Research Article

Gait recognition: highly unique dynamic plantar pressure patterns amongst 104 individuals

Todd C. Pataky[1,*], Tingting Mu[2] , Kerstin Bosch[3], Dieter Rosenbaum[4], John Y. Goulermas[5]
*Corresponding author

[1]Department of Bioengineering, Shinshu University, Tokida 3-15-1 Ueda, Nagano, 386-8567, Japan.

[2]School of Computer Science, University of Manchester, Oxford Road, Manchester, M13 9PL, UK.

[3]Movement Analysis Lab, Social Paediatric Centre, St.-Vincenz-Hospital, Suedring 41, Coesfeld, 48653, Germany

[4]Movement Analysis Lab, Institute of Experimental Musculoskeletal Medicine, University Hospital Münster, Domagkstr. 3, 48149, Münster, Germany.

[5]Department of Electrical Engineering and Electronics, University of Liverpool, Brownlow Hill, Liverpool, L69 3GJ, UK.

Word counts:
Summary: 177 words (200 words max)
Media summary: 98 words (100 words max)
Main text: 6971 words (2500-8000 words)

Submitted to: Royal Society Interface
Date: 17 August 2011   (Revision #2)

Running head:
Foot pressure-based biometric identification

Subject areas: Biometrics, Biomechanics, Bioengineering

Keywords: gait recognition, pedobarography, foot loading, image registration, dimensionality reduction, locomotion biomechanics

SUMMARY (177 words) (200 words max)
Everyone's walking style is unique, and it has been shown that both humans and computers are very good at recognising known gait patterns. It is therefore unsurprising that dynamic foot pressure patterns, which indirectly reflect the accelerations of all body parts, are also unique, and that previous studies have achieved moderate-to-high classification rates using foot pressure variables. However, these studies are limited by small sample sizes (N<30), moderate classification rates (CR=~90%), or both. Here we show, using relatively simple image processing and feature extraction, that dynamic foot pressures can be used to identify N=104 subjects with a CR of 99.6%. Our key innovation was improved and automated spatial alignment which, by itself, improved CR to over 98%, a finding that pointedly emphasises inter-subject pressure pattern uniqueness. We also found that automated dimensionality reduction invariably improved CRs. Since dynamic pressure data are immediately usable, with little or no preprocessing required, and since they may be collected discreetly during uninterrupted gait using in-floor systems, foot pressure-based identification appears to have wide potential for both the security and health industries.

MEDIA SUMMARY   (98 words) (100 words max)
Everyone's walking style is unique, and camera- and force-based systems are excellent at recognising known gait patterns. Here we show that foot pressure patterns can also be recognised, with relatively simple data processing, with an accuracy of 99.6% (104 subjects). The key processing step was spatial alignment which, by itself, improved accuracy to over 98%. Since foot pressure devices can be discreetly installed in the floor and are, in general, robust to multi-subject environments because two feet cannot be in the same place at the same time, foot pressure-based identification appears to have strong potential for security applications.

1. INTRODUCTION

When walking our feet interact with the ground in a stereotypical fashion: heel-strike, roll to the forefoot, then push-off with the distal forefoot and toes [1] (Fig.1a). This process takes about 0.7 s when walking at normal speeds of about 1.2 m/s. Is it possible that, within these stereotypical constraints, all individuals interact with the ground uniquely?

Based on the gait recognition literature this seems plausible: individuals move their bodies and limbs in highly unique and highly repeatable patterns [2], and camera-based computer systems can be trained to recognise these patterns [3], even in adverse conditions such as poor lighting and brief exposure [4]. We would therefore expect these highly unique movement patterns to be reflected, to a certain extent, in our mechanical interaction with the ground, and that computers could be similarly trained to recognise gait patterns from floor-based sensors. Indeed floor-based gait recognition has already been highly successful. Recent examples include use of ground reaction force (GRF) trajectories, wavelet decomposition and fuzzy set-based feature extraction to recognise individuals with classification rates (CR) of 97% [5] and 99% [6].

While both camera-based and GRF-based gait recognition have been widely successful, both also have certain practical limitations. Camera systems must overcome environmental noise, perspective, and other 3D calibration problems, which state-of-the-art systems can do impressively, but with only moderate accuracy (74%) [7]. Force plate-systems must be quite large, at least 0.5 m long for full foot contact during non-targeted gait, but multiple feet mustn't contact the plate at the same time, meaning that force plates cannot be positioned arbitrarily and also that they cannot be used in multi-subject environments.

An alternative is plantar pressure imaging (PPI) [8]. PPI systems typically consist of an array of hundreds or thousands of pressure-sensitive sensors which are capable of characterising plantar pressure distributions at spatial and temporal resolutions on the order of 5 mm and 100 Hz, respectively. There are a variety of PPI technologies [8], but in their final form most systems are thin, flat, relatively rigid boards that can be embedded in the floor to be flush with the walking surface. PPI systems do not suffer from environmental noise because the foot can be very easily isolated from the environment using low-pressure thresholding. Even though an individual may walk over a PPI plate at arbitrary angles, PPI systems also do not suffer from perspective problems because foot images may be spatially aligned using automated registration techniques [9,10]. Finally, high spatial and temporal resolutions mean that PPI systems can be used in multi-subject environments as all footsteps are, by nature, spatiotemporally isolated.

PPIs are qualitatively highly unique amongst different subjects (Fig.2), and PPI-based biometric identification has consequently also had varying degrees of success (Appendix A). Most of these studies report moderate accuracy (80-85%), and we are aware of only four that report accuracies greater than 90% for sample sizes of at least ten subjects: [11] – 98.6%, [12] – 96.0%, [13] – 93.1%, [14] – 92.3%. However, the maximum number of subjects tested in these studies was eleven, and in a variety of pilot tests we were unable to reproduce the best of these results [11], perhaps partly because we were unable to resolve certain ambiguities in the authors' algorithm descriptions. Only one study examined more than 11 subjects [15] (N=30), but accuracy was notably lower (86.1%) than a previous study by the same group with fewer subjects: [13] – 93.1% (N=10). To date high accuracies on samples notably larger than N=10 have only been achieved using complimentary information like high-resolution skin prints [16] – 99% (N=32) or 3D foot sole shape [17]– 98.7% (N=30), information which cannot be readily obtained during uninterrupted gait because of lengthy scanning durations.

Of the purely PPI studies, it is notable that many have employed spatial normalisation procedures; since the foot may adopt an arbitrary posture with respect to the PPI device, it seems logical to compensate for arbitrary postures using spatial normalisation. However, we note that most of these studies employed decorrelation (Appendix A), or equivalently: principal axis alignment [18], an approach which has been shown to yield much

poorer alignment than optimisation-based alignment procedures [10]. It is therefore conceivable that improved spatial alignment would yield improved biometric identification. It is also notable that previous PPI studies used a variety of pre-selected features to be extracted from the raw data (Fig.1), but none, to our knowledge, has conducted a systematic evaluation of the relative effectiveness of different features. The purposes of this study were thus: (1) to explore the feasibility of PPI-based gait recognition on a larger sample of subjects (N>100), (2) to systematically compare a variety of spatial alignment procedures, and (3) to systematically compare a variety of features and feature extraction procedures.

