

# Gait Recognition in Mobile Security

Chase R. Ottomoeller  
Division of Science and Mathematics  
University of Minnesota, Morris  
Morris, Minnesota, USA 56267  
ottom005@morris.umn.edu

I left some parts sparse as I'm still figuring out the amount of detail I want to put in that section.

## ABSTRACT

This paper discusses a new form of mobile security not just emphasizing security, but also usability. Having a form of security of a mobile device that is unobtrusive can make that device easier and a less of a hassle to use. This paper will discuss using built in accelerometers in smart phones as a way to collect biometric data. In this case the biometric data is a persons gait or walking pattern. This increases the security of the device by allowing the authentication to not be something you have or something you know, but something you are. There are two example that this paper will compare showing two forms of gait analysis.

## Keywords

mobile security, gait recognition, biometrics, accelerometer

## 1. INTRODUCTION

This section will be redone

As computational power in computers increases, the security risks grow higher for forms of authentication. One growing concern is security in mobile devices such as smart phones. Smart phones are also becoming increasingly more powerful, which results in more personal data being stored on the device. This means security for smart phones will become more important as time goes on. Currently security mainly consists of: pin, password, swipe lock, fingerprint. These forms are either something a person has(fingerprint) or something a person knows(password). They all also require a physical action to unlock the phone. A new form of mobile security involves a persons gait making it an unobtrusive form of security. This means that the built in accelerometer in the smart phone will analyze a persons walking pattern. If the phone tell that the current user should have

access the phone will be unlocked, otherwise it will switch to the normal pin password. This will increase the security of the mobile device because the form of unlocking the device is now something you are and not something someone else can take form you. This paper will analyse two different approaches for gait extraction/analysis. Each paper uses a general format of analysing the data which includes: Preprocessing, Feature Extraction, and Gait Analysis. Preprocessing is the process of capturing biometric data. Feature extraction is taking the raw biometric data and locating distinctive characteristics. Gait Analysis is taking the feature extracted data, and running algorithm on it which will return a positive or negative result. Though each paper follows this general outline, they did each step in a different way.

The introduction and background are similar and is being addressed.

## 2. BACKGROUND

Biometrics is used as a new form of security that provides unique identification to a mobile user. Now, instead of relying on a password of physical key, a person can gain access to their device by a physical trait specific to them. Some examples of biometric traits are: finger prints, facial features, eye, voice, hand, DNA, gait. In this paper I will focus on the gait recognition. The term gait refers to the walking pattern of a person. A persons walking pattern is cyclic in nature and may be composed of many gait cycles, where each gait cycle consists of at least two steps. [?] There are three types of biometric gait recognition: machine vision based, floor sensor based, and wearable sensor based. With mobile devices becoming more and more equipped with an array of sensors, the technique most suited for mobile devices is a wearable sensor based, using the mobiles devices built-in accelerometer. Data collected from the accelerometer will be processed in a set of operations later described. In the next sections I will be comparing two approaches that use accelerometers from a mobile device such as a smart phone. The primary way in which these two approaches differ is the placement of the mobile device on the body. One approach has the smart phone in a fixed location on the waist. I will refer to this method as the *fixed* method. The other method has the smart phone in a more natural position on the body such as in a pocket. This method will be referred to as the *unfixed* method. Though both papers use the general approach shown in figure 8 I will be simplifying them into three categories: Pre-Processing, Feature Extraction, and

This work is licensed under the Creative Commons Attribution-Noncommercial-Share Alike 3.0 United States License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/3.0/us/> or send a letter to Creative Commons, 171 Second Street, Suite 300, San Francisco, California, 94105, USA.

UMM CSci Senior Seminar Conference, December 2014 Morris, MN.

