

Gait Evaluation using Procrustes and Euclidean Distance Matrix Analysis

Arif Reza Anwary, Hongnian Yu and Michael Vassallo

Abstract—Objective assessment of gait is important in the treatment and rehabilitation of patients with different diseases. In this paper, we propose a gait evaluation system using Procrustes and Euclidean distance matrix analysis. We design and develop an android app to collect real time synchronous accelerometer and gyroscope data from two Inertial Measurement Unit (IMU) sensors through Bluetooth connectivity. The data is collected from 12 young (10 for modelling and 2 for validation) and 20 older subjects. We analyse the data collected from real world for stride, step, stance and swing gait features. We validate our method with measurements of gait features. Generalized Procrustes analysis is used to estimate a standard normal mean gait shape (NMGS) for 10 young subjects. Each gait feature of both young and older subjects is then converted to find the best match with the NMGS using ordinary Procrustes analysis. The shape distance between the NMGS and each gait shape is estimated using Riemannian shape distance, Riemannian size-and-shape distance, Procrustes size-and-shape distance and Root mean square deviation. A t-test is performed to provide statistical evidence of gait shape differences between young and older gaits. A mean form which is considered as a standard normal mean gait form (NMGF) and inter-feature distances are estimated from the set of 10 young subjects. The form difference is estimated between the NMGF and individual gaits of young and older. The degree of abnormality is then estimated for individual features and the result is plotted to visualize the feature in a gait. Experimental results demonstrate the performance of the proposed method.

Index Terms—Gait Analysis; Gait Assessment; Gait Features; Inertial Measurement Unit (IMU)

I. INTRODUCTION

HUMAN gait is the result of a series of rhythmic alternating movements of the arms, legs, and trunk which create forward movement of the body [1]. Its complex mechanisms depend upon the integrated actions of the musculoskeletal, nervous system, visual, vestibular, auditory systems leading to the smooth propulsive movement of the centre of gravity. Quantification of gait variabilities, kinematic and kinetic measurements, muscular measurements and energy expenditure, provide comprehensive locomotive gait information [2]. Gait quantification information is used to 1) distinguish the type of gait impairments and suggest possible

diagnoses; 2) measure and monitor the severity of an injury or a disease and determine the most appropriate treatment [3]; 3) be a determinant of progression in patients with medical conditions causing gait disorders [4, 5] monitor response to treatment in orthopaedic rehabilitation [6]; 4) monitor and improve an athlete's performance [7]; and 5) in biometrics and biomedical engineering areas, be an assistive tool to characterize human locomotion and have many applications [8]. Gait quantification information is important in elderly patient fall risk assessment [9] and also a predictor of functional and cognitive decline [10]. Therefore, the objective evaluation of gait and understanding the gait changes has many potential uses.

The paper is organized in the following sections. Section II introduces previous related work. Section III presents the proposed method. Section IV delivers the experimental results to demonstrate the proposed method. Section V presents the discussions. The conclusion is given in section VI.

II. RELATED WORK

The tools and methodologies used to assess human gait are often arbitrary and often studied in artificial controlled conditions. Gait abnormalities are generally assessed by physicians, physiotherapists and researchers in clinical settings or in gait laboratories. Clinical scales used to analyse gait parameters such as Gait Abnormality Rating Scale [11], Figure of 8 Walk Test [12], and Berg Balance Scale [13] are subjective or semi-subjective and a poor replacement to laboratory based methods. This may not satisfy scientific criteria of reliability and validity [14], which may affect the accuracy of diagnosis, follow-up and treatment [2]. There is no commonly accepted guideline, preferred methodology or protocol for gait changes evaluation. The European GAITRite Network Group, developed Guidelines for Clinical Applications of Gait Analysis [15], with the intention to facilitate collaboration and provide guidance to clinicians however there is no recommended systematic procedure in the guideline. The available common approaches [28] for gait quantification of temporal and spatial gait pattern are Symmetry index, Symmetry ratio, Ratio, Gait asymmetry, etc. The commonly used Symmetry Indices need to be normalized to a reference value [16, 17] and there is potential influence for artificial inflation as the normal values for young and older subjects are not the same [18]. Sometimes the mean value calculation used for quantifying gait asymmetry may lead to erroneous results as the mean measurements from two abnormal limbs may appear normal. For example, in a situation where a patient has asymmetry in the opposite direction of gait, the true magnitude

Arif Reza Anwary, Faculty of Science and Technology, Bournemouth University, UK, manwary@bournemouth.ac.uk

Hongnian Yu, Faculty of Science and Technology, Bournemouth University, UK, Correspondence Author: yuh@bournemouth.ac.uk; Tel: +44-01202 961150

Michael Vassallo, Royal Bournemouth Hospital, UK, CoPMRE Bournemouth University UK, Michael.vassallo@rbch.nhs.uk

of asymmetry for affected or unaffected limbs may be very small. The effect of the direction of gait asymmetry may be eliminated using absolute values in the symmetry indices [16]. There are methods [19, 20] which do not make it possible to identify the point during the gait cycle at which deviations occur. There are other approaches [21, 22] including principal component analysis, regions of deviation analysis, and paired t-test to quantify gait symmetry. However, the number of test subjects and experiments are important for these methods. These methods may also need normative data from able-bodied subjects as a reference [17]. Although gait asymmetry is frequently reported as present or not present which may not satisfy scientific criteria of reliability and validity [14], an arbitrary cut-off value of 10% deviation from perfect symmetry has been used as a criterion of asymmetry in gait assessment [23, 24]. This is later criticized due to its non-parameter specific nature [18]. Other previously used criteria to describe the absence or presence of gait asymmetry include sensitivity and specificity of parameter measurement [25], the use of 95% CI where gait asymmetry within the limits of a 95% CI obtained in a healthy population would define able-bodied gait, while gait asymmetry outside the 95% CI would define pathologic gait [18], and significant limbs difference [21] etc. Although there are many approaches for quantifying gait asymmetry, there is little research conducted on a gait quantification method based on overall gait features. Considering all the various parameters that constitute the gait cycle, we propose a novel gait quantification method which offers a simple and easily interpretable assessment of gait with good accuracy and comprehensive features.

In order to provide comprehensive gait information and evaluation in clinical screening and research, an affordable gait evaluation system is required which will provide the facility in clinic or at home. The aim of this study is to propose a novel method of gait evaluation using Procrustes superimposition [26] and Euclidian Distance Matrix Analysis (EDMA). To quantify individual gait based on all features four shape and size comparison techniques (Riemannian shape distance (RSD) [27], Riemannian size-and-shape distance (RSSD) [28], Procrustes size-and-shape distance (PSSD) [29] and Root mean square deviation (RMSD) [29]) are applied. We also aim to investigate how each feature impacts on a gait using EDMA. A high difference between the NMGF and each gait indicates a high degree of abnormality and a low value indicates close to a normal gait. To date, research on comprehensive understanding of gait quantification based on overall gait features to allow assessment and monitoring of gait changes from young and older adults has received little attention. Our method provides the facility to quantify gait and gait changes in both a clinic and at home which increases the availability and affordability of gait assessment.

