Gait-ID on the Move: Pace Independent Human Identification Using Cell Phone Accelerometer Dynamics

Felix Juefei-Xu[1,2], Chandrasekhar Bhagavatula[1,2], Aaron Jaech[1], Unni Prasad[1,2], and Marios Savvides[1,2]

[1]CyLab Biometrics Center
[2]Department of Electrical and Computer Engineering
Carnegie Mellon University, Pittsburgh, PA 15213, USA

`{juefeix,cbhagava,ajaech,uprasad}@andrew.cmu.edu, msavvid@ri.cmu.edu`

Abstract

*In this paper, we have proposed a robust, acceleration based, pace independent gait recognition framework using Android smartphones. From our extensive experiments using cyclostationarity and continuous wavelet transform spectrogram analysis on our gait acceleration database with both normal and fast paced data, our proposed algorithm has outperformed the state-of-the-art by a great margin. To be more specific, for normal to normal pace matching, we are able to achieve **99.4%** verification rate (VR) at 0.1% false accept rate (FAR); for fast vs. fast, we are able to achieve **96.8%** VR at 0.1% FAR; for the challenging normal vs. fast, we are still able to achieve **61.1%** VR at 0.1% FAR. The findings have laid the foundation of pace independent gait recognition using mobile devices with high accuracy.*

1. Introduction

Over the past years, more and more studies on various biometrics modalities have emerged. In addition to traditional biometrics traits like face, iris, fingerprints, *etc.*, novel biometrics traits such as gait, electroencephalography (EEG), electrocardiography (ECG), electromyography (EMG), ocular, *etc.* have drawn more and more attention recently. With the development of microchips, microelectromechanical systems (MEMS) and Cloud technology, the mobile devices have become much more powerful in computing, equipped with more sensors and transmit signal faster. Sophisticated recognition algorithms can now be run on board the mobile devices easily, making biometrics identification approachable to everyone. For example, by using the accelerometer sensor in the cell phone, we are able to capture a person's walking pattern. As it turns out, these patterns are very good biometric traits for people identification. Because it does not require any special devices, the gait biometrics of a subject can even be captured without

him or her knowing. This is a very good example of non-cooperative biometrics.

2. Related Work

Gait identification often falls into two categories, image based identification and accelerometer based identification. Image based approaches try to recognize the subject from a video of their gait, usually from some type of surveillance camera, whereas accelerometer based approaches rely on some type of accelerometer to be placed on the subject to collect data from. While image based approaches are attractive due to the fact that the subject can be identified without the subject needing to wear any type of sensor, these approaches suffer from many problems. One common way to identify gait patterns in images is to compute the Gait Energy Image (GEI) and to classify it [6]. However, one of the most prominent problems in using the GEI is recognizing a subject across multiple views. As the training often does not contain every view of a subject's gait, this is a difficult problem. Some methods have been proposed to account for viewpoint changes in gait recognition [9, 14]. These methods also suffer from some of the problems that any image processing algorithm suffers from like occlusions. Any occlusions, the subject holding a briefcase, for example, causes a significant change in the GEI leading to lower recognition rates. As a result, recognition rates decline in the presence of occlusions [14]. Low image resolution can cause problems as well though methods to perform recognition on low resolution images have been proposed [15]. Accelerometer based approaches do not have these problems as they are inherent to image based approaches.

Gait identification methods using accelrometers have performed fairly well in the past [5, 10, 11]. Pan *et al.* [10] attach five different accelerometers to a subject and show that gait recognition can be improved by fusing all of the data together. While this approach gives high accuracy results on a dataset of 30 subjects recorded multiple times, the

usefulness of such results are diminished by the extensive setup needed to collect the data. Similarly, Gafurov *et al.* [5] attach multiple sensors to a person, one at the hip and one on the ankle, to obtain results on a dataset of 50 subjects. The subjects are asked to walk normally with and without a backpack to see how it impacts the gait pattern. Rong *et al.* [11] only use one sensor attached to the waist over a dataset of 21 subjects. While these methods show promising results, they use special accelerometer sensors built for this purpose to record the data from the subjects. While these sensors can be very small, it does require the sensor to be attached to the subject, often with their knowledge. Most phones nowadays have accelerometers built in and as a result, they can be used to collect acceleration data from the person carrying it [3, 13]. The advantage being that the sensor, in this case, is a common device that most people have already which makes data collection easier. Chan *et al.* [3] illustrate that data collected from sensors in such phones successfully captures the gait pattern of a person. However, no identification is performed, only detection of such a pattern. Sprager *et al.* [13] use the data collected from the phone for identification of the five subjects to show that different solid surfaces do not seem to impact a subject's gait pattern. However, the phone is still attached to the hip as in many other methods. When attached to the hip, the phone does not shake as it does when in a pocket due to the motion of the pocket. Thus, a source of accelerometer noise is reduced. Shi *et al.* [12] create a system, SenGuard, in which they claim users can be identified from multiple sensor modalities on their smartphones. While accelerometer readings are used as one of the inputs, experiments are only run on a total of 8 subjects with accuracies only shown for 4. Additionally, the position of the phone is not standardized which, in such a small dataset, can result in a position dependent classifier rather than a biometric classifier. An aggregate classification of all the modalities is done at the end while we focus entirely on the accelerometer data. Kwapisz *et al.* [7] perform some experiments where a cell phone is placed in the pocket of the user and is identified from accelerometer readings. They do achieve a high identification rate but only when using segments of data ranging from 5 to 10 minutes. This is not feasible for device security as they claim as either the subject will not be able to access their phone for up to 10 minutes or an intruder will be able to. Authentication is only performed on a total of 5 subjects which is too small to make any claims of robustness. Additionally, in all of these previous works, query patterns are only compared to training patterns at the same pace which may not be known in real world scenarios.

