engineering proceedings

Proceeding Paper

Gait Segmentation and Grouping in Daily Data Collected from Wearable IMU Sensors †

Zhuoli Wang [1], Chengshuo Xia [2] and Yuta Sugiura [3,*]

[1] Keio University; zhuoli.wang@keio.jp
[2] Xidian University; xiachengshuo@xidian.edu.cn
[3] Keio University
* Correspondence: sugiura@keio.jp

† Presented at the 10th International Electronic Conference on Sensors and Applications (ECSA-10), 15–30 November 2023; Available online: https://ecsa-10.sciforum.net/.

Abstract: Gait analysis plays a vital role in medicine as it can help diagnose illness, monitor recovery, and measure physical performance. Related work in gait analysis has primarily utilized laboratory data due to its inherently low noise and ease of preprocessing. Daily data, gathered through wearable sensors, can also significantly impact medical care. Nonetheless, working with such data poses numerous challenges. This paper proposes an algorithm to solve the problems associated with gait segmentation in daily data obtained by inertial measurement units (IMUs) on wearable devices. The proposed algorithm can handle time-series data collected by wearable IMU sensors, including noise and different gaits. The proposed algorithm within this paper can identify the start and end points of each gait segment within the time series, and the same type of gait will be grouped together.

Keywords: gait analysis; step segmentation; wearable sensors; inertial measurement unit (IMU)

check for updates

Citation: Wang, Z.; Xia, C.; Sugiura, Y. Gait Segmentation and Grouping in Daily Data Collected from Wearable IMU Sensors . *Eng. Proc.* **2023**, *56*, 0. https://doi.org/

Academic Editor(s):

Published: 15 November 2023

1. Introduction

Gait detection technology plays a vital role in the field of medical health [1]. Gait serves as a valuable indicator of an individual's physical well-being, encompassing substantial information regarding their overall health [2]. Monitoring people's gait in daily life facilitates the timely identification of alterations in their underlying health status. The collection of human gait data can be achieved through the utilization of inertial measurement unit (IMU) sensors. With the increasing prevalence of smart wearable devices, IMU sensors are progressively gaining traction within everyday life.

Gait segmentation using daily gait data collected by IMU poses several difficulties and challenges. Firstly, the daily gait data often contains motion information that is not solely related to gait. Secondly, individuals employ various walking patterns throughout the daily life. Additionally, each person exhibits unique movement characteristics. The complexity and variability of human gait across individuals and activities further increase the difficulty of accurately segmenting the gait cycle.

The IMU is commonly employed for gait data collection. The segmentation of gait data is conducted by considering specific characteristics and regular patterns observed in human gait data obtained by the IMU [3–5]. Several studies have implemented a methodology for segmenting continuous gait data from a dataset according to the features of a set of stride [6]. This technological approach has been utilized in evaluating individuals with Parkinson's disease, enabling them to conduct the test in a home environment without the need for direct medical supervision [7].

The current existing solutions for gait segmentation mostly focus on processing data that only includes single gait patterns, mainly handling walking gait data. Additionally, these methods rely on data preprocessing techniques. Most of these approaches are based on laboratory data rather than real-world daily life data.

The processing of daily gait data presents some challenges, because the data will contain multiple gait types and noise. In order to overcome these challenges, a method for processing individuals' daily gait data is proposed in this paper. This proposed method effectively addresses the issue of noise within the data and achieves segmentation for different gait patterns, all without the need for pre-set template gaits or intensive model training. The output is a set of segmented gaits, categorized according to their respective gait types. It is worth noting that what this proposed method does is to find behavioral patterns that appear frequently multiple times throughout the time series. This study proposed a method for processing data that contains multiple gait types simultaneously, including walking, running and going up and down stairs. The results are presented for separate processing of data containing each of the four gaits mentioned above. Moreover, this study analyzes the influence of varying sensor numbers and sensor placements on the segmentation outcome.

