frontiers in Applied Mathematics and Statistics

ORIGINAL RESEARCH
published: 12 November 2020
doi: 10.3389/fams.2020.564935

From Learning Gait Signatures of Many Individuals to Reconstructing Gait Dynamics of One Single Individual

Fushing Hsieh and Xiaodong Wang*

Department of Statistics, University of California, Davis, CA, United States

Based on the same databases, we computationally address two seemingly highly related, in fact drastically distinct, questions via computational data-driven algorithms: 1) how to precisely achieve the big task of differentiating gait signatures of many individuals and 2) how to reconstruct an individual's complex gait dynamics in full. Our brains can "effortlessly" resolve the first question, but will definitely fail in the second one because many fine temporal scale gait patterns surely escape our eyes. Based on accelerometers' 3D gait time series databases, we link the answers toward both questions via multiscale structural dependency within gait dynamics of our musculoskeletal system. Two types of dependency manifestations are explored. We first develop simple algorithmic computing called Principle System-State Analysis for the coarse dependency in implicit forms. Principle System-State Analysis is shown to be able to efficiently classifying among many subjects. We then develop a multiscale Local-1st-Global-2nd Coding Algorithm and a landmark computing algorithm. With both algorithms, we can precisely dissect rhythmic gait cycles, and then decompose each cycle into a series of cyclic gait phases. With proper color-coding and stacking, we reconstruct and represent an individual's gait dynamics via a 3D cylinder to collectively reveal universal deterministic and stochastic structural patterns on centisecond (10 milliseconds) scale across all rhythmic cycles. This 3D cylinder can serve as "passtensor" for authentication purposes related to clinical diagnoses and cybersecurity.

Keywords: multidimensional time series, gait dynamics, algorithmic complexity, unsupervised learning, wearable sensors

OPEN ACCESS

Edited by:
Yajun Mei,
Georgia Institute of Technology, United States

Reviewed by:
Xin Guo,
Hong Kong Polytechnic University, Hong Kong
Effendi Dodi Arisandi,
National Institute of Aeronautics and Space of Indonesia, Indonesia

****Correspondence:***
Fushing Hsieh
fhsieh@ucdavis.edu

Specialty section:
This article was submitted to Mathematics of Computation and Data Science, a section of the journal Frontiers in Applied Mathematics and Statistics

Received: *22 May 2020*
Accepted: *24 August 2020*
Published: *12 November 2020*

Citation:
Hsieh F and Wang X (2020) From Learning Gait Signatures of Many Individuals to Reconstructing Gait Dynamics of One Single Individual. Front. Appl. Math. Stat. 6:564935. doi: 10.3389/fams.2020.564935

1. INTRODUCTION

It seems ordinary that we recognize our close friends and family members by their distinctive walking "styles," so-called signatures of gaits. With the complexity of neural and musculoskeletal systems in mind [1], the gait dynamics is not at all simple. Unlike high-speed camera, our eyes surely miss all gait patterns of fine temporal scales. So, this ability of ours is not at all ordinary. Although we humans are anatomically identical by sharing the same structural skeleton and muscle constructs, and any gait dynamics must obey the universal biomechanics governing our musculoskeletal system, what make up individual signatures of gaits as biometric traits is still not yet well understood.

Majority of gait related research works is in the category of modeling-based gait analyses. The whole gait dynamics is never the focus. Any model based on only a few characteristics of gait dynamics typically not only is prone to make mistakes, but also difficult to apply to large number of healthy people. For instance, many works mainly aim for either Parkinson disease predictions or risk

evaluations for the elderly [2–6]. Such top-down approaches are of limited use for surveillance, for example, because they do not embrace diverse spectra of gait characteristics. For instance, the fuzzy finite state machine [7] needs to incorporate expert opinions and judgments for specifying relevant states. Further transitions between states are governed by fuzzy logics [8].

Recently, data collecting technologies have drastically evolved with recent advances in microelectromechanical systems (MEMS), such as low-cost, lightweight, easy-to-use inertial measurement units (IMU), such as accelerometer and gyroscope sensors [9]. These sensors nowadays are integrated with mobile devices, which enable us to collect gait time series data outside of gait laboratory; see figures of human wearing sensors in Refs. 10, 11. However, the capacity of precisely differentiating many subjects' gait signatures and seeing a person's multiscale gait dynamics in full is not yet available in literature.

In this article, we develop computing and data-driven algorithms suitable for addressing two questions. 1) How to find and embrace large and diverse spectra of gait characteristics for identification purpose? 2) How to discover and recreate a person's gait dynamics in full?

The first theme of our data-driven developments is to compute and find many principle directions or vectors that implicitly capture many important aspects of above structural dependency-based heterogeneity across many people. We consider one manifestation of structural dependency through temporal patterns via a very simple and coarse coding scheme, called Principle System-State Analysis (PSSA). This dependency manifestation of coarse scale pattern is indeed very versatile for classifying among all subjects. We conjecture that this kind of dependency manifestation is potentially close to how our brains learn gait signatures.

As a complex system, the intelligence of musculoskeletal system is embraced by its multiscale heterogeneity [12]. It is well known that any real "rhythmic" biomechanics is far from being completely deterministic and it naturally embraces stochastic structures across all rhythmic cycles as well [13]. Here, it is worth emphasizing the evidently visible, but inexplicable stochasticity. Because this stochasticity is chiefly constrained by deterministic structures, it is not completely random. Therefore, extracting stochastic structures of gait dynamics is at least as equally important as extracting the deterministic counterparts.

For explicitly extracting such multiscale deterministic and stochastic information contents, we turn to and focus on the system's fundamental structural dependency among all observed gait time series. It is clear that such structural dependency is lost to a great degree in the so-called resultant acceleration signal [14, 15]:

$$A_{res}[t] = \sqrt{X^2[t] + Y^2[t] + Z^2[t]},$$

This fact is evident through our motivating Lempel-Ziv complexity experiments (see details in the next section). Results from such experiments imply how to build a symbolic coding scheme to retain structural dependency of multiple time series.

Based on such motivation, our second theme of data-driven computing paradigm is developed as an unsupervised learning-based multilayer coding scheme, called Local-first and Global-second (L1G2) coding scheme. We apply L1G2 to build a 2D code sequence pertaining to the [Left-foot + Right-foot] system. We also develop a landmark partition algorithm to dissect such a 2D code sequence into rhythmic cycles consisting of visible biomechanical states. Such rhythmic patterns confirm that this subsystem indeed dictates the contents of a rhythmic cycle, its period, and most importantly its evolving process. That is, the entire musculoskeletal system should function by coupling others subsystems upon [Left-foot + Right-foot] system.

To further show L1G2 effectively capturing multiscale gait dynamics, via graphic display, we simply stack all resultant color-coded rhythmic cycles aligned with the landmarks into a 3D cylinder. This rotatable 3D cylinder coherently reveals multiscale deterministic and stochastic rhythmic patterns as multiscale structural dependency across all rhythmic cycles. Such a 3D cylinder is the very foundation of further researches of gait-mimicking. It is also good for clinical diagnosis, and can be used as a "passtensor" for cybersecurity.

Two known gait time series databases are analyzed as the real data experiments: 1) MAREA database [10] with four sensors and 2) HuGaDB database [16] with six sensors. Both databases are created on healthy subject's gait when subjects wear with multiple sensors performing various activities on different kinds of surfaces. The sampling rate in MAREA is 128 Hz, and is less in HuGaDB. That is, the time series in these databases contains patterns of centisecond (10 minisecond) scale.

