View Invariant Gait Recognition

Richard D. Seely, Michela Goffredo, John N. Carter and Mark S. Nixon

Abstract Recognition by gait is of particular interest since it is the biometric that is available at the lowest resolution, or when other biometrics are (intentionally) obscured. Gait as a biometric has now shown increasing recognition capability. There are many approaches and these show that recognition can achieve excellent performance on large databases. The majority of these approaches are planar 2D, largely since the early large databases featured subjects walking in a plane normal to the camera view. To extend deployment capability, we need viewpoint invariant gait biometrics. We describe approaches where viewpoint invariance is achieved by 3D approaches or in 2D. In the first group the identification relies on parameters extracted from the 3D body deformation during walking. These methods use several video cameras and the 3D reconstruction is achieved after a camera calibration process. On the other hand, the 2D gait biometric approaches use a single camera, usually positioned perpendicular to the subject's walking direction. Because in real surveillance scenarios a system that operates in an unconstrained environment is necessary, many of the recent gait analysis approaches are orientated towards view-invariant gait recognition.

Richard D. Seely
University of Southampton, United Kingdom, e-mail: rds06r@ecs.soton.ac.uk

Michela Goffredo
University of Southampton, United Kingdom, e-mail: mg2@ecs.soton.ac.uk

John N. Carter
University of Southampton, United Kingdom, e-mail: jnc@ecs.soton.ac.uk

Mark S. Nixon
University of Southampton, United Kingdom, e-mail: msn@ecs.soton.ac.uk

1 Introduction

Much research has been done into identifying subjects by how they walk from two-dimensional video data[30]; with publications dating as far back as 1994 [31]. Gait has even found forensic use, securing the conviction of a bank robber based upon his gait[23].

Typically gait analysis techniques can described as either model based or appearance based. Model based analysis usually involves fitting a model representing various aspects of the human anatomy to the video data then extracting and analysing its parameters. Appearance based analysis often involves the analysis of a subject's silhouette shape and how it varies over time, or analysis can be carried out in a more direct fashion; considering the statistical distribution of pixels in the silhouette and how it varies over time.

The DARPA HumanID project spurred the creation of several large gait datasets[34, 41], each having over 100 subjects and a range of covariates such as different surfaces and footwear. Researchers were able to achieve extremely promising recognition rates on both datasets[46, 38], providing further proof that gait is a valid and usable biometric for identification purposes.

Most of the current approaches to gait biometrics process silhouettes which are obtained from images derived when a subject walks in a plane normal to the camera view. This derives a recognition metric which incorporates body shape and motion, but the process is largely dependent on the viewpoint. There is a much smaller selection of model-based approaches and these have some small angle invariance. There have been studies which have been aimed to improve viewpoint invariance [6, 23] using geometry considerations. In this Chapter we describe 2D view invariant approaches which use geometry considerations to derive a viewpoint invariant signature; we also describe developments in 3D approaches which are the complementary approach which has implicit viewpoint independence since arbitrary views can be synthesised from the 3D human figure.

2 2D View Invariant Gait Recognition

2.1 Overview

Markerless 2D view independent gait identification is a recent research area and the approaches found in literature can be can be broadly divided as pose-based and pose-free methods. The pose-based approaches aim at synthesizing the lateral view of the human body from any other arbitrary views [13, 19, 21]. On the other hand, the pose-free ones extract some gait parameters which are independent from the human pose [2, 7, 16, 18, 49, 50].

2.2 Pose-based methods

The methods which generate the lateral view from data acquired at different arbitrary views are the most recent approaches to 2D view point independent gait recognition. This choice is justified by the fact that the lateral view has proven the recognition capability in a great number of approaches [5, 15, 17, 30, 40, 47, 53, 54].

The biometrics research group of the University of Southampton has focused attention on 2D view invariant gait recognition from 1999 [6] where a trajectory-invariant gait signature was presented. The method of Carter and Nixon corrects the variations in gait data by knowing the walking trajectory and modelling the thigh as a simple pendulum.

The approach was then reformulated by Spencer and Carter [44] to provide a pose invariant biometric signature which did not require knowledge of the subject's trajectory. Results on synthesized data showed that simple pose correction for geometric targets generalizes well for objects on the optical axis.

More recently, these techniques have been refined for the analysis on walking subjects and applied with success on a larger number of video sequences acquired at 6 different point views [13, 43]. View independent gait analysis aims at synthesise the lateral view of a walking subject without camera calibration starting from the successful results on gait biometrics from a lateral view [4]. The dynamic parameters used for biometrics in [4]are the frequential characteristics of the angles that the upper and lower legs form with the vertical axes. Therefore, the view-point independent method aims at synthetise the projection of the principal joints (hips, knees, ankles) on a lateral plane from their positions in the image plane.

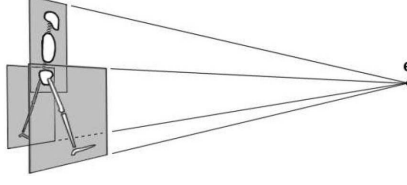The method is based on three main assumptions [29, 51]:

- the nature of human gait is cyclic;
- the distances between the bone joints are invariant during the execution of the movement;
- the articulated leg motion is approximately planar, since almost all of the perceived motion is contained within a single limb swing plane.

Considering a subject walking along a straight line, the multiple periods of linear gait motion appear analogous to a single period viewed from many cameras related by linear translation. Following this rationale, the positions of the points of interest, i.e. the leg joints, lie in an auto-epipolar configuration consistent with the imaged motion direction. The epipole is thus estimated by computing the intersection of the set of lines formed by linking the correspondent points of interest in each phase of the gait cycle (figure 1). Let $\mathbf{j}_i^{\ell}$ be the set of joints positions for each leg $\ell = \{1, 2\}$ at the ith frame in the image reference system. The relationship between $\mathbf{j}_i^{\ell}$ and the corresponding positions in the worldspace is

$$\mathbf{j}_i^{\ell} \times \mathbf{P}_i \cdot \mathbf{J}^{\ell} = 0 \tag{1}$$

where

$$\mathbf{P}_i = \begin{bmatrix} \mathbf{R}_{\mathbf{e}}^T & -i\mathbf{e}_0 \end{bmatrix} \tag{2}$$

Fig. 1 Epipolar configuration of the leg joints (source: Spencer and Carter [43])

and $\mathbf{R}_{\mathbf{e}}^T$ is the rotation matrix for aligning the epipolar vector $\mathbf{e}_0$ with the horizontal axis X. Then,

$$\mathbf{j}_i^\ell = \begin{bmatrix} \mathbf{R}_{\mathbf{e}}^T & -i\mathbf{e}_0 \end{bmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \mathbf{H}_{\mathbf{V}}^{-1} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \mathbf{H}_{\mathbf{V}} \end{pmatrix} = \mathbf{H} \cdot \mathbf{J}^\ell \tag{3}$$

having expressed the limb plane transformation matrix with $\mathbf{H}_{\mathbf{V}}$ so that the two cross section plane lines are parallel with the vertical axis Y, centred and normalised respect to Y and Z axes. By assuming the lengths of the articulated limbs $\mathbf{D}_\ell^2 = \Delta \mathbf{j}_i^{\ell\mathbf{T}} \Delta \mathbf{j}_i^\ell$ are constant over all the frames, the pose difference vectors for the limb segments at two consecutive frames, $\Delta \mathbf{j}_i^\ell$ and $\Delta \mathbf{j}_{i+1}^\ell$, are related by

$$\Delta \mathbf{j}_i^{\ell\mathbf{T}} \cdot \mathbf{H}^{\mathbf{T}} \cdot \mathbf{H} \cdot \Delta \mathbf{j}_i^\ell = \Delta \mathbf{j}_{i+1}^{\ell\mathbf{T}} \cdot \mathbf{H}^{\mathbf{T}} \cdot \mathbf{H} \cdot \Delta \mathbf{j}_{i+1}^\ell \tag{4}$$

