AN EXPERIMENTAL STUDY ON THE FEASIBILITY OF FOOTSTEPS AS A BIOMETRIC

Ruben Vera Rodriguez[1], *Nicholas W. D. Evans*[1,2], *Richard P. Lewis*[1], *Benoit Fauve*[1] *and John S. D. Mason*[1]

[1]School of Engineering, University of Wales Swansea, Swansea, SA2 8PP, UK
[2]LIA - University of Avignon, BP1228, 84911, Avignon Cedex 9, France
emails: [r.vera-rodriguez.405831, r.p.lewis, b.g.b.fauve.191992, J.S.D.Mason]@swan.ac.uk, nicholas.evans@univ-avignon.fr

ABSTRACT

This paper reports some experiments concerned with footsteps as a biometric. We present a comparison between different feature extraction techniques as well as classification methods, obtaining the best results with holistic feature extraction and a support vector machine (SVM) classifier. Results are reported in terms of detection error trade-off (DET) curves, showing minimum equal error rates of around 10%. Previous experimental work has limitation in sizes of databases; here we have over 3000 examples across 41 persons and are therefore able to design independent development and evaluation datasets. Once finished and validated, this database will be made freely available to the research community.

1. INTRODUCTION

Footstep recognition is a biometric that has been studied during the last 10 years but only by a relatively small number of researchers. As we review in Section 2, different techniques have been developed using different sensors, features and classifiers; however results are related to small databases in number of persons and footsteps and this is a limitation of the work to date.

In this paper we present results with a larger database and therefore we are able to present more statistically meaningful results and potentially more reliable predictions of performance. Footstep signals have been collected from two piezo electric sensors embedded into the underside of a rubber floor tile using a semi-automatic capture system. Details of the signal acquisition system are to be made available at [1]. This process is continuing and the results presented here relate to 3174 footsteps collected from 41 different persons. Once finished and validated this database will be made freely available to the research community in our web site [1]. This work serves to corroborate related work in the literature but in a more formal way. We have divided the data into independent development and evaluation datasets and adopted a standard verification assessment strategy. The current focus of our research is on feature extraction, where geometric and holistic approaches have been studied. Also, various classifiers are considered showing the benefit of statistical discriminative based classifiers in the form of a support vector machine (SVM) obtaining an equal error rate (EER) of 9.5% for development set and 11.5% for evaluation set for the holistic feature approach. In addition to this, we consider the different number of footsteps per person as a function of classification performance.

The remainder of the paper is organised as follows. Section 2 reviews related work in the literature. Feature extraction is covered in Section 3, experimental work and results are presented in Section 4 and finally our conclusions in Section 5.

2. REVIEW OF FOOTSTEPS AS A BIOMETRIC

When compared with other biometric modalities, footstep recognition is a relatively new biometric certainly judged in terms of published work. Table 1 summarises the material in the open literature.

As can be observed in the third column of Table 1, different sensor technologies have been used including load cells [2, 3, 8], switch sensors [5, 7], piezo force sensors [4] and electro mechanical film (EMFi) [6]. From Table 1 it is seen that the database sizes are relatively small, certainly in relation to other biometric evaluations where the persons are normally counted in hundreds or thousands and the number of tests perhaps in many thousands. A maximum number of 16 persons and 500 footsteps examples were gathered in all cases except in [3] which reports ID results on 1680 footsteps, but with only 15 persons. In all related work referenced in Table 1, databases were divided into training and testing sets, but none of them used development and evaluation sets, except [6] in which 5-fold cross-validation was used, but then with only a very small number of samples per subset.

As Table 1 indicates, different features are proposed, including subsamples from the ground reaction force (GRF) profile in [2], geometric features from the GRF in [3, 6, 8], the power spectral density in [4], position of several footsteps in [5], and stride length, stride cadence and heel-to-toe ratio in [7].