We focus only on accelerometer in this article. It picks up accelerations of linear motions of body parts, where the sensors are fixed, upon $X-$, $Y-$, and $Z-$axial orientations. The 3-dim measurements are referencing to the coordinate system of human body: anterior-posterior (forward vs. backward), superior-inferior (vertical up vs. down along gravity direction), and left-right [17]. Our developments can easily accommodate gyroscope-based time series. In MAREA database, each subject wore a 3-axes Shimmer3 (Shimmer Research, Dublin, Ireland) accelerometer (±8 g). In HUGA database, the information of accelerometer is described in the article [18].

2. REVELATIONS OF STRUCTURAL DEPENDENCY

To set the stage for our computational developments for exploring an individual's gait dynamics in full, we give an overview of the two contrasting manifestations of structural dependency contained in multidimensional gait time series. First by looking at an approximate 3 s recording of 12 dimensional time series of a MEARA subject's walking on indoor flat ground, as shown in **Figure 1**, we see that each sensor's triplet directional time series exhibit diverse scales of relational patterns, which evolve within each visible cycle, and recurrently appear across evident rhythmic cycles. Second, when

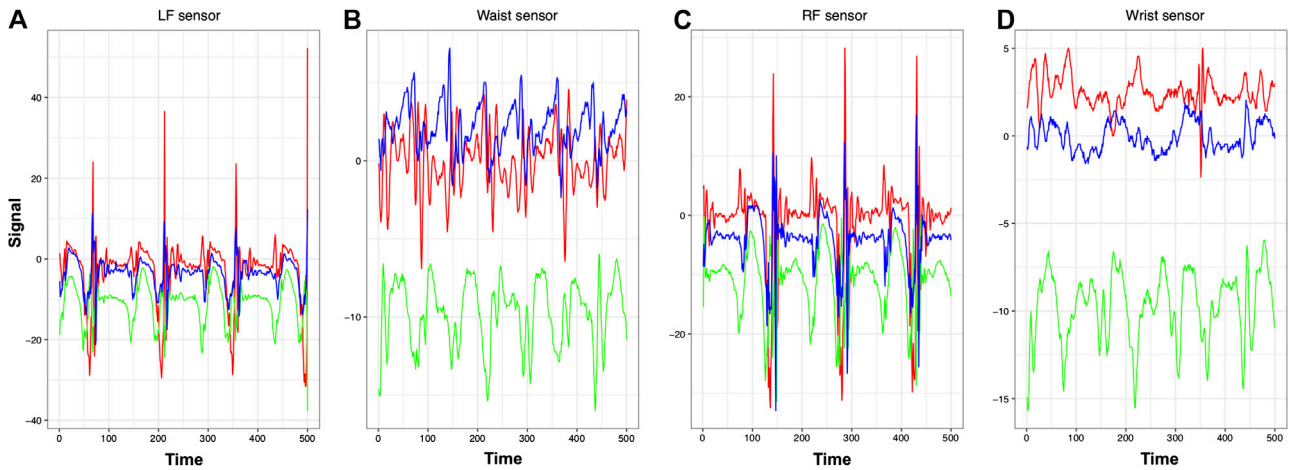

FIGURE 1 | Gait time series data of subject #5 from four sensors: **(A)** left foot, **(B)**waist, **(C)** right foot, and **(D)** wrist. X−dimension is red color-coded, Y−dimension is green, and Z−dimension is blue.

we compare such patterns across different sensors, we also discover various scales of recurrent pattern-to-pattern correspondences. Such pattern-to-pattern correspondences are especially evident between panel (A) of left-foot and panel (C) of right-foot of **Figure 1** across the evident cycles. Pattern-to-pattern correspondences between panel (B) of waist and either one of left-foot or right-foot are also apparent, but not between panel (D) of wrist with the rest of panels. These visible temporal-oriented relational patterns within cycles and complex pattern-to-pattern correspondences across cycles constitute multiscale structural dependency of gait dynamics contained in the 12 dimensional time series. This is the chief concern in this article.

In computational theory of computer science, the concept of Kolmogorov complexity is used in evaluating and exploring hidden structural patterns embraced within symbolic or digital time series. Its conceptual shortest universal computer program for regenerating a time series at hand is recognized to embrace all deterministic and stochastic structures. Unfortunately, Kolmogorov complexity cannot be calculated in general. We use Lempel-Ziv complexity to give an approximate measure by only using "copy" and "insert" two operations. This complexity can be efficiently computed, see Ref. 18. So, Lempel-Ziv is used in our complexity experiments. Before our complexity experiment, all the continuous time series must be categorized and transformed into a finite and discrete state sequence.

As shown in each panel of **Figure 1**, each triplet time series of (X, Y, Z) directions of an accelerometer reveal varying mechanism-specific gait dynamic patterns. Thus, we make use of this data transformation requirement to naturally link the concept of structural dependency among time series to system-states of its dynamics. The idea of system-state can be seen as follows. We develop two temposensitive digital-coding schemes upon gait time series along the temporal axis. The first scheme is to perform digital coding upon each of the triplet directional time series individually and then couple the three digital code sequences into one sequence of vectors. The second scheme is to apply hierarchical clustering algorithm on the temporal (column) axis of a data matrix representing the triplet time series with three rows. Based on the resultant clustering tree, a composition of clusters is chosen. A cluster of 3D vectors can be regarded as a symbolic code for a system state. Hence, the specific mechanism pertaining to an accelerometer along the temporal axis is represented by a 1D symbolic code sequence. Color-coded examples of such code sequences are given in **Figure 4**. The computing cost of the first approach is much less than that of second approach. But, unlike the second approach, the first approach can only capture relatively coarse structural dependency.

We compare these two coding schemes in a set of Lempel-Ziv complexity experiments based on a short temporal segment $[0, 300]$. Results of such experiments are summarized in **Figure 2**; also see **Supplementary Figures 10, 11** for more details. The top three panels of **Figure 2**, respectively, give the three directional symbolic code sequences. Each code sequence has three states and a value of Lempel-Ziv complexity. By coupling these three code sequences along the temporal axis, as shown in panel (D), the resultant code sequence with 27 state is seen nearly without any recognizable recurrent patterns. It has a complexity value 1,017. In comparison, the second scheme with 27 clusters results into code sequence, as shown in the panel (E), which shows very evident recurrent and rhythmic patterns with a complexity value 571. Furthermore, even if only 10 clusters are used to form the set of states, as seen in the bottom panel (F), the resultant code sequence is as evidently rhythmic as the one with 27 states in (E). With such rhythmic patterns in view, it is not surprising that its Lempel-Ziv complexity value is even lower. Evidently, it captures the rhythmic dynamics well. Such experimental results confirm the presence of structural dependency among the three directional gait time series, and at the same time imply that the second coding scheme is a way of extracting detailed dependency patterns in gait dynamics. Nonetheless, the first coding scheme has its own merit in identifying among many subjects as seen in the next section.

