Sensor data fusion

A Novel Pattern-Frequency Tree for Multisensor Signal Fusion and Transition Analysis of Nonlinear Dynamics

Cheng-Bang Chen, Hui Yang, and Soundar Kumara

Harold and Inge Marcus Department of Industrial and Manufacturing Engineering, Pennsylvania State University, University Park, PA 16802 USA

Manuscript received September 14, 2018; revised November 8, 2018; accepted November 25, 2018. Date of publication December 3, 2018; date of current version January 16, 2019.

Abstract—Many real-word systems exhibit nonlinear and nonstationary dynamics, which defy understanding based on the traditional reductionist's approach. However, traditional analytical methods designed to effectively handle nonlinear dynamics are not well integrated with multisensor data fusion for process monitoring and control objectives. Realizing full potentials of multiple sensor signals calls upon the development of new methods for anomaly detection and transition analysis of nonlinear dynamics in complex systems. This article presents a novel pattern-frequency tree (PFT) approach for multisensor signal fusion and dynamic transition analysis. We leverage both pattern and frequency information in the PFT model to develop efficient algorithms for modeling and analysis of abnormal transitions in the nonlinear state space. Experimental results demonstrate that the proposed PFT method achieves a superior performance for multisensor data fusion and anomaly detection in nonlinear dynamical systems.

Index Terms—Anomaly detection, nonlinear dynamics, pattern-frequency tree (PFT), state space, transition analysis.

I. INTRODUCTION

Real-world complex systems often exhibit nonlinear and dynamic behaviors, which pose significant challenges on process monitoring and anomaly detection. Nonlinear dynamic behaviors manifest in a variety of domains such as manufacturing (e.g., machining signals), as well as health care (e.g., cardiac signals and patient monitoring). The rise of complexity due to nonlinear dynamics is caused by interactions between system elements such as cooperation, collaboration, competition, and interference. To tackle this complexity, advanced sensing systems are increasingly employed to improve information visibility. It is not uncommon that multiple sensors are used to capture multifaceted dynamic behaviors, which leads to large amounts of data that are high dimensional, nonlinear, and nonstationary. Realizing the full potential of massive sensing data for process monitoring and dynamic transition analysis depends to a great extent on the development of new analytical methods to handle the nonlinear and nonstationary dynamics.

However, nonlinear dynamical systems defy understanding based on the traditional reductionist's approach, in which one attempts to understand a system's nonlinear behaviors by combining all constituent parts that have been analyzed separately. As such, traditional methods such as frequency decomposition and principal component analysis encounter significant difficulties in capturing nonlinear, nonstationary, and high-order variations. Poincaré's geometric thinking of nonlinear systems provides a new way for a dynamic transition analysis. One of the earliest efforts to capture higher order dynamics in quality control is reported by Kamarthi *et al.* [1]. Also, Bukkaptnam *et al.* developed local Markov models to predict system dynamics and future evolutions in the state space; they also proposed an adaptive wavelet method to represent nonlinear dynamic signals for feature extraction [2]–[4]. Yang *et al.* developed a multiscale method to characterize and quantify nonlinear recurrence dynamics in multiple wavelet scales [5]. Furthermore, Yang *et al.* [6], [7] developed a new heterogeneous recurrence approach for monitoring the nonlinear stochastic process. Chen *et al.* [8] proposed a recurrence network approach for the extraction of nonlinear patterns in spatial data. Nonetheless, the theory of nonlinear dynamics is primarily studied by physics and mathematics, and is not well integrated with multisensory data fusion for monitoring and control objectives. There is an urgent need to develop new nonlinear dynamic methods to realize the full potentials of multisensor signals for statistical process control.

In this article, we present a novel pattern-frequency tree (PFT) approach for dynamic transition analysis in nonlinear and nonstationary systems. This PFT method represents the signals from multiple sensors with a state space approach, which facilitates the geometric analysis of dynamic transition patterns. Using this representation, we develop the hyperoctree aggregate segmentation (HAS) approach to discretize the state space into subregions, each of which is assigned with an aggregated state index. When system dynamics go from one subregion to another, we characterize such transition patterns and then build a PFT model. Finally, we leverage both pattern and frequency information in the PFT model to develop efficient algorithms for dynamic transition analysis. Experimental results show that the proposed PFT model achieves better results regarding the detection accuracy and computational performance than conventional methods for statistical process control of nonlinear dynamical systems in multisensor settings.

