Machine Learning for Android Motion

--Overview and initial plans

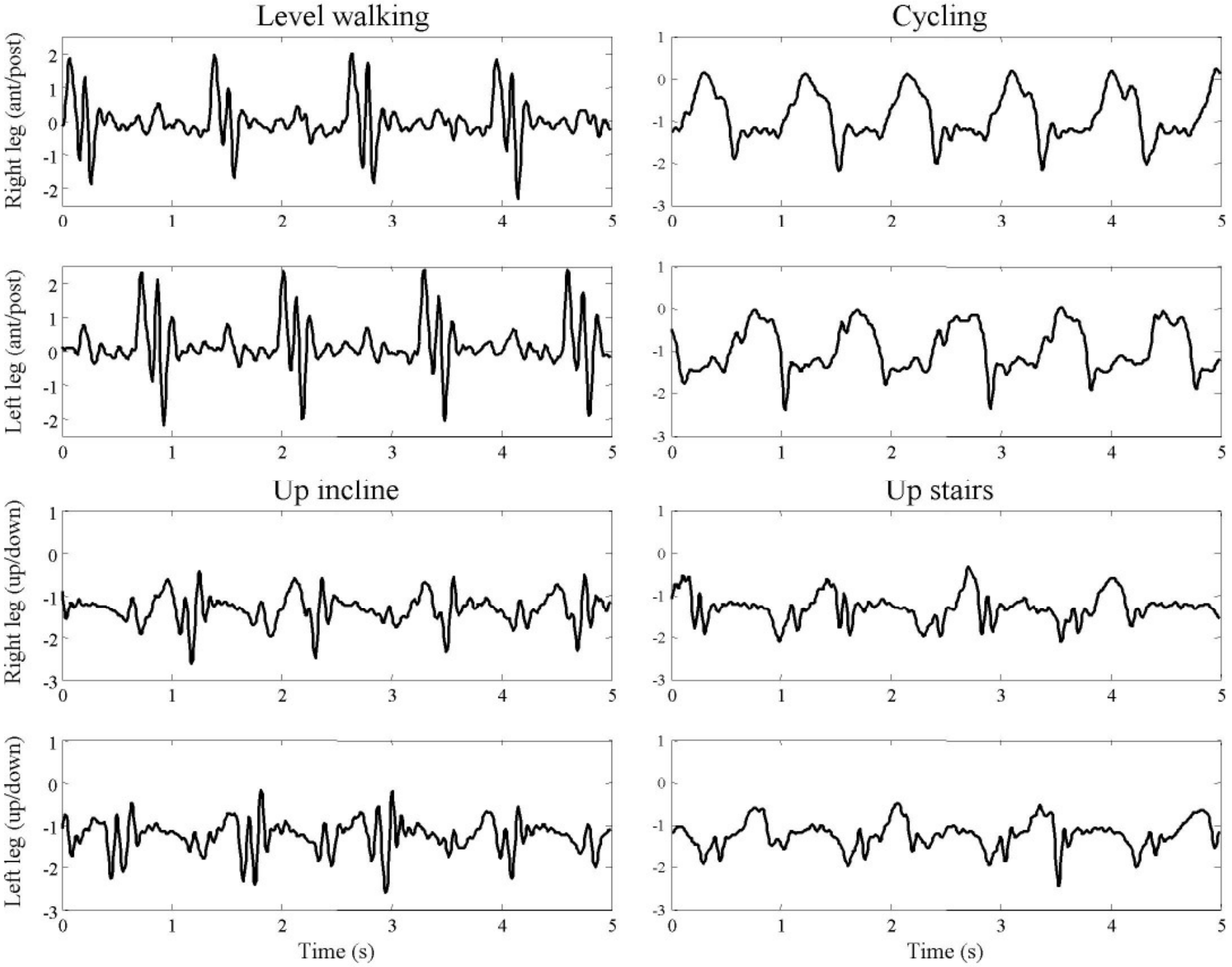
Dai Wei

Wenqi Yao

Doris Xin

Overview

- Classify accelerometer time series collected on Android phones into different categories based on the motions (e.g. walking, sitting, cycling, etc.) associated with the signature of the time series.
- Anomaly detection within each subcategory
  - e.g. The user is walking while the earthquake took place
  - Data collected from the Millikan Shakeout for training
- Intelligent division of maps to reflect propagation of seismic waves
  - Current version uses rectangular grids
  - Use information on epicenters to devise division models


Level walking
Cycling
Right leg (ant/post)
Left leg (ant/post)
Up incline
Up stairs
Right leg (up/down)
Left leg (up/down)
Time (s)
Time (s)