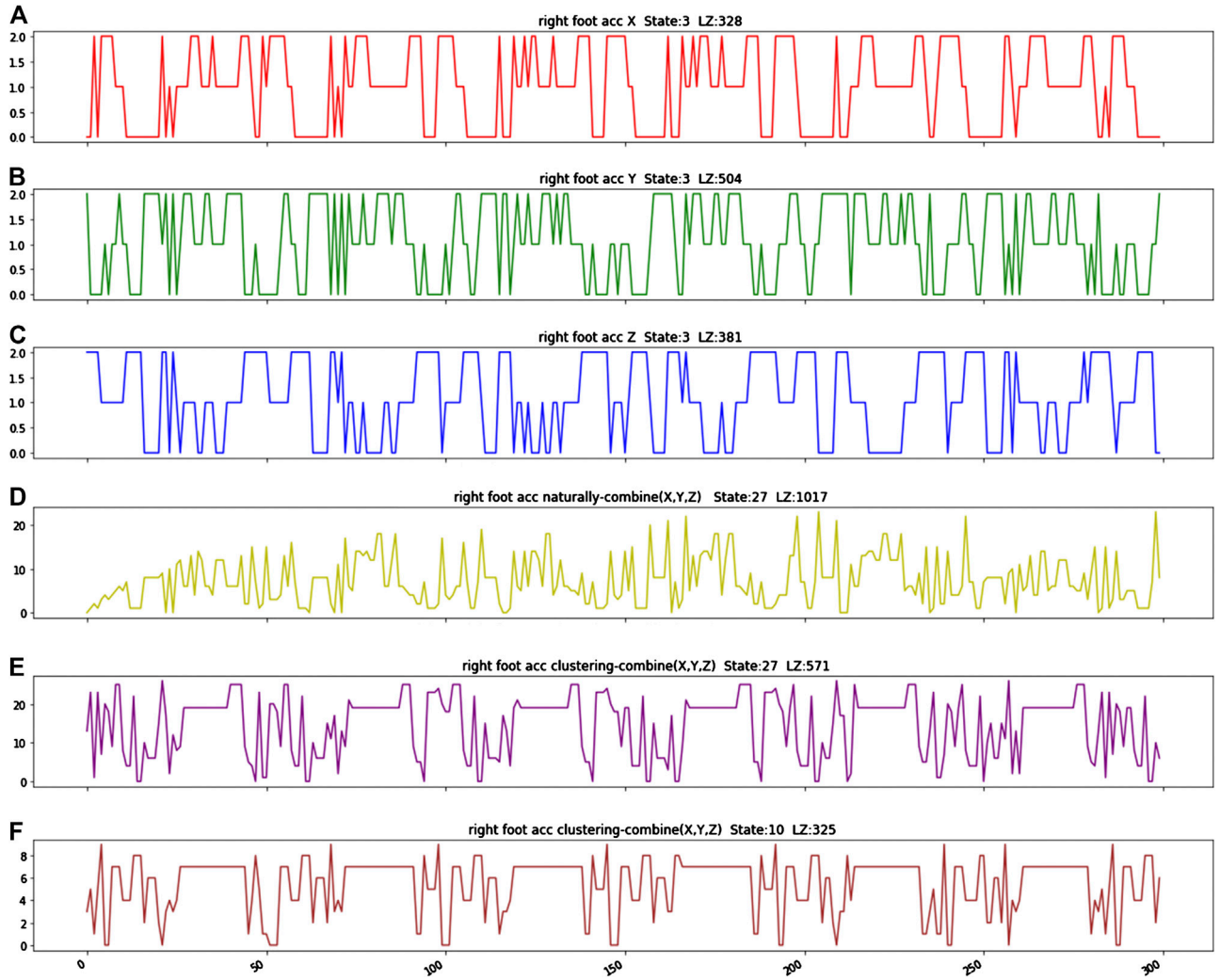

FIGURE 2 | (A–C) Three-state code sequences for X-,Y-,Z-accelerometer time series, respectively. **(D)** is a natural combination of X,Y,Z, and the resultant sequence is coded by 27 ($3 \times 3 \times 3$) states. **(E,F)** are sequences based on our clustering-based way of combination. **(E)** is coded by 27 states (clusters), the same number of states as **(D)**, whereas its LZ complexity reduces by half. **(F)** is a 10-states code sequence that can show the rhythmic pattern clear enough, and its LZ complexity is as low as that of one-dim time series case.

3. PRINCIPLE OF SYSTEM-STATE ANALYSIS (PSSA)

A simple way of having a glimpse of structural dependency among sensor-direction specific D dimensional gait time series is to transform and couple them into a D-dimensional digital vector trajectory. Here, D is equal to 12 for four sensors used in MAREA database and 18 in for six sensors used in HuGaDB database. This digital trajectory is to exhibit rough manifestations of rhythmic cycles. So, we manage to have a representation with relative small algorithmic complexity about the gait dynamics. This idea is simple and intuitive. Here, we develop data-driven computations via a coarse coding scheme to realize this concept. By doing so, we get away from the necessity of man-made system-states and requirements of their transition rules. The simple computational results are capable of identifying many subjects simultaneously on a single platform. Thus, we speculate such a simple algorithm is potentially what our brain actually performs in recognizing friends and relatives' gait signatures. To this aim, we propose an algorithm, called the Principle of System-State Analysis (PSSA), that attempts a single-layer coarse structural dependency among many individuals' D dimensional gait time series simultaneously.

3.1. PSSA Algorithm

For the purpose of identification, we expect to identify an individual by only glimpsing his/her short time of walking. Each individual's specific gait time series is subdivided into replicates of period in equal length l. We assume that in the test set, each unlabeled individual would have sample size exceeding l. The choice of l is supposed to be small while the signal is strong enough. Here, we set $l = 1,000$ time points, which

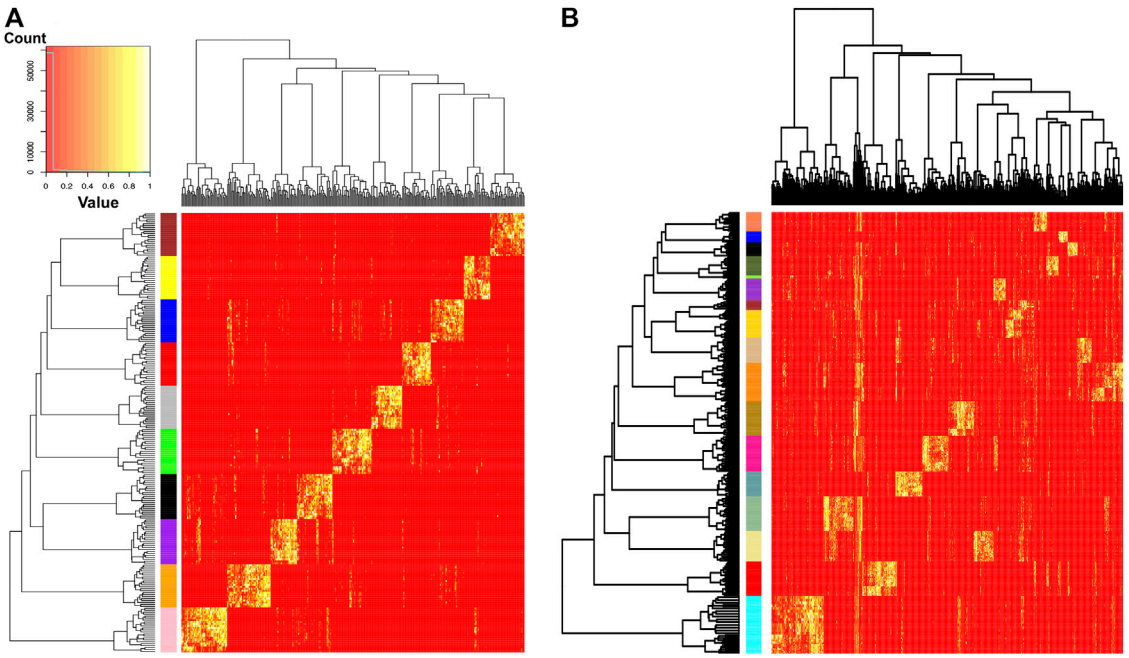

FIGURE 3 | Identification via heatmap of Σ_{PSS}. Each row indicates a segment of gait time, and rows from the same subject are labeled in the same color; each column indicates a selected PSS. **(A)** MAREA database: 10 subjects. The quantiles $\alpha = 0.3$ and $\beta = 0.7$. N^* (= 300) principle system-states based on nine dimensions of gait time series derived from three sensors fixed at the left foot, right foot, and wrist. **(B)** HuGaDB database: 17 subjects with six sensors tied to left and right thighs, shins, and feet. The quantiles $\alpha = 0.1$ and $\beta = 0.9$. N^* (= 500) principle system-states based on 18 dimensions of gait time series.

lasts about 7.8 s with respect to the sampling rate being set at 128 Hz. Consider that each individual at each time point has a D dimensional measurements (with the same unit m/s^2): three directional (x-, y-, z-) accelerations from each of accelerometer sensors. We stack such D dimensional vectors together across all individuals' time points into a large data matrix with 12 rows. After that, the PSSA algorithm is applied, which is described below.