In our approach, we use a commercially available off-the-shelf (COTS) Android smartphone placed in the pocket of the subject. With a reasonably sized dataset (36 subjects), we show preliminary results indicating that not only can smartphones be used to identify a person based on their normal gait but also that there is potential to match gait patterns across different speeds.

3. Android Acceleration Data Collection App

We have built an application using the Android platform to collect the gait acceleration data. We use the API's provided by the Android SDK to develop the application. The package used is *android.hardware.SensorManager* [1]. We attached event listeners to *Sensor.TYPE_ACCELEROMETER* for accelerometer data and *Sensor.TYPE_ORIENTATION* for gyroscope data. Both sensors are registered with *SensorManager.SENSOR_DELAY_FASTEST* for high sampling rate ($\sim$ 10ms). Gyroscope data is used to estimate the orientation of the phone in a user's pocket. This can be used to correct for different orientations of the phone to facilitate uncooperative data collection. However, for the preliminary experiments, this is unnecessary. Accelerometer data is available in x,y and $z-$axes. The data is logged at every sensor event occurrence. As the application is at the user level, it is common for it to be preempted by other high priority applications. Hence, the data collection isn't perfectly periodic. For this reason, we have logged the system clock for every log instance alongside the data. Figure 1(c) shows the screenshot of our Android App.

3.1. Accelerometer Data

The x-axis is horizontal and points to the right, the y-axis is vertical and points up and the z-axis points towards the outside of the front face of the screen. In this system, coordinates behind the screen have negative z values as illustrated in Figure 1(a) [1]. An example of gait data from x, y, and z channels are shown in Figure 1(d).

3.2. Gyroscope Data

x is defined as the vector product of y and z (It is tangential to the ground at the device's current location and roughly points East). y is tangential to the ground at the device's current location and points towards magnetic north. z points towards the sky and is perpendicular to the ground as illustrated in Figure 1(b) [1].

4. Database

In this section, we detail our gait acceleration data collection and pre-processing schemes.

4.1. Database Collection

We have collected gait acceleration data at two different walking speed (normal walking pace and fast walking pace). The subjects are asked to place the Android smartphone in their right pocket with the phone facing outwards and oriented vertically. With the gait data collection software launched, the subjects are asked to walk straight down a hallway and return. The hallway is 83 feet long resulting
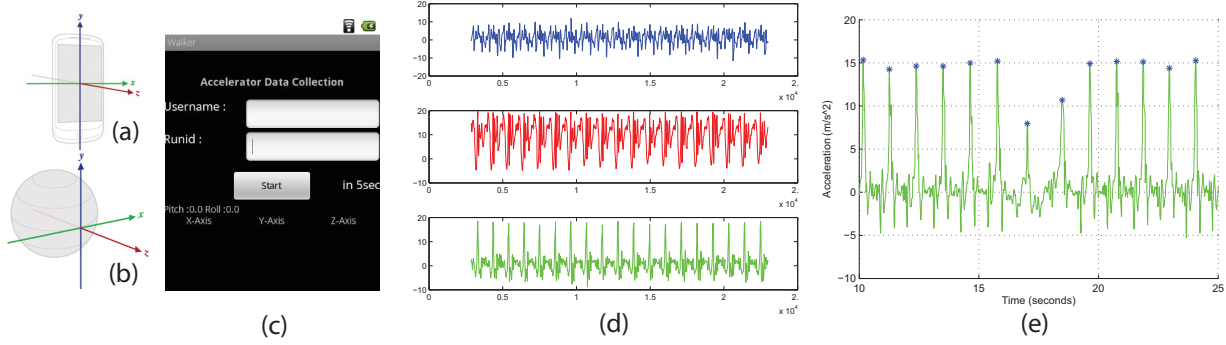
Figure 1: (a): Illustration of the accelerometer sensor tri-axial orientations. (b): Illustration of the gyroscope sensor tri-axial orientations. (c): Screenshot of Android App for acceleration data collection. (d): An example of gait data collected; from top to bottom, it's x, y, and z channel respectively; the horizontal axis is the sample size and the vertical axis is the acceleration (m/s^2). (e): Using gait acceleration data in z-direction for segmenting the steps.

in a total of 166 feet walked during each recording. The subjects are asked to walk at their normal pace for three recordings and at a faster pace for another three recordings. The normal pace speeds range from 3 to 5 feet per second while the fast pace speeds range from 4 to 6 feet per second. In total, 36 subjects are collected at their normal walking speed and 28 of the 36 have provided faster walking speed data. 33 out of the 36 subjects are within the age range of 22-28. There are 28 male subjects and 8 female subjects.

