Advances in Visual Computing: 6th International Symposium, ISVC 2010, Las Vegas, NV, (USA) November 29-December 1, 2010. Proceedings. Lecture Notes in Computer Science, Volumen 6453. Springer, 2010. 489-498

DOI: http://dx.doi.org/ 10.1007/978-3-642-17289-2_47

Analysis of Time Domain Information for Footstep Recognition

R. Vera-Rodriguez[1,2], J.S.D. Mason[2], J. Fierrez[1] and J. Ortega-Garcia[1]

[1] Biometric Recognition Group - ATVS, Universidad Autónoma de Madrid,
Avda. Francisco Tomás y Valiente, 11 - 28049 Madrid, Spain
[2] Speech and Image Research Group, Swansea University,
Singleton Park SA1 8PP, Swansea, UK
{ruben.vera, julian.fierrez, javier.ortega}@uam.es, j.s.d.mason@swansea.ac.uk

Abstract. This paper reports an experimental analysis of footsteps as a biometric. The focus here is on information extracted from the time domain of signals collected from an array of piezoelectric sensors. Results are related to the largest footstep database collected to date, with almost 20,000 valid footstep signals and more than 120 persons, which is well beyond previous related databases. Three feature approaches have been extracted, the popular ground reaction force (GRF), the spatial average and the upper and lower contours of the pressure signals. Experimental work is based on a verification mode with a holistic approach based on PCA and SVM, achieving results in the range of 5 to 15% EER depending on the experimental conditions of quantity of data used in the reference models.

1 Introduction

Footstep signals have been used in different applications including medicine [1], surveillance [2], smart homes [3] and multimedia [4]. Footstep recognition was first suggested as a biometric in 1977 [1], but it was not until 1997 when the first experiments were reported [5]. Since then the subject has received relatively little attention in the literature compared to other biometrics, even though it possesses some worthwhile benefits: unobtrusive, unconstrained, robust, convenient for users, etc.

Different techniques have been developed using different sensors, features and classifiers as described in [6]. The identification rates achieved of around 80-90% are promising and give an idea of the potential of footsteps as a biometric [7,8]. However, these results are related to relatively small databases in terms of number of persons and footstep signals, typically around 15 people and perhaps 20 footsteps per person [5]; this is a limitation of the work to date.

A database is an essential tool to assess any biometric; therefore, this paper reports experimental results of footsteps as a biometric on the largest footstep database to date, with more than 120 people and almost 20,000 signals, enabling assessment with statistical significance.

Regarding the sensors employed to capture the footstep signals, two main approaches have been followed in the literature: switch sensors [9,10,11] have been used with a relatively high sensor density (ranging from 50 to 1024 sensors per m^2) in order to detect the shape and position of the foot. On the other hand, different types of sensors that capture transient pressure [5,12,13,14,15] have been used with relatively low sensor density (typically 9 sensors per m^2), more focused in the transient information of the signals along the time course.

The capture system considered here uses a high density of approximately 650 piezoelectric sensors per m^2 which gives a good spatial information and measures transient pressure.

This paper is focused on the analysis of the temporal information of the footstep signals. In this sense the most popular feature extracted in the related works is the ground reaction force (GRF), in some cases used in a holistic manner [5], and in other cases geometric measurements are extracted from the GRF [12,15,16]. In our previous works [8,17,18] geometric and holistic features were compared obtaining in all cases better results for the holistic approach. In this paper, the GRF profiles are compared with other features in a holistic manner and also a fusion of them is carried out obtaining verification results in the range of 5 to 15% of EER depending on the experimental conditions. The experimental protocol is focused on the study of the influence of the quantity of data used in the reference models.

The paper is organized as follows. Section 2 describes the footstep signals used and the feature extraction process, focused on time information. Section 3 describes briefly the database, Section 4 presents the experimental results; and finally conclusions and future work are presented in Section 5.

2 Description of the Signals and Feature Extraction

2.1 Footstep Signals

As mentioned above, the capture system considered here uses piezoelectric sensors with a relatively high density, and therefore footstep signals collected contain information in both time and spatial domains. This is in contrast to previous related works, e.g. [12,10,11]. In fact, footstep signals collected here contain information in four dimensions namely: pressure, time, and spatial positions X and Y. The sensors are mounted on a large printed circuit board and placed under a conventional mat. There are two such mats positioned appropriately to capture a typical (left, right) stride footstep. Each mat contains 88 piezoelectric sensors in an area of 30×45 cm.