After recovering the fronto-parallel structure of subject gait (a more detailed description can be found in [13]) the representation of the leg joints function [$\mathbf{J}_x^\ell(t)$, $\mathbf{J}_y^\ell(t)$] is found by fitting a modified Fourier series to the data with fixed fundamental frequency f_0 and period T. In accordance with [10, 52], in fact, the hip rotation can be modeled by a simple pendulum

$$\mathbf{J}_x^\ell(t) = v_x t + \sum_{k=1}^n A_k \cos \left(2\pi k f_0 \left(t + \frac{(\ell-1)T}{2} \right) + \phi_k \right) + \mathbf{J}_{x0}^\ell \tag{5}$$

analogously for $\mathbf{J}_y^\ell(t)$. Considering $n = 5$ as the sufficient number of harmonics for describing the human walking [9], the projection of the leg joints on the 3D anterior-posterior plane can be expressed in the following way

$$\breve{\mathbf{J}}^\ell(t) = \begin{bmatrix} h_1 & h_2 & h_3 \end{bmatrix} g \left(t + \frac{(\ell-1)T}{2} : f_0, \mathbf{D}_\ell, v_x, v_y, F \right) \tag{6}$$

where $g(t)$ is the bilateral Fourier series function with coefficients F and h are the values of the inverse normalization transform matrix. The 6 results from an optimized procedure where the coordinate positions of limb points are computed and fitted to a linear velocity model with horizontal and vertical velocities equal to V_x and V_y.

Therefore, starting from a video sequence from a single camera and without cali-

bration, the method estimates the gait parameters projected on the lateral plane and their alignment makes them independent from the point view and allows their use for gait identification.

The first experimental tests of the method have been oriented towards the evaluation of its accuracy and therefore it has been applied on reliable and known limb trajectories extracted with reflective markers on the lower limbs. 30 video sequences along 6 different camera views have been analysed and the Mean Correlation Coefficient (MCC) along the i ($i = 1, ..., 6$) directions has been achieved in order to quantify the angle trends matching after the view synthesis. Figure 2 shows the MCC for the thight (γ) and shin (α) angles along the different walking direction.
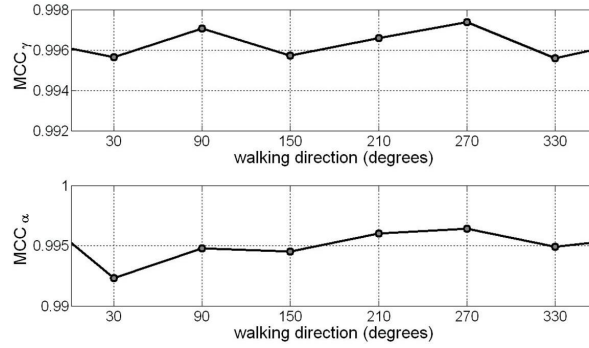

Fig. 2 Tests with reflective markers: mean Correlation Coefficient along the different walking directions (source: Goffredo et al. [13])

Furthermore, in [13] the method sensitivity with respect to the image resolution, camera frame and SNR of Gaussian noise added to the joints trajectories, has been extracted and the minimum MCC obtained is of 0.8.

More recently, the new method for view point independent gait analysis has been applied on video sequences without any markers attached to the subjects [12]. A markerless gait estimation method has been designed and the lower limbs' pose has been extracted over time. Therefore the lateral view synthesizer has been tested in a less constrained environment. Figure 3(a) shows an example of the variations of γ and α during two gait cycles for the six different camera positions. Predictably, the angles' trends are influenced by the subject's pose respect and they cannot be used directly for biometric identification. The angle trends after the application of the point view correction algorithm are shown in figure 3 (b). The slight variations between the resulting traces in figure 3(b) are consistent with intra-subject variation between the video acquisitions. Figure 4 shows the variation of the MCCs with respect to the camera azimuths and elevations. The results, with a mean value of 0.919, are consistent with the value of MCC obtained using reflective markers [13]. Therefore, the correlation values obtained with the angles' trends appear particularly encouraging for its application in the wider context of gait biometrics.
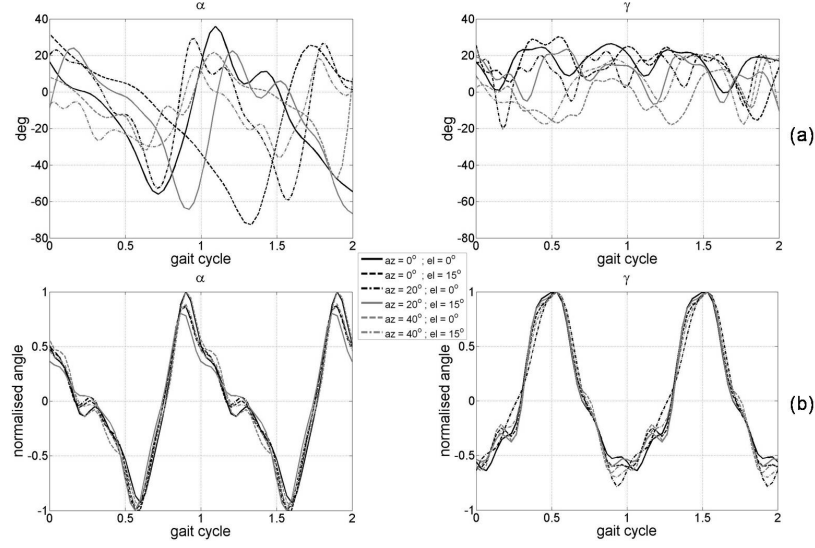
Fig. 3 Markerless tests: hip (γ) and knee (α) angles in different camera positions: (a) unprocessed; (b) corrected (source: Goffredo et al. [12])
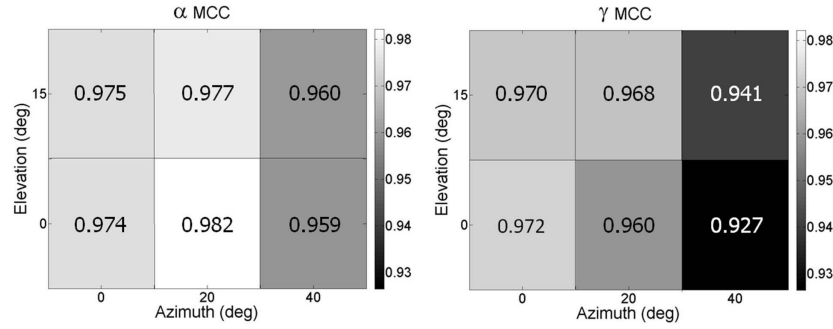
Fig. 4 Markerless tests: mean Correlation Coefficients (MCC) with respect to the different camera positions (source: Goffredo et al. [12])

In 2003 the group of the University of Maryland developed another gait recognition algorithm showing that if a person is far enough from a single camera, it is possible to synthesize the lateral view from any other arbitrary view by knowing the camera parameters [19].

