Gait Recognition in Mobile Security

Chase R. Ottomoeller

Division of Science and Mathematics
University of Minnesota, Morris
Morris, Minnesota, USA

December 6, 2014
Senior Seminar, Morris

The Big Picture

What is Mobile Security?

- Information Storage
- Device Access

How is mobile security evolving?

- No More Passwords
- Something You Are
- Unobtrusive Access

Outline

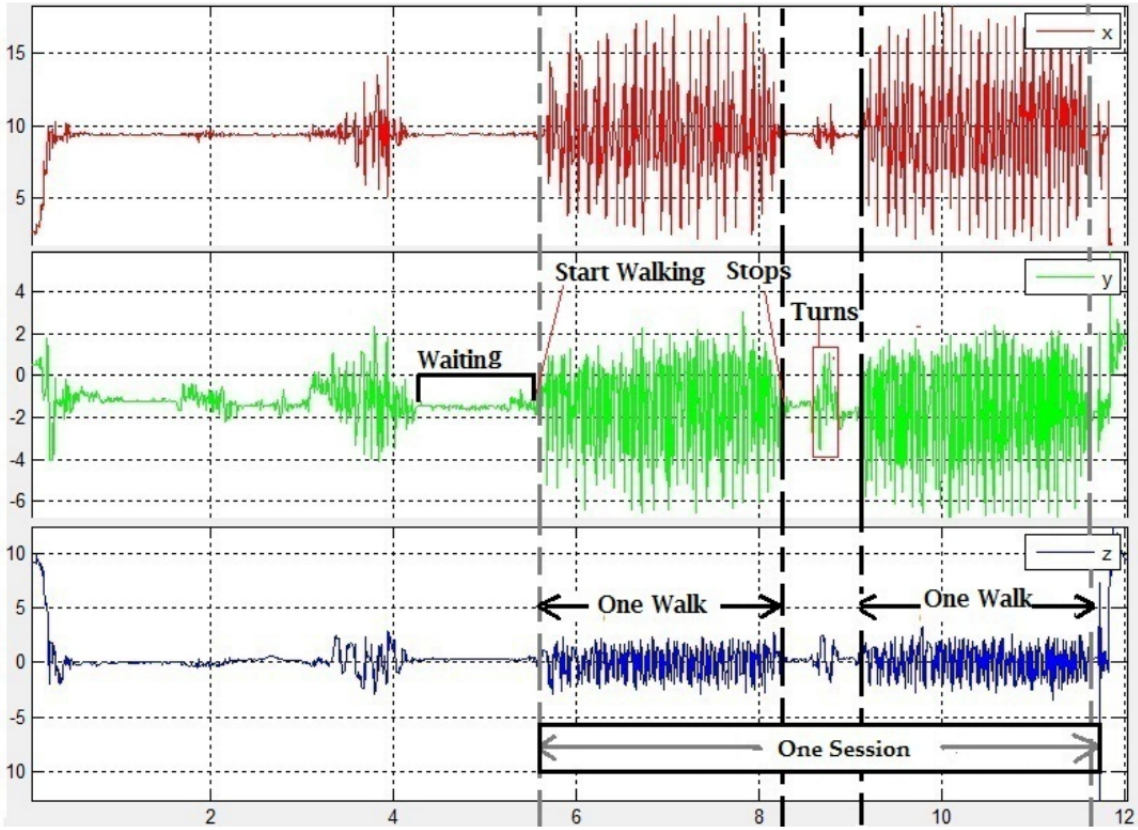
