, transferred or distributed, in any form or by any means, and no derivative works may be made based on the Wiley Materials without the prior permission of the respective copyright owner.**For STM Signatory Publishers clearing permission under the terms of the STM Permissions Guidelines only, the terms of the license are extended to include subsequent editions and for editions in other languages, provided such editions are for the work as a whole in situ and does not involve the separate exploitation of the permitted figures or extracts,** You may not alter, remove or suppress in any manner any copyright, trademark or other notices displayed by the Wiley Materials. You may not license, rent, sell, loan, lease, pledge, offer as security, transfer or assign the Wiley Materials on a stand-alone basis, or any of the rights granted to you hereunder to any other person.

- The Wiley Materials and all of the intellectual property rights therein shall at all times remain the exclusive property of John Wiley & Sons Inc, the Wiley Companies, or their respective licensors, and your interest therein is only that of having possession of and the right to reproduce the Wiley Materials pursuant to Section 2 herein during the continuance of this Agreement. You agree that you own no right, title or interest in or to the Wiley Materials or any of the intellectual property rights therein. You shall have no rights hereunder other than the license as provided for above in Section 2. No right, license or interest to any trademark, trade name, service mark or other branding

("Marks") of WILEY or its licensors is granted hereunder, and you agree that you shall not assert any such right, license or interest with respect thereto

- NEITHER WILEY NOR ITS LICENSORS MAKES ANY WARRANTY OR REPRESENTATION OF ANY KIND TO YOU OR ANY THIRD PARTY, EXPRESS, IMPLIED OR STATUTORY, WITH RESPECT TO THE MATERIALS OR THE ACCURACY OF ANY INFORMATION CONTAINED IN THE MATERIALS, INCLUDING, WITHOUT LIMITATION, ANY IMPLIED WARRANTY OF MERCHANTABILITY, ACCURACY, SATISFACTORY QUALITY, FITNESS FOR A PARTICULAR PURPOSE, USABILITY, INTEGRATION OR NON-INFRINGEMENT AND ALL SUCH WARRANTIES ARE HEREBY EXCLUDED BY WILEY AND ITS LICENSORS AND WAIVED BY YOU.

- WILEY shall have the right to terminate this Agreement immediately upon breach of this Agreement by you.

- You shall indemnify, defend and hold harmless WILEY, its Licensors and their respective directors, officers, agents and employees, from and against any actual or threatened claims, demands, causes of action or proceedings arising from any breach of this Agreement by you.

- IN NO EVENT SHALL WILEY OR ITS LICENSORS BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR ENTITY FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL, INDIRECT, EXEMPLARY OR PUNITIVE DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, PROVISIONING, VIEWING OR USE OF THE MATERIALS REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION SHALL APPLY NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.

- Should any provision of this Agreement be held by a court of competent jurisdiction to be illegal, invalid, or unenforceable, that provision shall be deemed amended to achieve as nearly as possible the same economic effect as the original provision, and the legality, validity and enforceability of the remaining provisions of this Agreement shall not be affected or impaired thereby.

- The failure of either party to enforce any term or condition of this Agreement shall not constitute a waiver of either party's right to enforce each and every term and condition of this Agreement. No breach under this agreement shall be deemed waived or excused by either party unless such waiver or consent is in writing signed by the party granting such waiver or consent. The waiver by or consent of a party to a breach of any provision of this Agreement shall not operate or be construed as a waiver of or consent to any other or subsequent breach by such other party.

- This Agreement may not be assigned (including by operation of law or otherwise) by you without WILEY's prior written consent.

- Any fee required for this permission shall be non-refundable after thirty (30) days from receipt by the CCC.

- These terms and conditions together with CCC's Billing and Payment terms and conditions (which are incorporated herein) form the entire agreement between you and WILEY concerning this licensing transaction and (in the absence of fraud) supersedes all prior agreements and representations of the parties, oral or written. This Agreement may not be amended except in writing signed by both parties. This Agreement shall be binding upon and inure to the benefit of the parties' successors, legal representatives, and authorized assigns.

- In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall prevail.

- WILEY expressly reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

- This Agreement will be void if the Type of Use, Format, Circulation, or Requestor Type was misrepresented during the licensing process.

- This Agreement shall be governed by and construed in accordance with the laws of the State of New York, USA, without regards to such state's conflict of law rules. Any legal action, suit or proceeding arising out of or relating to these Terms and Conditions or the breach thereof shall be instituted in a court of competent jurisdiction in New York County in the State of New York in the United States of America and each party hereby consents and submits to the personal jurisdiction of such court, waives any objection to venue in such court and consents to service of process by registered or certified mail, return receipt requested, at the last known address of such party.