### 3. PREPROCESSING

Once data is gathered it needs to be preprocessed into a usable form. For this paper a usable form is modified data that has been separated into frames. A frame is equally separated parts of data for feature extraction and classification. With two methods (fixed and unfixed), separating the data into modified frames was done differently. The fixed phone method used a strict test case that was modified using Linear Interpolation and Zero Normalization. The unfixed phone method used Framing and Projection. Framing separated the data into frames. These frames were then modified by using Projection. In the following sections I will explain the fixed and unfixed methods as well as compare them.

#### 3.1 Fixed Method Preprocessing

With the phone on the waist in a fixed position, uniform sets of data called walks were used. In this case one walk is represented by walking a measured distance down a hallway. The experiment was done in such a way that one set of data contained two "walks", or in other words, one set of data was walking a hallway; Once down and once back. After extracted, this data is modified to further extract features. These modifications include linear interpolation and zero normalization.

##### 3.1.1 Interpolation

Once the walks are separated from the initial data, they have to be formed into equal intervals. This is done by applying *Linear Interpolation*, which can reshape the data into equal intervals. In order to not lose data during this change up-sampling is applied. Up-sampling avoids data loss by producing "an approximation of the sequence that would have been obtained by sampling the signal at a higher rate" [?]. Once the data has been reshaped, the acceleration measurements have to be adjusted.

##### 3.1.2 Zero Normalization

The PMDs accelerometer being used measures 3 axis(X, Y, Z). the only axis needed, and the only axis affected by gravity is the X axis. Since the acceleration for the Y and Z axis are not stable over time, and they are not affected by earth's gravity, the acceleration data for the Y and Z axis have to be *zero normalized*. In other words the data for the Y and Z axis have to be set to zero. This is done by simply subtracting their total acceleration from their average acceleration. An overview of the stages explained above are shown in figure 1.

#### 3.2 Un-fixed Method Preprocessing

The previous method used a phone in a fixed location. This method tries to expand on this with a phone in a unfixed location, such as a pocket. Also, this method is less strict. It does not use a set length, such as a hallway, to mark each walk. Instead this method uses a process called *Framing*. Once the data is segmented into frames it is then projected onto a global coordinate system through a process known as *Projection*.

##### 3.2.1 Framing

During Framing, sensor data is segmented into uniform frames for feature extraction and feature classification. Un-

like the data for the fixed method, this data is segmented into 512 equal length samples. This is done by taking a sample every 5.12 seconds. They state the reason for 512 samples was "to balance between estimation accuracy and latency. Also, not all frames move onto feature extraction. Frames that are below a chosen threshold, representing no movement, are dropped.

##### 3.2.2 Projection

Each sample within a given frame contains an X, Y and Z coordinate. Each coordinate is then projected into a vertical and horizontal variable. Once each sample is projected, the direction of gravity is determined by using a mean filter. Unlike the previous example, this process doesn't assume a fixed location or orientation of the mobile device. If there is a significant change in the gravity variable then the orientation of the device has changed. This means the corresponding frame will be dropped and the horizontal and vertical axes will be adjusted accordingly.

Both the fixed and unfixed method split the data into equal intervals and then modify the data so that it is able to be worked with. The reason for the differences between stems from the differences in phone placement. The data collected in the fixed method is already more organized than the data from the unfixed method. This is because the unfixed method has to worry about all axis as well as organizing the data into neater parts. If the fixed method were to be implemented in more of a real life scenario then these methods become even more similar.

### 4. FEATURE EXTRACTION

*Feature Extraction* is extracting a set of data from a given frame to detect patterns of walking. Another way to define feature extraction is the process of extracting gait cycles from the data.

#### 4.1 Fixed Method Feature Extraction

The fixed method in order uses the following steps as seen in figure 2: Cycle length estimation, Cycle detection, Cycle length normalization, Omitting unusual cycles.

##### 4.1.1 Cycle Length Estimation

Once the raw gait data is preprocessed, the next step is extracting the biometric gait cycles. The first step in extraction is to be able to automatically detect gait cycles in the walk. This detection is done by estimating the cycle length, which is simply computing minimum salience vector. There is a minimum salience vector for each data point for a given walk. Each minimum salience vector is computed by counting the data values that are between that current data value and the following smaller value in the walk vector. If the current salience vector is greater than the minimal peak distance or height, it is a cycle start. This same process is used to determine the maximum peak and distance as well.

