PLoS COMPUTATIONAL BIOLOGY

Rhythmic Dynamics and Synchronization via Dimensionality Reduction: Application to Human Gait

Jie Zhang[1,2]*, Kai Zhang[3], Jianfeng Feng[1,4], Michael Small[2]

1 Center for Computational Systems Biology, Fudan University, Shanghai, People's Republic of China, 2 Department of Electronic and Information Engineering, Hong Kong Polytechnic University, Hong Kong, People's Republic of China, 3 Life Science Division, Lawrence Berkeley National Laboratory, Berkeley, California, United States of America, 4 Department of Computer Science and Mathematics, Warwick University, Coventry, United Kingdom

Abstract

Reliable characterization of locomotor dynamics of human walking is vital to understanding the neuromuscular control of human locomotion and disease diagnosis. However, the inherent oscillation and ubiquity of noise in such non-strictly periodic signals pose great challenges to current methodologies. To this end, we exploit the state-of-the-art technology in pattern recognition and, specifically, dimensionality reduction techniques, and propose to reconstruct and characterize the dynamics accurately on the cycle scale of the signal. This is achieved by deriving a low-dimensional representation of the cycles through global optimization, which effectively preserves the topology of the cycles that are embedded in a high-dimensional Euclidian space. Our approach demonstrates a clear advantage in capturing the intrinsic dynamics and probing the subtle synchronization patterns from uni/bivariate oscillatory signals over traditional methods. Application to human gait data for healthy subjects and diabetics reveals a significant difference in the dynamics of ankle movements and ankle-knee coordination, but not in knee movements. These results indicate that the impaired sensory feedback from the feet due to diabetes does not influence the knee movement in general, and that normal human walking is not critically dependent on the feedback from the peripheral nervous system.

Introduction

Complex physiological rhythms and synchronization processes are ubiquitous in biological systems and are fundamental to life [1]. The human heartbeat [2,3], walking [4], vocal cords vibration [5], blood pressure and respiration [6], white blood-cell count and tremor in patients [7,8], epidemic dynamics [9] all demonstrate a stable, possibly nonlinear, oscillatory pattern along with highly irregular fluctuations from period to period. Such signals are variously known as semiperiodic, approximately periodic or pseudoperiodic time series. The fluctuation overlying the oscillatory pattern, or specifically, the cycle-to-cycle variability, arises from the combined effects from the changing environment, the nonlinear nature inherent to biological systems, and noise of various sources. It contains a wealth of information regarding the health or disease status of an individual subject. Usually, little or no *a priori* knowledge or models that govern the underlying system are available. Therefore accurately characterizing and quantifying such biological rhythms through data-driven approaches contributes significantly to our understanding of complex biological control systems [10] and have important applications in disease diagnosis.

Traditionally, rhythmic signals are fruitfully analyzed by linear methods like the Fourier transform and power spectrum analysis. However, physiological signals as outputs of complex biological systems are typically nonlinear and non-stationary, and can not be properly characterized by linear methods. A number of new techniques based on nonlinear dynamical system theory [11,12] have also been intensively applied, like correlation dimension [13] and Lyapunov exponents [14]. Although the chaotic measures may provide new insights into the nonlinear nature of the system, they are severely hampered by the cyclic trend and noise in the system [15,16]. Recent attempts include producing pseudoperiodic surrogate data [17], or performing a transformation from time domain to dual complex network domain [18–24]. Generally, there still lacks systematic and robust approaches to handle such oscillatory, possibly nonlinear time series. The Fourier analysis decomposes the signal into harmonics that span over the entire time-line, thus information about how each period, or cycle, changes over time has been averaged out. Similarly, nonlinear measures are also based on averaged properties of phase space attractor reconstructed from the data. This lack of discrimination among the individual cycles calls for more advanced signal processing techniques. Inspired by the recent advances in the field of dimensionality reduction [25,26], we propose a novel and robust approach that can effectively capture the dynamics of cycle-to-cycle variation, which is especially suitable for analyzing approximately periodic, or semiperiodic data like human gait, ECG, respiration, and vowel data.

Human walking is a highly complex, rhythmic process which was found to exhibit long-range correlation and self-similarity, and has attracted sustained interest over the past decades [4,27–33]. The fluctuations overlying the cyclic trend in human walking may

Author Summary

Complex physiological rhythms arise from a large variety of biological systems that include natural pacemakers as well as feedback mechanisms, from the heartbeat to the rhythmic movement of human walking. Accurately extracting and characterizing the fluctuations underlying the biological rhythms is a fundamental problem which holds the key to understanding the mechanisms that govern the dynamics of biological systems. Usually such signals demonstrate certain oscillatory patterns, with each period displaying irregular fluctuation, or nontrivial dynamics, over time. This renders traditional spectral methods and nonlinear techniques less effective. We propose a novel approach to highlight the intrinsic fluctuations masked by the periodic component and noise through advanced dimension-reduction techniques, and apply it to human gait data from healthy subjects and diabetics. We find that this approach is capable of extracting the intrinsic dynamics and identifying the subtle synchronization pattern between knee and ankle. We find that although the two groups of individuals demonstrate remarkable differences in the dynamics of ankle movement and ankle-knee synchronization, the knee movement of both groups show similar dynamics. These results suggest that sensory feedback from a peripheral nerve system (like the feet) does not play an important role in regulating the motor control of human walking.

reflect valuable information about the neuromuscular processes responsible for normal and pathological locomotor patterns. In particular, the stride interval (SI) (the duration of each gait cycle) has been intensively studied to quantify the physiological or pathological state associated with walking. For example, it has been found that the long range correlation properties are altered with aging and disease [10,28]. The stride interval reflects the duration of each cycle, and a wealth of information contained in the waveform of the gait cycle is lost. Our approach utilizes the full waveform of each gait cycle, and is therefore expected to extract more information relevant to motor control of walking. We apply our method to the locomotion data collected from two groups of people — healthy subjects and neuropathic patients suffered from diabetes. We aim to find whether the extracted dynamical fluctuation of the knee and ankle movements as well as their synchronization pattern can vary between the healthy and diabetics group. Specifically, we want to find out whether the impaired sensory feedback due to diabetes can lead to different locomotion dynamics of the knee and ankle (and their synchronization pattern) compared with the healthy subjects.

Methods

Reconstructing Dynamics Underlying Cyclic Trend: Dimensionality Reduction

The general problem of dimension reduction has a long history. With advances in data collection, dimension reduction has reemerged as a prominent tool to unravel the high dimensional structure emerging in various disciplines. For example, it has been widely applied to gene and protein expression profiling for disease classification and prognostication [25,26]. Generally, the large number of dimension reduction approaches can be categorized into linear methods, including the principle component analysis, multidimensional scaling [34], and nonlinear methods such as state-of-the-art Isomap [34], laplacian eigenmaps [35] and local linear embedding [34]. Usually biomedical data process nonlinear

structures and that nonlinear dimensionality reduction methods might be more appropriate [26]. Here we use *Laplacian Eigenmaps* [35], which are based on *spectral graph theory* and projects the high-dimensional data into a low dimension so that two points nearby on the manifold are kept near to each other. We first illustrate with benchmark data from the chaotic Rössler system described by:

$$\begin{cases} X' &= -(Y+Z) \\ Y' &= X + 0.398\,Y \\ Z' &= 2 + Z(X-4) \end{cases} \quad (1)$$

The time series from the X-component (see Figure 1A), exhibits a strong periodic component along with irregular fluctuations, therefore it serves as an ideal example of an approximately periodic signal with nontrivial dynamics. Motivated by the fact that such data usually exhibit a highly redundant pattern in the form of repeated cycles, we can partition the time series into individual cycles C_i ($i = 1, \ldots, k$, which is the index of the cycle) at the peaks or troughs in the time series, see Figure 1A [36]. Each individual cycle can then be taken as a high dimensional vector x_i ($i = 1, \ldots, k$), whose dimension equals the number of points in that cycle. Our goal is to map these multiple, high-dimensional cycles to a set of new, low-dimensional (preferably one dimensional) representation, or embedding y_i's, such that the proximity relations among x_i's are maximally preserved in their low-dimensional counterparts y_i's. In the case that each cycle x_i is reduced to 1D (i.e., y_i being a scalar), the derived y_i constitute a new time series $(y_1, y_2, \ldots, y_k)$, which encodes the dynamics the original time series on the cycle scale.

