34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

# Real Time Gait Pattern Classification from Chest Worn Accelerometry During a Loaded Road March

Cynthia M. Clements, Mark J. Buller, Alexander P. Welles, William J. Tharion

*Abstract*— **Accelerometers, whether in smart phones or wearable physiological monitoring systems are becoming widely used to identify movement and activities of free living individuals. Although there has been much work in applying computationally intensive methods to this problem, this paper focuses on developing a real-time gait analysis approach that is intuitive, requires no individual calibration, can be extended to complex gait analysis, and can readily be adopted by ambulatory physiological monitors for use in real time. Chest-mounted tri-axial accelerometry data were collected from sixty-one male U.S. Army Ranger candidates engaged in an 8 or 12 mile loaded (35 Kg packs) timed road march. The pace of the road march was such that volunteers needed to both walk and run. To provide intuitive features we examined the periodic patterns generated from 4s periods of movement from the vertical and longitudinal accelerometer axes. Applying the "eigenfaces" face recognition approach we used Principal Components Analysis to find a single basis vector from 10% of the data (n=6) that could distinguish patterns of walk and run with a classification rate of 95% and 90% (n=55) respectively. Because these movement features are based on a gridded frequency count, the method is applicable for use by body-worn microprocessors.**

## I. INTRODUCTION

ACCELEROMETERS, whether in smart phones [1, 2] or wearable data logging systems [3, 4, 5] are becoming a foundation sensor for activity classification systems. Reviews by Godfrey et al. [6], Preece et al. [7], and Kavanagh and Menz [8] suggest current approaches tend to follow a standard formula: Segment the data, derive features (mostly frequency domain based), and use these features to develop a classification algorithm using a variety of classification approaches such as support vector machines, neural networks, and k-nearest neighbors. The derived features of these approaches are often not intuitive, require individual calibration, and require post-processing on computationally capable platforms.

With the development of recent ambulatory physiological status monitoring systems (e.g. Equivital EQ-02, Hidalgo Ltd., Cambridge UK, and BioHarness Zephyr Inc, Annapolis, MD) accelerometry is a promising additional tool in the determination of fatigue and thermal work strain in teams of emergency workers. Anecdotal evidence as early as 1877 from Company A of the 10th Cavalry describes dehydrated and heat stressed soldiers as exhibiting a "tottering gait" [9]. More recently, Goldman [10] lists unstable gait as a symptom of heat exhaustion. Patterns of acceleration associated with cyclical trunk movement during walking have been used to create indices of gait smoothness [11] and to distinguish between elderly individuals with and without gait stability problems by using trunk-mounted accelerometers [12].

Examining the temporal relationships between acceleration channels, distinct periodic patterns can be seen. Figure 1 shows the vertical axis and longitudinal axis of a trunk mounted tri-axial accelerometer plotted over a 4 second time frame during both walking and running.

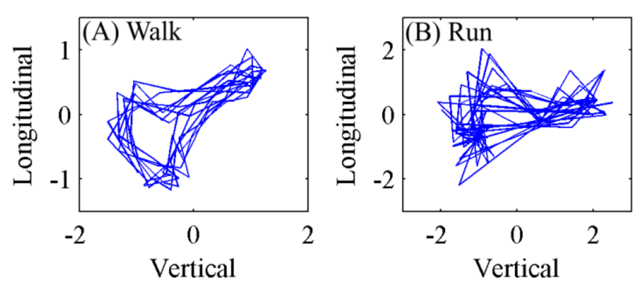


Fig. 1. Patterns of movement: vertical vs. longitudinal accelerations from a tri-axial chest mounted accelerometer shown over a 4s time frame during periods of **(A)** walking and **(B)** running.

These plots suggest that there is a distinct difference in gait between these two activities. Turk and Pentland [13, 14] demonstrated that complex images of faces could be broken down by the use of Principal Components Analysis (PCA) into components representing significant facial

C. Clements, is with the U.S. Army Research Institute of Environmental Medicine, Natick, MA 01760 USA. 508-233-6274, Cynthia.clements@us.army.mil.

M. Buller is with the U.S. Army Research Institute of Environmental Medicine, Natick, MA 01760 USA and the Computer Science Department at Brown University Providence RI.

