

Extraction of bodily features for gait recognition and gait attractiveness evaluation

Jie Hong · Jinsheng Kang · Michael E. Price

© Springer Science+Business Media New York 2012

Abstract Although there has been much previous research on which bodily features are most important in gait analysis, the questions of which features should be extracted from gait, and why these features in particular should be extracted, have not been convincingly answered. The primary goal of the study reported here was to take an analytical approach to answering these questions, in the context of identifying the features that are most important for gait recognition and gait attractiveness evaluation. Using precise 3D gait motion data obtained from motion capture, we analyzed the relative motions from different body segments to a root marker (located on the lower back) of 30 males by the fixed root method, and compared them with the original motions without fixing root. Some particular features were obtained by principal component analysis (PCA). The left lower arm, lower legs and hips were identified as important features for gait recognition. For gait attractiveness evaluation, the lower legs were recognized as important features.

Keywords Gait analysis · Gait features · Gait signatures · Gait attractiveness · Principal component analysis · Walking

1 Introduction

Human walking is a simple process but it contains a great deal of information, for example about gender, age, health, and emotion. This gender effect has been studied since as early as the 1970s [20]. A system has been proposed for the early automatic detection of health problems based on the gait of elderly people in their homes [27]. Extracting gait features is a common method in gait analysis [17, 26, 37]. There are different ways to extract features for different purposes, for example, gender recognition, age effect, individual identification, and medical condition analysis. Shoulder-hip ratio and hip rotation are considered important features for detecting gender based on gait [2, 9, 18, 29]. Many features such as step length, speed and double-support time have been analyzed in the gaits of elderly [23, 24]. In the medical application area, gait features usually depend on the disease that is being analyzed, for example, asymmetries in movement patterns were identified as important features in the analysis of developmental coordination disorder in children [30].

J. Hong · J. Kang (✉)
School of Engineering & Design, Brunel University, Uxbridge, Middlesex UB8 3PH, UK
e-mail: jinsheng.kang@brunel.ac.uk

M. E. Price
Department of Psychology, Brunel University, Uxbridge, Middlesex UB8 3PH, UK

Gait recognition is a challenging and active research topic. Many prior studies have focused on extracting gait features to identify individuals or explain individual differences in gait pattern. The difficulty is that the dimensionality of the feature space is much higher than the amount of sample space in the database. There have been no conventional ways to extract gait features up to now. In some reported work, features were extracted by mathematical methods, like principal component analysis (PCA) [1], general tensor discriminant analysis [34], eigen-space transformation with canonical space transformation [14], and wavelet based multi-scale analysis [19]. Some other extracted features were from human body segments, for example, leg angles based on regression analysis were used as gait signature [38]. Hip angle and angular velocity between human walking and passive dynamic walking were studied to compare different wavelet features [28], and seven components (head, arm, trunk, thigh, front-leg, back-leg, and feet) were used as features in silhouette gait recognition [21, 22]. Researchers have investigated soft biometrics, and the relevance to human identification of two novel soft biometric traits, namely weight and color of clothes [10]. Research has shown that PCA combined with LDA (linear discriminant analysis) can increase the accuracy of gait recognition [7], and DCT (discrete cosine transform) can be used for gait pattern classification [15].

Previous work has used various features to analyze gait, to classify subjects into different groups, and even to identify individuals. However, the reasons for choosing these features have received very little attention [28]. Answers to the questions of which features should be extracted, and why these features in particular should be extracted, are still not very clear. The goal of this paper is to provide analytical answers to these questions. Obviously, for different purposes, the answers to these questions should be different.

In this paper, we attempt to provide a solution to these questions by analyzing, via PCA, 3D gait data obtained from motion capture. We analyzed the relative motions from different body segments to root marker via the fixed root method (the root marker is located on the walker's lower back at the upper centre of the pelvis; see Fig. 1a, marker 24). We assumed that the root marker was virtually fixed, almost as if subjects were walking on a treadmill (but not exactly the same), and this helped us to analyze the relative motion of body segments, and compare it with the trajectory of whole body movement without fixed root. Based on the distribution of markers in PCA results, features for gait recognition were identified. We also used PCA and linear regression to identify some particular markers as important features in determining the attractiveness value of gait, and we verified the accuracy of these features.

2 Methods

2.1 Subjects and experimental protocol

Thirty male students at a British university (Mean age=20.83, $SD=3.12$) were recruited to participate in this study. The motion capture volume was 2 meters wide, 4 meters long and 2.2 meters high. Each subject wore a form-fitting motion capture suit, with 40 reflective markers placed as illustrated in Fig. 1a. Subjects were told to walk freely and naturally at normal speed, from one end of the capture volume to the other, and then walk back. The recorded root marker (on the back at the upper middle of pelvis) speed for 30 subjects ranged from 666.16 mm/s to 1255.48 mm/s with a mean of 1005.84 mm/s. The motion capture system used was from Motion Analysis Corporation, USA. Gait motion data were recorded by an Eagle digital system, which was constructed with seven digital cameras, the Eagle Hub, to which all of the cameras were connected and which uplinks to a computer terminal,