III. METHODS

A. Participants Selection

A convenience sample of 32 subjects are recruited: 12 healthy young subjects (9 male, mean age 25.4 years, standard deviation 4.64, range 19-35 years); 20 older adults (19 male, mean age 71.86 years, standard deviation 8.55, range 62-86 years). Among 12 young subjects, 10 are used for modelling

while an additional 2 are used for validation. Young subjects are selected with no signs of gait, balance or walking abnormalities. Older adults from a care home are invited to participate. They are a group of patients chosen with some having a normal and others an abnormal gait. It is coincidental that the majority of subjects are male.

B. Sensor placing location

In this study, the sensors are placed at the base of the first metatarsal of both feet. This position was previously determined and validated for collecting data since this can achieve the best performance compared to other foot locations [30] (Figure 1(a)).

C. Data collection

Our proposed android app for synchronous data collection from accelerometer and gyroscope is shown in Figure 1(b). The subjects perform a walk in a straight corridor comprising of 15 strides of normal forward walking, a turn-around and another 15 strides. The accelerometer and gyroscope raw data from young subject 1 is presented in Figure 1(c).

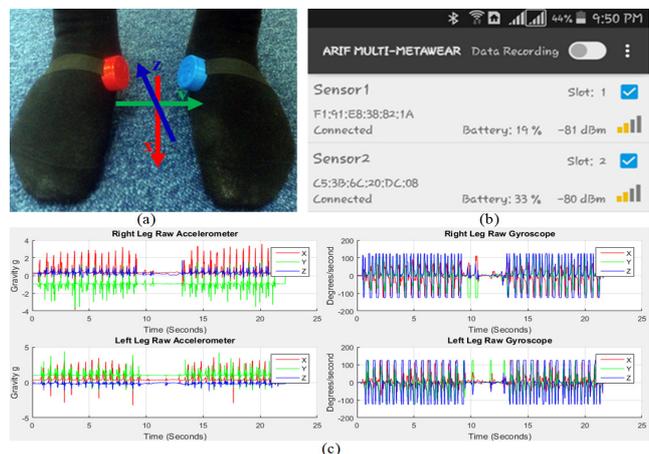


Figure 1. (a) IMU sensors placement in right and left metatarsal feet locations, (b) Android app for synchronous data collection from accelerometer and gyroscope, (c) Raw accelerometer and gyroscope data of young subject 1

D. Stride, stance, swing and step phase detection

Human walking can be described in the context of a gait cycle which has eight events shown in Figure 2 with stance and swing phases. A stride (whole gait cycle) is the distance between a point on one foot at the first foot contact and the same point on that foot at the next foot contact.

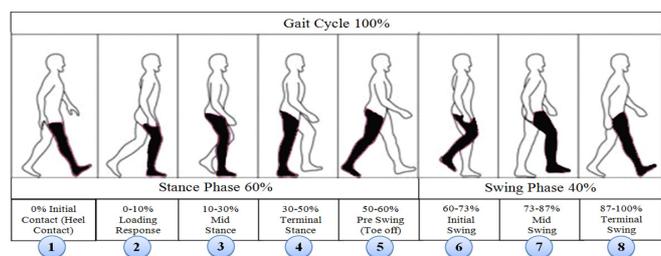


Figure 2: Normal human gait phases [31]

The stance phase shown in Figure 2 starts when the heel contacts the ground and the waist is in its lowest position during

the entire phase. There is deceleration of the leg towards the horizontal axis as the velocity moves to zero. The zero velocity remains until the terminal stance event where the foot is flat on the ground. In the pre-swing event, the toe is off the ground and starts forward movement demonstrating initial acceleration towards horizontal axis. The swing phase is when the heel moves off the ground. The acceleration interval corresponds to the change from the heel lift to the swing at the height point at mid-swing event. Deceleration starts during the terminal swing event from the highest point to the foot back flat on the ground. There is zero velocity again in the interval corresponding to the change from a flat foot to a heel lift. The eight events of a gait cycle presented in Figure 3 are identifiable from the IMU acceleration signal. The same phenomenon of human limb kinematic with accelerometer signal output during a typical walking cycle is identified in [31, 32]. Our gait cycle accelerometer signal (Figure 3) is agreed with the signal pattern in [31, 32]. Figure 3 shows the events of the gait cycle (Figure 2) with corresponding accelerometer signal.

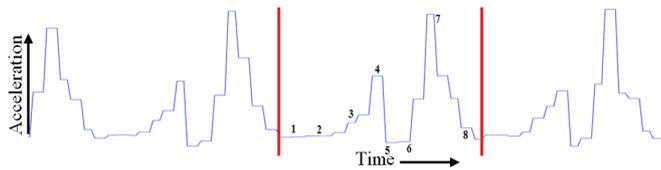


Figure 3: Eight different events of a gait cycle from accelerometer data

Figure 3 shows that at the start and end of each stride, the feet are stationary on the ground. Due to the walker's forward movement, the acceleration shows its high value in the swing phase. Based on these characteristics, we identify stride, stance and swing events from accelerometer signal shown in Figure 4.

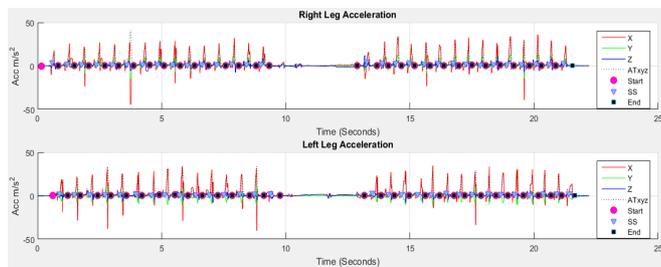


Figure 4. Result of stride, stance and swing event detection

Figure 4 shows the detected *Start* (purple circle) which is the foot's initial contact to the ground, the transition of stance-swing phase *SS* (cyan triangle) is the initial swing and *End* (black square) is the terminal swing of gait phases information of each stride for both legs where the stance phase information is provided by the difference between *Start* and *SS*; and the swing information is the difference between *SS* and *End*.

E. Velocity and distance estimation

The Madgwick quaternion technique [33] is applied for estimating the orientation followed by the trapezoidal double integral approach [34] applied to obtain the travelled distance from the user movement using accelerometer and gyroscope data. The input data are passed through a high-pass filter to remove the direct component of the acceleration signal.