A. Welles, is with the U.S. Army Research Institute of Environmental Medicine, Natick, MA 01760 USA.

W. Tharion, is with the U.S. Army Research Institute of Environmental Medicine, Natick, MA 01760 USA.

The opinions or assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the Army or the Department of Defense. The investigators have adhered to the policies for protection of human subjects as prescribed in Army Regulation 70-25, and the research was conducted in adherence with the provisions of 32 CFR Part 219. Citations of commercial organizations and trade names in this report do not constitute an official Department of the Army endorsement or approval of the products or services of these organizations.

 

characteristics that can be used for facial recognition. Similarly, our goal was to demonstrate that accelerometry signatures could be broken down into components that represent major characteristics of gait such as walk-run, load carriage, and wobble as examples. The specific aim of this paper was to demonstrate that this approach can break movement signatures into major components, the first of which can distinguish between walk and run. Our motivation for this paper was to develop a real-time gait analysis method that is intuitive, requires no individual calibration, can be extended to complex gait analysis, and can readily be adopted by ambulatory physiological monitors for use in real time.

## II. METHODS

### A. Subjects

Participants were 61 students enrolled in the Ranger Training Brigade's (RTB's) qualifying course. The two month course includes various qualifying events, one of which is a timed road march held on the 4$^{th}$ day. Volunteers were required to complete an 8 mile road march in 2 hours 10 minutes (Summer) or a 12 mile road march in 3 hours 15 minutes (Spring and Winter) to remain in the RTB qualifying course. Descriptive statistics are means $\pm$ standard deviations. Volunteers were relatively homogenous (Age: 25 $\pm$ 4 yrs; Ht: 175 $\pm$ 5 cm; Wt: 69 $\pm$ 8 kg; body fat 15 $\pm$ 4%). During the road march they carried 17.8 $\pm$ 1.8 kg.

### B. Equipment

Acceleration data were collected at 25.6 Hz using the Equivital EQ-01 (Hidalgo, Cambridge UK) physiological status monitor centered on the chest using a 3-axis micro electro-mechanical systems (MEMS) based accelerometer ± 3 g. Global Positioning System (GPS) data were collected on Fortrex 101 wrist watches (Garmin, Olathe, KS).

### C. Event

The required average pace of 16 min/mile was such that for most volunteers this represented a pace between "comfortable" walking and "comfortable" running. By "comfortable" we mean a speed of walk or run that during unloaded movement would be the preferred walk or run speed. Walking or running at a "non-preferred" pace can be less efficient [16, 17]. Instead of walking at this pace most volunteers chose to alternate between a "comfortable" walking pace and a "comfortable" jog or run. Figure 2 shows an example of a typical histogram of movement speeds from one volunteer for the entire event. Time to complete the road march was 3:03:11 $\pm$ 0: 12:27 (hr:min:sec), and 3:07:36 $\pm$ 0:06:28 for the 12-mile Winter and Spring marches respectively, and 1:59:19 $\pm$ 0:07:45 for the 8-mile Summer march. These road movements were very physically demanding given load, and the required pace.

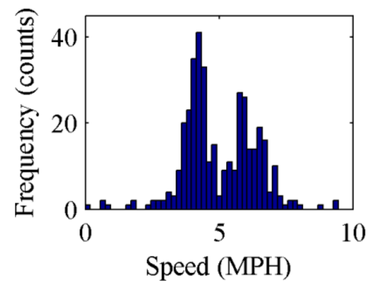


Fig. 2. Histogram of movement speeds for one subject obtained from the GPS, showing the typical bi-model distribution between walking and running speeds.

### D. Walk-Run Ground Truth

GPS location data were available for approximately 50% of the participants, and were collected every 15s. Since we wanted to use as many participants as possible and the 15s GPS data did not always accurately label the 4s periods, we chose to use a vertical acceleration threshold to identify each movement. In order to do this we determined a participant's individual vertical acceleration threshold based on known periods of walk and run identified by the GPS. A segment was classified as a run if this threshold was exceeded repeatedly. For a subset of 13 participants, we demonstrated that this threshold approach matched both the 1 minute GPS speeds and a frequency power content threshold derived for each subject for walk and run from a Fast Fourier Transform (FFT) of the vertical accelerations data. Figure 3, shows a spectrogram containing data for the entire road march. Red shows frequencies containing more power. The figure clearly shows the power differences during periods of walk (~1Hz) and run (~1.5Hz).