**WILEY OPEN ACCESS TERMS AND CONDITIONS**

Wiley Publishes Open Access Articles in fully Open Access Journals and in Subscription journals offering Online Open. Although most of the fully Open Access journals publish open access articles under the terms of the Creative Commons Attribution (CC BY) License only, the subscription journals and a few of the Open Access Journals offer a choice of Creative Commons Licenses. The license type is clearly identified on the article.

**The Creative Commons Attribution License**

The Creative Commons Attribution License (CC-BY) allows users to copy, distribute and transmit an article, adapt the article and make commercial use of the article. The CC-BY license permits commercial and non-

**Creative Commons Attribution Non-Commercial License**

The Creative Commons Attribution Non-Commercial (CC-BY-NC)License permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.(see below)

**Creative Commons Attribution-Non-Commercial-NoDerivs License**

The Creative Commons Attribution Non-Commercial-NoDerivs License (CC-BY-NC-ND) permits use, distribution and reproduction in any medium, provided the original work is properly cited, is not used for commercial purposes and no modifications or adaptations are made. (see below)

**Use by commercial "for-profit" organizations**

Use of Wiley Open Access articles for commercial, promotional, or marketing purposes requires further explicit permission from Wiley and will be subject to a fee.
Further details can be found on Wiley Online Library
http://olabout.wiley.com/WileyCDA/Section/id-410895.html

**Other Terms and Conditions:**

**v1.10 Last updated September 2015**

**Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.**

# References


[1] L. Kang and S. L. Albin, "On-line monitoring when the process yields a linear profile," *Journal of Quality Technology,* vol. 32, *No.4,* pp. 418-426, 2000.

[2] K. Kim, M. A. Mahmoud and W. H. Woodall, "On the monitoring of linear profiles," *Journal of Quality Technology,* vol. 35, pp. 317-328, 2003.

[3] C. Zou, Y. Zhang and Z. Wang, "A control chart based on a change-point model for monitoring linear profiles," *IIE Transactions,* vol. 38, *No.12,* pp. 1093-1103, 2006.

[4] D. W. Apley, "Discussion of nonparametric profile monitoring by mixed effects modeling," *Technometrics,* vol. 52, *No.3,* pp. 277-280, 2010.

[5] S. Zhou, Y. Ding, Y. Chen and J. Shi, "Diagnosability study of multistage manufacturing processes based on linear mixed-effects models," *Technometrics,* vol. 45, *No.4,* pp. 312-325, 2003.

[6] K. Paynabar, J. Jin and M. Pacella, "Monitoring and diagnosis of multichannel nonlinear profile variations using uncorrelated multilinear principal component analysis," *IIE Transactions,* vol. 45, *No.11,* pp. 1235-1247, 2013.

[7] Y. Ding, L. Zeng and S. Zhou, "Phase I analysis for monitoring nonlinear profiles in manufacturing processes," *Journal of Quality Technology,* vol. 38, *No.3,* pp. 199-216, 2006.

[8] H. Yang, "Multiscale Recurrence Quantification Analysis of Spatial Cardiac Vectorcardiogram Signals," *IEEE Transactions on Biomedical Engineering,* vol. 58, *No.2,* pp. 339-347, 2011.

[9] S. Zhou, B. Sun and J. Shi, "An SPC monitoring system for cycle-based waveform signals using Haar transform," *Automation Science and Engineering, IEEE Transactions on,* vol. 3, *No.1,* pp. 60-72, 2006.

[10] N. Ruschin-Riminia, I. Ben-Gal and O. Maimona, "Fractal geometry statistical process control for non-linear pattern-based processes," *IIE Transactions,* vol. 45, *No.4,* pp. 355-373, 2013.

[11] G. Liu and H. Yang, "Self-organizing network for group variable selection and predictive modeling," *Annals of Operations Research,* vol. 33, *No.9,* pp. 1399-1418, 2017.

[12] H. Sun, X. Deng, K. Wang and R. Jin, "Logistic regression for crystal growth process modeling through hierarchical nonnegative garrote-based variable selection," *IIE Transactions*, vol. 48, *No.8*, pp. 787-796, 2016.

[13] Y. Chen and H. Yang, "Heterogeneous recurrence representation and quantification of dynamic transitions in continuous nonlinear processes," *Eur. Phys. J. B*, vol. 89, pp. 155, 2016.