Outline

Biometrics

- Biometrics
- Gait Recognition
- Why Gait is Better

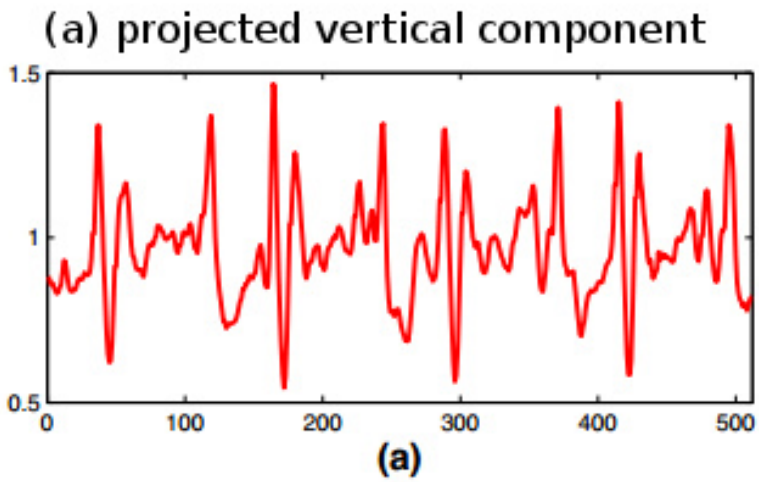
Fixed Method Approach

- Phone Clipped to Waist
- Walked Down 18.5 Meter Hallway
- Separated into "Walks"



Unfixed Method Approach

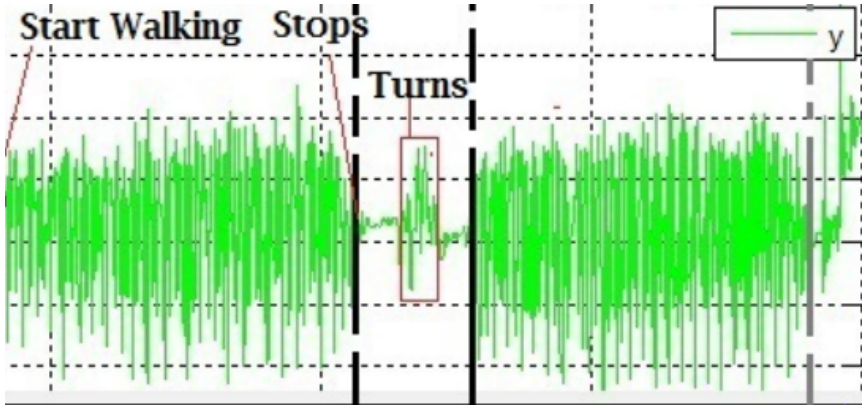
- Phone in more natural location (pocket, handbag, backpack)
- Performed in Real-world Environments
- Separated Into Frames