With respect to classifiers the majority used a simple nearest neighbour (NN) based Euclidean distance [3, 4, 6, 7, 8], perhaps because of the limited data sets which make statistical modeling difficult. However [2] uses an HMM classifier, [5] a multilayer-perceptron neural network and [6] learning vector quantization, even though such approaches normally demand large quantities of training data to see the true potential.

Identification, rather than verification, was the task considered in all but one of the cases, the exception being [4]. Identification has the benefit of utilizing the available data to a maximum but suffers from scalability in terms of the number of classes in the set.

3. FEATURES

Features play an important role in all biometrics and feature extraction is one of the key factors of our current work. Two

Group / Year	Database (total steps / persons)	Technology	Features	Classifier	Results
The ORL Active Floor / 1997 [2]	300 steps, 15 persons	Load cells	Subsampled GRF	HMM	ID rate: 91%
The Smart Floor (USA) / 2000 [3]	1680 steps / 15 persons	Load cells	Geometric feat. from GRF	NN	ID rate: 93%
ETH Zurich / 2002 [4]	480 steps / 16 persons	Piezo force sensors	Power Spectral Density	Euclidean distance	Verif. EER: 9.4%
Ubifloor (Korea) / 2003 [5]	500 steps / 10 persons	Switch sensors	Position of several steps	Multilayer-perceptron neural network	ID rate: 92%
EMFi Floor (Finland) / 2004 [6]	440 steps / 11 persons	Electro Mechanical Film	Geometric feat. from GRF	Learning vector quantization	ID rate: 70%
Southampton University (UK) / 2005 [7]	180 steps / 12 persons	Resistive (switch) sensors	Stride length, stride cadence and heel-to-toe ratio	Euclidean distance	ID rate: 80%
Southampton University (UK) / 2006 [8]	400 steps / 11 persons	Load cells	Geometric feat. from GRF	NN	ID rate: 94%

Table 1. A Comparison of different approaches to footstep recognition 1997 - 2006.

generic approaches have been considered, namely: geometric and holistic. The sampling frequency is 1024Hz and the maximum duration of a footstep equates to 1400 samples (1.37 seconds).

3.1 Geometric Approach

The majority of related work in the literature consider the GRF [2, 3, 4, 6, 8] as features. This is "the reaction produced by a measuring device to the weight and inertial forces of a body in contact with the device" [2].

The signals that our system produces relate to the instantaneous pressure for each sensor along the footstep. Figure 1 shows a typical footstep waveform. The relevant points, shown by numbers in Figure 1, were chosen as an indication of the signal's behaviour along time, similar to the work of [2, 3, 8]. These points are the relative and absolute maxima and minima present in the two sensors indicated above. Point 1 corresponds to the effect of heel pressure on the first sensor, the dashed profile in Figure 1. Points 2 to 5 correspond to the second sensor, the solid profile in Figure 1, and show the effect of the toe. Point 2 shows the initial pressure of the toe, point 5 shows the effect of the pushing off of the toe and points 3 and 4 mark the transition between points 2 and 5. The time and magnitude of these 5 points result in the first 10 features. Then, the inter-difference between each pair of points results in another 20 features (10 magnitude features and 10 time features). Finally, 6 additional features, the area, mean and standard deviation of both sensors, are concatenated to obtain a feature vector with a total of 36 geometric features for each footstep signal. These features were normalised with respect to the absolute maxima of the profile.

3.2 Holistic Approach

Holistic features have gained popularity in the field of face and handwritten word recognition with good results reported in [9, 10, 11, 12].

