Gait Analysis Techniques using Principle Component Analysis (PCA)

Kian Sek Tee[12*], Mohammed Awad[2], Abbas Dehghani[2], David Moser[3], Saeed Zehedi[3]

[1]Faculty of Electrical and Electronic Engineering (FKEE), University Tun Hussien Onn Malaysia, Parit Raja, 86400 Batu Pahat, Johor, Malaysia.
[2]School of Mechanical Engineering, University of Leeds, LS2 9JT, Leeds, United Kingdom.
[3]Chas A Blatchford & Sons Ltd., Lister Road, Basingstoke, Hampshire, RG22 4AH, United Kingdom.

*Corresponding author / email: tee@uthm.edu.my & mnkst@leeds.ac.uk

Abstract

A gait analysis procedure is derived to provide a reliable visual impression of gait performances. Kinematic information of the limb segments and body centre of mass (BCOM) are collected from three healthy subjects using an ambulatory gait monitoring system under several walking restrictions. The method includes multiple stages of gait data processing which eventually eliminates redundant or correlated variables and reduces vector dimensions. Clusters of walking performance under different walking restrictions are clearly separated in 2D and 3D plots of Principle Components (PCs). This method provides potential visual aids in decision making in many medical applications such as pathological gait analysis and lower limb prosthetic dynamic alignment.

Keywords: ambulatory, reliability, correlated, low dimension, PCA

1. INTRODUCTION

Instrumental gait analysis plays a crucial role in the study of human locomotion. Researchers [1, 2] have looked into human locomotion in the aspect of kinematic, kinetic, energy consumption or musculoskeletal system using electromyography (EMG) technique. A standard gait analysis system is confined in a room that allows limited walking distances. Commercial instruments such as camera system (Vicon) and force plate (Kistler) are amongst the popular yet expensive gait monitoring systems that provide reliable and accurate discrete gait information. Microelectromechanical system (MEMS) accelerometers and gyroscopes [3-8] provide another option for measuring body segment movements, usually as an ambulatory system. This system is much cheaper to build while providing reliable and consistent performance. Moreover, it is easily integrated with light-weight, thin film sensors [9, 10] such as force sensitive resistor (FSR) to study insole pressure or body load during dynamic movements.

One of the key procedures in gait analysis involves processing gait data using many gait variables [1] of interest. Very often, variables swell in high dimensions and size. It is hence a challenge to interpret them in an easy-to-understand form. Researchers have attempted many efforts [11, 12] including statistical methods, artificial neural network, wavelet method, fuzzy logic etc.

The authors hypothesized that since walking is a controlled sequence of falling and supporting, a gait performance is repetitive events with a statistical centre tendency. Appropriately chosen gait variables should enable the representation of this information. Objective and reliable feedbacks from the instruments are desired. These questions are asked in this paper.

1. Is the system reliable?
2. Which variables are to be chosen?
3. Which variables are correlated or redundant?
4. How to 'see' high dimension variables?

2. METHODOLOGY

2.1. AMBULATORY DATA COLLECTION DEVICE

A portable customized ambulatory system was designed to collect kinematic gait data from lower limb segments and body centre of mass (BCOM). The system includes a battery-powered datalogger, five units of IMUs and a vest. The datalogger consists of a Mbed NXP LPC1768 microprocessor, sampling at 200Hz and 32 analogue inputs being expanded using 10-Bit A/D Converters with SPI Interface (MCP3008, μchip). IMU-5DOF (ADXL330/ADXL335 and IDG300/IDG500) in PCB breakouts (SparkFun, Inc) were mounted on the body landmarks using Velcro straps. All IMUs were carefully calibrated using the method proposed by [13, 14].

2.2. EXPERIMENTAL SETUP

Three healthy subjects voluntarily took part in the trials. The trials received ethical approval from the research support unit of The University of Leeds. All subjects were required to walk under the following restrictions:

(S1). Walking normally on a flat level
(S2). Walking on a flat level with ankle-locked.
(S3). Walking on a 5^0 tilted manual treadmill.
(S4). Walking on a 5^0 tilted manual treadmill with ankle-locked.

During 'ankle-locked' walking, the subject's ankle was immobilized by putting on an ankle orthotic (Motion Walker, Physioterapystore Ltd.) locked at 90^0. Five to eight trials were carried out for each setup. Each subject was allowed to rest for a minute in between trials. All trials were repeated again in a week's time on the same subjects under the same criteria to ensure that repeatability has been taken into account.