Considering a subject walking with a translational velocity $\mathbf{V} = \left[v_x, 0, v_z\right]^T$ along a straight line which subtends an angle θ with the image plane (AC in Figure 5), if the distance Z_0 of the person from the camera is much larger than the width ΔZ of the person, then it is reasonable to approximate the actual 3D object as being represented by a planar one. Therefore the angle θ in the 3D world can be accurately

estimated in two ways: 1) by using the perspective projection matrix; 2) by using the optical flow based SfM equations. Both methods have been proposed in [19] and, by knowing the camera calibration parameters, the lateral view has been synthetised.

Fig. 5 Imaging geometry (source: Kale et al. [19])

In the perspective projection approach, the 3D line

$$Z = \tan(\theta) X + Z_0 \quad Y = k \tag{7}$$

that is at a height k from the ground plane and parallel to it, corresponds to a straight line in the image plane

$$y = \frac{kf}{Z_0} - k\frac{\tan(\theta)}{Z_0}x \tag{8}$$

where f is the focal length of the camera and

$$x = f\frac{X}{Z_0 + \tan(\theta)X} \qquad y = f\frac{Y}{Z_0 + \tan(\theta)X} \tag{9}$$

Therefore, if the slope of the line in the image plane is α, the angle θ can be estimated with

$$\tan(\theta) = \frac{1}{K}\tan(\alpha) \tag{10}$$

where K and f are obtained from the camera calibration procedure.
Conversely, the optical flow based approach estimate the angle θ with

$$\cot(\theta) = \frac{c(x,y) - \cot(\alpha(x,y))}{m(y,f)} \tag{11}$$

where $c(x,y)$ and $m(y,f)$ can be obtained from calibration data and

$$\cot(\alpha(x,y)) = \frac{p(x,y)}{q(x,y)} \tag{12}$$

where $p(x,y)$ and $q(x,y)$ represent the horizontal and vertical velocity field of a point (x,y) in the image plane.

Therefore, any point (x_α, y_α) on the subject's image, walking at an angle α to the image plane, can be projected on the lateral view:

$$x_0 = f \frac{x_\alpha \cos(\alpha) - f \sin(\alpha)}{-x_\alpha \sin(\alpha) + f \cos(\alpha)} \qquad y_0 = f \frac{y_\alpha}{-x_\alpha \sin(\alpha) + f \cos(\alpha)} \qquad (13)$$

After a camera calibration procedure for the parameters f, K, c and m, the method has been tested on 12 people walking along straight lines at different values of $\theta = 0, 12, 30, 45, 60$ degrees. Figure 6 (a) shows the quality degradation of the synthesized images as a function of angle θ.

In order to study the performance of gait recognition on the synthesized images and keeping in view the limited quantity of training data, the DTW algorithm [21] has been used for gait recognition. Considering a gallery of people walking at lateral view, the video sequences where people walks at arbitrary angles θ have been chosen as probes and the Receiver Operating Characteristic (ROC) has been computed for each θ (figure 6 b).
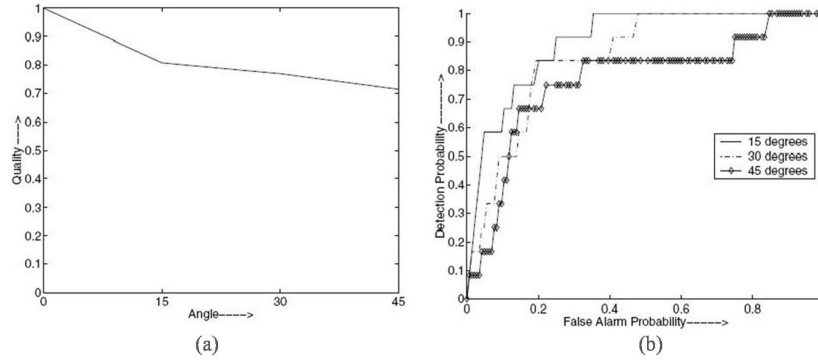

Fig. 6 (a): quality degradation of the synthesized images as a function of angle; (b): ROC curves (source: Kale et al. [19])

More recently, further experimental tests have been done by the same group of the University of Maryland [20] and the recognition performance has been extracted on two publically available gait databases: the NIST database [37] and the CMU database [14]. Some of the results of the lateral view synthesis are shown in figure 8. The gait recognition is depicted in figure 7 where a comparison with the non-normalised images has been included for both the NIST and CMU database.
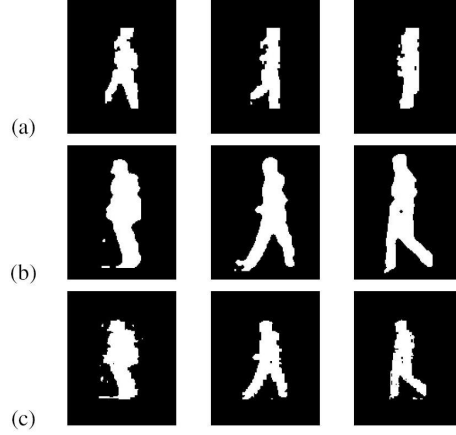
Fig. 7 Sample Images from the NIST Database: (a) gallery images of person walking parallel to the camera; (b) space un-normalized images of person walking at 33° to the camera; (c) space synthesized image for (b) (source: Kale et al. [20])
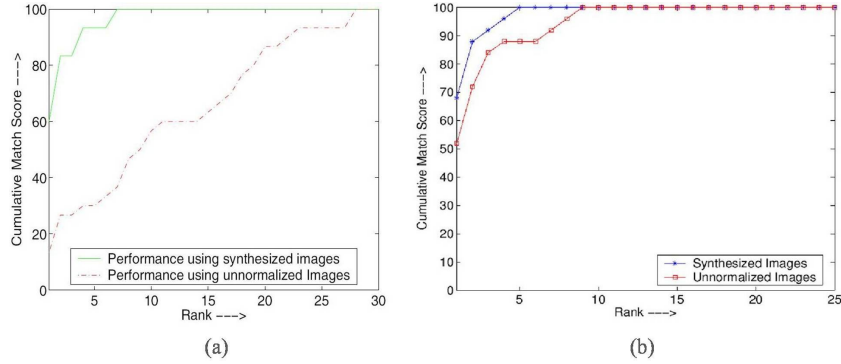
Fig. 8 Recognition results on: (a) NIST database; (b) CMU databases (source: Kale et al. [20])

2.3 Pose-free methods

While the pose-based approaches aim at synthesizing the lateral view of the human body from any other arbitrary views, the pose-free ones extract some gait parameters which are independent from the human pose.

