MEASURING THE MATURATION OF GAIT IN CHILDREN USING SYMBOLIC ENTROPY ANALYSIS

Samira Aziz
Department of Computer Sciences
Allama Iqbal Open University
Islamabad, Pakistan
paramount_sam@yahoo.com

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

Muhammad Shiraz
Department of Computer Sciences
Federal Urdu University
Islamabad, Pakistan
muh_shiraz@yahoo.com

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

***Abstract*— The development of neuromuscular control and maturation of locomotor function can be understood by analyzing the gait dynamics of children of different ages. The loco-motor system is made up of billions of interconnected and highly coupled cells whose response is non-linear and the physiological signals generated by such nonlinear dynamic systems are non-stationary and more often demonstrate complex structures which cannot be interpreted easily. Researchers has proposed different analysis techniques to extract information about the human gait dynamics. The linear methods cannot fully describe the underlying complex dynamics and thus may miss potentially useful information. For analysis of these complex and natural systems, non-linear signal processing methods are more appropriate, and these techniques may yield better results as compared to the linear techniques. In this paper, we used threshold based symbolic time series analysis method for the stride interval time series of 50 children (25 boys and 25 girls). The stride interval obtained from signals produced from walking of children of various ages were transformed into symbol sequences. Then Normalized Corrected Shannon Entropy (NCSE) was calculated from the symbol sequences as complexity value. Wilcoxon- rank-sum test (Mann–Whitney–Wilcoxon (MWW) test) was used to find the significant difference between the groups. After analyzing we concluded that the stride interval time series of children aged 3–4 years is less complex than the children of age 6–7 years, and 11–14 years), and gait dynamics get mature at the age of 11 to 14 years.**

Keywords*— *Maturation of Gait; Gait Analysis; Normalized Corrected Shannon Entropy; Stride interval; Symbolic Entropy

I. INTRODUCTION

The human body system is made-up of various complex physiological systems e.g. heart, brain, muscles, gait, etc. Signals are produced by the body during activity of these physiological systems. These signals could be electrical, chemical or acoustic in origin. These signals contain information about the condition and functioning of the respective system. It is often helpful to analyze these signals and dig out the information regarding the condition of the physiological system and discover its malfunctioning. In this paper, human locomotion system (gait dynamics) of children of various ages is analyzed. The ability to move from one place to another is called locomotion. Central nervous system controls and regulates the human locomotion function, and it is coordinated by the musculoskeletal system [1]. Gait analysis is the systematic measurement, description, and assessment of quantities that characterize human locomotion [2].

In general, a toddler is capable to sit at the age of about 6 months, able to move by crawling at the age of 9 months, and able to stand the age of 1 year. At this age the kids walk with immature control of posture [3]. They have unbalanced walking patterns and volatile posture because at that stage they are learning how to walk. It has found an immature motor control from one stride period (the time between two consecutive heel strikes of the same foot) to another, when children learn to walk [4]. When children jump from childhood to adulthood, the stride variability is decreased due to the development of motor skills and it results into a mature gait [5]. Some early studies by Beck et al. [6] shows that the relationships between temporal and distance parameters in children become stable by the age of 4. Sutherland [7] stated that the gait becomes relatively mature at the age of 4 years and walking patterns becomes more stable at this age. 230 subjects were studied by Norlin et al. [8] having ages from 3 to 16 years, and presented that the gait had not become mature until 8 years old. Another study by Menkveld et al. [9] shows that the temporal gait patterns presented apparent maturity but still continued developing in adolescence. Even though the research has not been made to fully distinguish changes in gait maturation within-subject changes yet, the discussions based on different observations have resulted a vast area to study and examine the matureness in gait [4]. Various studies has been made to analyse the human gait dynamics. Hausdorff [5] applied different variability analysis techniques (coefficient of variance, standard deviation, and fluctuation magnitude,) to measure the gait maturation of children. In anther study, Hausdorff [10] examined about the stride to stride variability during movement of the body. It offers a very good approach of quantifying locomotion & measures its variations with respect to disease and aging. He also monitored the healing interventions & rehabilitation.

Zhong [11] measured the stability of locomotors based on probability density functions & described characteristic of gait patterns in normal children using different methods based on statistics.

