

3rd CIRP Conference on BioManufacturing

# Image-guided Quality Control of Biomanufacturing Process

Chen Kan, Ruimin Chen, Hui Yang\*

*Complex Systems Monitoring, Modeling and Control Lab, Pennsylvania State University, University Park, PA 16802, USA*

\* Corresponding author. Tel.: +01-814-865-7397; fax: +01-814-863-4345. E-mail address: [huy25@psu.edu](mailto:huy25@psu.edu)

## Abstract

Technical advances in 3D metrology bring the increasing availability of imaging data, which are critical to quality inspection and process improvement. Dealing with 3D imaging data has become a general problem facing both traditional and next-generation innovation practices in biotechnology. Traditional methodologies in statistical quality control focus on key characteristics of the product, and are limited in the ability to model spatiotemporal patterns in imaging streams. This paper presents a dynamic network methodology for monitoring and control of high-dimensional imaging streams. The developed methodology is implemented and evaluated for process monitoring of living cells during the synthesis of bio-products.

© 2016 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Peer-review under responsibility of the scientific committee of the 3rd CIRP Conference on BioManufacturing 2017

**Keywords:** 3D imaging, quality control, dynamic network, process monitoring

## 1. Introduction

Advances in life science and engineering have spurred the use of biological systems to create bio-products, such as vaccines, biomaterials, and amino acids to add in medicine, biology, and industrial applications. The past decade has witnessed the booming of bio-tech industry and a fast growth of its global market. For example, biopharmaceuticals, one of the most important bio-products, not only replaces traditional small-molecule drugs with new therapeutics but also provides new types of treatment for previously untreatable diseases. According to [1], the global market of biopharmaceuticals will grow 12%, and it is estimated to reach \$344 billion in 2018. With the increasing understanding of complex biological systems in genetic and cellular levels and the development of single-use technology, biomanufacturing will become highly scalable and flexible and produces bio-products with high quality and controlled costs.

Biomanufacturing contains a broad range of areas. In mechanical engineering, researchers mainly focus on bio-mechatronics and bio-fabrication. For bio-mechatronics, machine-assisted systems, e.g., diagnostic equipment, robots, and surgical devices are designed to enhance clinical performance in medical interventions [2]. For example, Hansen

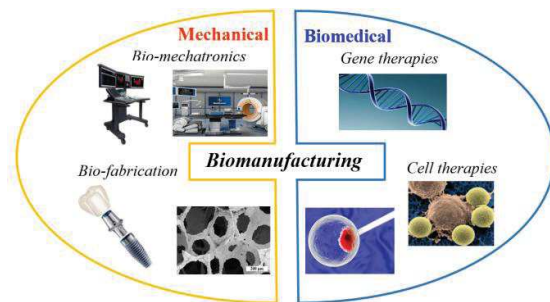


Fig. 1. Biomanufacturing defined from two perspectives: 1) mechanical: bio-mechatronics [1] and bio-fabrication [4-5]; 2) biomedical: gene and cell therapies (image from Penn Medicine).

Medical produced robotic catheters for cardiac arrhythmia treatment [3]. Incorporated with 3D mapping system and fluoroscopy, catheters are guided to the target region of the heart to deliver medications. Also, Medtronic developed surgical navigation systems to improve the precision of procedures. Advanced imaging technology and visualization solutions help surgeons to track surgical instruments during the operation. For bio-fabrication, investigations are conducted to produce natural (e.g., bones and tissues) and artificial (e.g., platinum group materials) biomaterials that can be used for

arthroscopy, oncology, and dental treatment [2]. For example, dental implants are forged with elevated temperatures [4]. Pressures are monitored to ensure the quality and reliability of the produced implant. Also, scaffolds are fabricated for bone tissue engineering. Superior permeability and high biocompatibility are required when selecting proper scaffold materials (e.g., porous titania) [5].