We obtain values for 13 spatial-temporal gait features separately from the right and left lower limbs previously validated with measurements of gait features collected in a laboratory environment using a Qualisys Motion Capture System [30, 35]. These include total distance (m), total time (s), velocity (m/s), swing length (m), swing velocity (m/s), stride length (m), stride time (s), stride velocity (m/s), step length (m), step time (s), step velocity (m/s), stance time (s), and swing time (s). From our evaluation, we conclude that the first five features are redundant since they can be estimated from the rest eight features. Therefore, we use the last eight features as these are all an average reading from 30 strides.

F. Understanding of shape, form and size

Both *shape* and *form* consisted of geometrical representation of an object can be represented by a set of points or landmarks. The form of an object may change when magnitude or volume changes along various axes and transforms from reference to a target form [36]. Figure 5 shows the relationship between shape, size and form changes.

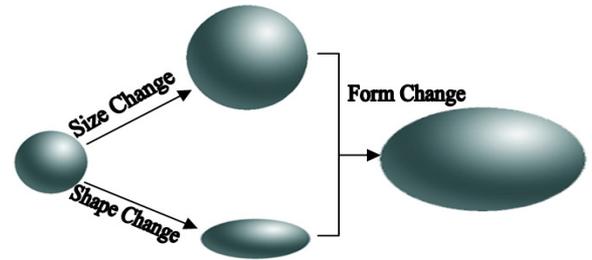


Figure 5: Geometric representation of form change relating with shape and size [36]

These landmarks remain invariant when an object is 1) moved within a given coordinate system (translation), 2) turned on any axis of a given coordinate system (rotation) and 3) flipped of a given coordinate system (reflection). For example, a triangle consists of three points considered to be landmarks. Any movement results in changes in the coordinate locations of the three points but no change to their relative positions. A new set of coordinates is therefore required to define the new position of the three points. The landmark coordinates matrix change upon reflection, translation or rotation even though the shape remained the same. In this research a total of eight gait features selected from the 13 extracted features are considered as landmarks. Procrustes analysis is used to analyse the distribution of features representing the gait shapes and Euclidean distance matrix analysis (EDMA) is used to analyse *form* difference between objects and the influence of each features in a gait.

G. Normal Mean Gait Shape estimation using Procrustes

In order to quantify and compare gait, a common procedure is to normalise the obtained gait features both in time and length. The eight gait features (stride length, stride time, stride velocity, step length, step time, step velocity, stance time and swing time) from right and left legs are presented in the Cartesian coordinate. The *x* and *y* axes represent the features of the right and left legs with the dimensionless numbers respectively. This coordinate represents the shape of gait features collected from both legs.

Procrustes analysis (a method of statistical analysis used to analyse the distribution of a set of shapes) describes curve shape and shape change in a mathematical and statistical framework, independently of time and size factors. Ordinary Procrustes analysis (OPA) finds the rotation matrix, translation vector and scaling factor to give the best match between two configurations [26]. Generalized Procrustes Analysis (GPA) is used to find the best fit among multiple objects [26, 29]. Instead of considering matching all possible independent matrix pairs, GPA is used in such a way that all matrices are simultaneously subjected to suitable rotation, translation and scaling transformations until a proper fit criterion is reached. For estimating NMGS using GPA, 10 young subjects gait information is used. GPA provides the least square correspondence of more than two data matrix configurations, X_i ($i=1,2,3,\dots,m$) be a series of m matrices that contain the coordinates of a set of p gait features called as landmarks on the m subjects called as number of shapes in k dimensions. Translation, rotation and scaling of a configuration are described [29] as

$$\hat{X}_i = c_i X_i O_i + j t_i^T \quad (1)$$

where \hat{X} is the new coordinate of the landmarks in the configuration. O_i is the rotation matrix, c_i is the scaling factor, t_i is the translation vector and j is the unit vector. Using GPA the configurations are translated, rotated and rescaled until the sum of the squares of the distances between the equivalent landmarks are minimized to give the best possible match between all configurations. Figure 6 shows the procedure where the individual configurations are translated, rotated and scaled so that they can be “superimposed” on each other to achieve a “best” fit.

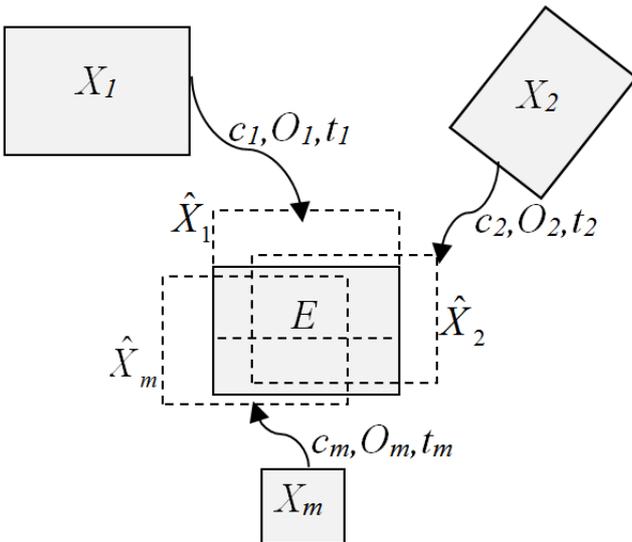


Figure 6: Concept of GPA

Iterative procedures are used for the minimisation process in GPA. The shapes are repeatedly scaled, rotated and translated until the sum-of-squares defining the distances between the equivalent landmarks on all shapes is minimised. The Procrustes derivation is described in [37].

The Procrustes superimposition computes a mean shape referred as NMGS for the young subjects based on gait features where scaling and reflection are not performed in this analysis. The shape of each subject’s gait is defined by its Procrustes residuals which are the deviation of the landmarks from the NMGS.

H. Gait shape comparison

RSD, RSSD, PSSD and RMSD are used to quantify a gait based on all gait features. In Riemannian geometry [38], a shortest curve between a pair of points on a curved surface is called a minimal geodesic. On some surfaces, there may be pairs of points which have more than one minimal geodesic between them (e.g., a sphere). RSD gives a measure of the relationship between the curvature of a space and its shape. Its parameter has a value between 0 and $\pi/2$; the smaller this value, the smaller the difference between the gaits. RSSD is the Riemannian distance between the size-and-shape of the configurations found by minimizing the Euclidean distance over rotations. The smaller the value is, the closer the configurations in size-and-shape distance. PSSD is defined as the distance between two shapes as the closest distance between the fibers on the pre-shape sphere in a non-Euclidean shape metric space. This allows us to compare two configurations which are independent of position, scale and rotation. RMSD is another measure of size-and-shape differences between configurations where the value is estimated from the square root of ordinary Procrustes sum of squares divided by the number of landmarks and number of dimensions. The small value means the small deviation between the configurations. RSD, RSSD, PSSD and RMSD are estimated for distinguishing degree of abnormality of each gait compared to NMGS. Each gait is translated and rotated to find the best match with NMGS using OPA and the distances are then estimated between NMGS and each best match gait.