2.3. DATA PROCESSING PROCEDURES

Gait data were processed in a multi-stage procedure as listed below:

1. Test-retest reliability
   Test-retest reliability [5, 6] checks the reliability of an instrument over time. All criteria are kept the same but repeated over a week. If an index called Cronbach's Alpha (CA) is above 0.7, the instrument is reliable and consistent.
2. Gait identification
   Critical gait events (Figure 1) such as Heel-Contact (HC), Toe-Off (TO) and Mid-Swing (MSW) are identified from gyroscope lateral axis [4, 7, 8, 10, 15] located at shanks.

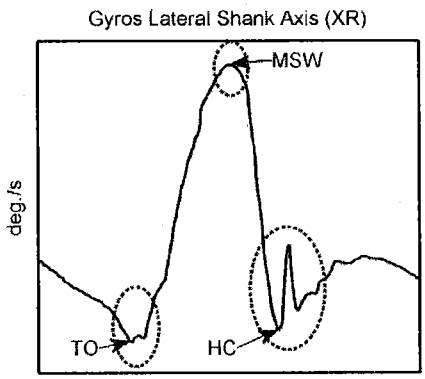

Figure 1: Gait Events for Identification (TO = Toe-Off, MSW = Mid-Swing, HC = Heel Contact)

3. Low Pass filter
   A zero phase low pass filter, f_c = 3Hz, smoothes out all IMUs outputs at the same time base as before.
4. Gait Cycle (GC) extraction
   GC [16] is defined as the period between two HC. GCs are extracted individually from time-series data using gait cycle indexes (GCI) generated from gait identification.
5. GC normalization
   For the sake of comparison, all time axes of extracted GCs are normalized from zero to one. However, they are not in same vector length.
6. Linear interpolation of normalized GC (LiNo GC)
   Normalized GCs are remapped into equal vector length using linear interpolation. Their time-axes are in the same range, same length and intervals as shown in Figure 2.
7. LiNo GC stacks and data structure
   LiNo GC are stacked in a structure (Figure 3) that branches into names using locations of IMUs with respect to the categories of walking restrictions. New structures are created according to subjects and weeks of trials taken. In this paper, six structures were created, i.e. 3 subjects x 2 weeks. At this stage, LiNo GC stacks are useful for many computational analysis.

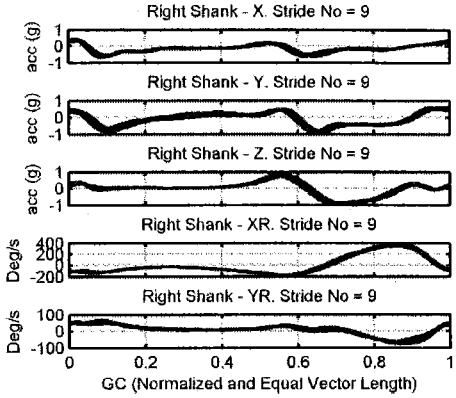

Figure 2: Example of Right Shank. Stacks of normalized and equal vector length GCs

S1	LiNo	Right Shank	X
S2	GC	Left Shank	Y
S3		Right Thigh	Z
S4		Left thigh	XR
		BCOM	YR

Figure 3: Structure Format

8. Features selection within a GC
   Several features are selected out from a GC of each side of the leg. Mid-Stance (MST) [16] is defined as 0.4 of a GC. The authors had selected four events from a GC, i.e. [HC, MST, TO, MSW]. This paper demonstrates 102 features selected from five IMUs, i.e. 5 IMU x 5 axes x 4 features + 2 Left/Right Stride Time. Figure 4 shows a sample of selected gait features marked on right shank-XR. For generality, the number of features could be any selections.

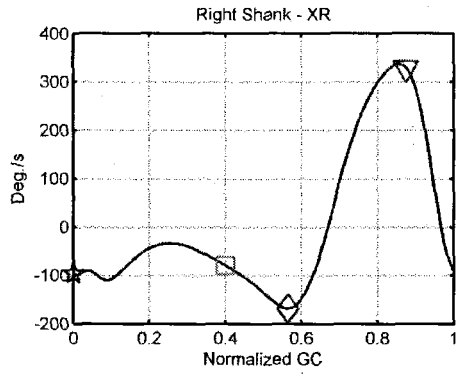

Figure 4: Features selection from a filtered GC. (☆ HC, □ MST, ◇ TO, ▽MSW). (HC = Heel Contact , MST = Mid-Stance, TO = Toe-Off, MSW = Mid-Swing)

