IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 16 November 2022, accepted 3 January 2023, date of publication 11 January 2023, date of current version 17 January 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3236000

RESEARCH ARTICLE

Mimicking Gait Dynamics: A Step Toward Precision Learning of Human Activities

XI YANG AND FUSHING HSIEH
Department of Statistics, University of California at Davis, Davis, CA 95616, USA
Corresponding author: Fushing Hsieh (fhsieh@ucdavis.edu)

ABSTRACT We develop a computational protocol for mimicking personal gait dynamics with 12-dimensional time series derived from 4 accelerometer sensors found in the MAREA database and then explore its utilities in line with precision learning of human activities. The foundation of mimicking high dimensional rhythmic dynamics is explicitly established upon deterministic and stochastic structures found on structural representations of evolving biomechanical states hidden within all computed gait cycles. Such a technique enables practitioners to detect and confirm minute structural changes that could last for only a few cycles with high precision. Our computational developments are step-by-step illustrated via one subject's data, while the other 8 subjects' data are also analyzed and compared accordingly. A common cyclic composition of evolving biomechanical states of various temporal scales emerges from the 9 subjects' comparisons. We conclude that mimicking an individual's gait dynamics offers precise detections of potential multiscale minute differences against gait dynamics of different time periods or of different persons, and further offers clues of efficiency on personal walking activity. This mimicking-based capability is a cornerstone for the proof of concept: dynamics mimicking enables precision learning by improving the efficiency of learning and performing human activities in competitive sports, social dancing, and physical rehabilitation, among many others.

INDEX TERMS Biomechanics, complex system, cyclic structural representation, musculoskeletal system, principal component analysis (PCA).

I. INTRODUCTION

The low-cost, lightweight, easy-to-use inertial measurement units (IMU), such as accelerometer and gyroscope sensors, can nowadays be found in many wearable devices, such as smartphones, smartwatches, and fitness trackers. Such sensors' data-collecting efficacy and precision have also drastically evolved and improved with recent technological advances in microelectromechanical systems (MEMS) [1]. These wearable sensors indeed liberate the domain of gait data-collecting regimes from a mechanics lab to wherever a subject wants to be. They also shift the focus of Gait Analysis [3] from medical care to all kinds of human activities, including all sports [2]. Nowadays gait data are collected in both indoor and outdoor environments.

However, even up to today, gait dynamics related research primarily focuses on an individual's gait recognition [4], [5], [6] and authentication [7], [8]. Our previous work [10], which analyzed data from the MAREA database [9] and another database HuGaDB [11], also focused on recognitions and authentication.

To see beyond surveillance purposes, viewing gait dynamics through the perspective of human activity's precision learning has not been conceived widely yet. Nonetheless, this perspective would be naturally recognized in near future. Originally, precision learning is a pedagogical terminology that represents a type of programmed instruction for under-represented learners. The teachers will adjust the instructional tactics and curricula by monitoring the learner-specific needs and the efficacy of the instruction with the ultimate goal of providing the best educational outcomes to the learner. Since nowadays gait data are popularly collected by healthy individuals who perform all sorts of activities. It will be time for healthy performers to try to improve their efficiency in performing activities they love and care about by applying personal precision learning protocols derived from their own and others' activity data.

Commercial wearable sensors typically have various sampling rates when collecting time series data. Some gait time series data are collected with sampling rates of more than

The associate editor coordinating the review of this manuscript and approving it for publication was Taous Meriem Laleg-Kirati.

100 Hz, that is, every data point is collected per less than 10 ms. For instance, the sampling rate in the MAREA database is 128 Hz [9]. It is critical to note that a gait pattern of 10 ms scale is invisible to human raw eyesight, as well as to video recording. That is, performers can not see that themselves, and neither could their coaches. This fact provides a major advantage in learning human activity, including walking, through analyzing data from wearable sensors.

To make the aforementioned advantages of wearable sensors real, it is necessary to conceptualize that a person's human activity consists of a steady, natural, but idiosyncratic series of evolving biomechanical states of varying temporal scales. Discovering such a personal collection of biomechanical states is a computing task that has not been well established in literature yet. In our previous study [10], we reported findings of alternating cyclic states. Such data-driven cluster-based states' variations are too big to be biomechanically coherent. That is, cyclic states resulting from any clustering algorithms need to be reorganized and recomposed to achieve the stability required by the nature of human biomechanics [13].

Biomechanical states are known as being not rigid and not varying too much. Since healthy human walks with coherent stability. This smooth and steady manner is significantly distinct from the robot's rigid and constant walking and from a drunken person's erratic walking. Here, the erratic walking of a drunken person is an example of the irregular and unrhythmic gait dynamics that can not be thoroughly analyzed and mimicked due to their instability and randomness. Therefore, computing coherent biomechanical states within stable cyclic series is one of the chief tasks in any study of gait dynamics. Before getting into computing, we first recognize that how human achieves such coherence and stability seemingly without effort is indeed an amazing mechanistic phenomenon.

This phenomenal mechanism involves multiple biomechanical states of large and fine scales within each cycle. Though each state varies, however, humans still achieve cycle stability when summing up varying durations incurred by these multiple biomechanical states in walking. This stable phenomenon is apparently achieved by some sort of automatic coordination. Such a coordinating mechanism implies complicated dependency among all involved states. This mechanism is supposed to be essential to the efficiency of a person's walking activity. To our knowledge, this coordinating mechanism of biomechanical states is still not well known in the literature. It is another chief task in our study of gait dynamics here.

To achieve the two aforementioned chief tasks in this paper, we formulate two chief computing tasks by developing: 1) a structural representation of gait cycles; and 2) a dynamics-mimicking protocol based on data from the MAREA database. We argue and demonstrate that these two computing tasks jointly serve as a step toward precision learning of human gait dynamics. In this paper, we do not analyze data from HuGaDB because of its low sampling rate.

FIGURE 1. Position and orientation of four accelerometer sensors used in MAREA database [9]. (Figure courtesy of Dr. Siddhartha Khandelwal and Prof. Nicholas Wickström.)

The design of collecting time series, surface conditions, and subject selections for the MAREA database were fully depicted in [9]. We only briefly mentioned sensor related information here. The position and orientation of the four accelerometer sensors attached to the Left-foot, Right-foot, Waist, and Wrist of each subject for the data collection of the MAREA database are shown in Fig. 1. Each accelerometer collected $x-$, $y-$, and $z-$ three mutually orthogonal directional time series.

We begin our outline of computational developments and lay out our contributions in this paper. We plan to first successfully capture the structural dependency of 6-dimensional (dim) time series pertaining to the Left-foot and Right-foot ([LF-vs-RF]) subsystem by applying Hierarchical Clustering (HC) algorithm on the whole set of data points without their temporal coordinate information. This [LF-vs-RF] subsystem is understood as being better at preserving gait rhythm embedded within the entire time series data. Here the HC algorithm makes good use of its capability in differentiating time points belonging to various cyclic components within each cycle. Each cyclic component realistically preserves a specific form of dependency among the 6 dimensions of the time series. Further, HC algorithm at the same time collects similar time points across different cycles. This fact was established in [10] via Lempel-Ziv complexity evaluations [12].

With resultant clusters being digital- and color-encoded, we then recover the temporal coordinates of all members of all encoded clusters. This simple procedure indeed transforms the observed 6-dim time series into a 1-dim color-code or digital-code time series that reveals recurrences of all digital (or color) codes with varying duration across the entire time series span. Some recurrent patterns are rather steady in their recurrences, some are not. We algorithmically discover subject-specific landmark that achieves the most regular recurrence. With this computed landmark in hand, we dissect the entire time series into stable cycles.

Our first major computing task is to develop a structural representation for all dissected gait cycles. Given that each cycle is represented by a 1-dim segment of transition of cluster-based states (without duration information), we stack these segments by aligning their landmarks, and then subsequently align with respect to the next discoverable common state shared by all cycles. With each row corresponding to a cycle, such an alignment likely generates empty cluster states between the landmark state and the first commonly shared state. After this alignment is manually carried out one by one, between the first commonly shared state and the second one, between the second commonly shared state and the third one, and so forth, the resultant array lattice is seen to embrace all cycles' state-transition representations. That is, this visible matrix format allows us to figure out the deterministic and stochastic structures in terms of states' evolution. Next, we recover each code entry's duration information and look for groups of adjacent columns that have more or less constant total temporal lengths. Each such column group is designated as a biomechanical state. This biomechanical state representation in matrix format is a structural representation of all cycles with respect to state-transitions as well as to state-duration.

Based on the above structural representation of personal gait dynamics, the second major computing task is to develop a mimicking protocol that primarily first identifies deterministic and stochastic structural patterns regarding the information on each biomechanical state's duration and the variations of all biomechanical states' collective behaviors, and secondly recreates these multiscale patterns accordingly. In other words, such a mimicking protocol is designed and constructed to explicitly preserve the evolving rhythm of biomechanical states and simultaneously conserve gait's dynamic multiscale deterministic and stochastic structures derived from both temporal and 6-dim numerical aspects. As such, our mimicking protocol produces mimicries of all observed cycles in one subject's entire gait dynamics under the [LF-vs-RF] subsystem. Such computational developments are likewise carried out by analyzing 3-dim time series from Waist and Wrist sensors, respectively. Indeed both structural representation and mimicking protocol computations with very mild modifications would be applicable to all rhythmic time series, such as personal heartbeat with multiple channels or vital signs with body temperature, heart rate, respiration rate, and blood pressure, among many others.

At the end of this section, we emphasize one of the chief merits of mimicking protocol: reliability and confidence evaluations. Since the human musculoskeletal system is a highly constrained physical system. Ranges of all biomechanical states are finite. Such a finiteness property renders the mimicking protocol to generate nearly 100% multi-dimensional manifolds for purposes of reliability and confidence evaluations because we can simulate as many mimicries as we wish. This chief merit in fact serves as the solid basis for precision learning of human activities. Specifically speaking of walking, from the perspective of Left-foot's or Right-foot's 3-dim acceleration time series data points in R^3 Euclidean space without temporal coordinates, for example, this chief merit of mimicking can be explicitly visualized through its capability of filling in spacing created within the two scatter plots. Such a filling-in function on open spaces renders that 3D scatter plots can be made into somehow smooth and solid manifolds by mimicking huge numbers of mimicries, while leaving out clear open spaces, including their outer space. Any open space defined by such manifolds is taken as foreign regions that do not belong to the subject's gait dynamics with nearly 100% confidence and reliability. By attuning to the issue of determining whether open spaces of the original manifold are occupied after adding new data, not only can we make subject-vs-subject comparisons in detail with nearly complete confidence, but we can also detect even minute structural shifts within a subject with almost certain reliability. This issue is fundamental to the precision learning of any human activity. By constructing resolutions of the above two computational tasks and revealing their merits, we project that the gait dynamics-mimicking protocol signals a step forward to precision learning of human activities. These are the contributions of this paper.

II. PRECISION LEARNING HUMAN PHYSICAL ACTIVITIES FROM COMPLEX SYSTEM PERSPECTIVE

Human physical activities manifested through sports, dances, and exercises have created many essential societal dynamics. For instance, the 36 professional soccer leagues consisting of 1018 clubs have generated social and economic activities that become essential parts of life across countries in Europe, so do 30 teams in Major League Baseball (MLB), 32 teams in National Football League (NFL), 30 teams in National Basketball Association (NBA) and hundreds of college sports teams across the USA, to name just a few. According to Statista, a German database company, the North American sports market had a value of about 71.06 billion U.S. dollars in 2018. This figure is expected to rise to 83.1 billion by 2023. As for dancing, the 778 ballet companies in the USA, not including contemporary dance companies, also offer a glimpse of professional dancing activities on top of many other styles of dances. As for exercises, including walking, their scales of societal dynamics and their impacts are even harder to be quantified because it involves almost everyone, including younger people and the elderly. However, the learning processes of these human activities for general people are still done via old fashion way: self-learning by practicing. Even professional athletes or dancers, who have private coaches, still more or less rely on this old fashion learning protocol. Should and could all people practice their activities effectively and scientifically via their own personal data-driven protocol?

Since practitioners and learners of physical activities spread across all ages as well as skill levels within each activity category. The young learners practice to improve and become skilled. The skilled ones practice to improve and become professional. The professional ones practice to improve and become the best ones. The essence underlying all these levels of practicing is the search for a personal learning path. It is because everyone learns differently.