To achieve this, a weighted matrix W is constructed, with each entry W_{ij} denoting the similarity between cycle x_i and x_j, which can be chosen conveniently as the correlation coefficient $\rho_{ij} = Cov(x_i, x_j)/(\sigma_{x_i}\sigma_{x_j})$ (under the circumstance that x_i and x_j differ in length, we shift the shorter vector along the longer one until ρ_{ij} maximizes). Then, the low-dimensional representation y_i's can be cast as the solution of the following optimization problem, $\min \sum W_{ij} \|y_i - y_j\|^2$, which penalizes those mappings where nearby points x_i's are relocated far apart in the space of y_i's. In case of univariate y_i's, the objective can be written as $\mathbf{y}^T L \mathbf{y}$, where $\mathbf{y} = [y_1, y_2, \ldots, y_k]^T$, $L = D - W$ is the *graph Laplacian*, and D is the diagonal degree matrix such that $D_{ii} = \sum_j W_{ji}$.

The above constrained minimization is solved by the generalized eigenvalue problem $L\mathbf{y}_i = \lambda_i D\mathbf{y}_i$, where λ_i's ($i = 1, 2, \ldots, k$) are eigenvalues sorted in an ascending order, and $\mathbf{y}_i$'s are the corresponding eigenvectors. The minimum eigenvalue λ_1 is zero, corresponding to an eigenvector ($\mathbf{y}_1$) whose entries are all 1. Therefore it is degenerate and the optimal solution is actually provided by $\mathbf{y}_2$, the eigenvector of the second smallest eigenvalue [35]. As is shown in Figure 1B, the eigenvector $\mathbf{y}_2(i)$ provides a cycle-scale representation of the original time series by reducing each full cycle to a single point. We use a general notation $c(i)$ (c for "cycle", and i indicates the ith cycle) for this simplified representation of the original time series. It is worthwhile to note that other dimension reduction schemes can also be adopted. For example, we can compute the Euclidian distance among cycles [36], and multidimensional scaling can be readily applied in this case to reduce the cycles to scalars. The $c(i)$ derived from multidimensional scaling and the Laplacian Eigenmap in fact yield quite similar results, see Figure 2.

The applicability of dimension reduction techniques is generally justifiable, considering the low correlation dimension of most real world pseudoperiodic data. For this kind of data, the trajectories of
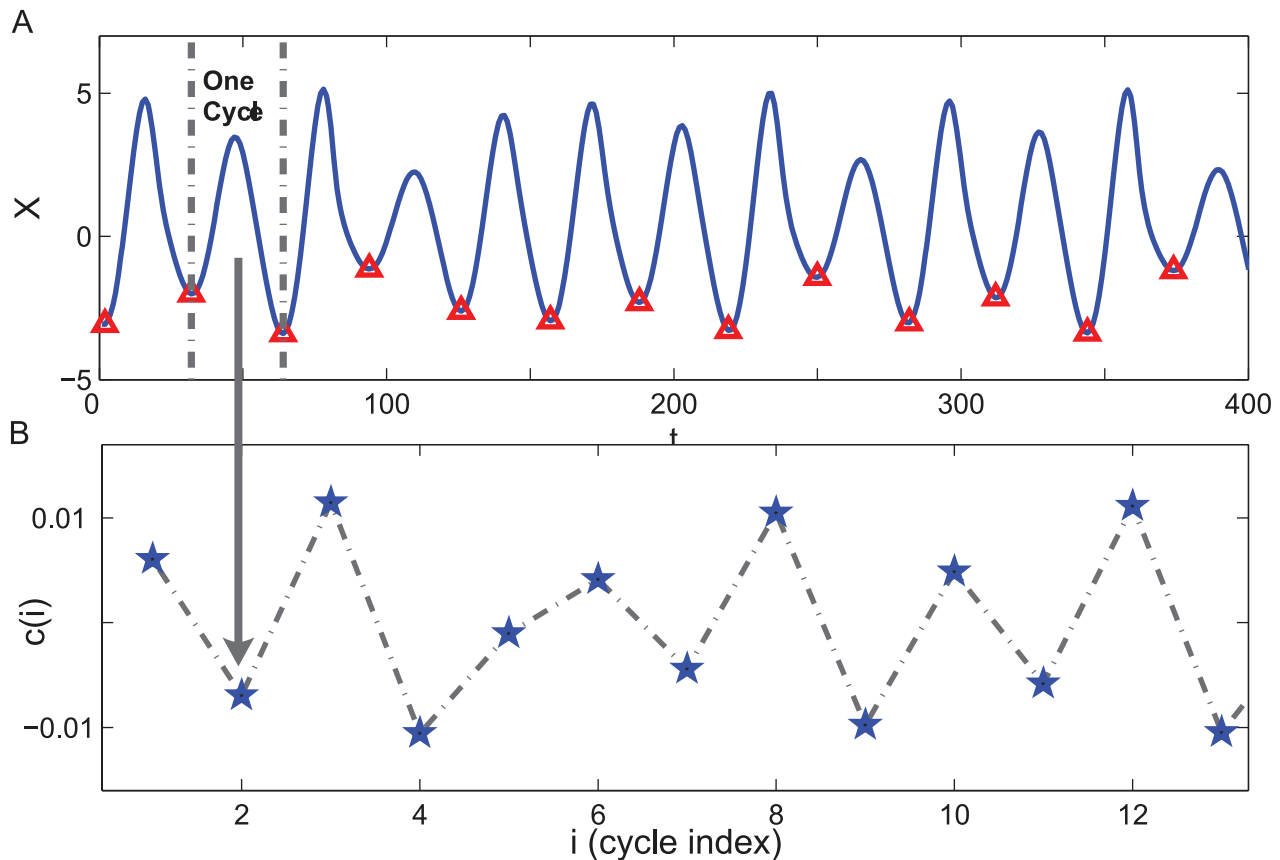
Figure 1. Illustration of transforming a pseudo-periodic time series $X(t)$ into a new series $c(i)$ by reducing each cycle in $X(t)$ to a point. (A) Time series form X-component of the chaotic Rössler system, which demonstrates obvious oscillatory pattern. It can be divided into consecutive cycles at local minimum points (denoted by triangles). (B) A new representation of the oscillatory time series $X(t)$ on the cycle scale, with each point in $c(i)$ corresponding to a cycle in $X(t)$.
doi:10.1371/journal.pcbi.1001033.g001

nearby cycles in phase space usually have similar orientations. Such redundancy can be effectively removed through dimension reduction, leaving only the useful degree of freedom. Finally, it is worthwhile to mention that for long time series with large number of cycles, the nyström method can be adopted to solve large scale spectral clustering problem [37,38].

The Advantage of Characterizing Dynamics on the Cycle Scale

A popular method in nonlinear time series analysis is to reduce a continues flow to a series of discrete points, called a Poincaré section. The Poincaré section is the intersection of flow data in the state space with a hyperplane transversal to the flow. Thus each cycle in the data is simplified into a single point on the Poincaré section, which preserves many properties of periodic or pseudo-periodic orbits. Now we compare $c(i)$ obtained by dimension reduction and the Poincaré section points $P(i)$ obtained by collecting the local minimum points in the data. As can be seen Figure 2, the return plot (i.e., plot of $x(i+1)$ versus its previous values $x(i)$) of $c(i)$ and $P(i)$ show similar quadratic form. Further calculation of the chaotic measure such as correlation dimension indicates that they have the same dynamical origin.