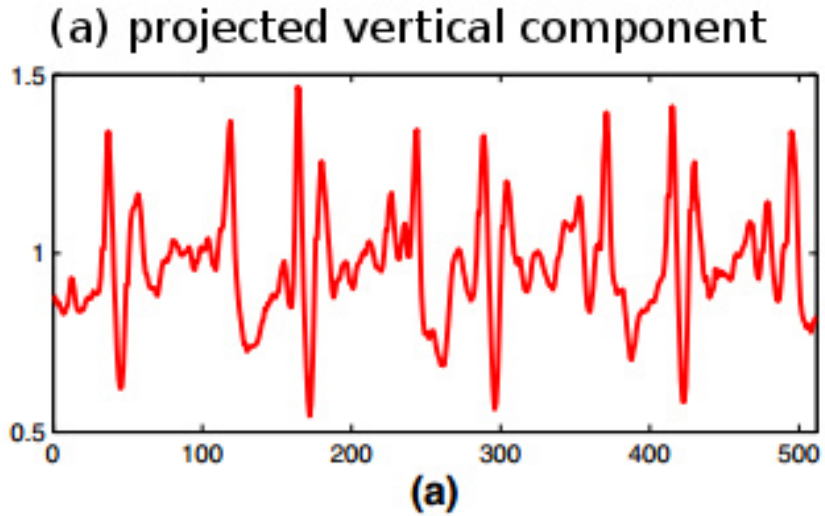


Outline

Preprocessing

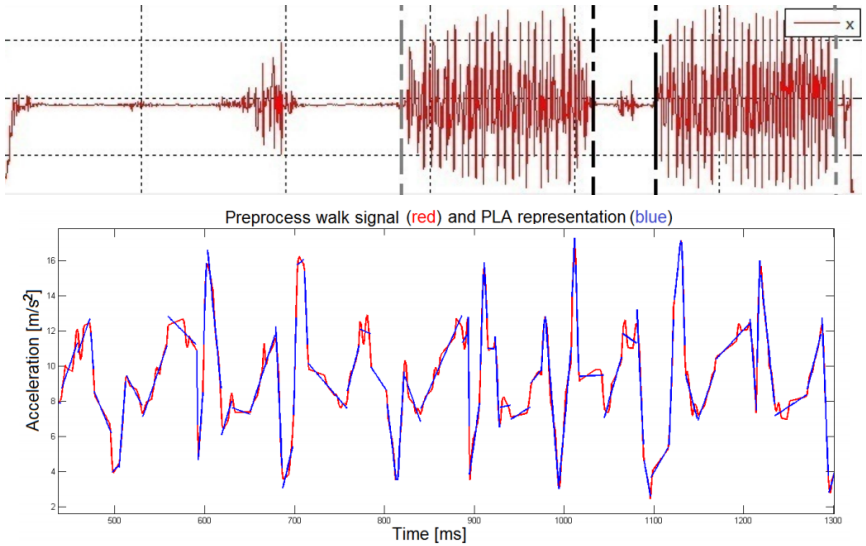
- Separates walking data from other noise
- Walks VS Frames





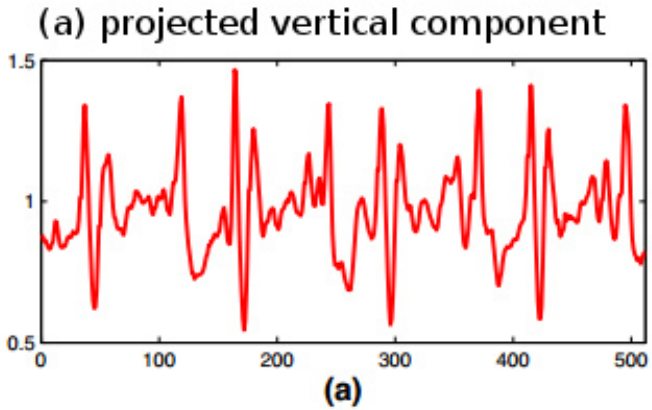
Linear Interpolation and Zero Normalization

- Walk Extraction
- Linear Interpolation (curve fitting)
- Zero Normalization



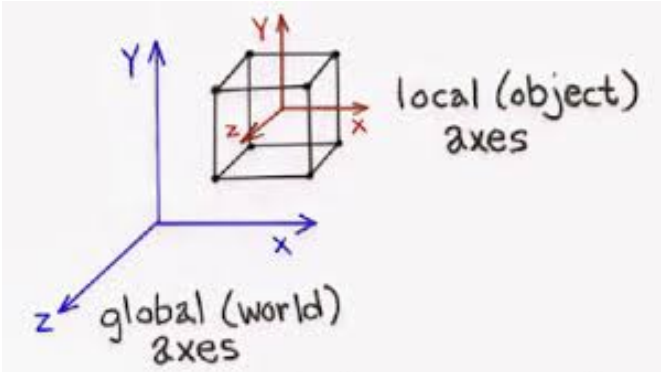
Framing

- Separating Data into Equal Sections
- Frame Length: 5.12 seconds
- Each Frame contains 512 Samples
- Stationary frames are dropped



Projection

- Each sample is projected onto a global coordinate system (sample = x, y, and z)
- Estimating direction of gravity with changes in x, y, and z.
- Each sample is split into two vectors:
  - Vertical (x)
  - Horizontal (y, z)
- Frame dropped if orientation is changed



Outline

What is Feature Extraction?