Cellular biomanufacturing, on the other hand, has been increasingly investigated in the biomedical community. Cellular biomanufacturing administers live cells or specific types of DNA for eliciting therapeutic efficacies and produces functional defined therapeutic cells, e.g., stem cells and T cells. Cellular biomanufacturing has a complex workflow, which involves cell isolation, activation, genetic modification, expansion and many others. Gene therapies and cell therapies produced from cellular biomanufacturing hold tremendous promise in treating complex disorders, such as cancer, cardiovascular diseases, and neurological diseases [6]. For example, normal genes of clotting factors are introduced to the liver of haemophilia patients to produce FVIII and FIX proteins for the correction of bleeding problems. Such gene therapy performs better than injected proteins that traditionally used to treat haemophilia. Also, T cells, a type of immune cells collected from the blood, are genetically engineered to produce chimeric antigen receptors (CARs). CARs enable T cells to recognize the antigen on tumors and further kill the cancer cells. Cellular biomanufacturing has a strong potential to revolutionize the healthcare industry in various contexts, including regenerative medicine, cancer treatment, and personalized medicine.

However, the inherent complexity of biological systems poses a great challenge for biomanufacturing. For example in bio-fabrication, the machining of a neural bone involves high-precision cutting and requires a careful control of multiple factors such as cut position, force, and heat generation [2]. Out of controls will significantly influence the effectiveness of treatment and even result in long-term problems. Also, cellular manufacturing is aimed at producing reproducible drugs with standardized characterization. Notably, large molecule drugs are most often produced using live microbial cells in the bioreactor. Because cells are alive and the intrinsically heterogeneous cell population, produced therapeutic cells have a large variability even within a single batch. As such, cellular biomanufacturing at large scale is expensive and unpredictable, which hampers the broad translation of cell therapies into clinical practice and results in the uncertainty in clinical trial outcomes. Hence, process monitoring and quality control of biomanufacturing are imperative to improve the productivity and ensure the integrity and efficacy of bio-products.

Existing quality control approaches of biomanufacturing rely on off-line, post-build inspection. Despite the long inspection procedure, such inefficient approaches result in a low yield and a high cost. In the past decade, advances in metrology bring the increasing availability of imaging data in biomanufacturing, which are critical to quality inspection and process improvement. For example in bio-fabrication, 2-D and 3-D microscopic images are collected to assess the microstructure of scaffolds [5]. Also, microscopic images were used to monitor crack propagation and chip formation in the

bone machining [7]. In cellular manufacturing, high-resolution video microscope is utilized for process monitoring and quality control. Notably, images provide anatomical information at subcellular level, which contributes to the understanding of structural and physical properties of cell-based products [8]. Furthermore, details of cellular dynamics are captured by the stream of dynamic images. This provides a great opportunity for the characterization of time-varying activities in biomanufacturing, such as the culture of cell lines and delivery of transgenes. This, in turn, enhances the capability of multi-stage and multi-scale process monitoring and quality control of biomanufacturing [9].

In this study, we develop a dynamic network approach for the image-guided monitoring and control of biomanufacturing processes. Specifically, we represent images as networks in the high-dimensional space and a graph-theoretic approach is introduced to characterize network topological structures (i.e., network communities). Further, statistics are extracted from network communities and a new control charting approach is developed for anomaly detection. The developed methodology is implemented and evaluated for process monitoring of living cells during the synthesis of bio-products. The remainder of this paper is organized as follows: Section II introduces the background of statistical process control (SPC) methods and control charts. Section III presents the developed methodology. Case studies are demonstrated in Section IV. Section V concludes and discusses this present investigation.

## 2. Research Background

Process control and quality improvement have been increasingly investigated over the past century. Early SPC methods primarily focused on key quality characteristics of product. For example, the diameter of holes drill in a certain metal clamp, the Rockwell hardness of the pin head, and the tensile strength of a material. Control charts were developed to evaluate the conformation to specifications by monitoring discrete data points. For example, two types of Shewhart charts, i.e., X-bar and R charts were widely used to monitor the mean and variance of quality characteristics. However, both charts used only the information in the most recently obtained sample, which limited their responsiveness to small process shifts. Thus, the cumulative sum (cusum) chart and exponentially weighted moving average (EWMA) chart were developed to address this challenge. The basic idea of cusum and EWMA was to incorporate the information of past observations. As such, the EWMA control chart was sensitive to small shifts that traditional Shewhart charts fail to capture.

