Journal of Multidisciplinary Engineering Science and Technology (JMEST)
ISSN: 2458-9403
Vol. 4 Issue 2, February - 2017

Quantification Of Human Gait Dynamics Using Multiscale Entropy Under Healthy And Diseased Conditions

Muhammad Junaid
Department of Computing and IT
Iqra University
Islamabad, Pakistan
itsrajajunaid@gmail.com

Adeel Khalid Siddiqui
Department of Computing and IT
Iqra University
Islamabad, Pakistan
adeel@iqraisb.edu.pk

Anees Qumar
Department of Computer Sciences & IT
Women University of AJ&K
Bagh, Pakistan
aneesqumar@gmail.com

Izza Naqvi
Iqra University
Islamabad, Pakistan
Izza.naqvi@iqraisb.edu.pk

Dr. Aihab Khan
Iqra University
Islamabad, Pakistan
aihabkhan@yahoo.com

Dr. Lal Hussain *
University of Azad Jammu and Kashmir
Muzaffarabad, Pakistan
lall_hussain2008@live.com
(* Corresponding author)

***Abstract*— Biological signals are generated during the activity of physiological systems. Hence these signals contain information that can be extracted to get information about the state and working of the physiological systems. In order to extract information about underlying complex dynamics of physiological system, sophisticated and robust analysis techniques are required. Various linear and non-linear methods have been proposed for analysis of these signals. In this research multi-scale entropy analysis (MSE) method (a complexity based method) is used for analyzing the human gait signals under various diseased and normal conditions. It is found that the complexity of the stride interval derived from controlled (normal) subject is higher than that of pathological/diseased subject.**

***Keywords*— Gait dynamics; Stride Interval; Approximate Entropy; Parkinson Disease**

I. INTRODUCTION

Variability analysis technique are applicable to any physiological time series but there has been significant interest in quantifying the dynamics of neurophysiologic control systems, particularly, human gait regulation and heartbeat regulation in diseased, health, and advanced age. Studies showed that the gait fluctuations present a complex and non-stationary behavior that may contain some hidden information, which may not be extractable with traditional analysis methods. Such information is of clinical values and is also relate to basic mechanism of healthy and pathologic function. Hausdorff and coworkers were the pioneers to show that the nature of stride to stride interval time series called stride rate variability (SRV), is similar to heart rate variability (HRV) time series [1]. The later studies showed that SRV time series is multi-fractal rather than mono-fractal [2].

During last three decades, various techniques for quantifying the variability of physiological signals has been developed. Though these techniques can be applied to any biological time signal, but considerable efforts have been made for quantifying the dynamics of neurophysiologic control systems like heartbeat regulation and human gait regulation in control, diseased and increased age [3].

Hausdorff and co-workers applied detrended fluctuation analysis (DFA) to analyze stride interval time series derived from the subjects who walked under different walking conditions and indicated that fluctuations of human gait cycle under free walking conditions exhibit long range co-relations with power law decay [4]. This means that any stride interval depends not only on the values of most recent stride intervals but on values of previous stride intervals also showing a memory affect. These findings are indicative of very complex dynamics. Costa and co-workers used Multiscale entropy method (MSE) to quantify the complexity of stride interval time series obtained from healthy subjects under constrained and unconstrained conditions walking at different speeds [5].

Recently Anees & Wajid applied threshold based symbolic time series analysis method to the human gait signals obtained from walking under normal and metronomic protocols at different speeds [6]. They also tested the hypothesis that the complexity of time series is encoded in the sequential order of the stride interval and does not result from stride interval histogram. They found that the complexity of the gait signals obtained under free conditions was more than that of signals obtained under metronomic walking conditions.

In the present study we have applied multiscale entropy to quantify the complexity of stride interval time series data derived from subjects having pathological subjects and normal subjects. The result indicates that diseased systems, when associated with the emergence of more regular behavior, show reduced entropy values compared to the dynamics of free-running healthy systems.