First, encode each sensor-direction specific 1-dim time series by using 3-digit alphabets.

$$S_d(t) = \begin{cases} 1 & X_d(t) \le \alpha \\ 2 & \alpha < X_d(t) \le \beta, \\ 3 & X_d(t) > \beta \end{cases}$$

where $X(t)$ is the variable at time stamp t and $d = 1, 2, \ldots, D$ indicating dimension. So, a D-dimensional digital system-state (vector), say $S(t) = (S_1(t), \ldots, S_D(t))'$, is formed at each time point t. The tuning parameters α and β ($\alpha < \beta$) are chosen based on the quantile of each 1-dim empirical distribution of pooled data across all involving subjects. Based on the consideration that the extreme values of each distribution played an important role in identifying different subjects, we choose $\alpha < 0.5 < \beta$ and α and β are closer to their extremes 0 and 1, respectively. As a result, the complexity of resultant digital code time series becomes smaller.

Second, collect all distinct system-states $S(.)$ and calculate their corresponding frequency f. There will be at most 3^D possibilities. Sort the distinct states with respect to frequency from the most frequent to the least $(S^{(1)}(.), \ldots, S^{(N)}(.))'$ with highest frequency $f^{(1)}$ to the lowest one $f^{(N)}$. Select a set of N^* states with top highest frequency as principle system-states (PSS).

Third, cut the gait time series from the training set into short-temporal segments in length l, and convert each segment to a N^*-vector of proportion of PSS occurring within the period. That is to say, we extract N^* from each of the segment, which represents the frequency of the appearance of the principle system-states.

Finally, an $m \times N^*$ rectangle matrix Σ_{PSS} is built by stacking all involving proportion vectors along the row-axis, where m is the total number of segments, and N^* is the number of principle components. The entry (i, j) of Σ_{PSS} can be explained as the frequency of the j-th principle state found in the i-th segment. Apply hierarchical clustering analysis on row and column axes of Σ_{PSS}, respectively. Find the corresponding "key" PSS for each individual such that the PSS can be used as a new feature (group) to exclusively identify the individual from others.

PSSA achieves a huge reduction on temporal dimensionality from $l = 1000$ to N^*. More importantly, such an N^*-dim vector is in the category of structural data, that is, each component can be treated as a feature variable. So, any classic machine learning techniques can come in and work on the structured matrix Σ_{PSS}.

With a chosen pair of tuning parameters α and β ($\alpha < 0.5 < \beta$), the complexity of digital coded D-dim time series can be seen via the curve of proportion of coverage on all involving trajectories as:

$$r(N^*) = \sum_{i=1}^{N^*} \frac{f^{(i)}}{N},$$

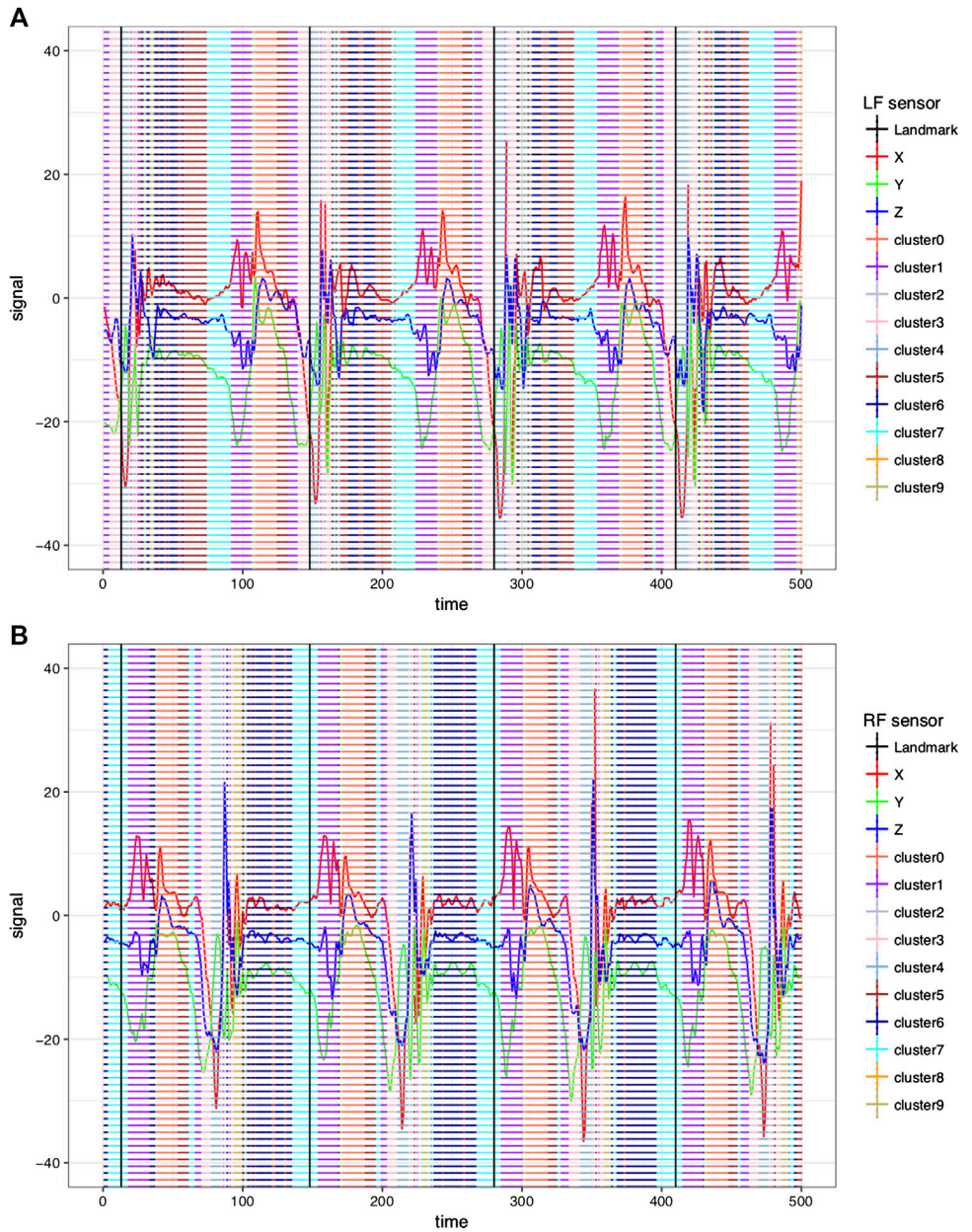

FIGURE 4 | 3D time series superimposed with color coding on temporal period [1, 500]. **(A)** Left-foot sensor and **(B)** right-foot sensor. Color coding of the 10 selected clusters are listed on the right hand side. The landmarks are calculated and marked with vertical black line.

The selection of N^* principle system-states $(S^{(1)}(.), \ldots, S^{(N^*)}(.))$ can be also based on this curve.

3.2. PSSA on Real Databases

Two examples of coverage proportion curves with respect to N^* principle system-states are given **Supplementary Figure 9** for MAREA database and HUGaDB database.

Both results in the training set are perfectly classified without any error among all 10 subjects' replicates in MAREA database, and 17 subjects' replicates in HuGaDB database, see **Figure 3**. By selecting one significant state's block or cluster for each individual, a simple decision tree can achieve perfect classification result in the test set. That is to say, the principle states take the shape of feature selection, and they are the keys in gait identification.

