Biometric Walk Recognizer

Maria De Marsico[✉] and Alessio Mecca

Sapienza University of Rome, Via Salaria 113, 00198 Rome, Italy
demarsico@di.uniroma1.it, alessio.mecca@hotmail.com

Abstract. In this paper we present a comparative test of different approaches to gait recognition by smartphone accelerometer. Our work provides a twofold contribution. The first one is related to the use of low-cost, built-in sensors that nowadays equip most mobile devices. The second one is related to the use of our system in identification mode. Instead of being used to just verify the identity of the device owner, it can also be used for identification among a set of enrolled subjects. Whether the identification is carried out remotely or even if its results are transmitted to a server, the system can also be exploited in a multibiometric setting. Its results can be fused with those from computer-vision based gait recognition, as well as other biometric modalities, to enforce identification for accessing critical locations/services. We obtained the best results by matching complete walk captures (Recognition Rate 0.95), but the implicit limitation is represented by the fixed number of steps in the walks. Therefore we also investigated methods based on first dividing the signal into steps. The best of these achieved a Recognition Rate of 0.88.

Keywords: Biometrics · Gait recognition · Accelerometer · Mobile devices

1 Introduction

Biometric recognition is attracting increasing interest, since in many environments it can provide efficient means of identity recognition. Some biometric traits can be considered as *strong*, since they allow robust authentication especially in controlled settings. Examples are face, iris, and fingerprints. Strong biometrics often concern physical traits. Other traits, mostly *behavioral* ones, can be less reliable, for example because they can be affected by emotional conditions. At present, gait recognition, i.e. recognizing people from the way they walk, is often classified in the latter category. It can be used to complement other strong or soft traits in a multibiometric approach, and is one of the recent attractive topics in biometric research. Gait recognition modalities can be categorized into three groups based on the technological setting they require: a) machine vision-based: these approaches suffer from typical image processing issues, e.g., occlusions and illumination variations; b) floor sensor-based: these approaches require specific ambient equipment, therefore cannot be used everywhere and require a preliminary complex set-up; c) wearable sensor-based: these approaches can be further distinguished according to the number and kind of sensors exploited (more accelerometers, accelerometers plus gyroscopes, etc.).

Given the limitations of the first two groups of techniques in terms of feasibility and ready availability, we will focus on the third group. Even in this case, we aim at exploiting the simplest possible set-up. It is easy to assume that such set-up should use off-the-shelf, widespread equipment, and possibly exploit only one sensor of one kind, to keep the method feasible on a wide range of devices. Moreover, it would be desirable to exploit the least computationally demanding procedure.

We present the results obtained using the accelerometer built in the OnePlus One mobile phone, and compare different matching algorithms to evaluate the better compromise between accuracy and computational cost. The latter would allow performing recognition both locally, e.g., on a smartphone, or remotely on a computer receiving gait data. Moreover, we will test these techniques in identification modality (no identity is claimed by the user). This also allows combination with other biometrics to secure the access not only to personal devices but also to locations or services. On the contrary, most approaches in literature use (implicit) verification: the identity claim is implicit in the preliminary enrollment of a single user, namely the owner of the device, and the algorithms aim at verifying if the user keeping the device is the owner. In identification, assuming open set mode (probe user may not be known to the system), a template distance threshold regulates acceptance as in verification; in addition the right subject has not only to be close enough, but also to be the closest one.

2   State of the Art for Accelerometer-Based Gait Recognition

During the last years, among the approaches mentioned above, the gait recognition based on wearable sensors is becoming increasingly popular in respect to the others. This is because, on one hand, the related techniques used for recognition are less computationally expensive, and wearable sensors have a cost much lower than cameras and floor sensors; on the other hand achieved results are similar or even better. Wearable sensors do not suffer from the problems normally raised by computer vision-related techniques, e.g., occlusion, and illumination. In addition, differently from approaches based on floor sensors, the subject can be followed in any place without the need to equipping the environment. Last but not least, nowadays the most common wearable sensor used for gait recognition, the accelerometer, is practically built in most portable devices such as mobile phones and tablets. This avoids the use of expensive ad-hoc equipment. The reader can find in [1] a comprehensive survey on present approaches to gait recognition divided into the three mentioned main groups; [2] provides a review of present vision-based technique, the most used so far.