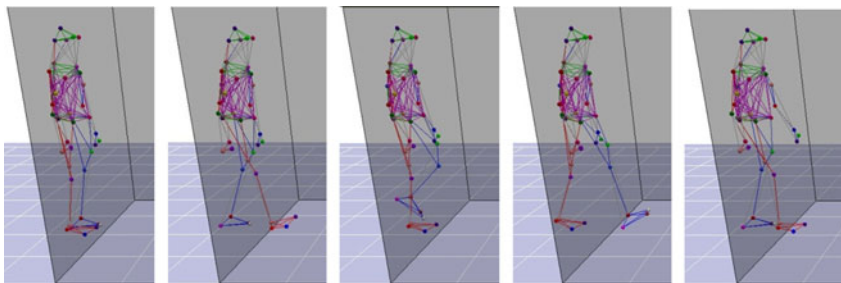
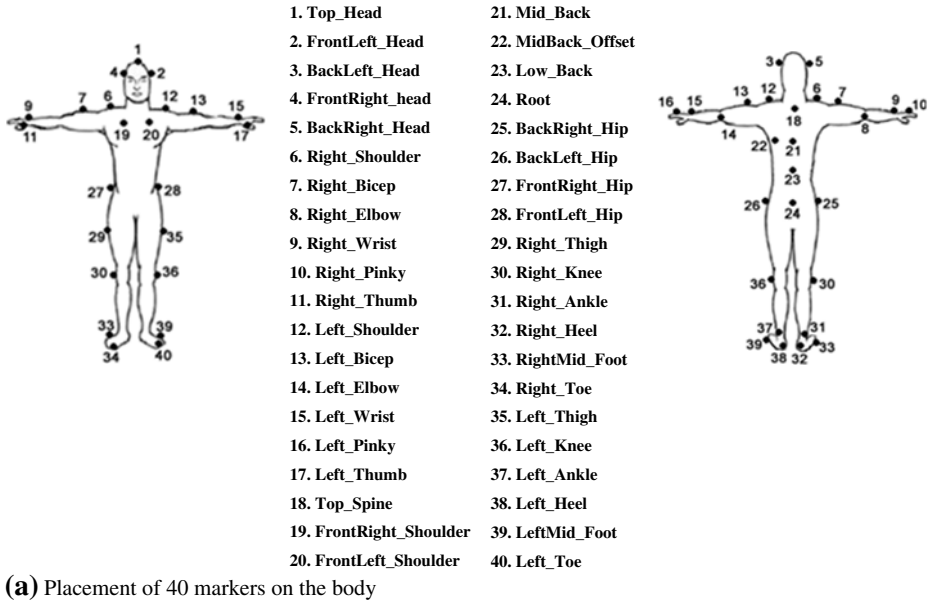


Fig. 1 (a) Placement of 40 Markers on the body (b) An example of a subject with fixed root

and EVaRT Real Time software. This software was used for recording, processing, displaying and post-processing data from the camera system.

2.2 Data collection

Gait data were collected in real time by the motion capture system described in Section 2.1 at a rate of 60 frames per second. The recorded data for each subject were 40 markers' coordinates, with an accuracy of less than 0.1 mm, in x, y and z directions in 3D space at each frame during walking inside the capture volume. These data were saved in the computer terminal as .trc files, and could be played back to show the gait motion video within the EVaRT software as 3D point clouds (to display the markers only), or 3D stick figures (to display the markers and the lines which join related markers together). In this research, the 3D stick figures of gait motion video were presented to evaluators to assess the gait attractiveness of each walker. Walkers were presented in random order. The gait motion video was presented on the computer screen with EVaRT software using a 360 degree rotation feature, so that different viewing angles could be

viewed by evaluators. The evaluators were 32 female students from a British university (Mean age=20.28, SD=3.38). They rated the attractiveness of each gait by drawing a line on a 100 mm scale on which 0 indicated “unattractive” and 100 indicated “attractive”. Since Cronbach’s α (a measure of agreement between raters) was reasonable (0.78), gait attractiveness ratings were averaged as a single value for each walker.

2.3 Data analysis

2.3.1 Fixed root method

One of the novelties of this research is that we adopted a fixed root method to analyse gait motion data. We chose this method for several reasons. First, we wanted to eliminate the influence of walking speed. As will be discussed in section 4.1, walking speed has some correlation with gait attractiveness. However, we wanted to find out apart from speed, which aspects of gait motion (which body segments) contributed most to attractiveness. Second, we wanted to use PCA to analyze the relative motion of each body segment to a specific body point, and this has not been done before, although PCA has been applied as a technique in gait analysis on many occasions. The relative motion of different body segments may provide useful information on gait features, that goes beyond tracking the trajectory of whole body motion; relative motion may illuminate some hidden natural gait features that otherwise would not be revealed. Finally, the root marker is the origin (grand—grand—grandparent) for all other markers placed on different body segments in the motion capture skeleton hierarchy. If we want to study the relative motion of each body segment, root is the one that should be fixed in order to preserve the correct relative motion.