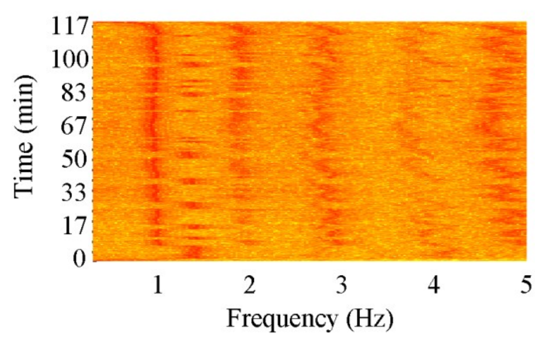


Fig. 3. Spectrogram from FFT of vertical axis data for one participant. Walks have high power at 1Hz and runs high power at 1.5Hz. Densities above 1.5Hz appear to reflect walk-run harmonics. Red shows more power, yellow less.

### E. Preprocessing

The vertical and longitudinal accelerations were chosen for this analysis as these components provided the greatest amount of variability during the loaded walk and run segments. The data were segmented into periods of 4s as a way to capture approximately two full step cycles. Each 4s segment was normalized to zero mean (by subtracting the mean of every 4s period) and unit variance. To avoid principal components being found based upon the phase of step accelerations, data from the vertical and longitudinal axes were binned into a 10 x 10 matrix. Our movement signature was then comprised of the frequency counts within each bin. Figure 4 shows an intensity plot of these frequency counts derived from the acceleration signals *shown in the examples from Figure 1, with blue showing 0* counts to red showing the most counts.

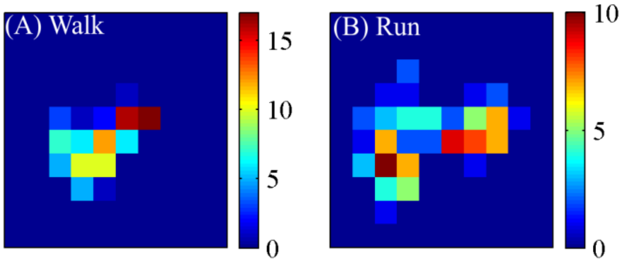


Fig. 4. Patterns of movement in normalized gridded form. Intensity plot of frequency counts of binned accelerometry data. Vertical vs. longitudinal accelerations from a tri-axial chest mounted accelerometer shown over a 4s time frame during periods of (A) walking and (B) running.

### *F. PCA and Likelihood Model*

To train our model we chose data from two volunteers at random from each of our three data collections (N=6, 10% of our population). The remaining participants (N=55, 90%) were used as a hold-out sample for validation. Ten-fold cross validation was used to ensure that our random selection of training participants did not unduly bias our model either positively or negatively. An equal number of run and walk segments were chosen at random from each training subject. Each movement segment (10x10 matrix) was flattened to a column vector (size = 100) and concatenated into a 100 x N (training segments) matrix **X**. Our goal was to develop a lower dimensional representation of each movement segment and from this develop a classification model for walk and run based upon the assumption that each class can be represented by a multivariate Gaussian probability density function.

To reduce the number of dimensions in our data we used Principal Component Analysis (PCA) [e.g. 18]. PCA determines a set of orthogonal basis vectors that allow us to represent the gait signature in a lower dimensional space. PCA solves the eignevalue problem:

$$\mathbf{\Lambda} = \mathbf{\Phi}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{\Phi} \qquad (1)$$

Where $\mathbf{\Lambda}$ is a diagonal matrix of eigenvalues, $\mathbf{\Phi}$ is the eigenvector matrix of training data ($\mathbf{X}$) covariance matrix $\mathbf{\Sigma}$. The eigenvector basis vectors are ordered such that the first basis vector accounts for the most variance in the data and so on. To represent a 4s signature in lower dimensions the original signature is mapped to low dimensions using a subset of the basis vectors determined by PCA thus:

$$\mathbf{y} = \boldsymbol{\phi}_M^T\tilde{\mathbf{x}} \qquad (2)$$

Where the vector $\mathbf{y}$ is the representation of the movement signature in low dimensional space; $\tilde{\mathbf{x}} = \mathbf{x} - \bar{\mathbf{x}}$, where $\mathbf{x}$ is a movement segment column vector and $\bar{\mathbf{x}}$ is the mean of all the training data; and $\boldsymbol{\phi}_M$ is a sub-matrix of the eigenvector matrix.