In most papers about gait recognition by accelerometer, the aim is to identify the owner of the phone. The recognition modality is verification: the identity claim is implicit, since only the owner is enrolled. Performances are usually measured by computing all-against-all matching results in a set of subjects. For verification, the performance measures are False Acceptance Rate (FAR), False Reject Rate (FRR) and Equal Error Rate (EER) given by the acceptance threshold were FAR=FRR.

In [3] the authors use the low sampling rate accelerometer that is built-in in Google G1 phone. Their system achieves an EER of 20% with a dataset of 51 volunteers. The algorithm uses cycle detection, computes average cycle and uses Dynamic Time

Warping for matching. The EER is not that low, however this is also due to the very low quality of the used accelerometer. As a matter of fact, in [4] a pair of the same authors improve their previous results using a better accelerometer, Motion Recording 100, which is not built in a mobile personal device. They perform some preprocessing operations, such us outlier steps removal, before matching performed by Dynamic Time Warping (DTW). The dataset size is increased to 60 users and the achieved EER is 5.7%. In comparison with this approach, we compute the average cycle to discard outliers, but we do not use it for matching. In addition, we also test our system in identification mode, which in some cases might be more challenging. While verification just requires sufficient similarity with the claimed identity, in identification one also has to avoid confusion between candidate identities, i.e., the right subject must not only be close enough, but also be the closest one. For both verification and identification the degree of similarity must be very high to obtain a correct answer of the system. However, consider a person that in identification modality is close enough but is not the closest one to the right one, and hence is not erroneously identified, e.g., as the owner of the device. This person, in verification modality, could be erroneously interpreted as the owner of the device if the similarity is just sufficiently good.

The recognition technique presented in [5] uses an accelerometer which is a standard in nowadays mobile devices, with sampling rate of 100 samples/second. They achieve a 5% of EER using histogram similarity and a 9% of EER using cycle group matching. However, though using the same accelerometer as in the preceding work, they use a smaller dataset (21 users). They do not report results for identification.

The work in [6] presents a framework for gait recognition using an Android mobile phone to get the accelerometer and gyroscope raw data. Verification is carried out on a desktop computer to recognize if the person who is walking with the phone is the enrolled user. The framework uses continuous wavelet transform time frequency spectrogram analysis for feature extraction and ciclostationarity analysis for matching. The tests achieve a 99.4% of verification rate at 0.1% of FAR for pace walk VS pace walk, 96.8% for fast walk VS fast walk, and a 61.1% for pace walk VS fast walk on a dataset of 36 users. This work achieves good results but uses the gyroscope too.

In [7] the authors use Hidden Markov Model (HMM). They achieve a 10.42% of False Non Match Rate and a 10.29% of False Match Rate on a dataset of 48 users. The system uses only one enrolled user for training the HMM. All the registrations for all users are used as probe to test performance of verification. The performances achieved by this method are not competitive with the above ones.

The work in [8] exploits a signature point based method for identification. The user is recognized among a set of enrolled subjects. The system exploits five accelerometers attached in various parts of the body. This method does not use the cycle division as the others discussed above. It achieves a high recognition rate (96.7%) with a dataset of 30 users. In [9] an evolution of the previous method is proposed, and is used for both identification and verification modalities. It uses clustering to preliminarily group the signature points. The authors have also created a huge dataset (with 175 different subjects) of raw acceleration data, which is publicly available for

non-commercial use[1]. On this dataset they achieve a Recognition Rate (RR) of 95.7% for identification and a 2.2% of EER for verification. This method achieves the best results among the mentioned ones, but uses a very complex equipment.

3   The Approaches We Tested

We now describe the different approaches we tested, starting from the best performing yet more constrained one, and releasing constraints afterwards. As for most papers in Section 2, the matching algorithm exploits DTW (in its basic version).

In order to capture the gait information, we chose 3-axis accelerometer as other related works. This sensor seems to better monitor the exact movement of a subject in the space, and nowadays is widespread and simple to set up, being built in most normal Android phones. The components exploited by our system are a OnePlus smartphone equipped with a 3-axis accelerometer and a desktop computer with Java Virtual Machine. We have not addressed interoperability issues yet, therefore the devices for enrollment and testing need to be of the same model.