This fixed root method allowed us to determine the motions of different body segments relative to the root marker. We denoted $M_j^i(t : x, y, z)$ as coordinates of number j marker in number i subject in the time of frame number t . Root is the 24th marker for every subject. $M_{24}^i(1 : x, y, z)$ means the initial coordinates of the root marker for number i subject. When we fixed the root marker, we obtained every markers’ new coordinates by the following formula.

$$Mfix_j^i(t : x, y, z) = M_j^i(t : x, y, z) - (M_{24}^i(t : x, y, z) - M_{24}^i(1 : x, y, z))$$

$$i = 1, \dots, 30; j = 1, \dots, 40; t = 1, \dots, tend$$

After obtaining the new coordinates of markers, we calculated speed and acceleration of all 39 markers except root for every frame. Then we averaged the speed and acceleration of all frames for every marker for each subject. We denoted them as \overline{Msf}_j^i and \overline{Maccf}_j^i , i for subject number, and j for marker number. Then we got two matrixes as follows:

$$\overline{Msf} = \begin{bmatrix} \overline{Msf}_1^1 & \overline{Msf}_2^1 & \dots & \overline{Msf}_{40}^1 \\ \overline{Msf}_1^2 & \overline{Msf}_2^2 & \dots & \overline{Msf}_{40}^2 \\ \dots & \dots & \dots & \dots \\ \overline{Msf}_1^{30} & \overline{Msf}_2^{30} & \dots & \overline{Msf}_{40}^{30} \end{bmatrix}_{30 \times 39}$$

$$\overline{Maccf} = \begin{bmatrix} \overline{Maccf}_1^1 & \overline{Maccf}_2^1 & \dots & \overline{Maccf}_{40}^1 \\ \overline{Maccf}_1^2 & \overline{Maccf}_2^2 & \dots & \overline{Maccf}_{40}^2 \\ \dots & \dots & \dots & \dots \\ \overline{Maccf}_1^{30} & \overline{Maccf}_2^{30} & \dots & \overline{Maccf}_{40}^{30} \end{bmatrix}_{30 \times 39}$$

\overline{Msf} and \overline{Maccf} are the average speed and average acceleration matrixes of 39 markers (except root) for the 30 subjects. Now we only have 39 markers instead of 40 markers. The 24th marker (root) has been removed because its displacement, speed and acceleration were all zero. An example of subject's gait after fixing the root is shown in Fig. 1b. We also have a data set which contains every subject's gait attractiveness average rating,

$$attract = [attract^1, attract^2, \dots, attract^{30}]^T_{30 \times 1}$$

The original average speed and average acceleration matrixes of all 40 markers for 30 subjects (without fixed root) were as follows, and designated as \overline{Ms} and \overline{Macc} .

$$\overline{Ms} = \begin{bmatrix} \overline{Ms}_1^1 & \overline{Ms}_2^1 & \dots & \overline{Ms}_{24}^1 & \dots & \overline{Ms}_{40}^1 \\ \overline{Ms}_1^2 & \overline{Ms}_2^2 & \dots & \overline{Ms}_{24}^2 & \dots & \overline{Ms}_{40}^2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \overline{Ms}_1^{30} & \overline{Ms}_2^{30} & \dots & \overline{Ms}_{24}^{30} & \dots & \overline{Ms}_{40}^{30} \end{bmatrix}_{30 \times 40}$$

$$\overline{Macc} = \begin{bmatrix} \overline{Macc}_1^1 & \overline{Macc}_2^1 & \dots & \overline{Macc}_{24}^1 & \dots & \overline{Macc}_{40}^1 \\ \overline{Macc}_1^2 & \overline{Macc}_2^2 & \dots & \overline{Macc}_{24}^2 & \dots & \overline{Macc}_{40}^2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \overline{Macc}_1^{30} & \overline{Macc}_2^{30} & \dots & \overline{Macc}_{24}^{30} & \dots & \overline{Macc}_{40}^{30} \end{bmatrix}_{30 \times 40}$$

2.3.2 PCA and linear regression

First, we conducted PCA analysis on 30 subjects to find out which markers are important as features in gait. We analyzed which markers should be extracted as features for gait recognition based on PCA results. Then we calculated the related coefficients between attractiveness $attract$ and markers' speed \overline{Msf} , and between attractiveness $attract$ and markers' acceleration \overline{Maccf} .

Second, we used natural logarithm of extracted principal components as independent variables and natural logarithm of attractiveness value as a dependent variable, to build a linear regression equation which can predict a subject's attractiveness value. In this part, two different methods were applied. The first method applied linear regression on the natural logarithm of attractiveness value and the natural logarithm of extracted principal components, which was based on matrix \overline{Msf} with fixed root. The other method applied linear regression on the natural logarithm of attractiveness value and the natural logarithm of extracted principal components, which was based on matrix \overline{Ms} without fixed root.

To verify the linear regression results, five subjects were randomly picked out for verification, and the other 25 subjects consisted of the sample database. The accuracy of these linear regression results was assessed by comparing the attractiveness value computed from linear regression equation and the real value. We repeated this random verification procedure eight times to investigate whether there is a systematic relationship between \overline{Ms} and attractiveness value in gait. After that, we analyzed which markers should be extracted as features for gait attractiveness. Finally, we compared the accuracy of predicting attractiveness values using all markers, versus using extracted feature markers only.

3 Results

3.1 Principal component analysis

We used PCA to find the feature markers among the \overline{Msf} matrix. Based on markers' speed \overline{Msf} after fixing the root, we obtained seven principal components. There were three principal components that each accounted for over 10 % of the total variance, and these first three principal components together accounted for 67 % of the total variance.

We compared the ten highest coefficients of the first three principal components, and they were listed as PC1, PC2 and PC3 (Table 1). Markers with the highest coefficients in PC1 are clearly concentrated around the left lower arm (pinky, wrist and thumb). In the ten highest coefficients of PC2, markers are concentrated around the lower legs, and included all markers on lower legs. In the ten highest coefficients of PC3, all four hip markers appear at the top of the list. This highly concentrated distribution of markers around different body parts with each principal component was extraordinary and very interesting. We also tried PCA on these gait data without fixing the root marker (\overline{Ms}) and found that the distribution was not as concentrated as that which is shown in Table 1. Fixing the root marker seemed to be an effective method of identifying the bodily features that are most important in gait. This PCA analysis was carried out only with respect to subjects' gait data, and it suggested that motions from the left arm, lower legs and hips are important variables for representing overall gait, and can be usefully extracted as features for gait recognition.