We assume that both walk and run classes in the low dimensional space can be represented by multivariate Gaussian probability density functions where the mean and covariance are learned from the training data mapped into low dimensional space. Thus, the likelihood that a given movement signature is in class Ω is given by:

$$P(\mathbf{y}|\Omega) = \frac{exp\left[-\frac{1}{2}(\mathbf{y}-\bar{\mathbf{y}}_\Omega)\mathbf{\Sigma}_\Omega^{-1}(\mathbf{y}-\bar{\mathbf{y}}_\Omega)\right]}{(2\pi)^{\frac{N}{2}}|\mathbf{\Sigma}_\Omega|^{\frac{1}{2}}} \qquad (3)$$

Movement signatures are classified based on the highest probability of class membership.

## III. Results

Figure 5 shows the mean run and walk gridded movement signatures from the training data, along with the first four basis vectors found from PCA. The first four basis vectors account for 33, 51, 62 and 70% of the cumulative variance respectively.

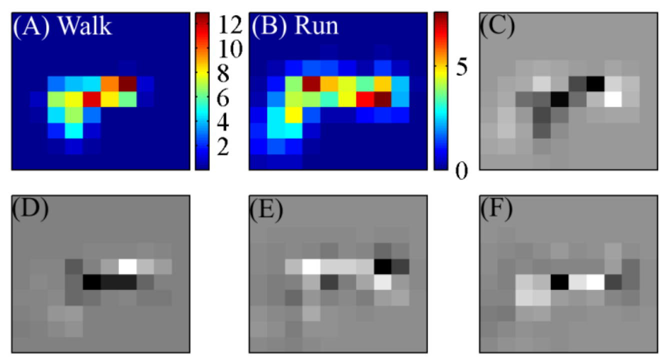


Fig. 5. Mean walk (A), mean run (B) gridded movement segments, and the first four basis vectors (C – F) from PCA shown as temperature plots.

Table 1 shows the classification rate for walks and runs using 1 to 4 basis vectors.

TABLE 1
CLASSIFICATION RATE

| # Basis. | Correct Classification (%) Walk (N=122303) | Run (N=8274) |
|---|---|---|
| 1. | 95.33 | 89.90 |
| 2. | 96.45 | 88.42 |
| 3. | 96.42 | 88.45 |
| 4. | 97.52 | 82.26 |

Using more than one basis for the classification model provides little gain over using a single basis vector which correctly classifies over 95% of walks and 90% of the runs. This suggests that the first principal component separates walk and run periods while the other principal components likely capture other features of movement. Classification errors occur across all participants and are not disproportionally due to any one group or volunteer.

Median classification rates for walk and run across the other cross validation models were 93 and 95% respectively. The mean likelihood for correct classifications (0.156) is significantly different from incorrect classifications (0.00049), $p<0.05$.

## IV. Discussion

Our approach has demonstrated that a walk-run classification can be learned from a small number of participants (N=6) and generalized to a much larger number of participants (N=55) without the need for any calibration.

The current model can only classify a movement segment as walk or run, and is forced to identify each movement as one of these two choices. For the majority of our misclassifications the likelihood of being either a walk or run is low. In these cases we suspect that the dichotomy of walk and run is an oversimplification. During the road march there were other segments of movement such as transitions (walk to run, run to walk), rests, stumbles, and jumps. This process did not take into account these possibilities.

Using only one dimension to differentiate between walk and run, our approach is viable for adoption on microprocessors and ambulatory physiological monitors. For each 4s segment the method needs the movement segment histogram be constructed, requiring the storage of only 100 integers. This histogram is then mapped to low dimensional space by computing the dot product between the histogram (as a vector) and one pre-stored eigenvector. Individual likelihoods can be extracted from a lookup table or log-likelihoods computed instead.