2. METHODS

2.1 Data
Plantar pressure data were collected from 104 healthy individuals at the University of Münster (Table 1). These data were previously used to compute a healthy 'average' pressure distribution [19]. Data were recorded for 1.0 s at 50 Hz using an EMED ST4 system (resolution: 5 mm) (Novel GmbH, Munich, Germany). Each subject performed a total of ten trials of self-paced walking, five for each foot, yielding a total of 1040 3D (x, y, time) images (Fig.1a). 'Follow-up' data from ten of these subjects were collected separately (Table 1); these data were obtained up to five years prior to the main data collection sessions. Prior to participation all subjects provided informed consent according to the policies of the University of Münster.

The left- and right-foot images were examined separately after finding that single-foot analyses yielded sufficiently high performance. This is justifiable, we believe, because (i) the literature shows that lower limb dominance is poorly defined [20], (ii) naturally occurring gait asymmetries tend to load left and right feet differently [21], and (iii) in *post hoc* analyses we found no systematic left-right asymmetries amongst subjects. We may thus justifiably regard the left- and right-foot datasets as essentially independent, at least for the purposes of validating our methods on the population from which the present subjects were drawn.

2.2 Image alignment
Images were spatially padded by adding at least 1 cm of zero pressure rows/columns to the foot periphery. They were then temporally aligned so that the first (x,y) time slice corresponded to initial heel-strike. Following padding all images were contained in a $65 \times 29 \times 50$ voxel grid (x, y, time) (94,250 voxels) of which an average of 8,291 voxels (8.8%) were non-zero for any given trial; across all subjects and trials 33,143 (35.2%) were non-zero. The raw data were quite smooth (Fig.1a,b) so images were neither spatially nor temporally filtered.

Subsequently three categories of spatial alignment procedures were tested (Table 2). The first: 'None' performed no alignment, passing raw images directly to feature extraction (below). The second: 'Decorrelation' performed a principal axis transformation to centre the pressure-weighted foot centroid and to vertically align the foot's minor principal axis. The third: 'Registration' [22] utilised a rapid frequency-based alignment procedure [9,23] to automatically align a given image to a foot template. The goal of the algorithm was to maximise cross-correlation, first in the frequency domain, to optimise horizontal and vertical foot translations, and then in the log-polar domain, to optimise foot rotation. Example data (Fig.3) reveal that both Decorrelation and Registration tended to improve alignment, although registration performed qualitatively better, agreeing with previous results [10].

For Registration, seven template images were tested including: (i) the morphologically average contralateral foot from the Münster data sample (RegMunCont) [19], (ii) an average foot from a separate study, a separate laboratory, and collected with a different manufacturer's equipment [10], and (iii-vii) average feet of the chronologically first five subjects from the cited study. Bilinear interpolation was used for all image transformations.

2.3 Pre-features

Since desktop computing memory was inadequate to submit the 3D images directly to classification routines, and since classifiers generally perform more poorly with increasing dimensionality [24], the images were first reduced to ten different 2D spatial (x,y) 'pre-features' by extracting specific characteristics of each pixel time series (Fig.1c,d) (Table 3); we refer to these as 'pre-features' to distinguish them from the final features upon which classification was based; the final features were extracted automatically from the pre-features using various dimensionality reduction techniques (Sect. 2.4).

Specific pre-features included those commonly used in the plantar pressure literature: 'peak pressure' or equivalently 'maximum pressure', or equivalently '100th percentile pressure' (P100). This 2D variable represents the maximum pressure experienced by each part of the foot over the course of stance, and is by far the most common variable seen in the plantar pressure literature, often used to check for plantar tissue overloading [8]. Other common variables analysed included: the pressure-time integral (PTI), contact duration (CD), and time-to-maximum (Tmax) [8]. The pressure-time integral represents the total loading during stance; areas of the foot with brief, high-pressure impulses may have a similar PTI value to areas with long, low-pressure impulses. Since the precise variable(s) regulating plantar tissue breakdown are unknown, PTI, which quantifies loading in a different way, has also been commonly analysed in the literature. CD is a PTI-like variable which considers only loading duration, not magnitude, and Tmax represents yet another loading feature: loading rate (with respect to initial heel contact). The point is that PPI data are complex, and that no single 2D variable can characterise the 3D loading profile.

In addition to these common variables, we also tested one that is less commonly used: time-to-first contact (Tfirst) [25] and others that, to our knowledge, have not been previously reported the 90th, 80th, 70th, 60th, and 50th percentiles (P90, P80, P70, P60, P50). Tfirst, like Tmax, represents a specific loading-rate feature: the speed with which one transitions to different parts of the foot. This is less common than the aforementioned variables, most likely because load magnitude is quite low at first-contact. The percentile variables, we believed, were also worth testing, partly because P100 is a maximum function, and therefore may be more susceptible to sensor noise than other percentiles, and partly to check if there was a systematic effect on the ultimate results as one considers relatively higher pressures. All aforementioned pre-features were tested either individually or in pairs, by vectorising then stacking 2D images. Since the full image time-series were too large for practical testing, 2D feature-pairing permitted inclusion of additional dynamic characteristics.

2.4 Dimensionality reduction

The second feature extraction phase used automated dimensionality reduction to further reduce the pre-features to a dimensionality most effective for classification. Reduction algorithms included (Table 4): Laplacian eigenmaps (LE) [26], normalised spectral clustering with a symmetric Laplacian (NCSL) [26], kernel principal component analysis (KPCA) [27], and locally linear embedding (LLE) [28]. Following semi-systematic analysis (Sect. 3.1) we found that reduction to a dimensionality of 70 (from 1885 dimensions for single pre-features and 3770 dimensions for paired pre-features) worked well for these data. Other reduction parameters were manually tuned for the left foot using 104-fold cross-validation (Sect. 2.5), and final performance was verified on the left-foot dataset and also with a separate (leave-one-out) validation scheme. As a baseline comparison we also used no dimensionality reduction, submitting pre-features directly to classification.

2.5 Classification

Classification of the final features was performed using nearest-neighbour (1NN) classification; this is the simplest possible classification scheme, detecting only the image most similar to the test image (i.e. minimum Euclidian distance) in reduced feature space. Although simple, 1NN was selected to emphasise the power of

automated dimensionality reduction for biometric-relevant feature extraction. Classifier performance was validated using 104-fold cross-validation (104-CV) and separately using leave-one-out cross-validation to ensure that 104-CV was not biased. We also employed a stratified 5-CV, wherein the first image of each subject was retained for testing, while the remaining four were used for training, and then repeated for the second images, third images, etc. This scheme (with a testing/training ratio of 25%) was adopted to ensure that the low testing/training ratio of 0.97 % in 104-CV was not a biasing factor.

2.6 *Algorithm evaluation*

A full-factorial evaluation of all aforementioned factors (alignment, pre-features, dimensionality reduction techniques, classification algorithms) would have required a prohibitively large number of iterative tests so we narrowed our focus by conducting semi-factorial evaluations in an *ad hoc* manner. For example, if variable P100 was found to perform generally better than other pre-features, then we used P100 to explore different alignment procedures, and the resulting best alignment procedures were used to re-test all pre-features. While incomplete, this approach proved to yield highly accurate classification performance.