[14] C. Kan, C. Cheng and H. Yang, "Heterogeneous recurrence monitoring of dynamic transients in ultraprecision machining processes," *Journal of Manufacturing Systems*, vol. 41, pp. 178-187, 2016.

[15] R. Jin, C. Chang and J. Shi, "Sequential measurement strategy for wafer geometric profile estimation," *IIE Transactions*, vol. 44, *No.1*, pp. 1-12, 2012.

[16] L. Zhang, K. Wang and N. Chen, "Monitoring wafers' geometric quality using an additive Gaussian process model," *IIE Transactions*, vol. 47, *No.1*, pp. 1-18, 2015.

[17] Z. Du, M. K. Jeong and S. G. Kong, "Band selection of hyperspectral images for automatic detection of poultry skin tumors," *Automation Science and Engineering, IEEE Transactions on*, vol. 4, *No.3*, pp. 332-339, 2007.

[18] P. Wilcox, T. Horton, E. Youn, et al, "Evolutionary refinement approaches for band selection of hyperspectral images with applications to automatic monitoring of animal feed quality," *Intelligent Data Analysis*, vol. 18, pp. 25-42, 2014.

[19] H. Yan, K. Paynabar and J. Shi, "Image-based process monitoring using low-rank tensor decomposition," *Automation Science and Engineering, IEEE Transactions on*, vol. 12, pp. 216-227, 2015.

[20] F. M. Megahed, L. J. Wells, J. A. Camelio and W. H. Woodall, "A spatiotemporal method for the monitoring of image data," *Quality and Reliability Engineering International*, vol. 28, *No.8*, pp. 967-980, 2012.

[21] B. Yao, F. Imani, A. Sakpal, E. Reutzel and H. Yang, "Multifractal analysis of image profiles for the characterization and detection of defects in additive manufacturing," *Journal of Manufacturing Science and Engineering*, 2017.

[22] F. Tao, Y. Cheng, L. Xu, L. Zhang and B. Li, "CCIoT-CMfg: Cloud computing and Internet of Things-based cloud manufacturing service system," *IEEE Transactions on Industrial Informatics,* vol. 10, *No.2,* pp. 1435-1442, 2014.

[23] D. Georgakopoulos, P. Jayaraman, M. Fazia, M. Villari and R. Ranjan, "Internet of Things and edge cloud computing roadmap for manufacturing," *IEEE Cloud Computing,* vol. 3, *No.4,* pp. 66-73, 2016.

[24] F. Tao, Y. Zou, L. Xu and L. Zhang, "IoT-based intelligent perception and access of manufacturing resource toward cloud manufacturing," *IEEE Transactions on Industrial Informatics,* vol. 10, *No.2,* pp. 1547-1557, 2014.

[25] K. Thramboulidis and F. Christoulakis, "UML4IoT—A UML-based approach to exploit IoT in cyber-physical manufacturing systems," *Computers in Industry,* vol. 82, pp. 128-137, 2016.

[26] G. Adamson, L. Wang and P. Moore, "Feature-based control and information framework for adaptive and distributed manufacturing in cyber physical systems," *Journal of Manufacturing Systems,* vol. 43, *No.2,* pp. 305-315, 2017.

[27] J. Qin, Y. Liu and R. Grosvenor, "A framework of energy consumption modelling for additive manufacturing using Internet of Things," *Procedia CIRP,* vol. 63, pp. 307-312, 2017.

[28] Y. Tan, Y. Ng and J. Low, "Internet-of-Things enabled real-time monitoring of energy efficiency on manufacturing shop floors," *Procedia CIRP,* vol. 61, pp. 376-381, 2017.

[29] Y. Xu and M. Chen, "Improving Just-in-Time manufacturing operations by using Internet of Things based solutions," *Procedia CIRP,* vol. 56, pp. 326-331, 2016.

[30] Y. Ding, P. Kim, D. Ceglarek and J. Jin, "Optimal sensor distribution for variation diagnosis in multistation assembly processes," *IEEE Transactions on Robotics and Automation,* vol. 19, *No.4,* pp. 543-556, 2003.

[31] C. Pasluosta, H. Gassner, J. Winkler, J. Klucken and B. Eskofier, "An emerging era in the management of Parkinson's disease: wearable technologies and the Internet of Things," *IEEE Journal of Biomedical and Health Informatics,* vol. 19, *No.6,* pp. 1873-1881, 2015.