4.2. Preprocessing

The acceleration data is irregularly sampled due to the inability of the sampling program to record when the scheduler gives control to other processes. The distribution of inter-sample times for the typical recording is multi-modal with approximately 25% of the inter-sample times less than $5ms$ and the remainder distributed around $16ms$ with a standard deviation of $4ms$. Since it is convenient in signal processing to work with regularly sampled signals we perform piece-wise cubic spline interpolation and re-sample at $2ms$ intervals or $500Hz$.

The periodic gait pattern is most obvious in the acceleration in the z-axis where there is a large spike in the signal corresponding to the time where the subjects' right foot contacts the floor. These locations, which we will call strike points, are marked with a peak detection function constrained to find local maxima separated by at least $600ms$ and with values greater than $4.5m/s^2$ above the mean as illustrated in Figure 1(e).

Since the gait pattern tends to repeat every step, we choose a feature vector that matches the length of a step as our fundamental unit for classification. In our database the average time between strike points, or when the right foot is raised until it next makes contact with the floor, is about 1 second. Thus each of our feature vectors represent approximately one second of recorded data.

The followings are the two normalization schemes we adopt in our experiments.

4.2.1. Normalization α: Centering Peak

One way to choose a normalization scheme is to take a fixed sized window centered around each strike point. This method avoids distorting the recorded signal but does not compensate for differences in walking speed. The strike points are located using the z-axis acceleration data alone but the feature vector is formed by concatenating the acceleration data from all three directions resulting in a 3-second feature vector. Before concatenating the three directions together each is made to have zero mean and unit norm. Figure 2(a) shows both the normal speed and fast speed walk from two subjects in our database under Normalization α and each subfigure is an ensemble plot of all the steps within one session with peaks centered.

4.2.2. Normalization β: Fixed Period Length

Alternatively, we can select our feature vector by taking the data in the interval between adjacent strike points (peak-to-peak period). The length of this interval is similar but different for every step. Using bi-cubic interpolation we scale the data in the x, y, and z-axes so that each are normalized to a length of 500 samples. Concatenating these together gives us our fixed length feature vector. Just like the previous case, the data from the x, y, and z-axes is adjusted to have zero mean and unit norm before concatenation. Figure 2(b) shows both the normal speed and fast speed walk from two subjects in our database under Normalization β and each subfigure is an ensemble plot of all the steps within one session with fixed period length.

Remark 1 *Using Normalization β makes it possible to match across different walking speed. This is true under the assumption that, within a reasonable pace range, the gait patterns (acceleration) from different paces should only be a stretching or compression over the time, and once they are normalized to be same length, the patterns should look similar. By comparing fast walking patterns and normal walking patterns in Figure 2(a) and 2(b), the patterns of the same subject with different pace do look similar while there is a good deal of dissimilarity among different subjects. However, by normalizing the walking patterns to be a fixed period length, we are losing biometrics information on how long it takes for a subject to finish one step. That is why we also look into Normalization α where the patterns are not scaled over time.*
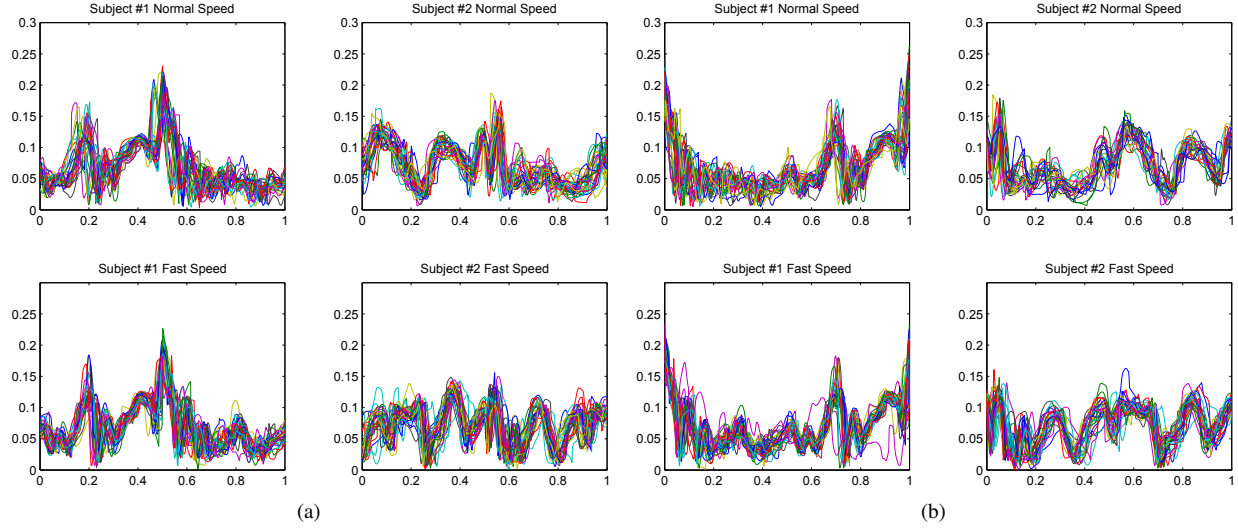
Figure 2: (a): Ensemble plot of steps within one session from two subjects with both normal and fast pace using Normalization α, centering peak. (b): Ensemble plot of steps within one session from two subjects with both normal and fast pace using Normalization β, fixed period length.