Therefore, from a learner's perspective, the central issue is: How to develop a precise personal path of learning a physical activity? While from the teacher and coach's perspective, the ultimate issue is: How to teach or coach every different individual differently? These two issues from distinct perspectives indeed converge to an intuitive concept of precision learning for physical activities. Previously, this term of precision learning has been specialized in education for children with special needs, such as autistic children.

To our limited knowledge, the domain of precision learning of human physical activities has not been seriously put into practice in real-world settings yet, either in academia or in the business world. A rigorous proof of concept is still waiting to be carried out: **Each individual needs a mimicking based person-specific precision learning to improve this individual's best efficiency in learning any physical activity**. The best efficiency here is referred to the concept: **The best athlete is the one who executes his or her skill the best**. After performing precision learning of a specific activity of any individual as well as the best athlete, we can tune and improve the activity performance of the individual by decreasing the disparity between the individual's and the best athlete's gait dynamics.

In this paper, we develop a computational protocol for mimicking human gait dynamics to serve as a basis for precision learning of walking. The reason for studying walking is obvious because of its fundamental role in all sports, dances, and exercises. The majority of humans indeed constantly generate and simultaneously make good use of their gait dynamics in every minute and second of their awake parts of daily life. However, healthy people hardly think about their own gait dynamics as if walking is performed by everyone's musculoskeletal system in an entirely automatic fashion. Further, since biomechanics, which refers to the study of the mechanical principles of living organisms, particularly their movement and structure, underlying gait dynamics seemingly has been well studied and known [13], and as if there is very little scientific interest left to be looked for.

However, the above viewpoint is a misconception. Gait dynamics is still somehow mysterious in our open eyes. It is still a mystery partly because we are indeed not able to visualize its collective whole of biomechanical operations. It is in part due to human raw eyesight being not capable of seeing the entire scale-spectrum of dynamic patterns that constitute trajectories of the human musculoskeletal system. Our eyes can't clearly detect patterns of 10 ms or finer scales. So, gait dynamics is basically hidden in plain sight, especially in sports biomechanics [14].

Further, a person-specific gait dynamics is an idiosyncratic complex system that involves multiple scales of spatial-oriented forces and angular-momentum of all directions [15], [16]. Such a multiscale nature indeed not only makes individuals' gait dynamics idiosyncratically distinct but also foretells that better knowledge of a person-specific gait dynamics, when this person is healthy, should be an invaluable personal asset. Since its beneficial values would suddenly become evident at the moment when this person is injured. It is the case because this person's idiosyncratic gait is likely forever altered at the moment in time. Apparently, this concept of preserving healthy persons' gait dynamics is not yet well perceived in comparison with personal DNA sequences in this Big Data era, see The Economist special issue (Feb. 27th, 2010) with the title "The Data Deluge."

Data pertaining to a person's gait dynamics is typically collected in a biomechanic laboratory within a biomechanic department, which is often an institute or department of a major medical center or school. In the lab, a person is placed in a room equipped with 8-12 cameras and sophisticated computer systems. As this person walks on a specially marked or constructed strip of floor, data are streamed through cameras into computers and then analyzed with a package that captures a fixed set of measurements. This data-collecting approach is constrained and limited from the perspective of performing human activities. Just like healthy persons hardly go to a hospital, hardly any healthy person's gait data are taken. So hardly data-driven patterns of a healthy person's idiosyncratic characteristics are extracted.

However, via recently developed wearable sensors, person-specific gait dynamics derived from the human's musculoskeletal system are going to be opened up for all persons' all activities under all out-of-lab environmental settings. Being free of constraints pertaining to biomechanics labs, healthy people can seriously and actively collect their own gait dynamic time series data when performing all sorts of activities.

From this perspective of data collection, the practical focus and scientific interests in gait dynamics are placed on discovering person-specific compositions of multiscale constituent mechanisms, ranging from the rhythmic cycle, all biomechanical states within a cycle, and all temporal patterns within each state, etc. These multiscale mechanisms are fundamental constructs of this person's gait dynamics as a complex system [17]. Thus, when putting such multiscale pattern-based compositions against the temporal axis, person-specific gait dynamics would become visible and explainable. To this goal, we explicitly construct such a display, called the structural representation of gait dynamics, in this paper. Further, by revealing such structural representations, healthy and active persons are able to recognize their own evolving gait dynamics from the beginning to the end of the activity with precise details. Such a structural representation is one major contribution of this paper. Since it resolves questions of great interest: How to characterize and compute biomechanical states? What are the precise characteristics of the

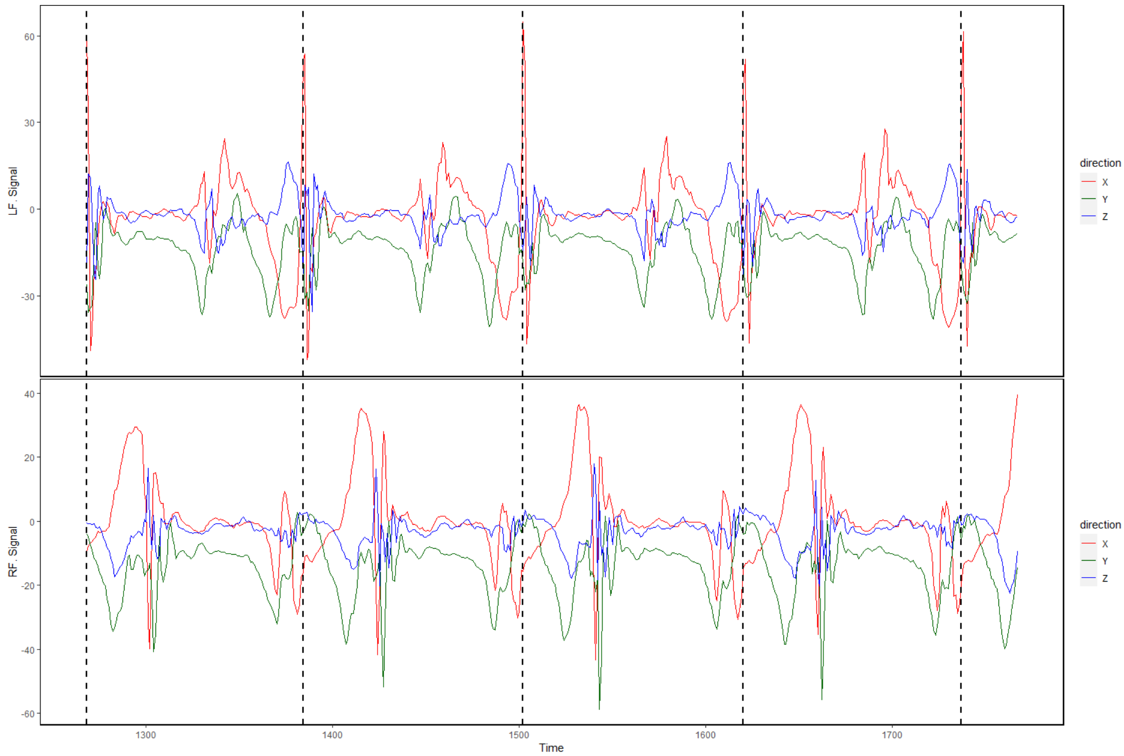
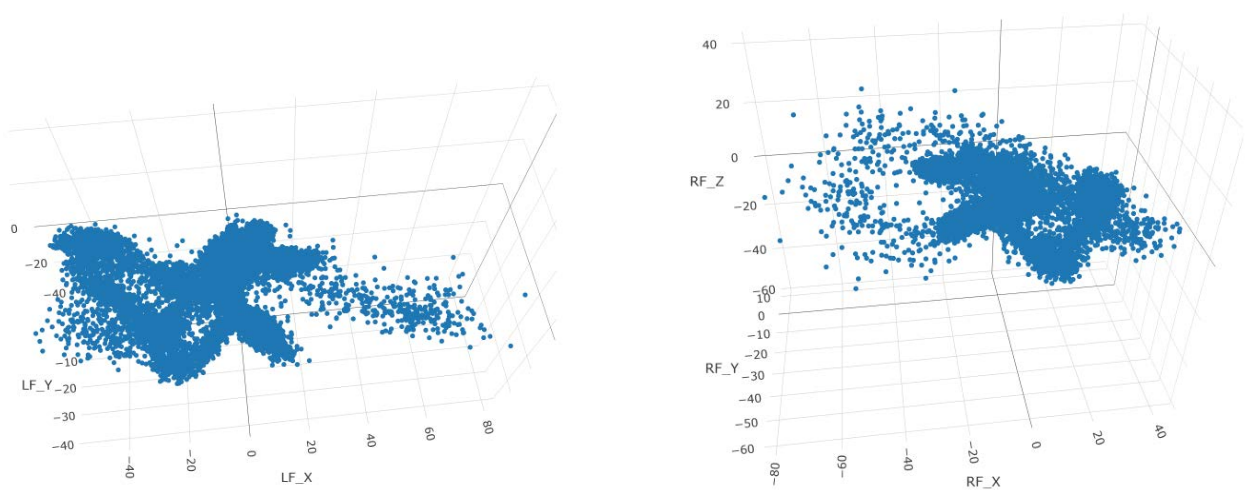


FIGURE 2. Time series representation with unmarked four cycles.



(a) (b)

FIGURE 3. Scatter plots of 3D manifolds based on 3-dim time series observed from the subject-12's: (a) Left-foot's and (b) Right-foot's accelerometer sensors. See rotatable 3D plots through the two links: (a) https://statistics2022.github.io/3Dplots/Figure2-PanelA; (b) https://statistics2022.github.io/3Dplots/Figure2-PanelB.

serial biomechanical states within a cycle and across a series of cycles?

Another major contribution of this paper is the computational protocol of mimicking a person's gait dynamics. Such a mimicking protocol can further address related fundamental and technical questions of great interest: How to prescribe stability and volatility within a person's gait dynamics? How to detect and confirm even minute structural changes? We explicitly illustrate our mimicking protocol in resolving these questions.

It is essential to note that mimicking based on structural representation indeed becomes a precision learning paradigm. For healthy and active individuals, these questions are performance oriented. They are far from medical ones.

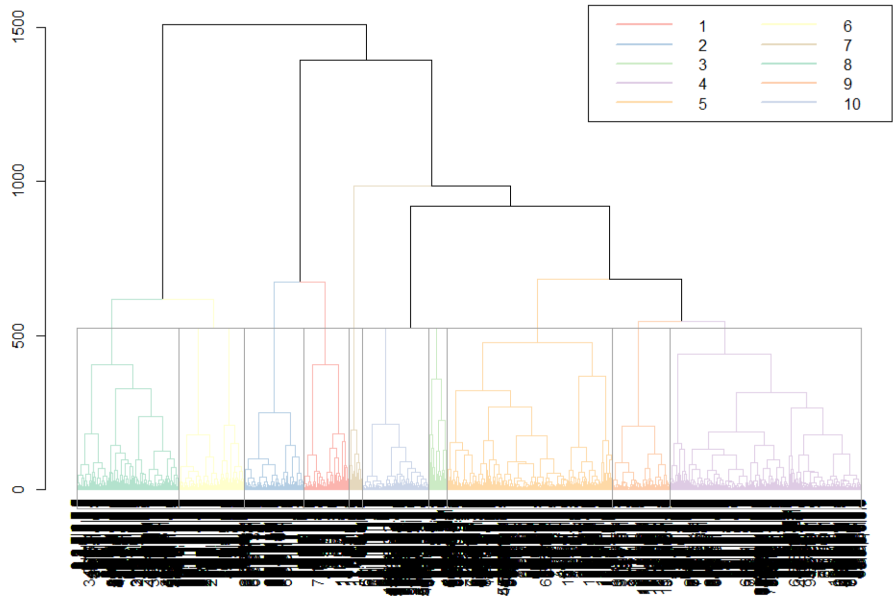
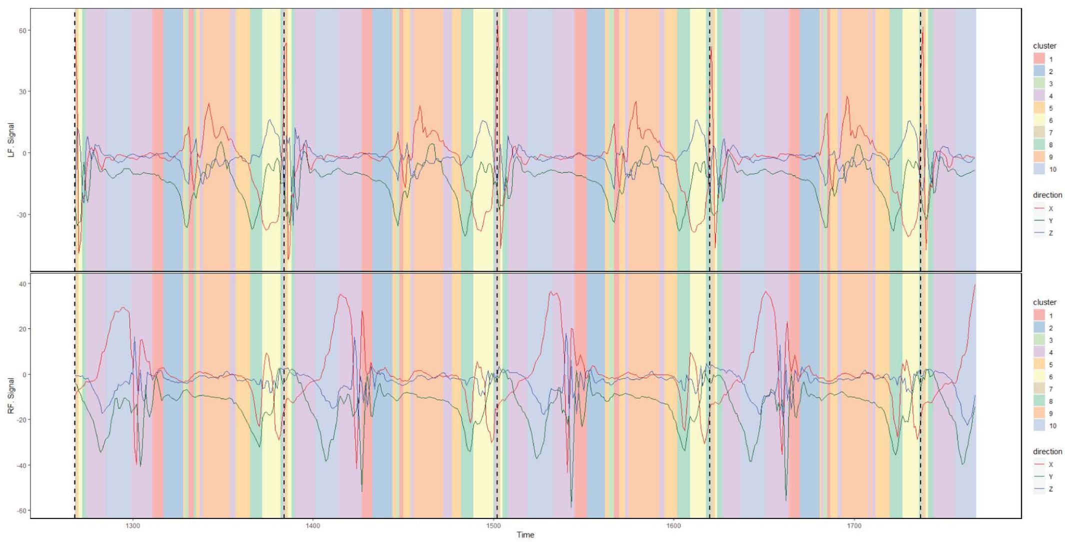
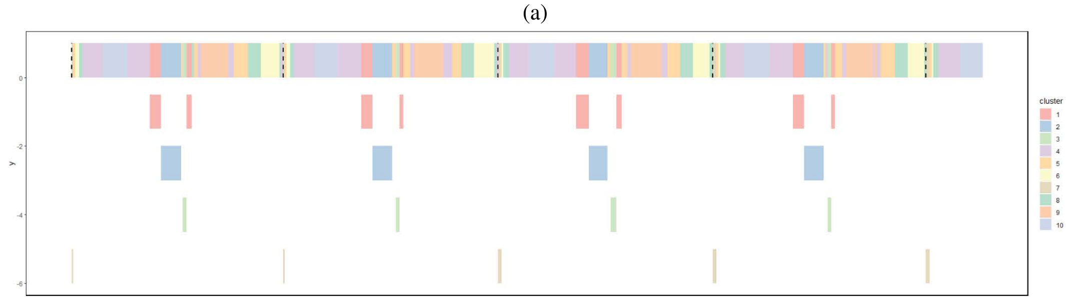