[32] M. Al-Taee, W. Al-Nuaimy, Z. Muhsin and A. Al-Ataby, "Robot assistant in management of diabetes in children based on the Internet of Things," *IEEE Internet of Things Journal*, vol. 4, *No.2*, pp. 437-445, 2017.

[33] S. Islam, D. Kwak, H. Kabir, M. Hossain and K. Kwak, "The internet of things for health care: a comprehensive survey," *IEEE Access*, vol. 3, pp. 678-708, 2015.

[34] H. Yang, S. T. S. Bukkapatnam and R. Komanduri, "Spatio-temporal representation of cardiac vectorcardiogram (VCG) signals," *Biomedical Engineering Online*, vol. 11, *No.16*, 2012.

[35] H. Yang, "Multiscale Recurrence Quantification Analysis of Spatial Cardiac Vectorcardiogram Signals," *Biomedical Engineering, IEEE Transactions on*, vol. 58, *No.2*, pp. 339-347, 2011.

[36] H. Yang, S. T. S. Bukkapatnam, T. Le and R. Komanduri, "Identification of myocardial infarction (MI) using spatio-temporal heart dynamics," *Med. Eng. Phys.*, vol. 34, *No.4*, pp. 485-497, 2011.

[37] H. Yang, S. Bukkapatnam and R. Komanduri, "Nonlinear adaptive wavelet analysis of electrocardiogram signals," *Physical Review E*, vol. 76, pp. 026214, 2007.

[38] G. E. Dower, A. Yakush, S. B. Nazzal, R. V. Jutzy and C. E. Ruiz, "Deriving the 12-lead electrocardiogram from four (EASI) electrodes," *J. Electrocardiol.*, vol. 21, *No.Supplement 1*, pp. S182-S187, 1988.

[39] D. Dawson, H. Yang, M. Malshe, S. T. S. Bukkapatnam, B. Benjamin and R. Komanduri, "Linear affine transformations between 3-lead (Frank XYZ leads) vectorcardiogram and 12-lead electrocardiogram signals," *J. Electrocardiol.*, vol. 42, *No.6*, pp. 622-630, 2009.

[40] Y. Jeong, M. K. Jeong and O. A. Omitaomu, "Weighted dynamic time warping for time series classification," *Pattern Recognit*, vol. 44, pp. 2231-2240, 2011.

[41] T. Kohonen, "The self-organizing map," *Proceedings of the IEEE*, vol. 78, *No.9*, pp. 1464-1480, 1990.

[42] M. Hagenbuchner and A. C. Tsoi, "A supervised training algorithm for self-organizing maps for structures," *Pattern Recog. Lett.*, vol. 26, *No.12*, pp. 1874-1884, 9, 2005.

[43] P. Rao, S. T. S. Bukkapatnam, O. Beyca, Z. Kong and R. Komanduri, "Real-time identification of incipient surface morphology variations in ultraprecision machining process," *ASME. J. Manuf. Sci. Eng.,* vol. 136, *No.2,* pp. 021008-021008-11, 2014.

[44] C. Park, J. Z. Huang, J. X. Ji and Y. Ding, "Segmentation, inference and classification of partially overlapping nanoparticles," *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* vol. 35, *No.3,* pp. 1-1, 2013.

[45] M. A. Mahmoud and W. H. Woodall, "Phase I analysis of linear profiles with calibration applications," *Technometrics,* vol. 46, *No.4,* pp. 380-391, 2004.

[46] M. A. Mahmoud, P. A. Parker, W. H. Woodall and D. M. Hawkins, "A change point method for linear profile data," *Quality and Reliability Engineering International,* vol. 23, *No.2,* pp. 247-268, 2007.

[47] C. K. H. Koh, J. Shi, W. J. Williams and J. Ni, "Multiple fault detection and isolation using the Haar transform, Part 2: Application to the stamping process," *ASME. J. Manuf. Sci. Eng.,* vol. 121, *No.2,* pp. 295-299, 1999.

[48] J. Jin and J. Shi, "Feature-preserving data compression of stamping tonnage information using wavelets," *Technometrics,* vol. 41, *No.4,* pp. 327-339, 1999.

[49] S. T. S. Bukkapatnam, A. Lakhtakia and S. R. T. Kumara, "Chaos theory based optimal cutting tool chatter control," *Speculations in Science and Technology,* vol. 19, pp. 137-148, 1996.

