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

ABSTRACT

This paper discusses a new form of mobile security, emphasizing both security and usability. Having a form of security of a mobile device that is unobtrusive can make that device easier to use. Using built-in accelerometers in smart phones as a way to collect biometric data is one of these forms. In this case the biometric data is a person's 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. Though both examples need further work, they both implement working security schemes allowing the phone to verify the users identity correctly.

Keywords

mobile authentication, gait recognition, biometrics, accelerometer

1. INTRODUCTION

As computational power in computers increases, the security risks for authentication grow higher. One growing concern is security in mobile devices such as smart phones. Smart phones are becoming increasingly more powerful, which results in more personal data being stored on the device. As time goes on, security for smart phones becomes more important. Common security authentication methods include: pin, password, swipe lock, and fingerprint. These methods are either something a person has (fingerprint) or something a person knows(password). They all require a physical action to unlock the phone. A new form of mobile security involves a person's gait (walking patter) making it an unobtrusive form of security. Gait data is collected via the built in accelerometer of the smart phone. The smart phone will use the collected data to analyze a persons walking pattern. If the phone can determines that the current user should have access the phone will be unlocked, otherwise it will switch to a different authentication method such as pin access. This

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

will increase the security of the mobile device because the form of unlocking the device is now something you are (a biometric) and not something that is likely stolen or forgotten. Also, unlike other biometrics such as a fingerprint, an imposter would have trouble replicating a gait pattern.

This paper analyzes two different approaches for gait extraction/analysis. Each paper uses a general format of analyzing the data which includes: Pre-processing, Feature Extraction, and Gait Analysis. *Pre-processing* is the process of capturing biometric data. *Feature Extraction* is taking the raw biometric data and locating distinctive characteristics within that data. *Gait Analysis* is taking the features extracted from the data, and running an algorithm on them. This algorithm will return a positive or negative result representing if the user is genuine or an imposter.

In the next sections I will be comparing the two approaches of gait recognition using 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 a complex approach shown in figure 1 I will be simplifying them into three categories: Pre-Processing, Feature Extraction, and Gait Analysis.

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. [3]

3. PRE-PROCESSING

Once data is gathered, it needs to be pre-procesed 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. In the following

Algorithm Process

Figure 1: Different Approaches

Figure 2: Fixed Method Pre-processing Step
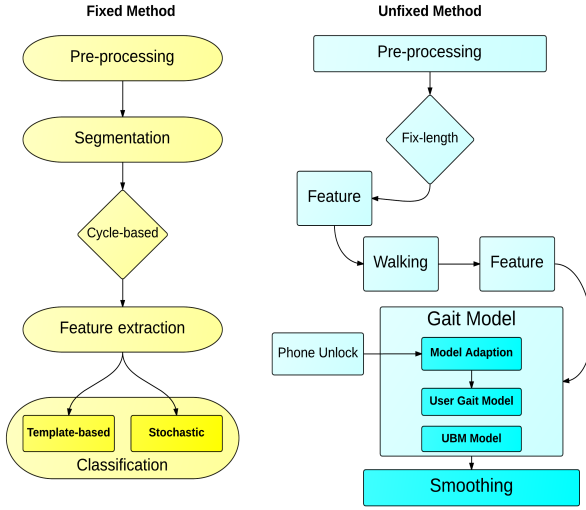
sections I will explain the fixed and unfixed methods as well as compare them.

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.

3.1 Fixed Method Pre-processing

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 is 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" [5]. 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 axes 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
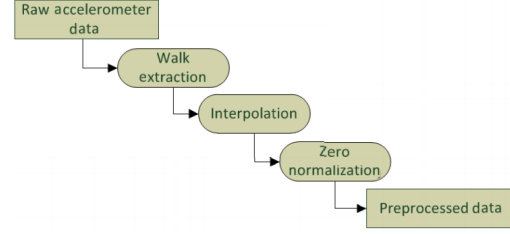
Y and Z axis have to be set to zero. This is done by simply subtracting their total acceleration from their average acceleration. An overview of the stages explained above are shown in figure 2.

3.2 Un-fixed Method Pre-processing

The previous method used a phone in a fixed location. This method expands on this with a phone in a unfixed location, such as a pocket. Also, this method is more flexible. It does no 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 know as *projection*.