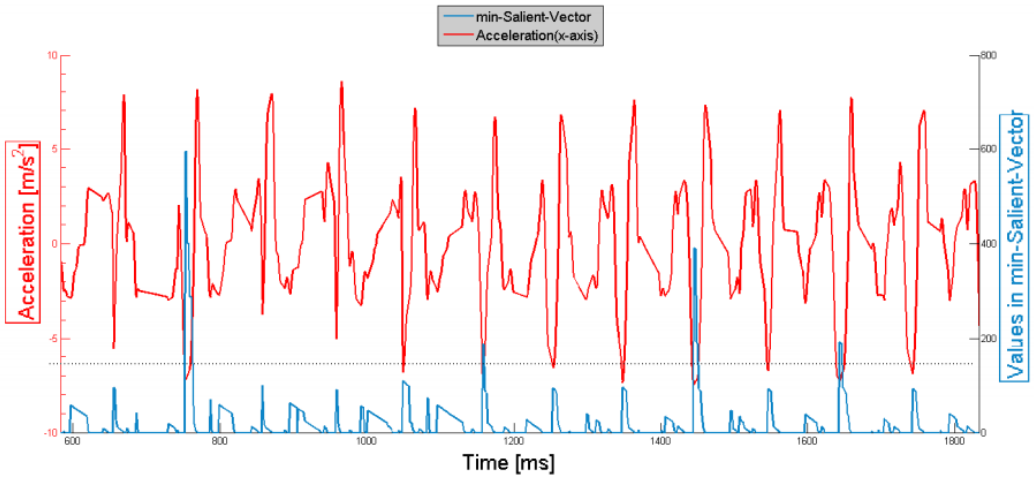
- Feature extraction separates "walking" cycles from "non-walking" cycles

Fixed Method Feature Extraction

- Four Steps:
  - Cycle Length Estimation
  - Cycle Detection
  - Cycle length normalization
  - Omitting Unusual Cycles

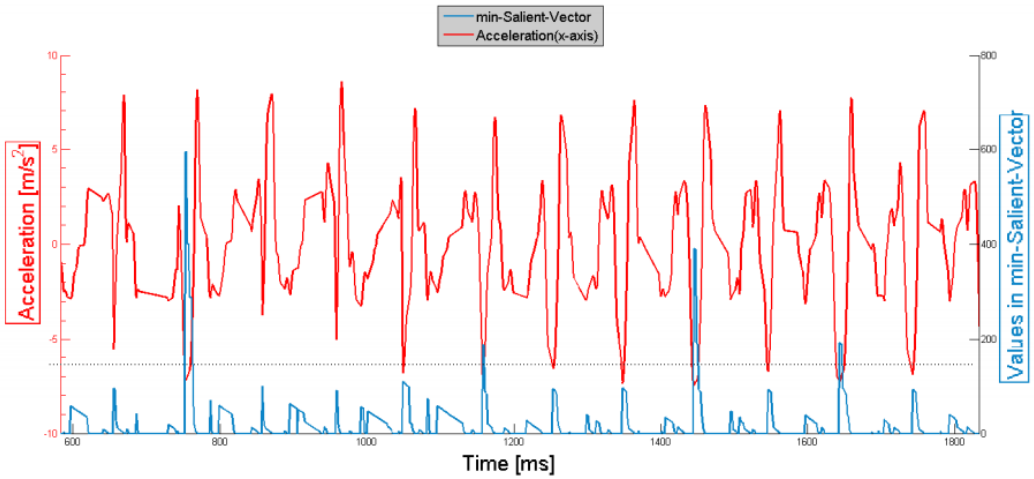
Cycle Length Estimation

- Estimate cycle lengths by computing Minimum Salience Vectors
- Minimum Salience Vector
  - Contains one entry for each data point
  - Each entry is the count of data values between the current value and following smaller value