I. Mean form and inter-feature distance estimation

EDMA [39] for comparing two shapes using landmark data is a method for comparing the forms of organisms that are measured using homologous landmarks. Homologous landmarks are those landmarks chosen to represent features on organisms that are similar due to a phylogenetic relationship. The organisms being compared thus share a common ancestor and the feature under study is present in all organisms under consideration due to each inheriting it from the common ancestor [40]. EDMA also allows form variation, shape or growth differences to examine through the comparisons of ratios of landmarks of equivalent configurations [39, 41].

The gait features extracted from each subject vary due to their walking style, speed and body characteristics etc. This variation is manifested as perturbations around the mean gait configuration. These perturbations vary in size and shape from feature to feature. Initially, the Euclidean distance between all pairs of features are estimated which is known as inter-feature distances [41]. The data is stored in an 8x8 symmetric matrix known as inter-feature distance matrix. The inter-feature distance matrix from all young subjects is then used to calculate the mean form matrix. The procedure of developing EDMA is given in [37].

J. Form matrix and form difference matrix estimation

$FM(A)$ is as the form matrix (FM) and returns all the relevant information about the form of an object as summarized by landmark coordinates.

For a form difference matrix (FDM), suppose the forms of two objects, A and B , each with K landmarks are to be compared. The forms of these two objects correspond to two points in an L -dimensional Euclidean space. If the forms are similar, then these two points lie on a ray going through the origin. If the above condition is true, then it can be concluded that the forms are different. A represents the form of the gait features each subject (including both young and older), B represents the mean form estimated from 10 young subjects using EDMA. $FM_{ij}(B)$ represents the reference form which is NMGS. $FM_{ij}(A)$ represents the real form measured from the individual. The ratios of corresponding linear distances from the two forms are calculated.

$FDMs$ contain all the relevant information (represented by the landmarks collected) regarding morphological distances between two forms (or sample of forms). Differences of form can reflect a simple difference in scaling of two forms (i.e. only in size), or a combination of difference in size and shape.

$FDM_{ij}(B,A)$ is then used to estimate the form difference from all subjects. The variance and covariance are estimated for individual features. Two gait features have the same form if their Euclidean matrixes are identical. Two gait features also have the same form if the Euclidean matrix describing one form is a constant multiple of the Euclidean matrix describing the second form. The procedure of developing FM and FDM using EDMA is given in [37].

IV. EXPERIMENTAL RESULTS

To verify the proposed gait quantification approach, we perform experiments to our collected gait features from all subjects. We also present detailed analysis on the experimental results using the statistical software R [42].

A. Data collection

A database is created for our experiment using the automatic gait feature extraction method presented in section III.D. The database consists of eight selected gait features among the 13 features extracted from both legs for all subjects shown in Figure 7. Eight features (stride length (m), stride time (s), stride velocity (m/s), step length (m), step time (s), step velocity (m/s), stance time (s), and swing time (s)) of all individual subjects are plotted and each of these points is notionally joined together to represent a shape.

Figure 7 shows that gait features of young subjects from the right and left legs are very similar, i.e., the features lying on or close to a hypothetical diagonal 45° line indicative of perfect symmetry (equal features arising from both legs). Conversely for the older subjects there is more variability in output of features from their legs. This results in a greater scatter in the output recorded, indicative of greater asymmetry shown in Figure 7. For this reason, we chose to perform our GPA on the young subjects who had a more normal gait than the older subjects with a view of developing a reference NMGS.

B. Estimating of Mean Normal Gait Shape (NMGS)

We perform GPA on the features (shapes) derived from 10 young subjects. To do this all 10 shapes of the young subjects obtained from both legs are plotted after GPA best fit alignment shown in Figure 8.

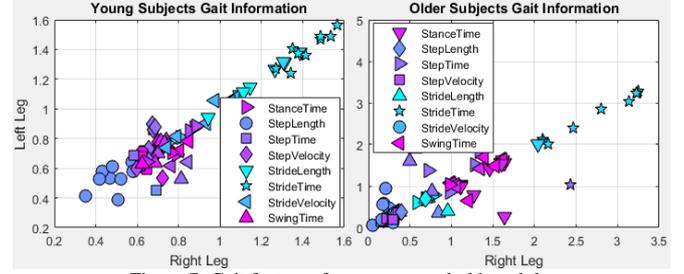


Figure 7: Gait features from young and older adults

This GPA translates and rotates each of the shapes to find the best fit. The mean of each shape of the features is then estimated and plotted generating the shape of NMGS shown in Figure 8 (black line).

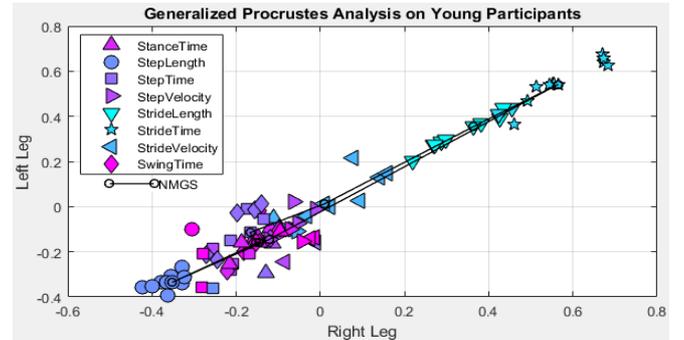


Figure 8 : Gait features from young after Generalised Procrustes analysis and the black line represents Normal Mean Gait shape (NMGS)

Figure 8 shows that gait features and NMGS obtained as the mean features from the individual young subjects are very close to the diagonal.

C. Gait quantification

Next, we determine the shape differences between each pair of shapes i.e. NMGS with the individual gait shapes. To quantify a gait based on all gait features we use four shape comparison techniques (RSD, RSSD, PSSD and RMSD) shown in Table 1. Results closer to 0 suggest a gait shape close to the NMGS gait.

