Accelerometer based Gait Recognition using Adapted Gaussian Mixture Models

Muhammad Muaaz
JRC u'smile & Johannes Kepler University Linz
Institute of Networks and Security
Altenbergerstr. 69, 4040 Linz, Austria
muaaz@ins.jku.at

René Mayrhofer
JRC u'smile & Johannes Kepler University Linz
Institute of Networks and Security
Altenbergerstr. 69, 4040 Linz, Austria
rene.mayrhofer@jku.at

ABSTRACT

Gait authentication using a cell phone based accelerometer sensor offers an unobtrusive, user-friendly, and a periodic way of authenticating individuals to their smartphones. In this paper, we present a GMM-UBM based gait recognition approach for a realistic scenario (when the phone is placed inside the trouser pocket and the user is walking) by using the magnitude data of a smartphone-based tri-axes accelerometer sensor. To evaluate our approach we use a gait data set of 35 participants collected at their respective normal walking pace in two different sessions with an average gap of 25 days between the sessions. We obtained EERs of 3.031%, 11.531%, and 14.393% for the same-day, mix-days, and cross-days, respectively.

CCS Concepts

•Human-centered computing → Smartphones; •Security and privacy → Biometrics;

Keywords

Accelerometer, gait recognition, Gaussian Mixture Models, segmentation, variance,

1. INTRODUCTION

Gait is defined as the way/style of walking, which is regulated by a complex biological process involving the brain, spinal cord, nervous, and musculo-skeleton systems. Humans walk bipedally, and any form of bipedal walking occurs as the result of repeated movement of each foot from one foothold to the next and the ground reaction forces experienced by the feet to balance the body weight while walking. Therefore, human gait is a cyclic process and the time interval between two consecutive occurrences of any of the periodic events of walking is called a gait cycle. Moreover, medical and psychological studies have shown that human gait patterns are unique. Over the past 30 years, researchers have explored gait as a biometric identifier to identify and verify individuals from their walking styles.

Currently, there exist three different approaches to gait recognition: video-based, floor sensors-based, and wearable sensors-based. The wearable sensors-based approach has drawn particular attention in this field of research with the development of (microelectromechanical systems) MEMS inertial sensors, especially their integration in smartphones. Since 2009, gait recognition using smartphone-based inertial sensors (accelerometers, gyroscopes) is being explored as one of the alternatives to traditional PIN/password based authentication systems. Previous studies on this subject have shown promising results [1, 6, 8, 10].

The gait recognition task can be seen as a typical pattern recognition problem and follows similar steps, such as preprocessing, segmentation, feature extraction, and classification. There exist two approaches to accelerometer based gait recognition on the basis of segmentation process: the cycle-based and frame-based [8]. In cycle-based approach, data is segmented from the start of every gait cycle and analysis is performed on these gait cycles [1, 3, 6]. However, the frame-base approach does not require gait cycle detection in the inertial data, but the data is segmented into either overlapping or non-overlapping segments (frames). Then features are extracted from these frames and machine learning algorithms are applied [8] for the genuine and impostor classification.

Over the past several years, adapted Gaussian Mixture Model-Universal Background Model (GMM-UBM) have become the state-of-the-art approach for speaker recognition systems. In this paper, we demonstrate the use of adapted GMM-UBM for cycle-based gait recognition. The outline of this paper is the following. In section 2, we briefly explain our data collection process. Data description and its processing steps are given in section 3. The GMM-UBM based classification framework is described in section 5 and finally, the experiment, results, and discussion are given in section 5.1.

2. DATA COLLECTION

In this paper, we use the same gait data set used in [6]. The data set comprises of gait data collected from 35 participants including 6 females and 29 males, using a Google Nexus Android phone. For data collection purpose, an Android application was developed which records three-dimensional (X, Y, and Z axis) accelerometer data at a sampling rate of 100 Hz and writes it to a text file with time stamps. Participants were instructed to wear trousers with not-too-

MoMM '16, November 28-30, 2016, Singapore, Singapore
© 2016 ACM. ISBN 978-1-4503-4806-5/16/11 . . . $15.00
DOI: http://dx.doi.org/10.1145/3007120.3007164

loose front pockets. For capturing a distinctive walking style, the phone or sensor must be placed close to the body otherwise it might pick up too much random noise. In the data recording phase, the phone was placed inside the trousers right side pocket. Participants were asked to walk at their normal pace in a 68 meters long straight corridor (with no stairs). They were told to wait for 1 second at the end of the walk then turn around and wait for another second before starting their new walk. In one session, every subject walked $4 \times 68 = 272$ meters or in other words completed two rounds of the corridor. For every subject, data recording was conducted in two different sessions. An average gap between the sessions is about 25 days. Eight walks were recorded for every subject in two different sessions.