One problem with the $P(i)$ series is that it is highly susceptible to noise that is inevitable in biological data. To see this, we plot the return map for the $c(i)$ and $P(i)$ series obtained from the noisy $X(t)$ data, see Figure 3. We find that although both return maps

display a clear quadratic form intrinsic to the chaotic Rössler system, the return map of $P(i)$ is more vulnerable to noise as the points are more dispersed than that of $c(i)$. We use the variance σ^2 of the least-square fit to the quadratic function ($y = ax^2 + bx + c$) to quantify the influence of noise, with $\sigma^2_{p(i)} = 4.5704$, almost 5 times larger than that of $c(i)$ ($\sigma^2_{c(i)} = 0.9360$). Now we explain why $c(i)$ is more robust to noise. First we take the "cycle" as the basic unit rather than a discrete point in the time series. Obviously the latter is more vulnerable to noise. Second, the acquisition of $c(i)$ is based on an optimization that preserves the proximity relation among all cycles simultaneously, while $P(i)$ is obtained by treating each cycle independent of one another. Our approach utilizes the richer information of pairwise cycle correlation, therefore it not only excavates the inherent dynamics obscured by the cyclic trend, but also offers an extra robustness to noise due to the global nature of this method.

Detecting Degree of Synchronization from Bivariate Oscillatory Data

Another interesting phenomena associated with rhythmic process is synchronization between self-sustained oscillators, which plays an important role in understanding coordination or cooperation in biological systems [39,40]. Several different types of synchronization have been observed, such as complete synchronization [41–43] generalized synchronization [44], and phase synchronization [45].
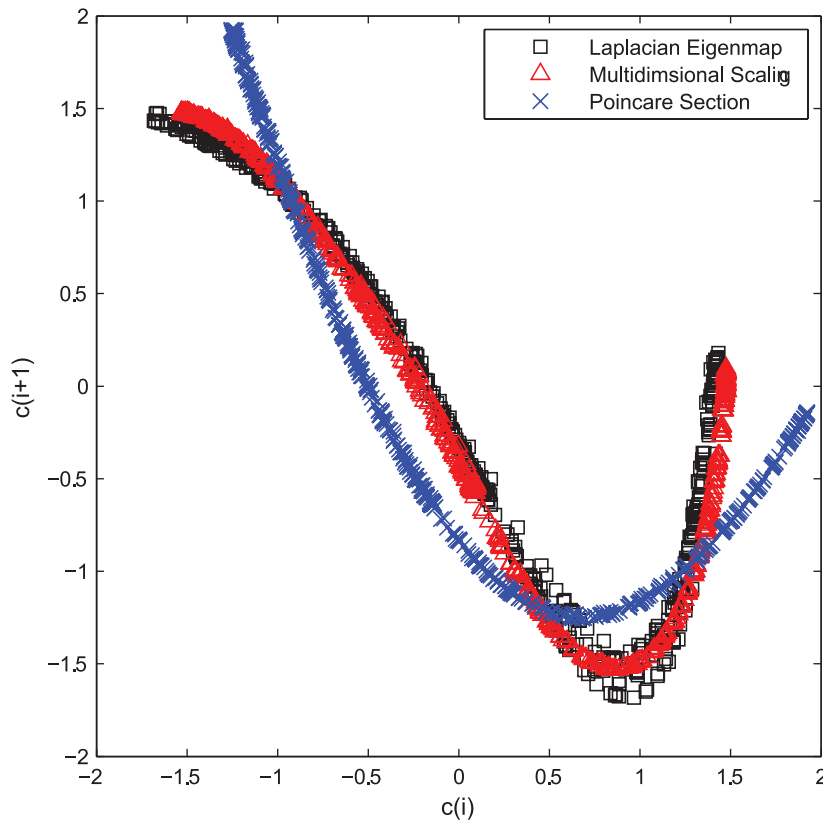
Figure 2. Return maps of $c(i)$ series and Poincaré section points $P(i)$. The time series $c(i)$ is extracted from the oscillatory data $X(t)$ using two dimension reduction techniques, i.e., the Laplacian Eigenmaps and Multidimentional scaling. The Poincaré section points $P(i)$ are extracted by collecting the local minimum points of $X(t)$.
doi:10.1371/journal.pcbi.1001033.g002

The evaluation of degree of synchronization from the outputs of coupled systems is of particular interest to study the interactions in biological systems. For example, the consistency of mutual nearest neighbors and the peakness of the phase difference distribution are used to characterize the dynamical interdependence [46] and phase synchronization [45], respectively. Here we are interested in the case where two processes are phase synchronized, but the synchronization strength is hard to estimate due to noise and non-phase-coherence, or is too subtle to be differentiated due to the mask of strong phase synchronization. For example, the knee and ankle move perfectly in phase during human walking. Under such circumstances the phase synchronization index will take on high values for both healthy subjects and diabetics and it may not probe the subtle difference in the degree of synchronization masked by the strong phase synchronization and noise.

To solve this problem, we propose to quantify the degree of synchronization between two noisy, phase-synchronized oscillatory processes *on the cycle scale* through the reduced representation $c(i)$ of the original data, thereby minimizing the influence of phase synchronization. We illustrate with the X and Y components of the noisy Rössler system, which are perfectly in phase and therefore serve as ideal benchmark data. The two time series are first segmented into cycles according to the local minimums of either X or Y time series, then we apply Laplacian eigenmap to both and obtain reduced representations, i.e., c_X and c_Y for X and Y time series. Finally we calculate the linear correlation coefficient ρ between c_X and c_Y as an indicator of the degree of synchronization between X and Y data. We find that the extracted c_X and c_Y series can successfully reveal the synchronization pattern in presence of noise, which is demonstrated by an increasing trend in the corresponding scatter plot (Figure 4B). The Poincaré section points, however, are less informative of the synchronization degree due to the presence of noise, as P_X and P_Y do not demonstrate clear correlation (Figure 4A).

Results

Data Description

Now we apply the method proposed in previous section to human gait data collected from two groups: the healthy controls (CO) and neuropathic patients (NP, with significant diabetic neuropathy), each with 10 subjects [4]. The kinematic data were collected from a portable data-logger equipped on the subjects during continuous overground walking for 10 minutes (sampled at 66.7 Hz). Three electrogoniometers were placed on the approximate joint centers of the hip, knee, and ankle joints of the right leg to measure their sagittal plane movements. Here we consider the signals measured from knee and ankle joints movement.

Characterizing Human Locomotion Dynamics

Human locomotion is a highly complex, rhythmic process that involves control from subcortical locomotor brain regions and feedback from various peripheral sensors. Typically, the human gait time series (see Figure 5 for knee and ankle movement) exhibit a stable frequency while irregular stride-to-stride fluctuation. For biological signals with a strong periodic component, vital
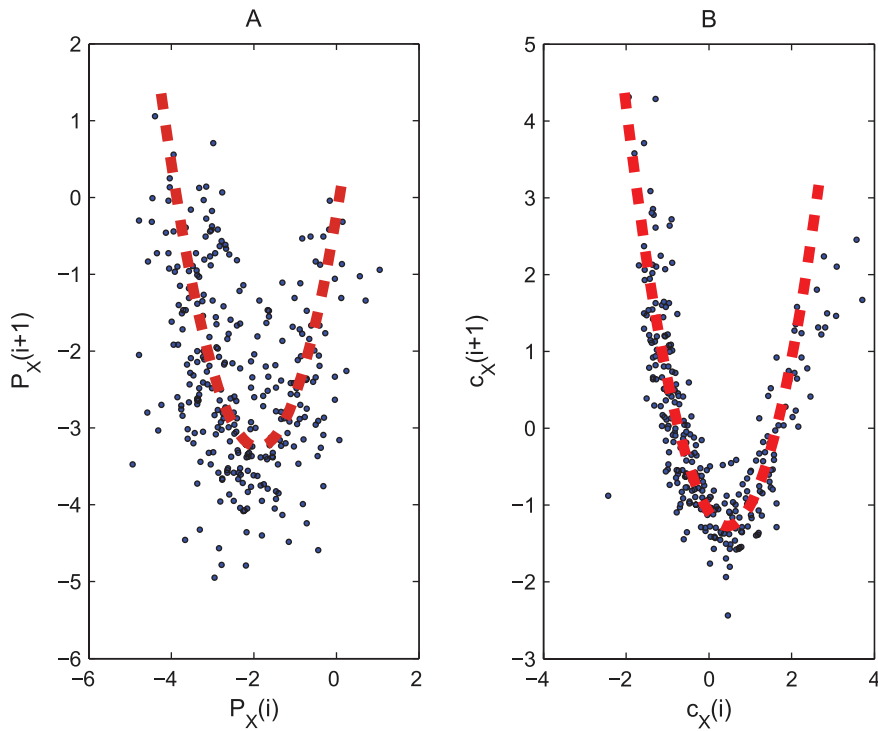
Figure 3. Return plots of Poincaré section points $P(i)$ and $c(i)$. (A) Return plot for $P(i)$ series that is obtained by collecting the local minimums X_{min}. (B) Return plot for $c(i)$ series that is extracted from $X(t)$ by dimension reduction. The Rössler system here is corrupted by 5% dynamical noise and 30% measurement noise. Obviously $c(i)$ is less influenced by noise than $P(i)$.
doi:10.1371/journal.pcbi.1001033.g003