Figure 1 shows three different 3D plots for an example of a footstep signal reflecting its three stages: Figure 1(a) shows the differential pressure for an instant in the first stage of the footstep, i.e. when the heel strikes the sensor mat, Figure 1(b) shows the same but for an instant in the second stage of the footstep, i.e. when the whole foot rests over the sensors, and Figure 1(c) the same but for an instant in the third stage of the footstep, i.e. when the heel leaves the surface

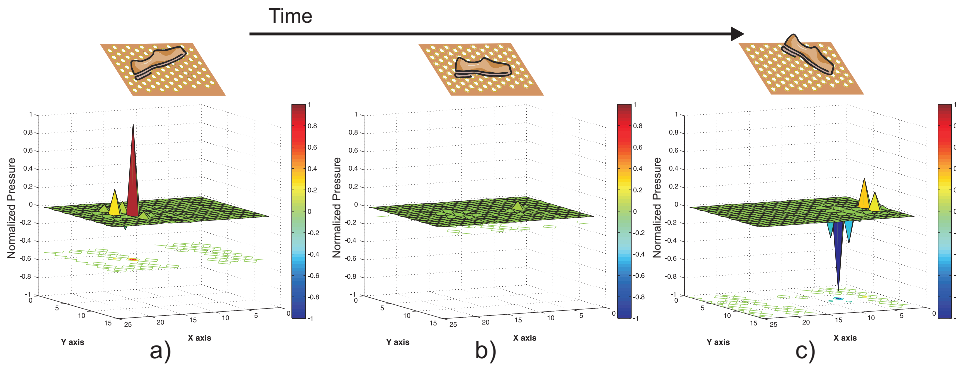


Fig. 1. Spatio-temporal footstep signal in the different stages. a) The derivative of the pressure against the position X and Y at the first stage of footstep. b) The same but for second stage of footstep signal. c) The same but for third stage of the footstep signal.

and the toes push off the sensor mat. It is worth noting that the output of the piezoelectric sensors is the differential pressure in time; thus, it can be seen in Figure 1(c) that there are negative values.

In this paper, the focus is on the analysis of the information of the footstep signals contained in the time domain, leaving the analysis of the spatial domain for further work.

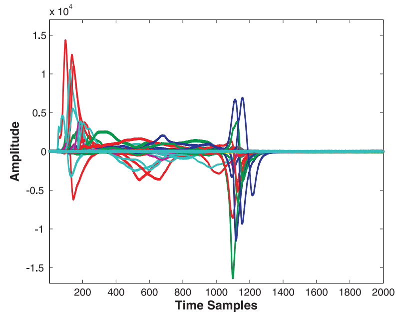
2.2 Feature Extraction and Matching

This section describes the time domain features that are used to assess footsteps as a biometric. Figure 2(a) shows an ensemble of signals from a single footstep. Each signal represents the differential pressure against time for each of the 88 sensors in one mat. An energy detector across the 88 sensors is used to obtain the beginning of each footstep to align the signals.

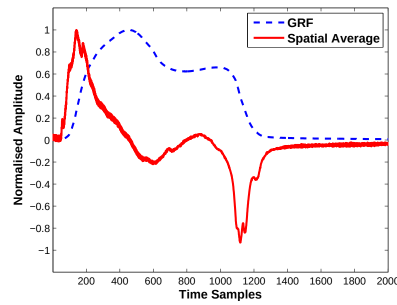
The most popular time domain feature in related works is the ground reaction force (GRF) [5,7,12,15,16]. Figure 2(b) shows the GRF profile for the example footstep considered here. In this case, as the piezoelectric sensors provide the differential pressure, the GRF is obtained by accumulating for each sensor signal across the time, and then an average of the 88 single profiles is computed to provide a global GRF. Formally, let $s_i[t]$ be the output of the piezoelectric sensors i, where t are the time samples being $t = 1,...,T_{max}$ and i are the sensors $i = 1,...,88$. Then the global GRF (GRF_T) is defined by:

$$GRF_T[t] = \frac{1}{88}\sum_{i=1}^{88}(\sum_{\tau=0}^{t}(s_i[\tau])) \quad (1)$$