Aziz & Arif [12] applied threshold based symbolic entropy (SyE) to characterize the dynamics of healthy and pathological subjects suffering from neurodegenerative diseases. They showed that stride interval time series of healthy subjects showed higher complexity as compare to neurodegenerative (ALS, Huntington and Parkinson) diseased subjects at various threshold values. Anees &

Wajid [13] used symbolic entropy analysis method on gait signals (normal walking, slow walking, fast walking, and metronmically controlled walking) and found out that the normal walking signals were more complex than the fast and metronmically controlled signals. In a subsequent study by Qumar et.al. [14], multi-scale entropy analysis and symbolic time series analysis methods were compared and it was found that the later technique (symbolic time series analysis) out performs the former in discriminating normal and metro-domically paced walking.

Goshvarpour [2] used non-linear techniques for analyzing gait signals. They used phase space, Hurst exponent, lyapunov exponents and Poincare plots for analyzing gait signals. They analyzed human gait signals in slow and fast paces. They found that complexity of gait signals was decline during slow and fast gait.

The linear methods cannot fully describe the underlying complex dynamics and thus may miss potentially useful information. For analysis of these complex and natural systems, non-linear signal processing methods are more appropriate, and these techniques may yield better results as compared to the linear techniques. In this study, we have used non-linear complexity based method i.e. Symbolic entropy(SyE) technique to analyze the gait signals of children of various ages.

II. MATERIAL AND METHODS

Usually it is assumed that everything in the universe changes from order to disorder eventually, and entropy is the measurement of that change. In case of a dynamical system like gait, heart or any other physiological system, it is rate at which the information is produced. Entropy based techniques are worthwhile to compute complexity of physiological signals which are noisy or of short duration; especially the stride rate and heart rat time series datasets [13]. There are many entropy based techniques available in literature for studying the gait signals with different merits and demerits.

Threshold based symbolic entropy analysis was proposed by Aziz and Arif [12] for analysis of the time series obtained from gait signals. In this method, the original time series is converted into a sequence of symbols and each value in the original time series is represented by a symbol. This process is called data symbolization. Different approaches has been proposed for data symbolization, the most popular approach is to convert the values into two symbols ‘0’ and ‘1’ (number of symbols used to represent the original values is called quantization level (QL)) on basis of some threshold value . This transformed series (original value to some symbol) is called symbol series and it will contain the same number of data points as it was in original. Next step is the symbol sequence formation. In this step, sequence of consecutive symbols of a certain length greater than one(called word length (WL)) are combined and a new series is formed called word series. The symbols are taken using overlapping window.

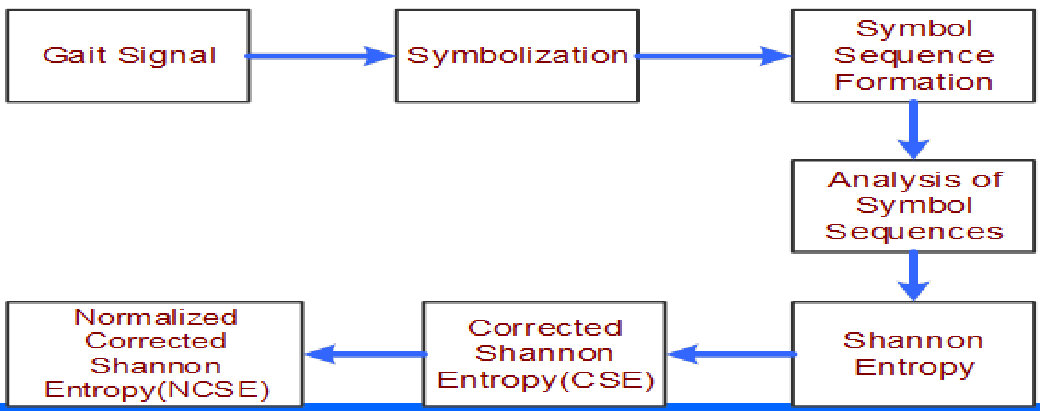

Fig 1. Conceptual Model of Proposed Technique

The length of this word series will be "N-WL+1" where N is total length of original time series. Finally code series is generated by converting the symbols into decimal values and histogram of the code series is plotted, and it is analyzed using some technique. The whole procedure explained above is illustrated with example by Qumar et.al [14]. The conceptual diagram of the proposed technique is presented in Fig. 1. The mathematical details of the SE, CSE, and NCSE are as under:

$$\text{Shannon Entropy (H)} = -\sum p_i log_2 p_i \qquad (1)$$

Where p shows the histogram of the symbol-sequence frequencies.