One of the first approaches has been presented by Johnson and Bobick [18], who developed a multi-view gait recognition method using static body parameters. The technique does not analyse the entire dynamics of gait, but uses the action of walking to extract some identifying properties of the individual behavior. Human body limbs are firstly labeled by analysing the binary silhouette of the subject. Subsequently, the head, pelvis and feet positions are extracted following the geometrical proportions

of human body segments. Moreover, two data compensations have been considered: the depth compensation and the camera view factor. The depth compensation is a conversion from pixels to centimeters via an hyperbola function dependent on y_b, the vertical location of the subject's feet:

$$CF\left(y_b\right) = \frac{A}{B - y_b} \tag{14}$$

where A is the vertical distance between the ground and focal point times the focal length and B is the vertical component of the optical center. A and B are estimated by fitting the conversion factor CF to some known subject locations by knowing the subjects' height in centimeters.
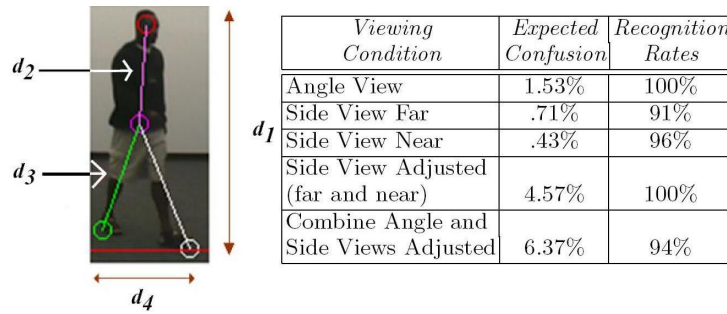
After body labeling and depth compensation, a 4D walk vector is computed as:

- d_1: height of the bounding box around the silhouette;
- d_2: distance (L2 norm) between the head and pelvis locations;
- d_3: maximum value of the distance between the pelvis and left foot location, and the distance between the pelvis and right foot location;
- d_4: distance between the left and right foot.

These parameters are measured only in the frame of maximum feet spread during the walking action in order to avoid self-occlusions.

The second data compensation aims at estimating 4 multiplicative factors to make d_1, d_2, d_3, d_4 independent to the camera view. At this purpose, ground truth motion capture data had been collected and analysed.

Experimental tests regarded 18 subjects walking in front of a single camera positioned at 45 and 90 degrees respect to the walking direction. The side-view data was captured at two different depths, 3.9 and 8.3 meters from the camera. The results had been reported by using an expected confusion metric in order to predict how the subject's parameters filter the identity in a large population. Figure 9 shows the extracted body parameters and the recognition rates at different angle-view, near-side-view and far-side-view. The first three rows represent the results obtained from

Viewing Condition	Expected Confusion	Recognition Rates
Angle View	1.53%	100%
Side View Far	.71%	91%
Side View Near	.43%	96%
Side View Adjusted (far and near)	4.57%	100%
Combine Angle and Side Views Adjusted	6.37%	94%

Fig. 9 Results of the multi-view gait-recognition method using the static body parameters d_1, d_2, d_3, d_4 (source: Johnson and Bobick [18])

the data without any compensation applied: the obtained recognition rate is higher than 91%. By considering the appropriate scale factor based on the viewing condition, the percentage for the side view goes to 100%. Including both the scale and the view-angle adjustments, the results on all the camera positions is 94% and the confusions rates are on the order of 6%.

While Johnson and Bobick proposed a gait analysis based on some static parameters, in 2002 BenAbdelkader et al. proposed a different approach for 2D view independent gait recognition where where the moving person is detected and tracked and an image template corresponding to the person's motion blob is extracted in each frame. [2].

Subsequently, a self-similarity plot from the obtained sequence of templates has been computed in the following way:

$$S(t_1, t_2) = \min_{|dx, dy| < r} \sum_{(x,y) \in B_{t_1}} |O_{t_1}(x + dx, y + dy) - O_{t_2}(x, y)| \qquad (15)$$

where N is the numbers of frames, $1 \leq t_1, t_2 \leq N$, B_{t_1} is the bounding box of the person blob in frame t_1, r is a small search radius, and $O_{t_1}, O_{t_2}, ..., O_{t_N}$ are the scaled and normalized image templates.

The similarity plot is a projection of the dynamics of the walking person that preserves the frequency and phase of the gait. Because gait consists of periodic contiguous steps, the similarity plot can be tiled into contiguous rectangular blocks, termed Units of Self-Similarity (USS), each of which consists of the person's self-similarity over two periods of gait (figure 10). For recognition, the method uses Principal Component Analysis to reduce the dimensionality of the USS's and the k-nearest neighbour rule for classification.

Experimental tests on outdoor sequences of 44 people with 4 sequences of each taken on two different days, and achieves a classification rate of 77%. It is also tested on indoor sequences of 7 people walking on a treadmill, taken from 8 different viewpoints (from 0° to 120°) and on 7 different days. A classification rate of 78% is obtained for near-fronto-parallel views, and 65% on average over all view. Figure 11 shows the classification rates for Dataset 3 for the 8 viewpoints both with absolute correlation of binary silhouettes and with normalized cross-correlation of foreground images. The method appears robust to tracking and segmentation errors and to variation in clothing and background. It is also invariant to small changes in camera viewpoint and walking speed.

A statistical shape analysis is the solution proposed by Wang et al. [50], where the temporal changes of the detected silhouettes are represented by a 2D vector, composed of the edge points distance to the centroid position over time (figure 12). The gait signature is then obtained via the Procrustes shape analysis [27]. A supervised pattern classification technique based on the full Procrustes distance measure has been adopted and the method has been tested on the NLPR gait database (240 sequences from 20 different subjects walking at 3 viewing angles in an outdoor environment). Figure 13 (a) shows the ROC curve, where the EERs (Equal Error Rate) are about 8%, 12%, and 14% for 0° , 90° , and 45° views respectively.
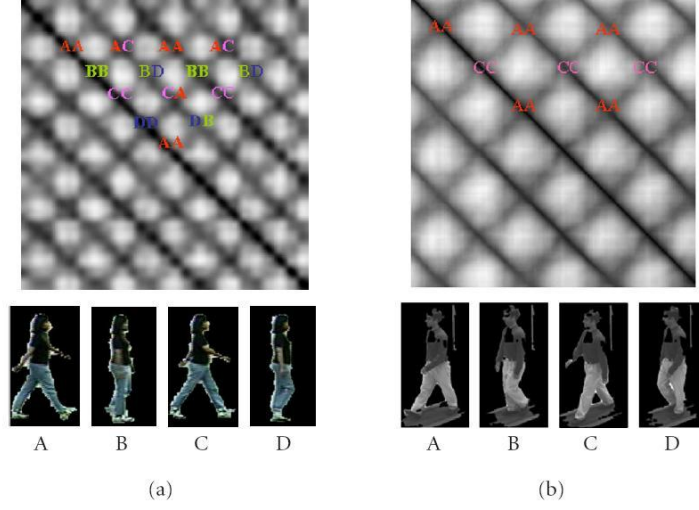
Fig. 10 Units of Self-Similarity for: (a) fronto-parallel sequence; (b) a non-fronto-parallel sequence. Similarity values are linearly scaled to the gray scale intensity range [0, 255] for visualization. The local minima of each SSP correspond to combinations of key poses of gait A, B, C, D (source: BenAbdelkader et al. [2])
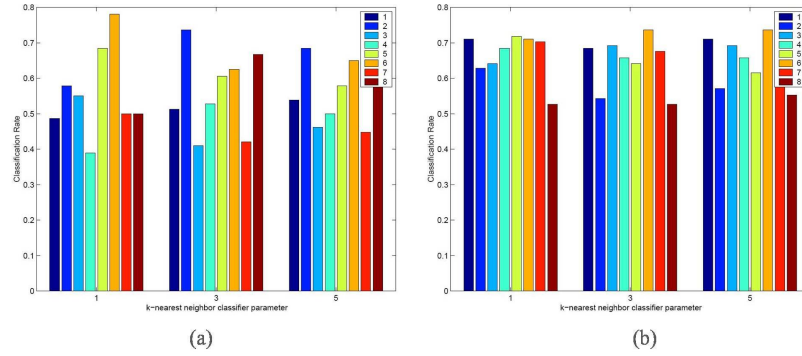

