PART 1

Honors Thesis

WALKING THE WALK: AN EXPLORATORY ANALYSIS IN BIOMETRIC GAIT RECOGNITION

Department of Mathematics

Brigham Young University

December 2016

Advisor: Jeffrey Humpherys Honors Coordinator: Paul Jenkins

ABSTRACT

WALKING THE WALK: AN EXPLORATORY ANALYSIS IN BIOMETRIC GAIT RECOGNITION

David S. Kartchner

Department of Mathematics

Bachelor of Science

With the increasing need for secure, electronic identification and authentication, many organizations are turning to biometrics (e.g. facial recognition, fingerprint scans, etc.) as means of identifying individuals. This thesis examines various techniques for biometric gait recognition using accelerometer data collected using ordinary pedometers. It expands on past literature by developing a method through which individual samples can be aggregated together to provide a more robust identification framework. We find that both this new framework and thoughtful feature engineering can produce better accuracies than have been described in the literature without significantly improving computational complexity.

ACKNOWLEDGEMENTS

Special thanks to Jeff Humpherys for being so flexible with me as I prepared my thesis, to Julie Radle for inspiring my intellectual curiosity, and to Thalia Pope for pushing me to write this thesis when I otherwise would have given up.

Contents

List of Figures

List of Tables

1 Introduction

"Biometric identification" refers to "the process by which a person's unique physical and other traits are detected and recorded by an electronic device or system as a means of confirming identity" [2]. Beginning with fingerprinting as a means of identifying individuals and catching criminals, biometrics have expanded to include numerous identification methods used for many purposes. DNA testing, for example, has been used in parentage tests to determine both guilt in rape cases and fathers who need to pay child support. Similarly, facial recognition has been used by police forces to identify criminals from photographs using databases of convicted felons.

While the term "biometrics" may not be widely known, humans use similar methods automatically and instinctually to determine how to interact with the world. The human brain, for example, uses facial features such as complexion, cheek bone structure, nose and mouth shape, and jawline to recognize faces and distinguish between friends, enemies, and unknown individuals. Similarly, it is usually takes only a few words to identify a close friend or family member on the telephone even in the absence of caller ID. While these tasks may seem insignificant, they are actually learned processes that enable the mind to quickly access relevant information (e.g. an individual's likes, dislikes, personality traits, marital status, etc.) in order to efficiently and productively navigate complex social situations with

multiple agents.

Biometrics are particularly useful for identification because they provide unique identification and are difficult to fake. It is much more difficult to steal a fingerprint than a pin, thus making biometric identification a powerful, secure means of unique identification. To quote one author, "biometrics are what you are, not what you have" [3].

Biometric identification has recently proliferated with advances in computational capacity. One potentially promising application of biometrics is in gait recognition, which is the process of uniquely identifying individuals by the way they walk, run, bike, etc. Though the word "gait" generally evokes an image of someone walking, gait more generally refers to "the coordinated, cyclic combination of movements that result in human locomotion" [4]. As it turns out, each individual produces slightly different patterns of muscle movement and acceleration while walking that potentially enable him or her to be uniquely identified. Such gait recognition offers large potential benefits to both law enforcement agencies, who could use it as a means of identifying criminals, and to medical companies, who could use it to create fall detection devices for elderly patients. Gait recognition poses similar potential benefits in improving the security of mobile devices by being used as an additional means of identification. In each of these cases, gait recognition

can be particularly attractive because it can be done passively and unobtrusively using something as simple as a pedometer or an ordinary smartphone.

2 Literature Review

By and large, the current work on biometric gait recognition can be subdivided into machine vision based and accelerometer based approaches. Machine vision based gait recognition usually uses a silhouette based approach in which a silhouette of the individual walking is segmented from his or her background and then compared for similarity against other samples generated by the various individuals[12]. This type of recognition, however, runs into problems when an individuals movement is obstructed by his or her clothing or shoe type, or by the terrain he or she is walking on. Further problems can result when an object being carried obstructs either gait or an individuals silhouette, or when the camera angle at test time is different than in previous samples for an individual. These challenges have been shown to reduce accuracy of computer vision based approaches to gait recognition. Despite these challenges, such gait recognition can still be used with reasonable effectiveness to identify individuals, especially when used in tandem with facial recognition or other identification methods. Using such dual identification has been shown in some studies to have accuracies well above 90% using a single camera

[12].