3. DATA PROCESSING

Figure 1 shows various activities performed in a data recording session. Approximately, the first 10-20 seconds of data is when the phone was being placed inside the pocket, and next 100 seconds are when the person is standing still and listening to the instructions. Then the participant starts walking and reaches the end point. This walking activity lasts around 50 seconds and varies from person to person as it highly depends on the walking pace of the participant. At the end of the walk participant waits for a second, turns around and waits for another second before the new walk, and so on a participant completes the session with four walks. Data processing begins by separating session-wise recorded walks and computing magnitude from the tri-axes accelerometer data.

3.1 Active Walk Segment Detection

Accelerometers are very sensitive to noise; even when a smartphone is in a steady state, acceleration measured along any axis is not stable over the time. For instance, when the smartphone is in the stable state, the squared sum of acceleration values of all three axes should be equal to the earth gravitational force ($9.81\frac{m}{s^2}$), but in practice this is not the case. Therefore, in the first step, recorded walk data is mean normalized. First, a mean acceleration value is computed for every axis from the acceleration data along that axis. Then from the acceleration data along every axis, its respective mean value μ is subtracted Afterwards, the process of extracting active walk segments begins. We define active segments as those sections of the recorded data when a user was walking. This is done by monitoring the variance of the acceleration magnitude of the tri-axes accelerometer in a sliding window, as stated in [5]. For our evaluation and implementation, we use a sliding window of two seconds. When the variance of acceleration magnitude rises above or drops below a variance-threshold ($0.8\frac{m}{s^2}$) this marks the start and the end of an active walk segment as shown in Fig 1. A grid search was performed on the recorded gait data to find an appropriate value of variance threshold.

3.2 Interpolation

The accelerometer sensor on Android phones does not output equidistant data. It only outputs data when Android API's onSensorChanged method is triggered. Therefore, the time interval between two consecutive sensor values is not equal. By applying interpolation, data can be reshaped in equal intervals of time and can also be up-sampled in order to avoid data loss of too many values, for this purpose, we have used linear interpolation.

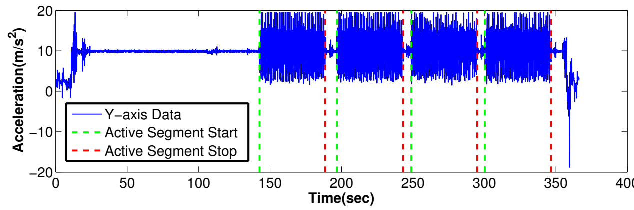

Figure 1: Detecting active walk segments.

3.3 Noise Removal

A Savitzky-Golay smoothing filter (also called the digital smoothing polynomial filter or least-square smoothing filter) is used to filter noise from the data. We preferred a S-G filter over the typical average moving filters because least-square smoothing not only reduces noise but also maintains the shape and height of waveform peaks. The basic idea behind S-G filter is to find a least-square fit with a polynomial of high degree for each data point, over an odd sized window centered around that data point.

4. SEGMENTATION

Human gait exhibits a cyclic pattern, and therefore measured acceleration is also periodic. The first step in the gait template generation process is to estimate the gait cycle length.