4.2.3. Outlier Removal

The performance is improved by removing data from abnormal or irregular steps. For this purpose we compute a similarity matrix using the normalized cosine distance across all the extracted features from each recording. Features whose median distance from the other features in the recording are greater than some threshold are discarded. The average number of steps per sample was 34.4 and the average number of outliers removed per recording 4.3.

5. Continuous Wavelet Transform Time Frequency Spectrogram Analysis

Another approach we investigate is using continuous wavelet transform (CWT) [8] features to identify a subject's gait pattern. The CWT decomposes signals into scaled and shifted versions of a mother wavelet function ψ. The transform is computed as:

$$X(a,b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} x(t)\psi^* \left(\frac{t-b}{a} \right) dt \qquad (1)$$

The Mexican Hat Wavelet has, so far, produced the best identification results. The Mexican Hat mother wavelet is defined as:

$$\psi(t) = \frac{2}{\sqrt{2}\sigma\pi^{\frac{1}{4}}} \left(1 - \frac{t^2}{\sigma^2} \right) e^{\frac{-t^2}{2\sigma^2}} \qquad (2)$$

The CWT is computed on the entire gait signal with portions corresponding to the steps in the signal being extracted. All portions are normalized to the same length. For each subject, the mean and variance of the subject's transform surfaces are computed. The mean and variance of all

of the transform surfaces from the other subjects are also computed. The Fisher ratio, defined as: $F = \frac{(\mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2}$ is a measure of the separation of two distributions. By computing the Fisher ratio of each feature in the transform surfaces between the subject's surfaces and the rest of the surfaces, we find the most discriminative features for each subject as illustrated in Figure 3.

The most discriminative features are selected to create a mask of features for each subject to use for identification shown in Figure 3(d). Each mask is applied to each query transform surface and the distance between the masked mean feature for the corresponding subject and the query surface is computed. The query surface is classified as the subject corresponding to the smallest distance.

6. Cyclostationarity Analysis

In this section, we detail the cyclostationarity analysis in the periodic gait acceleration data using three methods. Ideally, in order to characterize gait acceleration as a biometric trait, (1) steps within one session should be similar and (2) ensemble steps in one session should be similar to the ensemble steps in another session from the same subjects.

6.1. Covariance Method

We start by exploring the second-order statistics of the gait signal. A covariance matrix built upon all the steps in one session is a good indicator of how each dimension of the step feature (*e.g.*, $\Sigma_{ij} = \text{cov}(X_i, X_j) = E[(X_i - \mu_i)(X_j - \mu_j)])$, *each step is a 1500 dimension feature, so* $i, j = 1, 2, ..., 1500$) varies and covaries with corresponding dimensions. The covariance matrix captures the entire second-order statistics and can be used as a feature to inter-
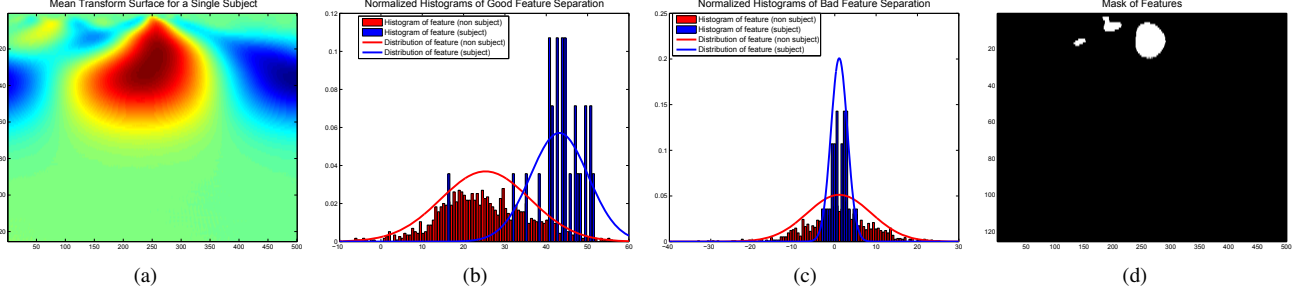
Figure 3: (a): Mean transform surface. (b): Separation of good feature used in classification of a subject. Note the separation of the two lines representing the distributions. (c): Separation of bad feature not used in classification of a subject. Note the lack of separation of the two lines representing the distributions. (d): Mask of 1000 most discriminative features for subject 1.

pret the walking sessions.

6.2. Variance Method

If we treat each dimension of the step feature independently, then we only look at the diagonal elements in the covariance matrix, which is the variance.

The gait recordings can be modeled as a periodic process with multiplicative stationary noise. Assuming multiplicative noise, the variance of the signal at any given phase angle is proportional to the square of the mean acceleration at that phase. Due to the nature of the human gait, small portions of the signal tend to be of much higher magnitude than the mean acceleration. Thus, the variance as a function of phase will be a sparse signal with large values (proportional to the square of the mean) near the peaks of the gait signal and near zero values everywhere else. If we use the variance as a feature for classification, we are essentially capturing information about the moments of the high acceleration portions of the subjects' gait.

Figure 4 shows the variance and corresponding step features. It is noted that the variance is sparse and its peaks correspond to high acceleration portion of the steps.