Figure 4. Correlation between X and Y component of noisy Rössler system as is revealed by Poincaré section points $P(i)$ and $c(i)$. (A) Correlation between $P_X(i)$ and $P_Y(i)$. (B) Correlation between $c_X(i)$ and $c_Y(i)$. The Rössler system is corrupted by 5% dynamical noise and 30% measurement noise.
doi:10.1371/journal.pcbi.1001033.g004

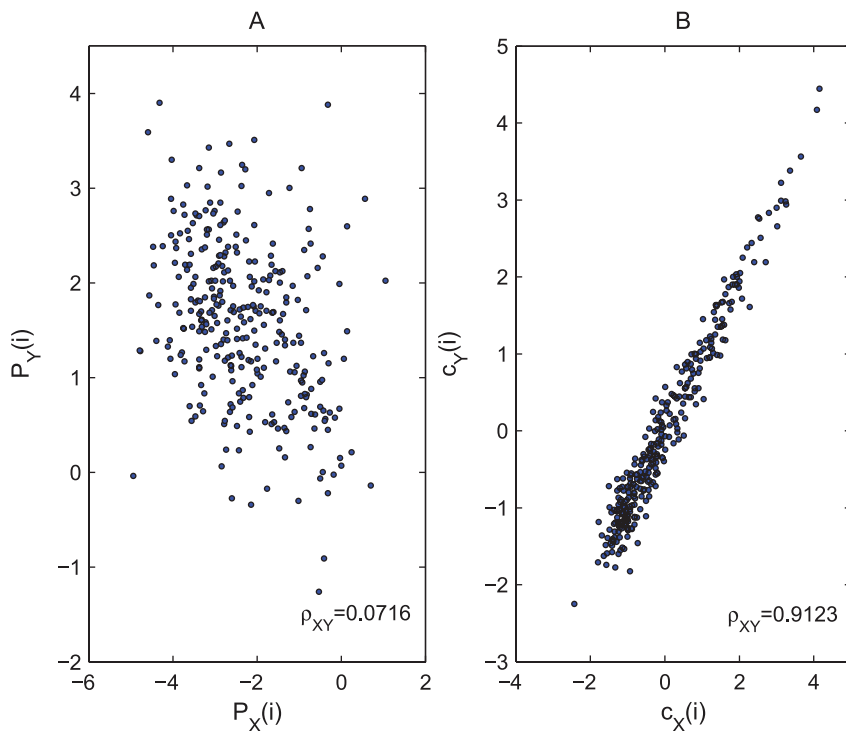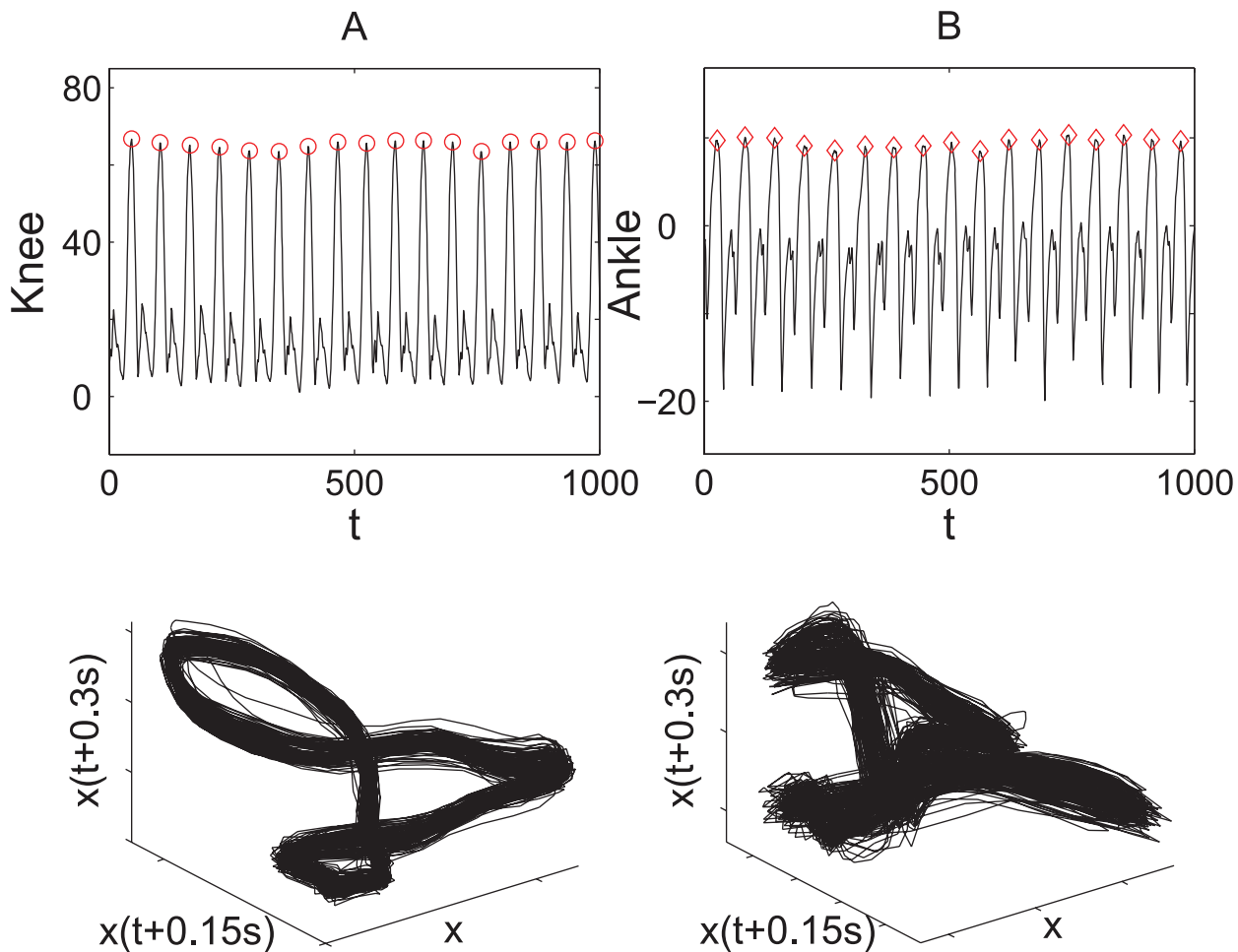**Figure 5. Time series (upper panel) and the corresponding phase space reconstructions (lower panel) of knee and ankle locomotion from a healthy subject.** (A) Knee locomotion data. (B) Ankle locomotion data. The time series are typically non-phase-coherent, demonstrating multi-oscillation within each cycle. This is also evident from the multi-center rotations of the attractor in phase space (lower panel). The two time series are divided into consecutive cycles by their respective local maximum points.
doi:10.1371/journal.pcbi.1001033.g005

information regarding the pathology and pathophysiology of the subject is hidden in the *cycle-to-cycle variation*. Accurately extracting this fluctuation and characterizing its dynamics are expected to provide important insights into the underlying neuromuscular control of walking and yield greater diagnostic information. For example, the stride interval (SI), defined as the time duration of each gait cycle, has been widely used to study the human gait [27,28,31–33]. It was shown that SI series displays long-range correlation intrinsic to the healthy locomotor system.