II. RESEARCH METHODOLOGY

A. State Space Representation

As shown in Fig. 1, we first represent multidimensional sensor signals in the nonlinear state space that will facilitate the geometric analysis. If each sensor provides a univariate time series $X_i = \{x_i(t)\}$, where i denotes the sensor i. Multiple sensors will lead to multivariate time series, $\boldsymbol{X} = \{\boldsymbol{x}(t) \in \mathbb{R}^d\} = \{[x_1(t), \ldots, x_d(t)]'\}$. In the state space representation, if each X_i is defined as one state variable, then, a state vector at t becomes $\boldsymbol{x}(t) = [x_1(t), \ldots, x_d(t)]'$. Note that continuous nonlinear processes are represented as $\frac{\partial \mathbf{X}}{\partial t} = \mathcal{G}(\mathbf{X}, \boldsymbol{\eta}, \boldsymbol{\varepsilon})$ and $\boldsymbol{x}(0) = \boldsymbol{x}_0$, where $\boldsymbol{\eta}$ is a process parameter vector varying over time and ε is system noises, and $\mathcal{G}(\cdot) \in \mathbb{R}^d \to \mathbb{R}^d$ captures the nonlinear dynamic relationship. A dynamic process at time t is, then, a point of

Corresponding author: H. Yang (e-mail: huy25@psu.edu).
Associate Editor: C.-A. Chou.
Digital Object Identifier 10.1109/LSENS.2018.2884241

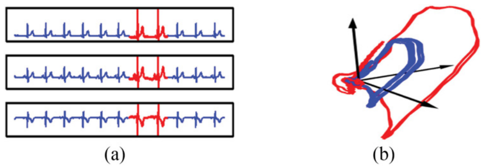


Fig. 1. Illustration of (a) multisensor signals and (b) nonlinear state space with dynamic transitions (i.e., finite-time detours in the trajectory).

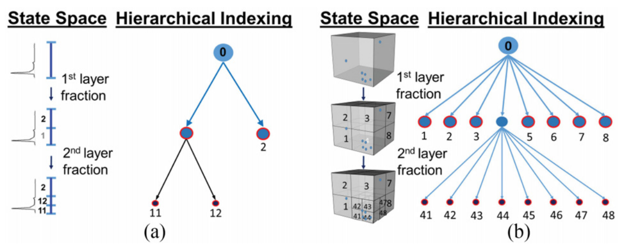


Fig. 2. Illustration of HAS tree-based scheme in (a) 1-D and (b) 3-D state space. Each time HAS seperates an over-capacity subregion into (a) 2^1, (b) 2^3 equal sized subregions, and labels with hierarchical indexes. E.g., in (a) region 1 is over-capacity; therefore, it is seperated into 11 and 12 subregions; in (b), region 4 is seperated into 41, 42, . . . ,48 subregions.

a geometric trajectory in the state space evolving from the initial state $\boldsymbol{x}(0)$. The evolution function $\boldsymbol{\eta}$ and the initial state $\boldsymbol{x}(0)$ delineate the dynamics of a nonlinear system. Therefore, complex system behaviors are represented as a nonlinear trajectory in the state space.

Based on the Poincaré's theorem, the detour of nonlinear trajectory in the state space indicates nonstationary transitions in the complex systems. The blue trajectory in Fig. 1 shows one type of system behaviors, which contains a specific pattern of recurrences. If a system does not contain any recurring pattern, the corresponding trajectory grows a random path without forming any particular route. Furthermore, the detour trajectories (red) in the state space, such as deviating from the original track (blue), imply that the anomalies occur in systems. Therefore, geometric properties of the nonlinear trajectory in the state space provide a new means to study and analyze dynamic transition behaviors of complex systems.

B. Hyperoctree Aggregate Segmentation (HAS)

To delineate local transition patterns, it is critical to segment the state space. A state space segmentation facilitates the analysis of states in collective sets and reduces the computational burden due to the large number of states. Note that it is desirable to partition the state space efficiently and make each subspace contain similar patterns or stationary characteristics.