II. MATERIAL AND METHOD

Entropy is basically related to information content. It can be defined as the rate of production of information with reference to dynamical system. Entropy methods are useful to calculate complexity for analysis of biological signals those involves noisy and short data sets specifically the heart rate and stride rate time series datasets.

MSE method to analyze physiological data is based on approximate entropy and the sample entropy (a revision of ApEn) [7]. In this method entropy is not computed directly by matching the patterns of fixed length. In this method original time series is transformed into a new coarse-grained averaged time series related to scale factor τ. If the scale is one, the coarse-grained time series will be exactly input (original)series will contain average of ith and (i+1)th, data item of original time series. The size of coarse-grained time series will be equal to size of original time series divide by „τ". Then Samp of new time series is computed and plotted against scale factor. The mathematical details and computation methods are described in below sections.

A. *Multi Scale Entropy Analysis (MSE)*

Multiscale entropy method is described by Anees & co-workers is reproduced as under [8]:

Let the data input is X = {X_1, X_2, X_3…….. X_N} (time series) where time series length is indicated with N. The first step in calculation of MSE is to construct a new time series called coarse-grained data series. This new time series is constructed by taking average of the continuous points in the series (As shown is Figure. 1). A single point in the newly constructed time series $Y_j^{(s)}$ is computed by applying following equation:

$$Y_j^{(s)} = \frac{1}{s} \sum_{i=(j-1)s+1}^{js} X_i \qquad (1)$$

In the above equation the scale factor is represented by 's' and $1 \le J \le N/s$, and N/s will be the length of each coarse-grained time series. If we use s=1, then the original time series and newly constructed time series will be same.

If Scale=2 then:

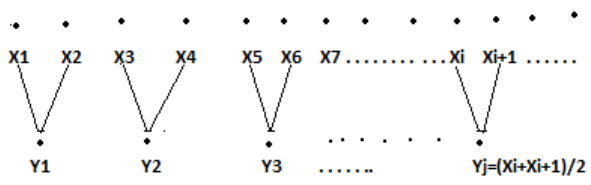


Fig. 1 Presentation of Coarse grained time series at s=2

After construction of coarse grained series, SampEn (Sample Entropy) is computed for new series and the result is plotted as function of 's'. The SampEn is based on Approximate entropy which is described in following sections.

B. *Approximate Entropy Analysis (ApEn)*

The summary of calculation method of ApEn presented by Pincus is reproduced as under[9]:

1. Given a sequence SN, consisting of instant Stride rate measurements, X(1),X(2) , X(N).

TABLE I. DESCRIPTION OF DATA SETS

2. In the second we decide the values of two input parameters *m* and *r* , *m* specifies the pattern length, and *r* defines the similarity criteria.

3. Another subsequence (pattern sequence) of stride rate measurements is formed, beginning at measurement within S_N, by the vector Pm(i).

4. The difference between two patterns Pm(i) and Pm(j) in pattern sequence (P) is calculated and both sequence are said to be similar if the difference is less than the similarity criteria(*r*), i.e., if

$$X(i+k)-X(j+k) < r \text{ for } 0<k<m \qquad (2)$$

5. Let Pm (pattern sequence) is the set of all patterns of length i.e.[Pm(1),Pm(2)……,Pm(N-m+1]], within Sn. In next step we calculate Cm(r) which expresses the prevalence of repetitive patterns of length in SN

$$C_{im}(r) = \frac{n_{im}(r)}{N-m+1} \qquad (3)$$

6. In last step the approximate entropy of S_N is calculated as under, which is natural logarithm of ratio of prevalence of repetitive patterns of length m and length m+1

$$ApEn(S_N, m, r) = \ln\left[\frac{C_m(r)}{C_{m+1}(r)}\right] \qquad (4)$$

C. *Sample Entropy Analysis (SampEn)*

Sample Entropy is modification of the ApEn algorithm in which the patterns are not matched with itself (self-matches are omitted). This technique is less dependent on the size of the data and shows consistency over broad ranges of possible m, r and N. The method for calculating Sample Entropy is same as of Approximate Entropy.