The second approach to biometric gait recognition, and the one which we address in this paper, is performed through very different means using an accelerometer placed somewhere on the subjects body. Rather than measuring visual indicators of gait, accelerometer based recognition identifies individuals by measuring acceleration along the three coordinate axes for each individual and identifying them by the particular patterns of acceleration they generate. This type of gait analysis can easily be done with a pedometer, Fitbit, or smartphone, each of which comes equipped with all of the hardware necessary to measure acceleration. Accelerometer based gait recognition also poses promising solutions to fitness tracking, as it would allow personal fitness devices to more effectively track what activity an individual was performing without requiring user input [11].

Two primary approaches to accelerometer based gait recognition are described in the literature. The first is a purely statistical approach in which the data are divided into samples of fixed duration (usually 2-3 seconds) and various features are generated from each sample [5]. Common features include the mean, standard deviation, minimum, maximum, value range (calculated as the difference between the max and min), and histogram distribution of acceleration along each axis, as well as the correlation between x, y, and z components of acceleration. Mannini and

Sabatini also note that in generating features, individuals should consider both the highly variable portions of acceleration associated with quick bodily movements (the "AC" component of acceleration) as well as the relatively constant portion of acceleration that comes primarily from gravity (the "DC" component of acceleration). They propose parsing these frequencies using the short-time frequency transform as well as either the continuous or discrete wavelet transform. They also note that one can use frequency domain entropy as a proxy for the complexity of a given signal, since signals with more underlying frequencies will have larger frequency domain entropy than those that have fewer underlying frequencies (e.g. walking would have higher frequecy domain entropy than biking, since walking generates sharp changes in acceleration when the foot strikes the ground whereas acceleration generated while biking is smooth and continuous) [11]. Nickel proposes using Mel-frequency and Bark-frequency cepstral coefficients to achieve similar results [5].

The second approach to gait recognition is a template based approach, which focuses on extracting gait cycles (i.e. two steps) from accelerometer data and comparing them based on magnitude and shape. To do so, the computer leans a gait template for each individual and compares extracted gait cycles to each pattern in the database to see whom they match most closely. Nickel noted that this

approach seems to be less accurate than the previously described statistical approach, but we note that her approach is limited in that it compares gait cycles to a single template cycle generated for each individual, rather than comparing them to a variety of samples using the L1 norm or a k-nearest-neighbors approach. [5] We address both this template based approach and the statistical approach described above in our paper.

3 Data

3.1 Data Description

To explore the question of biometric gait recognition, we used a dataset of accelerometer readings compiled by Matthew Webb, a masters student in the Brigham Young University Department of Mathematics. The data are comprised of accelerometer readings taken from a pedometer placed in a pants pocket while individuals engaged the following physical activities:

1. Regular Walking
2. Slow Walking
3. Speed Walking

4. Jogging

5. Stair Climbing

6. Biking

7. Faking (i.e. shaking pedometer to simulate walking)

Acceleration readings were taken at a rate of either 50 or 400 samples per second. Relevant data columns in the raw data, as well as the derived quantity A_{tot}, are summarized in Table 1.

3.2 Data Cleaning

Since our accelerometer data was collected at two different sampling rates, we began by downsampling the data to simulate sampling each individual 50 times/second. To do so, we separated the data by individual and took the arithmetic mean of each 8 consecutive observations for every individual sampled at a rate of 400 times/second. Doing so had a regularizing effect on the acceleration readings, significantly smoothing high-frequency oscillations and suppressing erratic spikes. A comparison of gait cycles before and after downsampling can be seen in Figure 1.

After normalizing the data to a single sample rate, we identified and deleted segments of the data where it was obvious that the individuals being measured were

Data Columns

PersonID	A label corresponding to the person being sampled. Contains 42 distinct values.
ActivityID	A numerical label corresponding to the activity being performed. Categories include walking, running, jogging, stair climbing, biking, and faking(shaking pedometer to simulate walking).
Samplerate	Number of acceleration measurements taken each second. Either 50 or 400.
A_x	Acceleration along pedometer's x axis at a given time.
A_y	Acceleration along pedometer's y axis at a given time.
A_z	Acceleration along pedometer's z axis at a given time.
A_{tot}	Magnitude of acceleration at a given time. Calculated as $A_{tot} = \sqrt{A_x^2 + A_y^2 + A_z^2}$.