##### 4.1.2 Cycle Detection

Once both minimum and maximum salience vectors are computed so they can be used to detect individual cycles. A good example can be seen in figure 4. In this figure the minimum salience vectors are seen in blue and the accelerometer data is seen in red. We can easily see where the start of each cycle is based on the minimum salience vector. Maximum salience vectors are also computed to help

determine the exact minima.

#### 4.1.3 Cycle Length Normalization

Using linear interpolation again, the determined cycles are normalized to a set length. These set lengths are later required in the Euclidean distance which requires vectors of the same length.

#### 4.1.4 Omitting Unusual Cycles

Some cycles compared to the majority are considered unusual. These could be caused by any motion or non-motion that isn't walking. This is done by "computing the distance using *Dynamic Time Warping* (DTW)." [?] Dynamic time warping is "an algorithm for measuring similarity between two temporal sequences which may vary in time or speed." [?] This means that cycles with a distance of half or more of the other cycles will be removed.

### 4.2 Unfixed Method Feature Extraction

A different form of extraction used by [?] uses a combination of time-domain, frequency-domain, and auto-correlation features. This example also did feature extraction in two stages: calculating the set of features for walking and calculating more precise gait features.

#### 4.2.1 Feature Extraction I

The first stage of the feature extraction is used to tell whether or not the data represents walking data. This is done to rule out the times when the mobile device is not collecting the right data. For example if the mobile device was collecting data when it was in a car going down the road, that data shouldn't be used because it isn't gait data. This also holds true for collecting data on someone who is standing still. The data collected wouldn't represent the data needed for a biometric. Telling whether the data collected is of someone walking is done by collecting vertical and horizontal features to produce motion patterns. Six time-domain features are used for both directions. These six features are: mean, variance, skewness, kurtosis, energy, mean-crossing rate. All these features are based on spectrum analysis. In other words, doing "different activities have different energy distributions over the frequency spectrum." [?] Walking can be measured around 1-2Hz while driving in a car will output a higher frequency band. Based on these frequency levels the data is put into one of three classes: walking, non-walking and motion (high frequency movement such as in a vehicle or running). [?]

#### 4.2.2 Walking Detection

Using data from feature extraction I, classification can be done using a decision tree. There are three activities that the data can be classified as: walking, non-walking, and random movements. Non-walking motion is running, biking, or moving in a vehicle. Random motion is a motion such as turning or skipping. A Markov Model smoother is applied to the new set of data to get rid of any outliers.

#### 4.2.3 Feature Extraction II

Feature extraction II is done once the data collected is determined to be gait data. In this stage more relevant features are extracted for analysis. The first analysis to be done is the Compressed sub-band cepstral coefficients (CSCC). This is done in the following three steps. First, the energy spectrum

is computed using the FFT spectrum. The Fast Fourier Transform (FFT) is an algorithm that computes the Discrete Fourier Transform (DFT) and its inverse. The study maps the energy spectrum into 26 bands using triangular overlapping bands and sum up the energy in each band. The processing then takes the discrete cosine transform of the sub-band energy to form a 12-dimension vector representation. This in all summarizes the fundamental frequency of the movements and the higher frequency vibrations in the data. [?]

This section will need further explanation, clarification and or simplification.

Both of these methods of feature extraction try to determine what parts of the data represent a walk cycle. The fixed method is simpler because it doesn't have to worry about different variables affecting the data. Though this method is simpler, it has the stipulation of the mobile device being clipped to your waste. The non-fixed method has to worry about more variables affecting the data, but implements a real life situation of just having the mobile device in a pocket.