3.2 Related coefficients

We calculated related coefficients between attractiveness ratings *attract* and marker speed matrix \overline{Msf} , and between attractiveness ratings *attract* and marker acceleration matrix \overline{Maccf} respectively. The results for the top ten high valued markers are listed in Table 2. The average related coefficient of speed and attractiveness is 0.349, the maximum is left_heel (0.688), and the minimum is right_shoulder (0.056). The average related coefficient of acceleration and attractiveness is 0.193, the maximum is left_wrist (0.451), and the minimum is low_back (0.002). It is obvious that marker speed is more related to attractiveness than is marker acceleration. Since the highest related coefficient between acceleration and attractiveness is below 0.5, we will not consider marker acceleration in the following analytical sessions. From the left side of Table 2, it can be seen that markers with higher related coefficients concentrated on lower legs, from knee to

Table 1 Ten markers with highest coefficients in PC1, PC2 and PC3 with fixed root

PC1		PC2		PC3	
Left_Thumb	0.845	RightMid_Foot	0.769	BackRight_Hip	0.776
Left_Wrist	0.82	Left_Toe	0.722	FrontLeft_Hip	0.753
Left_Pinky	0.801	Right_Ankle	0.72	FrontRight_Hip	0.688
BackRight_Head	0.799	Right_Heel	0.714	BackLeft_Hip	0.670
Right_Bicep	0.79	LeftMid_Foot	0.684	MidBack_Offset	0.628
BackLeft_Head	0.772	Left_Heel	0.666	Right_Thigh	0.305
FrontLeft_Shoulder	0.768	Left_Ankle	0.664	Low_Back	0.297
Left_Bicep	0.766	Right_Toe	0.635	Left_Elbow	0.177
Mid_Back	0.766	Left_Knee	0.592	Left_Toe	0.174
Top_Spine	0.753	Right_Knee	0.558	LeftMid_Foot	0.162

Table 2 Ten markers with highest related coefficients between speed/acceleration and attractiveness

Speed and attractiveness		Acceleration and attractiveness	
Left_Heel	0.688	Left_Wrist	0.451
LeftMid_Foot	0.662	Left_Toe	0.413
Left_Toe	0.662	Right_Thumb	0.411
Left_Ankle	0.658	Left_Knee	0.394
Left_Knee	0.654	Left_Ankle	0.368
Right_Toe	0.643	Right_Knee	0.354
Right_Knee	0.616	LeftMid_Foot	0.340
RightMid_Foot	0.571	Left_Heel	0.309
Left_Pinky	0.515	Left_Pinky	0.304
Right_Ankle	0.484	Left_Thumb	0.300

toe. This is similar to the distribution of markers on PC2’s coefficients (see the middle columns of Table 1).

3.3 Linear regression

3.3.1 Linear regression based on extracted principal components of \overline{Msf} with fixed root

In this section, we used 39 markers with fixed root in matrix \overline{Msf} and extracted seven principal components which occupied 89 % of the total variance. Linear regression was applied on the natural logarithm of attractiveness value and the natural logarithm of these seven principal components. One example of enter method linear regression is below:

$$Ln(attract) = 0.758Ln(PC1) + 0.239Ln(PC2) - 0.004Ln(PC3) + 0.006Ln(PC4) + 0.034Ln(PC5) - 0.096Ln(PC)6 - 0.005Ln(PC7) - 5.128 \tag{1}$$

We left out randomly five subjects to consist of the testing database, and used the other 25 subjects as the sample database to obtain a linear regression equation similar to the above. We repeated this process eight times. Three times we could not obtain effective linear regression results. Three times we obtained good linear regression results by the stepwise method, shown in Eqs. (2) and (4) as follows.

$$Ln(attract) = 0.851Ln(PC2) - 3.762 \tag{2}$$

(Average error in sample database is 8.35 %, in testing database is 5.57 %.)

$$Ln(attract) = 0.746Ln(PC2) - 2.878 \tag{3}$$

(Average error in sample database is 10.03 %, in testing database is 7.57 %.)

$$Ln(attract) = 0.694Ln(PC2) - 2.400 \tag{4}$$

(Average error in sample database is 9.68 %, in testing database is 9.97 %.)

On all the other occasions, we still obtained linear regression, but errors in the testing database were above 15 %. There were no stable linear relationships between the natural logarithm of principal components and the natural logarithm of attractiveness value after fixing the root marker. On the other hand, although regression results were not good, they still provided some useful clues. When using the stepwise method, all the good linear regression equations were related to PC2 only. This suggested that PC2 might be highly

related to attractiveness. It also provided some explanation about why the ten markers with the highest coefficients of PC2 (Table 1, middle columns) are similar to markers in related coefficients with attractiveness (Table 2, left side). These results strongly suggested that lower legs might be extracted as features for gait attractiveness.