Fig. 11 Classification rates for Dataset 3 for the 8 viewpoints with k = 1; 3; 5 and using: (a) absolute correlation of binary silhouettes; (b) normalized cross-correlation of foreground images (source: BenAbdelkader et al. [2])

The same authors also applied the Principal Component analysis on the 2D silhouette representation [49]. Tests on the same datatabase used in [50] are shown in figure 13 (b).

More recently, Hong et al. [16] introduced a new representation for human gait recognition, called mass vector and defined as the number of pixels with a nonzero value in a given row of the silhouette. Sequences of temporally ordered mass vectors have been used for gait representation and a dynamic time-warping (DTW) [22]
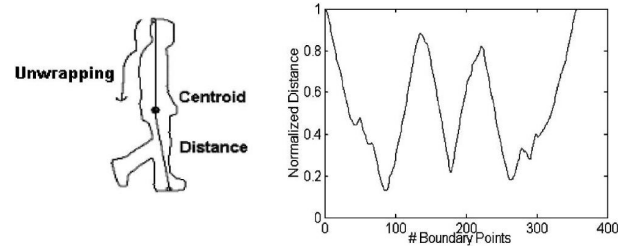
Fig. 12 2D silhouette representation (source: Wang et al. [49])
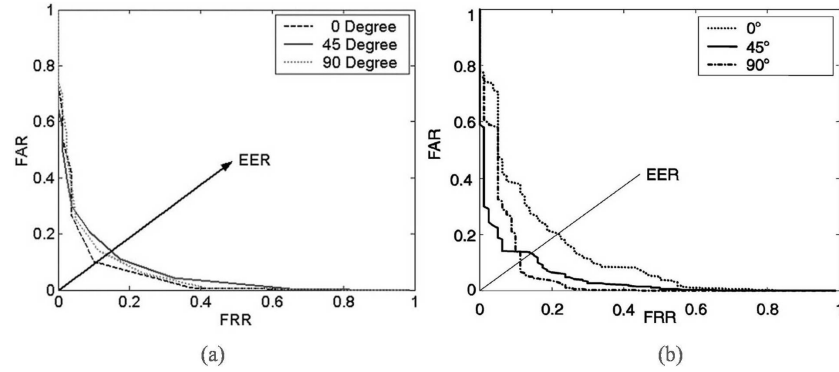
Fig. 13 ROC curves obtained with: (a) Procrustes distance measure; (b) Principal component analysis (source: Wang et al. [49] [50])

approach for the recognition phase.
Experimental tests of the NLPR gait database have given CCR higher of 95% as reported in table 1 where a comparison with the method of Kale et al. has been shown.
A polar representation of a walking subject has been proposed by Chen and Gao [7]

view	method	rank				
		1	2	3	4	5
0°	Kale et al [21]	82.5%	85%	87.5%	90%	92.5%
	Hong et al. [16]	96.25%	98.75%	98.75%	100%	100%
45°	Kale et al [21]	92.5%	95%	97.5%	97.5%	97.5%
	Hong et al. [16]	96.25%	98.75%	98.75%	98.75%	98.75%
90°	Kale et al [21]	77.5%	82.5%	83.75%	86.25%	88.75%
	Hong et al. [16]	88.75%	90%	90%	91.25%	95%

Table 1 CCR with respect to the view-points (source: Hong et al. [16])

in 2007, where a 2D polar gait is obtained by tiling one period gait subsequence in a 2D polar plane along a ring (figure 14 a). The gait characteristic is then achieved

by a combination of appearance models of individual silhouette and contextual silhouettes in the polarplane.

For the individual frame image X_n, a point sets $\{P_j\}$ sampled from a reference circle C are employed as control points. Silhouette description relative to the control point P_j is then provided by means of shape descriptors. Orderly concatenating the shape descriptors originating from $\{P_j\}$, a silhouette appearance model at the nth frame, denoted as h_n, has been obtained. Figure 14 (b and c) shows the silhouette descriptor with angle $\varphi = 32$ and radius $r = 6$. Therefore, the gait appearance model is

(a)   (b)   (c)
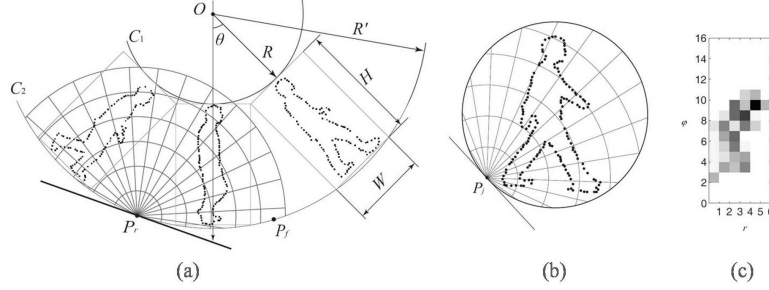
Fig. 14 (a) periodic gait images plane and histogram bins for the contextual silhouettes; (b) histogram relative to the control point P_j; (c) diagram of the polar histogram bins (source: Chen and Gao [7])

defined as a combination of histograms and is represented to be invariant to translation, rotation and scale by means of a shape descriptor and gait images plane. Jeffrey divergence [36] and DTW have been employed for measuring the similarity of gait appearance models between test and reference sequences.

Experimental results on CASIA database [1] demonstrate that the algorithm presents a total CCR of 85.42% and a comparison with other methods ([25, 33, 50]) demonstrated a better performace of the approach proposed by Chen and Gao only for the lateral view.

Angle	$0\,^{\circ}$	$-45\,^{\circ}$	$45\,^{\circ}$	$-90\,^{\circ}$	$90\,^{\circ}$
Sample					
λ	0.034	0.485	-0.592	-1	1
CCR(%)	92.5	95	85	75	65

Fig. 15 CCR with respect to the view angles (source: Chen and Gao [7])

3 3D Gait Recognition

3.1 Introduction

An alternative to using a two-dimensional viewpoint invariant model is to utilise a full three-dimensional analysis technique, as this will be inherently viewpoint independent. There are several different approaches to this; a 3D model can be applied to the video data collected from a single camera, multiple cameras for improved accuracy, or to 3D volumetric data reconstructed from multiple cameras. In this section we discuss several different 3D gait analysis methods proposed by the research community.