FIGURE 4. HC-tree marked with 10 clusters of 6-dim time series, each of which retains distinct dependency.



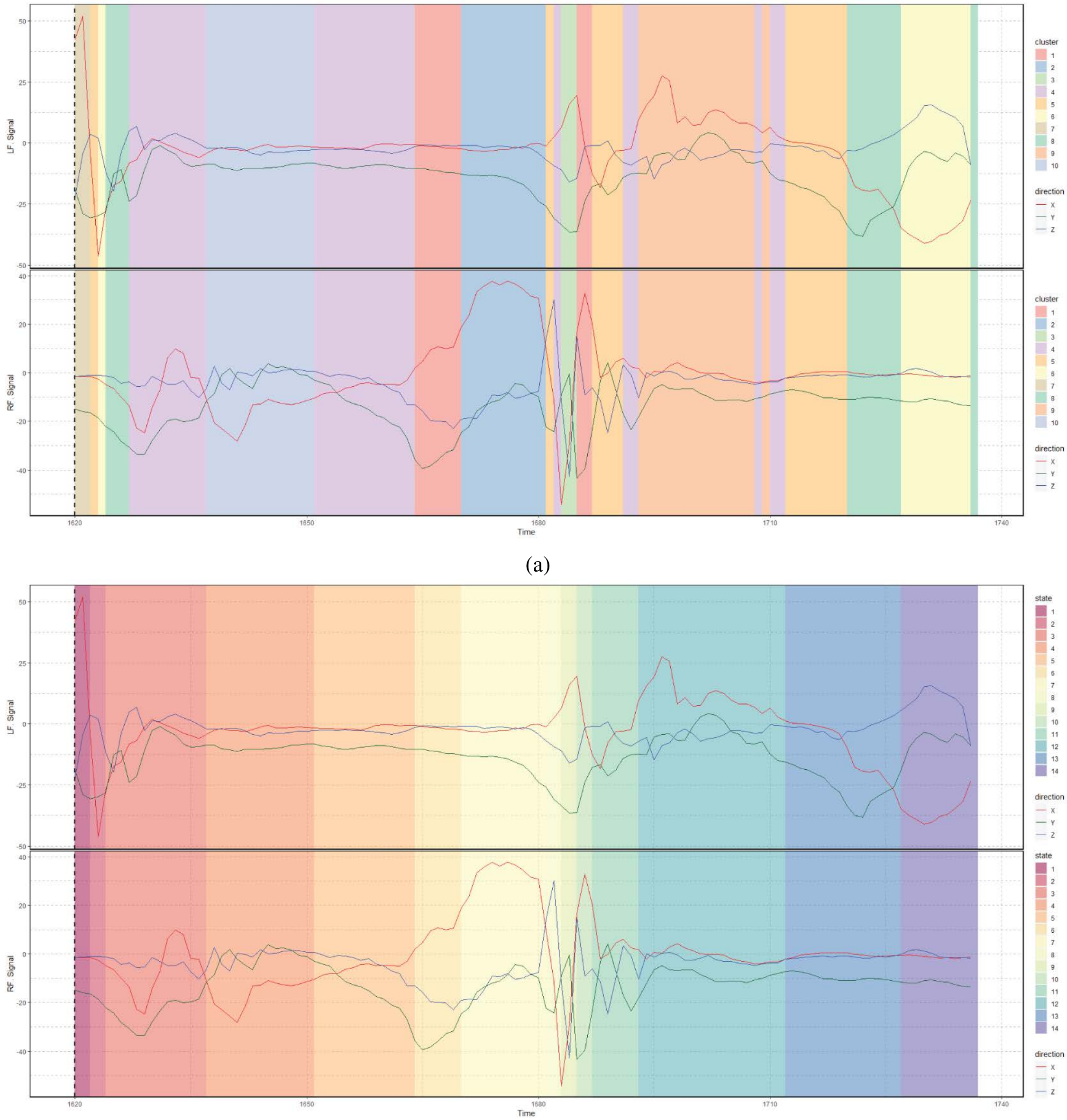
(a)



(b)

FIGURE 5. Run illustration (a) Time-series marked by color-coded clusters; (b) Illustrations of 4 run series w.r.t 4 clusters.

Based on many dimensional time series data collected from multiple sensors, we not only identify all visible and invisible biomechanical states involved within each cycle of gait dynamics of this person of interest, but we also infer multi-scale deterministic and stochastic structures from the entire sequence of cycles down to very detailed temporal patterns



(a)

(b)

FIGURE 6. Comparing cluster-based and biomechanical states on one cycle of 6-dim time series: (a) cluster-based states; (b) biomechanical states.

within each biomechanical state. These multiscale deterministic and stochastic structures collectively become a basis for mimicking a person's specific gait dynamics.

Such fundamental ideas underlying our mimicking protocol are also rather essential in following a mechanical sense. Since a deterministic structure would only allow a specific kind of stochasticity to occur, a multiscale composition of deterministic structures would constrain a specifically corresponding multiscale composition of stochasticity pertaining to this person's gait dynamics. Therefore, understanding where and what deterministic constraints and stochastic variations are across the entire collection of multi-dimensional time series is the essential computation in any complex system study [18]. Such computing indeed makes this study of gait dynamics another unique study of the complex system.

Thus, by adhering to the complex system perspective, our mimicking protocol would be a big step towards the precision learning of human activity, and at the same time would bring resolutions to all aforementioned questions regarding stability-volatility and detection of minute structural changes in gait dynamics.

III. STRUCTURAL DEPENDENCY AND CYCLES EMBEDDED WITHIN GAIT RHYTHMIC TIME SERIES

We begin this section by discussing structural dependency embedded within the 6-dim gait rhythmic times series

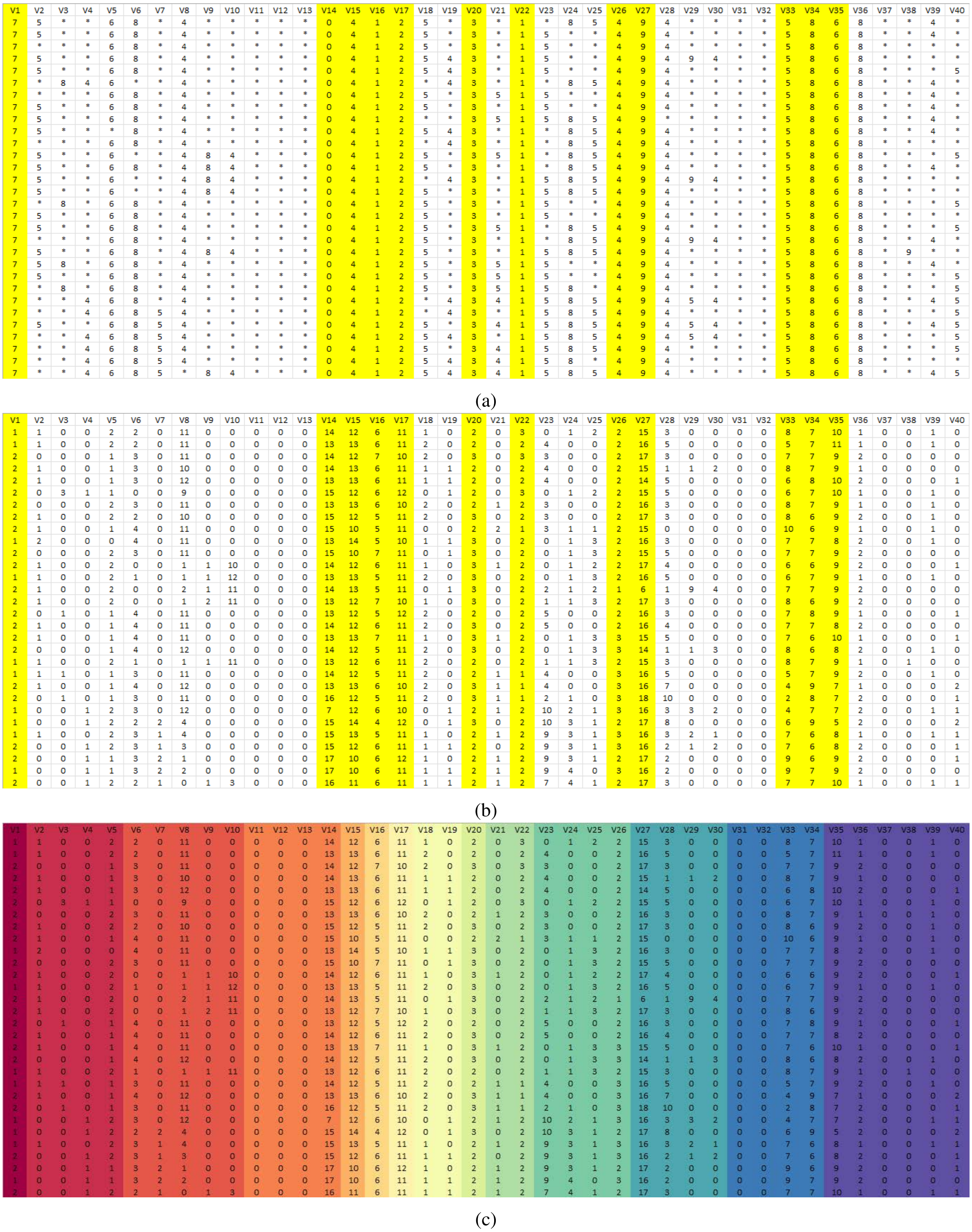


FIGURE 7. Computed and marked cycles of biomechanical states of [LF-vs-RF] subsystem. (a) Data-driven cyclic-state compositions are represented with identified deterministic and stochastic structures in 40-dim digital vectors; (b) Lengths of cyclic states are represented within a 40-dim digital vector; (c) Color-coded biomechanical states are identified as segments of digital coding having a "stable" duration across all cycles.

pertaining to the [LF-vs-RF] subsystem in general terms in the first subsection. Also, we explicitly show such evolving dependency structures through two representations of gait time series. Then, in the second subsection, we apply a clustering-based computational approach to encode the 6-dim time series into one 1-dim categorical time series, with whom

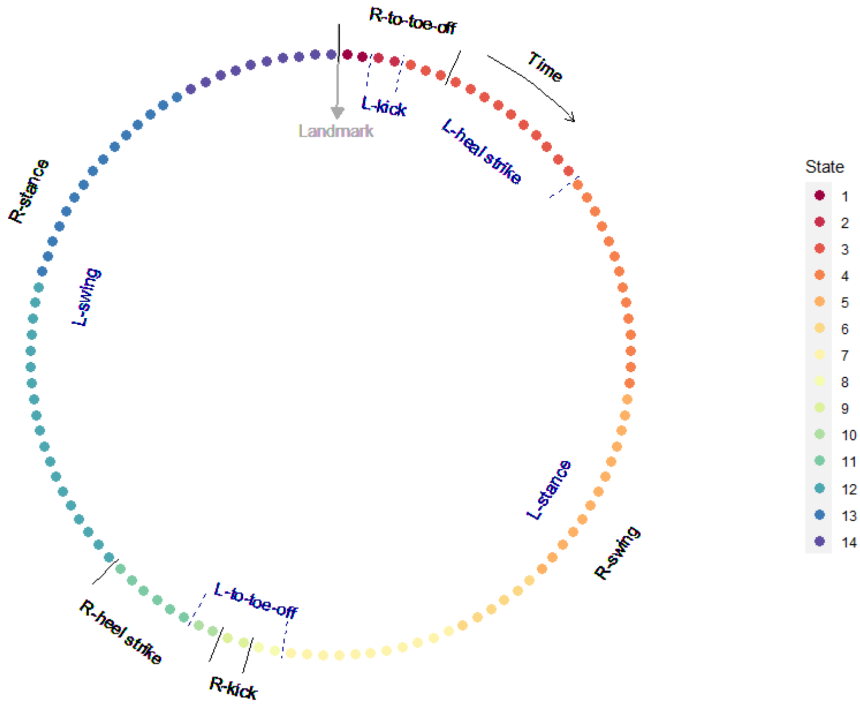


FIGURE 8. One cyclic representation of 14 biomechanical states.