Holistic features are comprised of the first 1400 samples of the Heel and Toe sensor (as the example of Figure 2 (a)

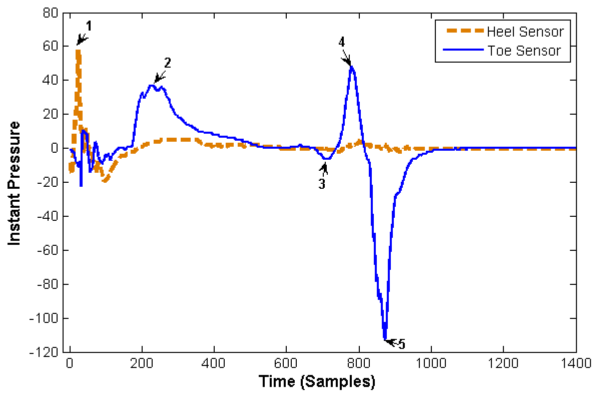

Figure 1. Instant pressure against time. Relevant points for geometric feature extraction are indicated.

and (b)), and also the first 1400 samples of the GRF (as in Figure 2(c)), calculated as the integration over time for these two sensors. In total 4200 holistic features have been obtained after normalization of each sensor and the GRF by its maxima. We refer to this configuration as holistic in terms of time domain as we introduce in the feature vector all the possible information available from a footstep.

Due to the high dimensionality of this holistic feature vector, principal component analysis (PCA) [13] was used to distil the information content. Thus, after PCA, a set of principal components is obtained, where each of them is a linear combination of the original feature set. Figure 3 shows how using the first 80 principal component, more than 96% of the original information is retained whilst achieving an 98% reduction in dimensionality.

4. EXPERIMENTAL WORK

Two sets of experiments are presented here. The first set compares the geometric and holistic feature extraction ap-

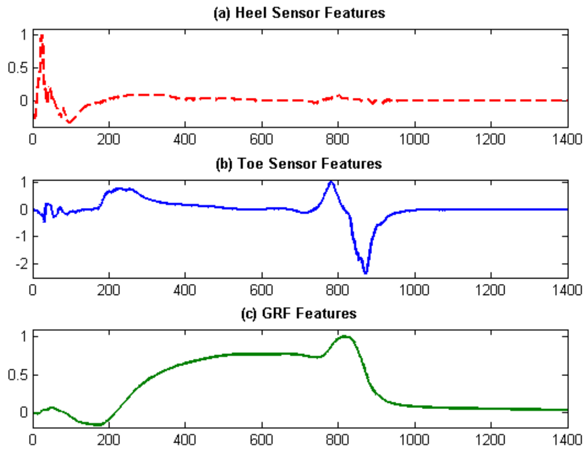

Figure 2. Holistic features used. (a) Heel sensor features. (b) Toe sensor features. (c) GRF features.

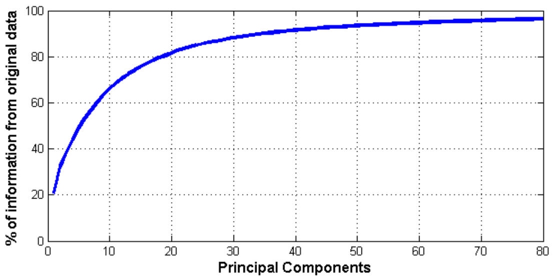

Figure 3. Percentage of information from original data against number of principal components.

proaches with K-NN and SVM classifiers. The second experiment aims to determine the number of footsteps needed per model. Results are presented with detection error trade-off (DET) curves [14, 15] as is popular in other biometric verification tasks.

4.1 The footstep database

The current footstep database is comprised of 3174 footsteps collected from 41 persons. Two subsets have been identified: a client set of 17 persons with an average of 170 footsteps per person (2884 total footsteps) and an impostor set of 24 persons with an average of 15 footsteps per person (290 total footsteps). Each person in the client set provides footsteps with at least two different shoes and the database is divided arbitrarily into independent development and evaluation datasets, each comprised of their own training and testing subsets.