Statistical hypothesis testing was conducted on a variety of classification-relevant metrics in an *ad hoc* manner as context demanded. For example, a paired-sample t test was used to test whether the difference between the None and Decorrelation alignment methods was different from zero; the motivation for this particular analysis was to examine whether Decorrelation, the predominant alignment procedure in the literature (Appendix A), is a better alignment choice than None. All aforementioned data processing was conducted in Matlab 7.10 (The MathWorks, Natick, MA, USA), and all figures were created using Matplotlib 0.99 as released with the Enthought Python Distribution 5.0 (Enthought Inc., Austin, TX, USA).

3. RESULTS

3.1 *Basic results*

With no image processing at all (except for image padding) nearest-neighbour classification identified individuals with an accuracy of 90.8% using the P100 pre-feature (Fig.4). Decorrelation surprisingly yielded a slightly lower average classification rate (CR) of 90.2%, while registration markedly increased the average CR to 98.9%. Dimensionality reduction also tended to improve CRs (Fig.4), albeit to a lesser extent than registration.

Across both feet the best-performing embedding dimension was 70 (Fig.5). Using this dimensionality, and following a systematic, semi-factorial study of the different alignment algorithms, pre-features, and dimensionality reduction schemes (Tables 5,6), the highest CR we were able to achieve in a single foot was 99.8% (519/520 correctly classified images). This was achieved on the right foot using RegMunCont alignment, the combined P100 and P80 pre-features, and LLE dimensionality reduction. For this set of parameters the left foot CR was 99.4% (517/520). Our semi-factorial analyses and manual parameter tuning were found to be unbiased as leave-one-out cross validation (Table 6b), as well as validation on the left foot yielded practically identical results (Table 6a). Additionally, we found that the low testing/training ratio of 0.97% in our validation scheme was not a biasing factor, as a 5-CV scheme (with a testing/training ratio of 25%) yielded CRs of 99.4% in both the left and right feet.

3.2 *Follow-up dataset*

Using the aforementioned 'best' parameters, CRs for the left and right feet were 98% (49/50) and 90% (45/50), respectively, for the ten-subject follow-up dataset (Fig.6a,b). We note, however, that one of the follow-up subjects had significantly higher right-foot metatarsal pressures in the 2007 'follow-up' trials than in the 2009 'original' trials (Fig.6c) ($p=0.005$, two-sample t test on extracted regional data [8]), and this led to 4/5

misclassifications for this subject's right foot. Upon questioning, this subject could not recall any orthopaedic condition that could explain the 2007-to-2009 metatarsal pressure difference. We also note that all five of this subject's left-foot follow-up images were correctly identified. If we exclude this subject's right foot data from follow-up analyses the CR across the nine remaining subjects would be 97.8% (44/45).

Once the classifier was trained on the 520 images from the original dataset, each follow-up image was read from disk and classified in 2.8 and 12.5 ms, respectively, as tested on a desktop computer (2.93 GHz dual-core processor, 4 GB memory, USB 2.0 connection to hardware) and averaged across the 100 follow-up images. Even though data transfer delays between the pressure measurement system and PC are longer than reading from disk (~64 ms, pilot results), a single footstep could still likely be identified within 100 ms of toe-off in a real-time implementation.

3.3 Decorrelation

Decorrelation *decreased* the average CR by 3.4% and 3.6% for no-reduction and LLE-reduced data, respectively, across all pre-features (Table 5) and both feet. After correcting for (two) multiple comparisons with a Bonferroni threshold of $p=0.025$ (family-wise Type I error rate: $\alpha=0.05$), paired t tests verified the significance of this decorrelation-induced CR drop ($p<0.001$ and $p=0.004$ for None and LLE, respectively). This finding was supported partially by root-mean-squared error (RMSE) results for the no-pre-processing, decorrelation, and registration (RegMunCont) conditions of: 22.4 ± 7.5, 18.0 ± 7.2, and 12.2 ± 5.7 kPa, respectively (mean ± st.dev., computed with respect to the intra-subject mean foot). It was further supported by ANOVA on no-alignment vs. decorrelation MSE; a significant SUBJECT effect was found ($p<0.001$), but no significant DECORRELATION effect was found for either the entire time series ($p=0.934$) or the P100 pre-feature ($p=0.339$). A marginal FOOT effect was found for the time series data ($p=0.070$) but not for the P100 pre-feature ($p=0.338$); since our best-performing classifier used only 2D pre-features (including P100) we may conclude that decorrelation's failure to reduce intra-subject MSE was similar in both feet. . In agreement with the present CR results (Fig.4), the present ANOVA results imply that decorrelation was not effective at reducing intra-subject variability. Therefore choosing decorrelation over no-alignment may not be statistically justified, in general, unless initial foot posture is highly variable. Indeed, over all tested parameter combinations registration invariably out-performed decorrelation.

3.4 Foot shape vs. pressure distribution

The best alignment and reduction schemes with a binary P100 pre-feature (i.e. a binary image defined by the inequality: P100>0) yielded CRs of 93.7% and 96.5% for the left and right feet, respectively. As compared with the continuous-pressure P100 pre-feature (Fig.1d) binary features reduced the CR by only 4.2%, suggesting that a large proportion of the present classification-relevant information was derivable simply from 5 mm-resolution foot shape. Nevertheless, in semi-factorial studies we were unable to achieve binary P100 performances greater than 97%, suggesting that pressure distribution information is necessary for optimal subject identification.

4. DISCUSSION

4.1 Classification

The fact that essentially no processing (except for zero padding) yielded CRs greater than 90% across 104 subjects, as well as the currently best results of CR>99%, strongly suggest that PPI data contain high-quality biometric information. This inter-subject uniqueness could only have been in embodied in plantar foot shape, dynamic plantar pressure distribution, or both, as these constitute the only subject-specific information sources in PPI data. The present binary image results of CR=~95%, which were very similar to previous binary image results of CR=94.6% [16] clarified that foot shape itself constituted a substantial source of classification-relevant

information in the current sample. Nevertheless, the original non-binary data pushed these CRs above 99%, suggesting that pressure patterns embody additional non-trivial inter-subject uniqueness.

In agreement with reports of high day-to-day PPI reliability [29], follow-up testing was also highly successful, yielding CRs of ~98%, despite fairly extensive delays of up to five years between testing sessions. Together with the presently estimated processing times of less than 100 ms per footstep these CR results suggest that PPI-based biometric identification may be suitable for real-world security applications.

Recent successes in PPI-based classification of healthy foot types [30], pathological state [31], and PPI-based fall detection [32] indicate that the current registration-based approach may also be useful for health-related applications. We hope to explore some of these applications in future work.

4.2 Previous studies

The current CR results are, to our knowledge, higher than previous purely PPI-based identification studies (Appendix A) except a previous five-subject study [33] (CR=100%). The best performing algorithm on a database of at least N=10 subjects was Jung et al. [11]: CR=98.6% (N=11), but a potential drawback to this study was that two steps were obtained on a short (80 cm) platform; given average foot lengths of 25.5 cm [34] and average stride lengths of 76 cm [35], subjects would have had to adopt unnaturally short strides to achieve two complete footfalls on the measurement platform. Regardless, Jung et al.'s results imply that a larger database of subjects may be identifiable even during unnatural or constrained gait. The remaining studies examined fewer than twelve subjects (except for [15]: CR=86.1%, N=32) and reported moderate CRs in the range 64-94%.