3.3.2 Linear regression based on extracted principal components of \overline{Ms} without fixed root

The only difference between this section and the previous section is that in section 3.3.1 we used 39 markers with fixed root in matrix \overline{Msf} and in this section we used 40 markers without fixed root in matrix \overline{Ms} . Two principal components were extracted, which accounted for 95.23 % of total variance. Based on these two principal components, we carried out the linear regression analysis to produce a linear expression of attractiveness value. The linear relationship between $\text{Ln}(PC1)$, $\text{Ln}(PC2)$ and $\text{Ln}(\text{attract})$ was highly significant, with the P (probability) value of regression below 0.001. The regression equation can be expressed as follows,

$$\text{Ln}(\text{attract}) = 0.829\text{Ln}(PC1) + 0.003\text{Ln}(PC2) - 5.044 \quad (5)$$

Next, we verified eight times by swapping the sample database and verification subjects in order to test the robustness of this predicting method. Each time, five subjects were randomly picked out for verification, and the other 25 subjects composed the sample database. After repeating this random verification procedure eight times, we found the regression results to be very similar. Each time, we first extracted two principal components (only two were produced) from the 40 marker speed matrix. The eigenvalues were close to those of the original principal components, and the percentages of total variance explained by these two components were all above 90 %. These results suggest that the markers have stable patterns regardless of sample differences. The resulting linear regression Eqs. (6)–(13) are very similar to each other as well as Eq. (5).

$$\text{Ln}(\text{attract}) = 0.879\text{Ln}(PC1) + 0.002\text{Ln}(PC2) - 5.525 \quad (6)$$

$$\text{Ln}(\text{attract}) = 0.875\text{Ln}(PC1) - 0.002\text{Ln}(PC2) - 5.520 \quad (7)$$

$$\text{Ln}(\text{attract}) = 0.815\text{Ln}(PC1) + 0.003\text{Ln}(PC2) - 4.892 \quad (8)$$

$$\text{Ln}(\text{attract}) = 0.862\text{Ln}(PC1) + 0.001\text{Ln}(PC2) - 5.391 \quad (9)$$

$$\text{Ln}(\text{attract}) = 0.891\text{Ln}(PC1) + 0.003\text{Ln}(PC2) - 5.677 \quad (10)$$

$$\text{Ln}(\text{attract}) = 0.833\text{Ln}(PC1) + 0.006\text{Ln}(PC2) - 5.081 \quad (11)$$

$$\text{Ln}(\text{attract}) = 0.850\text{Ln}(PC1) - 0.001\text{Ln}(PC2) - 5.244 \quad (12)$$

$$\text{Ln}(\text{attract}) = 0.802\text{Ln}(PC1) + 0.005\text{Ln}(PC2) - 4.756 \quad (13)$$

The linear regression results are much better than the ones using matrix \overline{Msf} with fixed root. In this case, we will apply PCA on \overline{Ms} without fixed root, and then use this type of linear regression equation to predict gait attractiveness values.

3.4 Verification of lower legs as features for gait attractiveness

To verify the correlation between lower leg motion and attractiveness values suggested in the previous sections (3.1 and 3.2), we compared the accuracy of predicting attractiveness values from the motions of all 40 markers as opposed to just ten markers from around the lower legs only. These ten markers were R/L knee, R/L ankle, R/L heel, R/L toe, and R/L mid_foot. The only difference is that we used ten markers on the lower legs without fixing the root marker in matrix $\overline{Ms10}$ in this section, whereas we used 40 markers without fixing the root marker in section 3.3.2. This time, we still extracted two principal components which accounted for over 97 % of the total variance, and we then used linear regression on the natural logarithm of these two principal components and the natural logarithm of gait attractiveness. The resulting squared multiple correlation coefficient was 0.546, and the Std error of the estimation was 0.115, so the regression equation was acceptable. The linear relationship between $\ln(PC1)$, $\ln(PC2)$ and $\ln(\text{attract})$ was highly significant, with the P (probability) value of the regression below 0.001. One example of the regression equation is shown below.

$$\ln(\text{attract}) = 0.794\ln(PC1) - 0.003\ln(PC2) - 3.507 \quad (14)$$

To test the robustness of the above regression equation, and make a comparison with the results in section 3.3.2, we verified the equation eight times. Each time, we left out randomly five subjects to constitute the testing database, and used the other 25 subjects as the sample database. Each time, the resulting linear regression equation was very similar to equation (14). These results suggested that the lower leg markers have stable patterns with gait attractiveness. We compared the results of using lower leg markers with the results of using 40 markers from the whole body, and they are listed in Table 3.

The left part of Table 3 shows the results of using all 40 markers, and the right part shows the results of using ten markers around the lower legs only. These results show that the error in the testing database was smaller using only lower leg markers, than using 40 markers, for every verification. The average error in predicting attractiveness was only 7.81 % when only

Table 3 Comparison of regression with all 40 markers and regression with only 10 lower leg markers

Verify time	Regression with 40 markers		Regression with only leg markers	
	Error in testing database	Error in sample database	Error in testing database	Error in sample database
1	7.76 %	9.09 %	7.49 %	9.08 %
2	5.45 %	8.62 %	5.48 %	8.63 %
3	9.24 %	8.58 %	8.71 %	8.83 %
4	8.09 %	8.99 %	7.16 %	9.15 %
5	6.89 %	9.24 %	5.74 %	9.43 %
6	10.98 %	8.20 %	9.59 %	8.66 %
7	10.89 %	8.80 %	8.98 %	8.76 %
8	9.38 %	8.49 %	9.30 %	8.72 %
Average	8.58 %	8.75 %	7.81 %	8.91 %

using leg markers. These results suggest that using lower leg markers as gait attractiveness features is adequate.