The fact that only one eigenvector dimension is needed to differentiate amongst walk and run, suggests that other elements of ambulation can also be identified in other PCA eigenvectors. For example components such as sway or wobble perhaps indicating fatigue or thermal work strain would ideally be identified. Future work includes collecting data sets where fatigue and thermal work strain are incorporated in the experimental design.

## V. Conclusion

Our method is computationally very simple and can be readily used by microprocessors in ambulatory monitoring systems. We were able to extract movement signatures that intuitively show movement patterns in accelerometry data. A movement model was learned from only six subjects and then applied successfully to over 55 subjects without any individual calibration with >95% accuracy. We conclude that this approach holds promise for more complex real time gait analysis.

## References


[1] S. S. Intille, F. Albinali, S. Mota, B. Kuris, P. Botana, and W. L. Haskell. "Design of a wearable physical activity monitoring system using mobile phones and accelerometers." In Proceedings of: *33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society* 2011.

[2] H. H. Wu, E. D. Lemaire, and N. Baddour. "Change –of-state determination to recognize mobility activities using a BlackBerry smart phone." In Proceedings of: *33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society* 2011.

[3] M. Patterson, D. McGrath, and B. Caulfield. "Using a tri-axial accelerometer to detect technique breakdown dues to fatigue in distance runners: A preliminary perspective." In Proceedings of: *33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society* 2011.

[4] A. Czabke, S. Marsch, and T.C. Lueth. "Accelerometer based real-time activity analysis on a microcontroller." In Proceedings of: *33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society* 2011.

[5] A. Reiss and D. Stricker. "Introducing a modular activity monitoring system." In Proceedings of: *33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society* 2011.

[6] A. Godfrey, Conway R, Meagher D, and G. O'Laighin G. "Direct measurement of human movement by accelerometry." *Medical Engineering and Physics* vol. 30, pp 1364-1386, 2008.

[7] S. J. Preece, J. Y. Gourlermas, L. P. J. Kenney, D. Howard, K. Meijer, and R. Crompton. "Activity identification using body-mounted sensors – a review of classification techniques." *Physiological Measurement*, vol. 30, pp R1-R33, 2009.

[8] J. J. Kavanagh and B. H. Menz. "Accelerometry: A technique for quantifying movement patterns during walking". *Gait & Posture* vol. 28, pp. 1-15, 2008.

[9] J. T. King. "The days gone by." *The Military Surgeon*. Vol 74, pp 193-197, 1934.

[10] R. P. Goldman. "Introduction to heat-related problems in military operations." In: K. B. Pandolf and R. E. Burr (Eds.) *Textbooks of Military Medicine, Vol. 1*. US Army Research Institute of Environmental Medicine, Natick, MA, pp. 3-49, 2008.

[11] K. M. Culhane, M. O. Connor, and G. M. Lyons. "Accelerometers in rehabilitation medicine for older adults." *Age and Ageing*. Vol. 34, pp 556-560, 2005.

[12] H. J. Yack and R. C. Berger. "Dynamic Stability in the Elderly: Identifying a Possible Measure." *Journal of Gerontology*. Vol. 48, No. 5, pp M225-M230, 1993.

[13] M. Turk and A. Pentland, "Face recognition using eigenfaces." In Proceedings of: *IEEE Conference on Computer Vision and Pattern Recognition*, Maui, Hawaii, 1991.

[14] M. Turk and A. Pentland, "Eigenfaces for recognition." *Journal of Cognitive Neuroscience*, Vol. 3, No. 1, pp. 71-86, 1991.

[15] R.P. Mello, M. M. Murphy, and J. A. Vogel. "Relationship between a two mile run for time and maximal oxygen uptake." *Journal of Applied Sports Science Research.* Vol. 1, pp 9-12, 1988.

[16] K.G. Holt, J. Hamill, and R. O. Andres. "Predicting the Minimal Energy Costs of Human Walking." *Medicine and Science in Sport and Exercise,* Vol. 23, No. 4, pp 491-498, 1991.

[17] K. G. Holt, S. F. Jeng, R. Ratcliffe, and J. Hamill. "Energetic Cost and Stability During Human Walking at the Preferred Stride Frequency." *Journal of Motor. Behavior.* Vol. 27, No. 2, pp 164-178, 1995.

[18] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY: Springer, 2006.