The higher current CRs can only be explained, we believe, by better data quality (spatiotemporal resolution, accuracy, precision, etc.), better feature selection, or both. Some studies, for example, used PPI systems with considerably less spatial resolution [12, 36, 37] (~35 mm). Others used relatively low-dimensional features like ~10-dimensional region of interest pressures [38] and ~100-dimensional centre of pressure (COP) trajectories [11,15,33,39-41]; this is contrasted with the current ~8000-dimensional pre-features. Thus compression of PPI data, either by sensor resolution or by lossy data reduction, likely sacrifices identification-relevant features. Automated dimensionality reduction, used also in previous investigations of biomechanical (kinematic) data [42,43], thus appears to be a more robust data compression tool.

4.3 Spatial alignment

Registration presently outperformed decorrelation over all tested parameter combinations, yielding CR improvements on the order of 10% despite moderately high pre-registration CRs of 85% or more. Registration's successes are somewhat unsurprising because registration's explicit goal is to minimise a dissimilarity metric which, by definition, reduces intra-subject variability. Its successes are also consistent with previous reports that a variety of registration approaches both qualitatively and quantitatively out-perform decorrelation [10].

It was more surprising that decorrelation performed worse than no spatial alignment in many cases. This can be partially explained by stereotypical foot postures adopted by subjects – particularly the angle of the foot's longitudinal axis with respect to progression direction [44]. Decorrelation removes this information because the foot becomes rotated to a 'vertical' posture. While registration to an arbitrary template would also remove some of this stereotypical posture information, registration achieves better intra-subject alignment [10], so postural information likely becomes less relevant once better alignment is achieved. Rather than registering to an arbitrary template, as was done currently, it would be interesting to test a registration scheme that iteratively registers a given PPI to a mean database image for each subject. This was not done currently because improvements would not be noticeable beyond the present CRs of 99.6%.

As an aside we note that many previous PPI-based identification studies used decorrelation for spatial alignment

[11,15,40,45]. Despite its prevalence in previous papers, the current results strongly suggest that decorrelation is a poor alignment choice. While we have speculated on potential mechanisms for decorrelation's poor performance (i.e. loss of stereotypical foot posture) it would be interesting to directly test this assertion by incorporating initial posture as an additional feature in a decorrelated dataset. However, since we had no reason to expect decorrelation's poor performance prior to the present results, we leave this hypothesis for future work.

We wish to emphasise that we do not believe that the current registration scheme [23] was particularly special in terms of generating higher CR; there are a plethora of registration algorithms in the literature [22], and indeed a variety of methods have been shown to yield similar results in plantar pressure data [10]. Furthermore, in *post hoc* analyses we employed a completely different registration scheme [10] and achieved similarly high, albeit slightly lower CRs of ~97.5%. The current algorithm was selected simply because it was fast and has worked well recently. To rule out a particular registration scheme as a limitation it would be prudent to evaluate other algorithms in future work.

4.4 Feature extraction

The best-performing single pre-features were P100, P90, P80 and PTI (Table 5), and the best pre-feature combination of P100,P80 only marginally improved ultimate CRs (Table 6). This gives anecdotal credence to the extensive use of P100 and PTI in the literature [8] as information-dense parameters. To our knowledge P90 and P80 have not been previously examined. One explanation for the success of the P100,P80 combination is that this essentially represents a dynamic gradient, albeit a low-frequency one, and that this low-feature gradient also contains subject-specific information. However it does not explain why P100&P80 was better than P100&P90. Regardless, since the performances of the P80, P90, and P100 pre-features were all quite high, a systematic exploration of their differences would not be possible without more data.

Moreso than particular pre-feature selections, and with the exception of KPCA, dimensionality reduction was found to invariably improve CR (Table 6), albeit to a smaller extent than registration. While the CR improvement was small it was nontrivial, pushing the average CR beyond what was achievable with raw-spatially aligned pre-features. We may thus conclude that while certain pre-features perform very well, only with dimensionality reduction can optimum CR be achieved. In other words, there are classification-relevant patterns in the pre-features that cannot be extracted in an *a priori* manner.

As an aside we note that the present percentile pre-features (P90, P80, …) were computed over all time frames (Fig.1c), and are therefore dependent on both the duration of supra-zero pressure and the recording duration (1 s). In *post hoc* analysis we also computed percentiles over contact duration, but we found little qualitative effect on the ultimate results: P100 was the best-performing percentile, and CR systematically reduced with percentile (Table 5).

We also wish to restate that we presently did not conduct temporal normalisation (aside from heel-strike alignment). This was done deliberately, to give time-related features like contact duration (CD) and time-to-max (Tmax) the maximum chance find temporal differences amongst subjects; if there were indeed significant temporal differences amongst subjects these features would be expected to yield higher CRs than if the data were temporally normalised. However, the fact that CD and Tmax performed relatively poorly (Table 5) suggests that inter-subject temporal differences were not as important as the pressure-related differences.

Finally, the present pre-feature list was incomplete. All 2D (x,y) pre-features the data were derived from the original 3D (x,y,time) image, but additional variables could have been analysed like the spatial pressure gradient [46] and the spatiotemporal (x,time) 100th percentile [31]. It may be informative to investigate such variables in future work.

4.5 Limitations

A major practical limitation of the current study is that we investigated only unshod walking. It is conceivable that shod walking considerably distorts classification-relevant pressure patterns and/or that subjects are not recognisable if they wear different shoes. A second key limitation of this study is that only natural self-paced walking data were collected; PPI data are known to change with walking speed [47], fatigue [48], and a variety of other factors [8], and we note that some previous PPI-based identification studies have indeed incorporated some of these factors in experimental classification tests [11].

Walking speed, in particular, would be interesting to consider; although general foot morphology does not change with speed, and thus binary features (Section 3.4) should be largely unaffected, the non-trivial pressure redistributions associated with walking speed [47] would likely affect subject separability, and it would be prudent to empirically define the walking speed limits that retain separability. However, since we can easily measure walking speed using cameras, and/or using foot-contact duration as a proxy, it may be possible to algorithmically compensate for walking speed variability, for example by introducing temporal normalisation, or by scaling pressures in certain foot regions.

Although PPI data can easily change, many gait-recognition applications involve desired identification, situations in which an individual *wants* to be identified (e.g. automated airport immigration control). For other applications it may be necessary to test the current algorithms on experimentally manipulated gait. Finally, we presently considered only particular testing/training ratios in our model assessment. It would be prudent to systematically explore testing/training ratios, with more images for each subject, to find the optimum number of images one should obtain if implementing a real-world plantar pressure-based identification scheme.

4.6 Conclusion

Normal self-paced unshod walking produced a high-quality plantar pressure-derived biometric, and the present identification implementation yielded classification rates of 99.6% in N=104 individuals. These results were largely driven by spatial image registration and, to enable finer subject differentiation, automated dimensionality reduction. Since plantar pressure data are highly unique amongst individuals, and since data can be easily collected and processed using commercial in-floor hardware, plantar pressure-based identification appears to have strong potential for a variety of security and health applications.