2. Methods

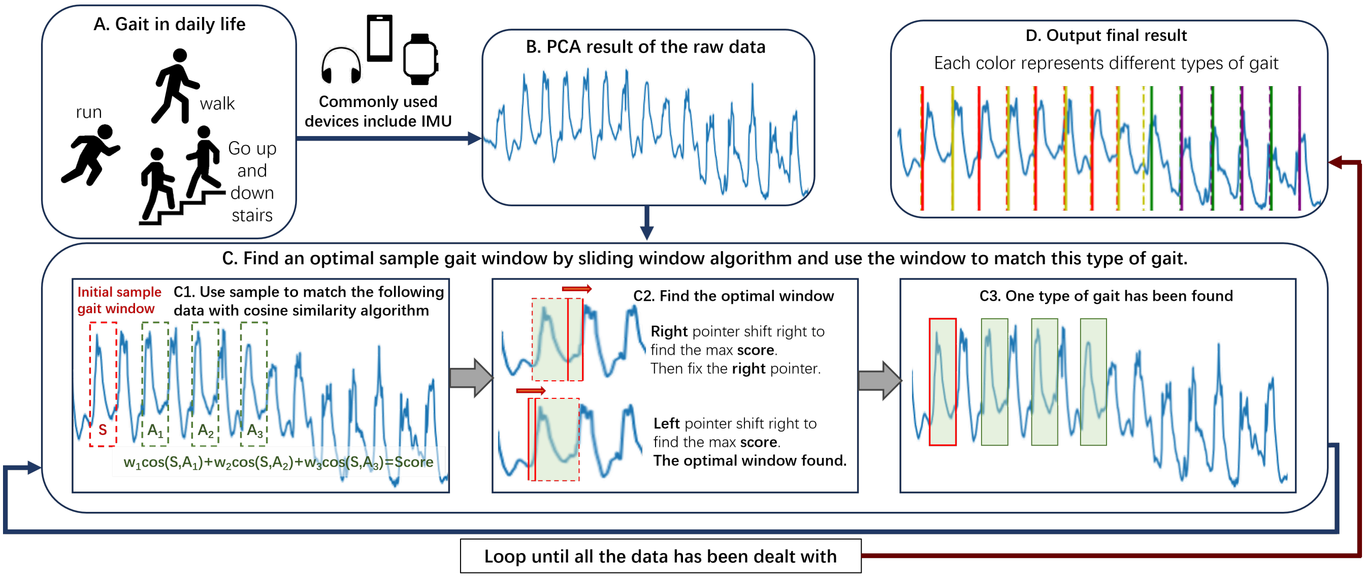


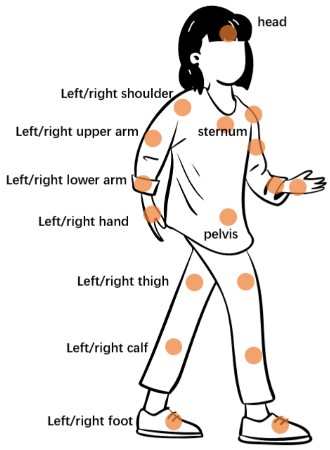
Figure 1. This figure illustrate the main processing step of the proposed method. The first step is collect data and use PCA method [8] to reduce dimension. Then sliding window algorithm is used to find the optimal sample gait window which can be used to match the same type of gait segmentation from the data set. In this experiment, the data includes walking, running, and stair-climbing. Please note that the graph displayed represents a small portion of the experimental data, and thus, the plotted line does not include all types of gaits mentioned before.

2.1. Data Collection

Four young people (2 males and 2 females, average: 25) were recruited to collect data. All participants signed the agreement and complied with the ethical review.

Participants in the study were outfitted with IMUs (MTw; Xsens Technologies Inc.), purposed to gather the necessary acceleration and angular velocity data. A total of seventeen sensors were placed in specific places on the person's body which can be seen in the Figure 2a.

Participants were instructed to engage in walking, running, and going up and down stairs, each at a self-determined comfortable pace. These four distinct gaits were selected to encompass a wide range of typical human locomotion patterns in their daily life. Each data point collected during each time frame included measurements for velocity, acceleration, angular velocity and angular acceleration at a frequency of 60 Hz. Each data point encompasses measurements for each IMU in the x, y, and z directions of global coordinates. Specifically, the x-axis indicates the positive direction towards the local magnetic (North), the y-axis aligns with the right-handed coordinates (West), and the z-axis indicates to be positive when pointing upwards [9] .