We propose an HAS scheme, inspired by the quadtree [9], to efficiently divide the state space into local and heteroegenous subregions. The HAS scheme tackles distribution dependent issues of traditional methods such as equal boxing (Eqbox) and symbolic aggregation approXimation (SAX) [10], [11]. Note that the Eqbox assumes the observations are uniformly distributed; the SAX assumes that the observations follow a normal distribution. Although they are efficient methods, most state distributions in complex systems are neither uniform nor normal. The HAS is a tree-based space segmentation method that overcomes the distribution dependent issue. It takes each subregion as a tree branch and then checks and partitions every branch layer-by-layer until every subregion is within the given capacity. Note that the HAS partitions each over capacity region into 2^d equal-sized subregions in the d-dimensional state space. After partitioning the state space, each subregion is indexed with a unique label. This method is generally applicable for different types of state distributions. Fig. 2(a) and (b) illustrates the HAS in 1-D and 3-D cases, the HAS recursively partitions each subregion into 2^1 and 2^3 subregions until all subregions are lower than a given capacity. Denoting the state distribution as $\boldsymbol{X}$ and the capacity as $\mathcal{u}$, the final subregion set $\boldsymbol{\mathcal{B}} = \{\boldsymbol{b}_j\}$ of HAS is a function of $\boldsymbol{X}$ and $\mathcal{u}$ represented as $\mathcal{HAS}(\boldsymbol{X}, \mathcal{u}) = \{\boldsymbol{b}_j\}$, where $\cup_j \boldsymbol{b}_j = \mathbb{R}^d$, $\boldsymbol{b}_i \cap_{i \neq j} \boldsymbol{b}_j = \varnothing$, and $|\boldsymbol{b}_j| \leq \mathcal{u}\ \forall j$.

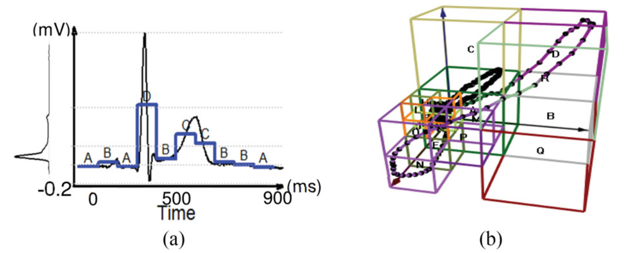


Fig. 3. Illustration of PAA for (a) 1-D signals and (b) 3-D state space.

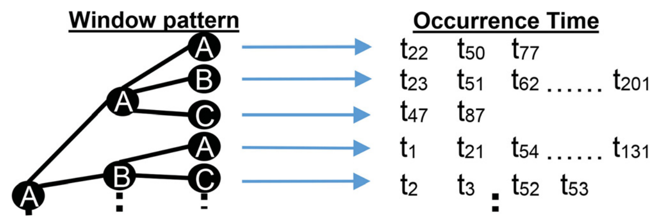


Fig. 4. PTF in the tree form.

C. Piecewise Aggregate Approximation (PAA)

To further improve the computational efficiency, we propose an augmented PAA extending from [10], for each subregion in the state space. Let $\boldsymbol{S}_i$ be a subset of $\boldsymbol{X}$ in the w-length sliding window, where the index i represents the window that starts from the time i, $\boldsymbol{S}_i = \{\vec{x}(i), \vec{x}(i+1), \ldots, \vec{x}(i+w-1)\}$, where $i \in \{1, 2, \ldots, T-w+1\}$. To derive the PAA approximation of $\boldsymbol{S}_i$, we first segment the $\boldsymbol{S}_i$ into K equal-length multivariate subsequence ($K \ll w$) and then use the aggregated approximation, such as the average, to represent each subsequence and finally form an PAA approximation as $\bar{\boldsymbol{S}}_i = \{\bar{\vec{S}}_i(1), \ldots, \bar{\vec{S}}_i(K)\}$, where $\bar{\vec{S}}_i(k) = \frac{K}{w}(\sum_{t=\frac{w}{K}(k-1)+i}^{\frac{w}{K}k-1+i} \boldsymbol{x}(t))$, and $k \in \{1, 2, \ldots, K\}$.