REFERENCES

1   Blanc, Y., Balmer, C., Landis, T., & Vingerhoets, F. 1999 Temporal parameters and patterns of the foot roll over during walking: normative data for healthy adults. *Gait Posture* **10**(2), 97-108.

2   Cutting, J. E. & Kozlowski, L. T. 1977 Recognizing friends by their walk: gait perception without familiarity cues. *Bull Psychon Soc* **9**(5), 353-356.

3   Nixon, M. S. & Carter, J. N. 2006 Automatic recognition by gait. *Proc IEEE* **94**, 2013-2024.

4   Stevenage, S. V., Nixon, M. S. & Vince, K. 1999 Visual analysis of gait as a cue to identity. *Appl Cogn Psychol* **13**(6), 513-526.

5   Yao, Z.-M., Zhou, X., Lin, E.-D., Xu, S. & Sun Y.-N. 2010 A novel biometric recognition system based on ground reaction force measurements of continuous gait. In *Proceedings of the IEEE International Conference on Human System Interactions*, pp. 452-458.

6   Moustakidis, S. P., Theocharis, J. B. & Giakas, G. 2009 Feature extraction based on a fuzzy complementary criterion for gait recognition using GRF signals. In *Proceedings of the Mediterranean Conference on Control and Automation*, pp. 1456-1461

7   Goffredo, M., Bouchrika, I., Carter, J.N. & Nixon, M.S. 2010 Self-calibrating view-invariant gait biometrics. *IEEE Trans Syst Man Cybern B Cybern* **40**, 997-1008.

8   Rosenbaum, D. & Becker, H. P. 1997 Plantar pressure distribution measurements: technical background and clinical applications. *Foot Ankle Surg* 3(1), 1-14.

9   Facey, O. E., Hannah, I. D. & Rosen, D. 1993 Analysis of the reproducibility and individuality of dynamic pedobarograph images. *J Med Eng Technol* **17**(1), 9-15.

10 Pataky, T. C., Goulermas, J. Y. & Crompton, R. H. 2008 A comparison of seven methods of within-subjects rigid-body pedobarographic image registration. *J Biomech* **41**(14), 3085-3089.

11 Jung, J.-W., Bien, Z. & Sato, T. 2004 Person recognition method using sequential walking footprints via overlapped foot shape and center-of-pressure trajectory. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences* **E87-A**(6), 1393-1400.

12 Yun, J., Woo, W. & Ryu, J. 2005 User identification using user's walking pattern over the ubiFloorII. In *Lecture Notes in Computer Science: Computational Intelligence and Security* **3801**, 949-956.

13 Yamakawa, T., Taniguchi, K., Asari, K., Kobashi, S. & Hata, Y. 2008 Biometric personal identification based on gait pattern using both feet pressure change. In *Proceedings of the IEEE World Automation Congress*.

14 Qian, G., Zhang, J. & Kidane, A. 2010 People identification using floor pressure sensing and analysis. *IEEE Sens J* **10**(9), 1447-1460.

15 Takeda, T., Taniguchi, K., Asari, K., Kuramoto, K., Kobashi, S. & Hata, Y. 2009 Biometric personal authentication by one step foot pressure distribution change by load distribution sensor. In *Proceedings of the IEEE International Conference on Fuzzy Systems*, pp. 906-910.

16 Uhl, A. & Wild, P. 2008 Footprint-based biometric verification. *J Electron Imaging* **17**(1), 011016.

17 Hild, M. 2008 Person identification based on barefoot 3D sole shape. In *Lecture Notes in Computer Science: Computational Forensics* (ed. Srihari, S. N. & Franke, K.) **5158**, 84-95.

18 Harrison, A. J. & Hillard, P. J. 2000 A moment-based technique for the automatic spatial alignment of plantar pressure data. *Proceedings of the Institute of Mechanical Engineers, Part H: Journal of Engineering in Medicine* **214**(3), 257–264.

19 Pataky, T. C., Bosch, K., Mu, T., Keijsers, N. L. W., Segers, V., Rosenbaum, D. & Goulermas, J. Y. 2011 An anatomically unbiased foot template for inter-subject plantar pressure evaluation. *Gait Posture*. **33**(3): 418-422.

20 Sadeghi, H., Allard, P., Prince, F. & Labelle, H. 2000 Symmetry and limb dominance in able-bodied gait: a review. *Gait Posture* **12**, 34-45.

21 Herzog, W., Nigg, B. M., Read, L. J. & Olsson, E. 1989 Asymmetries in ground reaction force patterns in normal human gait. *Med Sci Sports Exerc* **21**(1), 110-114.

22 Maintz, J. B. A. & Viergever, M. A. 1998 A survey of medical image registration. *Med Image Anal* **2**(1), 1-37.

23 Oliveira, F. P. M., Pataky, T. C. & Tavares, J. M. R. S. 2010 Registration of pedobarographic image data in frequency domain. *Computer Methods in Biomechanics and Biomedical Engineering* **13**(6), 731-740.

24 Raudys, S. & Pikelis, V. 1980 On dimensionality, sample size, classification error and complexity of classification agorithm in pattern recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2**(3), 242-252.

25 De Cock, A., De Clercq, D., Willems, T. & Witvrouw, E. 2005 Temporal characteristics of foot roll-over during barefoot jogging: reference data for young adults. *Gait Posture* **21**(4), 432-439.

26 Belkin, M. & Niyogi, P. 2002 Laplacian eigenmaps and spectral techniques for embedding and clustering. In *Advances in neural information processing systems vol. 14*, (ed. T. G. Dietterich, S. Becker, & Z. Ghahramani), pp. 585-591. Cambridge, MA: MIT Press.

27 Schölkopf, B., Smola, A. J. & Müller, K. R. 1998 Nonlinear component analysis as a kernel eigenvalue problem. *Neural Comput* **10**(5), 1299-1319.

28 Roweis, S. T. & Saul, L. K. 2000 Nonlinear dimensionality reduction by locally linear embedding. *Science* **290**, 2323-2326.

29 Gurney, J. K., Kersting, U. G. & Rosenbaum, D. 2008 Between-day reliability of repeated plantar pressure distribution measurements in a normal population. *Gait Posture* **27**(4), 706-709.

30 De Cock, A., Willems, T., Witvrouw E., Vanrenterghem J. & De Clercq, D. 2006 A functional foot type classification with cluster analysis based on plantar pressure distribution during jogging. *Gait Posture* **23**(3), 339-347.

31 Jeon, H.-S., Han, J., Yi, W.-J., Jeon, B. S. & Park, K. W. 2008 Classification of Parkinson gait and normal gait using spatial-temporal image of plantar pressure. In *Proceedings of the IEEE Engineering in Medicine and Biology Society Conference*, pp. 4672-4675.

32 Sazonov, E. S., Bumpus, T., Zeigler, S. & Marocco, S. 2005 Classification of plantar pressure and heel acceleration patterns using neural networks. In *Proceedings of the IEEE International Joint Conference on Neural Networks*, pp. 3007-3010.

