Contents

I. Programs

0.1. Cite as:

Cite the following sources if you use this code:

Clustering phase difference data in gait space
Kane, Suzanne Amador, Brooke L. Quinn, Xuanyi Kris Wu, Sarah Y. Xi, Michael F. Ochs, and S. Tonia Hsieh. "Unsupervised learning reveals rapid gait adaption after leg loss and regrowth in spiders." bioRxiv (2025): 2025-01.

Gait space metric
Wilshin, Simon, Paul S. Shamble, Kyle J. Hovey, Ryan Harris, Andrew J. Spence, and S. Tonia Hsieh. "Limping following limb loss increases locomotor stability." Journal of Experimental Biology 221, no. 18 (2018): jeb174268.

Phase difference calculation from foot motion data
Revzen, Shai, Samuel A. Burden, Talia Y. Moore, Jean-Michel Mongeau, and Robert J. Full. "Instantaneous kinematic phase reflects neuromechanical response to lateral perturbations of running cockroaches." Biological cybernetics 107, no. 2 (2013): 179-200.

Circular statistics calculations
Philipp Berens (2025). Circular Statistics Toolbox (Directional Statistics) (https://www.mathworks.com/matlabcentral/fileexchange/10676-circular-statistics-toolbox-directional-statistics), MATLAB Central File Exchange. Retrieved January 26, 2025.

0.2. Conventions for leg angles and phase differences

See this reference for more specifics: (Kane et al., 2025).

Units

The oscillation phase, ϕ_i, and phase differences, $\Delta\phi_i$, are computed in units of cycles, where 1 cycle = 2π rad = 360 deg.

Leg and phase difference numbers

We use the following convention for numbering the legs as i = 1, 2, … , N_{legs} to simplify computing the phase difference between adjacent legs. The index assumed in the code, i, starts with i = 1 at the left foreleg, then runs counterclockwise to the right foreleg.

For each case below, the oscillation phase differences are computed as $\Delta\phi_i = \phi_{i+1} - \phi_i$

Examples for intact legs

For a quadruped with all legs intact (N_{legs} = 4), the correspondences between *i* and leg are:

Leg	**Index i**
Left foreleg	1
Left hindleg	2
Right hindleg	3
Right foreleg	4

For a hexapod with all legs intact (N_{legs} = 6), the correspondences between *i* and leg are:

Leg	**Index i**
Left foreleg	1
Left midleg	2
Left hindleg	3
Right hindleg	4
Right midleg	5
Right foreleg	6

Example for missing legs

If a leg is missing, then it is skipped over in defining the index i. For example, for a quadruped is missing its left hindleg, N_{legs} = 3, the correspondences between *i* and leg are:

Leg	**Index i**
Left foreleg	1
Right hindleg	2
Right foreleg	3

1. Cluster multilegged phase difference data in gait space

Matlab code

GaitSpaceClusterer.m

Inputs

- Spreadsheet with phase difference samples (inputdata).csv (see below for simulated example data files)
- User selected number of legs, N_{legs}, which must agree with the N_{legs} -1 phase difference samples file (values of $3 \le N_{legs} \le 8$ currently supported; see below for how to modify for $N_{legs} > 8$).
- User selected number of clusters to assign data to.

Output

- (inputdata)_clust.csv file of original data with the assigned cluster number in the new rightmost column.

Description

Takes an input .csv file in which each row is a sample of N_{legs} -1 columns, each of which is a phase difference in cycles between adjacent pairs of N_{legs} legs (omitting one redundant pair; see above for numbering conventions). For each pair of samples, it computes the gait space distances between pairs of datasets and uses hierarchical clustering with Ward's linkage to create and display a dendrogram. The user then selects the number of clusters to assign samples to using the dendrogram. The final output is a .csv file containing the original data with a new column at the far right with the cluster assignments.

The input .csv file can also have additional columns with numerical or string entries for documentation. These could also be used to provide additional variables for alternative clustering approaches. (E.g., one could cluster based on both the gait space distances and the differences between kinematic variables, etc.)

Dependencies

1. Statistics and Machine Learning Toolbox – uses the following functions:
    a. `pdist`
    b. `linkage`
    c. `dendrogram`
    d. cluster
2. Circular Statistics Toolbox: Philipp Berens (2025). Circular Statistics Toolbox (Directional Statistics) (https://www.mathworks.com/matlabcentral/fileexchange/10676-circular-statistics-toolbox-directional-statistics), MATLAB Central File Exchange. Retrieved January 26, 2025.

2. Simulate phase differences for model gaits as input for clustering in gait space

Matlab code

GaitSimulator.m

Inputs

- Sample model gait spreadsheet (model gait).csv (see below for examples)
- User selected number of samples of each gait with added noise
- kappa (von Mises dispersion parameter)

Output

- (model gait)_sim.csv file of simulated model gait data with added noise.