Table 1: Data columns used for biometric gait recognition

not actually walking. To do so, we eliminated one-second segments of the smoothed acceleration signal during which the standard deviation of acceleration was less than 5% of the average standard deviation of acceleration across all such segments. We chose to remove segments with low standard deviation since these segments represent times when little apparent movement was taking place. Though we cannot ensure that our approach eliminated all segments of data where the individuals in

Figure 1: Three seconds of data from an individual in our dataset before and after sampling. Note that the curves are almost identical except at places where the acceleration spikes suddenly.

question were not walking due to lack of appropriate reference, our algorithm was able to identify and remove the segments in which it was immediately obvious that no activity occurred. After deleting these segments of inactivity, we deleted the beginning and ending 5 seconds of data from many of our samples to account for the time it may have taken for individuals to place the pedometer in their pocket at the start of the experiment and subsequently return it to the proctor after completion.

One limitation of our data is the lack of consistent orientation for acceleration along the x, y, and z axes. While many accelerometers have a gyroscope to keep

track of vertical and horizontal axes, our values of A_x, A_y, and A_z were determined arbitrarily by the orientation of the accelerometer in each individual's pocket. Moreover, though most machine learning models for gait recognition are trained on both horizontal and vertical components of acceleration, trying to do so with our model could lead to over-fitting since the sign and magnitude of each acceleration vector is more dependent on orientation than actual acceleration. To remedy this problem, we generated an additional data column A_{tot}, which is the magnitude of acceleration in each given sample. While training our models on total acceleration will miss the nuances of hip gyration, doing so allows us to produce more robust, generalizable results.

3.3 Feature Generation

3.3.1 Statistical Approach

After cleaning our data, we generated features with which to train machine learning classification algorithms. We first generated features for the purely statistical classification approach, which consists of splitting gait samples into into fixed length subsets and then extracting features consisting of relevant statistical moments and properties. These features are then used to train a classification model such as a support vector machine or random forest, which is then used to

classify test samples accordingly. For our model, we separated our data into three-second samples and generated the following features:

- **Mean:** The mean of the magnitude of acceleration in the sample.
- **Standard Deviation:** The standard deviation of the magnitude of acceleration in the sample.
- **Min:** The minimum value of the magnitude of acceleration in the sample.
- **Max:** The maximum value of the magnitude of acceleration in the sample.
- **Value Range:** The difference between the maximum and minimum value of the sample
- **Root Mean Square (RMS):** The root-mean-square of the acceleration samples in our dataset. Calculated as $RMS = \sqrt{\frac{A_{tot_1}^2 + A_{tot_2}^2 + \cdots + A_{tot_n}^2}{n}}$, where n is the size of our sample.
- **Sign Changes:** The number of sign changes in our sample. Calculated after centering sample around 0 by subtracting sample mean.
- **Histogram Distribution:** The histogram distribution of the normalized data. In order to use the same histogram bins across features, we scale the acceleration to have minimum 0 and maximum 1, then take the distribution of

this scaled data in 10 equally spaced bins. Thus, each histogram bin could be thought of as amount of times acceleration is in a certain quantile for a given sample. The count of each histogram bin is an element of the feature vector.

- **Frequency Domain Entropy:** The frequency domain entropy of each sample in our dataset. Calculated by performing the following steps:

  1. Take the Discrete Fourier Transform of our sample to give coefficients f_i
  2. Compute the unnormalized Power Spectral Densities (PSD's) by squaring the magnitude of each Fourier coefficient: $D_i = |f_i|^2$
  3. Normalize the PSD's to represent probabilities by dividing by the sum: $p_i = \frac{D_i}{\sum_i D_i}$.
  4. Use probabilities for standard entropy calculation: $S = -\sum_i p_i ln(p_i)$

3.3.2 Template-based Approach

In order to compare individual gait cycles based on shape, it was necessary to split data into individual gait cycles. Since walking generates repeated, cyclic movements in the legs, arms, hips, and torso, we could expect that the gait cycles generated by these movements would resemble a periodic function over time. However, discrete sampling and subtle variations in walking speed across time

prevent us from simply slicing the data gait cycles at fixed intervals, as each cut may not correspond to the same point in the gait cycle.

