Recurrence Quantification Analysis of Human Gait in Neurological Movement Disorders

Pooja Prabhu  
Biomedical Engineering  
Manipal Institute of Technology  
Manipal, India

Dr. N. Pradhan  
Department of Psychopharmacology  
NIMHANS  
Bangalore, India

Abstract—Various gait stability measures have been proposed in the literature, with the aim of quantifying gait abnormalities and balance disorders in healthy and impaired subjects. In this study, the conjecture is that there might be some underlying complex temporal structure in the step to step variations, that is, it may be non-linear or fractal in character. Recurrence quantification analysis helps to know the dynamics of human gait and quantify its parameters in both healthy and pathological subjects. Thirteen subjects were considered to analyze the gait of subjects with neurological disorders (Amyotrophic Lateral sclerosis, Huntington and Parkinson) and thirteen healthy control subjects. The gait parameters are recorded over a duration of 300 seconds per subject using force sensitive resistors. These gait parameters are then used for recurrence quantification analysis by estimating time delay, embedding dimension and threshold. The eight parameters obtained from analysis are then used for classification using probabilistic neural networks. The accuracy deduced from classification associated with stride, swing and stance phase is 50%, 37.5% and 54.16% respectively.

Keywords—Gait; Recurrence Quantification Analysis; Neurological Disorders; Sports Medicine

I. INTRODUCTION

Biological processes possess distinct properties of recursiveness like periodicity, deterministic behavior and irregular cycle. Fundamental property of non-linear or chaotic system is that at a given sufficient amount of time each state of the system may repeat or may be arbitrarily close to a previously observed state. This is called as recurrence property of a system. Fourier analysis of the physiological system like human gait which is non-linear, non-stationary and recurrent in nature, cannot be carried out. Since, the Fourier analysis assumes the signal to be linear, stationary, and it works well for long length datasets [1, 2], so it can only be used in analysis of complicated signals like electroencephalograms, electromyograms and electrocardiograms. To overcome this drawback, Recurrence Quantification Analysis (RQA) was developed, which rest on geometrical patterns derived from the study of non-linear systems. Geometrical patterns are visualized in the form of small scale textures (single dots, diagonal lines, vertical lines and horizontal lines), which are graphed as Recurrence Plots (RP) [1].

Gait cycle duration is rarely constant. Recently, these step to step fluctuations were just assumed to be random noise superimposed on a walking rhythm. The possibility of fluctuation is that there might be some underlying complex temporal structure which is non-linear or fractal phenomena. The gait cycle duration also known as stride time interval, fluctuates continuously in healthy subjects, which results in formation of complex pattern. Gait cycle duration varies from individual to individual and also it varies with the physiological conditions of an individual. Depending on this factor, gait parameters of healthy and impaired gait subjects are considered for the individuals mainly affected by neurological disorders. The gait parameter consists of temporal and spatial parameters. But in this study only temporal parameters are considered, since recurrence property of a system is temporal in nature [3]. Temporal gait measurements are repetitive in nature, because of which the clinicians were able to analyze the gait by quantifying the timings of critical events in cycle with a valuable tool. Stride time interval, swing time interval and stance time interval are the three main temporal gait parameters.

Using the recurrence property of gait parameters, RQA is carried out over gait parameters of every subject. The eight parameters (Recurrence Rate, Determinism, mean diagonal line, maximum diagonal line, entropy, Laminarity, Trapping time and maximum vertical line) yielded from RQA has its own significance over the gait topology. The Recurrence Plot (RP) obtained will exhibit several characteristic small scale structures (textures) and large scale structures (typology) [4], based on which the eight parameters of RQA are derived.

The eight parameters obtained by performing cross RQA between gait parameters associated with left leg and right leg, the classification of pathological subjects and normal subjects corresponding to stride, swing and stance phase is performed using Probabilistic Neural Network (PNN). PNN is suitable for this study because it has fast training process and also the samples used during the training process can be added or removed without extensive restraining the neural networks and also it is inherently parallel in structure.