TABLE 1
Gait Quantification Information

		RSD	RSSD	PSSD	RMSD
Young	1	0.129	0.152	0.129	0.054
	2	0.245	0.292	0.243	0.103
	3	0.304	0.364	0.299	0.129
	4	0.223	0.329	0.222	0.116
	5	0.367	0.467	0.359	0.165
	6	0.264	0.33	0.261	0.117
	7	0.204	0.237	0.202	0.084
	8	0.418	0.473	0.406	0.167
	9	0.38	0.441	0.371	0.156
	10	0.205	0.251	0.204	0.089
	11	0.270	0.324	0.267	0.156
	12	0.186	0.262	0.185	0.078

Older Adults	1	0.977	2.872	0.829	1.015
	2	0.922	1.885	0.797	0.666
	3	1.144	4.35	0.91	1.538
	4	0.905	1.79	0.786	0.633
	5	0.92	3.586	0.795	1.268
	6	0.886	3.104	0.775	1.097
	7	0.918	2.372	0.795	0.838
	8	0.874	1.711	0.767	0.605
	9	1.154	2.291	0.915	0.81
	10	0.959	3.417	0.819	1.208
	11	0.934	3.319	0.804	1.173
	12	1.058	3.755	0.872	1.328
	13	1.442	6.6	0.992	2.333
	14	1.018	2.989	0.851	1.057
	15	1.019	2.977	0.852	1.053
	16	1.173	5.084	0.922	1.798
	17	1.001	2.548	0.842	0.901
	18	0.94	1.848	0.807	0.653
	19	1.202	2.843	0.933	1.005
	20	1.01	3.809	0.847	1.347

Table 1 shows that variations of the distances of the young subjects are smaller than those of the older subjects. Therefore, Table 1 can help distinguishing different gait patterns in young and older adults.

We evaluate the data for statistical errors and assessed whether the estimated values are reasonable. A *t*-test comparing the mean values of RSD, RSSD, PSSD and RMSD values is carried out with a statistical significance level (alpha) of 0.05. The two sample unpaired *t*-test summary are given in Table 2.

TABLE 2
T-test for distances between MNGS and gaits

	MD	SD	t-value	p-value	df	95% Confidence Interval	
						Lower	Upper
Riemannian shape distance							
Young	0.274	0.092	9.441	0.000	9	0.208	0.340
Older	1.023	0.140	32.708	0.000	19	0.957	1.088
Riemannian size-and-shape distance							
Young	0.334	0.106	9.979	0.000	9	0.258	0.409
Older	3.158	1.199	11.775	0.000	19	2.596	3.719
Procrustes size-and-shape distance							
Young	0.270	0.088	9.706	0.000	9	0.207	0.332
Older	0.846	0.061	61.831	0.000	19	0.817	0.874
Root mean square deviation							
Young	0.118	0.037	10.013	0.000	9	0.091	0.145
Older	1.116	0.424	11.773	0.000	19	0.918	1.315

MD= Mean Difference

The *t*-tests indicate ($p < 0.05$) that there is a significant mean difference between the gait of young and older subjects for RSD, RSSD, PSSD and RMSD values.

The variability in gait shapes is shown in Figure 9 which determines the range of results. From the box plot and *t*-test above, it is clearly seen that the mean values of RSD, RSSD, PSSD and RMSD of the normal young is significantly lower than those of older adults.

Figure 9 shows that for young subjects, RSSD and RMSD are more consistent with less standard deviation (SD) than RSD and PSSD. For older subjects the opposite was identified with a wider SD for RSSD and RMSD than RSD and PSSD. The boxplot confirms the expected difference in gait shapes between young and older subjects. From Figure 9, we can observe that RSSD provides the best indication among the four

approaches since the variation of the older is large while the variation of the young is small. RMSD is the second best, followed by RSD and then PSSD. Next we determine what features of gait contribute to abnormality.

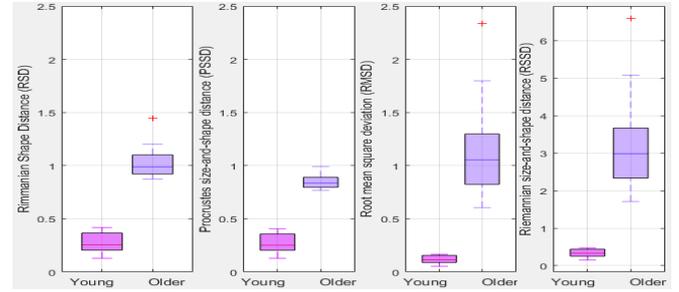


Figure 9: Boxplot of RSD, PSSD, RMSD and RSSD

D. NMGF and inter-feature distance estimation

The database created in section IV.A is used. The mean form based on these normal gaits is estimated and is considered as NMGF estimated directly from the unit less feature coordinate data using EDMA, which is shown in Table 3.

TABLE 3
Normal Mean Gait Form (NMGF) Information

Index	Feature	Right	Left
F1	Stride length	-0.48296	0.109081
F2	Stride time	-0.77767	-0.09042
F3	Stride velocity	-0.01428	0.002893
F4	Step length	0.489212	-0.05726
F5	Step time	0.192591	0.07357
F6	Step velocity	0.175095	-0.12865
F7	Stance time	0.211629	0.007791
F8	Swing time	0.206385	0.083

The Euclidean distance between all possible pairs of features are estimated from NMGF for the inter-feature distances which are stored in an 8x8 symmetric matrix. Table 4 presents the lower triangular part of the matrix.

TABLE 4
Inter-feature distances

	F1	F2	F3	F4	F5	F6	F7	F8
F1	0							
F2	0.356	0						
F3	0.481	0.769	0					
F4	0.986	1.267	0.507	0				
F5	0.676	0.984	0.219	0.324	0			
F6	0.700	0.954	0.231	0.322	0.203	0		
F7	0.702	0.994	0.226	0.285	0.068	0.141	0	
F8	0.690	0.999	0.235	0.316	0.017	0.214	0.075	0

Each cell in Table 4 of the inter-feature distance matrix shows the distance in two-dimensions that does not require a coordinate system. For example, the cell, that contains the number 0.356 in the mean form matrix of the young subjects, represents the distance between features F1 and F2. This is the distance estimated directly from the feature coordinate data. The inter-feature distance of NMGF is used to estimate the form difference matrix between NMGF and each gait to understand the degree of abnormality.

E. Form difference and form difference matrix between NMGF and each gait

Estimation of FDM is carried out for all gaits relative to NMGF. The sum of divergences to the median value for each feature is estimated considering the whole FDM matrix [43]. This is the matrix of the degree of abnormality where the higher the degree of difference the greater the abnormality. Lower values imply that the gait features of the individual are closer to NMGF and conversely higher values mean that there is greater abnormality as there is greater deviation from MNGF. To represent the degree of abnormality in a meaningful and easily interpretable way we propose a two dimensional plot to summarize, explore and interpret the FDM results. Figure 10 shows such a plot where x represents individual gait features and y represents the degree of abnormality in relation to the other features. The form difference for all eight gait features with respect to NMGF is plotted. For example in Figure 10 feature 1 has the highest difference with feature 2 but is very close to features 3-8. This analysis is applied to a set of 32 gaits (12 young and 20 older).