In the past decade, advances in sensing technology brought the era of profile monitoring. Instead of discrete data points, time-dependent functional data were obtained from online sensing and monitoring systems. To address the data complexity, multivariate control charts were more and more used in the process control. The most widely used multivariate chart was the Hotelling  $T^2$  chart, which was an extension of Shewhart chart for the monitoring of the mean vector of the process. Also,  $|S|$  chart with sample generalized variance was extended from the univariate  $s$  chart for variability monitoring in multivariate processes. In addition, EWMA chart was also

extended to multivariate quality control and multivariate EWMA chart was developed to provide more sensitivity to small shifts.

Besides the extension of univariate charts to multivariate charts, researchers addressed the data complexity by dimensional reduction algorithms. For example, Ding *et al.* [10] investigated both principal component analysis (PCA) and independent component analysis (ICA) for the Phase I monitoring of nonlinear profiles. Cycle-based forging tonnage signals were transformed into lower dimensional space by PCA and ICA to reveal pertinent quality information. Zhou *et al.* [11] decomposed cycle-based waveforms by Haar wavelet. The location of process shift and its magnitude were then detected by monitoring the wavelet coefficients.

It may be noted that most of above SPC methods and tools focus on key product characteristics, 1-D linear and nonlinear profiles. However, they are limited in their ability for the image-based process monitoring and control. With the increasing availability of images in biomanufacturing, there is a need to develop new approaches to leverage high-dimensional images for the quality control of bio-products. Most existing image-based SPCs, nevertheless, are limited in fault detection of snap-shot images. For example, Zhang *et al.* selected a number of sampling points from the wafer image to measure the variations of wafer thickness [12]. Wilcox *et al.* analyzed a subset of spectral bands from hyperspectral images for the monitoring of animal feed quality [13]. Thus, there is an urgent need to develop new SPC approaches that extract useful information from in-situ images in biomanufacturing and exploit the acquired information for process monitoring and control.

### 3. Research Methodology

In this study, a dynamic network scheme is developed to model the sequence of high-dimensional microscopic images collected in the biomanufacturing process. First, image profiles are represented as a network. Then, a graph-theoretic approach is introduced to detect community patterns in the network. Finally, a control chart is constructed to characterize variations in sub-graph structures, which, in turn, facilitates change-point detection in the image stream.

#### 3.1. Construction of Networks from Images

Image profiles can be represented as both unweighted and weighted networks. For an unweighted network, a binary variable, i.e., 0 or 1 is assigned between each pair of nodes to characterize their connectivity. In contrast, a weighted network also characterizes how close each pair of nodes are connected by assigning a weight. In the literature, a weighted network is usually represented by  $G = (V, E, W)$ , where  $V$ ,  $E$ , and  $W$  denote the set of nodes, edges and weights, respectively. In this study, we develop a regularized function to compute the weight  $W_{ij}$  between node  $i$  and  $j$  as:

$$W_{i,j} = \exp \left( \frac{\alpha \|x_i - x_j\| + (1-\alpha) \|s_i - s_j\|}{2 \times l^2} \right) \quad (1)$$

where  $l$  denotes the kernel bandwidth. It may be noted that the weight is calculated based on both intensity differences  $\|x_i - x_j\|$  and spatial distances  $\|s_i - s_j\|$  between node  $i$  and  $j$ . The regularization parameter  $\alpha \in [0, 1]$  balances these two terms. When  $\alpha$  is larger, the network primarily focus on intensity differences among pixels. When  $\alpha$  is smaller, the network focuses more on distances between spatial locations of pixels. As such, a small weight will be assigned between nodes  $i$  and  $j$  when their corresponding pixels have similar intensity values or they are located closely. Notably, Eq. (1) can be used to construct networks from not only gray-scale images, but also RGB images. For gray-scale images, the intensity value  $x_i$  is denoted by a scalar within  $[0, 255]$ . For RGB images, the intensity  $x_i$  is represented by a vector with 3 elements of primary colors, i.e., red, green and blue.