Here, we make a remark on how to scale a big ensemble of individuals via PSSA. When the ensemble of individuals is big in size, the PSSA needs a strategy to scale down the computing loading. That is, if such an ensemble is taken as being homogeneous, then PSSA will need a large collection of system-state vectors to cover enough complexity in identification task. Or the percentages α and β are chosen to

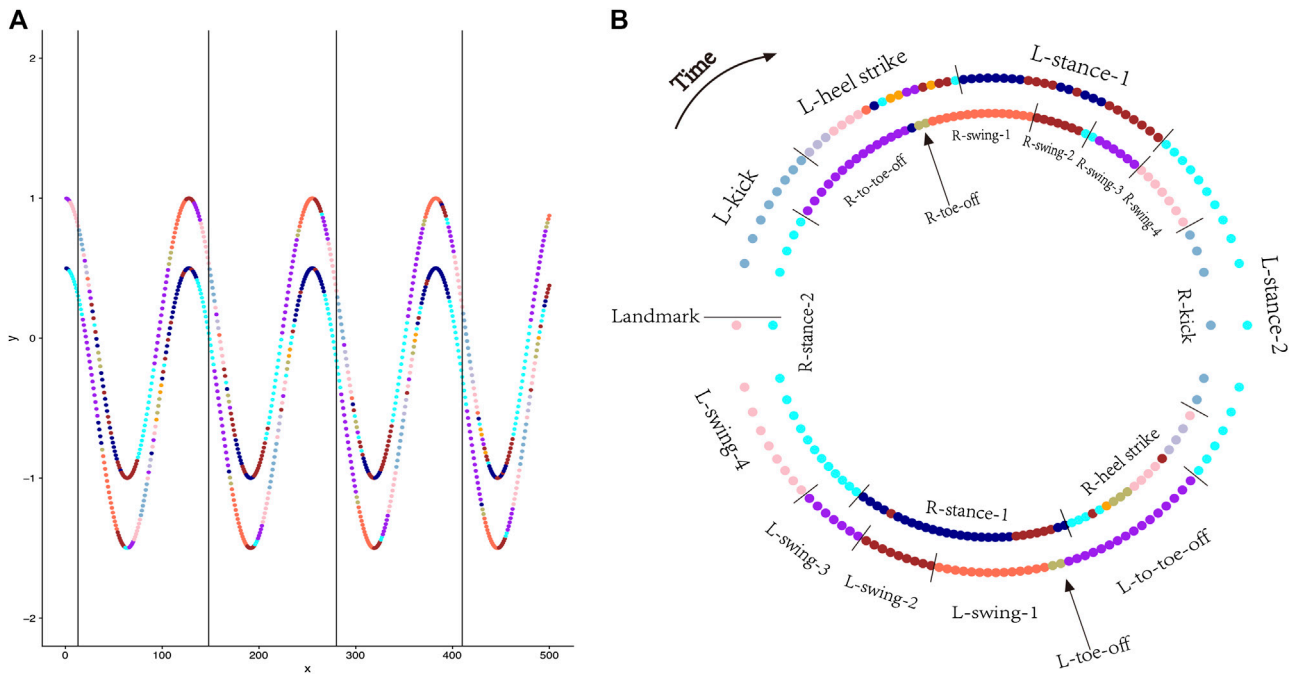

FIGURE 5 | Color-coded rhythmical cycles in L + R system of subject #5 marked with serial biomechanical phases. **(A)** The coupled color-coding time series on temporal period [1, 500] (upper curve for the left-foot, and lower curve for the right-foot). The landmarks are marked with vertical black lines. **(B)** Rhythmic cycle, the third one in panel **(A)**, is represented by two concentric rings (outer ring for the left-foot, and inner right for the right-foot). The temporal coordinates go clockwise.

be close their extremes. However, if heterogeneity is naturally present in any human ensemble, it implies the necessity of partitioning the whole ensemble into homogeneous subensembles, and then PSSA is applied respectively. This is a typical divide-and-conquer strategy. For instance, the database in Ref. 11 consists of more than 700 individuals. It is sensible to divide the whole ensemble with respect to available demographic information.

In summary, our PSSA algorithm apparently is able to identify a set of system-states as signatures for each individual subject via relatively easy computations, and then perfectly classify among these subjects. Such visible signatures are indeed between-subject characteristics in nature. Because the computing behind such signatures is so simple, it is postulated why our brain can capture such signatures seemingly with ease after lengthy observations.

4. AUTHENTICATION VIA STRUCTURAL DEPENDENCY

Here, if we agree that different sets of triplet time series from different sensors give rise to different aspects of gait dynamics pertaining to our musculoskeletal system, then to authentically recreate gait dynamics is equivalent to compute the multiscale structural dependency based on all available time series data.

Let the local scale refer to various body components of musculoskeletal system, such as left-foot, right-foot, waist, and wrist. Each component contributes a fixed series of nearly deterministic biomechanical phases. Each biomechanical phase involves with a specific type of stochasticity: either in lengths or compositional contents. It is worth noting that such stochastic structures are somehow constrained by deterministic structures.

Let the global scale refer to how different components of musculoskeletal system couple and work out gait dynamics. Due to their dual symmetry, we particularly focus on how left-foot relationally works with right-foot via an evolving process. The [Left-foot + Right-foot] subsystem is rather distinct from their relations to waist as the center of mass with the musculoskeletal system. That is, within the entire musculoskeletal system, the [Left-foot + Right-foot] system indeed functionally coordinates with different subsystems.

4.1. L1G2 and Landmark Partition Algorithms

We reiterate that left-foot and right-foot play dual roles, on one hand, and are comparable or even symmetric, on the other hand. Their two sets of triplet time series are highly associated. We denote the [Left-foot + Right-foot] as the L + R, for short. Thus, we will encode L + R system locally first, and then integrate L + R system with waist or wrist. That is, we make the L + R system a foundation to grow the integrated musculoskeletal system. For this integrative task, we develop a rather simple algorithm based "local-first and global-second (L1G2)" coding scheme in this section.

This L1G2 coding scheme is devised by first applying HC algorithm onto the stacked version of $X-$, $Y-$, and $Z-$directional time series from the left-foot and right-foot sensors to generate a clustering tree. Upon this tree, we pick a 10-cluster

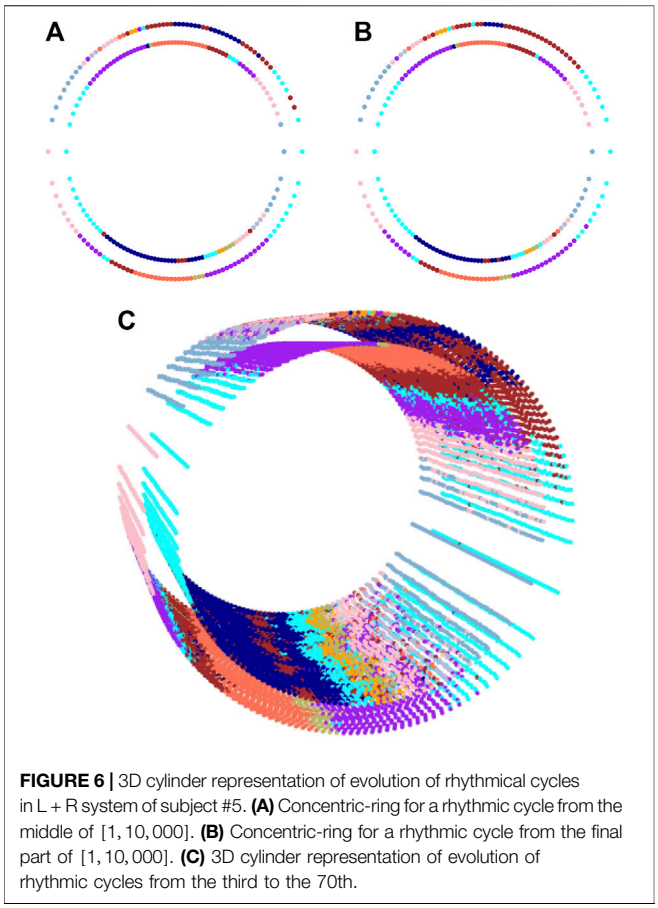

FIGURE 6 | 3D cylinder representation of evolution of rhythmical cycles in L + R system of subject #5. **(A)** Concentric-ring for a rhythmic cycle from the middle of $[1, 10, 000]$. **(B)** Concentric-ring for a rhythmic cycle from the final part of $[1, 10, 000]$. **(C)** 3D cylinder representation of evolution of rhythmic cycles from the third to the 70th.

composition to form a set of 10 codewords. Accordingly, left-foot's triplet time series are transformed into a 1D symbolic code sequence, so is the right-foot's. We then simply couple these two code sequences into a 2D L + R system-state trajectory. This choice of 10 codewords is supported by results of complexity evaluations in our Lempel-Ziv experiments in **Figure 2**.