## 5. GAIT CLASSIFICATION

This section will be reworded, and organized and added to

This last step is taking the extracted features and verifying that those features match a stored set of features; the stored set of features being the users. The fixed method uses Support Vector Machines (SVM's). The basic idea of SVM's is taking dimensional data and separating it into two classes. This involves the use of a Gaussian kernel defined in the following section. The unfixed method uses a Gaussian Mixture Model - Universal Background Model (GMM-UBM) framework to verify an individual's gait. The basic idea with this technique is with scoring. Verification of a user's identity is done by comparing the likelihood score from a user's gait pattern to the universal background model. The universal background model represents human gait patterns in general.

### 5.1 Template-based classification

This section will be redone

The template-based classification uses the same approach for gait extraction as shown in 3 with one extra step added shown in 6. Before the Cycle Length Estimation module is Piecewise Linear Approximation (PLA). A Piecewise Linear Function (PLF) is used for the PLA. A PLF is a function composed of straight-line sections connecting points on a graph. So, a PLA is an approximation of a function connecting significant points on the curve with straight-line sections. The approach used to produce the PLA representations is the Sliding Window And Bottom-up (SWAB). Once the unusual cycles are omitted, the best cycle is picked. The best cycle is determined by having the lowest DTW. One the best cycle is found the remaining cycles are called probe cycles. These two sets of cycles are compared using the DTW distance to compute the genuine and impostor class. The DTW of each cycle is computed and majority voting takes place. If 50% or more of the comparisons returns and accept then the cycle is a genuine, otherwise it's rejected as an impostor.

### 5.2 Fixed Method Gait Classification

This section will be reworked

SVMs are another way to classify gait data. SVMs are often used for biometric recognition, where imposter and genuine data have to be determined. First the SVM finds a *hyperplane*. A hyperplane is a subspace that is one dimension less than its “ambient space”. For example, there could be a set of data represented by 3 axis (X,Y,Z). The hyperplane would be the same data plotted on a 2 axis space. So, the SVM will take the data, and find a hyperplane that linearly separates  $\mathcal{D}$  dimensional data into two classes. The dimensional data has to be linearly separable for the SVM to work. Since the gait data isn’t linearly separable, a “kernel induced feature space” [?] is introduced in the SVM. A kernel function maps non linearly separable data to a high dimensional space. Once mapped to a high dimensional space the data is linearly separable and therefore can be used by the SVM to find a hyperplane. This means that depending on the kernel function you use, the better classification accuracy it will produce. There are kernel functions that would require the use of fixed length gait cycles. Since gait cycles are not usually of a fixed length, a different kernel is used. By using the DTW distance, the problem of fix-length input feature vectors can be solved as well as more accurate finding of similarity in the data. Two different approaches using SVM and DTW can be used to classify gait cycles: Pre-computed data matrix, and Pre-computed kernel. 7

### 5.2.1 Pre-computed data matrix

Gait cycles are represented by the DTW distance. The DTW distance is the distance between two gait cycles. The benefit of the DTW distance is that gait cycles with different length can be found. A DTW matrix is then computed by taking the sample DTW distance and compared to all other samples. This matrix is needed as an input for the SVM.

### 5.2.2 Pre-computed Kernel

Using GDTW kernel equation 1 as the kernel matrix also allows to classify the gait data. Though this kernel matrix is a symmetric matrix, it is not guaranteed that it has all positive eigenvalues.

$$K(x, z) = \exp(-\gamma \|DTW(x, z)\|^2) \quad (1)$$

## 5.3 Unfixed Method Gait Classification

The framework used for classifying gait cycles is a Gaussian Mixture Model-Universal Background (GMM-UBM). The algorithm can be split into three primary sections: UBM training, user gait model generation, runtime inference and adaptation.