Table 2 illustrates the distribution of the footstep data into the different datasets. It is worth noting that there is no data overlap between the Development Set and the two Evaluation Sets. The Development Set is comprised of footsteps from clients P1 to P8, each contributing 40 footsteps for training and another 40 footsteps for testing. Evaluation Set 1 is a balanced set comprised of footsteps from clients P1 to P17 where, for each client, there are 40 footsteps for training and another 40 for testing. Evaluation Set 2 uses all the footsteps available in the database that are not contained within the Development Set, and it is therefore an unbalanced set in terms of the number of footsteps per person. It is comprised of footsteps from client P1 to P17 with 45 footsteps per client for training, and an average of 87 footsteps per client for testing, the range being 40 to 170 footsteps per client. Thus Evaluation Set 1 is a subset of Evaluation Set 2.

	Devel	**Eval 1**	**Eval 2**
	Train / Test	*Train / Test*	*Train / Test*
Clients	*P1-P8 / P1-P8*	*P1-P17 / P1-P17*	*P1-P17 / P1-P17*
Footsteps/Client	*40 / 40*	*40 / 40*	*45 / 87*
Impostors	*P18-P41 / -*	*P18-P41 / -*	*P18-P41 / -*
Impostor Footsteps	*290 / -*	*290 / -*	*290 / -*
Subset Data	*610 / 320*	*970 / 680*	*1055 / 1479*
Total Set Data	*930*	*1650*	*2534*

Table 2. Distribution of footsteps in the datasets

As a part of the recognition system, the impostor footsteps are the same for all three datasets and come from persons P18 to P41 with a total number of 290 footsteps.

4.2 Features and Classifiers comparison

Both geometric and holistic features were assessed with K-nearest neighbour (K-NN) and support vector machine (SVM) classifiers. K-NN [16] is a simple method based on the Euclidean distance between the testing features and the K nearest training features. K was empirically optimised on the Development Set to K=20. The SVM [17, 18] is a statistical discriminative based classifier that finds an optimal hyperplane which maximizes the margin between in-class and out-of-class data. Results shown were obtained using a radial basis function (RBF) as a kernel.

Figure 4 illustrates DET curve results for the Development Set using both geometric and holistic features with K-NN and SVM classifiers. The first observation is that, as might be expected, better results are obtained using the SVM classifier than with the K-NN classifier. An interesting result is the superior performance of the holistic features. The best result is achieved when holistic features are applied to the SVM classifier where an equal error rate (EER) of 9.5% is observed. This contrasts with an EER of 17.5% when the geometric features are applied to the same classifier. This equates to a relative improvement of 46% in terms of EER with the holistic features. The same trend is observed with the K-NN classifier though the relative improvement is smaller (24% ERR with holistic features c.f. 28.5% with geometric features, relative improvement of 16%). In conclusion the SVM outperforms the K-NN, and holistic features outperform geometric features.

To assess the portability of the recognition system to new, unseen data it was then applied to Evaluation Set 1 without modification. The results are illustrated in Figure 5. Evaluation Set 1 has the same distribution as the Development Set, that is, there are 40 footsteps per client for training and another 40 for testing. The total number of footsteps is 1650. In all cases like-for-like results are marginally worse for the Evaluation Set 1 than for the Development Set however, reassuringly, the same trends are observed. The best result is again achieved using holistic features with an SVM classi-

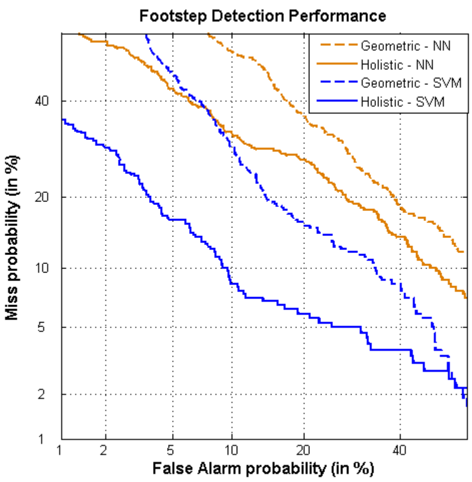

Figure 4. DET curves for Development Set with combination of K-NN/SVM classifiers and holistic/geometric features

fier (EERs of 12.5% for Evaluation Set 1 c.f. 9.5% for the Development Set).