II. RELATED WORKS

Recurrence plot is a tool which was first introduced by Eckmann et al (1987) in physics to locate hidden patterns in experimental time series. Later on, Zbilut et al (1992) made a significant development by introducing the Recurrence Quantification Analysis, which quantifies the Recurrence Plot. Recurrence Quantification Analysis is well-suited for short, non-stationary and non-linear physiological signals. Dr. Nobert Marwan et al, from University of Potsdam, Germany, has submitted a thesis on Recurrence Quantification Analysis and has also created a CRP (Cross Recurrence Plot) Toolbox.
which is developed under MATLAB software. Tufarelli et al. [5] used RQA in understanding the gait variability due to pathology affected to the different segmental links. Assessing the walking balance in normal and the subjects diagnosed by unilateral vestibular hypo-function by means of three-dimensional accelerations and angular velocities, helps to understand the complexity of head, trunk and pelvis during locomotion. Stagni et al. [6] used RQA in testing the performance of gait stability of individuals whose gait is assumed to be unstable like toddlers, by recording the acceleration and angular velocities using two tri-axial sensors of the trunk and leg. Methods like RQA, harmonic ratio and multi-scale entropy is applied to quantify the gait stability. Turvey et al. [7] used RQA to understand the postural fluctuations in participants with eye open and closed. The results suggested that vision affects the deterministic structure of walking pattern. Neghaban et al. [8] used RQA in understanding the postural sway in unilateral anterior cruciate ligament injury subjects. The result suggested that there is more regularity in double postural task than that of single postural task.

In this study, RQA is performed between gait parameters corresponding to left and right leg. The subjects diagnosed by neurological disorder (ALS, Huntington and Parkinson disease) and healthy subjects are considered and their corresponding stride time interval, swing time interval and stance time intervals are recorded. The eight parameters obtained from RQA, helps in understanding the dynamics of walking pattern.

III. METHODOLOGIES

A. Datasets Under Study

In this study, Neuro-Degenerative Disease Gait Dynamic Database (NDDGD) was used, which was published by Physionet [9]. The database includes the three gait parameters like stride time interval, stance time interval and swing time interval in seconds of every individual. 13 subjects under pathological category (Amyotrophic Lateral Sclerosis, Huntington and Parkinson subjects) and 13 healthy control subjects were considered for this study. The subjects were made wear force sensitive resistors and gait parameters were recorded over 300 seconds.

B. Human Gait

Gait cycle involves the sequence of rhythmical functions which begins when the reference foot contacts the ground and ends with its subsequent floor contact. The gait pattern is periodic in nature [3]. Fig. 1 shows entire gait cycle illustrated by gait phases. Stance phase starts when the reference foot contacts the floor and ends when it becomes contactless with ground. Swing phase is when the reference foot is not in contact with the ground.

Fig. 2, illustrates the seven components in top-down approach of gait like brain, peripheral nervous system (PNS), muscles, synovial joints, segmental links, movement and ground reaction force (GRF) [10] that form the functional basis for normal gait. If any of these seven components are affected by pathology or by a natural cause, it can affect the normal gait.

Gait results due to the co-ordination between the primary motor cortex, basal ganglia and the cerebellum. The pathology considered in the study involves neurological disorders like Amyotrophic Lateral Sclerosis (ALS), Parkinson’s disease and Huntington’s disease which mainly disturb normal gait. The symptom of neurological disorders includes paralysis, poor co-ordination, muscle weakness and loss of sensation [11].

Pre-processing of raw data is necessary in order to remove the spurious peaks and noise outliers. Since peaks occur randomly, the scope for using band-pass filters reduces. Thus a special type of filtering technique is incorporated in which a segment free from spurious peak is considered, and the average of its amplitude values is calculated, which is then used as threshold for the remaining time series.

C. Recurrence Quantification Analysis

The recurrence property of deterministic system or non-linear (chaotic) system can be visualized by plotting dots in the phase space coined as RP, proposed by Eckmann [12]. To build a recurrence plot, reconstruction of state space trajectory is a must. Since only few variables of a biological system are recorded, reconstruction is necessary in order to build the entire topology of dynamic system. The reconstruction is carried out using ‘Taken’s time delay embedding theorem’, which depends on the proper choice of user-adjustable parameters like time delay, embedding dimension and threshold [13]. According to Taken’s [14], the procedure for mapping a discrete signal with n terms, \( p_1, p_2, \ldots, p_n \), into phase space is by constructing vectors \( x_i \) of dimension \( m \) with time delay \( \tau \) as in (1). Mathematically, Taken’s theorem is given by:

\[
\mathbf{x}_i = \left\{ p_1, p_{1+\tau}, p_{1+2\tau}, \ldots, p_{1+(m-1)\tau} \right\}_{i=1,2,\ldots,(n-(m-1)\tau)} \tag{1}
\]

where, \( m \) is the embedding dimension obtained from False Nearest Neighbours (FNN), \( \tau \) is the time delay obtained from mutual information and \( n \) is length of time series.