4.1 Cycle Length Estimation

We begin by extracting a small subset of samples (for reference, see Fig. 2) from the center of the walk as it is the most stable section of the walk because few cycles in the beginning and ending of the walk may not adequately present the person's gait [1,6]. Then we compare this reference subset with the other subsets (of similar length) extracted from the same walk by moving one sample forward, towards the end of the walk as shown in Fig. 2. Selecting too few samples

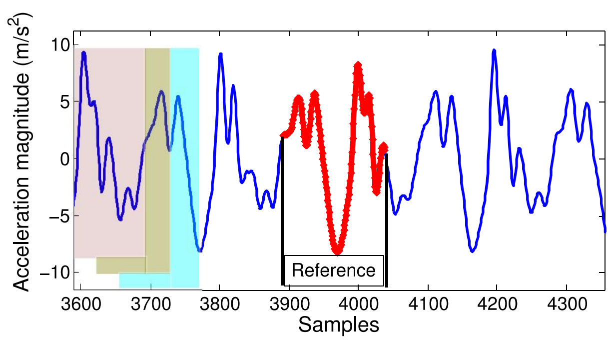

Figure 2: Estimating the gait cycle length.

for the reference subset will not reflect periodicity in the walk. Similarly, selecting too many samples for the reference subset will reduce the number of comparisons. From our experiments, we found that a reference window size equal to the sampling frequency not only reflects periodicity in the data, but also results in enough comparisons to estimate the gait cycle length. Comparing the reference subset with subsets extracted from the walk results in a distance vector as shown in Fig. 3. From this vector we find the indices of minima and store them in a minimum index vector. Later, we compute a difference vector which contains the difference of every two adjacent elements of the minimum index vector. Finally, the cycle length is computed by taking the mode of the difference vector. In those cases where mode does not exist (which means every step has different length e.g. if an individual is intentionally changing the walking pace) cycle length is computed by averaging the values of the difference vector.

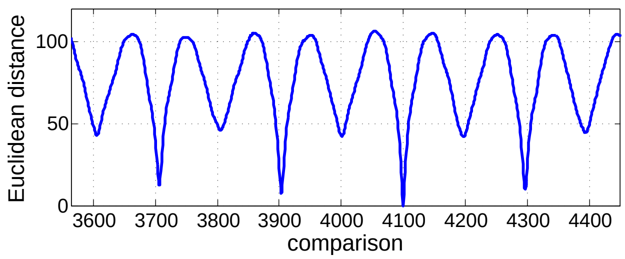

Figure 3: Estimating gait cycle length from detected minima.

4.2 Cycle Detection

Gait cycle detection starts by extracting a small segment (two times the estimated cycle length) from the center of the walk. Then we detect local minima in this extracted section of the walk. Sometimes interpolation errors could affect this area of the walk and we might detect a wrong minimum. To reduce this risk we use a segment size of double the cycle length. By doing so, we aim to pick two minimum values and we start cycle detection from the index of the most prominent minimum. From the index of this minimum point, cycle detection is done in a forward and backward direction by adding and subtracting the cycle length. From our experiments, we found that minima in the walk usually do not occur at equal intervals of gait cycle length. Therefore, we add and subtract a small offset value (0.2 times the estimated cycle length) to the index of the newly found end point and search for a local minimum in that region as shown in Fig. 4. Once all minima in both directions are found, the walk is segmented from the indices of these minima. Then all detected gait cycles are normalized to an equal length of 100

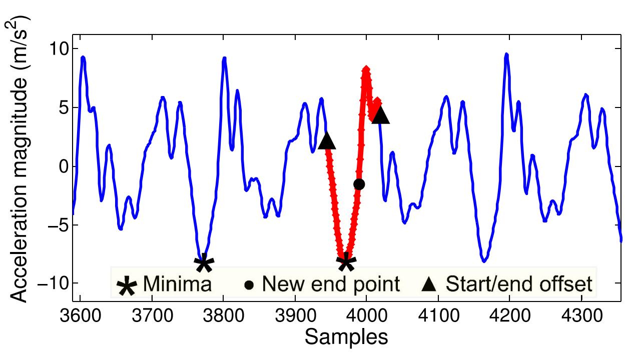

Figure 4: Estimating the start of a gait cycle.

samples because the similarity measures such as Euclidean distance only work on equal length data series.