6.3. Rayleigh Quotient Method

Since the covariance matrix is a good indicator for one walking session by capturing the entire second-order statistics, we employ a Rayleigh quotient method to determine how close one covariance matrix is to another, namely, how close one walking session is to another.

Since in our experiments, 3 sessions of subjects' gait acceleration data are recorded, we use covariance matrix built upon session-3 as the ground truth or the reference of each subject: $\mathbf{G}_1, \mathbf{G}_2, ..., \mathbf{G}_{36}$.

When an unknown query session comes in, we first generate the covariance matrix $\mathbf{C}$ from it. In order to compare covariance matrix $\mathbf{C}$ with ground truth covariance matrix $\mathbf{G}$, we follow Baada's proposed method [2]. Baarda propose to compare the variances of arbitrary function $f = \mathbf{e}^T\mathbf{x}$ determined with a given covariance $\mathbf{C}$ against variances determined by the ground truth covariance matrix $\mathbf{G}$.
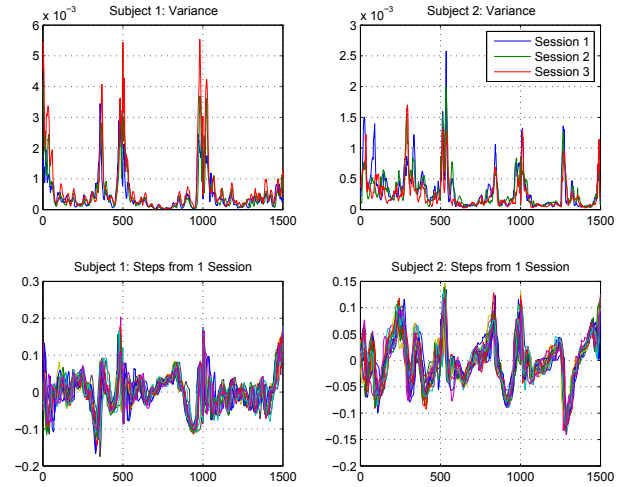

Figure 4: Top 2 subfigures: 3 variance plots from 3 sessions of 2 subjects. Bottom 2 subfigures: ensemble plot of corresponding sessions. The high peaks in the sparse variance correspond to the high acceleration portion in gait data.

One requirement is that the variance $(\sigma_{\mathbf{C}}^2(f))$ of f when calculated with $\mathbf{C}$ to be always smaller then the variance $(\sigma_{\mathbf{G}}^2(f))$ of f when calculated with $\mathbf{G}$:

$$\mathbf{e}^T\mathbf{C}\mathbf{e} \leq \mathbf{e}^T\mathbf{G}\mathbf{e} \iff 0 \leq \lambda(\mathbf{e}) = \frac{\mathbf{e}^T\mathbf{C}\mathbf{e}}{\mathbf{e}^T\mathbf{G}\mathbf{e}} \leq 1 \quad (3)$$

The maximum λ from the Rayleigh quotient can be found by taking $\frac{1}{2}\frac{\partial\lambda(\mathbf{e})}{\partial\mathbf{e}} = 0 \iff \lambda\mathbf{G}\mathbf{e} - \mathbf{C}\mathbf{e} = (\lambda\mathbf{G} - \mathbf{C})\mathbf{e} = 0$. This leads to the maximum eigenvalue $\lambda_{\max}(\mathbf{C}\mathbf{G}^{-1})$ from the generalized eigenvalue problem:

$$|\lambda\mathbf{G} - \mathbf{C}| = 0 \quad (4)$$

Note that $\lambda\mathbf{e}^T\mathbf{G}\mathbf{e} - \mathbf{e}^T\mathbf{C}\mathbf{e} = \mathbf{e}^T(\lambda\mathbf{G} - \mathbf{C})\mathbf{e} = 0$ for all $\mathbf{e} \neq 0$ is fulfilled if Equation 4 holds. This suggests that the eigenvalues of $\mathbf{C}\mathbf{G}^{-1}$ capture the difference of $\mathbf{C}$ and $\mathbf{G}$ completely.

We adopt the sum of the squared logarithms of the eigenvalues to be the distance measurement of query covariance matrix $\mathbf{C}$ and ground truth covariance matrix $\mathbf{G}$:

$$d(\mathbf{G}, \mathbf{C}) = \sqrt{\sum_{i=1}^{n} \ln^2 \lambda_i(\mathbf{G}, \mathbf{C})} \qquad (5)$$

with the eigenvalues $\lambda_i(\mathbf{G}, \mathbf{C})$ from $|\lambda \mathbf{G} - \mathbf{C}| = 0$.

Following this, we are able to calculate $d_1(\mathbf{G_1}, \mathbf{C})$, $d_2(\mathbf{G_2}, \mathbf{C})$,...,$d_{36}(\mathbf{G_{36}}, \mathbf{C})$ for query $\mathbf{C}$, and each d_i is the distance of the query to each ground truth. We then use the distances as a feature in classification tasks as will be discussed in Section 7.3.

7. Experimental Results

In this section, we detail the 3 different experiments on the gait acceleration dataset. Experimental setups and results are discussed in the following subsections individually. In the following experiments, all the receiver operating characteristic (ROC) curves are plotted in semi-logarithm on the axis of false accept rate (FAR) to better illustrate the details at low FAR. Normalized cosine distance is adopted to compute the similarity score between feature vectors.