As such, we can represent $\boldsymbol{S}_i$ in the ith sliding window as a "word," P_i, which is a symbolic string representing the PAA approximation $\bar{\boldsymbol{S}}_i$. Here, $P_i = \{p_{i1}, \ldots, p_{i_K}\}$ is a K length symbolic string. Let α_j denote the jth element of an alphabet. Then, the PAA maps $\bar{\boldsymbol{S}}_i$ to a word P_i as $p_{i_k} = \alpha_j$ iff $\bar{\vec{s}}_i(k) \in \boldsymbol{b}_j$. The HAS partitions the state space into subregions and then assign a unique label to all states in a subregion. As such, the PAA approximation $\bar{\boldsymbol{S}}_i$ is transformed into a symbolic string, which retains geometric features of system behaviors. Fig. 3 shows the PAA approximation and labeling for electrocardiogram (ECG) signals in one and three dimensions, respectively. In short, PAA converts $\boldsymbol{S}_i$, a subset of $\boldsymbol{X}$ in the w-length sliding window, into an approximated subsequence $\bar{\boldsymbol{S}}_i$ and then map to a word P_i, $P_i = \mathrm{PAA}(\bar{\boldsymbol{S}}_i, \boldsymbol{\mathcal{B}})$, with the label set $\boldsymbol{\mathcal{B}}$ generated through HAS.

D. Pattern-Frequency Tree (PFT)

After the state space trajectory from multisensor signals is converted to a symbolic string, we propose a tree structure to represent transition patterns in each sliding window and their occurrence frequencies. Because we represent $\boldsymbol{S}_i$ in each window as a word P_i, each word pattern is used as a tree branch. As shown in Fig. 4, we can, then, obtain a PFT that keeps records of not only the transition patterns but their occurrence time points and frequencies as well. Table 1

Table 1. PFT in the Table Form.

Time Index	Window Pattern	Frequency
t_1	ABA	13
t_2	ABC	4
t_3	ABC	4
...	...	...
t_{m-n+1}	CAB	21

illustrates the PFT in the table form, which provides the time indices and frequencies of word patterns in each sliding window.

Now, the dissimilarity level of $\boldsymbol{S}_i$ and $\boldsymbol{S}_j$ in two different windows can be represented as the distance between these two subsequences $\text{Dist}(\boldsymbol{S}_i, \boldsymbol{S}_j) = \sum_{t=0}^{w-1} \|\vec{s}(i+t) - \vec{s}(j+t)\|_2$, where $\|\cdot\|_2$ denotes the $L2$ norm. The word patterns and frequencies in the tree provide the first level of information to efficiently identify the transition patterns that occur frequently or rarely happen. In the next step, we propose a minimum deviation score ($\mathcal{MDS}$) to provide the level of irregularity for $\boldsymbol{S}_i$ in each window i, which is the minimum distance of all nonself-match in the state space trajectory, i.e.,

$$\text{MSD}(\boldsymbol{S}_i) = \min_{j \in \mathcal{T} \setminus |i-j|<w} \{\text{Dist}(\boldsymbol{S}_i, \boldsymbol{S}_j)\}.$$

The nonself-match is critical to avoid the $\mathcal{MDS}$ bias, because the self-match will give a zero distance and thus makes all minimum distances zero. As a result, given the state space $\boldsymbol{X}$, the discord is $\boldsymbol{S}_i$ with the biggest $\mathcal{MDS}$ score is $\max\{\mathcal{MDS}(\boldsymbol{S}_i)\}$. Thus, the mth largest discord is the subsequence $\boldsymbol{S}_i$ having the mth largest $\mathcal{MDS}$.

E. PFT-MDS Algorithm

One way to find the m largest $\mathcal{MDS}$ is the exhaustive search that compares all windows one-by-one through a double loop. Although an exhaustive algorithm can find the $\mathcal{MDS}$ for all windows, it is computationally inefficient and expensive. Instead, we propose to leverage the pattern and frequency information in PFT to improve the algorithmic efficiency. As shown in Algorithm 1, the PFT data structure helps speed up the processes without the need to go through the exhaustive search. The idea is to leverage the time locations frequency information of each transition pattern to reduce the number of $\mathcal{MDS}$ computations. For example, transition patterns that frequently occur in the trajectory has a smaller probability to be an anomaly, but those happen rarely will have a higher probability. As such, we modify two loops to speed up the PFT-MDS algorithm as follows.