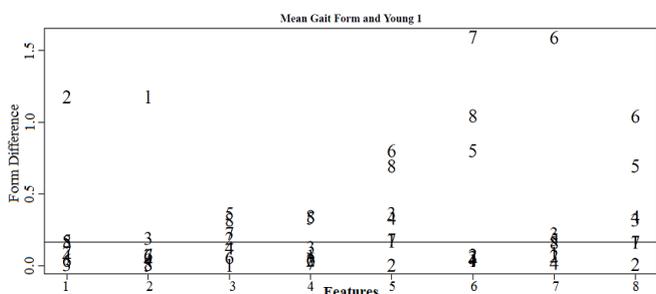


Figure 10: The degree of abnormality of Young Subject 1 with respect to the Mean Normal Gait Form for all the eight gait features.

V. DISCUSSION

This study demonstrates a detailed analysis of gait using Procrustes and EDMA methods. Procrustes is valuable in determining variation of gaits from NMGS while EDMA is useful in determining the degree of abnormality of the gait feature. We obtain the results using eight gait features collected automatically from both legs by adopting low cost IMU sensors synchronously. Our results have shown that a normal gait provides a set distribution of features. Any deviation from this distribution is identifiable as abnormal. This to our knowledge has not been done before. Although at this stage one cannot extrapolate this information to make accurate diagnoses the ability to identify such subtle differences in gait may have the potential to support specific diagnoses as well as treatment. This new method is more comprehensive than other methods that often rely on single or a smaller number of features [23, 44]. We also introduce a morphological analysis to the evaluation of gait where one can see a pattern of gait and identify where changes occur in the gait pattern. Different parameters of gait indicate different type of gait abnormalities.

Although our results are encouraging, there are a number of limitations. The number of subjects is relatively small (30) and no steps are taken to ensure a random sample. Coincidentally there is a gender bias with most subjects being male. The aim of the study is to see whether a Procrustes method can be used to analyse gait and not to study gait differences between the

genders. This gender bias is therefore unlikely to impact the value of our results and what they are trying to achieve. Other possible confounding factors are speed of walking as well as different height resulting in different gait parameters such as stride length. Our study was however intended to evaluate the normal baseline gait of our subjects only. The influence of these other factors will be studied in the future. Lastly, NMGS and NMGF are estimated using only 10 young subjects, while additional 2 young subjects are used for validation of our estimated NMGS and NMGF. There is the potential of a Type 1 error (false positive) in detecting an effect that is not there.

However, our work is a proof of concept study that has established our method for gait evaluation. Future work is to establish a database with a larger number of subjects which stores more medical and physical information as well as longitudinal data across a longer period of time. Such longitudinal information will demonstrate the potential for using our method in monitoring response to treatment in patient with gait disorders.

Normal gait is not determined by time and distance travelled. It is determined by the degree of variation in the gait features. While the time and distance can be assessed relatively easily using visual observation the variation is more difficult to determine. The Procrustes analysis uses translation and rotation among all gait feature shapes to find the best fit to identify such variation. We show how this normalization technique is used for a set of 10 normal young subjects to estimate NMGS. The RSD, RSSD, PSSD and RMSD distances between NMGS and all gaits are then calculated. We use the data from two additional young subjects to validate our results. This method has the potential to provide detailed analysis of gait on an individual basis. For example, from Table 1 we can see that the highest and lowest of RSD, RSSD, PSSD and RMSD distances are found in Y8 (young 8) and Y1, O13 (older 13) and O8 for older subjects respectively. From the individual gait features, the highest and lowest travelled distances are found from Y5 and Y10, the highest and lowest time are found from Y4 and Y8. Interestingly, considering all gait features, the highest variation lies in Y8. This is demonstrated in the Procrustes shape obtained in Figure 11a.

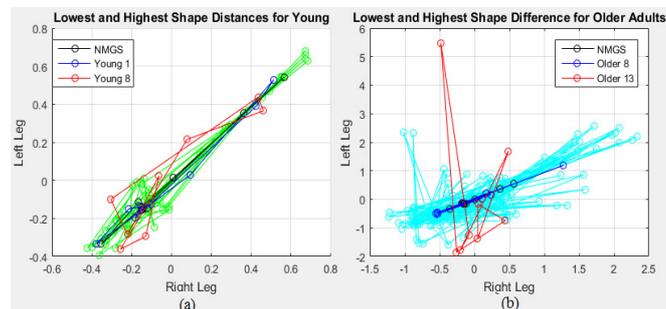


Figure 11: Lowest and highest shape differences from (a) young and (b) older subjects

Although, other young subjects travelled distance and time are higher than Y8, based on the overall gait features, the shape difference between NMGS, Y8 is the highest. Similar findings are also found for older subjects. The lowest and highest shape difference is found for O8 and O13 shown in Figure 11b.

Investigating the history of O13 helps explain the shape of the graph. In this case O13 had a stroke and numbness in the right leg. He was unable to move his right leg and used crutches for moving. Thus most of the movement during walking was covered by the left leg and crutches are used to keep body balance. In Figure 11b we can see that the normal left leg shows greater movement but the abnormal right leg has less movement detected. In the future we will investigate further the impact of specific diagnoses and patient health on these gait parameters by exploring gait patterns obtained in specific diagnoses such as Parkinson disease, Stroke, and other conditions causing abnormal gaits.

A *t*-test and Boxplots using RSD, RSSD, PSSD and RMSD distances show that the gait of young are distinguishable from older. The standard deviations are close to the mean indicating that the gait data distribution from young subjects' is more consistent than that from older. The Box plots of the four different distance approaches, RSD, RSSD, PSSD and RMSD, show different distributions. They indicate that for young subjects RSSD and RMSD provides more consistent results with less standard deviation (SD) than RSD and PSSD. For older subjects the opposite was identified with a wider SD for RSSD and RMSD than RSD and PSSD. This difference is likely to arise as a consequence of the different mathematical formulas involved in calculating these measurements. In the future we will explore the reasons for this in more detail.

To fully understand the degree of gait abnormality for older subjects, we use EDMA to locate the specific feature of the gait contributing to the abnormality. The process starts with estimating a mean form from a set of normal young gaits called as NMGF. It is then used to estimate the inter-feature distances that represents the distance between each feature from one to another. The form difference matrix is then estimated between NMGF and all gaits. Figure 12 shows the form difference of Y1, Y8, O8 and O13.