#### 3.2. Characterization of Network Community Patterns

A network community is a cluster of tightly connected network nodes. That is, nodes within a community have stronger internal connectivity than their connectivity to the remainder of the network. As network weights preserve the relationship (e.g., intensity differences and spatial distances) amongst image pixels, each community can be interpreted as a specific pattern in the image. Characterization of network communities provides a great opportunity to study patterns and their interactions in the image.

In this study, an entropy rate clustering approach is implemented to optimally partition the network to obtain compact and homogeneous communities [14]. Assume  $\Phi \subseteq E$  is a subset of edges selected from the full edge set  $E$ . The entropy rate of the random walk on the subset  $S$  is then:

$$\Gamma(\phi) = -\sum_i \eta_i \sum_j p_{ij}(\phi) \log(p_{ij}(\phi)) \quad (2)$$

where  $p_{ij}(\phi) = W_{i,j} / \sum_j W_{i,j}$  measures the transition probability from node  $i$  and  $j$  if the edge  $e_{ij} \in \Phi$ . Otherwise,  $p_{ij}(\phi) = 0$ .  $\eta_i$  characterizes the stationary distribution of transition probability, i.e.,  $\eta_i = W_i / \sum_i W_i$ .

Notably, it is important to maintain similar size of detected communities to facilitate the characterization of image patterns in different scales. Thus, a balancing term is further introduced to penalty the large difference in the size of communities. Assume the edge set  $\Phi$  is partitioned into  $K$  communities, i.e.,  $\Phi = \{\Phi_1, \Phi_2, \dots, \Phi_K\}$  with  $n_k$  denotes the number of nodes in the  $k^{th}$  community. The distribution  $D_\phi$  of cluster membership in  $\Phi$  can be represented as:

$$p_{D_\phi}(k) = \frac{n_k}{N} \quad (3)$$

where  $N$  is the total number of nodes in the network (i.e., the total number of pixels in the image). Hence, the balancing term can be formulated as:

$$B(\phi) = -\sum_i p_{D_\phi}(k) \log(p_{D_\phi}(k)) - k \quad (4)$$

The first term, i.e.,  $-\sum_i p_{D_\phi}(k) \log(p_{D_\phi}(k))$  computes the entropy of the membership distribution  $D_\phi$  and it favors communities with similar sizes.

Therefore, the objective function of the entropy rate clustering approach can be obtained by combining Eq. (2) and Eq. (4) as:  $E(\Phi) = \Gamma(\Phi) + \beta B(\Phi)$ . Here,  $\beta$  is a regularization parameter to give weights on the balancing term  $B(\Phi)$ . By maximizing  $E(\Phi)$ , the random graph walk has a large entropy rate and the cluster membership distribution has a large entropy. As such, strongly connected nodes will be assigned into the same community and the number of nodes in each community will be balanced.

Maximizing the objective function  $E(\Phi)$  involves a greedy heuristic. Starting from  $\Phi = \emptyset$ , edges in  $E$  are sequentially added into  $\Phi$ . In other words, the algorithm selects one edge in each step that best increases the value of objective function and adds it to the set  $\Phi$ . It iterates until the number of communities  $K$  is reached. Notably, the algorithm needs to evaluate the gain of objective function for every possible candidate in  $E$  at each iteration. Thus, it will be time consuming when partitioning a large network (i.e., the number of nodes  $N$  is large).

As both  $\Gamma(\Phi)$  and  $B(\Phi)$  are proven to be monotonically increasing submodular functions [14], a lazy greedy algorithm can be used to improve the computational efficiency [15]. The idea of lazy greedy algorithm is to exploit the submodularity and employ the diminishing return property. For the submodular function  $E$ , the diminish return property gives:

$$\begin{aligned} E(\phi \cup \{e_1\}) - E(\phi) \\ \geq E(\phi \cup \{e_1, e_2\}) - E(\phi \cup \{e_2\}) \end{aligned} \quad (5)$$

That is, the marginal gain obtained by adding the edge  $e_1$  to  $\phi$  is smaller if it is added later. Thus, instead of updating the gains for all remaining edges after picking one from  $E$  to  $\Phi$ , only the marginal gain of the edge that ranks on the top is updated. In this way, the computational complexity can be significantly reduced and it makes the algorithm well-suited for community detection of large networks.