1) *The Order of the Outer Loop:* Because the PFT records the occurrence time and frequency of each transition pattern, we compute the $\mathcal{MDS}$ with a sorted list of frequency of transition patterns. Note that if the frequency of a transition pattern is small, it will have a higher probability to yield a bigger $\mathcal{MDS}$. If a transition pattern occurs very frequently, then it is unlikely to be an anomaly. Therefore, we set the order of outer loop with an ascending order of frequencies of transition patterns. The PFT provides an effective and efficient means to quickly obtain the m largest $\mathcal{MDS}$.
2) *The Order of the Inner Loop:* Because the algorithm only keeps $\boldsymbol{S}_i$'s with the m largest $\mathcal{MDS}$'s, we use the minimum of these m largest $\mathcal{MDS}s$ as a criterion to eliminate unnecessary calculations. If the computed $\mathcal{MDS}(\boldsymbol{S}_p)$ is less than the smallest value of the m largest $\mathcal{MDS}$ so far, then $\boldsymbol{S}_p$ will not be among m candidates of anomalous transitions. Hence, the top m candidates for anomaly transitions remain the same. Note that time locations for the same word (or transition patterns) are recorded in the PFT structure. Therefore, the $\mathcal{MDS}$ computation can go through these locations for the transition pattern $\boldsymbol{S}_p$ and identify whether $\mathcal{MDS}(\boldsymbol{S}_p)$ is bigger than the smallest value of m largest $\mathcal{MDS}'s$. If not, the algorithm breaks and goes to the next loop. This will greatly help to increase the speed of search of m largest MDS values.

Algorithm 1: The Pseudocode Of The Pft-MDS Algorithm.

Procedure PFT-MDS COMPARISON ALGORITHM
Input:
$\{S$: a series of states, w: length of a subsequence,
m: number of best m results$\}$
Variable:
$mMDS$:= a m length vector to store the m largest $\mathcal{MDS}(\cdot)$
$\mathcal{L}_O$:= a series of ordered states for outer loop
$\mathcal{L}_O(i)$: the ith component of $\mathcal{L}_O$
$\mathcal{L}_I$:= a series of ordered states for Inner loop
$\mathcal{L}_I(j)$: the jth component of $\mathcal{L}_I$
$\boldsymbol{S}_t$:= the inspection window starting at time stamp t
Initialization:
Set $\min\{mMDS\} = \infty$
$\mathcal{L}_O \leftarrow$ **GetOuterLoopOrder**(S, w)
For $\boldsymbol{i} = \mathbf{1}$ to $|\mathcal{L}_O|\{$
  $\mathcal{L}_I \leftarrow$ **GetInnerLoop**$(S, w, \mathcal{L}_O(i))$
  For $j = 1$ **to** $|\mathcal{L}_I|$ $\{$
    If $|\mathcal{L}_0(i) - \mathcal{L}_I(j)| \geq \boldsymbol{w}$ **then** $\{$
      calculate dissimilarity level:
      $Dist(S_{\mathcal{L}0(i)}, S_{\mathcal{L}_I(j)})$
      If Dist $(S_{\mathcal{L}_O(i)}, S_{\mathcal{L}_I(j)}) < \min\{mMDS\}$
      then $\{$
        Break
$\}\}\}\}$
keep the m largest $\mathcal{MDS}(\cdot)$ as the final return
End Procedure.

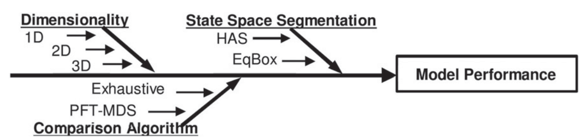


Fig. 5. Cause-and-effect diagram of experimental design.