The Data

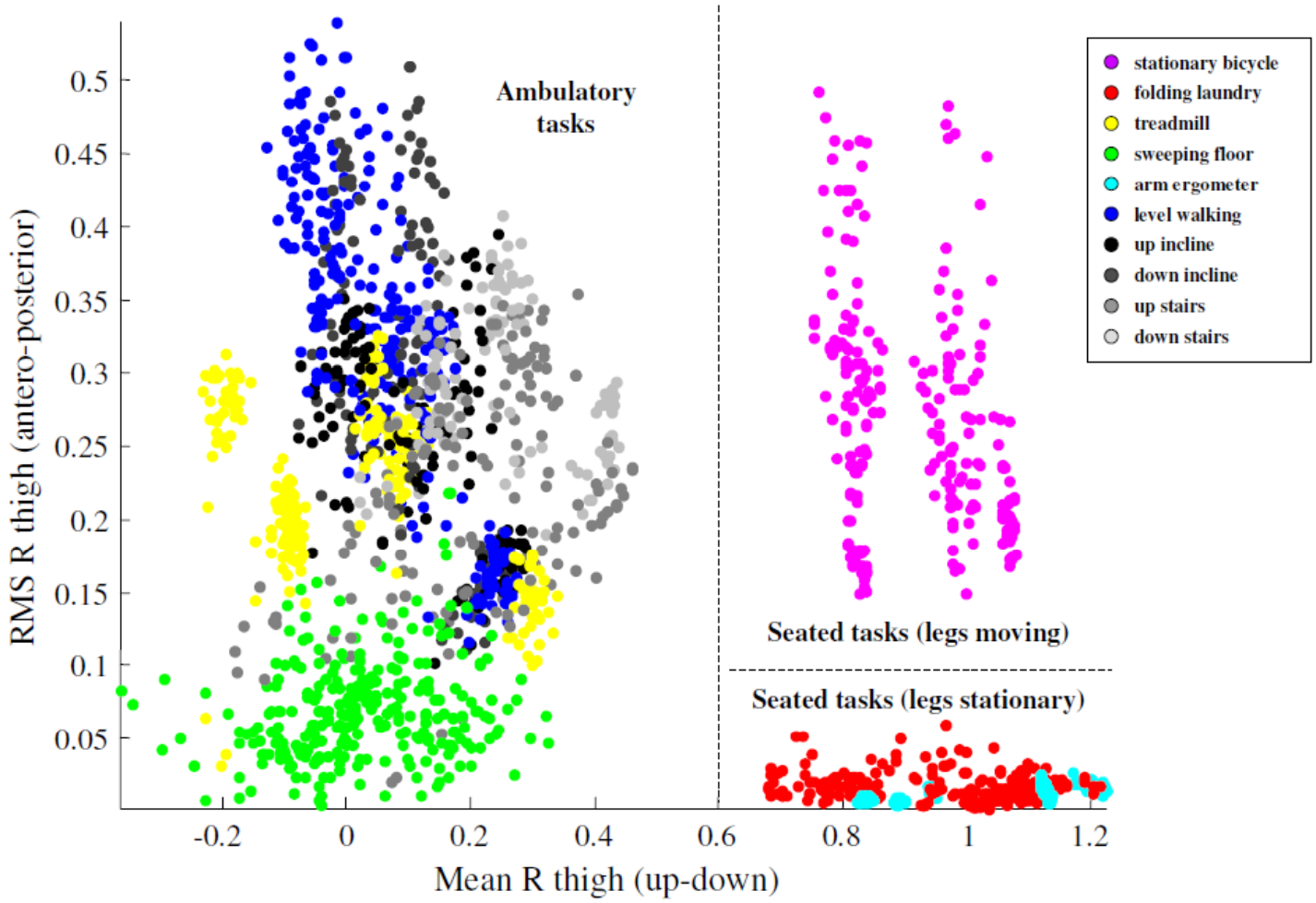
- Time series represented as vector of g-values
  - Either use all 3 axes (X, Y, Z) or only the up and down motion (Sherrill et al. 2005)
  - Discretize time and use the mean of each time bin (which can be very small)
  - Divide the time series into segments of equal length. From previous figure, 1s approximately corresponds to one cycle. Can have multiple cycles in one segment.
  - Incur penalty associated with higher dimensions with using all 3 axes (need to find the tradeoff for higher dimension and information gain experimentally)
- Available data includes data collected during Millikan Shakeout, shaketable experiments for anomaly detection, and ~ 60 GB of regular motion data.