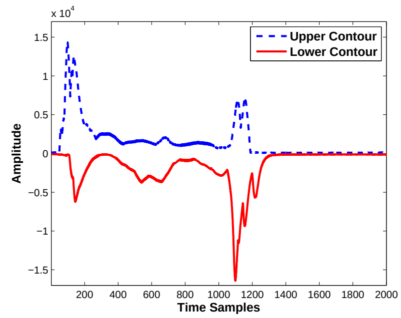
Apart from the GRF, two other feature approaches are studied here. The first comes from a spatial average [13,19] of the 88 sensors of the mat to produce a single profile. An example is shown in Figure 2(b).



(a) Footstep signal



(b) GRF/Spatial Average



(c) Upper/Lower contours

Fig. 2. Feature extraction in time domain for a footstep signal. (a) Differential pressure against time for the 88 sensors. (b) Normalised ground reaction force profile from (a) as defined in Equation 1, and normalised spatial average of the 88 sensors as defined in Equation 2. (c) Upper and lower contour profiles from (a) as defined in Equations 3 and 4 respectively.

$$s_{ave}[t] = \frac{1}{88}\sum_{i=1}^{88}(s_i[t]) \tag{2}$$

The second approach uses the upper and lower contour coming from the maxima and minima of the sensors for each time sample, as shown in Figure 2(c). These two signals are then concatenated into one contour signal.

$$s_{up}[t] = \max_{i=1}^{88}(s_i[t]) \tag{3}$$

$$s_{lo}[t] = \min_{i=1}^{88}(s_i[t]) \tag{4}$$

Equations 1 to 4 lead to a high dimensionality in the time domain with a vector of 8000 samples per footstep. Data dimensionality is further reduced using principal component analysis (PCA), retaining more than 96% of the original information by using the first 120 principal components for each feature approach. Regarding the classifier, a support vector machine (SVM) was adopted with a radial basis function (RBF) as the kernel, due to very good performance in previous studies in this area [7,8].

3 Database and Experimental Protocol

The database collected, apart from footsteps, contains another three biometric modes: speech, face and gait. These modes were included in order to assist in the labelling of the footstep signals, as the collection was an unsupervised process. The speech mode was used to carry out an automatic labelling of the database. A novel iterative process was developed using an identification strategy, labelling the data with the highest confidence first and leaving the data with less confidence for the last iterations [20].

Regarding the experimental protocol followed to assess footsteps as a biometric, special attention has been paid to the partitioning of the data into three sets, namely Reference data and two test sets. The first test set, called Development was used to set the parameters of the system such as the features, the PCA components and the SVM classifier. Then the unseen Evaluation test set is comprised of the last 5 signals collected from each person. It is worth noting that in this paper the data used in the different sets keeps the chronological time of the collection. Therefore, for each user the reference data is comprised of the first data provided, and the data used in the Evaluation set is the last collected. This is a realistic approach reflecting actual usage in contrast to previous related works [8,9,10].

The influence of the quantity of data used to train and test the system is a key factor in any performance assessment; while common in more established biometric modes this aspect is not considered in many cases of footstep studies, for example in [5,12,13], due to limited numbers of data per person in the databases. Different applications can be simulated using different quantities of data in the reference models. In the present work we simulate two important applications: smart homes and access control scenarios. In the case of a smart home there would be potentially a very large quantity of reference data available for a small number of persons, while in security access scenarios such as a border control, limited reference data would be available, but potentially for a very large group of people.

A characteristic of the database considered here is that it contains a large amount of data for a small subset of people (>200 signals for 15 people) and a smaller quantity of data for a larger group of people (>10 signals for 60 people). This reflects the mode of capture which was voluntary and without reward. The assessment of the system is carried out in several points or benchmarks considering different amounts of reference data.