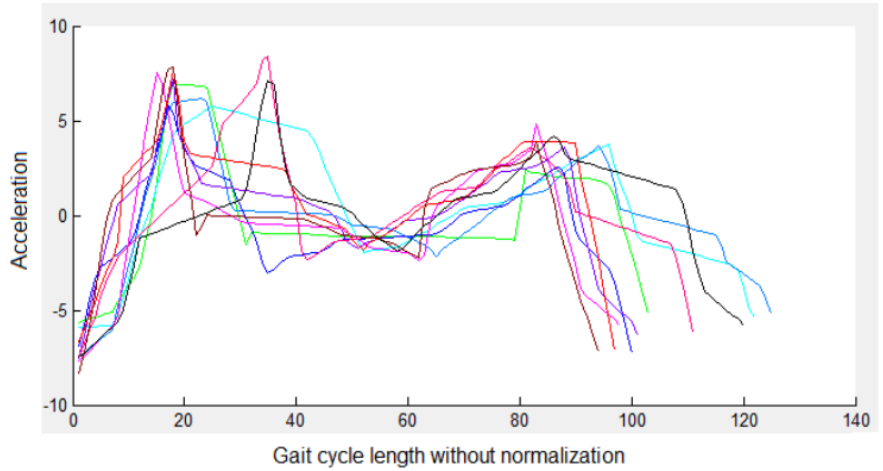
Cycle Detection

- Detecting Individual Cycles
- Start of each cycle is located using the entry with the greatest value
- Spike around points 750, 1150, 1450, 1650
- Long cycles are split again using the same method



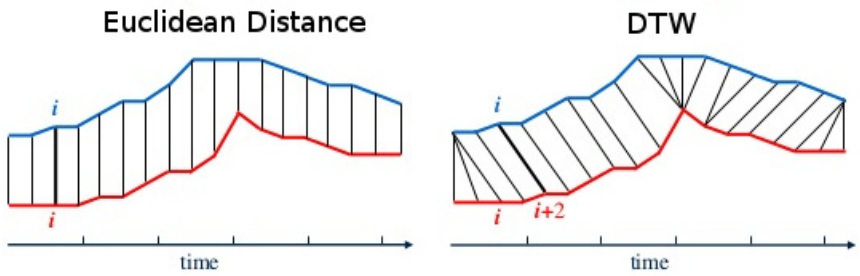
Cycle Length Normalization

- The distance of each cycle is measured from the start of one cycle to the start of the following.
- Cycles need to be of a set length for later Gait Analysis
- Linear Interpolation



Omitting Unusual Cycles

- Deleting Unusual Gait Cycles
- Dynamic Time Warping (DTW): An algorithm used to measure similarity between two sequences
- Cycles with an acceleration half that of the average are dropped



Unfixed Method Feature Extraction

- Three Steps:
  - Feature Extraction I
  - Walking Detection
  - Feature Extraction II

Feature Extraction I

- Determine differences between "walking" and "non-walking"
- Walking 1-2Hz vs Running >3Hz
- These features are used in Walking Detection

Walking Detection

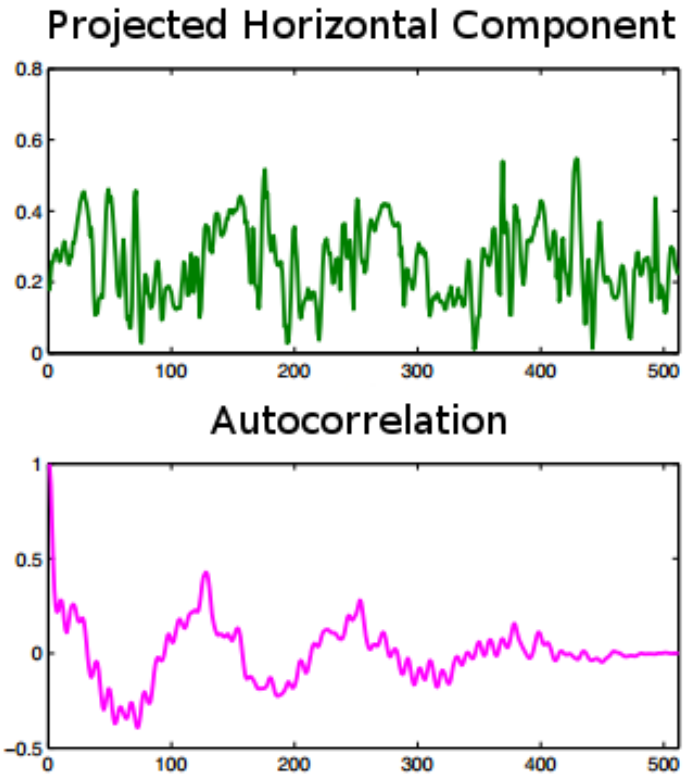
- Three classifications using a decision tree:
  - Walking: 1-2Hz
  - Non-Walking: >3Hz (running, biking, in moving vehicle)
  - Random Movements: >0Hz (transitional movements, short spikes)
- Cycles labeled as walking move onto the next step