### 3.3. Generalized Likelihood Ratio Chart for Process Monitoring

After community detection, homogeneous pixels in the image are clustered into one community. As such, it is proper to assume that pixel intensities within one community follow a normal distribution with mean  $\mu_k$  and variance  $\sigma_k^2$ . Thus, we can represent the  $i^{th}$  image profile as:  $\mathbf{y}^{(i)} = (\bar{\mathbf{x}}_1^{(i)}, \dots, \bar{\mathbf{x}}_K^{(i)}, s_1^{(i)}, \dots, s_K^{(i)})^T$  where  $\bar{\mathbf{x}}_k^{(i)}$  is the sample mean for the  $k$ th community,  $s_k^{(i)}$  is the sample standard distribution. Further, a multivariate normal distribution is assumed for  $\mathbf{y}^{(i)}$  as  $\mathbf{y}^{(i)} \sim \text{MVN}(\mu_0, \Sigma)$ . Here,  $\mu_0$  and  $\Sigma$  are in-control mean and covariance matrix of feature vectors. If there is a process shift

at the  $\tau^{th}$  image,  $\mu_0$  will be shifted to  $\mu_1$  and  $\Sigma$  is assumed as unchanged.

In this study, a network generalized likelihood ratio (NGLR) chart is used to monitor the change in  $\mathbf{y}^{(i)}$ . The NGLR chart calculates the likelihood function of both in-control and out-of-control distributions and computes the maximum likelihood ratio statistic at the  $m^{th}$  image:

$$R_{m,y} = \max_{0 \leq \tau \leq m} \frac{m-\tau}{2} (\mu_{m,\tau,y} - \bar{y})^T S^{-1} (\mu_{m,\tau,y} - \bar{y}) \quad (6)$$

where  $\hat{\mu}_{m,\tau,y} = \frac{1}{(m-\tau)} \sum_{i=\tau+1}^m \mathbf{y}^{(i)}$  is the estimation of shifted process mean and  $\bar{y}$  and  $S$  are sample mean and sample covariance estimated from in-control data.

However, Eq. (6) involves the calculation of  $S^{-1}$ , i.e., the inversion of sample covariance matrix. It may be noted that if two communities share similar distribution,  $S$  can be singular, which makes the computation of Eq. (6) impractical. As discussed in our previous study [16], this challenge can be addressed by transforming  $\mathbf{y}^{(i)}$  into the eigen space and using uncorrelated principal components  $\mathbf{Z}(i,:)^T$  to compute equivalent NGLR statistics  $R_{m,Z}$ :

$$R_{m,Z} = \max_{m-w \leq \tau \leq m} \frac{m-\tau}{2} (\mu_{m,\tau,Z})^T S_Z^{-1} (\mu_{m,\tau,Z}) \quad (7)$$

where  $\hat{\mu}_{m,\tau,Z} = \frac{1}{(m-\tau)} \sum_{i=\tau+1}^m \mathbf{Z}(i,:)^T$  is the estimation of shifted mean from the principal components. If we only keep the first  $d$  principal components, the estimation of shifted mean becomes:

$$\mu_{m,\tau,Z} = \frac{1}{m-\tau} \sum_{i=\tau+1}^m \mathbf{Z}(i,1:d)^T \quad (8)$$

By replacing  $\hat{\mu}_{m,\tau,Z}$  with  $\hat{\mu}_{m,\tau,\bar{Z}}$  in Eq. (7), we can have the NGLR statistics in the reduced dimension.

## 4. Experiments and Results

In this study, we first evaluate the proposed community detection algorithm on benchmark images. Then, we evaluate and validate the proposed dynamic network scheme for the monitoring of real-world biomanufacturing process. Our experiments and results are detailed in the following sections.