For example, Table 1 shows the quantity of data used in benchmark B1 (using 40 signals in the reference models and 40 models) for the different data sets of Development and Evaluation. Each signal from the test sets is matched

Benchmark B1	Reference Data	Test	
		Development	Evaluation
Clients	P1 – P40	P1 – P40	P1 – P40
Footsteps per client	40	170 (8-650)	5
Total for clients	1,600	6,697	200
Out of class users	P41 - P127	P41 - P78	P41 - P110
Total out of class	763	380	350
Total set	2,363	7,077	550
Total	9,990		

Table 1. Database configuration for benchmark B1 (40 signals per reference model).

against all the 40 reference models defined. As can be seen in the table, the total number of stride signals in the database is 9,990, i.e. 19,980 single (right and left) signals in total. As a result, the number of genuine matchings is 6,697 and 200 for Development and Evaluation respectively; and the number of impostor matchings are 276,383 and 21,800 for Development and Evaluation respectively. Similarly, other benchmark points have been defined with different number of models and signals per model. Profile results of these other benchmark points can be seen in Figure 3.

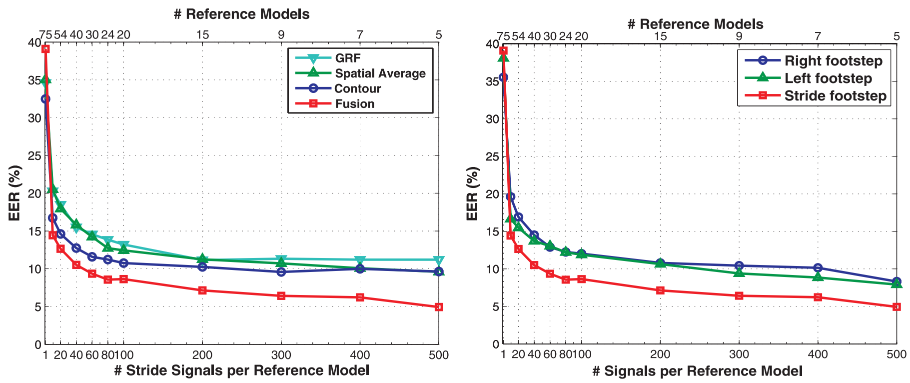
4 Experimental Results

This section describes the assessment of the time domain features described in Section 2 following the protocols defined in Section 3.

Figure 3(a) shows the EER against the different quantities of stride footstep signals used to train the reference models, bearing in mind that the number of reference models decreases as defined by the top abscissa axis. For example, the points on the left of the figure relate to 75 reference models using only 1 footstep signal to train each model; whereas points on the right relate to 5 reference models with 500 signals to train each model.

The figure shows EER results for the three feature approaches, i.e. the GRF, the spatial average, and the contour. Also a fourth plot in the figure shows the result of the fusion at the feature level of the three approaches, carried out concatenating the features of the single approaches after PCA. These results are generated for stride footsteps, which are comprised of concatenated right and left footstep signals.

All four plots have a similar overall shape with (i) an initial steep fall from approximately 35% EER to 15-20% EER when using 1 to 10 footsteps for training, (ii) a smooth knee curve when increasing the number of signals used in the reference models from 20 to 80 where the error rates change less rapidly from 18 to 13% for the cases of GRF and spatial average, from 15% to 11% for the case of the contour and from 13% to 9% for the case of the fusion of the three approaches; and (iii) relatively flat profiles where error rates are around 10% for



(a) Time Features

(b) Comparison Right/Left/Stride

Fig. 3. (a) Four plots of EER against number of stride signals used to train the reference models for concatenated (stride) footsteps. (b) EER against number of signals used to train the reference models with the fusion of the three feature approaches in the time domain for single (right and left) and stride footsteps.

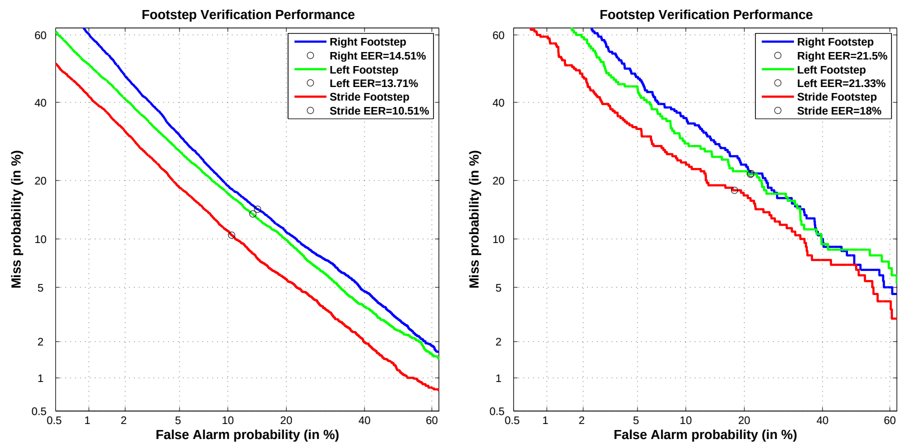
the three feature approaches (11% for the GRF) and around 5% EER for the case of the fusion when using 500 signals in the reference models.