[50] S. V. Kamarthi, S. R. T. Kumara and P. H. Cohen, "Flank wear estimation in turning through wavelet representation of acoustic emission signals," *ASME. J. Manuf. Sci. Eng.,* vol. 122, *No.1,* pp. 12-19, 2000.

[51] J. Kim, Q. Huang, J. Shi and T. Chang, "Online multichannel forging tonnage monitoring and fault pattern discrimination using principal curve," *ASME. J. Manuf. Sci. Eng.,* vol. 128, *No.4,* pp. 944-950, 2006.

[52] M. K. Jeong and J. Lu, "Wavelet-based SPC procedure for complicated functional data," *International Journal of Production Research,* vol. 44, *No.4,* pp. 729-744, 2006.

[53] Y. Chen and H. Yang, "Heterogeneous recurrence monitoring and control of nonlinear stochastic processes," *Chaos,* vol. 24, *No.1,* pp. 013138, 2014.

[54] M. Newman, "Modularity and community structure in networks," *Proc. Natl. Acad. Sci. USA,* vol. 103, *No.23,* pp. 8577-8582, 2006.

[55] S. Mankad and G. Michailidis, "Structural and functional discovery in dynamic networks with non-negative matrix factorization," *Physical Review E,* vol. 88, pp. 042812, 2013.

[56] D. Hu, P. Ronhovde and Z. Nussinov, "Replica inference approach to unsupervised multiscale image segmentation," *Physical Review E,* vol. 85, *No.1,* pp. 016101, 2012.

[57] G. Palla , I. Derényi , I. Farkas and T. Vicsek, "Uncovering the overlapping community structure of complex networks in nature and society," *Nature,* vol. 435, pp. 814-818, 2005.

[58] D. Martin, C. Fowlkes, D. Tal and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proceedings of the 8th International Conference on Computer Vision,* Vancouver, BC, July, 2001, pp. 416 - 423.

[59] S. Wang and M. Reynolds, "A GLR control chart for monitoring the mean vector of a multivariate normal process," *Journal of Quality Technology,* vol. 45, *No.1,* pp. 18-33, 2013.

[60] D. Du, H. Yang, S. A. Norring and E. S. Bennett, "Multi-scale modeling of glycosylation modulation dynamics in cardiac electrical signaling," in *Proceedings of 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC),* Boston, MA, 2011, pp. 104-107.

[61] D. Du, H. Yang, S. A. Norring and E. S. Bennett, "In-Silico Modeling of Glycosylation Modulation Dynamics in hERG Ion Channels and Cardiac Electrical Signals," *Biomedical and Health Informatics, IEEE Journal of,* vol. 18, *No.1,* pp. 205-214, 2014.

[62] Z. Wang, S. T. S. Bukkapatnam, S. R. T. Kumara, Z. Kong and Z. Katz, "Change detection in precision manufacturing processes under transient conditions," *CIRP Annals - Manufacturing Technology,* vol. 63, *No.1,* pp. 449-452, 2014.

[63] C. Cheng, Z. Wang, W. Hung, S. T. S. Bukkapatnam and R. Komanduri, "Ultra-precision machining process dynamics and surface quality monitoring," *Procedia Manufacturing,* vol. 1, pp. 607–618, 2015.

[64] J. Lee, Y. Shin, M. Kim, et al, "A simplified machine-tool power-consumption measurement procedure and methodology for estimating total energy consumption," *ASME Trans. Journal of Manufacturing Science and Engineering,* vol. 138, *No.5,* pp. 051004-1-051004-9, 2015.

[65] C. Kan, Y. Chen, F. Leonelli and H. Yang, "Mobile sensing and network analytics for realizing smart automated systems towards health internet of things," in *Automation Science and Engineering (CASE), 2015 IEEE International Conference on,* Gothenburg, Aug. 24-28, 2015, pp. 1072-1077.

[66] J. Jin and J. Shi, "Diagnostic feature extraction from stamping tonnage signals based on design of experiments," *ASME Trans. Journal of Manufacturing Science and Engineering,* vol. 122, *No.2,* pp. 360-369, 1999.

[67] J. Jin and J. Shi, "Press tonnage signal decomposition and validation analysis for transfer or progressive die processes," *ASME Trans. Journal of Manufacturing Science and Engineering,* vol. 127, *No.1,* pp. 231-235, 2005.

[68] P. Rao, S. Bukkapatnam, O. Beyca, Z. Kong and R. Komanduri, "Real-time identification of incipient surface morphology variations in ultraprecision machining process," *ASME Trans. Journal of Manufacturing Science and Engineering,* vol. 136, *No.2,* pp. 021008-1-021008-11, 2014.