The SI series contains the information of the duration of each gait cycle. Another source of information consists in the waveforms of the gait cycles, which is not reflected in SI series [29]. As can be seen in Figure 6, the stride interval from the gait data seems to contain insufficient information to reflect the intrinsic dynamics due to digitization. Therefore it is natural to expect that the $c(i)$ series, which is obtained by comparing the waveforms of each cycle (therefore preserves the full dynamical pattern within each cycle) will contain more information relevant to the locomotion dynamics. Meanwhile, $c(i)$ successfully removes the periodic trend that may obscure the underlying dynamics, and has the same advantage as SI series. In the following we will demonstrate that $c(i)$ obtained by Laplacian Eigenmap can reveal the dynamical fluctuation underlying the cyclic trend more effectively than the SI

series, so that we can distinguish clearly between the healthy and pathological groups and make inference about the neuromuscular control, especially on the role of sensory feedback from the feet in regulating dynamics of human walking.

First we check the ankle movement (see Figure 5B), and use the Laplacian eigenmap to extract the fluctuation $c(i)$ on cycle scale for the two groups. To quantitatively characterize the time evolution of $c(i)$, we furthermore compute its power spectrum density (PSD), see the top row in Figure 7. We find that most CO subjects demonstrate broad band spectrums (i.e., $1/f$ noise) that scale as $1/f^{\beta}$, with $\beta = 0.76(mean) \pm 0.23(std)$ (see Figure 8A), indicating the presence of long range correlation (i.e., the strides separated by a large time span are still statistically correlated). In comparison, the power spectrum of the diabetic patients are mostly flat resembling white noise processes ($\beta = 0.37 \pm 0.16$), which means that the strides at different times are mostly uncorrelated. The values of β in the two groups are statistically different ($P < 0.01$). This difference, however, has not been found with either the stride interval (SI) series (see the middle row in Figure 7) or the raw data (see the bottom row in Figure 7).

The absence of long range correlation in the ankle kinematics of the NP group suggests the alteration of the locomotor pattern in the lower limbs of neuropathy patients. This is due to the loss of
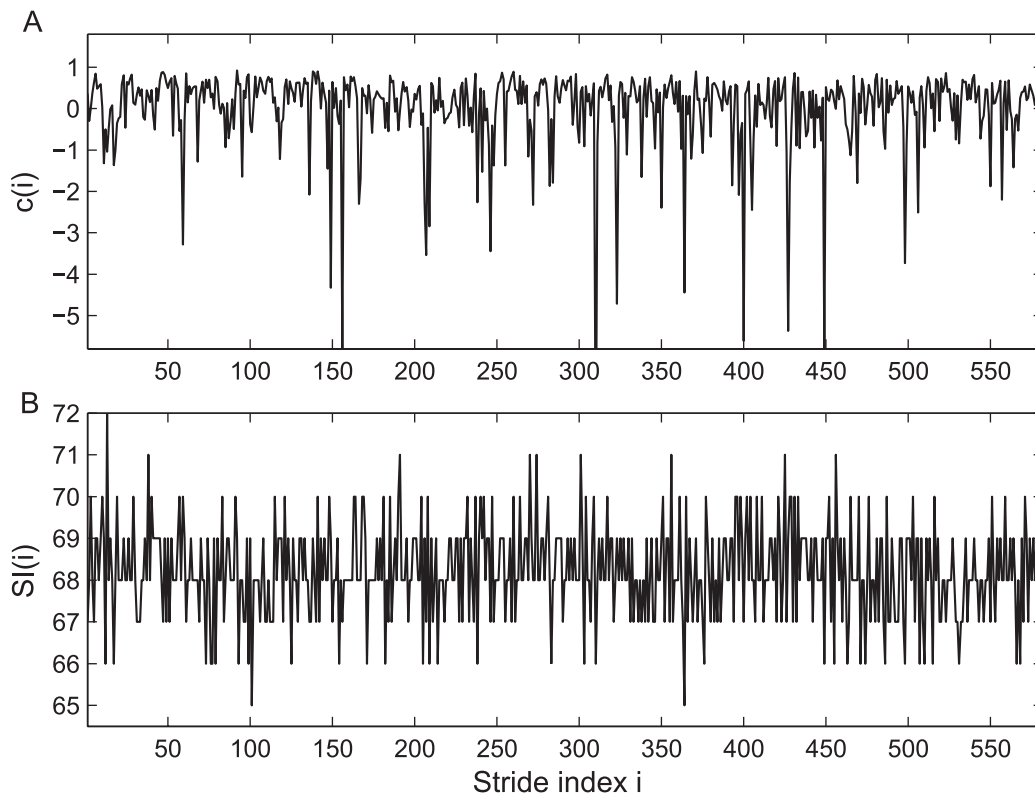
Figure 6. Time series of the extracted $c(i)$ and the stride interval (SI) from ankle movement data of a healthy subject. (A) $c(i)$ time series, which demonstrates significant fluctuation that is not observed in original ankle movement data. (B) Stride interval series extracted from the ankle movement data, which, due to digitization in data collection, loses much information about the original dynamics.
doi:10.1371/journal.pcbi.1001033.g006

peripheral sensation in the lower limbs, which arises from the gradual dying back of nerves from the fingers and toes typical of diabetes. Despite the deterioration in peripheral nervous system, the *knee* movement of the NP group is *still* found to demonstrate a stable long range correlation indistinguishable from the CO group ($\beta_{CO}=0.63(mean)\pm0.21(std)$, $\beta_{NP}=0.62(mean)\pm0.19(std)$, which are statitcally identical with $P=0.965$), see Figure 8B. These results suggest that the impaired peripheral feedback caused by the dying nerves in the feet does not influence the upper-limb dynamics, which leds us to another fundamental problem in human walking, i.e., what is the role of sensory feedback in adjusting the global locomotor dynamics? To understand this, we examine further the degree of synchronization between the knee and ankle movement using dimension reduction.

Assessing Synchronization between Knee and Ankle Movement

Human walking involves the coordination of two major joints, i.e., the knee and the ankle, whose movements during continuous walking are obviously in phase due to the physical connection between them. However, we find that correlation between knee and ankle movement for the two groups can hardly be distinguished by the phase index of the signal due to the presence of strong phase synchronization. Also, noise tends to destroy the local structure in phase space and thus hampers the dynamical dependence measures [47]. To circumvent these difficulties, we propose to compare the dynamics of the two time series by using their Laplacian eigenmap $c(i)$'s. Note that each time series can be

segmented by either its own local maximums, or those of its partner series (shown in Figure 5). Therefore we will segment each time series twice and compute the averaged correlation coefficients ρ_{ij}'s between c_{ankle} and c_{knee} for these two segmentation schemes.

Figure 9A shows the typical synchronization pattern between ankle and knee movement for healthy subjects. The scatter plot between $c_{knee}(i)$ and $c_{ankle}(i)$ demonstrates a significant increasing trend, indicating that the knee and ankle movements are highly synchronized. The correlation coefficient ρ between $c_{ankle}(i)$ and $c_{knee}(i)$ for the healthy group takes on a high value: $\rho=0.68(mean)\pm0.19(var)$, see Figure 8C. For diabetics, however, there is little correlation between $c_{ankle}(i)$ and $c_{knee}(i)$, as is manifested in randomly distributed points in Figure 9B. The correlation coefficient in this case is also low: $\rho=0.26(mean)\pm0.18(var)$, and the values of ρ is statistically different for the two groups ($P<0.001$). All the results of β and ρ values for the two groups of individuals are summarized in Supplementary Table S1. Again, the discrimination between CO and NP groups cannot be achieved by SI series, which always exhibits a strong correlation between the two joints (Figure 9, lower panel), corresponding to high degree of phase synchronization. Finally, we point out that a more comprehensive description of synchronization can be achieved by examining more Laplacian eigenvectors. In the current case the single eigenvector y_2 already encodes the primary variability and is thus sufficient for the discriminative task.