33 Jung, J.-W., Bien, Z., Lee, S.-W. & Sato, T. 2003 Dynamic footprint based person identification using mat-type pressure sensor. In *Proceedings of the IEEE Engineering in Medicine and Biology Society Conference*, pp. 2937- 2940.

34 Xiong, S. Goonetilleke, R. S., Witana, C. P. & Lee Au Y. 2008. Modelling foot height and foot shape-related dimensions. *Ergonomics* 51(8), 1272-1289.

35 Ostrosky, K. M., Van Swearingen, J. M., Burdett, R. G. & Gee, Z. 1994. A comparison of gait characteristics in young and old subjects. *Physical Therapy* **74**(7), 637-644.

36   Yun, J., Abowd, G., Woo, W. & Ryu, J. 2007 Biometric user identification with dynamic footprint. In *Proceedings of the IEEE International Conference on Bio-Inspired Computing: Theories and Applications*, pp. 225-230.

37   Yun, J., Abowd, G. D. & Ryu, J. 2008 User identification with user's stepping pattern over the UbiFloorII. *Int J Pattern Recognit* **22**(3), 497-514.

38   Yang, Y.-T., Lin, Y.-C. & Yang, B.-S. 2010 Gait recognition by features of plantar pressure. In *Proceedings of the World Congress of Biomechanics*, 243.

39   Jung, J.-W., Sato, T. & Bien, Z. 2002 Unconstrained person recognition method using dynamic footprint. In *Proceedings of the 18th Hungarian-Korean Seminar*, pp. 129-137.

40   Jung, J.-W., Sato, T. & Bien, Z. 2004 Dynamic footprint based person recognition using a hidden Markov model and a neural network. *Int J Intell Syst* **19**(11), 1127-1141.

41   Qian, G., Zhang, J. & Kidane, A. 2008 People identification using gait via floor pressure sensing and analysis. In *Lecture Notes in Computer Science: Smart Sensing and Context* (ed. Roggen, D., Lombriser, C., Tröster, G., Kortuem, G. & Havinga, P.) **5279**, 83-98.

42   Goulermas, J. Y., Findlow, A. H., Nester, C. J., Howard, D. & Bowker, P. 2005 Automated design of robust discriminant analysis classifier for foot pressure lesions using kinematic data. *IEEE Trans Biomed Eng* **52**(9), 1549-1562.

43   Mu, T., Pataky, T. C., Findlow, A. H., Aung, M. S. H. & Goulermas, J. Y. 2010 Automated nonlinear feature generation and classification of foot pressure lesions. *IEEE Trans Inf Technol Biomed* **14**(2), 418-424.

44   Taranto, J., Taranto, M. J., Bryant, A. & Singer, K. P. 2005 Angle of gait: a comparative reliability study using footprints and the EMED-SF®. *The Foot* 15(1), 7-13.

45   Nakajima, K., Mizukami, Y., Tanaka, K. & Tamura, T. 2000 Footprint-based personal recognition. *IEEE Trans Biomed Eng* **47**(11), 1534-1537.

46   Prabhu, K. G., Patil, K. M. & Srinivasan, S. 2001 Diabetic feet at risk: a new method of analysis of walking foot pressure images at different levels of neuropathy for early detection of plantar ulcers. *Med Biol Eng Comput* **39**(3), 288-293.

47   Rosenbaum, D., Hautmann, S., Gold, M. & Claes, L. 1994 Effects of walking speed on plantar pressure patterns and hindfoot angular motion. *Gait Posture* **2**(3), 191-197.

48   Stolwijk, N. M., Duysens, J., Louwerens, J. W. K. & Keijsers, N. L. W. 2010 Plantar pressure changes after long distance walking. *Med Science in Sports Exerc* **42**(12), 2264-2272.

TABLES

Table 1. Subject characteristics.

	Female	Male	Follow-up
N	64	40	10
Age (years)	30.0 (10.8)	34.4 (8.6)	36.7 (8.0)
Mass (kg)	63.2 (8.2)	80.6 (11.1)	67.6 (14.0)
Height (cm)	169.4 (6.3)	182.6 (7.1)	175.3 (9.7)
BMI (kg/cm^2)	22.0 (2.4)	24.1 (2.2)	21.9 (2.7)
Sport (hours/week)	3.2 [a] (2.2)	4.2 (3.1)	3.3 (1.8)

(Averages, with st.dev. in parentheses. 'Follow-up' data included five females and five males; main dataset: Spring 2009, two follow-up subjects: 1.5 years later, eight follow-up subjects: 1.5-5.0 years before. [a]Data for two female subjects were unavailable.)

Table 2. Spatial alignment methods.

#	Method	Description	Ref
1	None	No pre-processing (only image cropping and zero-padding)	×
2	Decorr	Decorrelation (principal axis transformation)	[18]
3	RegMunCont	Register to (flipped) Münster mean	[9,21]
4	RegIndepMean	Register to independent ipsilateral mean	"
5	RegIndep1	Register to independent arbitrary subject 1	"
6	RegIndep2	Register to independent arbitrary subject 2	"
7	RegIndep3	Register to independent arbitrary subject 3	"
8	RegIndep4	Register to independent arbitrary subject 4	"
9	RegIndep5	Register to independent arbitrary subject 5	"

Table 3. Pre-feature descriptions.

#	Pre-feature	Units	Description	Definition
1	P100	kPa	100th percentile (spatial maximum)	$Pp(x,y) \equiv I(k,y,k) + \dfrac{N}{100}(p - p_k)[I(x,y,k+1) - I(x,y,k)]$ where $p_k \leq p \leq p_{k+1}$
2	P90	kPa	90th percentile	"
3	P80	kPa	80th percentile	"
4	P70	kPa	70th percentile	"
5	P60	kPa	60th percentile	"
6	P50	kPa	50th percentile	"
7	PTI	kPa·s	Pressure-time integral	$PTI(x,y) = \sum_t I(x,y,t)$
8	Tfirst	s	Time to first contact (from heel strike)	$T_{first}(x,y) = \min_{I(x,y,t)>\varepsilon} t$
9	Tmax	s	Time to maximum (from heel strike)	$T_{max}(x,y) = \arg\max_t I(x,y,t)$
10	CD	s	Contact duration	$CD(x,y) = \sum_t f(x,y,t)$ where $f(x,y,t) = \begin{cases} 1 & \text{if} \quad I(x,y,t) > \varepsilon \\ 0 & \text{otherwise} \end{cases}$

Here $I(x,y,t)$ is the image time series, p denotes percentile, k indexes the ordered observations of a particular pixel's time series, p_k is the percentile of the kth ranked observation, N is the number of observations, and ε is a pressure threshold (manufacturer-set to $\varepsilon =5$ kPa in the current dataset)). Note: in the percentile equation k is not a time index, but rather indexes sorted observations, and k may be different for each pixel's time series.

Table 4. Dimensionality reduction methods.

#	Method	Description	Ref
1	None	No reduction	×
2	LE	Laplacian eigenmaps	[24]
3	NCSL	Normalised spectral clustering with symmetric Laplacian	[24]
4	KPCA	Kernel principal component analysis	[25]
5	LLE	Locally linear embedding	[26]

Table 5. Semi-factorial analysis: alignment and single pre-features, left foot.

