Human Perambulation as a Self Calibrating Biometric

Michela Goffredo, Nicholas Spencer, Daniel Pearce, John N. Carter, and Mark S. Nixon

ISIS, School of Electronics and Computer Science,
University of Southampton, SO17 1BJ, UK
mg2@ecs.soton.ac.uk

Abstract. This paper introduces a novel method of single camera gait reconstruction which is independent of the walking direction and of the camera parameters. Recognizing people by gait has unique advantages with respect to other biometric techniques: the identification of the walking subject is completely unobtrusive and the identification can be achieved at distance. Recently much research has been conducted into the recognition of fronto-parallel gait. The proposed method relies on the very nature of walking to achieve the independence from walking direction. Three major assumptions have been done: human gait is cyclic; the distances between the bone joints are invariant during the execution of the movement; and the articulated leg motion is approximately planar, since almost all of the perceived motion is contained within a single limb swing plane. The method has been tested on several subjects walking freely along six different directions in a small enclosed area. The results show that recognition can be achieved without calibration and without dependence on view direction. The obtained results are particularly encouraging for future system development and for its application in real surveillance scenarios.

Keywords: Human motion analysis, 3D modeling, Gait, Biometrics.

1 Introduction

Many biometrics require either a subject's cooperation or contact for data acquisition; vision-based systems usually require a chosen viewpoint. These methods cannot reliably recognize non cooperating individuals at a distance in the real world under changing environmental conditions. Gait, which concerns recognizing individuals by the way they walk, is a relatively new biometric without these disadvantages [1], [2]. There is a rich literature, including medical and psychological studies, indicating the potential for gait in personal identification [3], [4]. Moreover, early medical studies suggest that if all gait movements are considered then gait is unique [5].

There is a rich literature of various gait recognition techniques that can be broadly divided as model-based and model-free approaches. Model based approaches [6], [7] aim to derive the movement of the torso and/or the legs, recovering explicit features describing gait dynamics of joint angles. On the other hand, model-free approaches are mainly silhouette-based. The silhouette approaches [8], [9] characterize body movement by the statistics of the patterns produced by walking. These patterns

S.K. Zhou et al. (Eds.): AMFG 2007, LNCS 4778, pp. 139–153, 2007.

capture both the static and dynamic properties of body shape. A rich variety of data has been collected for evaluation. The widely used and compared databases on gait recognition include: the University of Maryland's surveillance data [10]; the University of South Florida's outdoor data [11]; Carnegie Mellon University's multi-view indoor data [12]; and the University of Southampton's data [13]. The majority of methods and databases found in the literature thus concern a person walking in fronto-parallel [6], [7], [9] or the use of several digital cameras acquiring the movement [7], [8] and thus the knowledge of the calibration parameters.

It appears obvious that for biometric aims the recognition system must be invariant to subject's pose or be able to reconstruct the canonical fronto-parallel view of the gait motion. Recently, in fact, novel approaches on biometrics based on gait are oriented towards synthesizing fronto-parallel views by use of structure from motion, but require some information about camera calibration [14], [15]. Moreover, considering the recent applications of gait recognition in criminal investigation, like the case of the murderer of Swedish Foreign Minister Anna Lindh [16], usually there is no access to the camera and generally only the recorded video sequences are available [17], [18]. Therefore, view-point independent reconstruction of gait would have a major impact on the viability of gait-based biometrics and a system for achieving this purpose is particularly attractive.

This paper presents a new method to reconstruct gait motion from monocular image sequences by taking advantage of the constraints of articulated limb motions. No prior knowledge of the camera calibration is necessary and the limbs landmark points are extracted over all frames in the sequence by tracking reflective markers on subject's legs. This work is part of a wider project (Fig. 1) where the single view human identification based on gait will be completely markerless. Therefore, the aim of this paper is to propose a novel method for reconstructing the gait parameters independently from the view-point and the subject pose.