The lack of significant synchronization between ankle and knee movements observed in diabetic patients suggests the "incoordi-
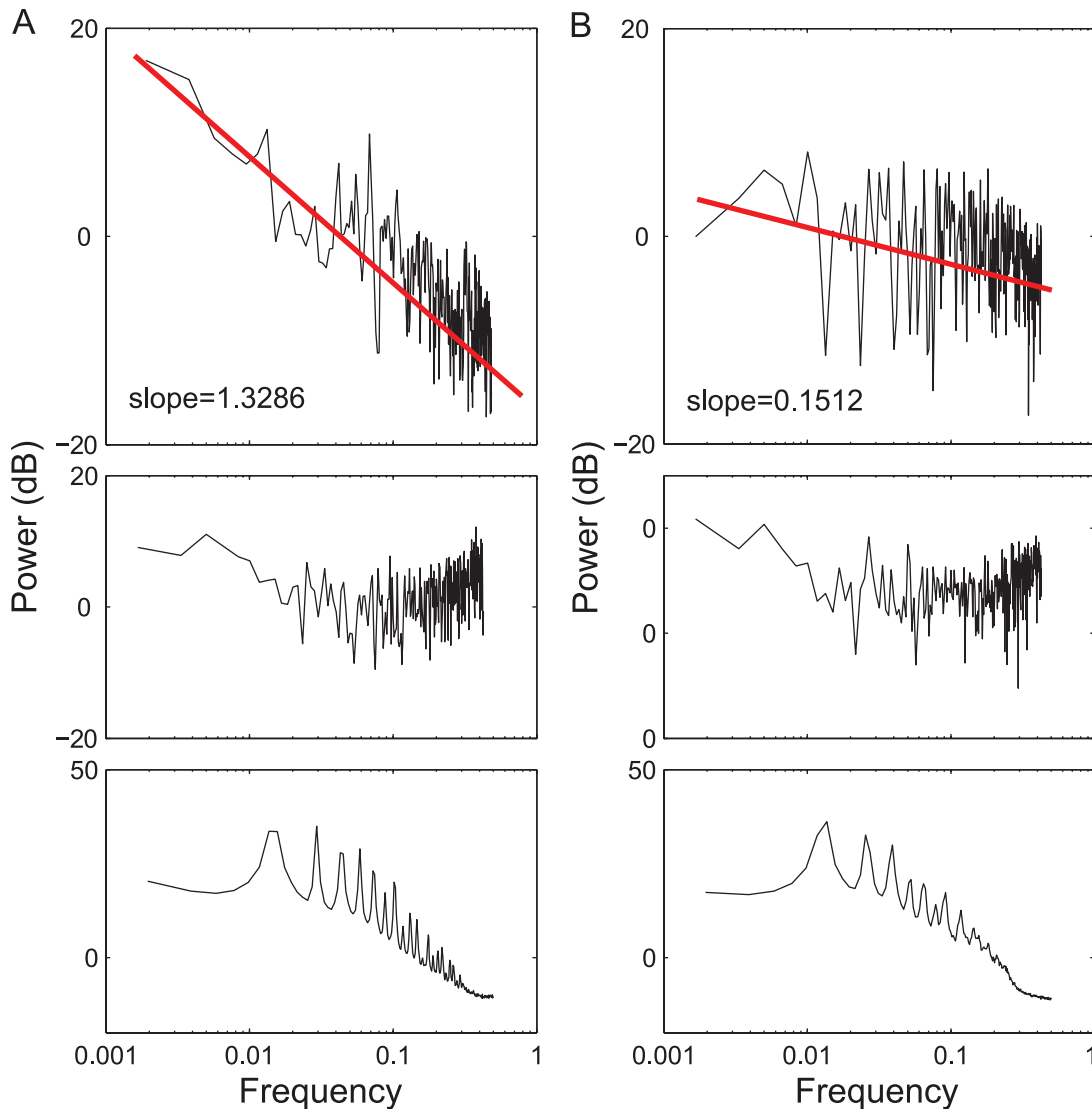
Figure 7. Power spectrum density (PSD) for a typical healthy subject (the left panel) and a diabetes patient (the right panel). (A) A healthy subject. (B) A diabetes patient. The top, middle and bottom rows are PSDs for the extracted $c(i)$, stride interval series, and the original ankle data, respectively. It is obvious that the PSDs for the stride interval series and the original data show no significant difference between the healthy subject and the diabetic patient. The sampling rate of $c(i)$ can be taken as the mean stride interval, and the log-log PSDs are then fitted with a linear function using least-square regression.
doi:10.1371/journal.pcbi.1001033.g007

nation" between the ankle and knee movements. This may arise from the gradual deterioration of the nerves in foot and toes of the diabetics, which is unable to produce sufficient neural feedback for the lower limb to be coordinated with the upper limb. In section "Characterizing Human Locomotion Dynamics" we have found that both the ankle and knee movements for healthy subjects demonstrate long range correlation, while for patients, only the knee movement show long range correlation. This finding is consistent with the result obtained here, i.e., the ankle and knee movement are more synchronized for healthy people than for the diabetics. Finally it should be noted that our study is limited by the relatively small sample size (each group has 10 subjects). Therefore significance tests are performed to verify the differences observed between the two groups of individuals. The current conclusion will be further validated on a larger data base available in the future.

Discussion

Central Nerve Control over Peripheral Nerve System

A fundamental question concerning human walking is the origin of the long range correlation (or $1/f$ noise) found in human gait data, the mechanism of which are not exactly clear [28]. Generally, the locomotor system incorporates inputs from both the central nervous system such as the motor cortex and basal ganglia, and peripheral inputs and sensory feedbacks. Both these two kinds of inputs are suggested to be possible reasons for the presence of the long range correlation in normal human walking [4].

In section "Characterizing Human Locomotion Dynamics", we found that although the locomotion dynamics of the ankle shows significant difference between the normal persons and the patients in terms of long range correlation, their knee movements
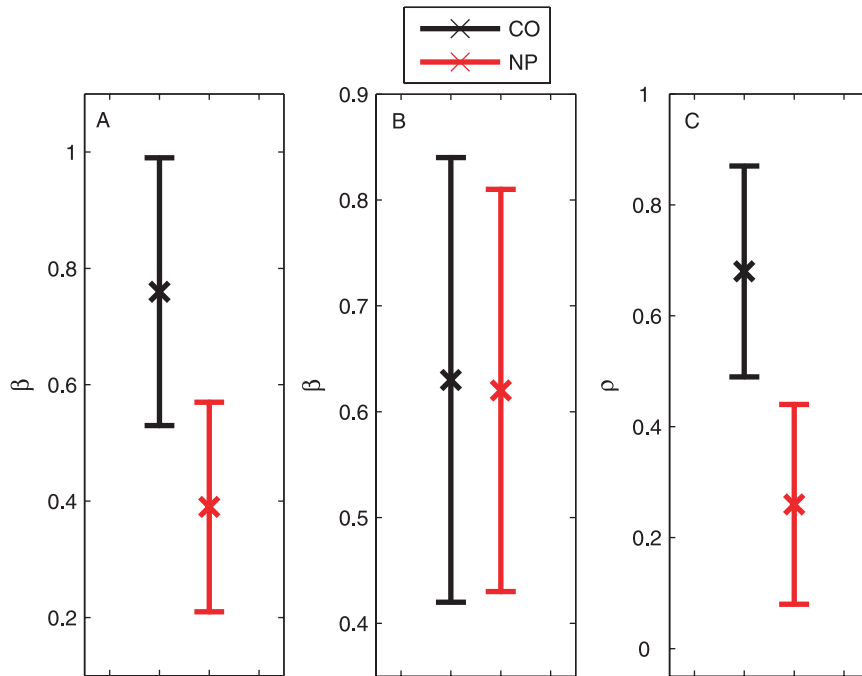
Figure 8. Mean and standard deviation of the derived statistics for control (CO) and neuropathic (NP) groups. (A) The slope β of the least-square-fit to the power spectrum density for the ankle movements. (B) The slope β of the least-square fit to the power spectrum density for the knee movements. (C) The correlation coefficient ρ between $c_{ankle}(i)$ and $c_{knee}(i)$.
doi:10.1371/journal.pcbi.1001033.g008
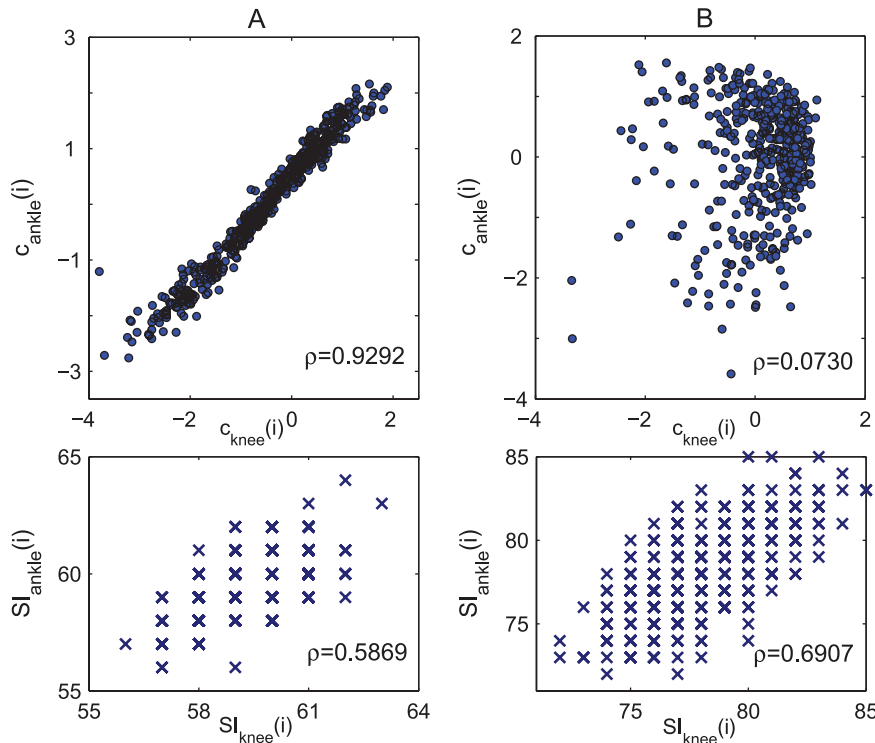
Figure 9. Synchronization pattern between knee and ankle locomotion revealed by $c_{knee}(i)$ and $c_{ankle}(i)$ (upper panel) and stride interval series $SI_{knee}(i)$ and $SI_{ankle}(i)$ (lower panel). (A) A healthy subject. (B) A diabetes patient. As can be seen in the lower panel, the stride interval series $SI(i)$ cannot distinguish the healthy from the diabetics, as the scatter plots between $SI_{knee}(i)$ and $SI_{ankle}(i)$ show similar increasing trends for both subjects.
doi:10.1371/journal.pcbi.1001033.g009