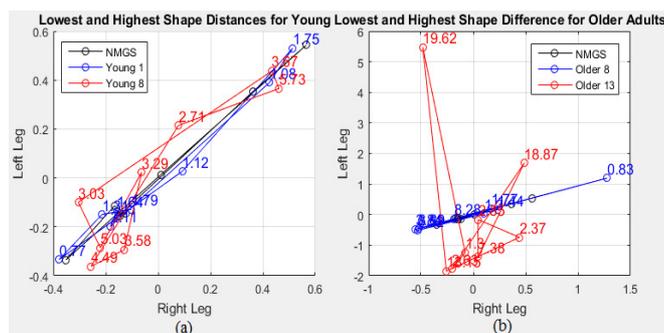


Figure 12: Degree of abnormality from (a) young and (b) older adults

Arguably, one can provide a multidimensional shape where each important feature is an axis which is to be explored in our future work. We have however deliberately chosen to use a two dimensional shape that is easy to visualize and interpret to start off. Gait is a complex interaction of all features together and giving individual RMSDs may be difficult to interpret in the context of the global picture of the persons' gait [45].

With an aging population and the increase in chronic illness such as poor mobility and falls there is an increasing drive for new technologies to support treatment of patients at their own home. Our proposed system can be used to monitor gait

abnormalities across a spectrum of diseases. A series of gait feature measurements on a regular basis can identify the progression or recession of changes in gait pattern as well as response to treatment with rehabilitation for these types of diseases and more. Growing young adults particularly if they have physical disabilities may develop gait abnormalities during puberty growth spurts. Periodic monitoring is becoming essential to make sure that such gait abnormalities are not progressing. Our method of gait evaluation can be used for such longitudinal monitoring for these cases. Our low cost gait evaluation system has the potential for widespread clinical use both at home and in a hospital setting. Using our method, it is possible to identify where in the gait cycle the abnormality lies and this enables therapists to identify problems to address these in a timely and in a more specific way. In future works, we plan to use our gait evaluation information to classify gait changes over time to identify abnormal gait patterns for the assessment of elderly fall risk, rehabilitation and sports applications.

VI. CONCLUSION

The aims of this study are to propose a novel method of gait evaluation using Procrustes superimposition and to investigate how each feature impacts on a gait using EDMA. We designed and implemented a portable system that can be used in both home and clinics without requiring access to a gait laboratory. Our method is objective and simple. It has three parts: we use 1) Procrustes for shape normalisation, 2) four techniques shown in Table 1 for gait quantification and 3) EDMA for identifying the degree of abnormality shown in Figures 10 and 12. This method also provides information to distinguish young from older gaits taking the full features distribution into account rather than relying on individual parameters such as specific length and time. EDMA can help to estimate and visualize the position of the gait abnormality. Our method offers several advantages: 1) it is easy to set up and implement; 2) it does not require complex equipment with segmentation of body parts required in a gait lab 3) it is relatively inexpensive and therefore increases its affordability decreasing health inequality; and 4) its versatility has the potential to increase its usability at home supporting inclusivity of patients who are home bound. Therefore, our method can help improve the accuracy of assessment and monitor the rehabilitation of patients with mobility problems.

ETHICAL APPROVAL

Ethical approval for this research was granted by the Bournemouth University ethical review committee and each subject was given a Participant Information Sheet and signed an informed Participant Agreement Form.

ACKNOWLEDGEMENT

This work was supported by European Commission ERASMUS MUNDUS FUSION project (545831-EM-1-2013-1-IT-ERAMUNDUSEMA21) and European Commission Marie Skłodowska-Curie SMOOTH (Smart robots for fire-fighting) project (H2020-MSCA-RISE-2016-734875). The authors would like to thank all participants that participated in the study.