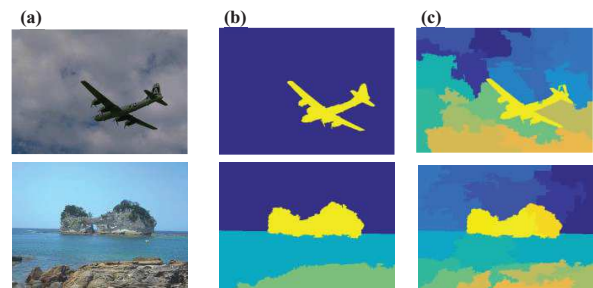


Fig. 2. Community detection using entropy rate clustering approach: (a) original RGB images; (b) smaller number of communities; (c) larger number of communities.



#### 4.1. Community Detection Results on Benchmark Images

We first evaluate the entropy rate clustering approach introduced in Section 3.2 on benchmark images from the Berkeley BSDS500 dataset [17]. It may be noted that a small number of communities highlights the main object from the background. For example, two communities separate the plane from the sky (see the top figure in Fig. 2b). Four communities depict the sky, island, water, and rock (see the bottom figure in Fig. 2b). On the contrary, a larger number of communities reveal detailed information from the image. For example, twenty communities not only characterize detailed patterns on the plane but also show the morphology of clouds (see the top figure in Fig. 3c).

#### 4.2. Process Monitoring of Biomanufacturing Process

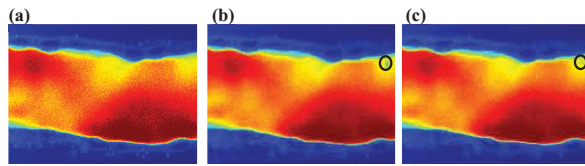


Fig. 3. (a) Cellular image in the biomanufacturing process; (b) Out-of-control image with a vesicle (intensity shift = 15 and size = 8); (c) Out-of-control image with a vesicle (intensity shift = 4 and size = 4).

Fig. 3a shows microscopic images of a living cell. Cytoplasm is shown in red and yellow, and extracellular fluids are in dark blue. It may be noted that the quality of biomanufactured cells directly impacts the effectiveness of gene and cell therapies. Thus, high-resolution video microscope is commonly used to monitor the biomanufacturing process for quality assurance in the cellular level. As opposed to static images, video microscopy brings a stream of dynamic images that describe time-varying activities of the living cells.

In biomanufacturing, a vesicle carrying transgenes is inserted into the cell to manipulate the phenotype. As it dissolves, the size and color of inserted vesicle vary with respect to time (see Fig. 3b and 3c). The goal of this present study is to capture the dynamic vesicle under different scenarios. In our previous study, a sequence of cellular images is collected from in-control processes (without inserted vesicles). Also, out-of-control imaging profiles with inserted vesicles were also collected. These images are used as the basis for computer experiments to evaluate and validate the proposed methodology.

The community detection result of in-control cellular images is shown in Fig. 4. Here, parameters of the entropy rate clustering algorithm are adjusted, and 15 communities are detected inside of the cell to reveal detailed patterns. It may be noted that when the inserted vesicle is significantly different from the background (i.e., both size and intensity shift are significant), it is likely to be isolated as a new community. As such, the vesicle can be detected by monitoring the change in the number of communities. However, more challenging

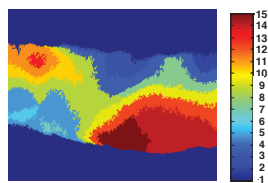


Fig. 4. Community detection result of in-control cellular images.

scenarios occur when the inserted vesicle is close to the background. As such, we need to extract community statistics and use the developed network GLR chart for the detection of inserted vesicles.

In this study, we implement the network GLR control chart to detect the inserted vesicle under two different scenarios: larger size and intensity shift (see Fig. 3b) and smaller size and intensity shift (see Fig. 3c). Notably, in both scenarios the vesicle is not significantly different from its background (i.e., the cell) to yield a new community.