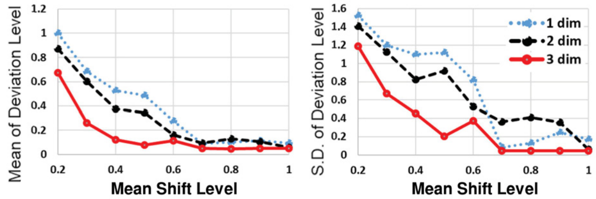
III. EXPERIMENTAL DESIGN AND RESULTS

As shown in Fig. 5, we evaluate and validate the proposed methodology with a three-factor experimental design, including the dimensionality factor (i.e., the number of sensors), spatial segmentation methods, and benchmark algorithms (i.e., exhaustive searching versus PFT-MDS algorithms).

In the present experiments, we use the 3-lead ECG signals from a human subject. The sampling frequency is 1 kHz, and the signal length is 6 s, i.e., 6000 data points. The dynamic transition is simulated by adding the finite time detour of 300 data points in the state space trajectory. Notice that we set the length of one sliding window is 400 data points, i.e., the cardinality of $|\boldsymbol{S}_t| = 400$. The performance metric of detection accuracy is the absolute difference between the time location of actual anomaly and predicted anomaly. The smaller the absolute difference is, the better the detection accuracy will be.

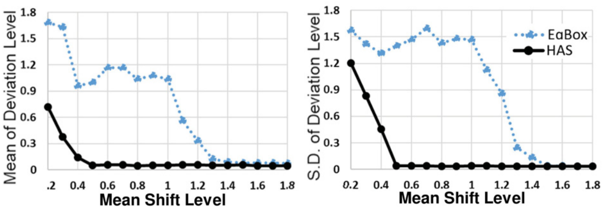
A. Dimensionality Effect

Because there are often multiple sensors involved in the system monitoring, we first evaluate the variations of detection accuracy for each scenario by varying the dimensionality of the state space. At each experimental scenario, we randomly add an anomaly to multidimensional signals, where the anomaly is a mean shift to the original signals. Each dimension has the same relative level of mean shift. This experiment is designed to test the hypothesis whether the dimensionality of state space impacts the detection performance at different levels of anomaly conditions.



(a) (b)

Fig. 6. Comparison of the mean (a) and the standard deviation (b) of detection accuracy concerning the dimensionality of state space.



(a) (b)

Fig. 7. Detection accuracy under using different spatial index methods (a) mean of deviation level and (b) standard deviation (S.D.) of deviation level.

We run each experimental scenario with 50 replications to obtain the mean and standard deviation of performance metrics. As shown in Fig. 6(a), the average of detection errors converges to zero when the mean shift level increases, but if the dimensionality of state space is higher, then the detection error is getting smaller at the same mean shift level. Fig. 6(b) shows that the standard deviations of detection errors also converge to zero when the mean shift level increases. However, the standard deviation is smaller if the dimensionality of state space is higher. In other words, the algorithm is more stable if we involve more sensors and use the multidimensional state space. In summary, Fig. 6 shows that the detection accuracy is impacted by the number of sensors used and the dimensionality of state space.

B. State Space Segmentation and Indexing

Also, there are different ways to segment the state space, including the EqBox and HAS methods. This experiment is designed to test the hypothesis whether the HAS method is better than the conventional EqBox method at different levels of anomaly conditions. Note that the HAS method equalizes the capacity of each subregion, while the EqBox method balances the size of each box, which refers to a subregion. In this study, we set the HAS capacity to be the same as the average capacity of all boxes in an EqBox approach. Again, we run each experimental scenario with 50 replications to obtain the mean and standard deviation of performance metrics. As shown in Fig. 7, HAS is better than EqBox for all examined scenarios regarding both the average and standard deviation of detection errors. The HAS for state space segmentation makes the algorithm yield better detection power with a smooth and robust converging trend.