demonstrate similar scaling properties. These results support the belief that the impaired peripheral feedback from the sensors in the feet of diabetics influences only the lower limb locomotion while not that of the knees. We therefore conclude that human walking is not critically dependent on the feedback from peripheral feedback of the lower-limb, and that the central nervous system is playing a major role in regulating locomotor dynamics. In fact it has been found that pathology in central nervous system, such as Huntington's disease, can result in a loss of long range correlation in the gait dynamics [28]. For diabetics, although the peripheral sensory feedback is weakened, their central nervous system is not damaged and still plays an important role in adjusting the locomotion dynamics. This is why their knee locomotion still demonstrate $1/f$ dynamics. Finally, it was pointed out that diabetics may still retain proximal somatosensory inputs, and visual or vestibular feedback information [33]. Further study need to be done to clarify the role of these factors in regulating the dynamics of human walking.

Possible Application to Other Areas

Our approach may be of great relevance, and is expected to provide more accurate and robust characterization and diagnostics to the complex oscillatory data observed in general biological and engineering fields. Reconstructing the dynamics on the cycle scale also brings new vitality to a number of other approaches which are otherwise not suitable for analyzing rhythmic data directly, such as detrended fluctuation analysis [48], recurrence plot [20–22,49], entropy measures [50–52], surrogate data method [17], causality analysis[53–56], and so on. The inherent periodicity in time series can cover up the intrinsic dynamical fluctuation, thus extracting the dynamics on a cycle scale is crucial for subsequent analysis. For example, human vowel data and sunspot number variation are typical oscillatory time series in biomedical and astrophysical fields. We can easily extract the corresponding $c(i)$ series from such signals and feed them into DFA algorithm. By this we are able to tell if the original data has long range correlation on the cycle scale more reliably than other approaches.

Another striking example of physiological rhythms and their interaction is the complex, human cardiovascular system (CVS) [57,58]. For example, the human heart is driving the blood circulation, in which one heartbeat corresponds to one cycle of blood pressure variation. Meanwhile, the heart rate is modulated by respiration through the so called Respiratory Sinus Arrhythmia (RSA). Analyzing the interaction among these biological rhythms from simultaneously measured ECG, blood pressure, and respiration force is crucial for understanding the cardiorespiratory control and disease diagnostics. Due to noise and non-phase-coherence of ECG, blood pressure and respiration signal, traditional measures may not be able to capture the subtle changes in the degree of synchronization, which could possibly be probed more accurately by our approach.

Finally, it is worthwhile to note that our approach can also be applied to time-course microarray data [59,60] as well as Functional Magnetic Resonance Imaging (FMRI) time series [61–63]. For example, genes related to cell-cycle control are always expressed periodically and subject to fluctuation, thus reliably characterizing the correlation or synchronization pattern among such genes using microarray data is expected to provide more insights into the corresponding regulatory network responsible for cell-cycle transcription and regulation. Similarly, FMRI experiments, which are unusually designed with periodic stimuli, always lead to periodic BOLD responses demonstrating large fluctuations [64–66]. Our approach is therefore expected to evaluate more reliably the functional connectivity among different regions in the brain for a deeper understanding of cerebral function.

Supporting Information

Table S1 β and ρ values for Control (CO) and Neuropathic (NP) groups.
Found at: doi:10.1371/journal.pcbi.1001033.s001 (0.03 MB PDF)

Acknowledgments

We thank Dr. J. B. Dingwell for providing the human gait data.

Author Contributions

Conceived and designed the experiments: JZ KZ JF MS. Analyzed the data: JZ KZ. Contributed reagents/materials/analysis tools: JZ KZ JF MS. Wrote the paper: JZ KZ JF MS.

References

1. Glass L (2001) Synchronization and rhythmic processes in physiology. Nature 410: 277–284.
2. Bezerianos A, Bountis T, Papaioannou G, Polydoropoulos P (1995) Nonlinear time series analysis of electrocardiograms. Chaos 5: 95.
3. Babloyantz A, Destexhe A (1988) Is the normal heart a periodic oscillator? Biol Cybern 58: 203–211.
4. Dingwell J, Cusumano J (2000) Nonlinear time series analysis of normal and pathological human walking. Chaos 10: 848.
5. Little M, McSharry P, Moroz I, Roberts S (2006) Testing the assumptions of linear prediction analysis in normal vowels. J Acoust Soc Am 119: 549.
6. Small M, Judd K, Lowe M, Stick S (1999) Is breathing in infants chaotic? Dimension estimates for respiratory patterns during quiet sleep. J Appl Physiol 86: 359.
7. Haurie C, Dale D, Mackey M (1998) Cyclical neutropenia and other periodic hematological disorders: a review of mechanisms and mathematical models. Blood 92: 2629.
8. Edwards R, Beuter A (2000) Using time domain characteristics to discriminate physiologic and parkinsonian tremors. J Clin Neurophysiol 17: 87.
9. Olsen L, Schaffer W (1990) Chaos versus noisy periodicity: alternative hypotheses for childhood epidemics. Science 249: 499.
10. Goldberger A, Amaral L, Hausdorff J, Ivanov P, Peng C, et al. (2002) Fractal dynamics in physiology: alterations with disease and aging. P Natl Acad Sci Usa 99: 2466.
11. Kantz H, Schreiber T (2004) Nonlinear time series analysis. Cambridge University Press 369 p.
12. Small M (2005) Applied nonlinear time series analysis: applications in physics, physiology and finance. World Scientific Pub Co Inc. 245 p.
13. Grassberger P, Procaccia I (1983) Characterization of strange attractors. Phys Rev Lett 50: 346–349.
14. Wolf A, Swift J, Swinney H, Vastano J (1985) Determining Lyapunov exponents from a time series. Physica D 16: 285–317.
15. Cazelles B (1992) How predictable is chaos? Nature 355: 25–26.
16. Abarbanel H, Brown R, Sidorowich J, Tsimring L (1993) The analysis of observed chaotic data in physical systems. Rev Mod Phys 65: 1331–1392.
17. Small M, Yu D, Harrison R (2001) Surrogate test for pseudoperiodic time series data. Phys Rev Lett 87: 188101.
18. Zhang J, Small M (2006) Complex network from pseudoperiodic time series: Topology versus dynamics. Phys Rev Lett 96: 238701.
19. Zhang J, Sun J, Luo X, Zhang K, Nakamura T, et al. (2008) Characterizing pseudoperiodic time series through the complex network approach. Physica D 237: 2856–2865.
20. Donner R, Zou Y, Donges J, Marwan N, Kurths J (2010) Recurrence networks¡ªa novel paradigm for nonlinear time series analysis. New J Phys 12: 033025.
21. Donner R, Zou Y, Donges J, Marwan N, Kurths J (2010) Ambiguities in recurrence-based complex network representations of time series. Phys Rev E 81: 015101.
22. Marwan N, Donges J, Zou Y, Donner R, Kurths J (2009) Complex network approach for recurrence analysis of time series. Phys Lett A 373: 4246–4254.
23. Yang Y, Yang H (2008) Complex network-based time series analysis. Physica A 387: 1381–1386.
24. Lacasa L, Luque B, Ballesteros F, Luque J, Nuño J (2008) From time series to complex networks: The visibility graph. P Natl Acad Sci Usa 105: 4972.