9. Extraction of feature vectors and array
   The features are extracted from LiNo GC and organized in a feature matrix such as equation 1. Each row of the matrix is corresponding to a specific observation and labelled. This paper demonstrates 102 feature vectors.

$$\begin{bmatrix} \vec{S}_1 \\ \vec{S}_2 \\ \vec{S}_3 \\ \vec{S}_4 \end{bmatrix} \rightarrow \begin{bmatrix} \vec{X}_{11} & \vec{X}_{12} & \cdots & \vec{X}_{1m} \\ \vec{X}_{21} & \vec{X}_{22} & \cdots & \vec{X}_{2m} \\ \vec{X}_{31} & \vec{X}_{32} & \cdots & \vec{X}_{3m} \\ \vec{X}_{41} & \vec{X}_{42} & \cdots & \vec{X}_{4m} \end{bmatrix} \quad (1)$$

where:

$\vec{S}_i$ labels of each walking restrictions

$\vec{X}_{im}$ selected feature vectors (m) as input variables in i restriction

10. Principle Component Analysis (PCA)
    PCA is an orthogonal transformation that transforms correlated variables to uncorrelated variables called principle components (PCs). Each column of the feature matrix must be standardized. The restrictions ($\vec{S}_1$ to $\vec{S}_4$) could be plotted in 2D or 3D graphs using first two or three PCs.

3. RESULTS

Results of test-retest reliability of LiNo GCs are shown in Table 1. In general, high CA are noticed ($\geq$ 0.78) but mostly above 0.9. Low standard deviation are noticed, indicating high repeatability.

Figure 5 shows clearly separated clusters (S1 to S4) of subject1 under different walking restrictions on 2D PCA plots over two weeks. Each restrictions have resulted in almost the same cluster locations over two weeks, i.e. S1 at the left side, S2 at the bottom, S3 at the top and S4 at the right side. S1 and S4 are two extreme setup. Their distance is the farthest. Additional information could be further interpreted from the figure. Two categories, i.e. ankle restriction and level restriction could be grouped in the plots. Clusters (S1, S3) represent the group of 'ankle-free' while clusters (S2, S4) represent the group of 'ankle-locked'. Clusters (S1, S2) represent the group of 'flat-level' while clusters (S3, S4) represent the group of 'tilted-level'. Clusters in the group of 'flat-level' appear to be apart but 'tilted-level' appear to be closer.

Table 1: CA of The Ambulatory System

IMU No	Loc.		X	Y	Z	XR	YR
1	R Shk	μ	0.98	0.98	0.98	0.99	0.96
		s	0.01	0.02	0.02	0.01	0.02
2	L Shk	μ	0.78	0.97	0.94	0.98	0.81
		s	0.05	0.01	0.02	0.01	0.02
3	R Thg	μ	0.93	0.98	0.96	0.99	0.95
		s	0.02	0.01	0.01	0.01	0.02
4	L Thg	μ	0.92	0.93	0.97	0.99	0.83
		s	0.02	0.02	0.01	0.00	0.03
4	BCOM	μ	0.96	0.97	0.98	0.81	0.93
		s	0.02	0.01	0.01	0.12	0.02

*μ = mean, s = standard deviation

Figure 6 lists the variance explained by each PCs and their cumulative percentages. Only the first 10 out of 102 PCs are shown. The first principle component (PC1) represents the most percentage of the variance explained after PCA is performed, followed by PC2 and PC3 and so on. A 2D plot (PC1, PC2) accounts for around 53% of the total variance explained while a 3D plot (PC1, PC2, PC3) accounts for around 62% of the total variance explained.

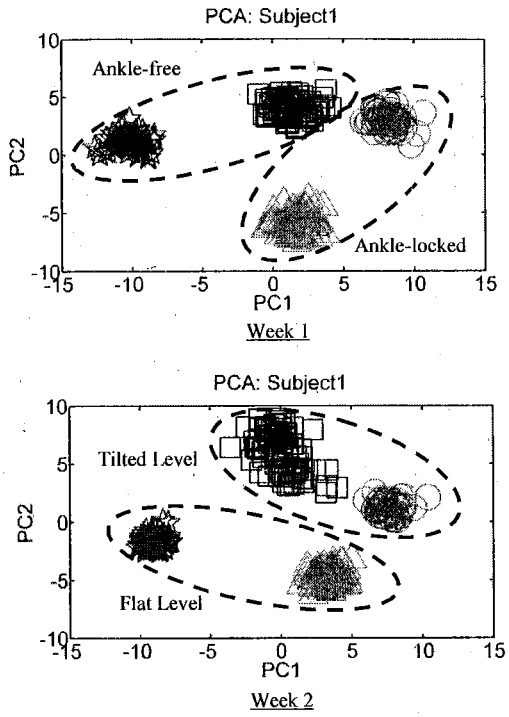

Figure 5: Clusters of Observations of Three Subjects Over Two Weeks. (

☆ (S1) Walking normally on flat level.