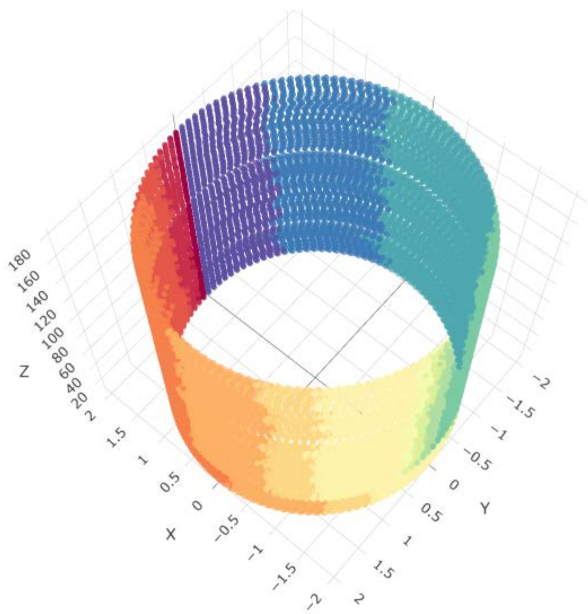


FIGURE 9. 3D cylinder of tempo-ordered cycles with 14 biomechanical states. The rotatable 3D plot can be seen through https://statistics2022.github.io/3Dplots/Figure%208.

an algorithm of computing landmark is developed and the whole time span is dissected into cycles.

A. DATA VISUALIZATIONS AND REPRESENTATIONS OF GAIT RHYTHMIC TIME SERIES

We consider two natural representations of directional forces time series recorded by accelerometer sensors to serve as the basis for exploring intrinsic patterns of gait dynamics. Such representations are not only for dynamic gait-pattern recognition but also for the system's exploratory computations. The first natural representation is a joint display of their trajectories, as seen in Fig. 2. As we look at the 6 trajectories of two triplets of time series from Left-foot and Right-foot in two separate panels of Fig. 2, one simple, but the crucial question arises: Should we explore these two triplets separately or jointly?

Ideally, if the triplet of 3-dim directional forces time series from the Left-foot sensor should and could reveal rhythmic state-patterns pertaining to the Left-foot of a subject's musculoskeletal system, then the triplet of 3-dim directional forces time series from the Right-foot sensor should and could reveal the complimentary rhythmic state-patterns pertaining to the Right-foot of a subject's musculoskeletal system. We can compute such sensor-specific patterns respectively and then align them together along their common temporal axis. That is, with respect to their common time axis, computed foot-specific cycles and their evolving states from the two sensors would be coupled together to show their interacting relational patterns. However, this coupling could be too rigid to reflect intrinsically steady interacting relational patterns. That is, too many interacting relational patterns are likely to arise because computed foot-specific states are varying in their durations. This phenomenon was shown and concluded in an experiment in [10] via Lempel-Ziv complexity evaluations.

On the other hand, since Left-foot and Right-foot are two highly coordinated components of the human musculoskeletal system. these 6 time series are supposed to be highly

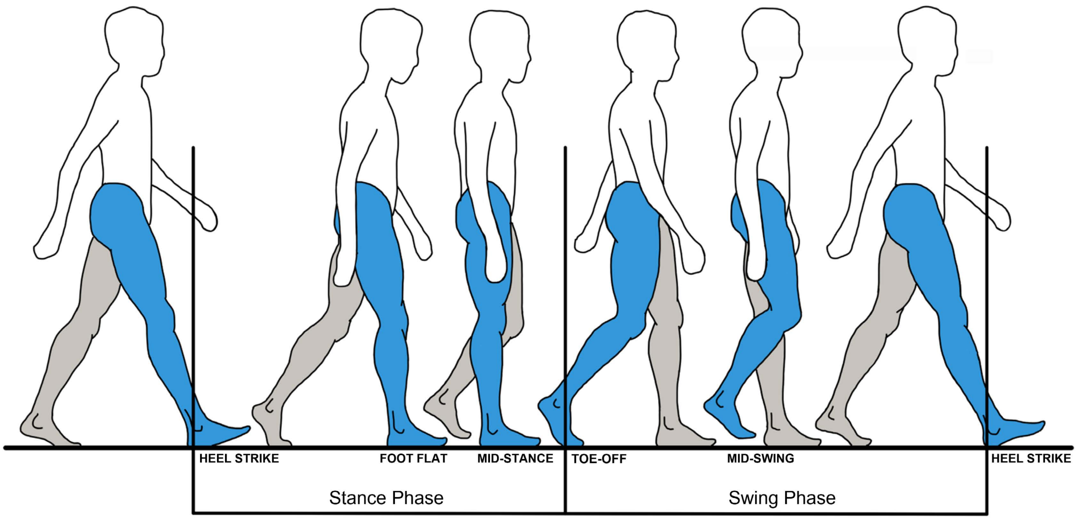


FIGURE 10. One typical gait cycle with "Sequential Movement".

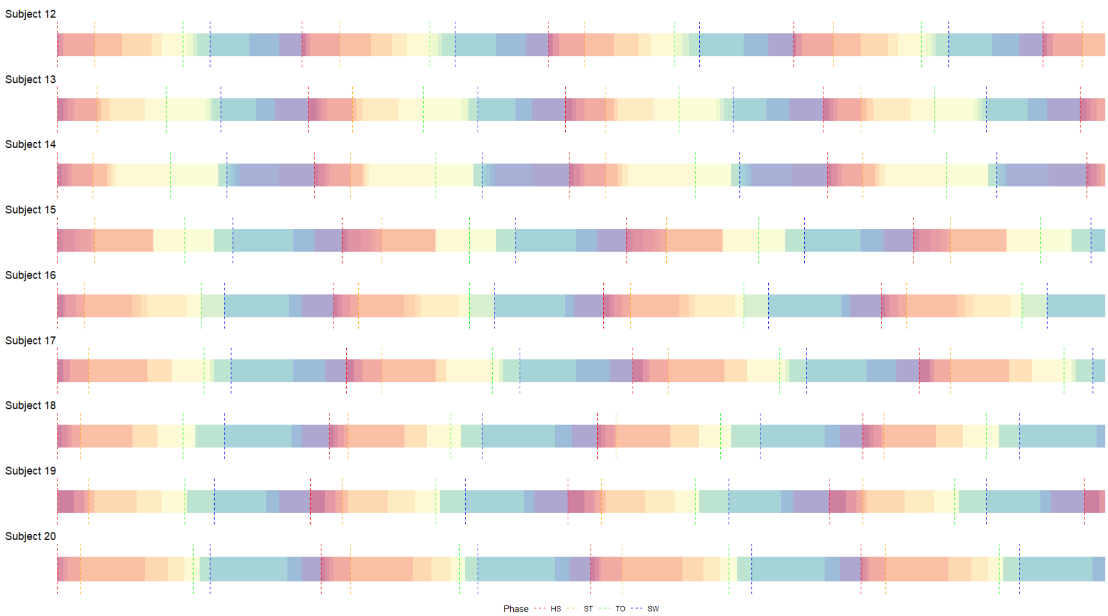


FIGURE 11. Comparing 9 subjects' gait dynamic rhythms of [LF-vs-RF] subsystem with respect to a biomechanical structural representation of their 4 gait cycles.

dependent in constituting the biomechanical essence of this person's gait dynamics. So, it is natural and intuitive that the 6-dim directional forces time series from the Left-foot and Right-foot sensors should constitute a full version of structural dependency pertaining to this subject's musculoskeletal [LF-vs-RF] subsystem. In order to capture such structural dependency well, it is the right choice to compute rhythmic patterns based on the 6 trajectories simultaneously. Not only we can visualize rhythmic state-patterns better based on 6-dim time series trajectories, but also we can compute these patterns more precisely. As such, idiosyncratic characteristics of person-specific gait dynamics can be more effectively computed and discovered.

In summary, from the perspective of [LF-vs-RF] subsystem, these 6-dim trajectories simultaneously traversing along the common temporal axis would jointly reveal various kinds of dynamic structural dependency accordingly embedded within all visible recurrent patterns (states) in an evolving fashion within each cycle and across all cycles. Furthermore, each of such structural dependency based recurrent states needs to be computationally extracted as one whole. Therefore, a feasible computational algorithm for such structural dependency must be equipped with the essential capability of preserving various kinds of dependency-constituted patterns all at once.

The second natural representation is done by giving up the temporal coordinates of data points via a 3D scatter plot for each sensor. Such a 3D plot reveals a unique sensor-specific geometric manifold that provides one sensor-specific view of the stability and volatility of gait dynamics in a visible fashion. Geometries of this kind indeed afford a global view of idiosyncratic characteristics of a person's gait dynamics, so via mimicking they play particularly essential roles in detecting minute structural changes.

This second natural representation of gait time series data is formally described as follows. An accelerometer sensor gives rise to $X-$, $Y-$, and $Z-$ directional acceleration measurements in time series format. For expositional simplicity, we denote the 3 time series from the left foot sensor as (L_x, L_y, L_z), and the 3 time series from the right foot sensor as (R_x, R_y, R_z) without using a temporal index, such as t. The 3D scatter plots of (L_x, L_y, L_z) and (R_x, R_y, R_z) display two geometric manifolds, as seen respectively in the two panels in Fig. 3 and two rotatable 3D plots through the two links provided in the legend of this figure. Such geometries evidently reveal aspects of structural dependency within each of the two triplets of measurements of "directional forces in R^3," which are very distinct from structural dependency revealed by 6-dim time series trajectories along their temporal axis.

B. COMPUTING CYCLES, LANDMARK, AND STATES WITHIN TIME SERIES

As seen in Fig. 2, Left-foot (LF) and Right-foot (RF) versions of structural dependency act in concert with the gait dynamics of the [LF-vs-RF] subsystem. Therefore, here the first computational task is to compute and analyze the subject's authentic structural dependency embraced in 6-dim time series $(L_x, L_y, L_z, R_x, R_y, R_z)$ as the first step of discovering rhythmic patterns of this subject's biomechanics.

The 6-dim trajectory of $(L_x, L_y, L_z, R_x, R_y, R_z)$ is rhythmic. This rhythmicity means that all involving biomechanical states in a subject's gait dynamics are all recurrent. That is, similar 6-dim data points must be seen across all cycles. Even within a cycle, some biomechanical states proceed smoothly, such as Stance, Toe-Off, and Swing, and some states proceed via more volatile fashions, such Heel-Strike. Thus, similar 6-dim data points are also seen within each cycle. Such recurrent and smoothly evolving similarity-based patterns of 6-dim data points indicate explicitly visible structural dependency. The intuitive idea is to collect, mark and encode all similar 6-dim data points across the entire temporal span.