3.2.1 Framing

During Framing, sensor data is segmented into uniform frames for feature extraction and feature classification. Unlike 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 does not 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 as well as modify the data so that it is in a useful form. The reason for the differences between the pre-processing methods 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 account for multiple axes as well as unorganized data. If the fixed method were to be implemented in more of a real life scenario then these methods become even more similar.
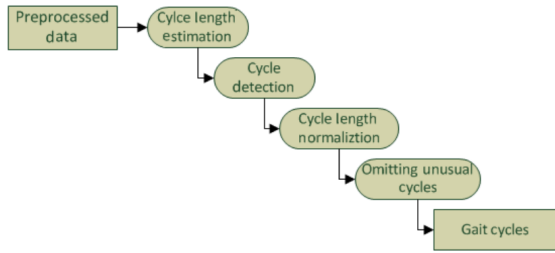
4. FEATURE EXTRACTION

Figure 3: 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 uses the following steps in order, as seen in Figure 3: Cycle length estimation, Cycle detection, Cycle length normalization, Omitting unusual cycles.

4.1.1 Cycle Length Estimation

Once the raw gait data is pre-processed, 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*. For this instance a salience vector is distance measurement from one point of a cycle to another. 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 the current data value and the following smaller value in the walk vector. The number of values counted is the current points distance measurement. The salience vector with the greatest distance measurement is considered a cycle start.

4.1.2 Cycle Detection

Once both minimum and maximum salience vectors are computed, they can be used to detect individual cycles. Figure 4 shows a set of gait cycles in black, and the salience vector distance measurement in grey. The start of each cycle is based on the minimum salience vector. Around times 750, 1150, 1450, and 1650 can be seen spikes where the distance measurement is larger than the rest. These spikes are the minimum salience vectors, marking the start of gait cycles. This same process is applied to determine the maximum salience vector as well. If there are some cycle that are much longer than others, these minimum and maximum points are used to determine more cycle starts. This splits too long cycles into smaller ones.

4.1.3 Cycle Length Normalization

Once each start of a cycle is determined, the cycle distance is the distance from the start of one cycle to the start of the following cycle. Each cycle does not have the same length. By using linear interpolation, the determined cycles are normalized to a set length. Equal cycle lengths are later required for some feature extraction techniques.
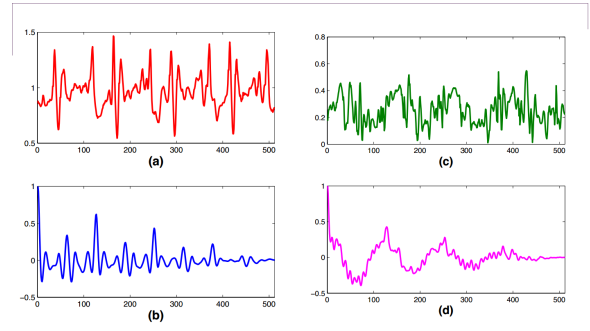

Figure 5: (a) projected vertical component. (b) normalized autocorrelation of the vertical component. (c) projected horizontal component. (d) normalized autocorrelation of the horizontal component

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 does not correspond to walking. Such patterns are detected by "computing the distance using *Dynamic Time Warping* (DTW)." [2] Dynamic time warping is "an algorithm for measuring similarity between two temporal sequences which may vary in time or speed." [4] This means that cycles with an acceleration of half or more of the other cycles will be removed

4.2 Unfixed Method Feature Extraction

A different form of extraction used by [1] is 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 should not be used because it is not gait data. This also holds true for collecting data on someone who is standing still. The data collected would not 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 patters. Six time-domain features are used for both directions. These six feature 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." [1] 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 a in a vehicle or running). [2]

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
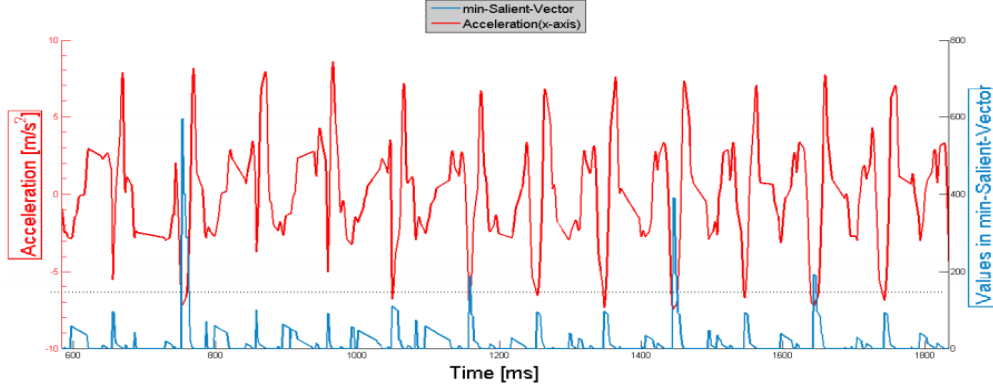
Figure 4: Minimum Salient Vectors