$$CSE = H + \frac{C_r - 1}{2Mln2} \qquad (2)$$

Where M is the total volume of symbol sequence and Cr is number of occurrence of words

CSE will be maximum for some word-length WL and quantization-level 'QL' (Possible number of symbols in a symbol sequence), when all words occur with identical distribution in the data series. It is given below:

$$\text{CSE}^{\max}(WL, QL) = -\log_2\left(\frac{1}{\text{M}}\right) + \left(\frac{\text{M}-1}{2\text{M}\ln 2}\right) \qquad (3)$$

The value of maximum CSE will not be analogous for two different word-lengths. When the value of L increases than value of M also increases. In response to changes in these values, the value of CSE will also increase. We cannot compare the values of CSE for two different word lengths, at same threshold value and quantization level. To resolve this dilemma, "Normalized Corrected Shannon Entropy (NCSE)" is proposed by Aziz and Arif (2006). In NCSE we are able to compare two different word lengths at same threshold and quantization level. The NCSE of word length WL and quantization level QL is given below:

$$NCSE(WL, QL) = \frac{\text{CSE}(WL, QL)}{\text{CSE}^{\max}(WL, QL)} \qquad (4)$$

The value of NCSE will be between 0 & 1 for every word-length and their quantization-level. We will use this value for measuring the complexity of the gait.

A. Data Sets

In this study we used dataset from the publicly available domain (www.physionet.org). The dataset is consists of fifty boys and girls of various ages with equal number of boys and girls. The datasets was classified into three main age groups: (1) 11 children of age three to four years, (2)20 children of age six to seven years (3) 12 children of age eleven to fourteen years. A small amount of some other subjects are also in the data set: 3 children of age five years, 1 children of age eight year and 3 children of age ten years. However, we did not use these small amount of subject.

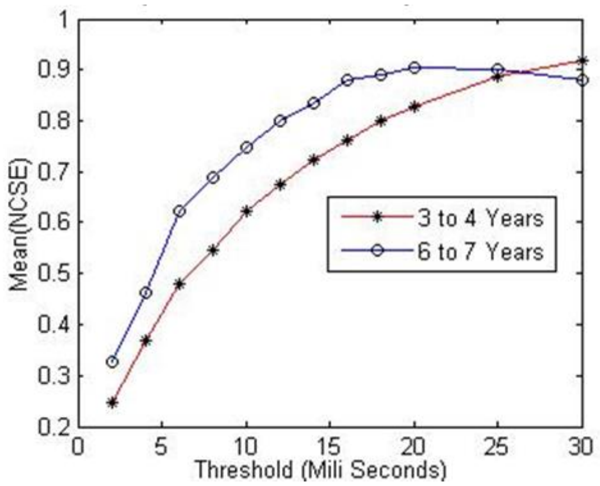

Fig 1. Results of Symbolic Analysis for age groups 3 to 4 years and 6 to 7 years

III. RESULTS AND DISCUSSION

The stride interval time series data corresponding to three different age groups (3 to 4 Years, 6 to 7 Years and 11 to 14 Years) was analyzed by using threshold based symbolic entropy method. For data symbolization, mean subtracted method as explained by Anees & Wajid [13] was used. We used quantization level 2 (symbols 0 and 1) at various threshold values.

Fig. 2 shows the comparison of symbolic entropy values (NCSE) for children of age group 3 to 4 years and 6 to 7 years. We can observe from the results that the complexity (NCSE value) of stride interval time series data obtained from the children of age group 6 to 7 year is greater than that of the children with age 3 to 4 years at shorter threshold values i.e. from 2ms to 25ms. It depicts that the gait dynamics of children of age 6 to 7 years is more mature than the gait dynamics of the children of age 3 to 4 years.

We next applied the SyE technique to the data sets derived from children of age 3 to 4 years and 11 to 14 years. The results are shown in Fig. 3. It can be observed from the figure that the complexity of the dataset of 11 to 14 years is greater than the complexity of the dataset of 3 to 4 years.