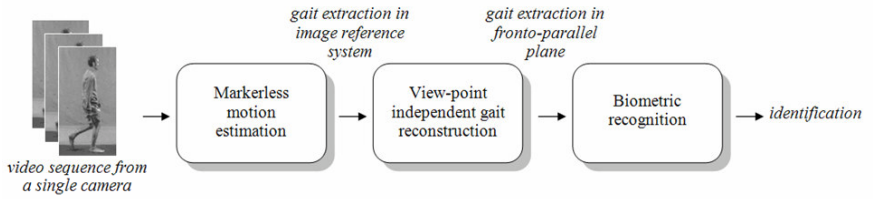

Fig. 1. Flow diagram of the overall project

2 Theory

Walking has been widely studied in medical research and biomechanics [19] and the characteristics of the gait cycle for subjects without any pathology are well-known. Each leg has two distinct periods: a stance phase, when the foot is in contact with the floor; and a swing phase, when the foot is off the floor moving forward to the next step. The time interval between successive instances of initial foot-to-floor contact 'heel strike' for the same foot is the gait cycle.

The proposed method for a pose invariant gait analysis is based on three main assumptions: the nature of human gait is cyclic; the distances between the bone joints are invariant during the execution of the movement; and 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. In order to find these correspondences, the gait periodicity is calculated by applying the stereopsis transformation that maps the epipole $\mathbf{e}$ to the ideal point $[1,0,0]^T$ and then by computing the cost based on dot product between matching limb segment vectors.

After estimating the periodicity of gait, assuming linear velocity between consecutive frames, the set of points of interest are recomputed in order to lie on straight lines starting from the epipole. At first the set of points is mapped to the unit square $\mathbf{x_0}=\mathbf{K_n x}$ with the matrix $\mathbf{K_n}$. Similarly the epipole $\mathbf{e_0}=\mathbf{K_n e}$ is re-normalized to the unit norm $\|\mathbf{e_0}\|=\mathbf{1}$. Subsequently, the optimal points are found by estimating the positions $\bar{\mathbf{x}}_i$ that lie on the epipolar line and that satisfies the condition

$$\bar{\mathbf{x}}_i^T [e_0]\times \bar{\mathbf{x}}_i = 0 \tag{1}$$

Therefore the back projected rays, formed from a set of optimal points, intersect in a single worldspace point: the epipole.

The back projection of all sets of points generates the cluster of 3D points for an assumed single period of reconstructed gait motion. The Direct Linear Transform, DLT [20], is thus used in order to triangulate each worldspace point $\mathbf{X}$

$$\left([\bar{\mathbf{x}}_k]\times \mathbf{P}_k\right)\cdot \mathbf{X} = 0 \tag{2}$$

with the set of camera projection matrices

$$\mathbf{P}_k = \left[\mathbf{R}_\mathbf{e}^T | -k\cdot \mathbf{e}_0\right] \tag{3}$$

where $\bar{\mathbf{x}}_k$ is the image of the worldspace point $\mathbf{X}$ in the k^{th} period image, $\mathbf{R_e}$ is the 3 by 3 rotation matrix that aligns the epipolar vector with the X (horizontal) axis, and k is an integer describing the periodicity of the subject's translation.

Considering the assumption that the articulated leg motion is approximately planar the 3D limb points can be reasonably fitted to two planes. Since the epipolar vector is aligned with the X axis, the ideal point $[1,0,0,0]^T$ do lie on each of the worldspace planes. Therefore, the pencil of planes that intersect this ideal point have the form $p=(0,v_2,v_3,v_4)^T$. Consequently the problem is reduced to finding two lines within the YZ plane cross section data.

After computing the mean $[y,z]^T$ of the point distribution, the translation $\mathbf{H_t}$ that maps this point to the origin is applied. The two cross section plane lines $l_1=[v_2,v_3,v_4]^T$ and $l_2=[v_2',v_3',v_4']^T$ are then achieved by orthogonal regression and then aligned parallel with the Y (vertical) axis by applying a rotation $\mathbf{H_r}$. The intersection point of the two lines is then called **u** and is given by the cross product between the two lines.