To collect various kinds of similar 6-dim data points, we stacked the 6-dim time series by arranging them onto 6 rows of a data matrix with all time points being arranged along the column-axis. We apply the HC algorithm among all 6-dim column vectors by simply using Euclidean distance and a Ward-D2 module. That is, we take off data points' temporal coordinates in this clustering step. The clustering results are reported in Fig. 4, for example.

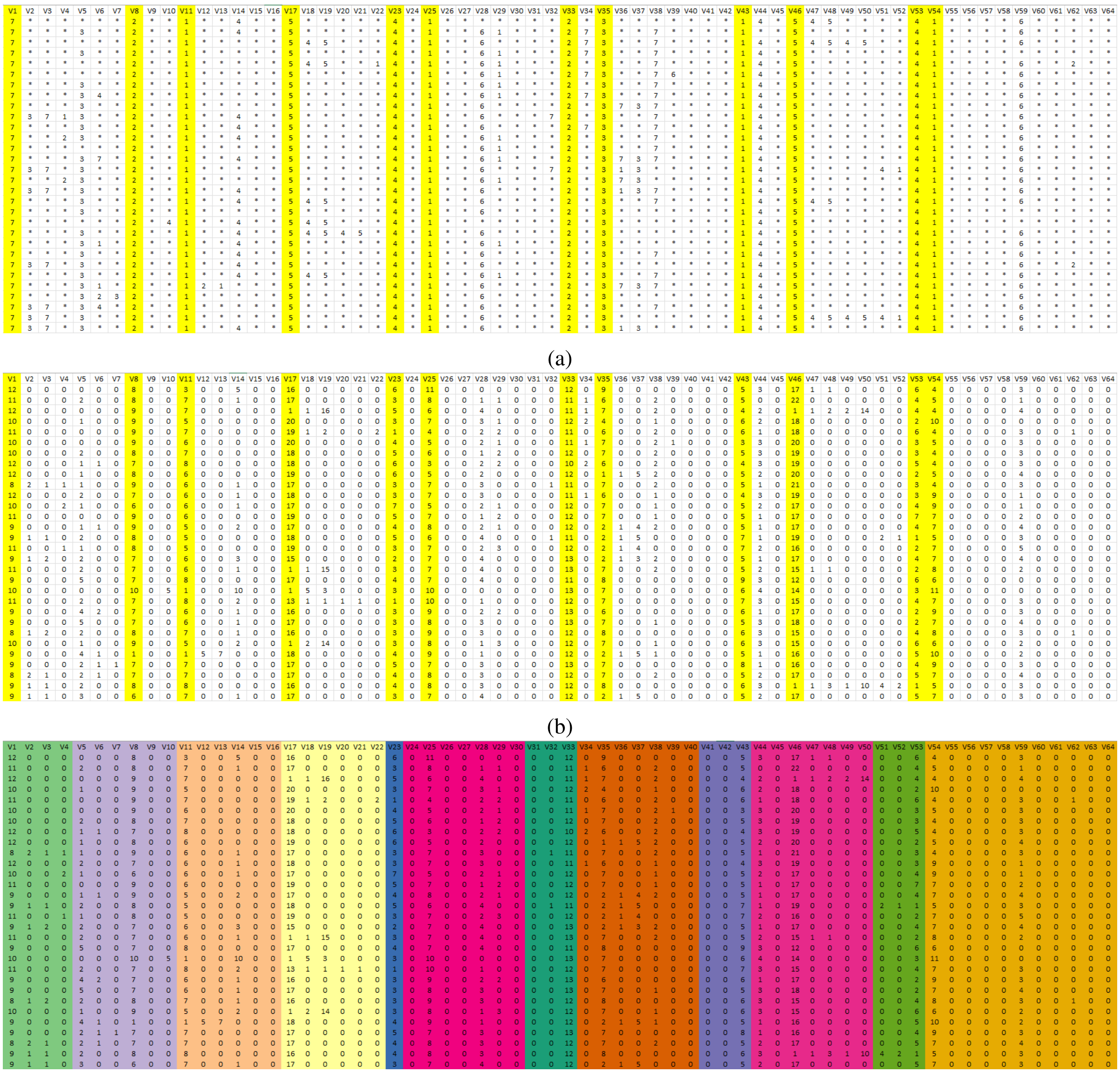
As shown in Fig. 4, there are 10 marked branches as 10 clusters for encoding purposes. When each cluster is encoded with a color code, each time point is encoded with a color code. Via this color-coding scheme, we reduce the 6-dim time series of $(L_x, L_y, L_z, R_x, R_y, R_z)$ into a 1-dim color-code time series, which exhibits color-codes' segment-wise duration and transitions, as illustrated in Fig. 5a. Each color-coded segment of this 1-dim time series in this panel is termed a cluster-based state. Though these recurrent cluster-based states along the temporal axis can preliminarily expose the hidden rhythm of gait dynamics, the chief utility of such a 1-dim color-coded time series is that it becomes the basis for computationally identifying a so-called landmark state, with which we dissect the whole time span into cycles.

In Fig. 5b, we illustrate 4 series of colored runs induced from the time series of color codes with respect to cluster-states {#1, #2, #3, #7}. Each run is a period of time of one single color continuum. They all seemingly reveal very regular cyclic patterns, but they come with different precisions. Here the precision can be defined via the variations of recurrence-time series of the beginning or end of a series of single-color runs. For this subject, the most stable cyclic pattern is calculated and found on the #7 cluster state. This is then called the landmark of the [LF-vs-RF] subsystem of this subject. It is essential to note that every subject has his/her own landmark. That is, it is far from being a universal trait of human gait dynamics. The chief merits of a computed landmark are rather diverse. It is used to dissect a subject's whole gait time series into cycles. The representation of such a cycle surely constitutes the first visible impression of a subject's gait dynamics, as seen in Fig. 6a.

A landmark also very importantly serves as the starting point for alignment across all cycles in order to bring out the biomechanical states, as would be discussed in the next section. For the contrasting and comparing purpose, the same cycle would be featured with would-be computed biomechanical states as shown in Fig. 6b.

IV. STRUCTURAL REPRESENTATIONS OF GAIT CYCLES

The human musculoskeletal system and its biomechanics are very robust even for walking on slightly uneven pavement. In sharp contrast, wearable sensors are relatively sensitive to even tiny deviations in forces and directions of walking. Though a computed landmark can manage to dissect rather stable cycles in length, all cycles' compositions of cluster-based states and states' duration could vary to some visible degrees, as shown in Fig. 7a and Fig. 7b. These kinds of variations are more related to sensors than to the subject's musculoskeletal system and biomechanics. In other words, the compositions of biomechanical states of cycles in human gait dynamics are expected to be stable in both aspects: cyclic evolving series and state-specific durations. That is, the collection of computed cluster-based coded states needs to be



(a) (b) (c)

FIGURE 12. Computed and marked cycles of biomechanical states of the waist.

transformed into a collection of ''biomechanical states'' that can retain stability in both aspects.

We develop an ''Biomechanical Structural Representation Algorithm'' to computationally construct and identify a composition of biomechanical states for each dissected gait cycle in this section. This algorithm is carried out via 2 steps of data-driven computations given as follows.

Biomechanical Structural Representation Algorithm

1) *Identifying and aligning conserved cluster-state-codes:* Display all dissected cycles' cluster-coding segments by aligning their landmark state and stacking them in a manner of ''one row one cycle'' with respect to their temporal ordering, and then proceed to identify the first conserved cluster-code across all cycles and then align them. This alignment operation might create empty spaces along the row axis for some cycles. Then we proceed to identify the 2nd conserved cluster code across all cycles and align them. We continue such operations of identifying conserved cluster code to the end of all cycles. Then, we mark the identified series of conserved cluster codes across all cycles, as demonstrated in Fig. 7a. At the end of the first step, all cycles' cluster-based code segments together with inserted empty spaces form a cluster-based code matrix;
2) *Transforming and grouping columns:* Transform the resultant cluster-based code matrix into a duration matrix by replaying each cluster-based code with its duration, as demonstrated in Fig. 7b. We then manually group neighboring columns when they achieve stable

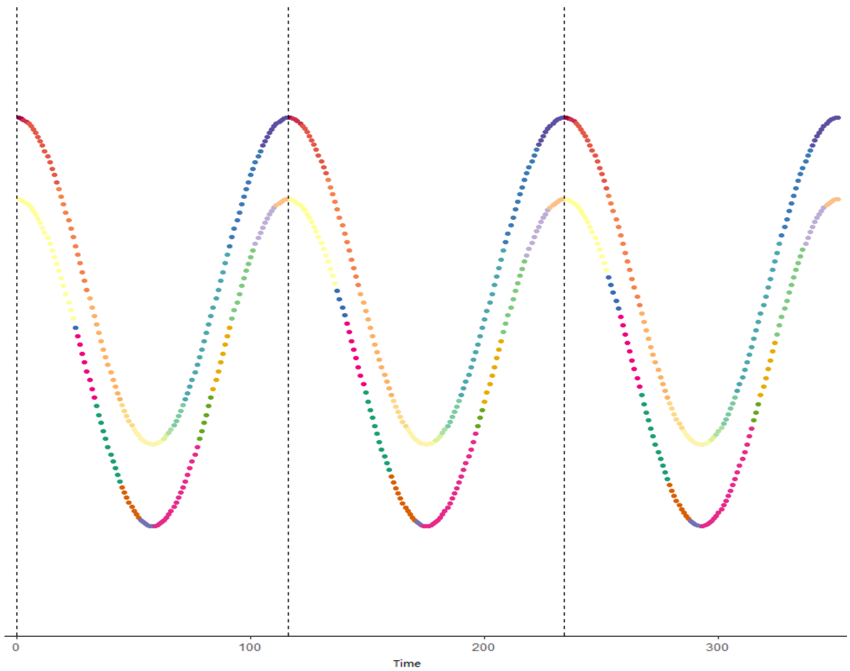


FIGURE 13. Cyclic correspondence between rhythms of [LF-vs-RF] subsystem and Waist.

sums of durations across all cycles. Each of such column groups is marked and termed one biomechanical state. They are demonstrated in Fig. 7c. This version of the duration matrix marked with biomechanical states is termed the structural representation of all involved cycles, and so is the version of the cluster-based code matrix marked with biomechanical states.

The structural representation of 30 cycles of subject-12's gait time series is displayed in Fig. 7c. One single cycle's composition of biomechanical states is displayed with respect to the 6-dim time series in Fig. 6b and also in a circular fashion in Fig. 8, while this subject's 163 cycles are stacked into a 3D cylinder in Fig. 9. Fig. 10 is an illustration of a typical gait cycle which can be compared with Fig. 8 for better understanding. From the latter, we can see that the cyclic presence of biomechanical states and their durations are indeed rather stable. Such stability is the consequence of the structural representation of gait dynamics. This biomechanical structural representation of gait dynamics will serve as the right platform for comparing gait dynamics across different subjects. As shown in Fig. 11, 9 subjects' [LF-vs-RF] gait dynamics are compared. This is one of several key merits of biomechanical structural representation.

A. COMPUTED RHYTHMIC DYNAMICS OF WAIST AND WRIST

We apply a similar computational protocol used for the 6-dim time series of [LF-vs-RF] subsystem and the "Biomechanical Structural Representation Algorithm" on the 3-dim time series observed through an accelerometer sensor tied to the waist of subject-12, as illustrated in Fig. 12. The waist's 12 biomechanical states are shown to be very much in synchrony with the computed rhythm of [LF-vs-RF] subsystem, as demonstrated in Fig. 13, via a sine-wave representation. This synchrony between the Waist and [LF-vs-RF] subsystem can be further viewed through time series presentation in the top three panel rows of Fig. 14. Such evident synchrony will serve as a basis to investigate the interacting patterns between the Waist and [LF-vs-RF] subsystem within subject-12's gait dynamics.

It is further noted that, when we apply the same computational protocol onto 3-dim time series observed from an accelerometer sensor tied to subject-12's wrist, we only find its synchrony with [LF-vs-RF] subsystem in one time period near the beginning of the recording, as shown in the bottom panel-row of Fig. 14. We found only an out-of-sync pattern with the rest of the period of recording.