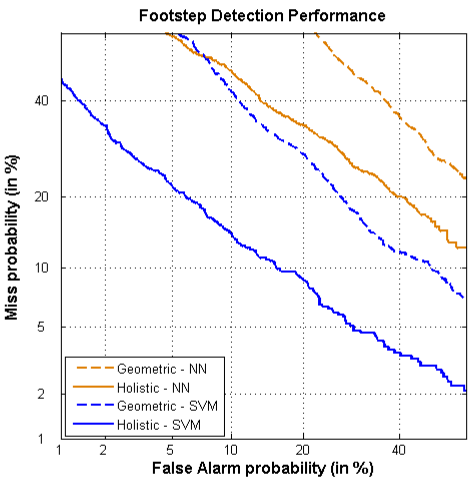

Figure 5. DET curves for Evaluation Set 1 with combination of K-NN/SVM classifiers and holistic/geometric features

Evaluation Set 2 is unbalanced and uses all the available data (2534 footsteps). As illustrated in Table 3 results are very similar to those obtained with Evaluation Set 1. Once again the same trends were observed and results correlate well with the results of Evaluation Set 1 (EERs of 9.5% for Dev. 12.5% for Eval. 1 and 11.5% for Eval. 2). The slight increase in performance in the case of holistic features with an SVM classifier can be attributed to the larger number of footsteps that are used to train each model. This raises an interesting question regarding the number of footsteps that are required to reliably train a model.

	Devel	Eval 1	Eval 2
Geometric - NN	*28.5%*	*38.5%*	*38%*
Holistic - NN	*24%*	*28%*	*28%*
Geometric - SVM	*17.5%*	*23.5%*	*23.5%*
Holistic - SVM	*9.5%*	*12.5%*	*11.5%*

Table 3. EER result of the experiments with the datasets

4.3 Data quantity for training models

A second experiment was carried out in order to measure the variation in performance against the number of footsteps per model. In general better recognition accuracy can be achieved with more data. However, the amount of data that can be used for training is limited by the number of footsteps that can be collected in practical enrolment. For this experiment, the whole database (3174 footsteps in total) was divided into training and testing sets. the size of the training set was varied from 1 to 63 footsteps per model (also for persons P1 to P17), adding one footstep per model for each iteration of the experiment whilst the testing set was comprised of 40 footsteps per client (P1 to P17). Both geometric and holistic features were assessed using an SVM classifier with an RBF kernel function and is exactly the same classifier as used for the experiments reported in section 4.2. Results are illustrated in Figure 6 and show that in both cases the EER (vertical axis) falls gradually as the number of footsteps per model (horizontal axis) increases. In all cases holistic features outperform the geometric features, the best results in each case being 10.5% and 21.5% for holistic and geometric features respectively. Recognition performance does not significantly increase beyond 40 footsteps per model and there is less than 2% improvement in EER when using 63 footsteps per model than when 40 are used.

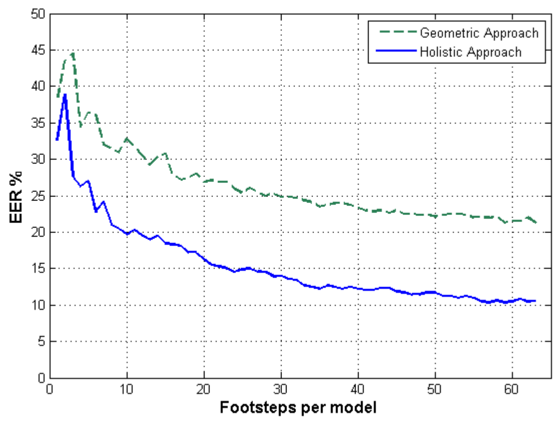

Figure 6. EER against number of footsteps per model present in training data. Testing data is comprised of 40 footsteps per each client.