Algorithm1: Local-first and Global-second (L1G2) coding

Input:
$\{(X_L(t), Y_L(t), Z_L(t)),\ 1 \le t \le T\}$ from Left-foot sensor
$\{(X_R(t), Y_R(t), Z_R(t)),\ 1 \le t \le T\}$ from Right-foot sensor
(1) Stack two time series and build a $3 \times 2T$ matrix
$\mathscr{M}_{L+R}[\cdot, 1:T] = \{(X_L(t), Y_L(t), Z_L(t)),\ 1 \le t \le T\}$
$\mathscr{M}_{L+R}[\cdot, (T+1):2T] = \{(X_R(t), Y_R(t), Z_R(t)),\ 1 \le t \le T\}$
(2) Apply HC on the temporal (column) axis of $\mathscr{M}_{L+R}$ to obtain H clusters, coded as $\{a_1, \ldots, a_H\}$, which represent local-system states
(3) Represent 3D time series $\{(X_L(t), Y_L(t), Z_L(t))$ and $\{(X_R(t), Y_R(t), Z_R(t))$ as 1D H-digital time sequence $\{S_L(t)\}$and $\{S_R(t)\}$, respectively
(4) Couple the two local system-state time series of left-foot and right-foot in a 2D L + R system-state time series with 2D vector $S_{L+R}(t) = (S_L(t), S_R(t))'$, for $t = 1, 2, \ldots, T$
(5) Integrate encoded waist and encoded L + R system by a 3D $(L + R) + W$ system-state time series with 3D vector $S_{(L+R)+W}(t) = (S_L(t), S_R(t), S_W(t))'$
Output: $S_{L+R}(t)$ and $S_{(L+R)+W}(t)$

Next, we develop a landmark algorithm to partition symbolic system-state trajectories into rhythmic cycles.

Algorithm 2: Landmark partition

Denote a Run_ibe a temporal segment that one specific state i consecutively repeats itself
Input: The 2D L + R system-state time series $\{S_{L+R}(t)\}$
(1) Calculate variance of the size of Run_i
(2) Calculate variance of the recurrence time of Run_i
(3) Choose the system-state i^* as a "landmark",
$i^* = argmin_i\{Var(\text{Size of } Run_i)$
$+ Var(\text{recurrence time of } Run_i)\}$
(4) Employ the landmark i^* to partition the entire system-state trajectory into pieces of rhythmic cycles
Output: A series of rhythmic cycles

Throughout our experimental explorations across many subjects, we found that rhythms in the L + R system are rather stable; although waist and wrist sensors' system-states are also rhythmic, their stability is weak. Further computed landmarks are found to coincide with the beginning of a system-state in L + R system, which is defined by a codeword pertaining to either left-foot or right-foot sensors, see **Figure 4**. This uncertainty is likely due to some degrees of asymmetry between left foot and right foot.

4.2. Color-Coded Rhythmic Cycles

We apply the L1G2 algorithm onto the L + R system of subject #5 on temporal period $[1, 10, 000]$. The local coding scheme is worked out on a stacked $3 \times 20,000$ matrix. The 10 codewords are color-coded, so that the identified system-states of L + R system are visible and readable with biomechanical meanings, as shown in **Figure 4**.

Each colored code sequence of left-foot and right-foot sensors, respectively, achieves a dimension reduction: from three to one. By coupling the two colored code sequences, as shown in **Figure 5A**, L1G2 algorithm results in cosine function-like rhythm under L + R system. The symmetry on both feet is also explicit. We then apply the landmark computing algorithm on such a 2D coupled colored code sequence on the temporal period $[1, 10, 000]$ to result 77 rhythmic cycles. The average period length and standard deviation are calculated as 127.56 ± 2.31.

To better visualize the progressing of system-state of L + R system via coupled colored codes, a rhythmic cycle is specifically represented by two concentric circles: Outer one for left-foot and inner one for right-foot, starting from the marked landmark located at the o'clock position, as shown in **Figure 5B**. Biomechanical phases on both feet are annotated. Indeed, the gait dynamics within a rhythmic cycle is evidently revealed with deterministic and stochastic structures as characterized as follows:

Deterministic Structures

A. The process of 2D coupling-phases as its state trajectory (with clockwise temporal coordinates) is nearly deterministic throughout all computed cycles:

Starting from "landmark" ⇒ (LF-Kick, RF-Stance2) ⇒ (LF-HeelStrike, RF-toToeOff) ⇒ (LF-HeelStrikeEnd, RF-ToeOff) ⇒ (LF-Stance1, RF-Swing1) ⇒ (LF-Stance1, RF-Swing2) ⇒ (LF-Stance1, RF-Swing3) ⇒ (LF-Stance2, RF-Swing4) ⇒ (LF-Stance2,