V. MIMICKING

In this section, we demonstrate the chief merit of the structural representation of all gait cycles as demonstrated in Fig. 6. The structural representation would serve as the base to build our mimicking protocol for any person-specific gait dynamics, at least for the [LF-vs-RF] subsystem. We organize this section as follows. In the first subsection, we build our mimicking protocol to first mimic the large and medium-scale structures visible across all computed gait cycles as shown in Fig. 9 in the previous subsection. In the second subsection, we mimic the fine-scale structures. In the third subsection, we apply our mimicking protocol to fill in the

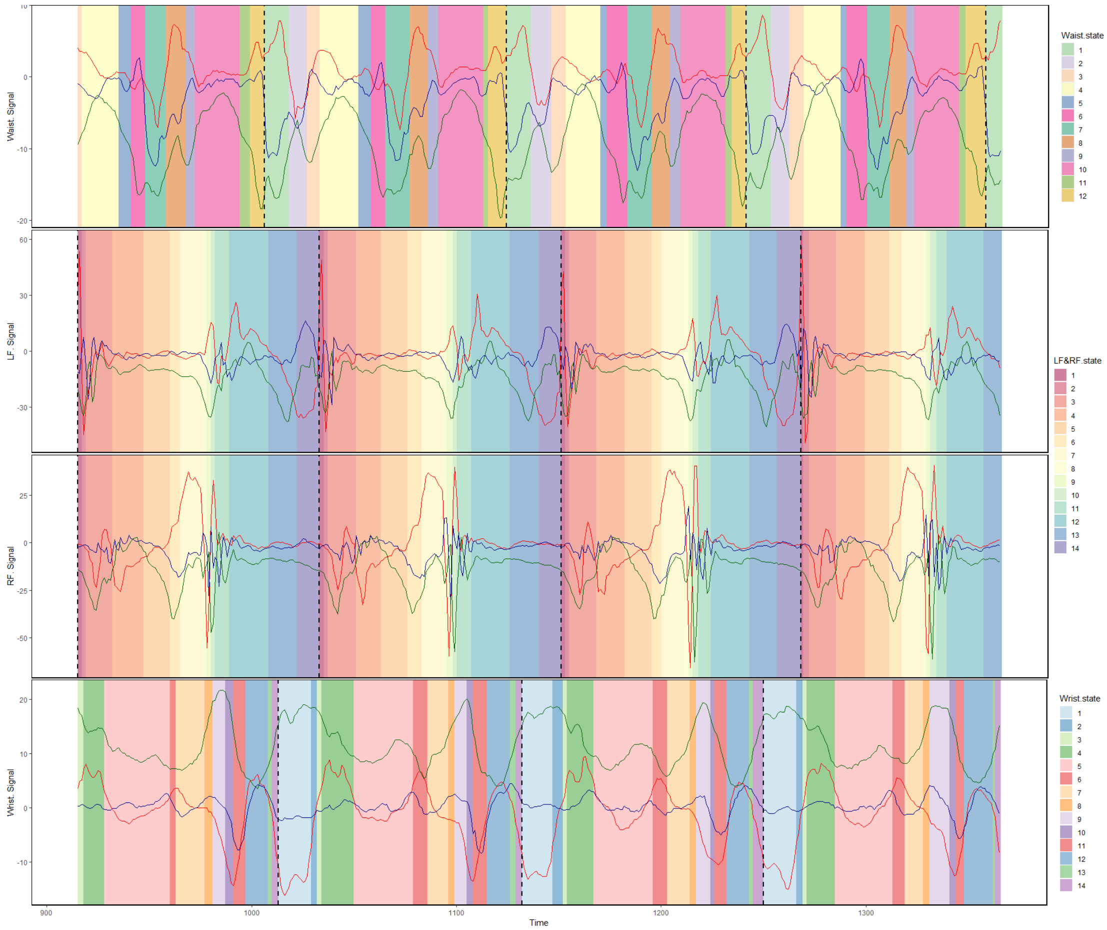


FIGURE 14. Drifting patterns observed on time series of Wrist with respect to rhythms of [LF-vs-RF] subsystem.

spacing of Left-foot and Right-foot specific scatter plots into manifolds, respectively, for the purposes of precise detecting and learning minute changes.

A. MIMICKING LARGE AND MEDIUM SCALE STRUCTURES OF RHYTHMIC GAIT CYCLES

It is intuitive and observable that some cycles are more similar to some cycles than other cycles. Such similarity with varying degrees among cycles is indeed fully reflected through the durations of 14 biomechanical states in a collective fashion. That is, if two cycles have similar durations across their 14 biostates, then they are similar. This similarity renders similar cycle lengths. To further confirm this intuition with observable patterns, it is natural to build a matrix with all involved cycles arranged along the column axis, and each cycle's 14 state-durations are arranged along the row axis. We term this matrix: biostate-duration matrix. We simply apply the HC algorithm with Euclidean distance and Ward-D2 module, then we build a HC-tree on the collection of cycles and superimpose this HC-tree on the column-axis of this biostate-duration matrix. Likewise, we also build a HC-tree on its row-axis. After respectively superimposing two HC-trees on the column- and row-axes, this biostate-duration matrix becomes a heatmap, as shown in Fig. 15. We then mark and map out the block patterns in this heatmap, which somehow reveal the deterministic patterns of biostate durations.

As one marked HC-tree branch on the column-axis reveals a cluster of similar gait cycles, its corresponding series of blocks as computed deterministic structures reveals the characteristics of durations across the 14 biomechanical states. In fact, each block found in this heatmap clearly embraces the stochastic patterns among similar cycles. It is essential to recognize that such stochasticity is indeed constrained by deterministic structures. Our mimicking protocol would be based on such deterministic and stochastic patterns.

Specifically speaking, if we want to mimic the k-th gait cycle, we need to locate which branch this cycle is from, then we know all cycles similar to the k-th cycle. This is structural information on a large scale. Further, this collection of similar gait cycles would offer proper stochasticity for each of the 14 biomechanical states. This is structural information of the medium scale.

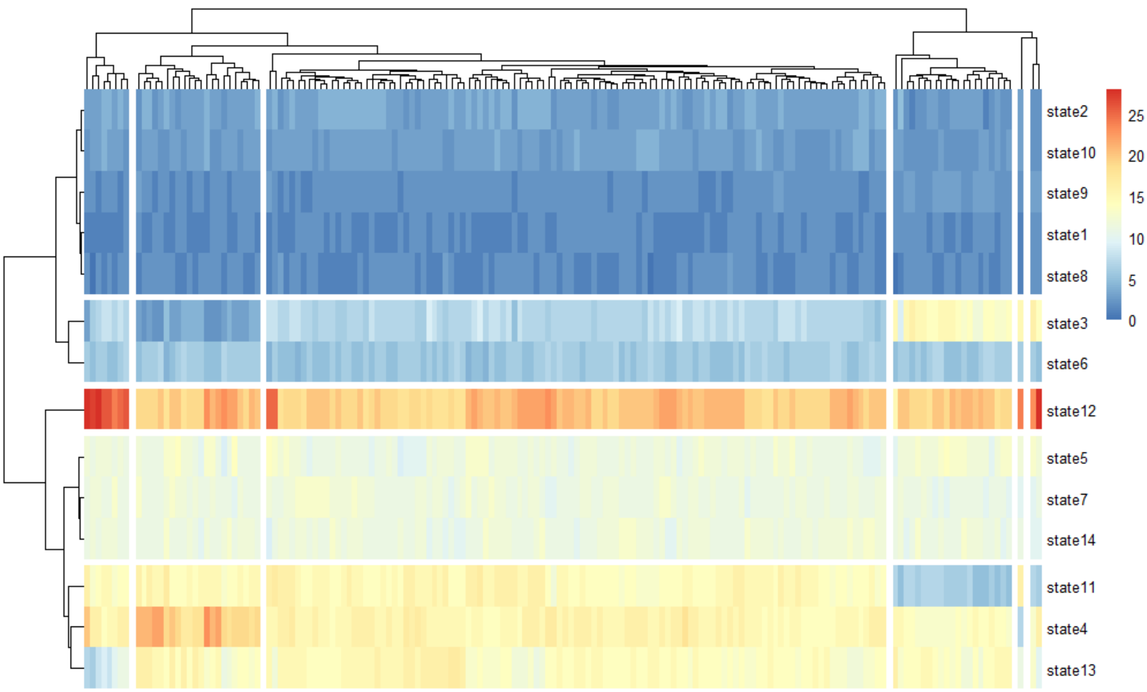


FIGURE 15. Heatmap of 14 biomechanical states' lengths of [LF-vs-RF] subsystem.

B. MIMICKING FINE SCALE TRAJECTORIES OF A BIOMECHANICAL STATE WITHIN A RHYTHMIC GAIT CYCLE

In Fig. 15, we explore and see the large-scale structures of the designated k-th cycle through the HC-tree branch it falls into on the column-axis. That is, all gait cycles falling into the same marked branch as this k-th cycle falling into are indeed structurally similar in the sense that they all have similar state-specific durations across all 14 biomechanical states. For instance, if $k = 11$, then it belongs to a branch that contains 9 structurally similar cycles as listed in Cycle-column of Table 1.

That is, we can recognize this k-th cycle's medium-scale structures of all its 14 biomechanical states' durations through a series of blocks under the marked HC-tree branch it belongs to. Specifically speaking, these block-specific or even state-specific structures of medium scales clearly indicate the information of which cycles do share a similar duration of each biomechanical state. Based on such available information, we further explore the temporal-related fine-scale deterministic and stochastic structures pertaining to each biomechanical state's 6-dim measurements.

Our explorations of such fine scale pattern information begin by collecting all state-specific time points' 6-dim measurements of all cycles falling into the same marked HC-tree branch. Since a state-specific trajectory in general consists of different temporal segments: beginning, middle, and ending parts. Thus, we make use of HC to perform data-driven partitions. We choose the number of focal clusters being equal to the median duration of this state, and we call them biostate-tempo-clusters.

For instance, Table 1 lists time-points' 6-dim measurements of a small part of #1-biostate-tempo-cluster #1. We see that some cycles' 1st and 2nd time points are falling into the same #1-biostate-tempo-cluster #1. In contrast, some cycles have their 1st and 2nd time points' 6-dim measurements falling into different #1-biostate-tempo-clusters. Therefore, these two examples of #1-biostate-tempo-clusters will together offer two binomial random variables, with which simulated 6-dim measurements of the 1st and 2nd time points will come from. Likewise, multiple multinomial random variables are accordingly created when a biomechanical state has multiple biostate-tempo-clusters.

To mimic by simulating a 6-dim measurement from a biostate-tempo-cluster, we construct a Principle Component Analysis (PCA) based protocol as follows. We apply the PCA algorithm onto a collection of 6-dim measurements, which is denoted as:

$$\{\boldsymbol{v}^i = (x^i_{(L)}, y^i_{(L)}, z^i_{(L)}, x^i_{(R)}, y^i_{(R)}, z^i_{(R)})|i = 1, \ldots, n\},$$

belonging to a generic focal biostate-tempo-cluster. Based on PCA, its 6 eigenvectors are extracted and denoted as $\{\boldsymbol{V}_1, \ldots, \boldsymbol{V}_6\}$. And let the 6×6 matrix $M[V]$ have the 6 eigenvectors as its row-vectors.

Denote $\varepsilon^i_j = \boldsymbol{V}^T_j \boldsymbol{v}^i$ with $j = 1, \ldots, 6$. That is, the 6-dim vector $\boldsymbol{\varepsilon}^i = (\varepsilon^i_1, \ldots, \varepsilon^i_6)$ is a projection of $\boldsymbol{v}^i$ with respect to

TABLE 1. The #1 biomechanical state or landmark state's temporal coordinates in [LF-vs-RF] subsystem and their corresponding cluster # of their 6-dim measurement vectors. This state is the color-coded cluster #7 in the HC-tree.