25. Tarca A, Carey V, Chen X, Romero R, Draghici S (2007) Machine learning and its applications to biology. PLoS Comput Biol 3: e116.
26. Lee G, Rodriguez C, Madabhushi A (2008) Investigating the efficacy of nonlinear dimensionality reduction schemes in classifying gene and protein expression studies. IEEE ACM T Comput Bi. pp 368–384.
27. Hausdorff J, Peng C, Ladin Z, Wei J, Goldberger A (1995) Is walking a random walk? Evidence for long-range correlations in stride interval of human gait. J Appl Physiol 78: 349.
28. Hausdorff J, Mitchell S, Firtion R, Peng C, Cudkowicz M, et al. (1997) Altered fractal dynamics of gait: reduced stride-interval correlations with aging and Huntington's disease. J Appl Physiol 82: 262.
29. Bartsch R, Plotnik M, Kantelhardt J, Havlin S, Giladi N, et al. (2007) Fluctuation and synchronization of gait intervals and gait force profiles distinguish stages of Parkinson's disease. Physica A 383: 455–465.
30. Dingwell J, John J, Cusumano J, Diedrichsen J (2010) Do Humans Optimally Exploit Redundancy to Control Step Variability in Walking. PLoS Comput Biol 6: e1000856.
31. Dingwell J, Kang H, Marin L (2007) The effects of sensory loss and walking speed on the orbital dynamic stability of human walking. J Biomech 40: 1723–1730.
32. Kang H, Dingwell J (2008) Separating the effects of age and walking speed on gait variability. Gait Posture 27: 572–577.
33. Gates D, Dingwell J (2007) Peripheral neuropathy does not alter the fractal dynamics of stride intervals of gait. J Appl Physiol 102: 965.
34. Roweis S, Saul L (2000) Nonlinear dimensionality reduction by locally linear embedding. Science 290: 2323.
35. Belkin M, Niyogi P (2003) Laplacian eigenmaps for dimensionality reduction and data representation. Neural Comput 15: 1373–1396.
36. Zhang J, Luo X, Small M (2006) Detecting chaos in pseudoperiodic time series without embedding. Phys Rev E 73: 016216.
37. Zhang K, Kwok J (2009) Density-weighted Nyström method for computing large kernel eigensystems. Neural Comput 21: 121–146.
38. Zhang K, Kwok J (2010) Clustered Nyström Method for Large Scale Manifold Learning and Dimension Reduction. IEEE T Neural Networ 21: 1576–1587.
39. Pikovsky A, Rosenblum M, Kurths J (2001) Synchronization A Universal Concept in Nonlinear Sciences. Cambridge University Press 432 p.
40. Arenas A, Daz-Guilera A, Kurths J, Moreno Y, Zhou C (2008) Synchronization in complex networks. Phys Rep 469: 93–153.
41. Pecora L, Carroll T (1990) Synchronization in chaotic systems. Phys Rev Lett 64: 821–824.
42. Lin W, He Y (2005) Complete synchronization of the noise-perturbed Chuas circuits. Chaos 15: 023705.
43. Lin W, Ma H (2010) Synchronization Between Adaptively Coupled Systems With Discrete and Distributed Time-Delays. IEEE T Automat Contr 55: 819–830.
44. Kocarev L, Parlitz U (1996) Generalized synchronization, predictability, and equivalence of unidirectionally coupled dynamical systems. Phys Rev Lett 76: 1816–1819.
45. Rosenblum M, Pikovsky A, Kurths J (1996) Phase synchronization of chaotic oscillators. Phys Rev Lett 76: 1804–1807.
46. Schiff S, So P, Chang T, Burke R, Sauer T (1996) Detecting dynamical interdependence and generalized synchrony through mutual prediction in a neural ensemble. Phys Rev E 54: 6708–6724.
47. Quiroga R, Arnhold J, Grassberger P (2000) Learning driver-response relationships from synchronization patterns. Phys Rev E 61: 5142–5148.
48. Peng C, Havlin S, Stanley H, Goldberger A (1995) Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. Chaos 5: 82.
49. Marwan N, Carmen Romano M, Thiel M, Kurths J (2007) Recurrence plots for the analysis of complex systems. Phys Rep 438: 237–329.
50. Amigó J, Zambrano S, Sanjuán M (2010) Permutation complexity of spatiotemporal dynamics. Europhys Lett 90: 10007.
51. Amigó J, Zambrano S, Sanjuán M (2007) True and false forbidden patterns in deterministic and random dynamics. Europhys Lett 79: 50001.
52. Amigó J, Zambrano S, Sanjuán M (2008) Combinatorial detection of determinism in noisy time series. Europhys Lett 83: 60005.
53. Brovelli A, Ding M, Ledberg A, Chen Y, Nakamura R, et al. (2004) Beta oscillations in a large-scale sensorimotor cortical network: directional influences revealed by Granger causality. P Natl Acad Sci Usa 101: 9849.
54. Kamiński M, Ding M, Truccolo W, Bressler S (2001) Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance. Biol Cybern 85: 145–157.
55. Guo S, Wu J, Ding M, Feng J (2008) Uncovering interactions in the frequency domain. Plos Comput Biol 4: e1000087.
56. Ge T, Kendrick K, Feng J (2009) A Novel Extended Granger Causal Model Approach Demonstrates Brain Hemispheric Differences during Face Recognition Learning. Plos Comput Biol 5: 172–181.
57. Chon K, Mullen T, Cohen R (1996) A dual-input nonlinear system analysis of autonomic modulation of heart rate. IEEE T Bio-Med Eng 43: 530–544.
58. Luchinsky D, Millonas M, Smelyanskiy V, Pershakova A, Stefanovska A, et al. (2005) Nonlinear statistical modeling and model discovery for cardiorespiratory data. Phys Rev E 72: 21905.
59. Rustici G, Mata J, Kivinen K, Lió P, Penkett C, et al. (2004) Periodic gene expression program of the fission yeast cell cycle. Nat Genet 36: 809–817.
60. Wichert S, Fokianos K, Strimmer K (2004) Identifying periodically expressed transcripts in microarray time series data. Bioinformatics 20: 5.
61. Friston K, Holmes A, Poline J, Grasby P, Williams S, et al. (1995) Analysis of fMRI time-series revisited. Neuroimage 2: 45–53.
62. Worsley K, Friston K (1995) Analysis of fMRI time-series revisited¡again. Neuroimage 2: 173–181.
63. Friston K, Harrison L, Penny W (2003) Dynamic causal modelling. Neuroimage 19: 1273–1302.
64. Bullmore E, Brammer M, Williams S, Rabe-Hesketh S, Janot N, et al. (1996) Statistical methods of estimation and inference for functional MR image analysis. Magn Reson Med 35: 261–277.
65. Marchini J, Ripley B (2000) A new statistical approach to detecting significant activation in functional MRI. Neuroimage 12: 366–380.
66. Guy C, ffytche D, Brovelli A, Chumillas J (1999) fMRI and EEG responses to periodic visual stimulation. Neuroimage 10: 125–148.