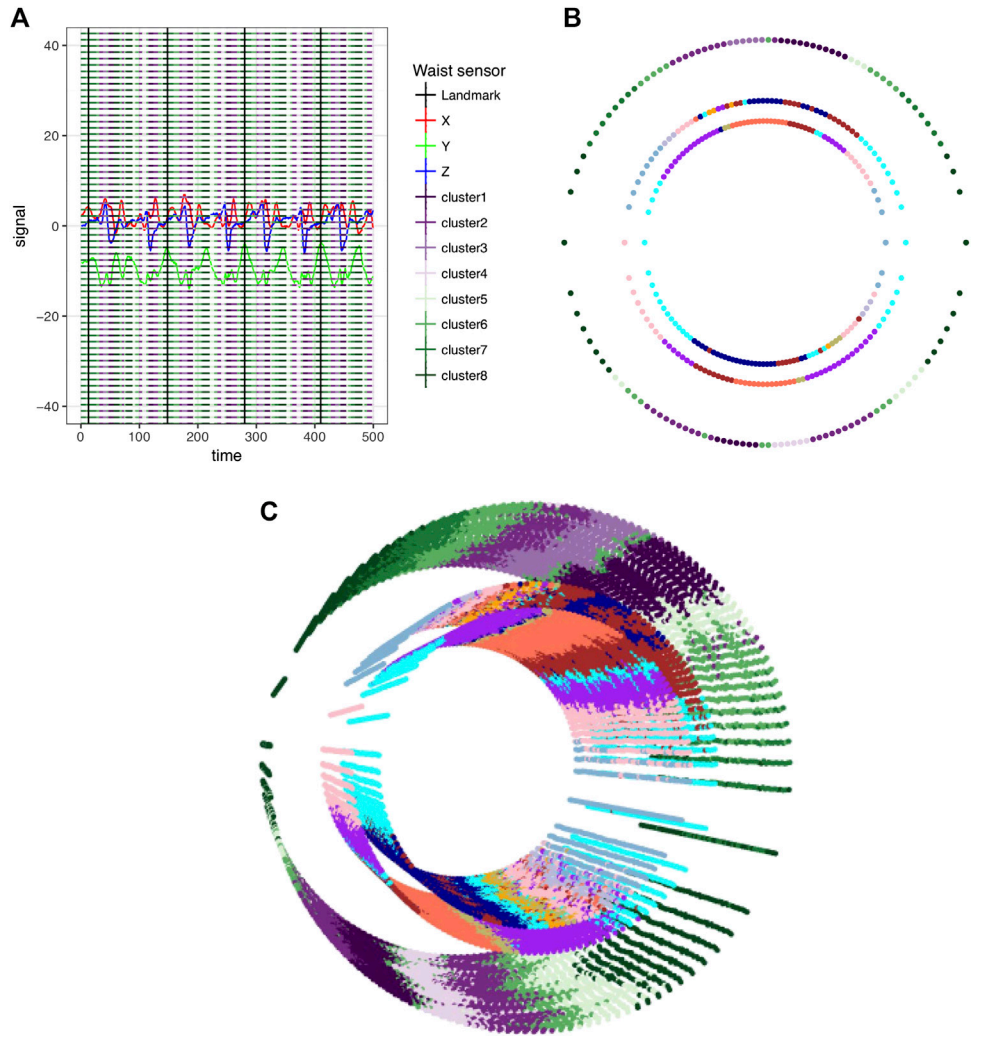

FIGURE 7 | Integrated gait dynamics of waist and L + R system. **(A)** Color-coded 3D time series from waist with eight clusters resulted from the local coding scheme of L1G2 algorithm. **(B)** Result of L1G2 algorithm represented by three layers of concentric-ring pertaining to the third rhythmic cycle on the temporal period $[1, 10, 000]$. **(C)** 3D cylinder representation of evolution of rhythmic cycles from the third to the 70th of this integrated system of three sensors.

RF-Kick) ⇒ (LF-ToeOff, RF-HealStrike) ⇒ (LF-ToeOff, RF-HeelStrikeEnd) ⇒ (LF-Swing1, RF-Stance1) ⇒ (LF-Swing2, RF-Stance1) ⇒ (LF-Swing2, RF-Stance1) ⇒ (LF-Swing3, RF-Stance2) ⇒ (LF-Swing4, RF-Stance2) ⇒ End at next "landmark";
B. A Toe-off phase of one foot has to happen after the end of heel-strike phase of the other foot;
C. The end of kick phase as the ending phase of swing process on one foot coincides with the beginning of "to-Toe-off" phase.

Stochastic Structures

A. Each 2D coupling-phase varies with lengths (seen through the 3D plot of rhythmic cycles from #3 to #70). This is the median-scale aspect of stochasticity within a rhythmic cycle;
B. The fine-scale stochasticity is seen in the phases of "heel-strike" of both left foot and right foot. The variations are far from being completely random;
C. There are some orders involving with a limited number of colored nodes. The large scale of stochasticity is seen via one or two distinct colored nodes being inserted between two phases specifically located at the two concentric circles;
D. There is also evident asymmetry on color coding of stance between the left foot and right foot.

5. GRAPHIC DISPLAY OF STRUCTURAL DEPENDENCY IN GAIT DYNAMICS

The explicit deterministic and stochastic structures in **Figure 5B** prescribe the structural dependency of gait dynamics in L + R system. Such a concentric-ring representation of a rhythmic cycle within L + R system is indeed very stable. Two more rhythmic cycles: one is from the middle and another one from the end of the temporal period $[1, 10, 000]$ among the 77 cycles, are rather similar, as shown in **Figure 6A,B**. The great degree of stability of gait dynamics

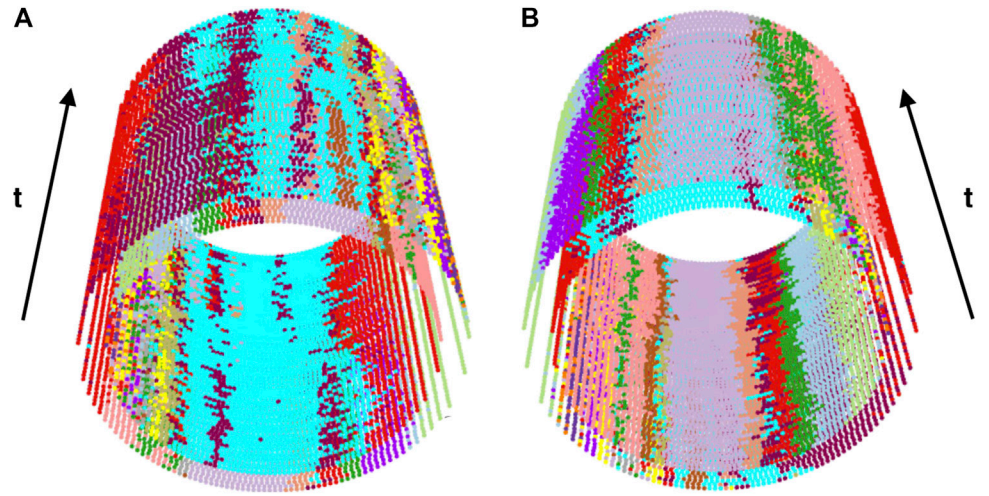

FIGURE 8 | Two angle-views of 3D passtensor constructed from subject #5's treadmill walking with slope changes in the middle of the temporal period in *t*. The slope changes cause very subtle change on **(A)**.

pertaining to the L + R system is also seen through a 3D cylinder representation in **Figure 6A**.

Such stability implies remarkable adaptability and precision of gait dynamics and its underlying structural dependency. The adaptability is primarily due to the interplay of deterministic and stochastic structures on the left foot and right foot. The deterministic structures give rise to a "typical" 2D coupling phase trajectory, whereas stochastic ones seemingly allow variations in lengths to happen among many components (or phases) of the typical cycle with total precision being about 36 ms (=:4,600/128). Such a precision is possible only when the deterministic structures are governed strictly by the biomechanics of human musculoskeletal system.

5.1. Integrating Waist Sensor Into L + R System

After constructing the rhythmic gait dynamics in L + R system, we then integrate it with the waist sensor. By applying the L1G2 algorithm on the 3D time series from waist sensor, the resultant local coding sequence is reported in **Figure 7A**, whereas the results derived from the global coding scheme is reported in **Figure 7B** for one rhythmic cycle with three layers of concentric circles. A 3D cylinder from 3rd to 70th rhythmic cycles is built and reported in **Figure 7C**. It is clear that 3D time series from waist sensor is rhythmic. But the rhythm is not symmetric with respect to dynamics in L + R system. Likewise, the wrist sensor can be integrated with L + R system as well.

5.2. Passtensors for Individual Authentications

The applications of coherently computed gait dynamics are rather wide and diverse. Here, we mention two essential ones in passing without going into details, and then focus on cybersecurity. The first comment is that this L1G2 algorithm will allow us to integrate acceleration sensors with gyroscope sensors. By combining the two kinds of sensors, the resultant gait dynamic system will be rather complex, but extremely interesting. The second comment is obvious that such a 3D representation can be utilized as a platform for mimicking the entire gait dynamics captured by time series data derived from the four acceleration sensors. Such a task of building realistic mimicry of a complex system is technically very challenging, although it is scientifically very important, for instance in robotics. Up to now, robots still walk in very unhuman-like fashions. This issue might be resolved to great extent by incorporating gait dynamics.