Consequently, the pair of transformed lines are mapped to $l_i'=\mathbf{H_r}l_i$ and the rotation matrix $\mathbf{H_r}$ and the perspective transformation $\mathbf{H_\alpha}$

$$\mathbf{H}_\alpha = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \alpha & 0 & 1 \end{pmatrix} \tag{4}$$

are applied to the point **u'** in order to transform it to the ideal point $[1,0,0]^T$.

Since **u'** lies on the Y axis and has the form $[y,0,w]^T$, the transformation $\mathbf{H_\alpha u'}$ gives $\alpha=-w/y$ and the corresponding line mapping $\mathbf{H_\alpha l_i'}$ effectively zeros the first component of the two normal lines. Since the lines are parallels, they are normalized

$$l_1'' = (0,1,-c_1)^T \qquad l_2'' = (0,1,-c_2)^T \tag{5}$$

so that is it possible to find the point (c_1,c_2) of intersection with Z (depth) axis. A further similarity transform $\mathbf{H_s}$ that translates the midpoint $(c_2+c_1)/2$ to the origin and scales in the Z direction to rectify the lines to the form $l=[0,1,\pm1]^T$ is then applied.

The translation by ±1 mapping the selected set of points onto the z=0 plane is then computed with the matrix $\mathbf{H_\beta}$. The combined set of transformations thus forms the limb plane transformation $\mathbf{H_v= H_\beta H_s H_\alpha H_r H_t}$. In order to change the matrix order, a similar set of transformation is constructed:

$$\begin{aligned} \mathbf{H_v} &= \mathbf{H_\beta H_\alpha (H_\alpha^{-1} H_s H_\alpha) H_r H_t} \\ \mathbf{H_v} &= \mathbf{H_\beta H_\alpha H_s' H_r H_t} \end{aligned} \tag{6}$$

Therefore, the projection transform mapping the back projected points into the image can be decomposed as:

$$\bar{\mathbf{x}}(k) = \left[\mathbf{R}_\mathbf{e}^T \mid -k\cdot \mathbf{e}'\right] \begin{pmatrix} 1 & 0 \\ 0 & \mathbf{H}_\mathbf{v}^{-1} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \mathbf{H}_\mathbf{v} \end{pmatrix} \mathbf{X} \tag{7}$$

where

$$\mathbf{H_v} = \begin{bmatrix} 1 & m_2 & m_3 & m_4 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -\beta \\ 0 & -\alpha & 0 & 1 \end{pmatrix} \tag{8}$$

The corresponding transformation of worldspace points $(u,v,0,w)^{\mathrm{T}}$ into the image is given by

$$\bar{\mathbf{x}}' = \mathbf{H}_{\mathbf{p}} \cdot (u, v, w)^{\mathrm{T}} \tag{9}$$

where

$$\mathbf{H}_{\mathbf{p}} = \left[\mathbf{e}' \quad m_2' - \alpha \cdot \left(m_4' - k \cdot \mathbf{e}' \right) \quad \left(m_4' - k \cdot \mathbf{e}' \right) - \beta \cdot m_3' \right] \tag{10}$$

with

$$m_i' = \mathbf{R}_{\mathbf{e}}^{\mathrm{T}} m_i \qquad \mathbf{e}' = \mathbf{R}_{\mathbf{e}}^{\mathrm{T}} [1,0,0]^{\mathrm{T}} \tag{11}$$

Finally the sets of optimal z=0 plane points is found by solution of the

$$\left(\mathbf{x}_{k,\beta}' \right) \times \mathbf{H}_{\mathbf{p}}(k, \beta) \bar{\mathbf{w}} = 0 \tag{12}$$

for each point **w** in order to minimize the reprojection error.