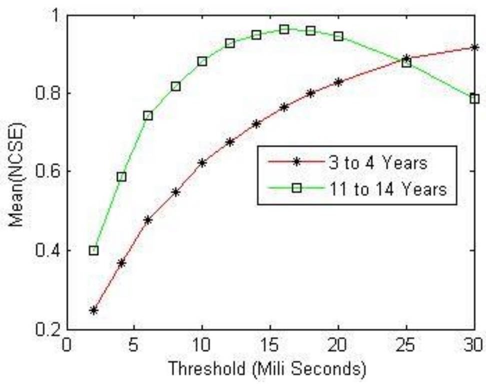

Fig 3. Results of Symbolic Analysis for age groups 3 to 4 years and 11 to 14 years

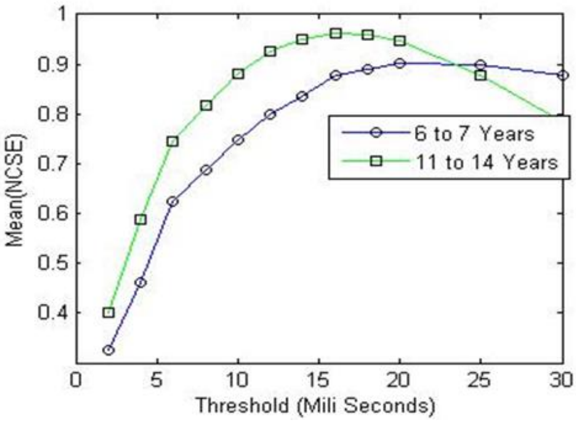

Fig 4. Results of Symbolic Analysis for age groups 6 to 7 years and 11 to 14 years

It compliment our hypothesis that the children of age 11 to 14 years have more mature walking pattern than the children of 3 to 4 years. The time series data of children of age 6 to 7 years was also compared with the time series data of children of age 11 to 14 years. The results are shown in Fig. 4. The results shows that the gait dynamics of the children of age 11 to 14 years are more complex than that of the children of age 6 to 7 years.

The overall results after applying symbolic time series analysis to data under various ages are shown in Table 1. The results presents that the maximum value of NCSE was "0.96±0.01" at threshold values 16ms and 18ms for the children of age 11 to 14 years. At other threshold values less than 25ms the NCSE values are higher for 11 to 14 years age group than other two age groups.

It confirms our idea that with the increase in age, the complexity of the time series increases hence it results into more mature gait dynamics and stable walking patterns.

To check the significant difference between the different age groups Wilcoxon- rank-sum test (Mann–Whitney–Wilcoxon (MWW) test) was applied to check weather different age groups are getting differentiated or not. P-values were calculated and the results are shown in Table II.

The findings showed that the p-values that describe the significance of results are lower when the NCSE values of 3 to 4 year age group was compared with the NCSE values of the age group 6 to 7 years or 11 to 14 years than the p-values of comparison of 6 to 7 year age group with 11 to 14 year age group. Although at almost all threshold values less than 25 ms, the result are statistically significant because we got p-value less than 0.05, b the maximum degree of separation (minimum-value i.e. 0.0001) between the agree groups was available at threshold values 10 ms to 18 ms while comparing 3 to 4 years age group with 11 to 14 years age group. It shows that, there is a significant difference between the NCSE values of these two age groups, and these both groups belongs to different

TABLE 1. VALUES OF NCSE MEAN ± SD OF DIFFERENT AGE GROUPS

Threshold(ms)	NCSE (Mean ±SD)		
	3 to 4 years	*6 to 7 years*	*11 to 14 Years*
2	0.25±0.02	0.33±0.02	0.4±0.03
4	0.37±0.03	0.46±0.03	0.59±0.03
6	0.48±0.03	0.62±0.03	0.74±0.03
8	0.55±0.03	0.69±0.04	0.82±0.02
10	0.62±0.03	0.75±0.04	0.88±0.02
12	0.67±0.03	0.8±0.04	0.93±0.02
14	0.72±0.03	0.84±0.03	0.95±0.02
16	0.76±0.03	0.88±0.03	0.96±0.01
18	0.8±0.03	0.89±0.03	0.96±0.01
20	0.83±0.03	0.9±0.03	0.95±0.01
25	0.89±0.02	0.9±0.03	0.88±0.02
30	0.92±0.02	0.88±0.03	0.78±0.03

group of data.

SyE analysis method has been used by the researchers in different fields of complexity analysis. All the studies show that the normal spontaneous walking has more complexity than walking affected with some pathology or due to age factor. Previously, some linear techniques were being applied on this data to measure the matureness of gait in children. But the built-in complexity and non-linearity in o the human gait system may not allow linear techniques to give more accurate results. Therefore we used Symbolic Entropy technique (A nonlinear symbolization based technique) to differentiate different age groups. This method was applied to 3 different data groups corresponding to 3 different age groups. The findings indicated that the time series data of adults (11 to 14 years) have larger complexity value than children of age 3 to 4 years or 6 to 7 years at certain range of threshold. Hence, it shows that, with the increase in age the walking patterns become more stable and the children get mature gait dynamics. In our case we got optimal values at threshold 16ms to 18ms.