4 Discussion and conclusion

4.1 Features for gait recognition

In principal component analysis, the most important criteria for determining the number of components to retain is the interpretability criteria. These criteria evaluate whether the variables in a component share the same conceptual meaning, whether variables in different components seem to be measuring different constructs, and if the results demonstrate a “simple” structure (which means that most variables have relatively large coefficients only for one component, and that most components have relatively large coefficients on some variables and small coefficients for remaining variables). In this research, we fixed the root marker and used PCA to investigate relative motion from different body segments, in order to reveal gait features. Our PCA results meet the above criteria very well. In PC1, the three variables with the largest coefficients were concentrated on the lower left arm. In PC2, all the ten variables based on lower legs and feet were in the list of the top ten largest coefficients. In PC3, all the four variables based on the hips had the largest coefficients, followed by the variable of *MidBack_Offset*; all remaining variables had coefficients of 0.305 or less and can be ignored. The PCA results in this study provided a simplified structure to reveal the most important features/characteristics for gait analysis.

PCA has been used for gait analysis in many occasions, for example in [1, 6, 25, 31], but to our knowledge, no previous work has applied PCA to study the relative motion of all body segments to a specific body point. Although previous research has applied PCA to gait on a treadmill [11], the normal gait for people naturally walking on the ground is different from walking on a treadmill. The gait we studied is the absolute relative motion which cannot be achieved by walking on a treadmill because there is no absolutely fixed point on the body. Through the fixed root method, the influence of walking speed was removed, and gait data were fully focused on the relative movement of each body segment, which helped illuminate natural gait features.

Much previous research on gait feature extraction has been based on video images, with features for gait recognition usually based on silhouette movement [4], for example, using moving shapes to get a sequence of silhouettes of walking subjects [13]. Using 3D motion capture, detailed gait data about body segments' movement and rotation can be obtained, for example, hip-knee angles have been used as features for gait recognition [3, 8], and hip flexion in swing and lower limb joint angles have been studied [35]. Movements from legs were identified as core features for gait recognition [11]. Movements from arms have received more attention recently, for example, swinging arm regions have been used for gait phase detection [36], the effect of arm swing on the local and global stability of steady-state gait has been studied [5], and through extra features produced from the motion of the arms, the discrimination capability of gait recognition has been considerably increased [33]. These features used in previous research are consistent with our findings based on PCA and the fixed root method. Furthermore, our findings provide reasons for choosing the motion of the left lower arm, lower legs and feet, and hips as features for gait recognition.

It is a surprising finding that motion from the left lower arm was identified as a predominant gait feature but motion from the right lower arm was not. The three markers associated with the left lower arm appeared at the top of the list of PC1 coefficients (Table 1). This is not biased by the starting posture of gait motion, since the gait motion was captured randomly (the subjects started walking in different postures, before they entered the motion

capture volume), and different subjects had different starting postures. In an associated study, we normalized the gait data and found that for most subjects, their two arms swung at different amplitude and speed, with one arm as leading swing arm, and the other one as a complementary follower (Hong et al., in preparation). The leading arm, with higher swing amplitude and higher speed, turned out to be the left one in most cases. Identification of the lower arm as a predominant gait feature is a new, interesting finding, which we wish to bring to the attention of researchers in the fields of gait recognition/identification, psychology, and physiology, for further verification and discussion.

4.2 Features related to gait attractiveness

Gait attractiveness is a fascinating issue. Although humans probably constantly perceive and evaluate, whether consciously or not, the gait attractiveness of others, the factors which influence gait attractiveness are not well-understood. In the field of psychology, research suggests that males with a higher social status tend to walk faster [16, 32]. If high status men walk faster, then it follows that faster male gaits should be more attractive to females because social status is one of the most important aspects of what makes a male attractive to females [12]. To verify this, we calculated the correlation coefficient between attractiveness ratings and subjects' average walking speeds (represented by the average speed of root marker for each subject). This coefficient was high and positive ($r=0.724$), indicating that female evaluators did prefer fast gait over slow gait.

In this research, we wanted to investigate gait attractiveness deeply to find out which specific body segment motions make one gait more attractive than another, apart from general walking speed. One possible approach to doing this would have been to normalise the gait data, that is, to make all gait data starting from the same posture and finishing at the same posture, in a complete gait cycle with the same number of frames after interpolation. If normalised gait were presented to evaluators, this approach would be appropriate for further investigation. In our case, the gait motion presented to evaluators was the original, and we employed a novel fixed root method to reveal the relative motion of each body segment. Thus the original relative body segment motion, as the evaluators rated it, was preserved. We applied PCA and linear regression on the original gait data after fixing the root markers.

By PCA and linear regression methods, it was found that PC2 is highly related to gait attractiveness. The ten markers with the highest coefficients of PC2 included all ten markers on the lower legs and feet. This is consistent with the related coefficients between marker speed and gait attractiveness (Table 2, left side). Of the listed ten markers with the highest related coefficients, nine of them are from the ten markers on the lower legs and feet. The results from PCA and linear regression, coupled with related coefficients, strongly suggest that the lower legs and feet could be extracted as features for gait attractiveness.