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 alogorithm that computes the Discrete Fourieer 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. [3] 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 does not 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 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 a individuals gait. The basic idea with this technique is with scoring. Verification of a users identity is done by comparing the likelihood score from a users
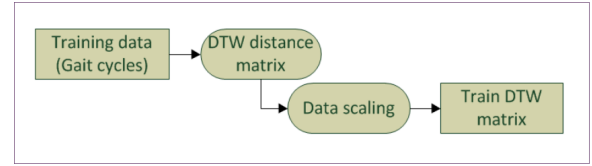

Figure 6: Preparing data for SVM)

gait pattern to the universal back ground model. The universal back ground model represents human gait patterns in general

5.1 Fixed Method Gait Classification

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 is not linearly separable, a "kernel induced feature space" [2] 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. 6

5.1.1 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.

5.1.2 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.

$$K(x,z) = exp(-\gamma \parallel DTW(x,z) \parallel^2) \tag{1}$$

5.2 Unfixed Method Gait Classification

The framework used for classifying gait cycles is a Gaussian Mixture Model-Universal Background Model (GMM-UBM). The algorithm can be split into three primary sections: UBM training, user gait model genertion, 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 ω represents the mixture weight, μ represents the mean, and Σ represents the covariance matrix. In other words ω is the prior distribution of waling patterns while μ and Σ 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 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. The individual user's gait model is a GMM, but instead of using EM training, a Bayesian adaptation is used. This is simply relating the odds of one event to the odds of another. The use of the Bayesian adaptation is both used for performance as well as for its ability to learn from the data. The Maximum-a-Posteriori (MAP) is the adaptation used. The MAP adaptation adjusts the Gaussian components and mixture weight to personalize the UBM model. It is during runtime when the extracted features are compared against the user gait model and the UBM creating a score. High scores result from the user model being higher than the UBM model and visa versa for low scores. The score are computed as

$$\Delta = log(p(X|\lambda_{user})) - log(p(X|\lambda_{ubm})) \tag{2}$$

In this equation Δ is equal to the users score, X is the feature vector, and λ represents the type of model the feature vector is being compared against. MAP is also ran at runtime to learn new gait patterns from the user. It does the from false negatives given to the device by an alternative form of authentication.

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
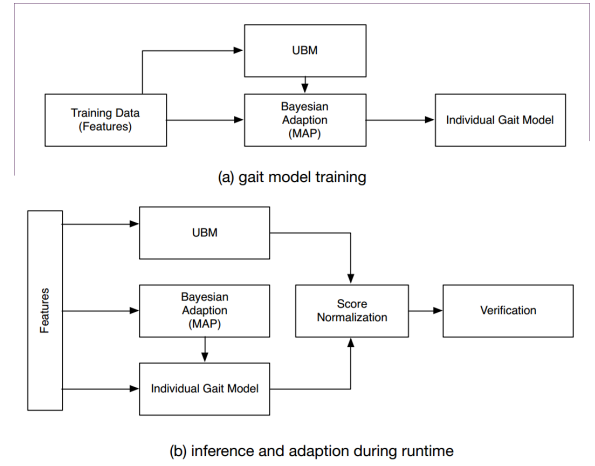

(a) gait model training

(b) inference and adaption during runtime

Figure 7: Algorithm work-flow for (a) generate user gait model by MAP adaptation (b) runtime inference and Individual model adaptation
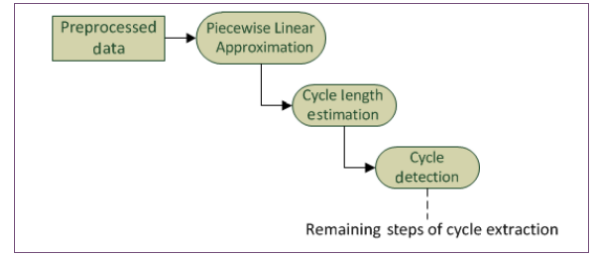

Figure 8: Gait cycle extraction steps with Piecewise Linear Approximation(PLA)

the device while the other used multiple locations. Both use different methods of extracting the gait data as well as well as analyzing the data, but each method does use a popular Android phone. Both of these experiments are not perfect yet but they both show promise in using walking patterns as a form of mobile security.