Among the apps for monitoring and recording the accelerometer data, we chose Physics Toolbox Accelerometer, available on Google Play[2]. This application has a number of advantages among which the instant saving of data in Comma Separated Values (CSV) format. Moreover, the app is light, very fast, and compatible with any device with an accelerometer, including less recent versions of Android or less performing devices. Android captures accelerometer information automatically in an intelligent way: once the designer has set a minimum threshold of time to acquire the information, the data capture along the three axes is done according to a sampling density that depends on the displacement of the device relative to the previous position, i.e., the lower the variation the sparser the sampling and vice versa. Many approaches include time interpolation to obtain fixed intervals, but we avoid this step.

3.1   Enrollment Phase

In enrollment phase, our goal is to obtain a uniform enough data capture, which is quite usual in most applications. Therefore we decided to standardize the movements for each recording according to the following procedure: 1) the subject is asked to put the phone in the belt, either on the right hip or on the left one, in vertical position with the screen facing out (because recording is ended by tapping on the screen), in about the same position for all users; 2) the subject is then asked to keep feet together and start walking by the leg opposite to phone location; 3) the system records 10 steps along a straight line in the most natural way. For each enrolled user a folder is created on the desktop computer, where data is transferred (manually by now), and results of possible processing are stored. At the moment, we acquire data for three different user walks that during tests are used from time to time as probe or as gallery (during identification, probe data

---

[1] http://www.ytzhang.net/datasets/zju-gaitacc
[2] https://play.google.com/store/apps/details?id=com.chrystianviey
ra.android.physicstoolboxaccelerometer&hl=it

is matched with data from two different instances of walk). However, all approaches can handle a different number of templates per enrolled user. At present the dataset has 26 users, wearing different shoes (but no high heels).

In order to have each walk signal starting from a relevant point, we discard all initial local maxima created by noise. To do this, we use a threshold experimentally set at 1.05 on y axis (this value is device dependent). After the procedure finds the first value greater than the threshold, the system searches for the first relative maximum from that point. Once this has been identified, it is saved and will be the starting point of the vector that will be used by matching. A similar procedure is used to identify the useful end of the signal. As for now, the same threshold is used.

The system uses the DTW separately for each axis. In order to investigate an effective fusion of the three results for each comparison into a final similarity, we preliminarily studied the plots returned and performed some tests. In most cases, and considering that axes depend from the device position, our y-axis is the most important for recognition, followed by z-axis. The resulting weights are very similar to those used in other works in literature. The x-axis of our device has very little impact on the recognition. All walks achieve a very similar value on that axis, mainly because it assumes specific meanings only in case of strong jerks or jumps, that are generally absent in natural walking. We perform a linear combination of the three values using experimentally determined weights that depend on the processing method (see below).

3.2   Testing Phase and Experimental Results

As for testing, we investigated different possibilities, that we will denote according to the way walk data is pre-processed and to the way matching is performed: Walk, Best Step and Best Step VS All, AllSteps VS All, and Steps Sliding Window. The pre-processing described above applies to all. We apply essentially two strategies. The first one exploits the overall signal for matching, and is implemented only in the Walk method. All the others rely on a preliminary segmentation of the walk signal into steps, and on the possible discarding of outliers. This further pre-processing phase is applied to both gallery templates and to the probe. We first identify the starting and ending points of the useful portion of the signal using the thresholds mentioned above. Afterwards, for each walk belonging to the same user, maxima are extracted and ordered by amplitude. We underline that, thanks to the way we ask the user to walk, we can assume a certain distribution of maxima within the signal. In particular, each step is generally characterized by a first higher peak, corresponding to moving the leg with the accelerometer, and a second higher peak corresponding to moving the other leg. Moreover, it contains many further local maxima corresponding to noise. For each walk in the gallery, local maxima are extracted and ordered. In doing this, we consider the *#walks* sequences captured for a user all together. The value of the local maximum in position 10*#*walks* of this ordered list is taken as the threshold for step segmentation for the user at hand. Maxima higher than such threshold are taken as the starting points of new steps. In this way, we obtain on average 10 steps per walk for each user, unless some error happens. Moreover, we discard outliers. After dividing the walk into steps, we use DTW to compute pairwise differences and discard steps showing a distance from the others higher than the average distance plus

the variance. In our case, we have an average of 20 steps per registered user. For each user, the threshold identified for step segmentation is stored together with the whole walk as well as the identified steps, with and without the outliers. At present, we assume closed set identification (the user is in the database) but setting a threshold is sufficient to pass to open set identification (the user might be unknown).