5. CONCLUSIONS

A semi-automatic system for capturing footsteps has led to a database comprised of more than 3000 footsteps, and an objective of this project is to release this database to the research community, making it publicly available in the near future.

We have described in this paper the first experiments and results carried out on our system in the context of footsteps as a biometric. We have combined techniques used in other biometrics and have devised tests that give some confidence in performance assessment. Interestingly, verification error rates as low as 10% are predicted from our experiments and these results would seem to agree with previously reported work.

We also show that performance improves steadily when more footsteps are used in training, with only small improvements after about 40 examples per person.

Whilst the results are still limited compared to other biometric assessments they are based on more than 3000 footsteps from 41 persons overall and this is significantly beyond other results reported to date.

6. ACKNOWLEDGEMENTS

We gratefully acknowledge the support of the UK Engineering and Physical Science Research Council (EPSRC) grant and the European Social Funding (ESF).

REFERENCES

[1] Footstep recognition at Universtity of Wales Swansea available at http://eegalilee.swan.ac.uk

[2] M. D. Addlesee, A. Jones, F. Livesey and F. Samaria, "The ORL Active Floor," in *IEEE Personal Communications*, vol 4 ,1997, pp. 35-41.

[3] 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.

[4] Philippe C. Cattin, "Biometric Authentication System Using Human Gait," *PhD Thesis*, Zurich, 2002.

[5] Jae-Seok Yun, Seung-Hum Lee, Woon-Tack Woo and Je-Ha Ryu, "The User Identification System Using Walking Pattern over the ubiFloor," in *Proc. International Conference on Control, Automation, and Systems*, 2003, pp. 1046-1050.

[6] J. Suutala and J. Rning, "Toward the adaptive identification of walkers: automated feature selection of footsteps using distinction-sensitive LVQ," in *Proc. International Workshop on Processing Sensory Information for Proactive Systems*, 2004, pp. 61-67.

[7] Lee Middleton, Alex A. Buss, Alex I. Bazin and Mark S. Nixon, "A floor sensor system for gait recognition," in *Proc. Fourth IEEE Workshop on Automatic Identification Advanced Technologies*, 2005, pp. 61-67.

[8] 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. 9th International Conference on Recent Advances in Structural Dynamics*, 2006.

[9] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, vol. 3, pp. 72-86, Jan. 1991.

[10] W. Zhao, A. Krishanaswamy and R. Chellappa, "Discriminant Analysis of Principal Components for Face Recognition," in *Proc. International Conference on Automatic Face and Gesture Recognition*, 1998, pp. 336-341.

[11] W. Zhao, R. Chellappa, A. Rosenfeld and P. Phillips, "Face recognition: A literature survey," *UMD CfAR Technical Report CAR-TR-948*, 2000.

[12] S. Madhvanath and V. Govindaraju, "The Role of Holistic Paradigms in Handwritten Word Recognition," in *IEEE transactions on Pattern Analysis and Machine Intelligence*, 2001.

[13] I.T. Jolliffe, *Principal Component Analysis*. Springer Series in Statistics, 2002.

[14] A. Martin, G. Doddington, T. Kamm, M. Ordowski and M. Przybocki, "The DET curve in assessment of detection task performance," *Eurospeech*, vol. 1, pp. 1895-1898, 1997.

[15] NIST, "DET-Curve Plotting software for use with MATLAB," *Software available at http://www.nist.gov/speech/tools*.

[16] T. Denoeux, "A k-nearest neighbor classification rule based on Dempster-Shafer theory," *IEEE transactions on System, Man and Cybernetics*, vol. 25(05), pp. 804-813, 1995.

[17] V.N. Vapnik, *Statistical Learning Theory*. New York: Wiley, 1998.

[18] C.J.C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowl. Discov.*, vol. 2, no. 2, pp. 1-47, 1998.