Feature Extraction II

- Once Walking Detection confirms that the frame contains walking data, more relevant features are extracted
- Some features extracted using Autocorrelation

Autocorrelation

- Useful to find periodicity and cadence of the gait
- Collecting data from a phone inside a pocket
- Jostling of the phone can create spikes
- Segmentation, like minimum salience vectors, cannot be used
- Autocorrelation can reveal features even with noise



Outline

What is Gait Classification?

- Gait Classification determines if the user is "genuine" or an "impostor"
- Fixed Method Gait Classification
  - Template-based
  - Machine Learning
- Unfixed Method Gait Classification
  - Gausian Mixture Model-Universal Background Model

Template-based

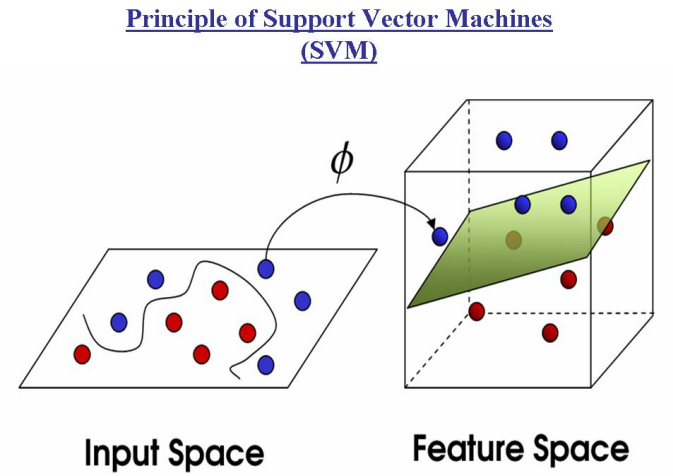
- Feature Cycle (the cycle with the lowest DTW distance)
- Probe Cycles (the remaining cycles)
- After computing probe and reference cycles for all walks two classes are made:
  - Genuine
  - Impostor
- Genuine and Impostor are made by comparing the DTW distance of all the reference and probe cycles
- 50% of the Probe cycles must vote genuine

Machine Learning

- Data is split into two groups:
  - Training (80%)
  - Testing (20%)
- Support Vector Machines (SVMs) are used for biometric classification
- A SVM finds a hyperplane that linearly separates data into two classes: genuine and impostor

Machine Learning

- The data is not usually linearly separable. Therefore, a kernel function is used.
- A kernel function maps non linearly separable data to a high dimension space.
- These data points are now compared to the Testing data set. Again, the class with the maximum votes wins.



Unfixed Method

- The use of more than one training model is used to help classify gait cycles
- Universal Background Model (UBM) is used to train a large source of data
- User's gait model is generated relating the odds of one event to the odds of another
- Maximum-a-Posteriori (MAP) is used to adjust Gaussian components and mixture weight to personalize the UBM model
- The current user's gait cycle is compared to the personalized UBM model and either accepts or rejects.
- MAP is also used by recording false negatives

Outline

Conclusions

- The unfixed method is developed more and is better for real-life situations
- Given time, the fixed method can perform just as well as the unfixed method

References