To solve this problem, we noted that gait cycles follow a pattern alternating peaks and troughs, generated as the feet of an individual being sampled repeatedly impact the ground (causing high acceleration readings) and then move smoothly through the air (causing low acceleration readings). Unfortunately, simply finding local minima and maxima is still insufficient, since other subtle movements may create numerous local extrema over the course of a gait cycle, making them useless to split on. However, inspection of our data shows that while there are many local extrema over a single gait cycle, there are also particularly salient peaks and troughs, presumably corresponding to the impact and freefall stages of a gait cycle. Therefore, we choose to specify these salient troughs as the start of our gait cycles and identify them using a variation of the min salience vector, as described by Nickel [5]. To do so, we first smooth the signal slightly to reduce the impact random shocks which appear in our data, but appear to result from erroneous acceleration measurements rather than actual movement. We then find all salient minima by calculating all points x_j in our smoothed signal satisfying:

$$x_j = \min_t \{x_{j+t} : t \in [-0.6, 0.6]\}$$

In other words, we say that x_j is a salient minimum if it is lower than all points

gathered within 0.6 seconds before or after x_0. The time interval of 0.6 seconds is chosen specifically to ensure that each salient minimum is smaller than all points in the previous step and subsequent step, since 0.6 seconds is almost always longer than it takes for an individual to take a single step while still being too small a time interval to complete two full steps. We can similarly extract single steps by choosing an interval of 0.3 seconds on each side of x. Examples of gait cycles extracted from various individuals, as well as a visualization depicting the calculation of salient minima, are pictured in Figure 2 and Figure 3. Once cycles our data has been split into gait cycles, each is normalized to lie on the unit interval and interpolated to approximate a gait function.

We note that while our means of cycle segmentation was reasonably robust, it did not perfectly identify gait cycles. Sometimes our algorithm identified a single, long step or a series of 3-4 steps as a cycle, rather than a 2-step cycle as we would expect. To correct this error, we simply limited our samples to gait cycles between 0.8 seconds and 1.5 seconds in length.

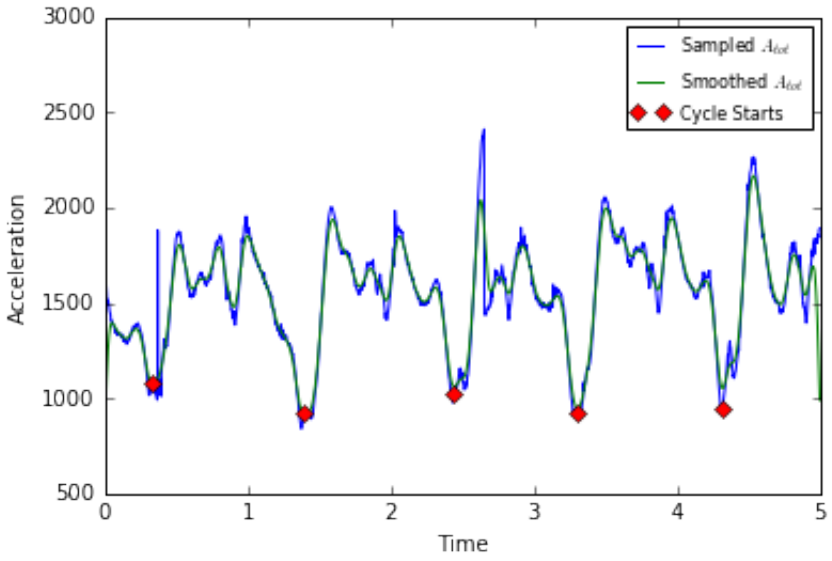

Figure 2: Four consecutive walking gait cycles from an individual in our dataset. Each red dot corresponds to a salient minimum at the beginning of a new cycle.

4 Methodology and Results

4.1 Classification Techniques

4.1.1 Random Forest Classifiers

One of the most common and most robust methods for multi-class classification is random forest classification. To understand random forest classifiers, we must first understand how decision tree classifiers work. A decision tree is a means of classifying categorical data by sorting them into smaller homogeneously labeled subsets based on their attributes. This is done by performing successive,

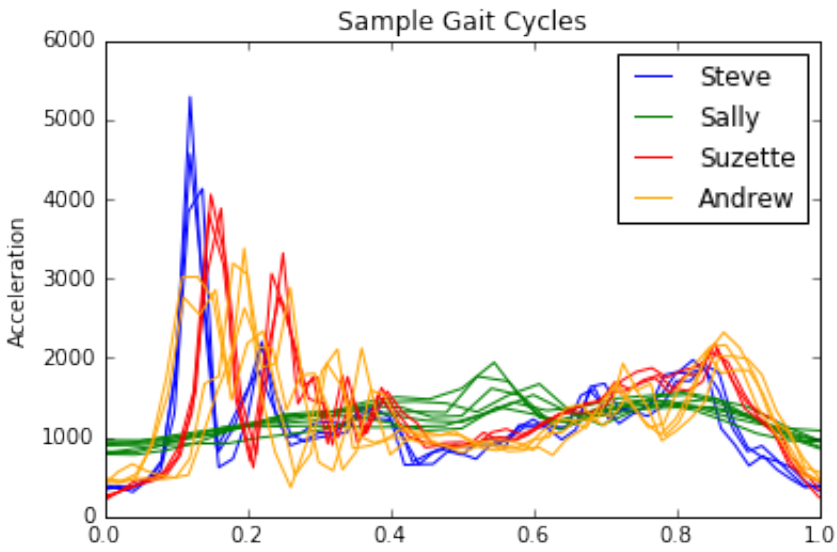

Figure 3: A comparison of the shape of walking gait cycles for four different individuals, each represented by a different color.