Bhanu and Han proposed a novel model-based technique[3], utilising a three-dimensional model with thirty-three degrees of freedom. Several assumptions were made in order to simplify matters; the camera was stationary, subjects walked in a constant direction and the swing of the limbs was parallel to the direction of walking. The model was fitted to silhouette data using a genetic algorithm based on the least squares technique. Recognition performance was evaluated on the stationary and kinematic features separately and combined.

The task of fitting a three-dimensional model to a human subject in a single monocular viewpoint can prove to be quite a difficult task, as there are many degrees of freedom compared to the number of constraints created by a silhouette from a single viewpoint. Even so, it is a very popular research area with many excellent techniques proposed[48, 26, 11]. The use of multiple viewpoints greatly eases the task of fitting a model to a subject, as the set of possible configurations for the model is greatly reduced as the model is simultaneously constrained by multiple silhouettes.

There is a wide variety of datasets containing human motion from multiple viewpoints available to the research community, the most popular one amongst the gait community is that from Carnegie Mellon University; the Motion of Body (MoBo) database[14]. The dataset was recorded indoors using a treadmill, this allowed them to record subjects walking and running at varying gradients and also requires less space. Six cameras were positioned around the subject, and the video data was captured simultaneously. The database contained twenty five subjects, each having a total of twenty-four samples.

Orrite-Uruuela et al devised a technique for fitting a 3D skeletal model to multi-view video data from the CMU database[32]. This involved fitting point-distribution models to the silhouette from each frame, which were similar to the Active Shape Model[8]. The skeletal model is then derived from the set of point-distribution models.

A similar method of fitting a three-dimensional model was proposed by Zhao et al[55], where multiple views were used to improve model fitting performance. A skeletal model is initially fitted to the first frame in a sequence, with the position, orientation, body geometry and joint angles being manually chosen. Tracking

was then performed on the subsequent frames to find the variation in the model's parameters, which could then be used for recognition.

Using video data from multiple viewpoints, it is possible to create a three-dimensional volumetric reconstruction of the subject. One of the most popular methods of doing this is the Visual Hull[24]; the silhouettes from each view are re-projected from the viewpoints, and the intersection of the projections gives the Visual Hull; the Visual Hull provides the maximal shape achievable in the bounds of the silhouettes. This means that the derived shape is often larger than the original, another artifact often experienced is that concave areas are always filled in. More sophisticated techniques consider the correspondence of colour information between viewpoints, in order to provide more realistic reconstructions of the original object[42].

One of the earliest uses of 3D volumetric reconstruction for gait analysis was that by Shakhnarovich et al[39]. A small dataset was created, comprised of 3D visual hull data for 12 subjects walking in arbitrary directions through the target area, with the number of samples per subject varying between two and eight. The visual hulls were constructed using silhouettes generated from four video cameras. The trajectory of the subject was found, and 2D canonical view silhouettes were synthesised by placing a virtual camera into the volumetric space. Two dimensional gait analysis techniques were then used to analyse the synthesised silhouettes. This approach allowed the use of 2D gait analysis techniques in a scenario where the orientation of the subjects is unknown.

Another way of deriving three-dimensional information from multiple views is to use a stereo depth reconstruction technique, where the distance of a point from a pair of cameras can be found from the point's disparity. Stereo vision is used by Urtasun and Fua to aid fitting a three-dimensional model to a subject[45]. A sophisticated deformable model[35] was used for representing the shape of the human subjects and a motion model describing the deformation of the shape model was produced from data collected using a marker based computer-vision system to capture four subjects walking. The shape and motion models were then fitted to three-dimensional data collected from a multi-view stereo vision system using the least squares method.

## 3.2 The University of Southampton Multi-Biometric Tunnel

Up until now, very little research has been carried out into the use of true three-dimensional gait analysis techniques on volumetric data in comparison to two-dimensional gait analysis techniques. The most likely reason for this is the difficulty in acquiring volumetric (3D) data compared to 2D video data. Also the storage capacity required to hold such volumetric data is typically much larger than video data, which in the past made handling and analysing such datasets a very large undertaking. With the ever increasing capacity of storage devices and the growth in processing power, it is likely that the area of 3D gait analysis will start to receive much more attention from the research community.

(a) Inside the tunnel (b) Placement of gait cameras
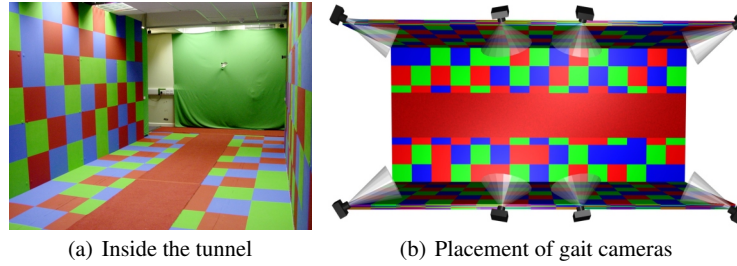
Fig. 16 The University of Southampton Multi-Biometric Tunnel

The currently available three-dimensional gait datasets do not have enough sub-jects or samples to use for any thorough performance analysis, making the com-parison of competing 3D gait analysis techniques difficult. Work has started at the University of Southampton creating a multi-biometric dataset using the University's Multi-Biometric Tunnel[28]. The Multi-Biometric Tunnel is a unique research facil-ity, which facilitates the capture of biometric data from a large number of subjects in a fast and efficient manner. The tunnel area is painted with bright saturated colours in a non-repeating pattern, as shown in Figure 16(a). This simplifies the task of cam-era calibration and also makes it easier to segment the subjects from the background, resulting in high quality silhouettes.

(a) Face (b) Ear (c) 2D Gait (d) 3D Gait
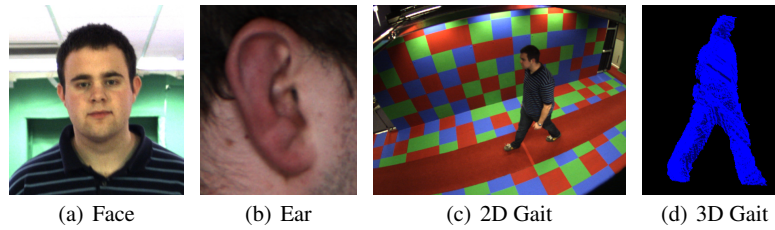
Fig. 17 Example data from Multi-Biometric Tunnel

The dataset will contain several of the most common non-contact biometrics; face, ear and most importantly multiple viewpoint 3D gait, as shown in Figure 17. The subject's face is captured using a UXGA (1600×1200)resolution IEEE1394 video camera at a rate of 7.5 frames per second. The ear is captured by another UXGA video camera assisted by a pair of photographic strobe units. The subject is captured by eight VGA (640×480) resolution IEEE1394 cameras placed around the perimeter of the tunnel, as shown in Figure 16(b); the cameras acquire video at a rate of 30 frames per second. The captured 2D video data and silhouettes will be available along with 3D visual hull reconstructions. Initial tests on a dataset of ten people gave a 100% correct classification rate when synthesising canonical view silhouettes and calculating the average silhouette.