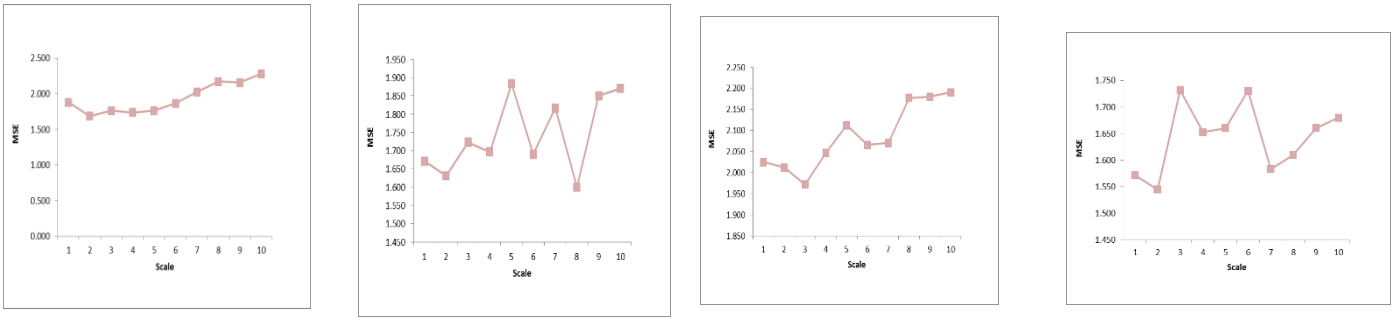
For all cases presented in the study of Costa et. al. are , m = 2 and r = 0:15 [5]. In general, Pincus suggested m=2 and r=0:2 for the analysis of heart rate data [9].

D. *Data Sets*

The Data sets used in our study were obtained from neurodegenerative database available at www.physionet.org [10]. The total data sets are 64 which are categorized as below:

Description/ Subject Groups	No. of Subjects	Male	Female	Age Range

				(Years)
Healthy Subjects	16	2	14	20-74
Huntington Disease Subjects (HD)	20	14	6	29-71
Parkinson Disease Subjects (PD)	15	10	5	44-80
Amyotrophic Lateral Sclerosis (ALS) Disease Subjects	13	10	3	36-70



a. Controlled Subject b. Parkinson Disease subject c. Huntington Disease subject d. ALS Disease subject

Fig. 2 (a-d). MSE Values of different signals at Scales =1 to 10 and (m2, r=0.15)

III. RESULTS

The complexity of stride interval time series data derived from the subjects who walked freely (normal pace) was calculated using the multiscale entropy technique. The number of scales for which the original time series is coarse grained was used between 1 to 10.

Fig. 2(a-d) shows, the values of multiscale entropy (MSE) across different scales (1 to 10) at pattern length m=2 and similarity criteria r=0.15 for time series data derived from different conditions. It can be observed that the values of MSE are maximum at scale s=10.

The comparison of MSE values for all types of subjects (controlled, ALS, Huntington, Parkinson) is shown in Fig. 3. It can be observed that the maximum value of multiscale entropy is obtained for the stride interval time series data derived from normal (controlled) subject. At some scales other types of subject also have higher MSE values but the maximum value is for controlled subject at scale 10. This shows that the complexity of controlled subject is higher than that of diseased subjects.

The resultant MSE data was not normally distributed, so we used Wilcoxon- rank-sum test (Mann–Whitney–Wilcoxon (MWW) test) to check the significant difference between the subjects. The degree of separation between groups at different scale values was quantified by using the p-values.

