AUTOMATED MULTI-FEATURE SEGMENTATION OF TREADMILL RUNNING

Mathew Schwartz[1], Todd Pataky[2], CHUA Sui Geok Karen[3], ANG Wei Tech[4], Cyril J. Donnelly[4]

New Jersey Institute of Technology, Newark, NJ, USA[1]
Department of Human Health Sciences, Kyoto University, Japan[2].
Tan Tock Seng Hospital Rehabilitation Centre, Singapore[3].
Rehabilitation Research Institute of Singapore, NTU, Singapore[4].

The definition of gait events and phases have been well established in the literature through the use of qualitative movement descriptors. The repeatable, objective definitions of gait events and phases is the cornersone of sucess when performin a multi-center trial. A correlation-based multi-feature automated segmentation algorithm was developed and applied to treadmill running data. The features used were soley from 3D kinematic marker trajectory data, including generated features such as vectors between kinematic markers. The algorithm was compared against a trained tester who used visual inspection and threshold limits of the vGRF to segment stance. The automated segmentation approach was shown to consistently identify the same gait events as the trained tester, representing a significant time savings for the signal processing of large volume treadmill running data.

KEYWORDS: analysis, reliability, validity, rolling correlation, gait.

INTRODUCTION: With a trend towards the development of large volume, multi-centre, clinical movement databases within the Human Movement Sciences, large volume databases of various movements are beginning to emerge in the literature. Though a positive step forward for the Human Movement Sciences, it can be argued it is critical that the signal processing and modelling of these enormous datasets are automated and standardised. This will help ensure the data being used is reliable, and the manpower needed for its collection, processing and modelling can be performed by any research laboratory. We propose an automated algorithmic approach for the detection of gait events (i.e., endpoint identification), which in turn define gait regions of interest within any standard motion capture file (i.e., c3d). This will be accomplished with a rolling correlation of reference data to a target file.

Segmenting human motion data is a topic studied in a variety of fields such as robotics, biomechanics, and computer graphics (e.g., gaming and animation). In this paper, the term segmentation refers to a contiguous time series, which is a subset of another larger time series (Nguyen et al., 2010). Generally, segmentation is when the start and end points (commonly referred to as gait events within the biomechanics literature) of a desired subsequence of motion (Lin et al., 2016). In our work, this subsequence is from a reference file and the time series we are interested in segmenting is from a different target file. When using the segments, or regions of interest (ROI) for statistical analysis, accurate segmentation may be important, particularly when investigating ROI's with a brief duration (e.g., weight acceptance or impact phase of stance). To what resolution this accuracy must be defined before the statistical analysis of a ROI is influenced is currently unknown.

The importance of clinical gait analysis for the planning and monitoring of an athlete's or patient's rehabilitation pathways has led to the development of numerous automated segmentation algorithms. In many cases, the complete volume of data collected during testing is not used for analysis, particularly when the ground reaction forces (GRF) is used to segment phases of gait cycle (e.g., weight acceptance, stance, swing, flight). Some research has integrated foot-switches to capture direct information on foot contact time, regardless of force plate location (Agostini et al., 2013). Though an interesting approach for the segmentation of large volumes of data, it is not a robust solution as it requires the use of a specialised hardware, coupled with a bespoke algorithm. Therein, this approach does not translate to movement data collected without the device. Using, among other methods, cross-correlation to help detect peaks in acceleration signals, Yoneyama *et al.* (2013) found gait sequences with a reliance on periodicity in the data. In contrast, our work uses a rolling correlation window rather than cross-correlation and does not explicitly rely on gait cycles to identify similar movements from a given target or reference motion. Most recently, Kidziński *et al.* (2019) demonstrated a real-time algorithm with high accuracy in foot-contact and foot-off event detections by using deep neural networks trained on a large dataset. In comparison, our work aims to limit the need for large datasets as input for segmentation as manual segmentation of the dataset the first time is time consuming. Importantly, we aim to develop a method that only uses kinematic marker data, with no reliance on external transducers (i.e., foot-switch, GRF measures). To demonstrate the algorithm's segmentation validly, we compared the gait events defining the stance phase of treadmill running as defined by a trained tester, with the use of GRF traces (gold standard), versus those defined using 3D kinematic data and our rolling correlation algorithm.