Time-pt	LF_X	LF_Y	LF_Z	RF_X	RF_Y	RF_Z	Cycle	State	Tempo-Coord	Cluster #
1268	59.608	-26.196	-20.706	-0.627	-14.902	-1.725	11	1	1	1
1384	51.608	-15.059	-22.275	-1.098	-15.059	-1.569	12	1	1	1
1502	64.784	-18.353	-14.275	-1.882	-14.118	-2.196	13	1	1	1
1503	48.627	-28.078	-6.431	-1.882	-14.745	-1.569	13	1	2	1
1620	42.039	-18.667	-22.588	-1.569	-14.902	-1.412	14	1	1	1
1621	51.922	-28.863	-4.235	-1.569	-16	-1.255	14	1	2	1
1737	39.686	-20.235	-16.471	-0.627	-14.275	-1.569	15	1	1	1
1738	61.333	-27.765	-18.824	-0.471	-14.431	-1.412	15	1	2	1
1857	57.098	-17.725	-29.49	-1.412	-15.373	-1.725	16	1	1	1
1858	69.961	-21.961	-39.216	-2.039	-16.784	-0.784	16	1	2	2
1974	48.471	-14.902	-18.039	-1.569	-16	-2.039	17	1	1	1
1975	51.451	-29.647	-8.627	-1.569	-16.941	-1.725	17	1	2	1
2091	39.216	-21.804	-9.255	-1.412	-14.902	-2.039	18	1	1	1
2092	67.451	-30.275	-20.235	-1.569	-15.059	-1.725	18	1	2	1
2209	38.275	-21.49	-8.157	-1.098	-13.804	-3.137	19	1	1	1
2210	46.118	-28.706	-10.51	-0.784	-14.431	-2.51	19	1	2	1

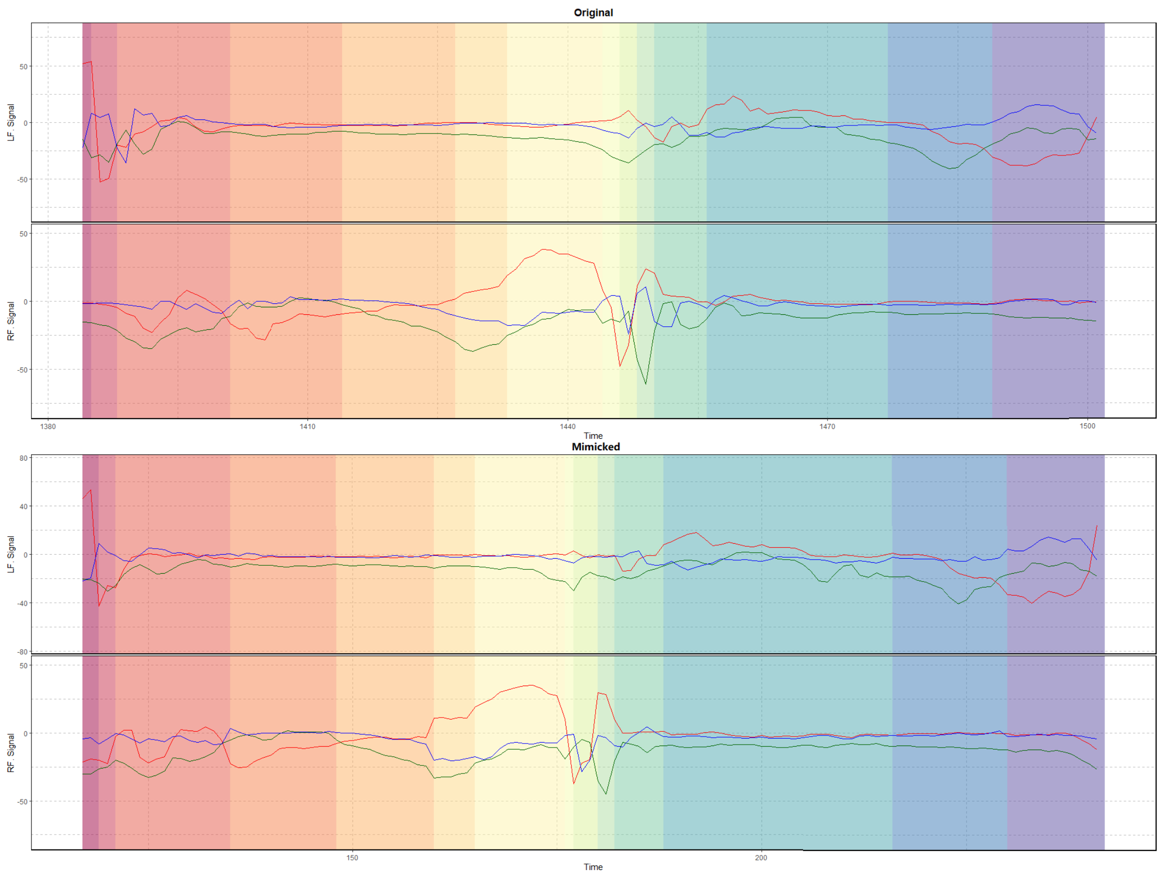


FIGURE 16. One observed and mimicked the cycle of Left-foot and Right-foot rhythm.

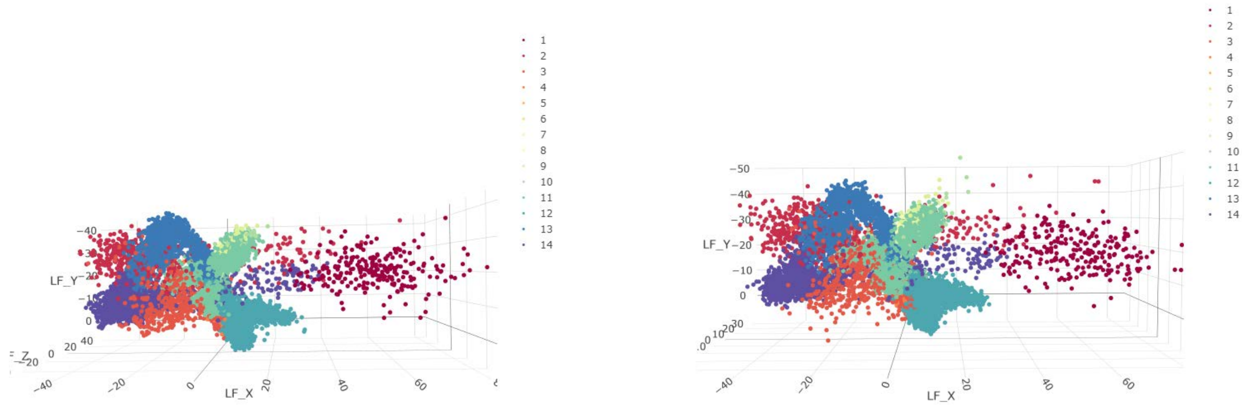
the 6 eigenvectors, that is, for all $i = 1, \ldots, n$, we have

$$\boldsymbol{\varepsilon}^i = M[V]\boldsymbol{v}^i. \tag{1}$$

Then, we build one histogram, say H_j, based on the collections: $\{\varepsilon_j^i | i = 1, \ldots, n\}$ with $j = 1, \ldots, 6$. It is essential to note that these 6 histograms define 6 stochastically independent random variables. Thus, we can simulate 6 components of 6-dim vector $\boldsymbol{\varepsilon}^{i*}$ from the 6 histograms $\{H_j | j = 1, \ldots, 6\}$, respectively. Then, we convert $\boldsymbol{\varepsilon}^{i*}$ into a mimicked 6-dim $\boldsymbol{v}^{i*}$ as follows:

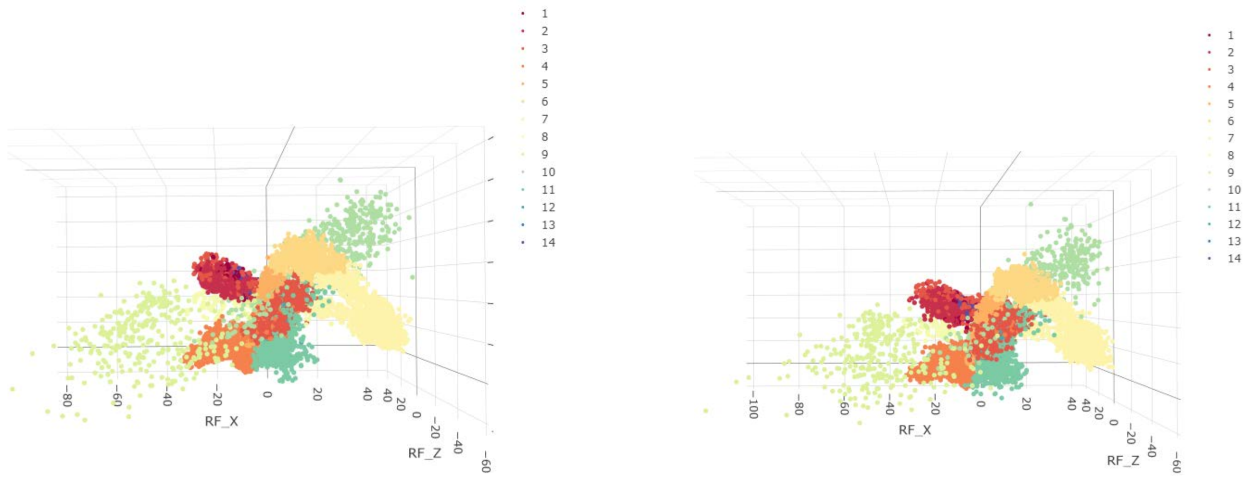
$$\boldsymbol{\varepsilon}^{i*} = M[V]\boldsymbol{v}^{i*}. \tag{2}$$

With this PCA-based protocol, we can simulate as many copies of $\boldsymbol{v}^{i*}$ as we wish to generate. Therefore, by applying



(a) (b)

FIGURE 17. Observed (a) and its mimicked (b) 3D manifolds of Left-foot rhythm. The corresponding rotatable 3D plots: (a) https://statistics2022.github.io/3Dplots/Figure15-PanelA; (b) https://statistics2022.github.io/3Dplots/Figure15-PanelB.



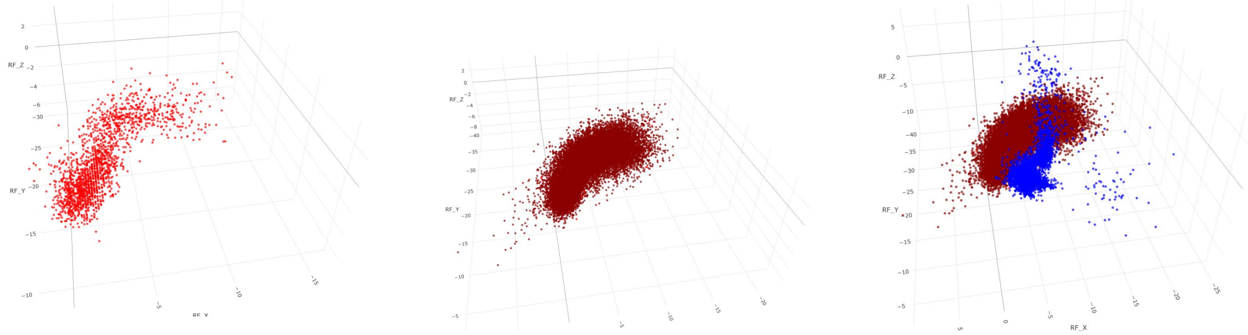
(a) (b)

FIGURE 18. Observed (a) and its mimicked (b) 3D manifolds of Right-foot rhythm. The corresponding rotatable 3D plots: (a) https://statistics2022.github.io/3Dplots/Figure16-PanelA; (b) https://statistics2022.github.io/3Dplots/Figure16-PanelB.

the PCA-based protocol on the state-specific biostate-tempo-clusters' multinomial random variable, we can mimic this focal state's trajectory as many times as we wish to generate. We demonstrate the results of our mimicking protocol via three versions of mimicries. In Fig. 16, we compare the 6 Left-foot and Right-foot time series of an observed cycle with 6 time series of its mimicked cycle. We can see that all large- and medium-scale characteristics are almost equal, while some small deviations are visible on fine-scale structures. In Fig. 17 and Fig. 18, we compare the observed 3D manifold of subject-12's 14 biomechanical states with the mimicked 3D manifold on the Left-foot and Right-foot, respectively. We demonstrate the very close similarity between the observed and mimicked manifolds and this high-level similarity indeed proves the accuracy of our developed computational protocol.