The UBM is a large Universal background GMM that is trained from a large source of data. This means a UBM represents the general walk pattern of humans. The UBM is defined as  $\lambda(\omega, \mu, \Sigma)$  where  $\omega$  represents the mixture weight,  $\mu$  represents the mean, and  $\Sigma$  represents the covariance matrix. In other words  $\omega$  is the prior distribution of waling patterns while  $\mu$  and  $\Sigma$  are different walking patterns given by the population in different conditions. For efficiency the *covariance matrix* is used. A covariance matrix “generalizes the notion of variance to multiple dimensions.” Also, to avoid over fitting the training data, a standard expectation maximization (EM) algorithm is applied. The EM is what

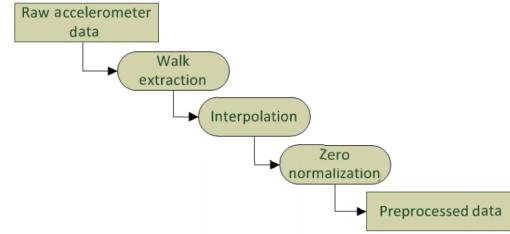


Figure 1: Preprocessing Step

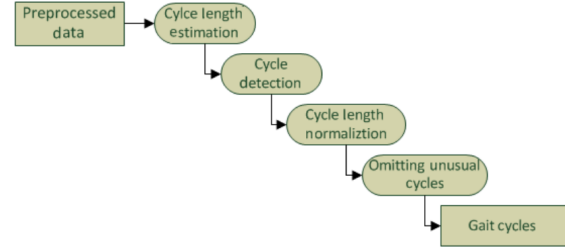


Figure 3: Feature Extraction

trains the UBM models. For example, if given a collection of training vectors, the EM estimates the parameters to most likely to occur. The EM algorithm does this through iteration. Each time the algorithm iterates it refines the GMM parameters, increasing the likelihood of the estimated model for the given feature vectors.

## 6. PUTTING IT ALL TOGETHER

I have compared two forms of using built gait as a form of mobile security. One experiment used a fixed location for the device while the other used multiple locations. Both use different methods of extracting the gait data as well as analyzing the data, but each method works. Both of these experiments are not perfect yet but they both show promise in using walking patterns as a form of mobile security. With more sampling/testing as well as improving efficiency with algorithms will improve these mobile security methods and show potential to become implemented in a large scale.