Alignment	Pre-feature									
	P100	P90	P80	P70	P60	P50	PTI	Tfirst	Tmax	CD
None	94.6	94.8	90.6	85.2	72.9	9.8	94.0	88.5	81.0	93.7
Decorr	93.1	93.7	91.0	85.4	73.3	2.1	92.9	79.0	66.9	88.3
RegMunCont	98.8	98.8	98.5	96.0	86.9	2.1	98.1	89.6	79.0	94.0
RegIndepMean	92.9	93.3	93.7	90.8	79.4	1.5	92.9	81.0	71.2	89.6
RegIndep1	99.0	99.2	99.2	97.5	87.1	1.0	99.6	87.7	77.5	95.8
RegIndep2	96.2	96.2	95.4	93.1	78.1	1.3	96.0	85.0	74.0	91.2
RegIndep3	93.5	94.2	94.2	91.0	82.7	1.3	94.2	85.6	72.7	91.0
RegIndep4	96.7	96.9	95.4	93.1	86.9	3.8	96.7	87.3	73.3	94.0
RegIndep5	95.8	95.8	96.0	92.5	86.0	2.9	96.0	89.4	77.9	93.8

(Data are classification rates, %. Data reduction: LLE. The alignment methods and features yielding CR>90% are highlighted in gray.)

Table 6. Semi-factorial analysis: pre-processing and dimensionality reduction methods.

(a) 104-fold cross-validation

Foot	Pre-processing	Dimensionality reduction				
		None	LE	NCSL	KPCA	LLE
Left	None	91.4	95.4	95.6	68.7	95.8
	Decorr	92.3	96.2	96.0	67.9	95.8
	RegMunCont	99.0	99.4	99.4	89.8	99.4
	RegIndepMean	94.2	94.2	94.0	88.9	94.0
Right	None	92.9	94.8	94.0	68.3	94.6
	Decorr	90.2	93.9	94.0	63.4	93.5
	RegMunCont	99.2	99.8	99.8	91.3	99.8
	RegIndepMean	95.0	93.7	93.9	87.7	94.2

(b) Leave-one-out cross-validation

Foot	Pre-processing	Dimensionality reduction				
		None	LE	NCSL	KPCA	LLE
Left	None	91.4	95.4	95.6	68.7	95.6
	Decorr	92.5	96.2	95.8	67.9	95.8
	RegMunCont	99.0	99.4	99.4	89.8	99.4
	RegIndepMean	94.2	94.4	94.0	88.9	94.2
Right	None	93.1	95.0	94.6	68.1	94.8
	Decorr	90.4	93.9	93.7	65.2	93.5
	RegMunCont	99.2	99.8	99.8	91.4	99.8
	RegIndepMean	94.2	94.4	94.4	88.9	94.2

(Data are classification rates, %. Combined pre-features: P100 and P80. The best-performing methods are highlighted.)
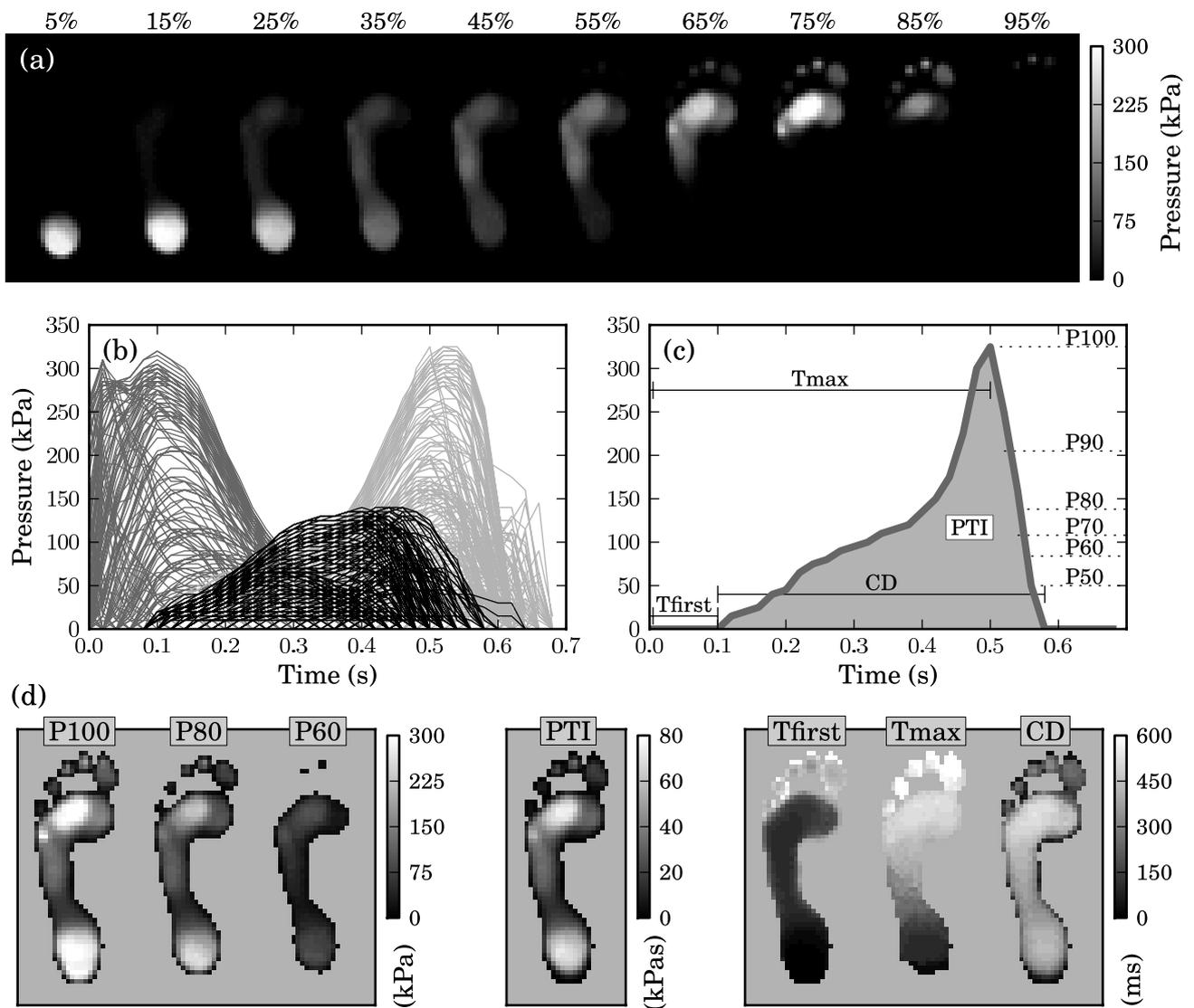
FIGURE CAPTIONS

Figure 1. Description of plantar pressure data for a single step. (a) Pressure image time series; percentages indicate normalised time (% stance). (b) Pixel time series; dark grey, black, and light grey trajectories indicate pixels whose maxima were reached in the first, second, and final thirds of stance phase, respectively. (c) Pre-features for an example pixel time series (see Table 2 for variable descriptions). (d) Pre-features, when computed across all pixels.
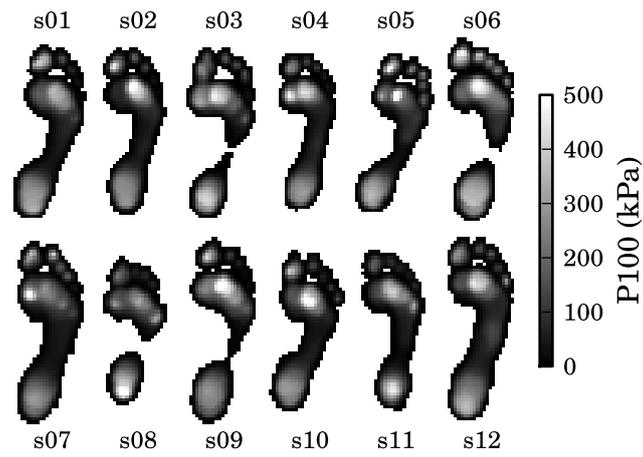
Figure 2. Maximal pressures (P100) for the chronologically first twelve subjects; averaged across five trials.