A simple proof of concept experiment was recently conducted using the Multi-Biometric Tunnel to establish whether it is possible to measure a change in an individual's gait when they are concealing an object, such as a bomb. In order to ensure that the analysis only considers the subject's gait and not any change in shape, the participants were asked to wear a harness with eight half-litre water bottles attached to it and tests were performed with the bottles empty and full; creating a change in weight of four kilograms without changing the subject's shape. Four participants were used in the experiment, each providing eight samples with and without the concealed weight. A simple classifier was created, which projects the 3D data into a two-dimensional canonical view, which is then used to calculate the average silhouette. Analysis of Variance and Canonical analysis was then used to improve the discriminatory ability of the classifier. A leave one out strategy was used with the dataset for testing the performance of the classifier. A correct classfication rate of 73.4% was acheived; demonstrating that it is possible to detect concealed weight on one's person by the change in their gait, even when using a very simple gait analysis technique.

4 Conclusions

To deploy gait biometrics in unconstrained views, it is necessary to relieve viewpoint dependence in recognition. To achieve this, we can incorporate geometrical considerations, and the inherent periodicity of gait, and thus achieve a technique which can correct for the effects of viewpoint. In a limited study, this has been shown to achieve viewpoint independent recognition, from image data derived from a single camera view. Conversely, 3D data acquisition requires multiple cameras and then synthesises a solid model of the walking subject. This can achieve viewpoint independence since arbitrary views can be synthesised. A selection of approaches has been developed for these purposes and in limited evaluations has been shown to achieve recognition capability. Now that computational power is abundant, and sufficient for the processing of multi-view data acquired at video rate, we anticipate more development in 3D recognition capability. Equally, as surveillance technology continues to increase, the new techniques are ready for deployment of 2D viewpoint invariant approaches.

References

1. Casia gait database. online (2006). URL http://www.sinobiometrics.com
2. BenAbdelkader, C., Davis, L.S., Cutler, R.: Motion-based recognition of people in eigengait space. In: FGR, pp. 267–274 (2002). URL http://csdl.computer.org/comp/proceedings/fgr/2002/1602/00/16020267abs.htm
3. Bhanu, B., Han, J.: Human recognition on combining kinematic and stationary features. In: Proceedings of Audio- and Video-Based Biometric Person Authentication, *Lecture Notes in*

Computer Science, vol. 2688, pp. 600–608. Springer-Verlag (2003)

4. Bouchrika, I., Nixon, M.S.: Model-based feature extraction for gait analysis and recognition. In: Mirage: Computer Vision / Computer Graphics Collaboration Techniques and Applications, vol. 4418, pp. 150–160 (2007)

5. Boyd, J.E.: Synchronization of oscillations for machine perception of gaits. Comput. Vis. Image Underst. **96**(1), 35–59 (2004)

6. Carter, J.N., Nixon, M.S.: On measuring gait signatures which are invariant to their trajectory. Measurement and Control **32**(9), 265–269 (1999)

7. Chen, S., Gao, Y.: An invariant appearance model for gait recognition. Multimedia and Expo, 2007 IEEE International Conference on pp. 1375–1378 (2-5 July 2007)

8. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active shape models — their training and application. Comput. Vis. Image Underst. **61**(1), 38–59 (1995)

9. Cunado, D., Nixon, M.S., Carter, J.N.: Automatic gait recognition via model-based evidence gathering. In: L. O'Gorman, S. Shellhammer (eds.) Proceedings AutoID '99: IEEE Workshop on Identification Advanced Technologies, pp. 27–30. IEEE (1999)

10. Cunado, D., Nixon, M.S., Carter, J.N.: Automatic extraction and description of human gait models for recognition purposes. Comput. Vis. Image Underst. **90**(1), 1–41 (2003)

11. Fua, P.: Markerless 3d human motion capture from images. In: Encyclopedia of Biometrics. Springer (2008). Forthcoming

12. Goffredo, M., Seely, R.D., Carter, J.N., Nixon, M.S.: Tech. Rep. n.1456, University of Southampton (2007)

13. Goffredo, M., Spencer, N., Pearce, D., Carter, J.N., Nixon, M.S.: Human perambulation as a self calibrating biometric. In: S.K. Zhou, W. Zhao, X. Tang, S. Gong (eds.) AMFG, *Lecture Notes in Computer Science*, vol. 4778, pp. 139–153. Springer (2007). URL http://dx.doi.org/10.1007/978-3-540-75690-3_11

14. Gross, R., Shi, J.: The cmu motion of body (mobo) database. Tech. Rep. CMU-RI-TR-01-18, Robotics Institute, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213 (2001). URL http://www.ri.cmu.edu/pubs/pub_3904.html

15. Hayfron-Acquah, J.B., Nixon, M.S., Carter, J.N.: Automatic gait recognition by symmetry analysis. Pattern Recogn. Lett. **24**(13), 2175–2183 (2003)

16. Hong, S., Lee, H., Nizami, I., Kim, E.: A new gait representation for human identification: Mass vector. Industrial Electronics and Applications, 2007. ICIEA 2007. 2nd IEEE Conference on pp. 669–673 (23-25 May 2007)

17. Huang, P., Harris, C., Nixon, M.: Recognising humans by gait via parametric canonical space. Journal of Artificial Intelligence in Engineering **13**(4), 359–366 (1999)

18. Johnson, A.Y., Bobick, A.F.: A multi-view method for gait recognition using static body parameters. In: AVBPA '01: Proceedings of the Third International Conference on Audio- and Video-Based Biometric Person Authentication, pp. 301–311. Springer-Verlag, London, UK (2001)

19. Kale, A., Chowdhury, A., Chellappa, R.: Towards a view invariant gait recognition algorithm. Proceedings. IEEE Conference on Advanced Video and Signal Based Surveillance, 2003. pp. 143–150 (21-22 July 2003)

20. Kale, A., Roychowdhury, A., Chellappa, R.: Fusion of gait and face for human identification. Acoustics, Speech, and Signal Processing, 2004. Proceedings. (ICASSP '04). IEEE International Conference on **5**, V–901–4 vol.5 (17-21 May 2004)

21. Kale, A.A., Cuntoor, N.P., Yegnanarayana, B., Rajagopalan, A.N., Chellappa, R.: Gait analysis for human identification. In: AVBPA, pp. 706–714 (2003). URL http://springerlink.metapress.com/openurl.asp?genre=article&issn=0302-9743&volume=2688&spage=706

22. Keogh, E., Ratanamahatana, C.A.: Exact indexing of dynamic time warping. Knowl. Inf. Syst. **7**(3), 358–386 (2005)

23. Larsen, P.K., Simonsen, E.B., Lynnerup, N.: Gait analysis in forensic medicine. In: J.A. Beraldin, F. Remondino, M.R. Shortis (eds.) Videometrics IX, vol. 6491, p. 64910M. SPIE (2007). DOI 10.1117/12.698512. URL http://link.aip.org/link/?PSI/6491/64910M/1