6.1 Fixed Method Experiment

The Fixed method did two experiment setups: Template-based classification and Machine learning based classification. The data set was made from 51 subjects (39 males, 12 females) that walked down a 18.5 meter long corridor.

6.1.1 Template-based classification

For the template-based classification a interpolation of 100 Hz was used. Also, for this experiment a Piecewise Linear Approximation (PLA) was used before the cycle length estimation step. A PLA is a function that is used to approximate a curve connecting points along the curve with straight lines (interpolating linearly). Also, the PLA uses the Sliding Window And Bottom-up (SWAB) approach. The SWAB approach starts at the leftmost detected segment, applies the PLA function, and moves onto the new set of data using the sliding window approach. It is this PLA representation of the walks that will undergo the remaining steps. From the final processed data, the best cycle with the lowest DTW distance is chosen. This cycle is known as the feature cycle.

the remaining cycles are known as probe cycles. Once all the reference and probe cycles are computed two classes are made: genuine and impostor. These classes are made by comparing the two classes against one another in respect to their DTW distances. If at least 50% of one walk votes for the genuine class then accept, else reject. The accuracy of this method is measured by the Equal Error Rate (EER). This is the rate ate which both the acceptance and rejection errors are equal. For this template-based classification the EER was 22.49% for same-day, 29.4% for different-day and 33.3% for mixed days.

6.1.2 Machine learning based classification

The machine learning classification uses the normal approach to processing the data along with the added approaches described in section 5.1. Once the gait cycles are extracted they are separated, 80% go into a training data set while 20% go into a testing data set. The DTW distance matrix is computed using the training data set.

6.2 UnFixed Method Experiment

The Unfixed method experimented using 3 data sets all using android phones. The first data set included 49 people, 19 female and 28 male. These subjects performed various tasks that included: standing, sitting, walking, biking, running, driving, and random movements. The second data set was made up of 12 people, 5 female and 7 male. These subjects did tests in a more controlled environment with different phone placements, walking distances, and walking speeds. The third data set was to collect data in a real-world setting. This involved 8 people that collected data for 5 hours each.

7. CONCLUSION

I have compared two forms of gait recognition methods using very similar strategies to create an unobtrusive security process for mobile phones using stock android phones. The difference between these two methods are: placement of the mobile device, and their different implementations of of the strategies to which their algorithms were built. Though the fixed method was more geared to a controlled setting and the unfixed method was geared to more real world implementations, they both show promise for future work.

The fixed method used a new method to extract gait cycles using Piecewise Linear Approximation. The fixed method also used two new approaches using Support Vector Machines and Dynamic Time Warping to classify gait cycles. Though only gait cycles of the same length were used, DTW allows for future work with gait cycles of variable lengths. Future work will involve recording at higher sampling rate (varying frequencies) as well as placing the phone in more realistic positions on the body.

The unfixed method used a different approach when dealing with gait data. There was the use of MAP and Gaussian Mixture Models, customizing the gait model. Future work will involve training the UBM better by sampling from a larger data set. There are also trade-offs to be made balancing performance over usability. Using a longer smoothing window could help improve accuracy but will take longer to verify the user.

8. ACKNOWLEDGMENTS

9. REFERENCES

[1] H. Lu, J. Huang, T. Saha, and L. Nachman. Unobtrusive gait verification for mobile phones. In *Proceedings of the 2014 ACM International Symposium on Wearable Computers*, ISWC '14, pages 91–98, New York, NY, USA, 2014. ACM.

[2] M. Muaaz and R. Mayrhofer. An analysis of different approaches to gait recognition using cell phone based accelerometers. In *Proceedings of International Conference on Advances in Mobile Computing & Multimedia*, MoMM '13, pages 293:293–293:300, New York, NY, USA, 2013. ACM.

[3] M. Sujithra and G. Padmavathi. Next generation biometric security system: An approach for mobile device security. In *Proceedings of the Second International Conference on Computational Science, Engineering and Information Technology*, CCSEIT '12, pages 377–381, New York, NY, USA, 2012. ACM.

[4] Wikipedia. Dynamic time warping — wikipedia, the free encyclopedia, 2014. [Online; accessed 25-October-2014].

[5] Wikipedia. Upsampling — wikipedia, the free encyclopedia, 2014. [Online; accessed 25-October-2014].