##### 1) Detection of vesicles with larger size and intensity shift

We first evaluate our proposed network monitoring scheme under the scenario that inserted vesicles are with larger size and intensity shift (e.g., intensity shift = 15 and size = 8, as shown in Fig. 3b). In total, we have 70 in-control images and 30 out-

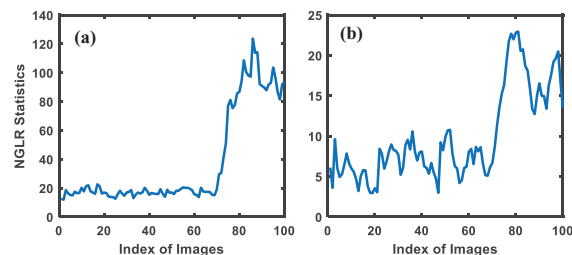


Fig. 5. NGLR chart for the detection of vesicles with intensity shift = 15 and size = 8, (a) keep all principal components; (b) keep first 10 principal components.

control ones. In other words, the change point  $\tau = 71$ . As shown in Fig. 5a, the NGLR statistic increases dramatically after the 70<sup>th</sup> image. That is, the control chart signals immediately when the vesicle is inserted. Here, all principal components in  $\mathbf{Z}$  are kept for the calculation of NGLR statistic. If we only keep the first 10 principal components (i.e.,  $d = 10$  in Eq. (8)), the NGLR chart is shown in Fig. 5b. It may be noted that the control chart is still effective in the detection of inserted vesicles.

##### 2) Detection of vesicles with smaller size and intensity shift

Furthermore, we conducted experiments to investigate the performance of proposed NGLR under the scenario that inserted vesicles are with smaller size and intensity shift (e.g., intensity shift = 4 and size = 4, as shown in Fig. 3c). In this case, it becomes difficult to distinguish the vesicle from the background by eyes. Fig. 6 shows the distribution of sample-to-sample mean intensity of the community that contains the inserted vesicle. The distribution of mean intensity for the first 70 in-control images (without vesicle) are shown in blue and the distribution of mean intensity for the 30 out-of-control images (with vesicle) are shown in orange. It may be noted that the two distributions are overlapped and there is no distinct

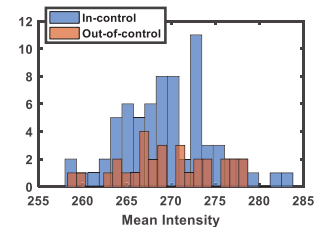


Fig. 6. Histogram of sample-to-sample mean intensity of the community that contains the inserted vesicle.

difference. Thus, it is difficult to detect the inserted vesicle in this case.

As shown in Fig. 7a, the developed NGLR chart still yields a large jump at the change point  $\tau = 71$ . This indicates that the developed NGLR chart is effective for the detection of very small changes in the biomanufacturing process. If we further reduce the dimensionality and only keep the first 10 principal components, the NGLR chart will encounter difficulty to detect the out-of-controls.

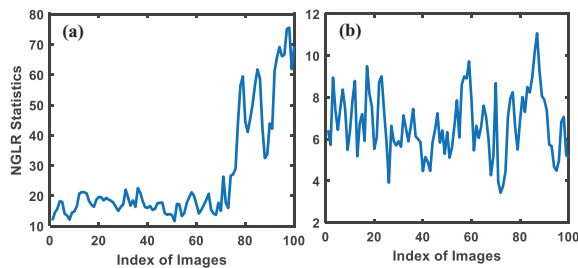


Fig. 7. NGLR chart for the detection of vesicles with intensity shift = 4 and size = 4, (a) keep all principal components; (b) keep first 10 principal components.

## 5. Discussion and Conclusions

The advancement of sensing technology brings the increasing availability of imaging data, which are critical to quality inspection and process improvement of biomanufacturing. However, traditional SPC methodologies focus on key characteristics of the product and are limited in the ability to model spatiotemporal patterns in imaging streams. Little work has been done to develop SPC methods for monitoring high-dimensional dynamic imaging profiles.