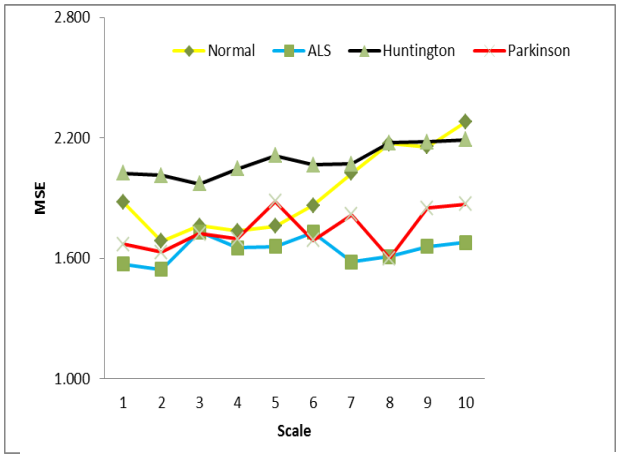


Fig 3. Comparison of MSE values for different subjects at Scales =1 to 10 and (m=2, r=0.15)

The comparison of MSE values of controlled subjects and diseased subjects and their corresponding p-values at various scales are presented in Table 2.

The p-values show that the results are not statistically significant for Normal and Parkinson Diseased subjects at scales less than 8 and 9 to 10 (because p-value is greater than 0.05). However, the maximum degree of separation is available at scale=8(2.170, 1.600and p-value=0.019). This shows that controlled and Parkinson diseased subjects show the same behavior along smaller scales. On comparison of Normal Subjects with Huntington Diseased subjects, the p-values show that the results are not statistically significant at scales greater than 6 (because p-value is greater than 0.05). We got the statistically significant results at scale 1 to 6. However, the

maximum degree of separation is available at scale=5(1.761, 2.112and p-value=0.004). This shows that controlled and Huntington diseased subjects show the same behavior along larger scales.

Next, the MSE values for controlled subjects were compared with the MSE values for the subjects with ALS disease. The p-values show that the results are not statistically significant at scales 1-6 and 8-9 (because p-value is greater than 0.05).

We got the statistically significant results at scale 7 and 10. However, the maximum separation is available at scale=10 (2.278, 1.680 and p-value=0.015). This shows that controlled and ALS diseased subjects show the same behavior along shorter scales. The hypothesis that the complexity of these time series is encoded in the sequential ordering of the stride intervals and does not result from stride interval histograms was further tested. For each time series, a surrogated time series was obtained by shuffling (randomly reordering) the sequence of data points.

TABLE II. COMPARISON OF MSE VALUES FOR NORMAL AND DISEASED SIGNALS AND CORRESPONDING P-VALUES

Scale	MSE		p-value	MSE		P-value	MSE		P-value
	Normal	ALS		Normal	HD		Normal	PD	
1	1.879	1.572	0.361	1.879	2.025	0.032	1.879	1.671	0.253
2	1.685	1.545	0.748	1.685	2.012	0.006	1.685	1.631	0.950
3	1.763	1.732	0.980	1.763	1.972	0.029	1.763	1.723	0.787
4	1.736	1.653	0.902	1.736	2.047	0.021	1.736	1.696	0.755
5	1.761	1.660	0.902	1.761	2.112	0.004	1.761	1.883	0.394
6	1.864	1.730	0.537	1.864	2.066	0.067	1.864	1.690	0.289
7	2.025	1.583	0.028	2.025	2.070	0.399	2.025	1.817	0.280
8	2.170	1.610	0.415	2.170	2.177	0.514	2.170	1.600	0.019
9	2.155	1.660	0.401	2.155	2.180	0.899	2.155	1.850	0.371
10	2.278	1.680	0.015	2.278	2.190	0.577	2.278	1.870	0.560

Correlations among the stride intervals were destroyed while preserving the statistical properties of the distribution. The results are shown in Table III. These results show that physiologic time series are more complex than surrogate ones.