Names are chosen randomly and do not reflect actual names of participants.

binary splits on our training data, each which seeks to minimize the label impurity at in each of its child subsets. This creates a binary tree of data with a subset at each node. We define the *impurity* of a given node to be a measure of how likely an item is to be misclassified at our node. Given that our set contains n distinct categories with respective relative probabilities $p_1, p_2, \cdots, p_n$, the impurity I of our set S is given by:

$$I(S) = 1 - \sum_{i=1}^{n} p_i^2$$

To determine where to split the data at a given node, our tree chooses the split that will minimize the weighted sum of the impurities at the child nodes created by the split. This procedure is repeated on each of the child nodes until either some max tree depth is reached or no significant improvements in node purity can be attained. To classify a test sample, a decision tree simply filters it into one of its leaf nodes and assigns it the most common label at that node.

Using our gait recognition problem as an example, it is reasonable to assume that some individuals walk with heavier steps than others, thus causing them to have higher variance in their accelerometer readings. Accordingly, our decision tree may note that certain individuals tend to have higher variance samples, and would split our data into high variance and low variance groups. These two groups would be chosen to make each as homogeneous as possible. Our tree would then look at each of the newly created subsets and seek to classify them based on other features, such as the number of sign changes.

Random forest classifiers are a means of classifying categorical data using a bootstrapped ensemble of decision trees. Each bootstrapped tree is constructed from a random subset of the training data, where each sample only has a subset of the features. Each tree is then trained on this subset of the data. When a random forest classifies a test sample, each bootstrapped tree classifies the test sample

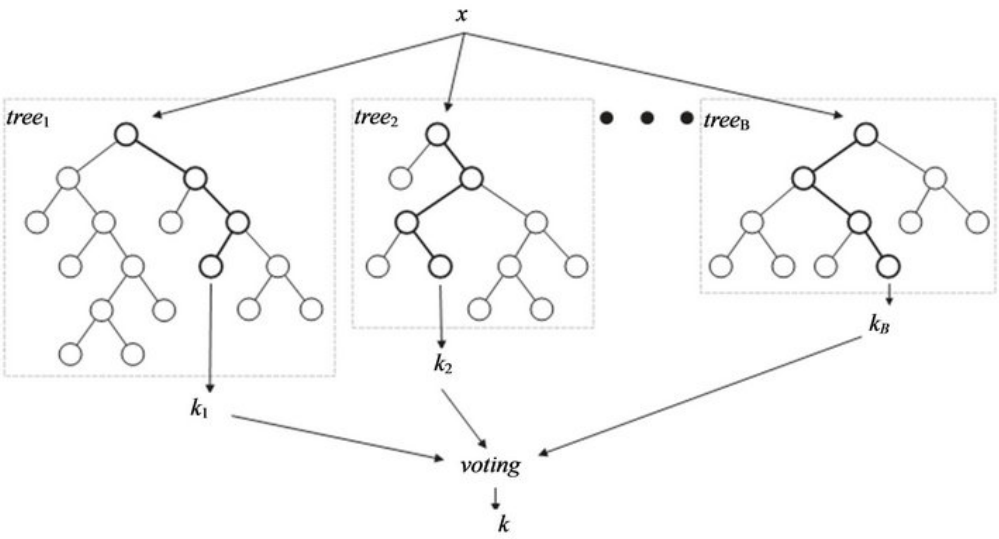

Figure 4: A visualization of how a random forest classifies a new data point with feature vector x. Image courtesy of The Journal of Biomedical Science and Engineering.

independently and then the trees vote on which category to assign the new sample. The sample is then classified as whichever category received the most votes. A visualization of a random forest can be found in Figure 4.