△ (S2) Walking with ankle-locked on flat level.

□ (S3) Walking normally on a 5^0 tilted manual treadmill.

○ (S4) Walking with ankle-locked on a 5^0 tilted manual treadmill.)

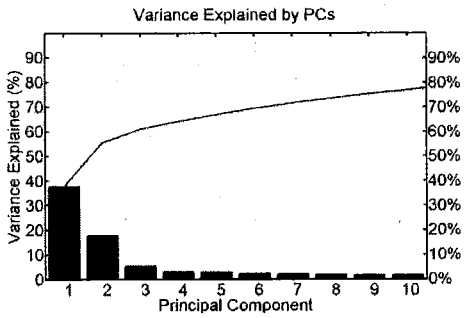

Figure 6: Variance Explained by Each Principle Components

4. DISCUSSION

The accuracy and repeatability of the gait data are largely dependent on the instrument reliability. Test-retest reliability shows that the instrument exhibits reliable and consistent performance.

Selection of IMU and IMU axis for gait identification have raised a few attentions. Firstly, IMU outputs are attenuated [3] as mounted IMU position approximates proximally. In this paper, their outputs at shanks are highest and clearly spotted from a time-series plot in corresponding axes. Secondly, shank IMUs provide choices of outputs from its accelerations and angular speeds at vertical, anterior/posterior (A/P) and medial/lateral (M/L) axes respectively. However gyroscope output at M/L axis (sagittal plane) exhibits larger excursions and relatively clearer signals during walking as compared to A/P and accelerations. Thirdly, gait events (HC, TO) under other walking restrictions besides normal walking at flat level (S1) are unclear. Researchers [4, 8, 10] had proposed gait identification (HC, TO) using a gyroscope shank lateral axis under normal level walking. However IMUs outputs varies substantially according to different walking restrictions. Features as shown in Figure 1 might not be clearly noticeable in other waking restrictions. Assistive aids using shank accelerations [15] during gait event identification provide secondary suggestions. At HC or TO, time-series shank accelerations always exhibit spikes which will coincides with the definition of M/L gyroscope outputs. Lastly, it is a tedious and time-consuming process to identify gait events manually from all collected gait data. An automatic and robust algorithm is urgently needed to identify gait events from all walking terraces and restrictions regardless of the subjects.

The choices of filter types and cut-off frequencies vary amongst researchers [5, 17]. A typical spectral analysis will reveal the signal spectrum. The noises originate from muscle movements and vibrations, impact during HC, abrupt changes of lower limb during falling and supporting. In this paper, it is highlighted that gait events (HC, TO, MSW) are identified using unfiltered gait data because all high frequency spikes that carry the features are removed in filtered gait data. To eliminate different time delay, this paper has introduced a zero phase low pass filter (LPF). Further analysis using unfiltered GC might cause unsatisfactory results due to greater variability and uncertainties caused by the noises.

Adaptive to walking restrictions, human walking is a series of controlled falling and supporting events. These controlled events exhibit a centre tendency which could be classified by its categories and hence revealed in selected feature variables. The selection of feature variables are unlimited. However, correlated variables will cause unnecessary analytical noises. High dimension variables could not be easily interpreted using a plot. PCA is the simplest way to eliminate correlated noises and provides an option to plot in low dimension with certain extent of confidence. Besides providing low dimension plots, PCA also provides clean, uncorrelated variables in preparation before other analysis such as Self-Organizing Map (SOM) and Back-Propagation Artificial Neural Network (BPANN).

Stacks of normalized and equal length GCs have opened a number of potential analysis [11, 12]. For example, symmetry index [18] of the left and the right legs and step-to-step variations [19]. Multi-variants statistical analysis such as MANOVA, factor analysis and pattern recognition algorithms using neural network are highly applicable.