Description

Use inputs a .csv file with various model gaits for a given number of legs and computes a dataset that includes a user-selectable number of samples for each model gait with added random noise. The user selects kappa, the parameter that determines dispersion in the von Mises distribution (similar to a normal distribution for circular statistics), to control how variable the phase differences generated are. Here, kappa = 2.9 is a value measured from the original study and higher (lower) values of kappa correspond to distributions with lower (greater) dispersion.

Dependencies

1. Circular Statistics Toolbox: Philipp Berens (2025). Circular Statistics Toolbox (Directional Statistics) (https://www.mathworks.com/matlabcentral/fileexchange/10676-circular-statistics-toolbox-directional-statistics), MATLAB Central File Exchange. Retrieved January 26, 2025.

3. Convert foot x,y coordinates vs time into oscillation phase differences

Matlab code

PhaseDiffCalculator.m

Description

Converts data in a format compatible with the tracking code DLTdv (Hedrick, 2008) into velocity along the y direction using a running quadratic fit over a user selected time window. Uses y and y velocity for each foot to compute and save the oscillation phase differences between adjacent legs, as defined above. (Sample input data for testing can be generated using *FootYGaitMotionSimulator.m*)

Inputs

- DLTdv compatible (corefilename)_data_xypts.csv spreadsheet of foot x,y coordinates (sample spreadsheets provided; see description of *FootYGaitMotionSimulator.m* for how to compute samples for testing)
- Video framerate (frame/s) (used in y velocity calculations)
- Time window in ms over which to compute the y velocity.

Output

- (corefilename)_phasediff.csv file of phase differences between each pair of adjacent legs for each frame.

4. Simulate foot x,y coordinates vs time given model gait oscillation phase differences

Matlab code

FootYGaitMotionSimulator.m

Description

Given an input model gait specified by a spreadsheet of oscillation phase differences between adjacent feet, computes y vs frame for user selected number of strides and stride period.

Inputs

- Sample model gait spreadsheet (model gait).csv (see below for examples)
- User selected number of stride periods to simulate
- User selected number of frames (samples) per stride period

Output

- DLTdv compatible (model gait name number of legs)_data_xypts.csv spreadsheet of foot x,y coordinates (can be used as input data to *PhaseDiffCalculator.m*)
- Plots of all y vs frame number.

5. Compute the gait space metric for > 8 legs

Matlab code

gaitspacemetricsolver.py

Description

Sample python code adapted from toroidalGeometry.py (Supplementary Information, (Wilshin et al., 2018)).

As written, the MATLAB script *GaitSpaceClusterer.m* supports between 3 and 8 legs (the gait space method is not informative for biped locomotion, for which a separate literature already exists).

If you wish to use it for $N_{legs} = N_{new} > 8$ (i.e., locomotion with more than 8 legs), you will need to compute and add a new gait space metric to the MATLAB function `gaitspacedistance`

1. First use **gaitspacemetricsolver.py** to compute the new gait space metric for $N_{legs} = N_{new}$:

   This generates gait space metric tensors for N_{legs} = 3, 4, 5, 6, 7, 8.

   The relevant calculations are extensible to more legs by analogy to how the various options are currently computed.

2. Next, add this new gij for $N_{legs} = N_{new}$ to the MATLAB function `gaitspacedistance` in the MATLAB script `GaitSpaceClusterer.m` under the switch Nlegs … case loop:

```
switch Nlegs

        case Nnew

                gij = [ … ];
```

3. Add the new supported value to this line in `GaitSpaceClusterer.m`

```
if Nlegs > 8 || Nlegs < 3
    display('Only 3-8 legs supported -- see documentation for how to add
option for more');
```

II. Spreadsheet files

A. Sample model gait files for GaitSimulator.m

2. quadruped_gaits.csv
3. hexapod_gaits.csv
4. octapod_gaits.csv

Description

Spreadsheet files that give phase differences in cycles between N_{legs} – 1 pairs of adjacent legs (using conventions described above) for several model gaits for 4, 6 and 8 legged locomotion. In each file, the first row must be a header. Columns 1 to N_{legs} – 1 are the N_{legs} – 1 phase differences (in cycles) and the last column is the model gait name, which can be a string.

B. Sample simulated randomly perturbed model gait data for GaitSpaceClusterer.m

1. quadruped_gaits_sim.csv
2. hexapod_gaits_sim.csv
3. octapod_gaits_sim.csv

Description

Spreadsheet files with randomly perturbed samples of model gaits for a given number of legs (see GaitSimulator.m entry above) provided to enable testing the GaitSpaceClusterer code.

III. References

Hedrick, T. L. (2008). Software techniques for two- and three-dimensional kinematic measurements of biological and biomimetic systems. *Bioinspir. Biomim.* **3**, 034001.

Kane, S. A., Quinn, B. L., Wu, X. K., Xi, S. Y., Ochs, M. F. and Hsieh, S. T. (2025). Unsupervised learning reveals rapid gait adaption after leg loss and regrowth in spiders. 2025.01.23.634080.

Wilshin, S., Shamble, P. S., Hovey, K. J., Harris, R., Spence, A. J. and Hsieh, S. T. (2018). Limping following limb loss increases locomotor stability. *Journal of Experimental Biology* **221**, jeb174268.