To verify this, we compared the effectiveness of predicting attractiveness by using only lower legs and feet markers as opposed to using all 40 markers in linear regression. This comparative analysis showed that attractiveness could be predicted slightly better by only using legs and feet markers than by using all 40 markers. This means that instead of using 40 markers, ten markers from the lower legs and feet can be used to fully represent and predict attractiveness values. The relationship between the movement from the lower legs and feet and attractiveness could not have been revealed without using the fixed-root method.

4.3 Conclusion

The goal of this research was to provide non-subjective, analytical solutions to the problems of which gait features should be extracted, and why those features in particular should be

extracted, using accurate 3D gait data obtained from motion capture. A novel fixed root method was employed to reveal the hidden relative motion of each body segment in gait, and then PCA and linear regression were applied to identify the significant gait features. It was found that lower legs are significant features in the evaluation and prediction of gait attractiveness, and that the left lower arm, lower legs and hips could be important features in gait recognition. In addition, this research produced the surprising new finding that the motion of the left lower arm, but not of the right lower arm, constitutes a predominant gait feature. We think that these findings will be of interest to gait researchers in a variety of fields, and hope that they will be subjected to further verification and discussion in future research.

Acknowledgments We thank the Dorothy Hodgkin Postgraduate Award to Jie Hong and HEFCE SRIF2 project BRUN 07/033 funding for motion capture system.

References

1. Arantes M, Gonzaga A (2011) Human gait recognition using extraction and fusion of global motion features. *Multimed Tool Appl* 55(3):655–675
2. Barclay CD, Cutting JE, Kozlowski LT (1978) Temporal and spatial actors in gait perception that influence gender recognition. *Percept Psychophys* 23(2):145–152
3. Barton JG, Lees A (1997) An application of neural networks for distinguishing gait patterns on the basis of hip-knee joint angle diagrams. *Gait Posture* 5(1):28–33
4. Boulgouris NV, Hatzinakos D, Plataniotis KN (2005) Gait recognition: a challenging signal processing technology for biometric identification. *Signal Process Mag IEEE* 22(6):78–90
5. Bruijn SM, Meijer OG, Beek PJ, Van Dieën JH (2010) The effects of arm swing on human gait stability. *J Exp Biol* 213(23):3945–3952
6. Carriero A, Zavatsky A, Stebbins J, Theologis T, Shefelbine SJ (2009) Determination of gait patterns in children with spastic diplegic cerebral palsy using principal components. *Gait Posture* 29(1):71–75
7. Cho C, Chao W, Lin S, Chen Y (2009) A vision-based analysis system for gait recognition in patients with Parkinson's disease. *Expert Syst Appl* 36(3, Part 2):7033–7039
8. Cunado D, Nixon MS, Carter JN (2003) Automatic extraction and description of human gait models for recognition purposes. *Comp Vision Image Underst* 90(1):1–41
9. Cutting JE, Proffitt DR, Kozlowski LT (1978) A biomechanical invariant for gait perception. *J Exp Psychol* 4(3):357–372
10. Dantcheva A, Velardo C, D'Angelo A, Dugelay J-L (2011) Bag of soft biometrics for person identification. *Multimed Tool Appl* 51(2):739–777
11. Das SR, Wilson RC, Lazarewicz MT, Finkel LH (2006) Two-stage PCA extracts spatiotemporal features for gait recognition. *J Multimed* 1(5):9–17
12. Davies APC, Shackelford TK (2008) Two human natures: how men and women evolved different psychologies. In: Crawford CB, Krebs DE (eds) *Foundations of evolutionary psychology*. Lawrence Erlbaum, New York, pp 261–280
13. Foster JP, Nixon MS, Prugel-Bennett A (2003) Automatic gait recognition using area-based metrics. *Pattern Recogn Lett* 24(14):2489–2497
14. Huang PS, Harris CJ, Nixon MS (1999) Human gait recognition in canonical space using temporal templates. *IEEE Vis Image Signal Process* 146(2):93–100
15. Ibrahim RK, Ambikairajah E, Celler BG, Lovell NH (2008) Gait pattern classification using compact features extracted from intrinsic mode functions. *Engineering in Medicine and Biology Society, EMBS 2008. 30th Annual International Conference of the IEEE: Vancouver, BC: 3852–3855*
16. Jahoda M, Lazarsfeld P, Zeisel H (1933) *Die Arbeitslosen von Marienthal*. S. Hirzel, Leipzig
17. Ji X, Liu H (2010) Advances in view-invariant human motion analysis: a review. *IEEE Trans Systems Man Cybern Part C Appl Rev* 40(1):13–24
18. Johnson KL, Tassinari LG (2005) Perceiving sex directly and indirectly - meaning in motion and morphology. *Psychol Sci* 16(11):890–897