Walk. In Walk method, it is assumed that the probe contains the same number of steps as the gallery templates. The probe walk is matched against all stored walks (at present, two for each enrolled user). The system returns the labels of the candidates ordered by decreasing distance of each walk from the probe (the two walks for the same user are not necessarily close in the returned list). We use the following weights to combine results from the axes: x-axis=0, y-axis=0.8, z-axis=0.2.

   Notwithstanding the simple approach, Walk modality achieves very interesting identification results. Notice that Recognition Rate (RR) is above 0.95 in the best test. However, the limitations given by the strictly controlled acquisition, especially the fixed number of steps, make this method not much appealing for real-world settings.

Best Step and Best Step VS All. These methods are a first attempt to overcome the limitations of Walk. They use step segmentation and outlier elimination, also for the probe. A further information is stored during the enrolling phase: the centroid of steps for each enrolled user is computed using DTW-based distances, and is saved in the gallery. It should represent the *average* user's step. During testing by Best Step, this *centroid step* of each user in the gallery is matched against the *centroid step* of the probe. For each matching operation, the probe is segmented into steps using the segmentation threshold stored in the gallery for the user to match. This algorithm achieves much lower performances with respect to Walk, with a RR of only 38.5%. It is reasonable to assume that this is due to the reduced size of the matched signal (a single step) which is the only information used to determine the final distance. We also tried to match the centroid step of each enrolled user to all steps of the probe, to take the minimum distance as result (Best Step VS All). However, this achieved the worst result of all algorithms we tried (a very poor RR of 23.1% in the best test).

AllSteps VS All. In AllSteps VS All method (AllSteps for short), we perform segmentation as described above and discard outliers, also for the probe. The identification phase is more articulated in this modality. For each gallery user, we divide the probe walk according to the user-specific threshold. For each step so obtained, we compute the minimum distance from the steps of that user. The final distance will be the average of all probe step minimum distances. The linear combination of results from the three axes is performed using the following weights: x-axis=0.1, y-axis=0.7, z-axis=0.2 (yet with very little performance decrease with the same Walk weights).

   The reverse of the medal for all methods allowing much more freedom in the probe walk, i.e., that avoid setting in advance any number of steps in testing operations, is a decrease in performance, which is lower for this one, i.e., a RR of above 0.88 is achieved in the best test that corresponds in any case to a very promising result.

Steps Sliding Window. This method tries to recover some of the accuracy achieved by Walk. It uses segmentation just to compute the number of steps in the signal to

match, and of course does not discard outliers, since matching is performed on se-
quential portions of the signals. We consider a window of x steps, where x is the min-
imum between the total number of steps of the probe and the total number of steps for
the enrolled user being matched. The method applies DTW by sliding the shortest
window over the longest one, and returning the minimum distance found. The weights
assigned to axes are: x-axis= 0, y-axis=0.6, z-axis=0.4. This is the computationally
most expensive among the tested algorithms and achieves a RR of 84.6% in the best
test. In fact, it is slightly worse than AllSteps and also more onerous.

Figure 1 summarizes the achieved results in terms of Cumulative Match Curves
(CMCs) in the best test. Since Walk returns ordered walks, while the others return
ordered identities, the plot for Walk spans more ranks. We also performed   a kind of
3-fold validation, with each gallery template playing in turn the role of probe. To
stress the Walk experimental setting, we did not repeat it three times with a balanced
number of templates between the identity corresponding to the probe and those of the
others: in each experiment we picked one template per user as probe (so that only two
were left in the gallery) and matched it against all three templates of the other users.
Results were only slightly worse, with similar relative performance of methods.

We tested the system for verification too using an all-against-all procedure. We as-
signed to each probe the identity of all enrolled users in turn (claim). Impostors were
obviously much more than genuine users (|probe|*|gallery-1| impostors and |probe|
genuines). The results confirm the relative performance for the different methods. We
achieved the following ERR values: 7.69% for Walk, 30.46% for Best Step, 30.92%
for Best Step VS All, 10.46% for AllSteps, and 15.38% for StepSlidingWindow. This
confirms that the use of the centroid step is the worst strategy for verification too.
Walk confirms to be the best method while the two methods with less restriction, i.e.,
AllSteps and StepSlidingWindow, achieve an acceptable result.