4.3 Omitting Unusual Cycles

Detected cycles are cleaned by deleting unusual cycles. These outliers may occur e.g. if a person has stumbled while walking, stopped for moment to open the door, or quickly turned. To remove outliers, pairwise distances are computed between all detected cycles using DTW. This results in a matrix $D_{n \times n}$, where n is the number of detected cycles. If $A = (a_1, a_2, a_3, ..., a_n)$ and $B = (b_1, b_2, b_3, ..., b_n)$ are two gait cycles then $DTW(A, A) = 0$ and $DTW(A, B) = DTW(B, A)$. This means that we only need to compute lower or upper triangular elements of the matrix D. Finally, those cycles which have 50% of their pairwise distances greater than a threshold (0.6) are removed. Remaining cycles as shown in Fig. 5 represent the gait template of an individual. These gait

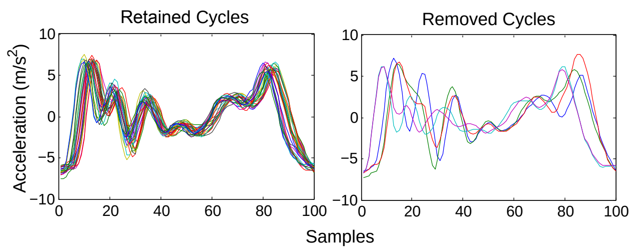

Figure 5: Example of an extracted gait template.

cycles can be used for template based classification as used in [6], but in this study, we extract following 88 features from every gait cycle including 80 Discrete Cosine Transformation (DCT) coefficients and spectral (roll-off, kurtosis, centroid, skewness, slope, decrease, flatness, and spread).

5. CLASSIFICATION

We use a GMM-UBM verification framework, a state-of-the-art approach for speaker verification [9]. A GMM-UBM approach can be divided into three parts: off-line training of a subject independent UBM via the Expectation Maximization (EM) algorithm, subject specific gait model generation, and likelihood ratio-based subject verification. A UBM is a subject independent large universal background GMM trained from gait data pooled from lots of subjects, intended to represent how humans walk in general. This UBM-GMM is a multivariate Gaussian distribution which is parameterized as $\lambda_{ubm}(w_i, \mu_i, \Sigma_i)$, where w_i is mixture weights (which satisfy the constraint $\sum_{i=1}^{M} w_i = 1$), μ_i and Σ_i are the mean and covariance matrices of each mixture or each Gaussian component. It is empirically observed that a diagonal covariance matrix out-performs a full covariance matrix [9], therefore, in our approach we also use diagonal covariance matrices. Further, to avoid over-fitting the training data, a variance limiting threshold is applied. Given the collection of training vectors, maximum likelihood model parameters are estimated by using the EM algorithm. The EM algorithm iteratively refines the GMM parameters to monotonically increase the likelihood of the estimated model for the observed features. The subject specific gait model is also a GMM model but instead of using iterative (EM) algorithm it is derived from the universal background model by using the Maximum-a-Posteriori (MAP) adaptation. This provides a tighter coupling between the UBM and subject models. Further, it is less computationally intensive compared to the EM algorithm. Given a background model and training vectors from a user, we first determine the posterior probability of each mixture as given in equation 1 and then compute sufficient statistics for the weight, mean and variance as given in equations 2, 3, and 4.

$$p(i \mid x_t) = \frac{w_i p_i(x_t)}{\sum_{j=1}^{M} w_j p_j(x_t)}; \tag{1}$$

$$n_i = \sum_{t=1}^{T} P(i \mid x_t); \tag{2}$$

$$E_i(x) = \frac{1}{n_i} \sum_{t=1}^{T} P(i \mid x_t) x_t; \tag{3}$$

$$E_i(x^2) = \frac{1}{n_i} \sum_{t=1}^{T} P(i \mid x_t) x_t^2. \tag{4}$$

Finally, the parameters of the adapted subject specific model (w_i, μ_i, Σ_i) are updated as given in equations 5, 6, and 7.

$$w_i' = [\frac{\alpha_i n_i}{T} + (1 - \alpha_i) w_i]\gamma; \tag{5}$$

$$\mu_i' = \alpha_i E_i(x) + (1 - \alpha_i)\mu_i; \tag{6}$$

$$\sigma_i^{2'} = \alpha_i E_i(x^2) + (1 - \alpha_i)(\sigma_i^2 + \mu_i^2) - \mu_i^2. \tag{7}$$