(**a**)

(**b**)

Figure 2. The (**a**) displays the position of the 17 IMUs [10]. The two photographs in the (**b**) illustrate the participants after the completion of wearing the devices.

2.2. Data Processing

First, the modules of velocity, acceleration, angular velocity and angular acceleration in the horizontal plane of the obtained data are determined. This can be accomplished by employing the following formula: $|V| = \sqrt{V_x^2 + V_y^2}$. The input data for the four modules, along with their respective vertical dimensions, will be considered in this study.

This is due to the fact that for human gait, vertical speed, acceleration and angular velocity are important features [11]. In the field of gait analysis, the specific direction of human locomotion holds no significance. When individuals move with the same gait in different directions, these movements should not be considered as different gaits. Subsequently, the dataset was standardized to ensure that the features have comparable scales [8]. This was done to ensure that each feature carries equal importance during the subsequent step of principal component analysis (PCA). Then, the PCA method [8] was to reduce the dataset containing all IMUs' information to one dimension.

The program will remove the data in static state, based on the given value. If a continuous sequence of data points at static state, reaching a certain threshold (in this experiment, it is set as 15), this segment of data will be excluded.

2.3. Initialize Sample Gait Window and Use It to Match the following Data

The initial sample gait window is selected from the beginning of the current valid data, based on the given parameter (in this experiment, it is set as 18 time frames). The sample gait window gradually slides backwards and is compared with the valid data using the cosine similarity algorithm. When the calculated similarity exceeds the given threshold (in this experiment, it is set as 0.8), the corresponding data will be marked as matching and the corresponding similarity value will be recorded. This will be used to score the current sample gait window. Previously compared data is not re-evaluated to avoid duplication.

2.4. Find the Optimal Sample Gait Window

The proposed method employs the sliding window algorithm to identify the optimal sample gait window. The continuous occurrence of matched segmented data with high similarity values indicates that the sample gait window demonstrates a higher degree of accuracy or quality, resulting in a higher score. The right pointer traverses a given range, and a score is computed in accordance with the aforementioned method. Finally, the position of the right pointer is located at the point that receives the highest score. The same process applies to the left pointer. By determining the positions of the left and right pointers, the current optimal sample gait window can be determined. The threshold of range is set to prevent windows with excessively large size. The threshold is determined as twice the average step length base on experiment. Furthermore, if the initial window contains content that is not related to any gait such as noise, it will be unable to match similar data, resulting in a significantly low score. The program will filter out the sample windows with low score.

2.5. One Group of Gait Was Found

The data that matches with the optimal sample gait window will be considered as a group of gait. They will be labeled as invalid during subsequent iterations, in order to prevent redundant marking.

The previous two steps will be iterated to search for additional gaits. The iteration will end when all data has been traversed. Then output obtained gait groups.

3. Results and Discussion

The ground truth is annotated by the 3D model movements generated by analyzing the data of human movements through the supporting software of the IMU used in this study. The visualization results of ground truth are shown in the Figure 3.

A visualization of the results is shown in Figure 4. It illustrates the results obtained by the proposed method. The gait segmentation results are represented on the figure. It can be noticed that each detected gait segmentation is marked with its starting and ending points. The corresponding gait type is also labeled in Figure 4 according to the ground truth.

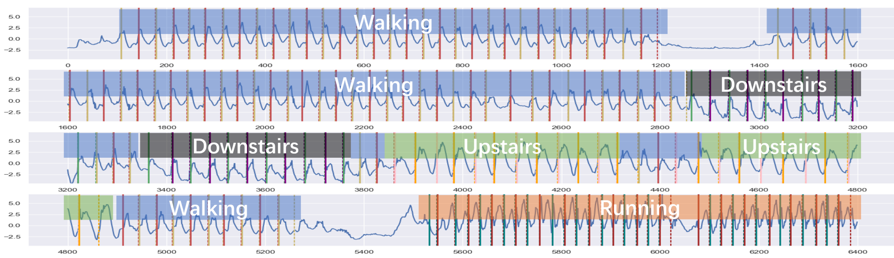


Figure 3. Here is the figure of ground truth. The time series gait data post-PCA is denoted by the blue line on the graph. The start of a gait is indicated by a solid line, while the termination of a gait is indicated by a dashed line. The different color of lines represents different groups of gaits.