Structure on the z=0 plane has been recovered up to an affine ambiguity $\mathbf{H}_{\boldsymbol{\mu}}$ that maps the imaged circular points $(1,\mu \pm i\,\lambda,0)^{\mathrm{T}}$ back to their canonical positions $(1,\pm i,0)^{\mathrm{T}}$:

$$\mathbf{H}_{\boldsymbol{\mu}} = \begin{pmatrix} 1 & 0 & 0 \\ -\dfrac{\mu}{\lambda} & \dfrac{1}{\lambda} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{13}$$

For estimating the metric structure, the lengths of the articulated limbs is assumed to be known and constant over all the frames. Thus the squared distance between two points $\mathbf{x_0}$ and $\mathbf{x_1}$ can be written

$$d^2 = \boldsymbol{\Delta}\mathbf{x}^{\mathrm{T}} \boldsymbol{\Delta}\mathbf{x} \tag{14}$$

where

$$\boldsymbol{\Delta}\mathbf{x} = \left[u_1 - u_0, v_1 - v_0 \right]^{\mathrm{T}} \tag{15}$$

If $\boldsymbol{\Delta}\mathbf{x_1}$ and $\boldsymbol{\Delta}\mathbf{x_2}$ are the pose difference vectors for a limb segment at two consecutive frames, then the equal limb length constraint can be written

$$\boldsymbol{\Delta}\mathbf{x}_1^{\mathrm{T}} \mathbf{H}^{\mathrm{T}} \mathbf{H} \boldsymbol{\Delta}\mathbf{x}_1 = \boldsymbol{\Delta}\mathbf{x}_2^{\mathrm{T}} \mathbf{H}^{\mathrm{T}} \mathbf{H} \boldsymbol{\Delta}\mathbf{x}_2 \tag{16}$$

Therefore, writing $\boldsymbol{\Delta}\mathbf{x_i} = [\delta x_i, \delta y_i]$ and the element of the matrix $\mathbf{M}=\mathbf{H}^{\mathrm{T}}\mathbf{H}$ as $\mathrm{m}=\left[\mathrm{M}_{11}, \mathrm{M}_{12}, \mathrm{M}_{22} \right]^{\mathrm{T}}$, the equation is

$$\left|\delta x_1^2 - \delta x_2^2 \quad 2\cdot\left(\delta x_1^2 \delta y_1^2 - \delta x_2^2 \delta y_2^2\right) \quad \delta y_1^2 - \delta y_2^2\right| \mathrm{m} = 0 \tag{17}$$

Since m is defined up to scale then a minimum of two corresponding pose constraints are required. All constrains formed from all sets of combinations of same limb frame poses are stacked on each swing plane.

The rectification matrix $\mathbf{H_\mu}$ is formed from the extracted parameters of $\mathbf{H^T H}$, where $\mu = -m_2/m_3$ and

$$\lambda = \sqrt{m_1 / m_3 - \mu^2} \tag{18}$$

The ideal epipole $[1,0,0]^T$ is then mapped by $\mathbf{H_\mu}$ to $[1,-\mu/\lambda,0]^T$ so a rotation $\mathbf{H_r}$ is necessary in order to align the epipole back along the X axis such that $\mathbf{H_a = H_r H_\mu}$ is the affine transform that recovers metric angles and length ratios on both planes.
Points on the metric plane **w** are then mapped into the image as:

$$\bar{\mathbf{x}} = \mathbf{H_p H_a^{-1}}\left(\mathbf{H_a}\bar{\mathbf{u}}\right) = \bar{\mathbf{H}}\bar{\mathbf{w}} \tag{19}$$

Scaling is then applied to both planes in order to transform each first limb segment to unit length. The mean set of limb lengths for both planes is estimated as **d**, **d'**. These lengths are related by the inter-plane scaling: $\mathbf{d_i = \tau d_i'}$. A minimal solution to this trivial set of linear equations requires at least one valid length correspondence within the set of limb segments. With $\mathbf{H_\tau}$ now known the optimal first limb segment length $\mathbf{D_1}$ on the first plane can be evaluated. The scaling transform $\mathbf{H_s}$ that maps $\mathbf{D_1}$ to the unit length and update both sets of points and projection homographies is then calculated.