Initial Plans—Learning Algorithms

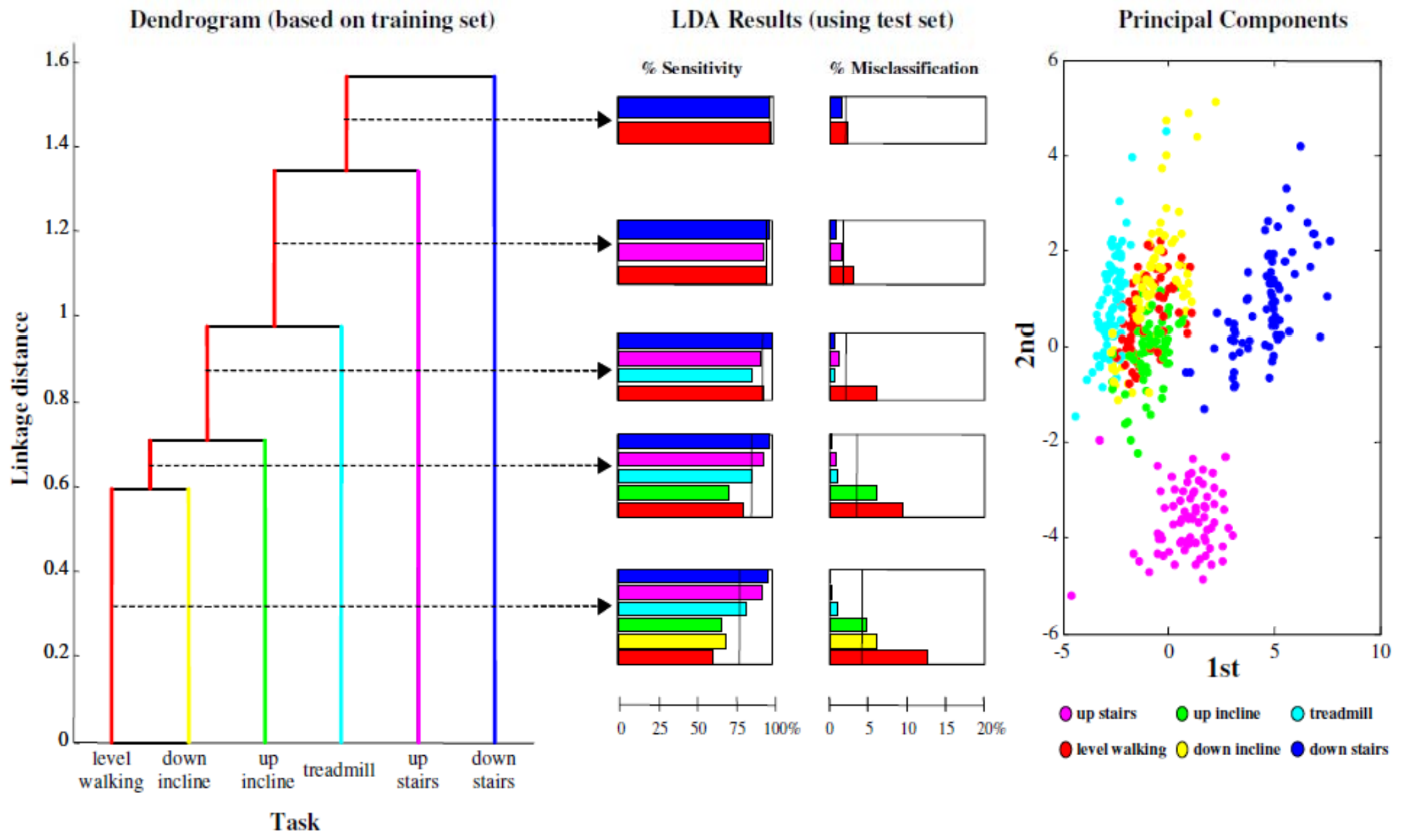
- Supervised Learning:
  - One-vs-One and One-vs-All online SVMs (support vector machines) for motion classification.
    - One-vs-One: distinguish between 2 specific types of motions. e.g. running vs. walking
    - One-vs-All: determine whether the motion is of a specific type. e.g. running vs. not running
    - Online: constantly acquiring training data from the user
  - Backpropagation with online learning.
    - 3-layer network with only one hidden layer
    - Online learning requires frequent updates of weight vectors

Initial Plans—Learning Algorithms

- Unsupervised Learning:
  - Hierarchical clustering in 2D (Sherrill et al 2005) or higher dimensions



Sherrill et al 2005



Initial Plans—App Development

- Collect accelerometer data in the background
- Allow users to identify their activities for labeled training data
- Performs learning tasks locally
- Upload training data to the server at a regular frequency or upon request
- Alert server when anomalous events are detected
- Provide simple statistics on the type of activities the user engages in on a daily basis

Initial Plans—Server

- Use information on epicenters to create division of maps that reflect the probable propagation of seismic waves
  - Draw rings around the epicenters and use data collected on phones in each ring to determine whether an earthquake is occurring in the ring.
  - Boost confidence in detection when rings around an epicenter sequentially “light up” with respect to the distance from the epicenter.

Questions? Comments?