In this study, we developed a dynamic network monitoring scheme for the process control of biomanufacturing. First, each image is optimally represented as a weighted network. Then, an entropy rate clustering approach is implemented to characterize community structures in the network. Statistics are further extracted from detected communities to characterize spatiotemporal patterns in the image. Finally, a network generalized likelihood ratio chart is constructed for the change-point detection of the biomanufacturing process. The developed algorithm is implemented and validated in real-world applications of biomanufacturing. Experimental results have shown that the developed NGLR chart effectively characterizes variations in the complex structure of high-dimensional images and has high sensitivity for small process shifts. The proposed image-guided methodology has strong potential to be used for online monitoring of biomanufacturing processes and can be extended to 3D, 4D and even higher dimensional images.

## Acknowledgements

This work is supported in part by the National Science Foundation (CMMI-1646660, CMMI-1617148, CMMI-1619648, and IOS-1146882). The authors also thank Harold

and Inge Marcus Career Professorship (HY) for additional financial support.

## References

- [1] Zhang, J. The global biomanufacturing outsourcing market. *Biopharm International* 2015;39:28-9.
- [2] Mitsuishi M, Cao J, Bartolo P, et al. Biomanufacturing. *CIRP Annals - Manufacturing Technology* 2013;62: 585-606.
- [3] Chun K, Schmidt B, Köktürk B, et al. Catheter ablation - new developments in robotics. *Herz* 2008;33:586-9.
- [4] Arentoft M, Eriksen R, Paldan N. Towards the first generation micro bulk forming system. *CIRP Annals - Manufacturing Technology* 2011;60:335-8.
- [5] Stipniece L, Narkevica I, Sokolova M, Locs J, Ozolins J. Novel scaffolds based on hydroxyapatite/poly(vinyl alcohol) nanocomposite coated porous TiO<sub>2</sub> ceramics for bone tissue engineering. *Ceramics International* 2016; 42:1530-7.
- [6] Roh K, Nerem R, Roy K. Biomanufacturing of therapeutic cells: state of the art, current challenges, and future perspectives. *Annual Review of Chemical and Biomolecular Engineering* 2016;7:455-78.
- [7] Sugita N, Osa T, Aoki R, Mitsuishi M. A new cutting method for bone based on its crack propagation characteristics. *CIRP Annals - Manufacturing Technology* 2009;58:113-8.
- [8] Van Der Hofstadt M, Hüttener M, Juárez A, Gomila G. Nanoscale imaging of the growth and division of bacterial cells on planar substrates with the atomic force microscope. *Ultramicroscopy*. 2015;154:29-36.
- [9] Li F, Vijayasankaran N, Shen A, Kiss R, Amanullah A. Cell culture processes for monoclonal antibody production. *MAbs* 2010;2:466-77.
- [10] Ding Y, Zeng L, Zhou S. Phase I analysis for monitoring nonlinear profiles in manufacturing processes. *Journal of Quality Technology* 2006;38:199-216.
- [11] Zhou S, Sun B, Shi J. An SPC monitoring system for cycle-based waveform signals using Haar transform. *Automation Science and Engineering, IEEE Transactions* 2006;3:60-72.
- [12] Zhang L, Wang K, Chen N. Monitoring wafers' geometric quality using an additive Gaussian process model. *IEEE Transactions* 2015;47:1-18.
- [13] Wilcox P, Horton T, Youn E, et al. Evolutionary refinement approaches for band selection of hyperspectral images with applications to automatic monitoring of animal feed quality. *Intelligent Data Analysis* 2014;18:25-42.
- [14] Liu M, Tuzel O, Ramalingam S, Chellappa R. Entropy-rate clustering: cluster analysis via maximizing a submodular function subject to a matroid constraint. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2014;36:99-112.
- [15] Leskovec J, Krause A, Guestrin C, Faloutsos C, VanBriesen J, Glance N. Cost-effective outbreak detection in networks. *Proc. 13th ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining* 2007;420-9.
- [16] Kan C, Yang H. Dynamic network monitoring and control of in-situ image profiles from ultraprecision machining

and biomanufacturing processes. Quality and Reliability International. Accepted.

- [17] Martin D, Fowlkes, C, Tal D, Malik J. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring

ecological statistics. Proceedings of the 8th International Conference on Computer Vision, Vancouver, BC, July, 2001;416-23.