24. Laurentini, A.: The visual hull concept for silhouette-based image understanding. IEEE Transactions on Pattern Analysis and Machine Intelligence **16**(2), 150–162 (1994). DOI 10.1109/34.273735
25. Lee, L., Grimson, W.: Gait analysis for recognition and classification. Automatic Face and Gesture Recognition, 2002. Proceedings. Fifth IEEE International Conference on pp. 148–155 (20-21 May 2002)
26. Liu, H., Chellappa, R.: Markerless monocular tracking of articulated human motion. In: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 1, pp. 693–696 (2007)
27. Mardia, K., Jupp, P.: Directional Statistics. New York, Wiley (2000)
28. Middleton, L., Wagg, D.K., Bazin, A.I., Carter, J.N., Nixon, M.S.: A smart environment for biometric capture. In: IEEE International Conference on Automation Science and Engineering, pp. 57–62 (2006). URL http://eprints.ecs.soton.ac.uk/12914/
29. Murray, M.P., Drought, A.B., Kory, R.C.: Walking patterns of normal men. The Journal of Bone and Joint Surgery **46**, 335 (1964)
30. Nixon, M.S., Carter, J.N.: Automatic recognition by gait. Proceedings of the IEEE **94**(11), 2013–2024 (2006)
31. Niyogi, S.A., Adelson, E.H.: Analyzing and recognizing walking figures in xyt. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 469–474 (1994)
32. Orrite-Uruñuela, C., del Rincón, J.M., Herrero-Jaraba, J.E., Rogez, G.: 2d silhouette and 3d skeletal models for human detection and tracking. In: Proceedings of the 17th International Conference on Pattern Recognition, vol. 4, pp. 244–247 (2004)
33. Phillips, P., Sarkar, S., Robledo, I., Grother, P., Bowyer, K.: Baseline results for the challenge problem of humanid using gait analysis. Automatic Face and Gesture Recognition, 2002. Proceedings. Fifth IEEE International Conference on pp. 130–135 (20-21 May 2002)
34. Phillips, P.J., Sarkar, S., Robledo, I., Grother, P., Bowyer, K.W.: The gait identification challenge problem: data sets and baseline algorithm. In: Proceedings of The 16th International Conference on Pattern Recognition, vol. 1, pp. 385–388 (2002)
35. Plänkers, R., Fua, P.: Articulated soft objects for video-based body modeling. In: Proceedings. Eighth IEEE International Conference on Computer Vision, vol. 1, pp. 394–401 (2001). DOI 10.1109/ICCV.2001.937545
36. Puzicha, J., Buhmann, J., Rubner, Y., Tomasi, C.: Empirical evaluation of dissimilarity measures for color and texture. Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on **2**, 1165–1172 vol.2 (1999)
37. Sarkar, S., Phillips, P., Liu, Z., Vega, I., Grother, P., Bowyer, K.: The humanid gait challenge problem: data sets, performance, and analysis. Transactions on Pattern Analysis and Machine Intelligence **27**(2), 162–177 (Feb. 2005)
38. Sarkar, S., Phillips, P.J., Liu, Z., Vega, I.R., Grother, P., Bowyer, K.W.: The humanid gait challenge problem: data sets, performance, and analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence **27**(2), 162–177 (2005)
39. Shakhnarovich, G., Lee, L., Darrell, T.: Integrated face and gait recognition from multiple views. In: Computer Vision and Pattern Recognition, *Proceedings of the 2001 IEEE Computer Society Conference on*, vol. 1, pp. 439–446 (2001)
40. Shutler, J., Nixon, M.S.: Zernike velocity moments for sequence-based description of moving features. Image and Vision Computing **24**(4), 343–356 (2006)
41. Shutler, J.D., Grant, M.G., Nixon, M.S., Carter, J.N.: On a large sequence-based human gait database. In: Proceedings of Fourth International Conference on Recent Advances in Soft Computing, pp. 66–72 (2002). URL http://eprints.ecs.soton.ac.uk/7901/
42. Slabaugh, G.G., Culbertson, W.B., Malzbender, T., Stevens, M.R., Schafer, R.W.: Methods for volumetric reconstruction of visual scenes. Int. J. Comput. Vision **57**(3), 179–199 (2004). DOI http://dx.doi.org/10.1023/B:VISI.0000013093.45070.3b
43. Spencer, N., Carter, J.: Towards pose invariant gait reconstruction. Image Processing, 2005. ICIP 2005. IEEE International Conference on **3**, III–261–4 (11-14 Sept. 2005)

44. Spencer, N.M., Carter, J.N.: Viewpoint invarience in automatic gait recognition. In: Proceedings of Third IEEE Workshop on Automatic Identification Advanced Technologies, AutoID'02, pp. 1–6 (2002)

45. Urtasun, R., Fua, P.: 3d tracking for gait characterization and recognition. In: Proceedings. Sixth IEEE International Conference on Automatic Face and Gesture Recognition, pp. 17–22 (2004). DOI 10.1109/AFGR.2004.1301503

46. Veres, G.V., Gordon, L., Carter, J.N., Nixon, M.S.: What image information is important in silhouette-based gait recognition? In: Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 776–782 (2004)

47. Wagg, D.K., Nixon, M.S.: Automated markerless extraction of walking people using deformable contour models. Computer Animation and Virtual Worlds **15**(3-4), 399–406 (2004)

48. Wang, L., Hu, W., Tan, T.: Recent developments in human motion analysis. Pattern Recognition **36**(3), 585–601 (2003)

49. Wang, L., Tan, T., Hu, W., Ning, H.: Automatic gait recognition based on statistical shape analysis. Image Processing, IEEE Transactions on **12**(9), 1120–1131 (Sept. 2003)

50. Wang, L., Tan, T., Ning, H., Hu, W.: Silhouette analysis-based gait recognition for human identification. IEEE Transactions on Pattern Analysis and Machine Intelligence **25**(12), 1505–1518 (2003)

51. Winter, D.A.: Biomechanics and Motor Control of Human Movement. Wiley (2004)

52. Yam, C., Nixon, M.S., Carter, J.N.: Extended model-based automatic gait recognition of walking and running. In: Proceedings of 3rd Int. Conf. on Audio- and Video-Based Biometric Person Authentication, AVBPA 2001, pp. 278–283 (2001)

53. Zhang, R., Vogler, C., Metaxas, D.: Human gait recognition at sagittal plane. Image Vision Comput. **25**(3), 321–330 (2007)

54. Zhang, Z., Troje, N.F.: View-independent person identification from human gait. Neurocomputing **69**(1-3), 250–256 (2005). URL http://dx.doi.org/10.1016/j.neucom.2005.06.002

55. Zhao, G., Liu, G., Li, H., Pietikäinen, M.: 3d gait recognition using multiple cameras. In: Proceedings of the Seventh IEEE International Conference on Automatic Face and Gesture Recognition (FG '06), pp. 529–534. IEEE Computer Society, Los Alamitos, CA, USA (2006). DOI http://doi.ieeecomputersociety.org/10.1109/FGR.2006.2