METHODS: The motion capture data used was a subset of data from a previous published study (Stearne et al., 2014). Sixteen male runners (1.9±2.6m, 75±6.3kg, 22±3.7yrs) were recorded on an instrumented split-belt treadmill at 4.5ms^{-1} (Bertec, Columbus, Ohio). Participants ran naturally, and the cohort included both rearfoot (RFS) andforefoot (FFS) strikers. To delineate the events that defined the stance phase of running, the vertical ground reaction force vector (vGRF) recorded at 2,000 Hz, with visual inspection was used. Stance was defined as the labelled events within the files recorded from (Stearne et al., 2014). The vGRF was visually inspected within the software package Nexus (Vicon, Oxford, UK). Foot strike events were also externally verified by a trained independent tester via high-speed video (100 Hz) in the sagittal plane of the runner (Basler A602fc-2, Ahrensburg, Germany).

Of the 16 subjects (9 RFS/7 FFS) used for analysis, each trial contained between 2-3 full strides of treadmill running; there was 1 natural trial per subject. To demonstrate the segmentation capability of the algorithm the data was split into a reference and a target group. The reference group consisted of a randomly selected 3 subjects (1 RFS/2 FFS), with 3 observations each. From the 3 observations only 2 were used, which generated a total of 6 trials or observations, which were passed to the algorithm, and entered as the reference dataset for the rolling correlation window. The remaining 13 trials (1 trial from 13 participants, 5 FFS/8 RFS) were used as the targets to be segmented automatically by the algorithm. The reference movements are based on the trained tester's manually labelled events, with the start of stance delineated 'Right Foot Strike' and the end of stance delineated 'Right Foot Off'. As the motion capture data recorded at 200Hz, one-step (a frame in the data) in the rolling window correlation corresponds to 5ms of time.

Segmentation was performed using a rolling correlation window across multiple features of a group of reference data (more than 1) of the same motion against the same features in a trial that contains that desired motion. The process is inspired by shape-matching techniques in which a spline or curve is compared against another for similarity in shape. While the algorithm has parameters set initially, there are few settings required for this general comparison using the tool. A user can set custom ranges or use default values for the comparisons. The values are: 1) the step size for the rolling window (where 1 is a check at each frame and 2 is a check at every other frame) 3) the range to scale the reference data where 50% is half the original duration and 200% is twice as long as the original duration 4) the threshold for features to use in the correlation comparison.

The segmentation process is as follows: 1) Define a collection of observations corresponding to the same movement as a reference data group. 2) Use cubic spline interpolation to up-sample shorter segments of the reference group. 3) For all data, the pelvis direction in each frame is defined by the ASIS and PSIS markers. At each frame the lab coordinates are rotated to align with pelvis direction (unifies the trajectory data when comparing between trials). 4) Create additional features to marker trajectories by subtracting marker positions from each other, for every frame, to create vectors representing the change in relationship over the length of the movement. 5) Calculate average correlation of all features in reference data to itself and store results. 6) Iterate through each frame in the data to be segmented. 7) Scale reference data to varying window sizes. 8) Reduce reference features to a given cut-off from the feature correlations between the reference data items. 9) Average correlation of data window to each item in the reference group, multiplying by feature weight. 10) Find peaks of correlation over time and use clustering to identify which peaks are of the similar motion. 11) Store highest correlations and remove time-frame from possible segments.

After applying the kinematic rolling window approach to the 13 trials, the algorithm identified 37 unique segments. Each segment was composed of a starting frame, ending frame, and all frames in-between. The algorithm performance was assessed in two ways. First, we assessed how the algorithm compared to the trained testers definition of a segment. Segment definitions alignment was classified as either in agreement or not between the algorithm and trained tester. Second, the difference between the events defining stance identified by the 3D kinematics and algorithm, and the trained tester and vGRF were compared. Descriptive statistics included minimum, maximum, mean, and root mean squared error (RMSE) differences. The total duration of the segmented stance phase was also compared.

RESULTS: All segments identified by the trained tester matched the algorithm (i.e., true positive). There were no disagreements in segment definitions between the two methods. The minimum, maximum, mean, and RMSE errors of each segment for each subject is shown in Table 1. The last row in Table 1 shows mean errors and statistical differences of the four dependent variables.

Table 1: Difference in duration (ms) of segments identified by the algorithm and the expert. Each subject had multiple observations (2 or 3) in which the results are calculated.