C. Computational Efficiency

The PFT provides an effective data structure for both transition patterns and frequency information, which greatly helps to improve the efficiency of detection algorithms. For the benchmarking purpose, we compare the computational time between the fast PFT-MDS algorithm and the conventional exhaustive method at each experimental scenario. As shown in Fig. 8, the PFT-MDS algorithm yields at least a 6-fold decrease in the computational time. Also, as the dimensionality increases, more time is needed for the computation. Nonetheless, it

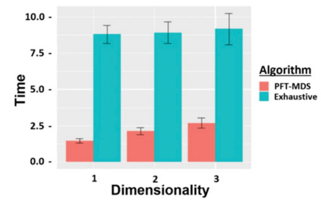


Fig. 8. Comparison of computational time between the PFT-MDS algorithm and the exhaustive comparison method.

is evident that the PFT-MDS algorithm is more efficient than the conventional exhaustive method.

IV. CONCLUSION

This letter presents a novel PFT method for a dynamic transition analysis in nonlinear and nonstationary systems. As opposed to the traditional reductionist approaches such as frequency analysis or principal component analysis, we propose a state space representation of signals from multiple sensors that leverages the geometric analysis of nonlinear dynamics and then use the HAS approach to discretize the state space into subregions, each of which is assigned with an aggregated state index. Whenever system dynamics transit from one subregion to another, we build a PFT model to characterize both transition patterns and frequency information for dynamic transition analysis. Experimental results show that the proposed PFT method achieves better results regarding the detection accuracy and computational efficiency than conventional methods for multisensor data fusion and anomaly detection in nonlinear dynamical systems.

ACKNOWLEDGEMENT

This work was supported by the National Science Foundation under Grant CMMI-1617148.

REFERENCES

[1] S. V. Kamarthi, S. R. T. Kumara, and P. H. Cohen, "Wavelet representation of acoustic emission in turning process," *Intell. Eng. Syst. Through Artif. Neural Netw.*, vol. 5, pp. 861–866, 1995.

[2] S. T. S. Bukkapatnam, H. Yang, and F. Madhavi, "Towards prediction of nonlinear and nonstationary evolution of customer preferences using local markov models," in *Product Research: The Art and Science Behind Successful Product Launches*. New York, NY, USA: Springer, 2009, pp. 271–287.

[3] S. T. S. Bukkapatnam, S. R. T. Kumara, A. Lakhtakia, and P. Srinivasan, "The neighborhood method and its coupling with the wavelet method for signal separation of chaotic signals," *Signal Process.*, vol. 82, no. 10, pp. 1351–1374, 2002.

[4] H. Yang, S. T. S. Bukkapatnam, and R. Komanduri, "Nonlinear adaptive wavelet analysis of electrocardiogram signals," *Phys. Rev. E, Stat. Nonlin. Soft Matter Phys.*, vol. 76, no. 2, 2007, Art. no. 026214.

[5] H. Yang, "Multiscale recurrence quantification analysis of spatial cardiac vectorcardiogram signals," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 2, pp. 339–347, Feb. 2011.

[6] Y. Chen and H. Yang, "Heterogeneous recurrence representation and quantification of dynamic transitions in continuous nonlinear processes," *Eur. Phys. J. B*, vol. 89, no. 6, 2016, Art. no. 155.

[7] C. Kan, K. P. Yip, and H. Yang, "Two-phase greedy pursuit algorithm for automatic detection and characterization of transient calcium signaling," *IEEE J. Biomed. Health Inform.*, vol. 19, no. 2, pp. 687–697, Mar. 2015.

[8] C.-B. Chen, H. Yang, and S. Kumara, "Recurrence network modeling and analysis of spatial data," *Chaos, Interdiscip. J. Nonlin. Sci.*, vol. 28, no. 8, Aug. 2018, Art. no. 085714.

[9] A. Ayyagari, S. Kumara, and I. Ham, "Robot path planning in 2-d, using modified quad-tree approach," in *Recent Developments in Production Research*, A. Mital, Ed. Amsterdam, the Netherland: Elsevier, 1988, pp. 707–713.

[10] J. Lin, E. Keogh, L. Wei, and S. Lonardi, "Experiencing SAX: A novel symbolic representation of time series," *Data Min. Knowl. Discovery*, vol. 15, no. 2, pp. 107–144, 2007.

[11] E. Keogh and J. Lin, "Clustering of time-series subsequences is meaningless: Implications for previous and future research," *Knowl. Inf. Syst.*, vol. 8, no. 2, pp. 154–177, 2005.