4.1.2 K-Nearest Neighbors

Unlike many machine learning algorithms, nearest neighbors classification schemes seek to classify data points by storing training data rather than building a mathematical model. In the case of K Nearest Neighbors (KNN) classification, each

point x to be classified is represented as a n-dimensional vector in $\mathbb{R}^n$, following which we calculate the k points (where k is a user-specified positive integer) of training data closest to x in our metric of choice, usually the 1-norm or 2-norm. The label of point x is then decided to be the most commonly occurring label among the k nearest points. When a tie occurs between two classes, the chosen class is determined by which class has a smaller mean distance to the point x. Some approaches to k nearest neighbors always use a weighted vote among the k nearest points, where each point y_i's vote weight is inversely proportional to the distance between y_i and x [15]. A 2-dimensional visualization of how K-Nearest classification can be used for multiclass labeling is given in Figure 5.

The KNN approach to data classification is a natural and intuitive choice for template based classification, since the distance between y_i and x is a discrete approximation of the norm of the difference between the gait cycles represented by x and y_i under the L_1 or L_2 norm. In calculating this metric, we consider two different theories regarding individual classification. In the first, simply calculate the the distance between the two gait cycles using the raw acceleration data. In the second, we normalize the data so that acceleration takes on values between 0 and 1 to strictly compare the shape of the gait cycles.

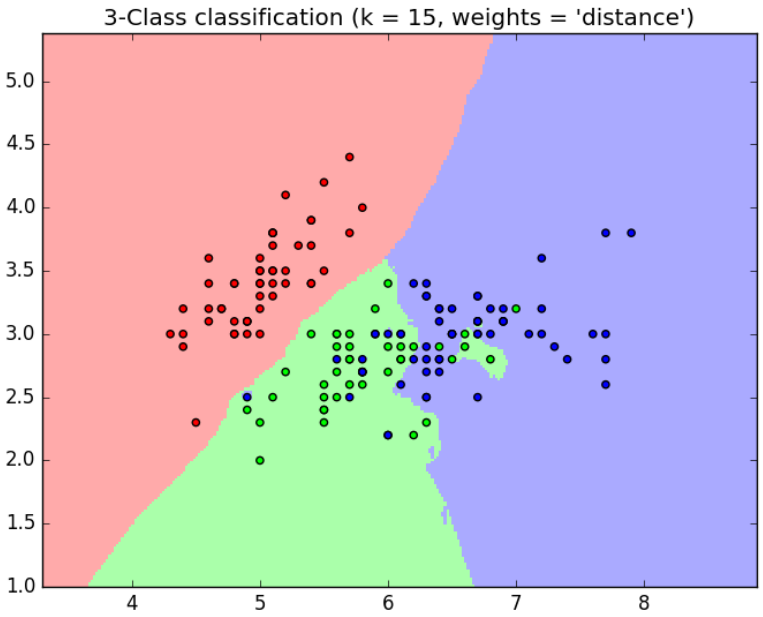

Figure 5: 3-Class identification using K-Nearest Neighbors. Image courtesy of Scikit Learn.

4.2 Person Identification

Now that we have identified pertinent features in our data, we begin by trying to identify individuals who are all participating in the same activity. For our training data, we use walking samples, as this will give us the largest sample size.

4.2.1 Statistical Approach

To build our statistically based model, we began by selecting all of our walking data, splitting it into 3-second chunks, and calculating the statistical properties described in Section 3.3.1. This left us with a total of $6,120$ labeled samples with corresponding feature vectors. We then split the data randomly a training set containing 70% of the data and and test set containing the remaining 30%. In doing so, we ensured that this 70-30 split ratio is maintained for each individual, so that individuals with a small number of samples are not accidentally excluded from our training set. After splitting the data, we used our training data to fit a scikit-learn random forest model containing 100 trees. We used this model to predict labels for our test data and determine our prediction accuracy. For robustness, we repeated this process 10 times and averaged our results. Surprisingly, our model correctly classifies the forty-two individuals in our dataset over 99% of the time.

In light of the suprisingly high classification accuracy of the above trials, we note that we achieved these results by testing a number of approaches in the features engineering process. In calculating the mean, min, max, RMS, and frequency domain entropy of our signal, for example, it was critical that we did so without first normalizing our data. Similarly, in calculating the number of sign changes, we found that it was important to subtract the mean of each sample from

the raw acceleration to allow sign changes to act as a proxy for the amount of rapid oscillation in our signal. These techniques enabled us to distill the maximum possible information content into our features, boosting our classification accuracy from approximately 80% before performing these modifications.

We can determine the most important classification features by looking at how often our forest chose a particular feature to split on. The more often a feature was chosen as a splitter, the more important it is in predicting our outcome. A graph of the relative importance of features in our forest is given in Figure 6.