Now, we turn to cybersecurity, clinical diagnosis, and self-evaluating individual health statuses. It becomes clear that, based on our 3D graphic displays of gait dynamics, an individual's process of rhythmic cycle is characterized by the evolution of cyclic deterministic phases with individual specific twists as well as idiosyncratic stochastic deviations associated with all phases. Hence, a 3D cylinder indeed becomes a basis for authenticating this particular individual. For this use, such a 3D cylinder is called "passtensor." More specifically speaking, an L + R system's deterministic cycle of 2D biomechanical phases: from one landmark proceeding to the next one, indeed provides a rigid frame, whereas the stochastic phases' lengths and presence or absence of some color codes between adjacent phases provide the soft frames for the purposes of authentications. This authentication capacity further illustrated as follows. For instance, consider the subject #5 in MAREA walked on a treadmill with slope change: from horizontal (0°) to 5° during a recording period. This person's 3D passtensor corresponding to this period is shown in **Figure 8** with two views from two different angles. The angle-specific view in **Figure 8A** reveals visible changes. Such changes are likely critical patterns for authentication purposes.

Here, we briefly reiterate the practical uses of our 3D cylinder graphic display of gait dynamics in self-evaluating individual health statuses. By stacking two temporal segments of gait time series from two different temporal periods, we can examine the degrees and aspects of similarity and differences regarding deterministic and stochastic structures between these two temporal segments. This is

an effective way of finding out subtle and minute discrepancies to serve the early warning purposes.

6. CONCLUSION

6.1. Conclusion in System Complexity

Our first theme of data-driven computing paradigm, PSSA, allows us to include many principle gait states as a collective of key characteristics for identifying as many people as we want. From many aspects, this identification approach is indeed very distinct from identifications based on facial and voice recognitions, fingerprint, or retina scanning. It is much easier to achieve social unbiasedness. It is much more difficult to imitate or to fake.

Our second theme of data-driven computing paradigm, consisting of L1G2 coding and landmark algorithms, enables us to explicitly manifest multiscale dynamic patterns of gait dynamics. The graphic displays of single rhythmic cycle and collective 3D passtensor clearly demonstrate how the deterministic circle of biomechanical phase couples with stochastic variations sprinkling between consecutive phases, and offer a whole view of an individual's gait dynamics. Such intricate coupling relations between deterministic and stochastic structures are the backbones of structural dependency of gait dynamics. They retain essential basis for mimicking an individual's gait dynamics in animation. Its practical uses in clinical diagnosis and cybersecurity are also evident. In fact, the original motivations of this gait study is aiming at detecting relative minor changes in gait dynamics for healthy peoples and gesture tuning for athletes. These two topics require very detailed structures within personal dynamics.

From a computational science perspective, our PSSA and L1G2 coding algorithms rest on the crucial fact that different time series have different functions linking to different subsystems of a complex system of interest, so they should not be treated equally and uniformly. Such a rationale is a key for revelations of multiscale structural dependency. It is also the key rationale for recreating a system's authentic dynamics. Overall, good design of graphic displays definitely paves avenues for true understanding onto a complex system.

6.2. Conclusion in Security Issue

PSSA is purely developed for individual identification within a close community, such as a company or agency that needs a high degree of security, because the data are collected through multiple sensors placed on body parts. Hence, an individual's consent has to be in place first before data collection. Within a close community or company, PSSA is an effective alternative to facial recognition, because it does not suffer from problems due to shading on images or shadowing and cause social biases. And any individual outside of this community will be identified as outliers. Its application beyond a close community is still in a stage of theoretical research. In theory, it might be possible to convert a 3D video recording data into an accelerometer-based data format. But this technique is still not yet available. In fact, at the current state of technologies, any real-world recording via one camera, for example, CCTV, is unlikely to create an authentic 3D recording because of missing data.

For individual gait dynamics, our developments are geared to help individuals to do self-detections for minor gesture changes when walking or doing activities. Such analysis and results are highly personal. So, they are intended to be kept and used only by the owner of data. Our potential role would be limited to pointing out where minor changes might have taken place. Even this step is still under intensive researches.

DATA AVAILABILITY STATEMENT

All datasets presented in this study are included in the article/ **Supplementary Material**.

AUTHOR CONTRIBUTIONS

FH designed the study. XW preprocessed the data for analysis. FH and XW analyzed the data and interpreted the results. Both authors gave final approval for publication.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fams.2020.564935/full#supplementary-material

REFERENCES

1. Winter DA. *Biomechanics and motor Control of human gait: normal, Elderly and pathological.* Canada, ON: Waterloo (1991) 143 p.
2. Ailisto HJ, Lindholm M, Mantyjarvi J, Vildjiounaite E, Makela SM. Identifying people from gait pattern with accelerometers. *Proc SPIE* (2005) 5779:7–14. doi:10.1117/12.603331
3. Gafurov D, Kirsi H, Torkjel S. Biometric gait authentication using accelerometer sensor. *J Comput* (2006) 1:51–9. doi:10.4304/jcp.1.7.51-59
4. Lai DTH, Begg RK, Palaniswami M. Computational intelligence in gait research: a perspective on current applications and future challenges. *IEEE Trans Inf Technol Biomed* (2009) 13:687–702. doi:10.1109/titb.2009.2022913.
5. Trivino G, Alvarez-Alvarez A, Bailador G. Application of the computational theory of perceptions to human gait pattern recognition. *Pattern Recogn* (2010) 43:2572–81. doi:10.1016/j.patcog.2010.01.017.
6. Michael W. *WhittleGait analysis: an introduction*. 4th ed. Edinburgh, UK: Butterworth-Heinemann, Elsevier (2008) 192 p.
7. Alvarez-Alvarez A, Trivino G, Cordon O. Human gait modeling using a genetic fuzzy finite state machine. *IEEE Trans Fuzzy Syst* (2012) 20:205–23. doi:10.1109/tfuzz.2011.2171973.
8. Zadeh LA. Towards a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic. *Fuzzy Set Syst* (1997) 90:111–27. doi:10.1016/S0165-0114(97)00077-8
9. Sprager S, Juric M. Inertial sensor-based gait recognition: a review. *Sensors* (2015) 15:22089–127. doi:10.3390/s150922089.

10. Khandelwal S, Wickström N. Evaluation of the performance of accelerometer-based gait event detection algorithms in different real-world scenarios using the MAREA gait database. *Gait Posture* (2017) 51:84–90. doi:10.1016/j.gaitpost.2016.09.023.
11. Ngo TT, Makihara Y, Nagahara H, Mukaigawa Y, Yagi Y. The largest inertial sensor-based gait database and performance evaluation of gait-based personal authentication. *Pattern Recogn* (2014) 47:228–37. doi:10.1016/j.patcog.2013.06.028.
12. Anderson PW. More is different. *Science* (1972) 177:393–6. doi:10.1126/science.177.4047.393.
13. James P. Crutchfield between order and chaos. *Nat Phys* (2012) 8:17–24.
14. Karantonis DM, Narayanan MR, Mathie M, Lovell NH, Celler BG. Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring. *IEEE Trans Inf Technol Biomed* (2006) 10:156–67. doi:10.1109/titb.2005.856864.
15. Sant'Anna A, Wickström N. Developing a motion language: gait analysis from accelerometer sensor systems. In: 3rd international conference on pervasive computing technologies for healthcare; 2009 Apr 1–3; London, UK (2009) p. 1–8.
16. Chereshnev R, Kertesz-Farkas A. HuGaDB: human gait database for activity recognition from wearable inertial sensor networks. In: W. Van der Aalst, et al., editors. Analysis of images, social networks and texts; 2018 Jul 57; Moscow, Russia. Switzerland, AG: Springer, Cham (2017) *Lecture notes in computer science*, 10716.
17. Gietzelt M, Schnabel S, Wolf K-H, Büsching F, Song B, Rust S, et al. A method to align the coordinate system of accelerometers to the axes of a human body: the depitch algorithm. *Comput Methods Progr Biomed* (2012) 106:97–103. doi:10.1016/j.cmpb.2011.10.014.
18. Kaspar F, Schuster HG. Easily calculable measure for the complexity of spatiotemporal patterns. *Phys Rev A* (1987) 36(2):842–8. doi:10.1103/physreva.36.842.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.