Subject	Min	Max	Mean	RMSE
1	15.0	30.0	20.0	21.2
2	20.0	25.0	22.5	22.6
3	40.0	50.0	43.3	43.6
4	10.0	15.0	11.7	11.9
5	5.0	30.0	13.3	17.8
6	5.0	20.0	13.3	14.7
7	0.0	20.0	8.3	11.9
8	5.0	5.0	5.0	5.0
9	15.0	15.0	15.0	15.0
10	5.0	10.0	6.7	7.1
11	0.0	15.0	5.0	8.7
12	5.0	35.0	15.0	20.6
13	30.0	30.0	30.0	30.0
Overall	0.0	50.0	16.2	20.5

The events defining the start (i.e., foot contact) and end of stance (i.e., toe off) were also analyzed. Figure 1 shows a histogram of event time difference when identified by a trained tester and vGRF versus 3D kinematic data and the algorithm. Overall, results showed the algorithm with a mean value for start of stance earlier but with insufficient evidence to conclude consistency (-0.3±11.3ms; $p = 0.887$) and ending later (8.9±11.3ms, <0.001) (Mean±SD). While the maximum error in the duration of a segment from one of the observations was 50ms the average was only 16.2ms, which is approximately 3.25 frames of data. In a practical setting, this margin would be unlikely to change the results of a kinematic analysis, however future research is recommended (i.e., sensitivity analysis) to verify this statement.

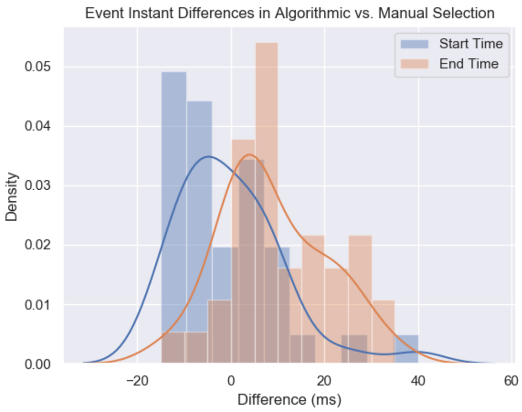

Figure 1: Histogram of difference between predicted (3D kinematics + algorithm) and original (trained tester + vGRF) events defining the stance phase of running. A –ve, value indicates prediction was early, +ve means prediction was late.

DISCUSSION: Using only 6 3D kinematic trials from 3 subjects, the algorithm successfully classified the stance phase of treadmill running, with a reasonable amount of error. Further testing is necessary to understand when the algorithm is likely to fail or incur larger errors and also to understand the practical implications of the identified errors. Due to the method used for segmentation, the resulting accuracy on other types of data would be dictated by the similarity between movements, more than the speed of the movements. After further testing and validation, the automated segmentation algorithm may be used to standardise the signal processing of large volume, multi-centre human movement databases. Importantly, the automated segmentation algorithm requires small volumes of 3D kinematic training, while achieving the segmentation performance similar to that of a trained tester expert using high fidelity vGRF traces in a laboratory setting.

CONCLUSION: The algorithm presented identified heel contact and toe off gait events, solely from 3D kinematic data with minimal error when compared to a trained tester using vGRF measures. The algorithm presented shows potential in becoming a standard for the segmentation of gait events for trials with long time series or experiments with large sample populations, within both laboratory (i.e., 3D kinematics) and field based settings (i.e., IMU's).

REFERENCES

Agostini, V., Balestra, G., & Knaflitz, M. (2013). Segmentation and classification of gait cycles. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *22*(5), 946-952.
Kidziński, Ł., Delp, S., & Schwartz, M. (2019). Automatic real-time gait event detection in children using deep neural networks. *PloS one*, *14*(1).
Lin, J. F. S., Bonnet, V., Joukov, V., Venture, G., & Kulic, D. (2016, November). Comparison of kinematic and dynamic sensor modalities and derived features for human motion segmentation. In *2016 IEEE Healthcare Innovation Point-Of-Care Technologies Conference (HI-POCT)* (pp. 109-112). IEEE.
Nguyen, N. T., & Bui, T. D. (2010, August). Automated posture segmentation in continuous finger spelling recognition. In *2010 3rd International Conference on Human-Centric Computing* (pp. 1-5). IEEE.
Stearne, S., Alderson, J., Green, B., Donnelly, C.J., Rubenson, J. Joint kinetics in rearfoot versus forefoot running: implications of switching technique. *Medicine & Science in Sports & Exercise*. ISSN: 0195-9131. 2014, 26(8): 1578-1587.
Yoneyama, M., Kurihara, Y., Watanabe, K., & Mitoma, H. (2013). Accelerometry-based gait analysis and its application to Parkinson's disease assessment—part 1: detection of stride event. *IEEE Transactions on neural systems and rehabilitation engineering*, *22*(3), 613-622.

ACKNOWLEDGEMENTS: We thank Agency for Science, Technology and Research (A*STAR), Nanyang Technological University (NTU) and the National Health Group (NHG) for funding this research (RRG3: 2019/19002). We thank Patrick Kwong, Ananda Sidarta, and Choon Kong Yap for their thoughtful feedback and conversations on definitions of endpoints and segments of different movements.