In order to explore the hidden patterns of dynamic system, the distance between each datum of the vector \( x_i \) with that of another vector \( y_j \) or with itself is calculated, which give rise to correlation matrix or distance matrix \( D_{ij} = \| x_i - y_j \| \). The distances are then normalized by means of threshold. The
normalized values are then mapped to either 1 or 0 based on Heaviside function as in (2) [15]. The corresponding matrix with 1’s and 0’s is called recurrence matrix. Each element of the recurrence matrix is mapped to black or white pixel at the corresponding location, where black and white pixels represent the presence and absence of recurrences respectively, which is then called as Recurrence plot (RP). RP’s visualize spatial and temporal correlations within the sequence of data, which is defined by the matrix,

\[ R_{ij}(\varepsilon) = H(x_i - \|x_j - y_j\|) \]  

(2)

where, \( \varepsilon \) is the predefined threshold, \( x_i \) and \( y_j \) are the two reconstructed vectors and \( H \) is the Heaviside function i.e \( H(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{otherwise} \end{cases} \). When \( i=j \) the line obtained is called line of identity.

Based on the recurrence points, recurrence plot exhibits certain patterns categorized as typology and textures. Large scale structures like typology offer a global impression about the system indicating whether the system is homogenous, periodic, drift or abrupt [16]. The closer inspection of RP gives the small scale patterns as textures, which include single isolated dots, diagonal lines and vertical or horizontal lines.

Qualitative analysis (Recurrence plot) alone may not be sufficient in practical applications, since RP often contains subtle patterns that are not easily established by visual inspection. Zbilut and Webber introduced a new concept termed as Recurrence Quantification Analysis (RQA) [17]. RQA is a quantitative analysis method, which quantifies the small scale structures detected in RP. The eight parameters yielded from RQA are:

1) **Recurrence Rate (RR):** RR is the % of the plot occupied by the recurrence points. The more periodic the signal, higher is the %RR value. Recurrence Rate is calculated as in (3).

\[ RR(\varepsilon) = \frac{1}{N^2} \sum_{i,j=1}^{N} R_{ij}(\varepsilon) \]  

(3)

2) **Determinism (DET):** Determinism discloses the percentage of recurrence points contributing to form the diagonal lines which are longer than the threshold \( l_{min} \). Processes with stochastic behavior (for example: random numbers) produce none diagonals or very short diagonals, but periodic signals (for example sine waves) produce very long diagonal lines. Determinism is calculated using (4).

\[ DET = \frac{\sum_{l=l_{min}}^{N} l P^c(l)}{\sum_{l=l_{min}}^{N} R_{ij}(\varepsilon)} \]  

(4)

where, \( l \) is the length of diagonal line and \( P^c(l) \) is the diagonal line length distribution parallel to the line of identity.

3) **Average diagonal line (L):** Average diagonal line gives the average time the two segments of trajectory were close to each other. Average diagonal line is calculated using (5).

\[ L = \frac{\sum_{l=l_{min}}^{N} l P^c(l)}{\sum_{l=l_{min}}^{N} P^c(l)} \]  

(5)

4) **Maximal diagonal line (** \( L_{max} **)**: The maximal diagonal line is the longest line measured parallel to the main diagonal or line of identity. It determines the chaotic behavior of the system. Periodic signals produce long diagonal lines [18]. Maximal diagonal line is estimated using (6).

\[ L_{max} = \max ((l_i, i=1,2,...,N_i)) \]  

(6)

where, \( N_i \) is the number of diagonal lines parallel to line of identity.

5) **Entropy (ENT):** Entropy quantifies the distribution of recurrent points that form the diagonal lines. More the deterministic structure of RP, smaller the number of bits required to represent the structure [19]. Shannon entropy of distribution of diagonal line is calculated as in (7).

\[ ENT = -\sum_{l=l_{min}}^{N} p(l) \ln(p(l)) \]  

(7)

where, \( p(l) \) is the probability of finding a diagonal of length \( l \) which can be calculated as \( P(l)/k \). \( P(l) \) is the histogram of diagonal lines and \( k \) is the total number of diagonals.

6) **Laminarity (LAM):** Laminarity describes the contribution of recurrence points to form vertical lines of length greater than \( v_{min} \). Laminarity is calculated as in (8).

\[ LAM = \frac{\sum_{v=v_{min}}^{v_{max}} v P^c(v)}{\sum_{v=1}^{N} P^c(v)} \]  

(8)

where, \( P(v) \) is the histogram of vertical lines with length ‘\( v \’ \).

7) **Trapping Time (TT):** Trapping time gives the average length of the vertical structures. The parameter measures the time for which the state is trapped [19]. Trapping time is calculated using (9).

\[ TT = \frac{\sum_{v=v_{min}}^{v_{max}} v P^c(v)}{\sum_{v=1}^{N} P^c(v)} \]  

(9)

8) **Maximal vertical line (Vmax):** Maximal vertical line is given by,

\[ V_{max} = \max ((V_i, i=1,2,...,N_i)) \]  

where, \( N_i \) is the number of vertical lines.