$$\begin{aligned}
\mathbf{H_1} &= \bar{\mathbf{H}}_\mathbf{p}\bar{\mathbf{H}}_\mathbf{s}^{-1} = \begin{bmatrix} \frac{p_1}{s} & \frac{p_2}{s} & p_3 \end{bmatrix} \\
\bar{\mathbf{H}}_\mathbf{2}\mathbf{H}_\tau^{-1}\mathbf{H}_\mathbf{s}^{-1} &= \begin{bmatrix} \frac{p_1}{s}\cdot\tau & \frac{p_2}{s}\cdot\tau & p_3' \end{bmatrix} \\
\mathbf{H_2} &= \begin{bmatrix} \frac{p_1}{s} & \frac{p_2}{s} & \tau\cdot p_3' \end{bmatrix}
\end{aligned} \tag{20}$$

where

$$\begin{aligned}
p_2 &= m_2' - \alpha\cdot m_4' \\
p_3 &= m_4' - m_3' \\
p_3' &= m_4' + m_3'
\end{aligned} \tag{21}$$

defined in equation 10.

The true metric structure $\mathbf{w_i}$ is then recomputed from the real normalized image points $\mathbf{x_i'}$ by applying the inverse mappings

$$\mathbf{w_i} = \mathbf{H_1}^{-1}\mathbf{x_i}^{'} \qquad \mathbf{w_i}^{'} = \mathbf{H_2}^{-1}\mathbf{x_i}^{'} \tag{21}$$

Therefore, the four-fold X,Y reflection ambiguity of the metric plane is resolved by consideration of the gross spatiotemporal motion structure. Two smoothed data vectors $\tilde{u}_0$, $\tilde{u}_0^{'}$ generated from the mean X coordinate positions of limb points over a centred 3 frame window, are computed and fitted to a linear velocity model with a pair of simultaneous equations:

$$\tilde{u}_i = v_x \cdot i + u_0 \qquad \tilde{u}_i^{'} = v_x \cdot i + u_0^{'} \tag{22}$$

The gait sequences is then normalized in order to emulate a left to right walk, so ensureing that v_x is positive by applying a reflection about the Y axis and updating both points $\mathbf{w_i}$, $\mathbf{w_i}^{'}$ and the homographies. The reflection about the X axis, to ensure that the sky is upward, is determined from the Y coordinate ordering of the means of each limb point over all frames. The only remaining ambiguity is then the translation between both sets of plane points.

Since normal gait is bilaterally symmetric with a half phase shift, for each limb segment both plane limb angle sets and their corresponding time sample vectors are computed. Therefore, the angle vectors can be concatenated as **A=[a|a']** and the time sample vector as **S=[t|t'+T**/2].

With the knowledge of the normalized limb lengths **D** both sets of origin limb points **o**, **o'** can be found by back substitution. We then compute two vectors of smoothed X origin limb data generated from the mean positions over a centred 3 frame window, and fit the linear velocity model to the pair of simultaneous equations in **t**. This gives a reasonable estimate of the linear velocity component and initial X offset points [$\mathbf{u_0}$,$\mathbf{u_0}'$] of gait on the metric plane. We now compute a partitioned bilateral Fourier series representation of the origin point displacement function with sample data **o**, **o'** and fixed fundamental frequency f_0

$$\begin{aligned} u(t) &= v_x \cdot t + \sum_{k=1}^{n} A_k \cos\left(2\pi k f_0 t + \phi_k\right) + u_0 \\ u^{'}(t) &= v_x \cdot t + \sum_{k=1}^{n} A_k \cos\left(2\pi k f_0 \left(t + \frac{T}{2}\right) + \phi_k\right) + u_0^{'} \end{aligned} \tag{23}$$

The initial first harmonic is firstly computed by partitioning the parameter vector $\mathbf{P_1}$= [v_x,A_1, ϕ_1 | u_0, $u_0^{'}$]^T.