[69] O. Beyca, P. Rao, Z. Kong, S. Bukkapatnam and R. Komanduri, "Heterogeneous sensor data fusion approach for real-time monitoring in ultraprecision machining (UPM) process using non-parametric Bayesian clustering and evidence theory," *Automation Science and Engineering, IEEE Transactions on,* vol. 13, *No.2,* pp. 1033-1044, 2016.

[70] P. Rao, J. Liu, D. Roberson, Z. Kong and C. Williams, "Online real-time quality monitoring in additive manufacturing processes using heterogeneous sensors," *ASME Trans. Journal of Manufacturing Science and Engineering,* vol. 137, *No.6,* pp. 061007-1-061007-12, 2015.

[71] C. Shao, K. Paynabar, T. Kim, et al, "Feature selection for manufacturing process monitoring using cross-validation," *Journal of Manufacturing Systems,* vol. 32, pp. 550-555, 2013.

[72] M. Seera, C. Lim, D. Ishak and H. Singh, "Fault detection and diagnosis of induction motors using motor current signatue analysis and a hybrid FMM-CART model," *IEEE Transactions on Neural Networks and Learning Systems,* vol. 23, *No.1,* pp. 97-108, 2012.

[73] S. Emec, J. Kruger and G. Seliger, "Online fault-monitoring in machine tools based on energy consumption analysis and non-invasive data acquisition for improved resource-efficiency," *Procedia CIRP*, vol. 40, pp. 236-243, 2016.

[74] Y. Liu and X. Xu, "Industry 4.0 and cloud manufacturing: a comparative analysis," *ASME Trans. Journal of Manufacturing Science and Engineering,* vol. 139, *No.3,* pp. 034701-1-034701-8, 2017.

[75] P. Daugherty, P. Banerjee, W. Negm and A. Alter, "Driving unconventional growth through the industrial internet of things," *New York: Accenture*, 2014.

[76] S. Kumara, L. Cui and J. Zhang, "Sensors, networks and internet of things - research challenges in health care," in *IIWeb '11 Proceedings of the 8th International Workshop on Information Integration on the Web: In Conjunction with WWW 2011,* Hyderabad, India, 2011, .

[77] S. Yang, B. Bagheri, H. Kao and J. Lee, "A unified framework and platform for designing of cloud-based machine health monitoring and manufacturing systems," *ASME Trans. Journal of Manufacturing Science and Engineering,* vol. 137, *No.4,* pp. 040914-1-040914-6, 2015.

[78] N. Raghavan and T. Waghmare, "DPAC: An object-oriented distributed and parallel computing framework for manufacturing applications," *IEEE Transactions on Robotics and Automation*, vol. 18, *No.4,* pp. 431-443, 2002.

[79] D. Mourtzis, E. Vlachou, N. Milas and N. Xanthopoulos, "A cloud-based approach for maintenance of machine tools and equipment based on shop-floor monitoring," *Procedia CIRP*, vol. 41, pp. 655-660, 2016.

[80] F. Tao, Y. LaiLi, L. Xu and L. Zhang, "FC-PACO-RM: a parallel method for service composition optimal-selection in cloud manufacturing system," *IEEE Transactions on Industrial Informatics,* vol. 9, *No.4,* pp. 2023-2033, 2012.

[81] H. Yang, C. Kan, G. Liu and Y. Chen, "Spatiotemporal Differentiation of Myocardial Infarctions," *Automation Science and Engineering, IEEE Transactions on,* vol. 10, *No.4,* pp. 938-47, 2013.

[82] H. Yang and G. Liu, "Self-organized topology of recurrence-based complex network," *Chaos*, vol. 23, *No.4,* pp. 043116, 2013.

[83] Y. Chen and H. Yang, "Sparse modeling and recursive prediction of space-time dynamics in stochastic sensor networks," *IEEE Transactions on Automation Science and Engineering,* vol. 13, *No.1,* pp. 215-226, 2016.

[84] C. Kan, F. M. Leonelli and H. Yang, "Map reduce for optimizing a large-scale dynamic network – internet of hearts," in *Proceedings of 2016 IEEE Engineering in Medicine and Biology Society Conference (EMBC),* Orlando, FL, 2016, pp. 2962-2965.

[85] P. Demartines and J. Herault, "Curvilinear component analysis: a self-organizing neural network for nonlinear mapping of data sets," *Neural Networks, IEEE Transactions on,* vol. 8, *No.1,* pp. 148-154, 1997.