The results show that the complexity of original time series is more than the complexity of shuffled time series. The category to category description is listed below:

A. Controlled Original Vs Controlled Surrogated

The p-values show that the results are statistically significant for controlled subjects (Normal and shuffled/ surrogated data) at scale 2 to 7 while they are not statistically significant at other scales. However, the maximum degree of separation is available at scale=5 (1.761, 1.430and p-value=0.0001).

B. ALS Original Vs ALS Surrogated

The p-values show that the results are statistically significant at scale 7 while they are not statistically significant at other scales. The maximum degree of separation is available at scale=7(1.583, 1.992and p-value=0.0451)

C. Huntington Original Vs Huntington Surrogated

The p-values show that the results are statistically significant at scale 5,6 and 8 while they are not statistically significant at other scales. The maximum degree of separation is available at scale=5(2.112, 2.205and p-value=0.0179).

D. Parkinson Original Vs Parkinson Surrogated

The p-values show that the results are not statistically significant at scale 1, 5 and 10 while they are statistically significant at other scales. The maximum degree of separation is available at scale=6 (1.690, 2.107and p-value=0.0014).

IV. DISCUSSION

The extraction of information related to the physiological behavior of system by analyzing the biological signals is an interesting and imperative research field. Because of the nonlinearities and complexness present in the underlying physiological systems, the linear analysis cannot provide the complete information about the state of the system. This limitation of linear techniques makes the non-linear measures an important way to analyze the physiological signals. More robust techniques with ability of classification of signals are needed to quantify the dynamics of biological signals in normal

and abnormal conditions. In this thesis, multiscale entropy analysis is used to study the human gait dynamics of controlled and diseased subjects.

The stride interval time series of controlled subjects was compared with the stride interval time series of neurodegenerative disease subjects. The MSE values for controlled subject were large than the MSE values of ALS diseased subjects at all scale values At a scale of 10, maximum value of MSE (hence maximum complexity) is obtained. While comparing the controlled subjects with Huntington diseased subjects, it was found that at shorter scales the complexity of Huntington diseased subjects is higher than that of controlled subject but at scale value of 10, the maximum value of MSE is obtained for controlled subjects. The MSE values for controlled subject were large than the MSE values of Parkinson diseased subjects at all scale values. At a scale of 10, maximum value of MSE (hence maximum complexity) is obtained.

The neurodegenerative diseased subjects were compared with each other also to find out the classification in the sub groups. On comparing the ALS subject with Huntington and Parkinson subjects, it was found that the Huntington diseased subjects are more complex than ALS and Parkinson diseased

TABLE III. COMPARISON OF ORIGINAL SIGNAL AND SURROGATED SIGNALS

scale	MSE(Control Subject)		p-value	MSE (ALS Subject)		p-value	MSE (Huntington Subject)		p-value	MSE (Parkinson Subject)		p-value
	original	surrogated		Original	surrogated		original	surrogated		original	surrogated	
1	1.879	1.104	0.1011	1.572	1.809	0.3933	2.025	2.091	0.5792	1.671	1.835	0.2064
2	1.685	1.139	0.0005	1.545	1.863	0.1679	2.012	2.237	0.3648	1.631	2.016	0.014
3	1.763	1.233	0.0006	1.732	2.047	0.5545	1.972	2.157	0.2393	1.723	2.108	0.0456
4	1.736	1.379	0.0003	1.653	2.092	0.5114	2.047	2.322	0.062	1.696	2.049	0.0508
5	1.761	1.430	0.0001	1.660	1.996	0.4307	2.112	2.205	0.0179	1.883	2.102	0.2064
6	1.864	1.391	0.0002	1.730	2.108	0.2122	2.066	2.441	0.0215	1.690	2.107	0.0014
7	2.025	1.306	0.0044	1.583	1.992	0.0451	2.070	2.300	0.0856	1.817	2.150	0.0054
8	2.170	1.338	0.1871	1.610	2.100	0.6928	2.177	2.350	0.0199	1.600	2.190	0.0028
9	2.155	1.400	0.1316	1.660	2.170	0.2094	2.180	2.410	0.3939	1.850	2.200	0.0033
10	2.278	1.485	0.6509	1.680	2.200	0.0994	2.190	2.380	0.8176	1.870	2.210	0.3218