Figure 4. Here is the figure of the result. The legend refers to the figure above.

The subsequent figures show the validation results obtained from different combinations of sensors across various types of datasets, includes multiple distinct gaits or singular gaits exclusively. The validation results include accuracy and rand index. The accuracy of gait segmentation was computed using.

$$Accuracy = \frac{Accurate \quad segmentation \quad number}{Ground \quad truth \quad segmentation \quad number} \tag{1}$$

The result segmentations were compared with the ground truth gait segmentation and were marked as accurate. Rand index is a method used to compare the similarity of two clusters [8]. A value closer to 1 denotes a higher level of similarity between the compared clusters. In this study, the rand index was employed to assess the resemblance between the grouping outcomes attained by the algorithm and the ground truth.

The outcomes obtained from the dataset contains four distinct gaits are presented as follows. Across various sensor combinations, the accuracy values range from 0.82 to 0.90, while the rand index is between 0.76 to 0.83. It is notable that that among all sensor combinations, the combination of three IMUs on the feet and head has the best results, with accuracy and Rand Index being 0.9 and 0.83 respectively.

The "corresponding equipment" in the table represents the IMU put at specific positions during the experiment, which may correspond to devices used in daily life.

Body Part	Corresponding Equipment	Number of sensors	Accuracy	Rand index
Double feet	IMUs on feet	2	0.89	0.78
Double feet + head	IMUs on feet and ear phones	3	0.9	0.83
Double feet + Right hand	IMUs on feet and smart watch	3	0.83	0.76
Double feet + Left hand	IMUs on feet and smart watch	3	0.82	0.76
Double feet + Pelvis	IMUs on feet and IMUs in the phone put in the user's pocket	3	0.88	0.77

Figure 5. Testing result for all types of gait used with accuracy and rand index.

The validation results for data segmentation, focusing on the dataset contains a single gait, are illustrated in the following Figure 6. Notably, the utilization of different sensor combinations leads to improved final outcomes.

Body Part	Corresponding Equipment	Number of sensors	Accuracy	Rand index
Double feet	IMUs on feet	2	0.97	1
Double feet + head	IMUs on feet and ear phones	3	0.97	1
Double feet + Right hand	IMUs on feet and smart watch	3	0.97	1
Double feet + Left hand	IMUs on feet and smart watch	3	0.97	1
Double feet + Pelvis	IMUs on feet and IMUs in the phone put in the user's pocket	3	0.97	1

Figure 6. Testing result for only walking gait used with accuracy and rand index.

The Table 1 below compares the proposed method in this paper with the existing alternatives. It can be found that the proposed method performs comparably well with the existing methods in terms of accuracy when dataset containing only walking gaits.

Table 1. Comparative analysis of our proposed methodology with the existing approaches.

Reference	Methodology	Result
Jens B et al. [12]	Subsequence Dynamic Time Warping	97.7% accuracy in walking gait segmentation
A. R. Anwary et al. [13]	Peak detection	95.47% accuracy in walking gait segmentation
Jagos H et al. [14]	Autocorrelation	96% accuracy in walking gait segmentation
Proposed method	**Matching gait by cosine similarity in sliding window**	**97% accuracy in walking gait segmentation.** **90% accuracy in mixed gait segmentation.**

4. Conclusions

This study introduces a practical methodology capable of segmenting and grouping daily gait data obtained by IMUs. Data containing acceleration, angular velocity, etc. of various parts of human body was collected. Then, the method proposed in this paper was validated. The results indicated that the method performed accordingly to expectations and demonstrated the feasibility of both gait segmentation and grouping. The findings showed a greater performance when the sensors placed on both feet and the head. It was observed that the separate gaits tended to offer better performance compared to mixed gaits in the same dataset.

In conclusion, this study proposes a method that establishes a foundation for analyzing daily gait patterns. It has potential applications in the detection of anomalies, monitoring of sports-related activities, and other areas. This includes the development of a more accurate pedometer for different gaits, saving time on manual data annotation for large datasets, and aiding in the rehabilitation monitoring of Parkinson's patients through gait segmentation and recording.