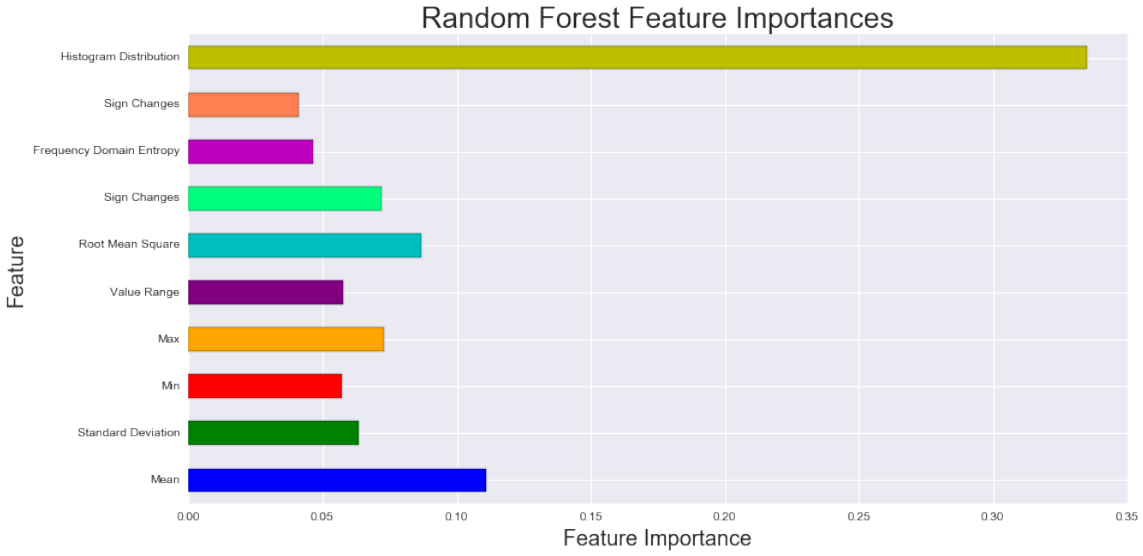

Figure 6: Relative importance of features in our random forest model for person identification.

4.2.2 Template Based Approach

To build our template based model, we split our data into gait cycles as described in 3.3.2 and calculated the mean, min, max, standard deviation, and length of each cycle to use as potential additional features for cycle classification. Before fitting our KNN model, we noticed that a number of classes had less than 100 gait cycle instances while others had well over 1000. This could create classification problems by allowing the class with the most instances to dominate matches by sheer force of numbers. To correct this problem, we trimmed each class to have less than 150 instances in order build a more balanced KNN model. Doing so left us with $3,016$ gait cycles, which we again split 70-30 into train and test data, ensuring that the 70-30 ratio was maintained for each class in our data.

After splitting our data, we fit a KNN and a random forest on the interpolated points from our cycle templates. We repeated this procedure for all odd values for k under 25 and discovered that KNN classification accuracy decreases more or less monotonically with the number of neighbors used, with every trial having the best accuracy when only using one neighbor. We also found that weighting neighbors based on distance from point to be classified improved the accuracy of KNN classifiers with more neighbors, but did not beat the classification accuracy of classifiers using only a single neighbor. KNN and random forest

classifiers had attained similar levels of accuracy.

We repeated our KNN classification trials in using the 1-norm and 2-norm as distance metrics. We also tested a KNN using cycles normalized to have min 0 and max 1 to strictly compare the shape of cycles. We found that classification accuracy changes only marginally with by changing the choice of norm and that normalizing data significantly decreases classification accuracy. In each case, we performed 10 random splits and averaged our results, which are summarized in table 2.

Template-Based Classification Results

KNN w/ 1-norm	76.0%
KNN w/ 2-norm	74.2%
Random Forest	74.5%
KNN w/ 1-norm, Normalized Gait Cycles	70.5%
KNN w/ 2-norm, Normalized Gait Cycles	70.3%
Random Forest, Normalized Gait Cycles	69.5%

Table 2: Classification Accuracy for Template-Based Gait Recognition

4.3 Probability Updating

While our template-based model only achieves mediocre results on individual gait cycles, we note that any individual carrying a cell phone or pedometer would provide many samples which could be used to identify that individual. Accordingly,

it is worth considering how well the results of our individual cycle identification can be aggregated to identify individuals. To test whether individual identification is more successful than single cycle identification, we took 11 gait cycles from our test data for each individual (or fewer if our test data contained less than 11 for that individual) and predicted a class label for each of them. We then took the mode of these predictions and assumed that the individual belonged to the most commonly occurring class. We said that our model correctly predicted the individual if they were both placed in the correct class and if more than half of the individual labels corresponded to the correct category. We repeated this procedure 50 times with random train/test splits and averaged the results for both our best KNN and best random forest parameters. These results are summarized in Table 3.