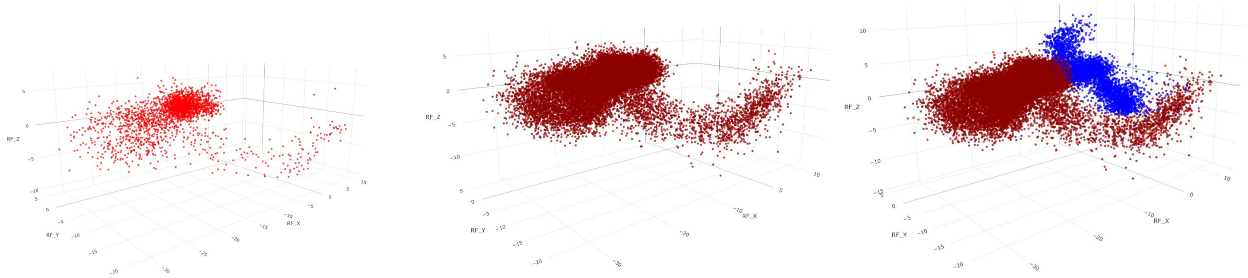
C. PRECISION-DETECTIONS OF FINE-SCALE-DEPARTURES VIA MIMICKED MANIFOLDS

In this subsection, we discuss and illustrate one of the chief merits of mimicking gait dynamics: precision detections of fine-scale departures within any biomechanical state. The idea behind precision detection is the fact that mimicking can fill in all stochasticity-related spacings allowed by the constraints of deterministic structures. That is, all filled-in (solid) spaces supported by stochasticity are taken as confidence regions, while any empty spaces are out-of-confidence regions allowed by the deterministic constraints. With this rationale, we build a subject-specific 100% 3D confidence manifold to facilitate our precision-detection of exogenous time series from "different" subjects. In this fashion, any minute departures within any biomechanical state will be detected with precision. We illustrate two such departures



(a) (b) (c)

FIGURE 19. Observed manifold of Right-foot's "before Heel-Strike" state of subject-12 (in Red) (a) and its mimicked manifold (in Dark Red) (b) and comparing subjects-19's observed manifold of the same state (in Blue) (c). See the rotatable 3D plot through the link: https://statistics2022.github.io/3Dplots/Figure17.



(a) (b) (c)

FIGURE 20. Observed manifold of Right-foot's "after Heel-Strike" state of subject-12 (in Red) (a) and its mimicked manifold (in Dark Red) (b) and comparing subjects-19's observed manifold of the same state (in Blue) (c). See the rotatable 3D plot through the link:https://statistics2022.github.io/3Dplots/Figure18.

during two biomechanical states: i) before Heel-strike in Fig. 19; ii) after Heel-strike in Fig. 20, of subject-12.

In Fig. 19a, the 3D scatter plot presents data points of the biomechanical state of ''before Heel-strike'' on the Right-foot. We see the spacing among all observed data points, and then wonder where are covered by stochasticity and where are not. In Fig. 19b, after mimicking, we realize and confirm the domain of stochasticity. With such confirmation, the mimicked manifold precisely marks where spacing are belonging to this biomechanical state of subject-12, and where it is not. In Fig. 19c, we clearly see the departure when superimposing data points belonging to the same state of a different subject. A similar illustration of the biomechanical state of ''after Heel-Strike'' is revealed in Fig. 20 for the Right-foot.

This mimicking-based approach for precision-detection of fine-scale departures for any state of any subject is an important step towards precision learning of walking. It is not only suited for checking whether one subject's gait activities recorded at two or multiple time periods are the same or not, but also for figuring out minute differences in all biomechanical states belonging to two subjects. This approach certainly can be expanded and applied to other kinds of activities far beyond walking, such as dancing and others. And we believe that its merits include far-reaching potentials and impacts through precision learning in many competitive sports, such as track and field. A companion work on Tango dance is ongoing in the author's group.

VI. CONCLUSION

Based on many dimensional subject-specific acceleration time series of walking derived from wearable sensors found in the MAREA database, we develop computational methodologies for the structural representation of cycles of biomechanical states and for mimicking personal gait dynamics to facilitate very precise detections for minute structural changes on the personal walking activity as well as on person-vs-person activity comparison. This capability lays the fundamental step toward precision learning of human walking activity. Recently, this step is further expanded in a series of ongoing precision learning projects in our research group. We mention two projects here. One project involves two very experienced tango dancers: the reasonably good tango dancer tries to learn and catch up with a very good tango dancer who is also a professional dancer by training. Another project involves two student-athletes in track-and-field running: the good runner tries to improve and catch up with the better runner. As far as our research shows us, we envision that such a dynamics-mimicking based precision learning will work for a wide spectrum of human activities in the not-distant future.

Further, we are confident that the concept and technique of our computational developments on the multiscale structural representation of personal biomechanical gait cycles could have significant merits in scientific areas beyond human activities. This confidence is indeed emboldened by the results of an application of similar computational developments to bring out characteristics of a person's heartbeat dynamics. In our recent study on computational cardiology, based on multiple heartbeat time series recorded by multiple channels, once again the structural representation computations effectively show very promising deterministic and stochastic structures regarding the periodic R-R interval evolution, which is a critical feature of any heartbeat dynamics. This experience makes us to further believe that this concept and technique could be proved very effective in dealing with very general rhythmic dynamics.

Furthermore, the concept of mimicking observed data can be traced back to Kolmogorov complexity: the shortest universal computer program to regenerate the observed data. While this complexity is uncomputable, mimicking is practical and applicable in all data analyses, see details in [19]. The nature of the concept of mimicking is also very distinct from bootstrapping in statistics since it requires analysts to recognize the deterministic and stochastic structures embedded within data. This requirement needs serious, but doable efforts to achieve. In the setting of rhythmic dynamics, the structural representation of rhythmic cycles provides the platform for fulfilling this requirement.

One chief merit of mimicking gait dynamics, as mentioned in the Introduction section, is to provide reliability and confidence evaluations with nearly complete certainty by simulating as many mimicries as we can afford. The resultant characteristic manifolds allow us to build a mimicking-based precision learning protocol that can detect and confirm minute structural changes with temporal persistence lasting only for a few cycles in a row. Once again, we reiterate that this capability is a proof of concept toward precision learning of any human activities, in which achieving perfection and efficiency are critical. It is worth reemphasizing that many key and critical patterns of human activities are of 10 ms or finer scales. These patterns are beyond the raw eyesight of persons of interest and their coaches. Thus, our precision learning protocol can be expected to outperform the old-fashioned self-learning by practicing as well as video annotation and analysis specialized and popularized nowadays in sports technology aiming at improving sports performance with efficiency and accuracy. This is one unique perspective to view the potential of precision learning based on wearable sensors.

Though our computational developments of structural representation of gait dynamics are illustrated step-by-step through analysis of only one subject's time series data, the results of structural representations of the remaining 8 subjects contained in the same database are also obtained and show almost universal cyclic evolutions of biomechanical states of gait dynamics. We project this universal cyclic dynamics to be valid for all healthy humans' gait dynamics. With such universality of structural representation of gait dynamics, we project our mimicking protocol developed coherently based on 6-dim gait time series data from the [LF-vs-RF] subsystem to be valid for all healthy humans' gait dynamics as well. Thus, the merits of such a scientific precision learning protocol are far wider and beyond machine learning based gait identification and authentication.

Lastly, though our precision learning protocol is tuned to achieve efficiency and even excellence in performing human activities throughout this paper, we also confidently envision that our computational precision learning protocol still can offer a very wide spectrum of applications in the medical field. That is, its merits in medical precision care are also clear and evident. Since this protocol can be geared to help detect minor gesture changes in individuals' walking and other activities as well as transient structural changes in vital signs and beyond, such as physiological and circadian signals, to serve as early warning signals of certain diseases [2].

ACKNOWLEDGMENT

The authors would like to thank Dr. Siddhartha Khandelwal and Prof. Nicholas Wickström for permitting the re-use of Fig. 1.

REFERENCES

[1] S. Sprager and M. B. Juric, "Inertial sensor-based gait recognition: A review," *Sensors*, vol. 15, no. 9, pp. 22089–220127, Sep. 2015, doi: 10.3390/s150922089.

[2] D. A. Winter, "Biomechanics as an interdiscipline," in *Biomechanics and Motor Control of Human Gait: Normal, Elderly and Pathological*, 2nd ed. Waterloo, ON, Canada: UWaterloo Press, 1991, pp. 1–12.

[3] W. M. Whittle, "Normal gait," in *Gait Analysis: An Introduction*, 4th ed. Edinburgh, U.K.: Butterworth, 2008, pp. 48–90.

[4] H. J. Ailisto, M. Lindholm, J. Mantyjarvi, E. Vildjiounaite, and S.-M. Makela, "Identifying people from gait pattern with accelerometers," *Proc. SPIE*, vol. 5779, pp. 7–14, Apr. 2005, doi: 10.1117/12.603331.

[5] G. Trivino, A. Alvarez-Alvarez, and G. Bailador, "Application of the computational theory of perceptions to human gait pattern recognition," *Pattern Recognit.*, vol. 43, no. 7, pp. 2572–2581, Jul. 2010, doi: 10.1016/j.patcog.2010.01.017.

[6] T. T. Ngo, Y. Makihara, H. Nagahara, Y. Mukaigawa, and Y. Yagi, "The largest inertial sensor-based gait database and performance evaluation of gait-based personal authentication," *Pattern Recognit.*, vol. 47, no. 1, pp. 228–237, 2014, doi: 10.1016/j.patcog.2013.06.028.

[7] D. Gafurov, K. Helkala, and T. Sondrol, "Biometric gait authentication using accelerometer sensor," *J. Comput.*, vol. 1, no. 7, pp. 51–59, Nov. 2006, doi: 10.4304/jcp.1.7.51-59.

[8] D. T. H. Lai, R. K. Begg, and M. Palaniswami, "Computational intelligence in gait research: A perspective on current applications and future challenges," *IEEE Trans. Inf. Technol. Biomed.*, vol. 13, no. 5, pp. 687–702, May 2009, doi: 10.1109/TITB.2009.2022913.

[9] S. Khandelwal and N. Wickström, "Evaluation of the performance of accelerometer-based gait event detection algorithms in different real-world scenarios using the MAREA gait database," *Gait Posture*, vol. 51, pp. 84–90, Jan. 2017, doi: 10.1016/j.gaitpost.2016.09.023.

[10] F. Hsieh and X. Wang, "From learning gait signatures of many individuals to reconstructing gait dynamics of one single individual," *Frontiers Appl. Math. Statist.*, vol. 6, Nov. 2020, Art. no. 564935, doi: 10.3389/fams.2020.564935.

[11] R. Chereshnev and A. Kértesz-Farkas, "HuGaDB: Human gait database for activity recognition from wearable inertial sensor networks," in *Proc. AIST*, May 2017, pp. 131–141, doi: 10.1007/978-3-319-73013-4_12.

[12] F. Kaspar and H. G. Schuster, "Easily calculable measure for the complexity of spatiotemporal patterns," *Phys. Rev. A, Gen. Phys.*, vol. 36, no. 2, pp. 842–848, Jul. 1987, doi: 10.1103/PhysRevA.36.842.

[13] H. Hatze, "The meaning of the term biomechanics," *J. Biomech.*, vol. 7, no. 2, pp. 189–190, Mar. 1974, doi: 10.1016/0021-9290(74)90060-8.

[14] R. Bartlett, "Movement patterns—The essence of sports biomechanics," in *Introduction to Sports Biomechanics*. New York, NY, USA: Routledge, 1997, pp. 1–58.

[15] M. Gell-Mann, "What is complexity?" *Complexity*, vol. 1, no. 1, pp. 16–19, Sep. 1995, doi: 10.1002/cplx.6130010105.

[16] C. Adami, "What is complexity?" *BioEssays*, vol. 24, no. 12, pp. 1085–1094, Nov. 2002, doi: 10.1002/bies.10192.

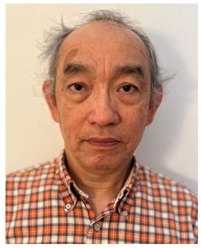
[17] P. W. Anderson, "More is different: Broken symmetry and the nature of the hierarchical structure of science," *Science*, vol. 177, no. 4047, pp. 393–396, Aug. 1972, doi: 10.1126/science.177.4047.393.

[18] J. P. Crutchfield, "Between order and chaos," *Nature Phys.*, vol. 8, no. 1, pp. 17–24, Jan. 2012, doi: 10.1038/nphys2190.

[19] T.-L. Chen, E. P. Chou, and H. Fushing, "Categorical nature of major factor selection via information theoretic measurements," *Entropy*, vol. 23, no. 12, p. 1684, Dec. 2021.

XI YANG received the B.S. degree in biological sciences from Zhejiang University, Zhejiang, China, in 2016, and the M.S. degree in statistics from Rutgers University, The State University of New Jersey, New Brunswick, NJ, USA, in 2018. She is currently pursuing the Ph.D. degree in biostatistics with the University of California at Davis, Davis, CA, USA.

Her research interests include data visualization, gait dynamics, machine learning, and statistical modeling.



FUSHING HSIEH received the Ph.D. degree in statistics from Cornell University, Ithaca, NY, USA, in 1990.

He is currently a Professor in statistics with the University of California at Davis, Davis, CA, USA. He develops conditional entropy and mutual information based on categorical exploratory data analysis (CEDA) and major factor selection as foundations for analyzing data from complex systems with or without rhythm.