19. Khandoker AH, Lai DTH, Begg RK, Palaniswami M (2007) Wavelet-based feature extraction for support vector machines for screening balance impairments in the elderly. *Neural Syst Rehabil Eng IEEE Trans* 15 (4):587–597
20. Kozlowski LT, Cutting JE (1977) Recognizing the sex of a walker from dynamic point-light display. *Percept Psychophys* 21(6):575–580
21. Li X, Maybank SJ, Tao D (2007) Gender recognition based on local body motions. *Systems, Man and Cybernetics, ISIC. IEEE International Conference: Montreal, Que.* 3881–3886
22. Li X, Maybank SJ, Yan S, Tao D, Xu D (2008) Gait components and their application to gender recognition. *Syst Man Cybern Part C Appl Rev IEEE Trans* 38(2):145–155
23. Menant JC, Steele JR, Menz HB, Munro BJ, Lord SR (2009) Effects of walking surfaces and footwear on temporo-spatial gait parameters in young and older people. *Gait Posture* 29(3):392–397
24. Menant JC, Steele JR, Menz HB, Munro BJ, Lord SR (2009) Rapid gait termination: effects of age, walking surfaces and footwear characteristics. *Gait Posture* 30(1):65–70
25. Muniz AMS, Nadal J (2009) Application of principal component analysis in vertical ground reaction force to discriminate normal and abnormal gait. *Gait Posture* 29(1):31–35
26. Nixon MS, Carter JN (2004) Advances in automatic gait recognition. *IEEE Face and Gesture Analysis 2004 (FG'04): Seoul Korea* 139–144
27. Pogorelc B, Bosnić Z, Gams M (2012) Automatic recognition of gait-related health problems in the elderly using machine learning. *Multimed Tool Appl* 58(2):333–354
28. Preece SJ, Goulermas JY, Kenney LPJ, Howard D (2009) A comparison of feature extraction methods for the classification of dynamic activities from accelerometer data. *Biomed Eng IEEE Trans* 56(3):871–879
29. Roislien J, Skare Ø, Gustavsen M, Broch NL, Rennie L, Opheim A (2009) Simultaneous estimation of effects of gender, age and walking speed on kinematic gait data. *Gait Posture* 30(4):441–445
30. Rosengren KS, Deconinck FJA, DiBerardino LA III, Polk JD, Spencer-Smith J, De Clercq D, Lenoir M (2009) Differences in gait complexity and variability between children with and without Developmental Coordination Disorder. *Gait Posture* 29(2):225–229
31. Samantha MR, Ryan BG, Patrick AC (2010) Differentiation of young and older adult stair climbing gait using principal component analysis. *Gait Posture* 31(2):197–203
32. Schmitt A, Atzwanger K (1995) Walking fast-ranking high: a sociobiological perspective on pace. *Ethol Sociobiol* 16:451–462
33. Tafazzoli F, Safabakhsh R (2010) Model-based human gait recognition using leg and arm movements. *Eng Appl Artif Intel* 23(8):1237–1246
34. Tao D, Li X, Wu X, Maybank SJ (2007) General tensor discriminant analysis and Gabor features for gait recognition. *Pattern Anal Mach Intell IEEE Trans* 29(10):1700–1715
35. Vrieling AH, van Keeken HG, Schoppen T, Otten E, Halbertsma JPK, Hof AL, Postema K (2008) Uphill and downhill walking in unilateral lower limb amputees. *Gait Posture* 28(2):235–242
36. Wang K, Ben X, Zhao Y (2009) Gait period detection based on regional characteristics analysis. *Pattern recognition, 2009. CCPR 2009. Chinese conference on:* 1–6
37. Wang L, Hu W, Tan T (2003) Recent developments in human motion analysis. *Pattern Recogn* 36(3):585–601
38. Yoo JH, Nixon MS, Harris CJ (2002) Extracting human gait signatures by body segment properties. *Fifth IEEE Southwest Symposium on Image Analysis and Interpretation:* 35–39



Ms Jie Hong is a PhD candidate at the School of Engineering and Design, Brunel University. She obtained her BSc in Applied Mathematics in 2000, and MSc in Mathematics in 2003, both from Zhejiang University, China. She was a Research Assistant at the National Research Center for Science and Technology, Beijing, China from 2003 to 2006. She started her PhD research in Brunel University, UK in 2008 under the Dorothy Hodgkin Postgraduate Award. Her research was focused on individual identification from gait based on 3D motion capture techniques.



Dr Jinsheng Kang is a Senior Lecturer at the School of Engineering and Design, Brunel University, UK. He is a Chartered Engineer (CEng), Member of IMechE, and Fellow in The Higher Education Academy. He obtained his BEng degree in Mechanical and Manufacturing Engineering in China in 1982 and his PhD degree in Mechanical Engineering in the UK in 2001. After completing his PhD, he was a Postdoctoral Research Fellow in Bournemouth University from 2001 to 2003 before he joined Brunel University. He was an Assistant Lecturer (1982–1986), Lecturer (1988–1993), Associate Professor (1993–1996) in the Department of Mechanical Engineering, Xinjiang Institute of Technology (Now, School of Mechanical Engineering, Xinjiang University) in China. He was an Academic Visiting Scholar in the Department of Mechanical Engineering, University of Washington, Seattle, USA from 1986 to 1988, in the Department of Manufacturing Engineering, Loughborough University, UK from 1996 to 1997, and in the Center for Human Modelling and Simulation (HMS), University of Pennsylvania, USA in summers 2008 and 2009. His research interest is currently focused on human modelling and simulation.



Dr Michael E. Price has a BA (1995) in psychology from Duke University and a PhD (2003) in biosocial anthropology from the UC Santa Barbara Center for Evolutionary Psychology. His PhD dissertation focused mainly on group cooperation among indigenous Amazonian hunter-horticulturalists. His first job was a postdoctoral fellowship sponsored jointly by the Indiana University Workshop in Political Theory and Policy Analysis and the Santa Fe Institute, and he was then appointed Visiting Lecturer in Organizational Behavior at Washington University in St. Louis. Since 2006 he has been Lecturer in Psychology at Brunel University, where he is now co-Director of the Centre for Culture and Evolutionary Psychology and Deputy Head of Department. His current research focuses on various areas of evolutionary social psychology.