Figure 3. Spatial alignment example, first subject. Top, middle, and bottom rows depict the original, decorrelated, and registered images, respectively; here the registration template was RegMunCont (Table 2). The thick dark outline depicts the cross-trial mean.

Figure 4. Classification rate (CR) for all 104 subjects using the P100 pre-feature. See Tables 3 and 4 for alignment and dimensionality reduction and alignment method descriptions.

Figure 5. Classification rate (CR) as a function of embedding dimension (alignment: RegMunCont, pre-feature: P100, reduction: LLE).

Figure 6. 'Follow-up' test results. (a-b) Number of correctly classified images (out of five) for the left and right feet (light and dark bars, respectively) and for all ten follow-up subjects (s01…s10). Numbers in the white boxes indicate the number of years between collection of the follow-up and main datasets. (c) Left foot P100 images for subject 6, mean across five trials. The 'Target' image is from the main dataset.

Figure 1. Description of plantar pressure data for a single step. (a) Pressure image time series; percentages indicate normalized time (% stance). (b) Pixel time series; dark grey, black, and light grey trajectories indicate pixels whose maxima were reached in the first, second, and final thirds of stance phase, respectively. (c) Pre-features for an example pixel time series (see Table 2 for variable descriptions). (d) Pre-features, when computed across all pixels.

Figure 2. Maximal pressures (P100) for the chronologically first twelve subjects; averaged across five trials.
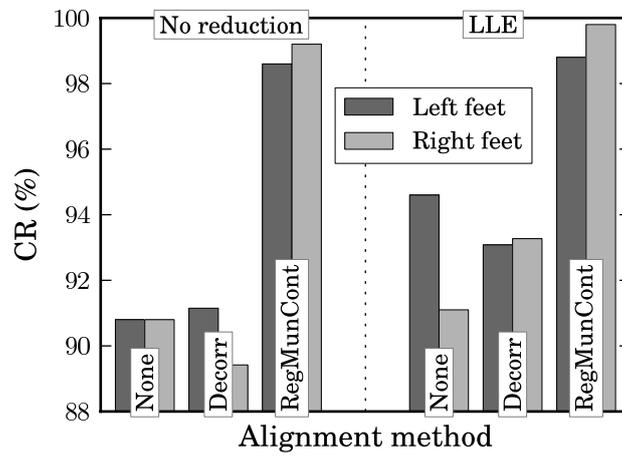
<u>Figure 3</u>. Classification rate (CR) for all 104 subjects using the P100 pre-feature. See Tables 3 and 4 for alignment and dimensionality reduction and alignment method descriptions.
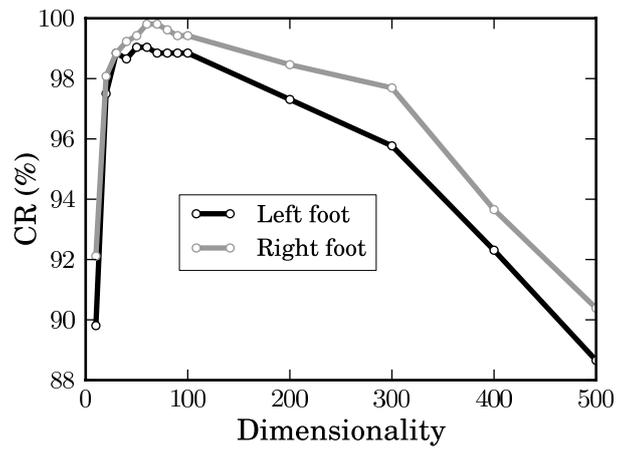
Figure 4. Classification rate (CR) as a function of embedding dimension (alignment: RegMunCont, pre-feature: P100, reduction: LLE).
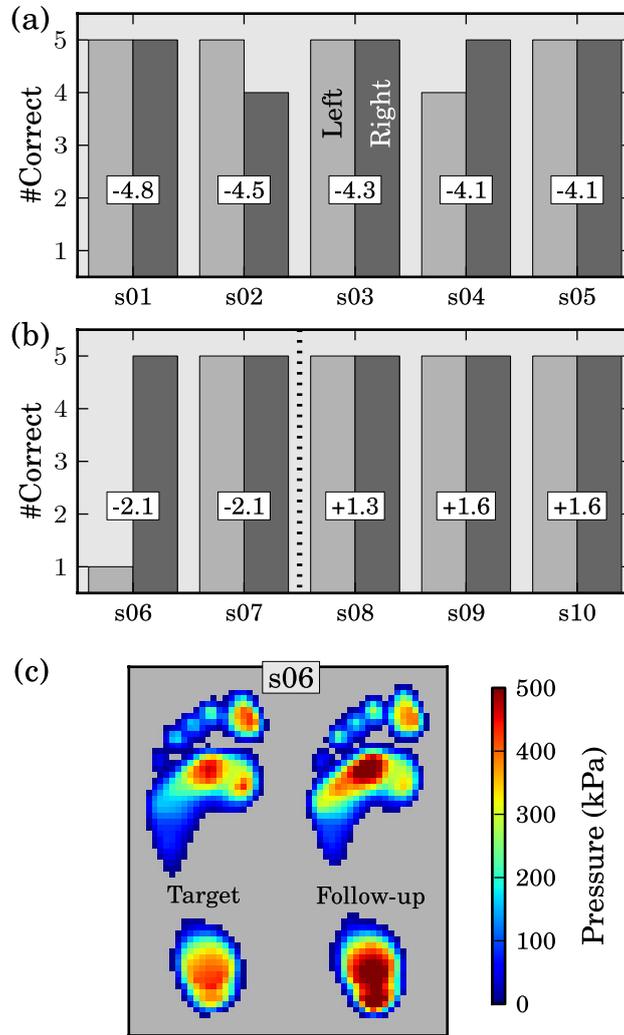
Figure 5. 'Follow-up' test results. (a-b) Number of correctly classified images (out of five) for the left and right feet (light and dark bars, respectively) and for all ten follow-up subjects (s01…s10). Numbers in the white boxes indicate the number of years between collection of the follow-up and main datasets. (c) Left foot P100 images for subject 6, mean across five trials. The 'Target' image is from the main dataset.