subjects, while Parkinson subject are more subjects, while Parkinson diseased subjects are more complex than ALS subjects at almost all scale values At a scale of 10, maximum value of MSE (hence maximum complexity) is obtained. The neurodegenerative diseased subjects were compared with each other also to find out the classification in the sub groups. On comparing the ALS subject with Huntington and Parkinson subjects, it was found that the Huntington diseased subjects are more complex than ALS and Parkinson diseased subjects, while Parkinson subject are more complex than ALS subjects at almost all scale values. The Multiscale entropy showed a significant difference between controlled subjects and diseased subjects. Our study showed that the dynamics of normal subjects are more complex than the subjects who suffered from neurodegenerative diseases.

We noted that our findings complement those obtained from previously reported Multiscale Entropy analysis of gait data by Costa *et al* [5].

V. CONCLUSION

Multiscale entropy analysis has widely been used to analyze the signals generated from biological systems. In this study, multiscale entropy analysis was applied to compare the complexity of human gait time series from healthy subjects neurodegenerative diseases subjects. We note that stride interval time series of human gait presents a complex behavior and the complexity of spontaneous output of the human locomotors system during normal walking is higher than diseased conditions. The main limitation of the study is that, the number of subjects is modest and we were unable to study the subtle distinctions among the sub-groups on gender and age basis.

REFERENCES

[1]. Hausdroff . et.al., "GAIT DYNAMICS, FRACTALS AND FALLS: FINDING MEANING IN THE STRIDE-TO-STRIDE FLUCTUATIONS OF HUMAN WALKING," National Center for Biotechnology Information, 1995

[2]. Scafettaa N, Griffin L and West B J (2003). Holder exponent spectra for human gait. Physica A 328: 561-583

[3] Aziz W. (2006). Variability Analysis of ysiological Signals using Non Linear Time

Series Analysis Techniques. (Unpublished) PhD Thesis PIEAS Pakistan 150 pp.

[4] Peng, C. K., Hausdorf, J., & Goldberger, A. L. (2000). Fractal mechanisms in neuronal control: human heartbeat and gait dynamics in health and disease. In J.Walleczek (Ed.), Self-organized biological dynamics and nonlinear control (pp. 66–96). Cambridge: Cambridge University Press

[5] Costa,., Penga, C. K., Goldbergera, L., & Hausdorf, J. M. (2003). Multiscale entropy analysis of human gait dynamics. Physica A, 330, 53–60.

[6] A.Q. Abbasi ., W.A .Loun, "Symbolic time series analysis of temporal gait dynamics. Journal of signal processing systems", 2014, 74, 3, 417−422

[7] Costa M, Goldberger A L and Peng C K (2002b). Multiscale entropy analysis of complex physiologic time series. Physical Review Letters 89: 068102. doi: 10.1103/PhysRevLett.89.068102.

[8] A. Qumar ., W. Aziz,., S. Saeed,., I. Ahmed,., and L. Hussain., "Comparative study of multiscale entropy analysis and symbolic time series analysis when applied to human gait dynamics.", 2013 International Conference on Open Source Systems and Technologies (ICOSST), 2011, 126−132

[9] Pincus, S. M. (1991). Approximate entropy as a measure of system complexity. Proceedings of the National Academy of Sciences of the United States of America, 88, 2297–2301.

[10] Goldberger A L, Namaral L A, Glass L, Hausdorff J M, Ivanov P C, Mark R G, Mietus J, Moddy G B, Peng C K and Stanley H E (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals Circulation 101: 215-220.

.