This work is licensed under a Creative Commons Attribution 4.0 International License.
The mutual information is performed between two time series corresponding to left and right leg stride interval, left and right leg swing interval and left and right leg stance interval. The time delay $\tau$ is yielded by the lag corresponding to the first minimum of mutual information [13]. The phase dimension, referred to as embedding dimension has been selected by using FNN method. FNN is obtained by considering particular time series and its corresponding time delay $\tau$. The corresponding value of dimension for which the value of FNN is minimum or zero is considered as embedding dimension. The maximum value of embedding dimension $m$ obtained from the two time series is then used for cross recurrence quantification [20]. The third user-adjustable parameter is threshold $\epsilon$. According to Zbilut, the choice of threshold $\epsilon$ must be such that, the value of $\epsilon$ is 10% of mean or maximum phase space diameter, i.e. at least 10% of recurrence points must be present in the threshold region [13]. After obtaining the three user-adjustable parameters specifically delay $\tau$, embedding dimension $m$ and threshold $\epsilon$, cross RQA is carried out. The cross RQA yields the eight parameters.

IV. RESULT ANALYSIS

A. Preprocessing of Gait Time Series

The raw data collected contains spurious peaks and noise outliers (as shown in Fig.43(a)), that cannot be directly used in RQA. Thus, the raw data has to be filtered using a special type of filter which can remove the spurious peaks as shown in Fig. 3 (b).

![Fig. 2. Left leg stride time series of ALS subject (a) raw data (b) preprocessed data](image)

B. Analysis and Quantification of Cross Recurrence

The cross RQA of left and right leg stride interval is carried out for ALS, control, Huntington and Parkinson subjects. The result obtained has the eight parameters of RQA corresponding to each subject. Table I tabulates the mean $\pm$ standard deviation values of the eight parameters associated with stride phase corresponding to ALS, control, Huntington and Parkinson subjects.

From Table I, it is observed that recurrence rate (RR) and the periodic nature (indicated by DET, $L$ and $L_{max}$) is minimal in control subjects as compared to the other category. This is due to the wide range of motion of the segmental links essential for walking in healthy subjects. Since the entropy value of the control subjects is minimum compared to other classes, the gait pattern of normal subject is said to be more complex than that of pathological subjects.

Table II implies that the recurrence behavior in the swing phase is least in healthy subjects and highest in Huntington subjects, as these subjects show jerky movements or sudden actions. Parkinson subjects have the second highest value, as they exhibit a postural sway. Also, observing the values of the maximum vertical line ($V_{max}$), it can be concluded that the gait parameters of Huntington subjects remain in the same state for some particular time or it gradually changes over time, owing to the highest value of $V_{max}$ observed in the result.

Table III indicates that recurrence and periodicity of control group is least even in the stance phase. The highest $TT$ value observed in ALS subjects indicates that they are trapped in the same state for a long time, this could be due to the stiffness in segmental links and muscle atrophy.

C. Classification using PNN

PNN consists of training sets, which includes the eight parameters of RQA corresponding to the seven subjects belonging to each class. The class vector consists of classes labeled as class 1 for ALS, class 2 for control, class 3 for Huntington and class 4 for Parkinson. The testing set includes eight parameters of RQA corresponding to the six subjects belonging to each class. The accuracy associated with stride phase is 50%, swing phase is 37.5% and stance phase is 54.16%. The stance phase is better classified, since the parameters of RQA are significantly different for all the classes. The accuracy can still be improved by considering more number of subjects for classification.

V. CONCLUSIONS AND FUTURE WORK

The gait parameter recorded in the study is one-dimensional, but if the entire topology of gait has to be considered, then the gait parameters have to be measured in all possible dimensions. This makes the study complicated, since sophisticated machine needs to be developed to analyze the gait variables with all possible dimensions. In this study, using ‘Taken’s time delay embedding theorem’, the variable of one-dimension was converted to higher dimensions to get the entire topological pattern of gait. The time series now changes to the embedded dimension and by exploring the hidden patterns, the entire dynamics of gait and the correlation between left and right leg gait parameter could be visualized. By quantifying the hidden patterns, the eight parameters of...
RQA can be obtained. Incorporating the RQA and classification technique in the gait recording machine, helps in measuring the gait parameters of the subjects and classifies them into the respective classes. Development of gait machine with these criteria helps in sports medicine to know the rate of strains in a particular muscle. The work also has a scope in spatial tracking of human body, since the values corresponding to eight parameters gives some insight into the control process of equilibrium and balancing of posture taking place in brain. The work also helps to distinguish the gait of myopathy and neurologically disordered subjects.

**CONFLICT OF INTEREST:** NONE

**REFERENCES**