[86] M. Li, T. Zhang, Y. Chen and A. Smola, "Efficient mini-batch training for stochastic optimization," in *KDD '14 Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,* New York, 2014, pp. 661-670.

[87] D. Mozaffarian, E. Benjamin, A. Go, et al, "Heart disease and stroke statistics - 2016 update: a report from the American Heart Association," *Circulation,* vol. 136, *No.18,* 2017.

[88] G. De Luca, H. Suryapranata, J. Ottervanger and E. Antman, "Time delay to treatment and mortality in primary angioplasty for acute myocardial infarction: every minute of delay counts," *Circulation,* vol. 109, *No.10,* pp. 1223-1225, 2004.

[89] Y. Zhang, Y. Zheng, W. Lin, H. Zhang and X. Zhou, "Challenges and opportunities in cardiovascular health informatics," *IEEE Transactions on Biomedical Engineering,* vol. 60, *No.3,* pp. 633-642, 2013.

[90] V. Elmberg, J. Almer, O. Pahlm, G. Wagner, H. Engblom and M. Ringborn, "A 12-lead ECG-method for quantifying ischemia-induced QRS prolongation to estimate the severity of the acute myocardial event," *Journal of Electrocardiology,* vol. 49, *No.3,* pp. 272-277, 2016.

[91] M. Meo, V. Zarzoso, O. Meste, D. Latcu and N. Saoudi, "Spatial variability of the 12-Lead surface ECG as a tool for noninvasive prediction of catheter ablation outcome in persistent atrial fibrillation," *IEEE Transactions on Biomedical Engineering,* vol. 60, *No.1,* pp. 20-27, 2012.

[92] O. Perlman, A. Katz, G. Amit and Y. Zigel, "Supraventricular tachycardia classification in the 12-Lead ECG using atrial waves detection and a clinically based tree scheme," *IEEE Journal of Biomedical and Health Informatics,* vol. 20, *No.6,* pp. 1513-1520, 2016.

[93] C. Cheng, C. Kan and H. Yang, "Heterogeneous recurrence analysis of heartbeat dynamics for the identification of sleep apnea events," *Computers in Biology and Medicine,* vol. 75, pp. 10-18, 2016.

[94] Y. Chen and H. Yang, "Self-organized neural network for the quality control of 12-lead ECG signals," *Physiol. Meas.,* vol. 33, *No.9,* pp. 1399, 2012.

[95] H. Yang and F. Leonelli, "Self-organizing visualization and pattern matching of vectorcardiographic QRS waveforms," *Comput. Biol. Med.,* vol. 79, pp. 1-9, 2016/10, .

[96] S. Man, C. Rahmattulla, A. Maan, et al, "Role of the vectorcardiogram-derived spatial QRS-T angle in diagnosing left ventricular hypertrophy." *Journal of Electrocardiology,* vol. 45, *No.2,* pp. 154-160, 2012.

[97] M. Yasin, T. Tekeste, H. Saleh, B. Mohammad, O. Sinanoglu and M. Ismail, "Ultra-low power, secure IoT platform for predicting cardiovascular diseases," *IEEE Transactions on Circuits and Systems-I: Regular Papers,* vol. 64, *No.9,* pp. 2624-2637, 2017.

[98] G. Yang, M. Mantysalo, X. Zhou, et al, "A health-IoT platform based on the integration of intelligent packaging, unobtrusive bio-sensor, and intelligent medicine box," *IEEE Transactions on Industrial Informatics,* vol. 10, *No.4,* pp. 2180-2191, 2014.

[99] C. Kan and H. Yang, "Internet of hearts - large-scale stochastic network modeling and analysis of cardiac electrical signals," in *Stochastic Modeling and Analytics in Healthcare Systems*, J. Li, N. Kong and X. Xie, Eds. World Scientific, 2017, pp. 211-251.

[100] Z. Eapen, M. Turakhia, M. McConnell, et al, "Defining a mobile health roadmap for cardiovascular health and disease," *Journal of American Heart Association*, vol. 5, *No.7*, pp. 1-6, 2016.

[101] K. Jang, H. Chuang, S. Xu, et al, "Soft network composite materials with deterministic and bio-inspired designs," *Nature Communications*, vol. 18, *No.6*, pp. 1-11, 2015.

[102] S. Israel, J. Irvine, A. Cheng, M. Wiederhold and B. Wiederhold, "ECG to identify individuals," *Pattern Recognition*, vol. 38, pp. 133-142, 2005.