Probability Updating Accuracy	
KNN, Aggregated Probabilities	91.8%
Random Forest, Aggregated Probabilities	88.3%

Table 3: Individual Identification Accuracy after Probability Aggregation on Data Samples

5 Conclusion

Biometric gait recognition is a powerful tool for identifying individuals. While our template based classification performed significantly worse than our statistical classification, the results of our probability aggregation show that purely template based classification can still be used rather effectively to identify individuals when multiple samples are available. We feel that this technique—a technique we have found nowhere in the literature—is a notable step towards effective gait-based identification and authentication systems.

Such probability aggregation would be especially useful when combined with our statistical approach to gait recognition. We find it very significant that our statistical model was able to identify individuals with over 99% accuracy using only 18 features generated with less than 1 hour of data on each individual. We attribute this accuracy to thorough exploration of the feature space, which enabled us to construct relevant data columns that could be effectively used by our random forest classifier. Given the classification accuracies that our statistical model achieved on a dataset of 42 individuals, we feel that investigating the plausibility of identifying individuals in larger populations (e.g. thousands of individuals) is a promising area of future research.

References

[1] Boulgouris, Nikolous; Hatzinakos, Dimitrios; Konstantinos, Plataniotis. *Gait Recognition: A challenging signal processing technology for biometric identification.* IEEE Signal Processing Magazine, 22(6):78-90.

[2] *Dictionary.com.* "Biometrics." http://www.dictionary.com/browse/biometrics

[3] *Find Biometrics: Global Identity Management.* "Applications." http://findbiometrics.com/applications/

[4] Boyd, Jeffrey; Little, James. *Biometric Gait Recognition.* University of Calgary. http://pages.cpsc.ucalgary.ca/ boyd/papers/biometric-summerschool.pdf

[5] Nickel, Claudia. *Accelerometer-based Biometric Gait Recognition for Smartphones.* Universität Darmstadt. http://tuprints.ulb.tu-darmstadt.de/3014/4/20120620_Dissertation_Nickel_final.pdf

[6] *Decision Tree Learning.* Wikipedia. https://en.wikipedia.org/wiki/Decision_tree_learning

[7] *Random forest classifier combined with feature selection for breast cancer diagnosis and prognostic.* Journal of Biomedical Science and Engineering, 6(5).

[8] Clark, Bryan. *The History of Biometric Security, and How It's Being Used Today*. Mar 04 2015. http://www.makeuseof.com/tag/the-history-of-biometric-security-and-how-its-being-used-today/.

[9] Mayhew, Stephen. *History of Biometrics*. Jan 11, 2015. http://www.biometricupdate.com/201501/history-of-biometrics

[10] Bayat, Akram; Pomplun, Marc; Tran, Duc. *A Study on Human Activity Recognition Using Accelerometer Data from Smartphones*. 11th International Conference on Mobile Systems and Pervasive Computing. http://www.cs.umb.edu/ duc/publications/papers/mobispc14.pdf

[11] Maninni, Andrea; Sabatini, Angelo. *Machine Learning Methods for Classifying Human Physical Activity from On-Body Accelerometers*. Sensors 2010, 10(2), 1154-1175.

[12] Gafurov, Davrondzhon. *A Survey of Biometric Gait Recognition: Approaches, Security and Challenges*. Gjvik University College.

[13] Seely, Richard; Goffredo, Michela; Carter, John; Nixon, Mark. "View Invariant Gait Recognition." *Handbook of Remote Biometrics*, edited by Massimo Tistarelli, Stan Z. Li, Rama Chellappa, 2009, pp. 61-81.

[14] Pedregosa *et al. Scikit-Learn: Machine Learning in Python.* JMLR 12, pp. 2825-2830, 2011.

[15] *Nearest Neighbors*, Scikit-Learn Documentation. http://scikit-learn.org/stable/modules/neighbors.html

[16] Kwapisz, Jennifer; Weiss, Gary; Moore, Samuel. *Cell phone based biometric identification.* 2010 Fourth IEEE International Conference on Biometrics: Theory, Applications and Systems. Sept 27, 2010. http://www.cis.fordham.edu/wisdm/includes/files/btas10.pdf