7.1. Support Vector Machines Experiments

A linear Support Vector Machine (SVM) is trained for each class. As training data for the positive class we take the first 18 features from the input recording. For the negative training data we randomly select four times as many (72) features from the collection of training recordings of the remaining subjects. The trained SVM model also fits a distribution to the distances to the margin enabling us to compute the probability of class membership. Thus given an SVM for class i and a feature s_j we can estimate the likelihood that s_j belongs to $class_i$: $p(class_i|s_j)$. We used the libSVM library [4]. A test recording is classified by taking the class with the maximum log likelihood over all the features extracted from the test recording. $\max_i \sum k \log(p(class_i|s_k))$.

7.1.1. Normal Pace vs. Normal Pace

There are 3 recordings each from 36 subjects in the normal walking speed data set. For the normal speed experiment we train on one randomly selected recording from each class. The test set is comprised of the remaining two recordings from each class. Results reported here are validated with 10-fold repeated random sub-sampling.

Table 1 shows the equal error rate (EER) and verification rate (VR) at 1% false accept rate (FAR) of normal pace vs. normal pace under both normalization schemes. For Normalization β, we also study how each of the x, y, and z channel plays a role in gait recognition tasks. Results from experiments using x, y, and z channel alone, using

	EER	VR
Normal / Normal, α	0.0527	0.9431
Normal / Normal, β	0.0514	0.9417
Normal / Normal, β, x	0.0503	0.9056
Normal / Normal, β, y	0.0527	0.9153
Normal / Normal, β, z	0.0388	0.9097
Normal / Normal, β, xyz	0.0458	0.9417
Normal / Normal, $\beta, |xyz|$	0.0206	0.9458
Normal / Fast, α	0.2099	0.4852
Normal / Fast, β	0.1484	0.8173

Table 1: EER and VR at 1% FAR for experiments using linear SVMs. α and β denote two normalization schemes discussed in Section 4, and x, y, z uses only one channel, xyz uses 3 channels concatenated, and $|xyz| = \sqrt{x^2 + y^2 + z^2}$ is the magnitude norm from three channels.

xyz channel combined and using the magnitude norm of $|xyz| = \sqrt{x^2 + y^2 + z^2}$ are also reported in Table 1. Figure 5(a) and 5(b) show the ROC curves of normal vs. normal pace experiments. For matching normal pace to normal pace, we are able to achieve 94.3% VR at 1% FAR and only 5.1% EER.

7.1.2. Normal Pace vs. Fast Pace

There are 28 subjects recorded at a fast walking speed. For this experiment one normal speed recording from each of these subjects is used for training. Testing is performed on the 82 fast recordings. Two of the 28 subjects only have 2 fast recordings since their third recordings are thrown out due to failure to acquire a usable signal. Table 1 shows the EER and VR at 1% FAR of normal vs. fast pace under both normalization schemes. Figure 5(c) shows the ROC curves of normal vs. fast pace experiments. For matching normal pace to fast pace, we obtain 81.7% VR at 1% FAR and 14.8% EER using Normalization β, fixed period length. Matching across different paces using Normalization β gives much better results than Normalization α, which is expected. We have stated the reason in Remark 1.

7.1.3. Discussions

From Table 1 and Figure 5, we have the following observations. Considering more channel information is better than just using one channel alone, and there is an obvious crossing-over between combining xyz channels and the magnitude norm of xyz at around 1% FAR, so if the application environment can relax the FAR to be greater than 1%, we prefer to use the magnitude norm because by taking the norm, some level of noise is smoothed out.

7.2. CWT Time-Frequency Spectrogram Analysis Experiments

We experiment with varying numbers of features to select from the surfaces. We find a similar pattern emerges in both test cases, normal speed vs. normal speed, and fast vs. fast. In both cases, the accuracies seem to plateau after a sharp rise in accuracy after a number of features are selected as shown in Figure 6.
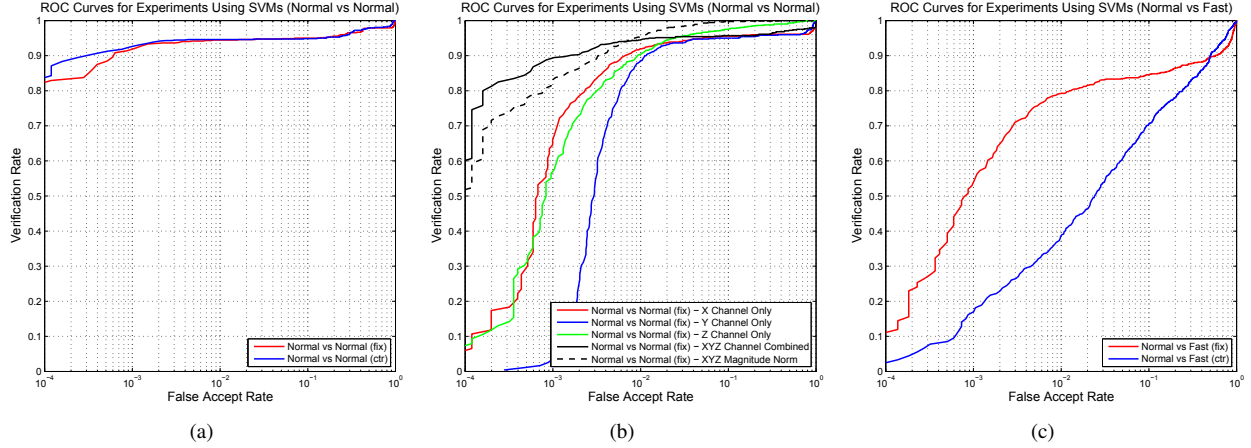
Figure 5: ROC curves for experiments using SVMs under Normalization α (ctr) and β (fix).
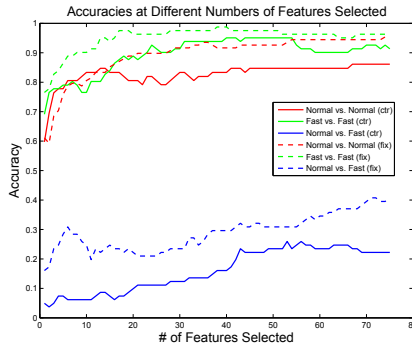