Author Contributions: Conceptualization, Z.W.; Formal analysis, Z.W.; methodology, Z.W.; software, Z.W.; writing—original draft preparation, Z.W.; writing—review and editing, C.X.; supervision, Y.S. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: Not applicable.

Acknowledgments: This work was supported in part by JSPS KAKENHI (grant numberJP21H03485) and JST SPRING (Grant Number JPMJSP2).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Anwary, A.R.; Yu, H.; Vassallo, M. Gait quantification and visualization for digital healthcare. *Health Policy Technol.* **2020**, *9*, 204–212.
2. Maki, B.E. Gait Changes in Older Adults: Predictors of Falls or Indicators of Fear? *J. Am. Geriatr. Soc.* **1997**, *45*, 313–320. https://doi.org/10.1111/j.1532-5415.1997.tb00946.x.
3. Jain, R.; Semwal, V.B.; Kaushik, P. Stride segmentation of inertial sensor data using statistical methods for different walking activities. *Robotica* **2022**, *40*, 2567–2580.
4. Gujarathi, T.; Bhole, K. Gait analysis using imu sensor. In Proceedings of the 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), IEEE, Kanpur, India, 6–8 July 2019; pp. 1–5.
5. Ullrich, M.; Küderle, A.; Hannink, J.; Del Din, S.; Gassner, H.; Marxreiter, F.; Klucken, J.; Eskofier, B.M.; Kluge, F. Detection of gait from continuous inertial sensor data using harmonic frequencies. *IEEE J. Biomed. Health Inf.* **2020**, *24*, 1869–1878.
6. Han, J.; Jeon, H.S.; Jeon, B.S.; Park, K.S. Gait detection from three dimensional acceleration signals of ankles for the patients with Parkinson's disease. In Proceedings of the IEEE The International Special Topic Conference on Information Technology in Biomedicine, Ioannina, Epirus, Greece, Citeseer, 2006; Volume 2628.
7. Ullrich, M.; Mücke, A.; Küderle, A.; Roth, N.; Gladow, T.; Gaßner, H.; Marxreiter, F.; Klucken, J.; Eskofier, B.M.; Kluge, F. Detection of unsupervised standardized gait tests from real-world inertial sensor data in Parkinson's disease. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2021**, *29*, 2103–2111.
8. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
9. Roetenberg, D.; Luinge, H.; Slycke, P. Xsens MVN: Full 6DOF human motion tracking using miniature inertial sensors. *Xsens Motion Technol. BV Tech. Rep.* **2009**, *1*, 1–7.
10. Xia, C.; Maruyama, T.; Toda, H.; Tada, M.; Fujita, K.; Sugiura, Y. Knee Osteoarthritis Classification System Examination on Wearable Daily-Use IMU Layout. In Proceedings of the 2022 ACM International Symposium on Wearable Computers, 2022; pp. 74–78.
11. Van Nguyen, L.; La, H.M. Real-Time Human Foot Motion Localization Algorithm With Dynamic Speed. *IEEE Trans. -Hum.-Mach. Syst.* **2016**, *46*, 822–833. https://doi.org/10.1109/THMS.2016.2586741.
12. Barth, J.; Oberndorfer, C.; Kugler, P.; Schuldhaus, D.; Winkler, J.; Klucken, J.; Eskofier, B. Subsequence dynamic time warping as a method for robust step segmentation using gyroscope signals of daily life activities. In Proceedings of the 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, Osaka, Japan, 3–7 July 2013; pp. 6744–6747.
13. Anwary, A.R.; Yu, H.; Vassallo, M. Optimal foot location for placing wearable IMU sensors and automatic feature extraction for gait analysis. *IEEE Sens. J.* **2018**, *18*, 2555–2567.
14. Jagos, H.; Reich, S.; Rattay, F.; Mehnen, L.; Pils, K.; Wassermann, C.; Chhatwal, C.; Reichel, M. Determination of gait parameters from the wearable motion analysis system eSHOE. *Biomed. Eng.* **2013**, *58*, 000010151520134241.