VII. REFERENCES

- [1] M. P. Murray, "Gait as a total pattern of movement: including a bibliography on gait," *American Journal of Physical Medicine & Rehabilitation*, vol. 46, no. 1, pp. 290-333, 1967.
- [2] A. Muro-de-la-Herran, B. Garcia-Zapirain, and A. Mendez-Zorrilla, "Gait Analysis Methods: An Overview of Wearable and Non-Wearable Systems, Highlighting Clinical Applications," *Sensors*, vol. 14, no. 2, pp. 3362-3394, 2014-02-19, 2014.
- [3] R. Baker, "Gait analysis methods in rehabilitation," *Journal of NeuroEngineering and Rehabilitation*, vol. 3, no. 1, pp. 4, 2006-03-02, 2006.
- [4] M. D. Lewek, C. E. Bradley, C. J. Wutzke *et al.*, "The relationship between spatiotemporal gait asymmetry and balance in individuals with chronic stroke," *J Appl Biomech*, vol. 30, no. 1, pp. 31-36, Feb, 2014.
- [5] H. Böhm, and L. Döderlein, "Gait asymmetries in children with cerebral palsy: Do they deteriorate with running?," *Gait & Posture*, vol. 35, no. 2, pp. 322-327, 2012.
- [6] M. P. M. Steultjens, J. Dekker, M. E. v. Baar *et al.*, "Range of joint motion and disability in patients with osteoarthritis of the knee or hip," *Rheumatology*, vol. 39, no. 9, pp. 955-961, 2000-09-01, 2000.
- [7] Y. Wahab, and N. A. Bakar, "Gait analysis measurement for sport application based on ultrasonic system." pp. 20-24.
- [8] N. M. Bora, G. V. Molke, and H. R. Munot, "Understanding human gait: A survey of traits for biometrics and biomedical applications," in International Conference on Energy Systems and Applications, 2015, pp. 723-728.
- [9] G. Yogevev, M. Plotnik, C. Peretz *et al.*, "Gait asymmetry in patients with Parkinson's disease and elderly fallers: when does the bilateral coordination of gait require attention?," *Experimental Brain Research*, vol. 177, no. 3, pp. 336-346, 2007.
- [10] M. Plotnik, Y. Dagan, T. Gurevich *et al.*, "Effects of cognitive function on gait and dual tasking abilities in patients with Parkinson's disease suffering from motor response fluctuations," *Experimental Brain Research*, vol. 208, no. 2, pp. 169-179, 2011.
- [11] J. S. Brach, and J. M. VanSwearingen, "Physical Impairment and Disability: Relationship to Performance of Activities of Daily Living in Community-Dwelling Older Men," *Physical Therapy*, vol. 82, no. 8, pp. 752-761, 2002-08-01 00:00:00, 2002.
- [12] R. J. Hess, J. S. Brach, S. R. Piva *et al.*, "Walking Skill Can Be Assessed in Older Adults: Validity of the Figure-of-8 Walk Test," *Physical Therapy*, vol. 90, no. 1, pp. 89-99, 21/04/2010 received 06/09/2010 accepted, 2010.
- [13] K. O. Berg, S. L. Wood-Dauphinee, J. I. Williams *et al.*, "Measuring balance in the elderly: validation of an instrument," *Can J Public Health*, vol. 83, no. 2, pp. 7-11, Jul-Aug, 1992.
- [14] K. R. Archer, R. C. Castillo, E. J. MacKenzie *et al.*, "Gait symmetry and walking speed analysis following lower-extremity trauma," *Physical therapy*, vol. 86, no. 12, pp. 1630, 2006.
- [15] R. W. Kressig, and O. Beauchet, "Guidelines for clinical applications of spatio-temporal gait analysis in older adults," *Aging clinical and experimental research*, vol. 18, no. 2, pp. 174-176, 2006.
- [16] R. A. Zifchock, I. Davis, J. Higginson *et al.*, "The symmetry angle: a novel, robust method of quantifying asymmetry," *Gait & posture*, vol. 27, no. 4, pp. 622-627, 2008.
- [17] M. Błażkiewicz, I. Wiszomirska, and A. Wit, "Comparison of four methods of calculating the symmetry of spatial-temporal parameters of gait," *Acta of bioengineering and biomechanics*, vol. 16, no. 1, pp. 29-35, 2014.
- [18] W. Herzog, B. M. Nigg, L. J. Read *et al.*, "Asymmetries in ground reaction force patterns in normal human gait," *Med Sci Sports Exerc*, vol. 21, no. 1, pp. 110-114, 1989.
- [19] S. J. Crenshaw, and J. G. Richards, "A method for analyzing joint symmetry and normalcy, with an application to analyzing gait," *Gait & Posture*, vol. 24, no. 4, pp. 515-521, 2006/12/01/, 2006.
- [20] A. Sant'Anna, A. Salarian, and N. Wickstrom, "A new measure of movement symmetry in early Parkinson's disease patients using symbolic processing of inertial sensor data," *IEEE Trans Biomed Eng*, vol. 58, no. 7, pp. 2127-35, Jul, 2011.
- [21] H. Sadeghi, P. Allard, F. Prince *et al.*, "Symmetry and limb dominance in able-bodied gait: a review," *Gait & Posture*, vol. 12, no. 1, pp. 34-45, 2000.
- [22] F. P. Carpes, C. B. Mota, and I. E. Faria, "On the bilateral asymmetry during running and cycling – A review considering leg preference," *Physical Therapy in Sport*, vol. 11, no. 4, pp. 136-142, 2010/11/01/, 2010.
- [23] R. Robinson, W. Herzog, and B. Nigg, "Use of force platform variables to quantify the effects of chiropractic manipulation on gait symmetry," *Journal of manipulative and physiological therapeutics*, vol. 10, no. 4, pp. 172-176, 1987.
- [24] C. K. Balasubramanian, R. R. Neptune, and S. A. Kautz, "Variability in spatiotemporal step characteristics and its relationship to walking performance post-stroke," *Gait & posture*, vol. 29, no. 3, pp. 408-414, 2009.
- [25] A. L. Leddy, B. E. Crouner, and G. M. Earhart, "Functional gait assessment and balance evaluation system test: reliability, validity, sensitivity, and specificity for identifying individuals with Parkinson disease who fall," *Physical Therapy*, vol. 91, no. 1, pp. 102, 2011.
- [26] C. Goodall, "Procrustes methods in the statistical analysis of shape," *Journal Royal Statistical Society, Series B-Methodological*, vol. 53, no. 2, pp. 285-339, 1991.
- [27] D. G. Kendall, "Shape manifolds, Procrustean metrics and complex projective spaces," *Bulletin of the London Mathematical Society*, vol. 16, no. 2, pp. 81-121, 1984.
- [28] H. Le, "Mean size-and-shapes and mean shapes: a geometric point of view," *Advances in Applied Probability*, vol. 27, no. 1, pp. 44-55, 2016.
- [29] I. L. Dryden, and K. V. Mardia, *Statistical Shape Analysis*: John Wiley & Sons, Chichester, 1998.
- [30] A. R. Anwary, H. Yu, and M. Vassallo, "Optimal foot location for placing wearable IMU sensors and automatic feature extraction for gait analysis," *IEEE Sensors Journal*, vol. 18, no. 6, pp. 2555-2567, 2018.
- [31] D.-X. Liu, X. Wu, W. Du *et al.*, "Gait Phase Recognition for Lower-Limb Exoskeleton with Only Joint Angular Sensors," *Sensors*, vol. 16, no. 10, pp. 1579, 2016.
- [32] M. Patterson, E. Delahunt, K. Sweeney *et al.*, "An Ambulatory Method of Identifying Anterior Cruciate Ligament Reconstructed Gait Patterns," *Sensors*, vol. 14, no. 1, pp. 887, 2014.
- [33] S. O. Madgwick, "An efficient orientation filter for inertial and inertial/magnetic sensor arrays," *Citado*, vol. 5, pp. 9-19, 2010.
- [34] Y. K. Thong, M. S. Woolfson, J. A. Crowe *et al.*, "Numerical double integration of acceleration measurements in noise," *Measurement*, vol. 36, no. 1, pp. 73-92, 2004.
- [35] A. R. Anwary, H. Yu, and M. Vassallo, "An Automatic Gait Feature Extraction Method for Identifying Gait Asymmetry Using Wearable Sensors," *Sensors*, vol. 18, no. 2, pp. 676, 2018.
- [36] J. T. Richtsmeier, J. Cheverud, and S. R. Lele, "Advances in Anthropological Morphometrics," *Annual Reviews in Anthropology*, vol. 21, pp. 283-305, 1992.
- [37] A. R. Anwary, "Statistical Shape Analysis for the Human Back," 2012.
- [38] M. d. Carmo, *Riemannian Geometry*: Birkhäuser Basel, Englewoods Cliffs, 1992.
- [39] S. Lele, and J. Richtsmeier, "Euclidean distance matrix analysis: a coordinate-free approach for comparing biological shapes using landmark data," *American Journal of Physical Anthropology*, vol. 86, no. 3, pp. 415-427, 1991.
- [40] S. Lele, and J. Richtsmeier, *An Invariant Approach to Statistical Analysis of Shapes*, First ed., London: Chapman and Hall-CRC Press, 2001.
- [41] S. Lele, "Euclidean distance matrix analysis (EDMA): estimation of mean form and mean form difference," *Mathematical Geology*, vol. 25, no. 5, pp. 573-602, 1993.
- [42] R. D. C. Team, "R: A language and environment for statistical computing," R Foundation for Statistical Computing, 2017.
- [43] J. Claude, *Morphometrics with R*, New York: Springer-Verlag, 2008.
- [44] V. Agrawal, R. Gailey, C. O'Toole *et al.*, "Symmetry in External Work (SEW): A Novel Method of Quantifying Gait Differences Between Prosthetic Feet," *Prosthetics and Orthotics International*, vol. 33, no. 2, pp. 148-156, 2009.
- [45] A. R. Anwary, H. Yu, and M. Vassallo, "Wearable sensor based gait asymmetry visualization," in The 24th Americas Conference On Information Systems, AMCIS 2018, New Orleans, LA, USA, 2018.