This shows that in all cases the performance saturates when the number of signals per model exceeds approximately 80 footsteps. Also, errors as low as 5% are viable, especially with further system optimisation. It should be emphasized that: (i) these results relate to features extracted from the time domain only, no spatial information is considered here, and (ii) the number of trials varies along the abscissa axis.

It is interesting to note that the GRF and the spatial average features give similar performance, while the contour features provide the best results of the three approaches more accentuated in the left part of the figure, i.e. when using 10 to 100 signals to train the reference models. Also, the fusion outperforms the three single approaches. This approach provides the best results for footstep recognition using time domain features. The following results of this section relate to the fusion of the three feature approaches in the time domain.

Figure 3(b) shows the EER against different quantities of reference data for the case of the single footsteps (right and left) and the stride for the fusion of the three time domain feature approaches. As can be seen the three plots follow the same trend, but there is a significant improvement of performance when using the stride compared to the single footsteps (right and left), reducing the EER by an average of 3%.

Figure 4 analyses in more detail the case of benchmark B1 (i.e. using 40 signals per reference models and 40 models) comparing results obtained for the Development (Fig. 4(a)) and Evaluation (Fig. 4(b)) sets. In both cases there is a superior performance for the case of the stride footstep with an average relative improvement of 25%. It is interesting to see such a significant perfor-



(a) Development (b) Evaluation

Fig. 4. DET curves for fusion of the three features approaches. (a) Results for benchmark B1 for the Development set, and (b) for the Evaluation set.

mance degradation between the Development and the Evaluation. As described in Table 1, in both datasets there is a common reference data. As described, in these experiments the time sequence of the collection is kept, i.e. data used in the test sets was collected later in time than data used to train the reference models. Therefore data used in the Development set is closer in time to the reference than data used in the Evaluation set, and therefore more likely to be more similar. This explains the degradation observed in the Evaluation set.

Results achieved here are better compared to those obtained in the related works. Also, it is worth noting that the experimental setup here is the most realistic at least in two factors: (i) it considers the largest footstep database to date, and (ii) it keeps the time lapse between reference and test data, in contrast to most previous works, for example [8,9,10], which randomize the time sequence of the data in the experiments. The randomization makes reference and test datasets more similar and therefore it is possible to achieve artificially good results.

5 Conclusions and Future Work

This paper studies footstep signals as a biometric based on the largest footstep database to date with more than 120 people and almost 20,000 signals. Footstep signals collected contain information in both time and spatial domains, in contrast with previous related works.

This paper focuses on the analysis of the time information of the signals. Features such as the popular ground reaction force, together with two others approaches named the spatial average and the contour are compared and fused following a holistic approach with PCA and SVM.

The experimental protocol is designed to study the influence of the quantity of data used in the reference models, simulating conditions of possible extreme applications such as smart homes or border control scenarios. Results in the range of 5 to 15% EER are achieved in the different conditions for the case of the stride footstep for the fusion of the three feature approaches, which are better than previous works, and with a much more realistic experimental setup.

The time gap between reference data and test is an important point to consider in further work as we have observed a significant degradation of the performance in the Evaluation set which is comprised of the last data collected in the database.

Also, the analysis of the spatial information of the footstep signals and a fusion with the time domain information are very interesting lines for further research.

Acknowledgements

R.V.-R., J.F. and J.O.-G. are supported by projects Contexts (S2009/TIC-1485), Bio-Challenge (TEC2009-11186), TeraSense (CSD2008-00068) and "Cátedra UAM-Telefónica". The multimodal footstep database was collected with support from UK EPSRC and in this context the Authors wish to acknowledge the major contributions of Nicholas W.D. Evans and Richard P. Lewis.