[103] C. Kan and H. Yang, "Dynamic network monitoring and control of in-situ image profiles from ultraprecision machining and biomanufacturing processes," *Quality and Reliability Engineering International*, 2017.

[104] A. Goldberger, L. Amaral, L. Glass, et al, "PhysioBank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 23, pp. e215-e220, 2000.

[105] F. Jager, A. Taddei, G. Moody, M. Emdin, G. Antolic and R. Dorn, "Long-term ST database: a reference for the development and evaluation of automated ischaemia detectors and for the study of the dynamics of myocardial ischaemia," *Medical & Biological Engineering & Computing*, vol. 41, *No.2*, pp. 172-183, 2003.

[106] D. Horgan and C. Murphy, "On the convergence of the chi square and noncentral chi square distributions to the normal distribution," *IEEE Communication Letters*, vol. 17, No. 12, pp. 2233-2236, 2013.

[107] G. Box, J. Hunter and W. Hunter, *Statistics for experimenters*: *design, innovation and discovery*, 2nd ed. New York: Wiley, 2005.

[108] S. de Waele and P. Broersen, "Order selection for vector autoregressive models," *IEEE Transactions on Signal Processing*, vol. 51, *No.* 2, pp. 427-433, 2003.

# Vita

**Chen Kan**

## EDUCATION

**Ph.D.** **Industrial and Manufacturing Engineering,** Pennsylvania State University, May 2018
**MSc.** **Industrial and Management Systems Engineering**, University of South Florida, May 2012
**B.Sc.** **Electrical Engineering**, China University of Mining and Technology (Beijing), May 2010

## SELECTED PUBLICATIONS

1. **C. Kan**, Y. Chen, and H. Yang, "Multiscale quality control of telemedicine ECG signal acquisition," Telehealthcare Computing and Engineering: Principles and Design, CRC Press, Eds: Hu Fei, ISBN: 978-1578088027, March 2012.
2. **C. Kan** and H. Yang, "Internet of Hearts – large-scale stochastic network modeling and analysis of cardiac electrical signals," Stochastic Modelling and Analytics in Healthcare Systems, Eds: Jingshan Li, in press, 2017.
3. **C. Kan**, C. Cheng and H. Yang, "Heterogeneous recurrence monitoring of dynamic transients in Ultra-precision machining processes," ***Journal of Manufacturing Systems***, Vol. 41, pp. 178-187, 2016.
4. C. Cheng, **C. Kan** and H. Yang, "Heterogeneous recurrence modeling and analysis of heartbeat dynamics for the identification of sleep apnea events," ***Computers in Biology and Medicine***, Vol. 75, No. 1, pp. 10-18, 2016.
5. **C. Kan**, K.P. Yip, and H. Yang, "Two-phase greedy pursuit algorithm for automatic detection and characterization of transient Calcium signaling," ***IEEE Journal of Biomedical and Health Informatics***, Vol. 19, No. 2, pp. 687-697, 2015.
6. H. Yang, **C. Kan**, G. Liu, and Y. Chen "Spatiotemporal differentiation of myocardial infarctions," ***IEEE Transactions on Automation Science and Engineering***, Vol. 10, No. 4, pp. 938-947, 2013.
7. **C. Kan** S. Kumara and H. Yang, "Fast parallel computing algorithms for large-scale machine network modeling and condition monitoring," under 2nd round review, ***Journal of Manufacturing Science and Engineering***, 2017.
8. **C. Kan** and H. Yang, "Dynamic network modeling of in-situ image profiles for statistical process control – applications in ultra-precision machining and bio-manufacturing process," ***Quality and Reliability Engineering International***, Vol. 33, No. 8, pp. 2003-2022, 2017.

## SELECTED HONORS & AWARDS

- **2016, Best Paper Finalist**, Quality Control and Reliability Engineering Division (QCRE), IIE Annual Conference, Anaheim, CA.
- **2015, NSF Travel Grant** for presenting at 2015 IIE Annual Conference, Nashville, TN.
- **2015, Best Paper Finalist**, Process Industries Track, 2015 IIE Annual Conference, Nashville, TN.
- **2015, Featured Article**, IEEE Journal of Biomedical and Health Informatics.
- **2014, The 2nd Place** in IIE Mobile App Competition, ISERC 2014, Montreal, Canada.
- **2014, Entrepreneurial Lead**, NSF IIP-1447289, NSF I-Corps Team for the market research and commercialization of the Mobile E-Network Smart Health (MESH) technology.