Figure 6: Accuracies for varying number of features training and testing on normal speed data, training and testing on fast speed data, and cross-speed matching with a global training set from both speed. Both of the two Normalization schemes α (ctr) and β (fix) are evaluated.

For normal vs. normal pace experiments, the accuracy converges at about 86% under Normalization α, and 95% under β. For fast vs. fast pace experiments, the accuracy converges around 96% under α and 98% under β. For normal vs. fast pace experiments under Normalization β converges at 40%, much better than under α as expected for the same reasons discussed in Remark 1.

7.3. Cyclostationarity Analysis Experiments

The covariance cyclostationarity analysis has been detailed in Section 6 and using the (1) variance method, (2) covariance method, and (3) Rayleigh quotient method for feature extraction, we are able to run various verification experiments on gait acceleration data and search for the best feature built upon second-order statistics of the data. For the variance and the covariance method, all the sessions from all the subjects are matched against each other in 1-to-1 verification fashion. For the Rayleigh quotient method, we use

	Var.		Whole Cov.		RQ Method	
	VR	EER	VR	EER	VR	EER
N/N, α	0.914	0.046	0.914	0.056	0.620	0.152
N/N, β	0.994	0.036	0.914	0.038	0.605	0.136
F/F, α	0.952	0.031	0.968	0.015	0.754	0.079
F/F, β	0.921	0.031	0.944	0.022	0.702	0.076
N/F, α	0.381	0.381	0.421	0.206	0.079	0.369
N/F, β	0.611	0.141	0.587	0.157	0.361	0.206

Table 2: VR at 0.1% FAR and EER for experiments on cyclostationarity analysis. The first portion of the experiments uses only the variance as feature (diagonal of the covariance matrix), the second portion uses the whole covariance matrix as features, and the third portion uses Rayleigh quotient (RQ) method to extract feature.

session 3 as ground truth and all session 1 and session 2 from all the subjects are matched against each other also in 1-to-1 verification fashion.

Table 2 shows the VR at 0.1% FAR and EER of normal vs. normal pace, fast vs. fast pace, and normal vs. fast pace. All the experiments are carried out under both normalization schemes α and β using all 3 feature extraction methods. Figure 7 shows the ROC curves for the entire cyclostationarity analysis experiments.

For matching normal to normal pace, we are able to achieve 99.4% VR at 0.1% FAR and only 3.6% EER. For fast vs. fast pace, we are able to achieve 96.8% VR at 0.1% FAR and only 1.5% EER. For the challenging normal vs. fast experiments, we are still able to achieve 61.1% VR at 0.1% FAR and 14.1% EER.

7.3.1. Discussions

Cyclostationarity analysis experiments give very good results as shown in Table 2 and Figure 7. The second-order statistics of the gait time series captured by the covariance matrix is a very good feature in distinguishing among different subjects both with same pace and different pace. Normalization α is better at same-pace gait recognition, while
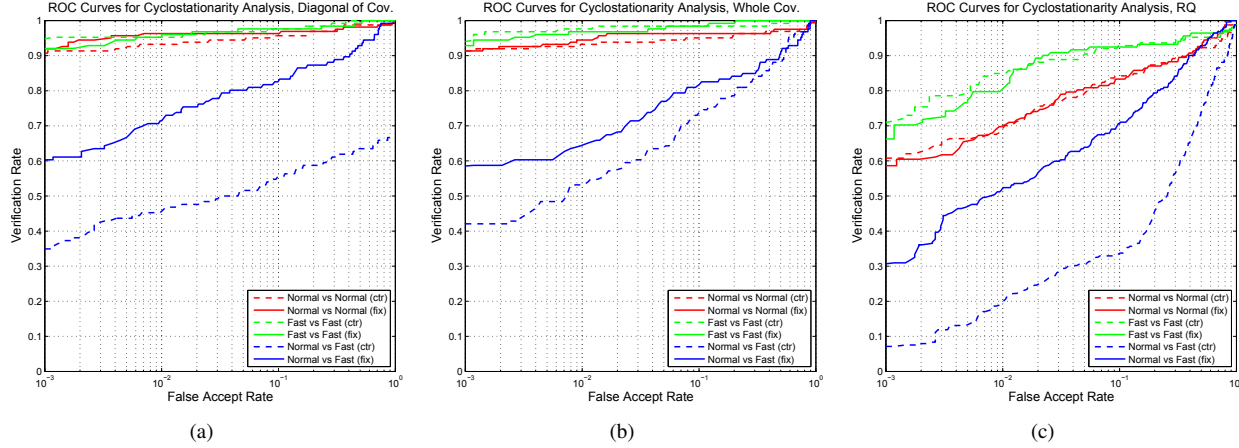
Figure 7: ROC curves for cyclostationarity analysis experiments under Normalization α (ctr) and β (fix).