IV. CONCLUSION

Complexity analysis is an interesting filed of research. Signals obtained from biological systems have built-in complexity due to a very coupled and integrated system. Complexity analysis can be used to check the pathology in some system because whenever any disease affect the system, the complexity will reduce. The low value of complexity may be a representative of some disease or it may be aging factor. This study shows the use of symbolic entropy analysis method to measure the matureness of gait dynamics in children and we concluded that the complexity of the stride interval data of the children with age 11 to 14 years is greater than the complexity of stride interval data of children with age 3 to 4 years of 6 to 7 years. We can relate this value of complexity with matureness level. An increased complexity value indicates more mature gait while decrease in complexity value indicates less mature gait.

Lastly, It was found that our findings complement the previous studies reported by different researchers [5, 6, 7,8]. However because of non-linearity of the physiological systems, the complexity based values may perform well as compare to other nonlinear techniques.

TABLE II. P-VALUES of COMPARISON NCSE VALURES FOR DIFFERENT AGE GROUPS

Threshold(ms)	p Values		
	3 to 4 years Vs 6 to 7 years	*3 to 4 years Vs 11 to 14 years*	*6 to 7 years Vs 11 to 14 years*
2	0.0244	0.0028	0.054
4	0.041	0.0004	0.0061
6	0.0012	0.0001	0.0086
8	0.0041	0.0002	0.0096
10	0.0053	0.0001	0.0048
12	0.0041	0.0001	0.002
14	0.0077	0.0001	0.0022
16	0.0047	0.0001	0.0167
18	0.0068	0.0001	0.0645
20	0.0088	0.0028	0.4715
25	0.4208	0.5588	0.167
30	0.445	0.0028	0.0042

References

[1] N. Xiang, . S. Cai, . S. Yang, S.; Z. Zhong, Z.; Zheng, F.; He, J.. Statistical Analysis of Gait Maturation in Children Using Nonparametric Probability Density Function Modeling. Entropy 2013, 15, 753-766.

[2] A. Goshvarpour., & Goshvarpour, A. Nonlinear analysis of human gait signals. International Journal of Information Engineering and Electronic Business, 2012 4(2), 15.

[3] D. Sutherland, The development of mature gait. Gait Posture 1997, 6, 163–170.

[4] [S. J. Hillman., B.W. Stansfield, , A.M. ichardson, , J.E. Robb,. "Development of temporal and distance parameters of gait in normal children.", Gait & Posture, 2009, 29(1), 81-85.

[5] J. M. Hausdorff, L. Zemany, C.K. Peng, A.L Goldberger,.. "Maturation of gait dynamics: stride-to-stride variability and its temporal organization in children." Journal of Applied Physiology, 1999, 86(3), 1040-1047.

[6] RJ Beck,.; T.P Andriacchi, ; K.N Kuo,.; R.W/ Fermier, ; Changes in the gait patterns of growing children. J. Bone Joint Surg.Am. 1981, 63, 1452–1457.

[7] D. H Sutherland,., R.I.C.H Olshen,., Cooper, L., & Woo, S. L.. "The development of mature gait.", J Bone Joint Surg Am, 1981, 62(3), 336-53.

[8] R. Norlin, .; P. Odenrich, .; Sandlund, B. Development of gait in the normal child. J. Pediatr. Orthop. 1981, 1, 261–266.

[9] S.R. Menkveld, .; E.A. Knipstein,; J.R. Quinn, . Analysis of gait patterns in school-aged children. J. Pediatr. Orthop. 1988, 8, 263–267.

[10] [10] J.M Hausdorff,. Gait variability: methods, modeling and meaning. Journal of NeuroEngineering and Rehabilitation, 2005, 2, 19. http://doi.org/10.1186/1743-0003-2-19

[11] [Wu, Y.F.; Zhong, Z.T.; Lu, M.; He, J. Statistical analysis of gait maturation in children based on probability density functions. Proc. 33rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. 2011, 1652–1655.

[12] W. Aziz , and M. Arif. Complexity analysis of stride interval time series by threshold dependent symbolic entropy, European Journal of Applied Physiology, 2006, vol. 98(1), pp. 30-40.

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

[14] 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