Here, γ is a scaling factor, used to ensure that the sum of adapted mixture weights equates to unity and $\alpha_i = \frac{n_i}{n_i + r}$, where r is a relevance factor. Finally, features are extracted from testing data and evaluated against the UBM and subject specific GMMs and the decision is made by comparing the log likelihood of the subject's adapted gait model and the UBM model as given in equation 8. For the genuine users the log likelihood from their adapted models will be higher than the UBM and for impostors the log likelihood of UBM model will higher.

$$\Lambda = log(p(X \mid \lambda_{user})) - log(p(X \mid \lambda_{ubm})) \tag{8}$$

5.1 Experiment, Results, and Discussion

As mentioned in section 2, in total we record eight walks (four walks in both sessions) for every user. Features extracted from the gait cycles of the first two walks of every user are used for training the UBM-GMM model ($mixtures = 256$, $r = 20$, $variance\ limiting\ threshold = 0.005$). Features from the gait cycles of next 1 walk of every user is used for adapting the user specific GMM models. Features from the remaining 1 walk and the 4 walks of the 2nd session are used as testing features for same-day and cross-day performance of the system. Results of this study in comparison to other studies are given in table 1.

If we compare the results given in table 1 with previous studies, we notice that this approach outperforms other approaches. However, results also indicate a reasonable difference in same-day and cross-day performance, that supports the argument that gait varies over the period of time, therefore, in our future work we are looking forward to search and extract features which are more session invariant and introduce on-line learning methods to cope with gait aging factor.

Table 1: Comparison of results with other studies, s stands for same, c for cross and m for mixed session.

Study	Placement	Subjects	Settings	Best EER
[2]	trouser pocket	25	s	100% CCR
[4]	trouser pocket	5	s	100% CCR
[1]	waist	48	m	20.1
[7]	waist	48	s	16.26
[7]	waist	48	c	29.39
[6]	trouser pocket	35	s	7.051
[6]	trouser pocket	35	c	18.965
this study	trouser pocket	35	s	3.031
this study	trouser pocket	35	m	11.531
this study	trouser pocket	35	c	14.393

6. ACKNOWLEDGMENTS

This work has been carried out within the scope of *u'smile*, the Josef Ressel Center for User-Friendly Secure Mobile Environments. We gratefully acknowledge funding and support by the Christian Doppler Gesellschaft, A1 Telekom Austria AG, Drei-Banken-EDV GmbH, LG Nexera Business Solutions AG, and NXP Semiconductors Austria GmbH.

7. REFERENCES

[1] M. O. Derawi. *Smartphones and Biometrics: Gait and Activity Recognition*. PhD thesis, Gjøvik University College, November 2012.

[2] J. Frank, S. Mannor, and D. Precup. Activity and gait recognition with time-delay embeddings. In *AAAI*, 2010.

[3] D. Gafurov. *Performance and Security Analysis of Gait-based User Authentication*. PhD thesis, Universitas Osloensis, 2004.

[4] 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. IEEE, 2010.

[5] R. Mayrhofer and H. Gellersen. Shake well before use: Intuitive and secure pairing of mobile devices. *IEEE Transactions on Mobile Computing*, 8(6):792–806, 2009.

[6] M. Muaaz and R. Mayrhofer. Orientation independent cell phone based gait authentication. In *Proceedings of the 12th International Conference on Advances in Mobile Computing and Multimedia*, MoMM '14, pages 161–164, New York, NY, USA, 2014. ACM.

[7] M. Muaaz and C. Nickel. Influence of different walking speeds and surfaces on accelerometer-based biometric gait recognition. In *Telecommunications and Signal Processing (TSP), 2012 35th International Conference on*, pages 508–512, 2012.

[8] C. Nickel. *Accelerometer-based Biometric Gait Recognition for Authentication on Smartphones*. PhD thesis, TU Darmstadt, June 2012.

[9] D. A. Reynolds, T. F. Quatieri, and R. B. Dunn. Speaker verification using adapted gaussian mixture models. *Digital signal processing*, 10(1):19–41, 2000.

[10] M. Tamviruzzaman, S. I. Ahamed, C. S. Hasan, and C. O'brien. epet: When cellular phone learns to recognize its owner. In *Proceedings of the 2Nd ACM Workshop on Assurable and Usable Security Configuration*, SafeConfig, pages 13–18. ACM, 2009.