## 7. ACKNOWLEDGMENTS

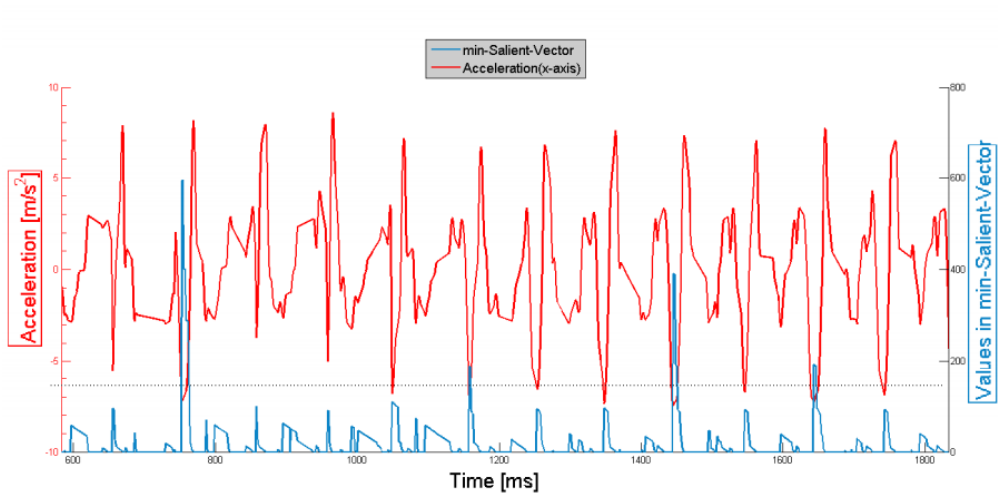


Figure 2: Minimum Salient Vectors

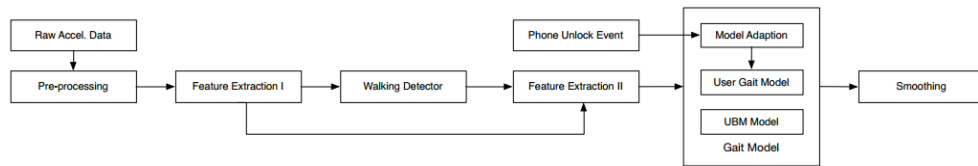
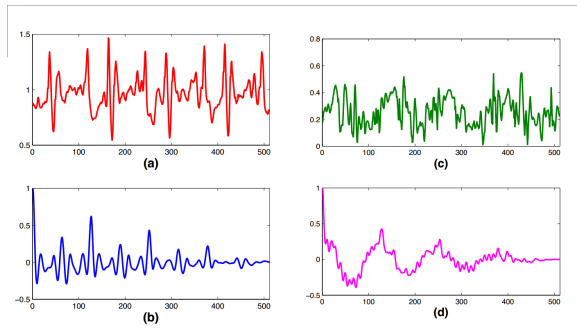
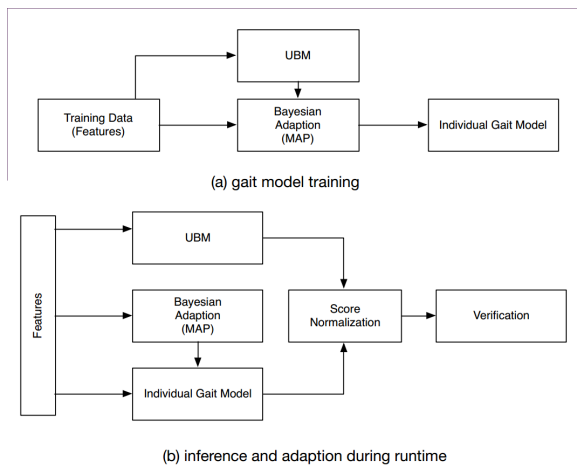


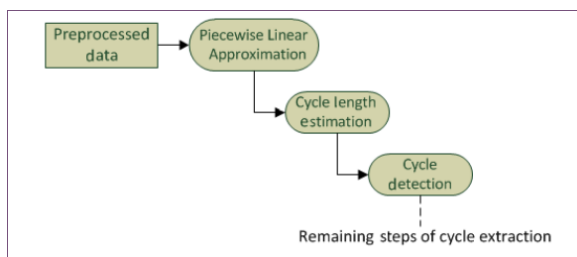
Figure 9: Algorithm Overview



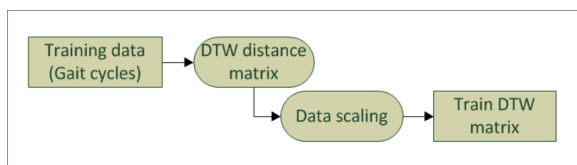
**Figure 4:** (a) projected vertical component. (b) normalized autocorrelation of the vertical component. (c) projected horizontal component. (d) normalized autocorrelation of the horizontal component



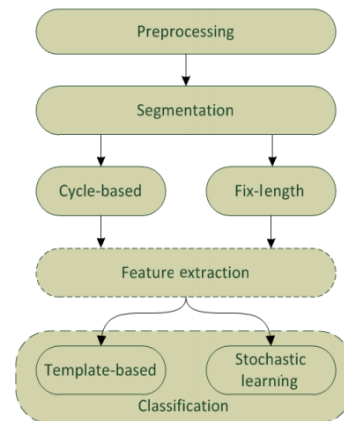
**Figure 5:** Algorithm work-flow for (a) generate user gait model by MAP adaptation (b) runtime inference and Individual model adaptation



**Figure 6:** Gait cycle extraction steps with Piecewise Linear Approximation(PLA)



**Figure 7:** Preparing data for SVM)



**Figure 8:** Algorithm Overview