References

1. A. Pedotti, "Simple Equipment Used in Clinical Practice for Evaluation of Locomotion," *IEEE Trans. on Biomedical Engineering*, vol. 24, no. 5, pp. 456–461, 1977.
2. Shoji, Y., Takasuka, T., and Yasukawa, H., "Personal Identification Using Footstep Detection," in *Proc. ISPACS*, pp. 43–47, 2004.
3. Liau, W. H., Wu, C. L., and Fu, L. C., "Inhabitants Tracking System in a Cluttered Home Environment Via Floor Load Sensors," *IEEE Trans. on Automation Science and Engineering*, 5(1):10–20, 2008.
4. Srinivasan, P., Birchefield, D., Qian, G., and Kidane, A., "A Pressure Sensing Floor for Interactive Media Applications," in *Proc. ACM SIGCHI*, 2005.
5. M. D. Addlesee, A. Jones, F. Livesey, and F. Samaria, "The ORL Active Floor," *IEEE Personal Communications*, vol. 4, pp. 235–241, 1997.
6. R. Vera-Rodriguez, N. Evans, and J. Mason, *Encyclopedia of Biometrics*, Springer, 2009, ch. Footstep Recognition.
7. J. Suutala and J. Roning, "Methods for person identification on a pressure-sensitive floor: Experiments with multiple classifiers and reject option," *Information Fusion*, vol. 9, no. 1, pp. 21 – 40, 2008.
8. R. Vera-Rodriguez, R. Lewis, J. Mason, and N. Evans, "Footstep recognition for a smart home environment," *International Journal of Smart Home*, vol. 2, pp. 95–110, 2008.
9. J. S. Yun, S. H. Lee, W. T. Woo, and J. H. Ryu, "The User Identification System Using Walking Pattern over the ubiFloor," in *Proc. ICCAS*, 2003, pp. 1046–1050.

10. L. Middleton, A. A. Buss, A. I. Bazin, and M. S. Nixon, "A floor sensor system for gait recognition," in *Proc. AutoID*, 2005, pp. 171–176.
11. J. Suutala, K. Fujinami, and J. Röning, "Gaussian process person identifier based on simple floor sensors," in *Proc. tEuroSSC '08*, 2008, pp. 55–68.
12. R. J. Orr and G. D. Abowd, "The Smart Floor: A Mechanism for Natural User Identification and Tracking," in *Proc. Conference on Human Factors in Computing Systems*, 2000.
13. C. Cattin, "Biometric Authentication System Using Human Gait." *PhD Thesis.*, 2002.
14. J. Suutala, S. Pirttikangas, J. Riekki, and J. Roning, "Reject-optional LVQ-based Two-level Classifier to Improve Reliability in Footstep Identification," *LNCS. Springer*, vol. 3001, pp. 182–187, 2004.
15. Y. Gao, M. J. Brennan, B. R. Mace, and J. M. Muggleton, "Person recognition by measuring the ground reaction force due to a footstep," in *Proc. RASD*, 2006.
16. J. Suutala and J. Roning, "Combining classifiers with different footstep feature sets and multiple samples for person identification," in *Proc. ICASSP*, vol. 5, 2005.
17. R. Vera-Rodriguez, N. W. D. Evans, R. P. Lewis, B. Fauve, and J. S. D. Mason, "An experimental study on the feasibility of footsteps as a biometric," in *Proc. EUSIPCO*, 2007, pp. 748–752.
18. R. Vera-Rodriguez, J. Mason, and N. Evans, *Handbook of Remote Biometrics*, Springer, 2009, ch. Assessment of a Footstep Biometric Verification System.
19. J. P. Stevenson, S. L. Firebaugh, and H. K. Charles, "Biometric Identification from a Floor Based PVDF Sensor Array Using Hidden Markov Models," in *Proc. SAS'07*, 2007.
20. R. Vera-Rodriguez, J. Mason, and N. Evans, "Automatic cross-biometric footstep database labelling using speaker recognition." in *Proc. ICB*, , Springer LNCS, vol. 5558/2009, 2009, pp. 503–512.