Normalization β is better at cross-pace gait recognition. We can also observe that fast vs. fast matching yield better verification results than normal vs. normal matching although they are within the same-speed matching framework. This is because, when subjects are asked to walk faster, their gait acceleration data become more distinct at the peaks where strides are made. If the subjects are walking extremely slowly, the accelerometer would barely catch any peaks, thus making it almost impossible to perform matching task. Still, if subjects walk too fast, his or her walking patterns will significantly change, thus our assumption in Remark 1 using Normalization β will no longer hold.

8. Conclusions

In this work, we have developed an Android App for gait acceleration data collection. Using the Android smartphone, we have collected normal pace and fast pace gait acceleration data from 36 subjects with multiple sessions. Then two normalization schemes are adopted to normalize the segmented gait data. In Normalization α, peaks are centered, and in Normalization β, the peak-to-peak period length is fixed. As mentioned in Remark 1, Normalization β is for the purpose of matching gait patterns across different paces. We have conducted 3 different experiments using (1) an SVM method, (2) a CWT time-frequency spectrogram analysis method, and (3) a cyclostationarity analysis method. Our best results have shown that for matching normal to normal, we are able to achieve **99.4%** VR at 0.1% FAR; for fast vs. fast, we are able to achieve **96.8%** VR at 0.1% FAR; for the challenging normal vs. fast, we are still able to achieve **61.1%** VR at 0.1% FAR. If the FAR is even relaxed, better VR can be achieved as shown in all the ROC curves. So far, we can safely draw the conclusion that, with gait data collected from a COTS Android smartphone, a robust acceleration based, pace independent gait recognition is possible. The findings of this paper can also serve as an asset for future pedo-biometrics research.

References

[1] Android Developers. API (public class:sensormanager). http://developer. android.com/reference/android/hardware/SensorManager.html. 2

[2] W. Baarda. S-transformations and criterion matrices. *Band 5 der Reihe 1. Netherlands Geodetic Commission*, 1973. 5

[3] H. Chan, H. Zheng, H. Wang, R. Gawley, M. Yang, and R. Sterritt. Feasibility study on iphone accelerometer for gait detection. In *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2011 5th International Conference on*, pages 184–187, may 2011. 2

[4] C.-C. Chang and C.-J. Lin. Libsvm: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.*, 2(3):27:1–27:27, May 2011. 6

[5] D. Gafurov, E. Snekkenes, and P. Bours. Gait authentication and identification using wearable accelerometer sensor. In *Automatic Identification Advanced Technologies, 2007 IEEE Workshop on*, pages 220–225, june 2007. 1, 2

[6] J. Han and B. Bhanu. Individual recognition using gait energy image. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 28(2):316–322, feb. 2006. 1

[7] J. R. Kwapisz, G. M. Weiss, and S. A. Moore. Cell phone-based biometric identification. In *Biometrics: Theory Applications and Systems (BTAS), 2010 Fourth IEEE International Conference on*, pages 1–7, sept. 2010. 2

[8] W. C. Lang and K. Forinash. Time-frequency analysis with the continuous wavelet transform. *American Journal of Physics*, 66(9):794, Sept. 1998. 4

[9] N. Liu, J. Lu, and Y.-P. Tan. Joint subspace learning for view-invariant gait recognition. *Signal Processing Letters, IEEE*, 18(7):431–434, july 2011. 1

[10] G. Pan, Y. Zhang, and Z. Wu. Accelerometer-based gait recognition via voting by signature points. *Electronics Letters*, 45(22):1116–1118, 22 2009. 1

[11] L. Rong, Z. Jianzhong, L. Ming, and H. Xiangfeng. A wearable acceleration sensor system for gait recognition. In *Industrial Electronics and Applications, 2007. ICIEA 2007. 2nd IEEE Conf. on*, pages 2654–2659, may 2007. 1, 2

[12] W. Shi, J. Yang, Y. Jiang, F. Yang, and Y. Xiong. Senguard: Passive user identification on smartphones using multiple sensors. In *Wireless and Mobile Computing, Networking and Communications (WiMob), 2011 IEEE 7th International Conference on*, pages 141–148, oct. 2011. 2

[13] S. Sprager and D. Zazula. Impact of different walking surfaces on gait identification based on higher-order statistics of accelerometer data. In *Signal and Image Processing Applications (ICSIPA), 2011 IEEE International Conference on*, pages 360–365, nov. 2011. 2

[14] D. Tao, X. Li, X. Wu, and S. Maybank. General tensor discriminant analysis and gabor features for gait recognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29(10):1700–1715, oct. 2007. 1

[15] J. Zhang, J. Pu, C. Chen, and R. Fleischer. Low-resolution gait recognition. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 40(4):986–996, aug. 2010. 1