3.3   Discussion and Conclusions

This work presented a preliminary investigation on biometric identification by gait
recognition via smartphone accelerometer. We tested different methods, starting from
Walk that requires a fixed number of steps, but achieved the best results (RR above
0.95, EER 7.69%). The best compromise between low constraints and good accuracy
is AllSteps (RR above 0.88, EER 10.46%). Identification results are confirmed in a
kind of 3-fold validation. Methods achieving better results for verification used a
more complex equipment and/or more expensive computational methods.

The only method in related work achieving better results in identification modality
(RR 95.5% with no limitation on walk length) uses 5 accelerometers, therefore it is
not suited for use with personal mobile devices. Moreover, the reported results using
only the accelerometer located on pelvis, which is the same setting as our, only reach
an RR of 73.4%, yet with a larger gallery. We will test further modalities and future
work will address a number of issues. We will study how to make the method intero-
perable, by devising a kind of signal normalization procedure allowing to match sig-
nals from different accelerometers. We will also extend the database in less controlled
or adverse conditions, i.e., by shoes with heels, and at different walking pace.
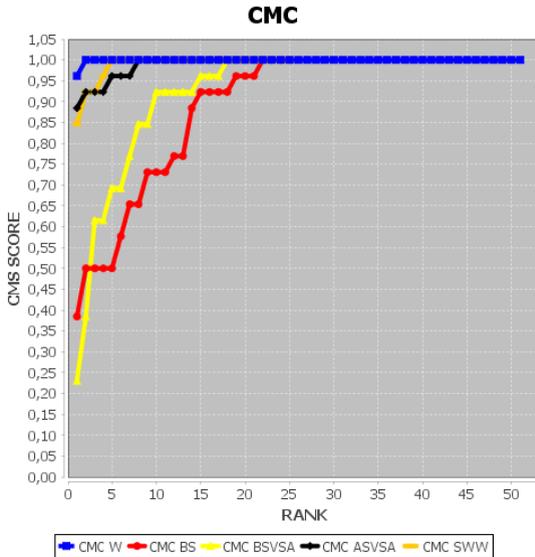
Fig. 1. CMCs for best tests in all modalities

References

1. Gafurov, D.: A survey of biometric gait recognition: approaches, security and challenges. In: Annual Norwegian Computer Science Conference, pp. 19–21 (2007)
2. Lee, T.K., Belkhatir, M., Sanei, S.: A comprehensive review of past and present vision-based techniques for gait recognition. MTAP **72**(3), 2833–2869 (2014)
3. Derawi, M.O., Nickel, C., Bours, P., Busch, C.: Unobtrusive user-authentication on mobile phones using biometric gait recognition. In: 2010 Sixth Int. Conf. on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), pp. 306–311 (2010)
4. Derawi, M.O., Bours, P., Holien, K.: Improved cycle detection for accelerometer based gait authentication. In: 2010 Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), pp. 312–317. IEEE (2010)
5. Gafurov, D., Helkala, K., Søndrol, T.: Biometric gait authentication using accelerometer sensor. Journal of Computers **1**(7), 51–59 (2006)
6. Juefei-Xu, F., Bhagavatula, C., Jaech, A., Prasad, U., Savvides, M.: Gait-ID on the move: pace independent human identification using cell phone accelerometer dynamics. In: IEEE BTAS 2012, pp. 8–15. IEEE (2012)
7. Nickel, C., Busch, C., Rangarajan, S., Mobius, M.: Using hidden markov models for accelerometer-based biometric gait recognition. In: 2011 IEEE 7th International Colloquium on Signal Processing and its Applications (CSPA), pp. 58–63. IEEE (2011)
8. Pan, G., Zhang, Y., Wu, Z.: Accelerometer-based gait recognition via voting by signature points. Electronics Letters **45**(22), 1116–1118 (2009)
9. Zhang, Y., Pan, G., Jia, K., Lu, M., Wang, Y., Wu, Z.: Accelerometer-based gait recognition by sparse representation of signature points with clusters. IEEE Transactions on Cybernetics, November 2014. doi:10.1109/TCYB.2014.2361287, http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6963443&tag=1