5. CONCLUSION

A reliable customized ambulatory system is proposed to collect kinematic gait data from three healthy subjects under four walking restrictions. Meanwhile, a gait analysis procedure is introduced to provide comprehensive results in low dimension plots. Visualization in low dimension plots provide a potential objective aids in many medical applications such as pathological gait analysis and lower limb prosthetic dynamic alignment. Gait data explorations using multi-variant statistics and neural networks are potential future applications.

REFERENCES

[1] D. Winter, *Biomechanics and motor control of human movement*, 4 ed. Hoboken, New Jersey: John Wiley & Sons, 2009.

[2] M. W. Whittle, *Gait analysis: an introduction*, 3 ed. Oxford: Butterworth-Heinemann, 2001.

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

[4] K. Aminian, B. Najafi, C. BulaBula, P. F. Leyvraz, and P. Robert, "Spatio-temporal parameters of gait measured by an ambulatory system using miniature gyroscopes," *Journal of Biomechanics,* vol. 35, pp. 689 - 699, 2002.

[5] R. Moe-Nilssen, "Test-retest reliability of trunk accelerometry during standing and walking," *Archives of Physical Medicine and Rehabilitation,* vol. 79, pp. 1377-1385, 1998.

[6] M. Henriksen, H. Lund, R. Moe-Nilssen, H. Bliddal, and B. Danneskiod-Samsøe, "Test-retest reliability of trunk accelerometric gait analysis," *Gait & Posture,* vol. 19, pp. 288-297, 2004.

[7] J. Rueterbories, E. G. Spaich, B. Larsen, and O. K. Andersen, "Methods for gait event detection and analysis in ambulatory systems," *Medical Engineering and Physics,* vol. 32, pp. 545-552, 2010.

[8] K. Tong and M. H. Granat, "A practical gait analysis system using gyroscopes," *Medical Engineering & Physics,* vol. 21, pp. 87-94, 1999.

[9] M. Raggi, A. G. Cutti, A. Giovanardi, L. Chiari, D. Orlandini, and A. Davalli, "Wearable sensors for the real-time assessment of gait temporal symmetry in above-knee amputees: The 'SEAG' protocol," *Gait & Posture,* vol. 28, pp. S26-S27, 2008.

[10] I. P. I. Pappas, T. Keller, S. Mangold, M. R. Popovic, and V. M. Dietz, M., "A reliable gyroscope-based gait-phase detection sensor embedded in a shoe insole," *Sensors Journal, IEEE,* vol. 4, pp. 268-274, 2004.

[11] T. Chau, "A review of analytical techniques for gait data. Part 2: neural network and wavelet methods " *Gait and Posture,* vol. 13, pp. 102 - 120 2001.

[12] T. Chau, "A review of analytical techniques for gait data. Part 1: fuzzy, statistical and fractal methods " *Gait and Posture,* vol. 13, pp. 49-66, 2001.

[13] K. S. Tee, M. Awad, A. Dehghani, D. Moser, and S. Zahedi, "Comparison of Two Static Calibration Methods of an Inertial Measurement Unit," in *Biomedical Engineering*, Innsbruck, Austria, 2011.

[14] K. S. Tee, M. Awad, A. Dehghani, D. Moser, and S. Zahedi, "Triaxial Accelerometer Static Calibration," in *The 2011 International Conference of Mechanical Engineering*, London, UK, 2011.

[15] J. M. Jasiewicz, J. H. J. Allum, J. W. Middleton, A. Barriskill, P. Condie, B. Purcell, and R. C. T. Li, "Gait event detection using linear accelerometers or angular velocity transducers in able-bodied and spinal-cord injured individuals," *Gait and Posture,* vol. 24, pp. 502-509 2006.

[16] J. Perry, *Gait analysis : normal and pathological function*. Thorofare, N.J.: SLACK inc, 1992.

[17] R. C. González, A. M. López, J. Rodriguez-Uría, D. Álvarez, and J. C. Alvarez, "Real-time gait event detection for normal subjects from lower trunk accelerations," *Gait & Posture,* vol. 31, pp. 322-325, 2010.

[18] R. E. Hannah, J. B. Morrison, and A. E. Chapman, "Prostheses alignment: effect on gait of persons with below-knee amputations," *Archives of Physical Medicine and Rehabilitation,* vol. 65, pp. 159-162, 1984.

[19] M. S. Zahedi, W. D. Spence, S. E. Solomonidis, and J. P. Paul, "Repeatability of kinetic and kinematic measurements in gait studies of the lower limb amputee," *Prosthetics and Orthotics International,* vol. 11, pp. 55-64, 1987.