The estimation of $\mathbf{P_1}$ is then used to bootstrap the full partitioned parameterization:

$$\mathbf{P} = \left[v_x, A_1, \phi_1, ..., A_n, \phi_n \mid u_0, u_0^{'}\right] \tag{24}$$

The Y component origin limb point displacement function is similar, though v_y is held fixed (zero). Both are computed using a partitioned Levenberg-Marquardt algorithm [21] with fixed fundamental frequency f_0. The translations $\mathbf{H_o}$, $\mathbf{H_o}^{'}$ then map the starting origin limb point displacements [u_0, v_0]^T , [$u_0^{'}$, $v_0^{'}$]^T to the origin.

We finally apply the inverse normalization transform to the updated homography mappings $\mathbf{K_n}^{-1}\mathbf{H_i}'$ in order to map the metric plane points to image points:

$$\begin{aligned}\mathbf{x}_1(t) &= \begin{bmatrix} h_1 & h_2 & h_3 \end{bmatrix} g\left(t : f_0, \mathbf{D}, V_x, V_y, F\right) \\ \mathbf{x}_2(t) &= \begin{bmatrix} h_1 & h_2 & h_3' \end{bmatrix} g\left(t + T/2 : f_0, \mathbf{D}, V_x, V_y, F\right)\end{aligned} \tag{25}$$

where $g(t)$ is the bilateral Fourier series function, V_x and V_y are the velocity and Fourier coefficients of the metric plane origin limb displacement functions and F are the Fourier coefficients of the set of limb pose angle functions. As a final optimization step we perform a bundle adjustment procedure that minimizes reprojection error with respect to all parameters **P** of the gait projection function.

3 Experimental Tests

The method has been tested on 5 subjects walking freely along 6 different directions in a 7x7m^2 area. The video sequences were acquired with a digital camera FLEA IEEE-1394 Digital Camera (Point Grey Research) with a spatial resolution of 1024x768 pixels and 30 fps. Two 575 Watt lights illuminated the scene and 6 reflective markers (Vicon® 14mm) have been applied on the lower limbs (3 on the shank and 3 on the thigh). Fig. 3 shows the experimental setup.

3.1 Markers and Limb Pose Extraction

The aim of this paper is to test the proposed view-point independent gait reconstruction method on reliable and known limb trajectories. For this reason a set of markers has been applied on the subjects' legs and thus the first block of Fig. 1 has been replaced with the marker extraction (Fig. 2). Moreover, this is only the first step of a wider project where the extraction of leg pose will be completely automatic by developing a specialized image processing technique.

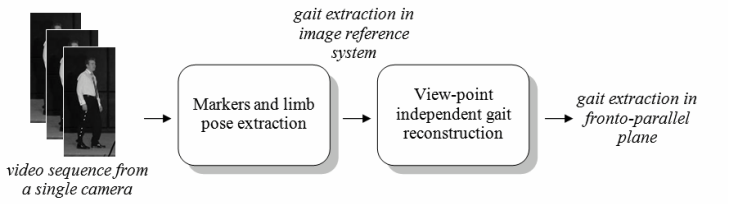

Fig. 2. Flow diagram for testing the proposed method

For the extraction of limb poses frame by frame an algorithm based on color segmentation and blob analysis has been designed.

Because of the special reflective structure of the markers with respect to the background and to the subject clothes, the marker extraction can be easily and robustly achieved by applying an RGB threshold. Robustness to noise and lighting changes has been achieved by comparing the roundness of the objects obtained from the color segmentation. The centroid of the remaining objects is then estimated and a vertical separation of these points is used to classify them into two subsets representing the thigh and the shank (lower leg) points. The individual thigh and shank points are then matched between consecutive frames on the assumption that the same point in the next frame has not moved more than the still closest point, in 2D Euclidian norm, to the point in the present frame. This is of course dependent on the frame rate of the camera system. The frame rate of 30 fps that is used here is enough to make this classification stable for the type of movement captured.

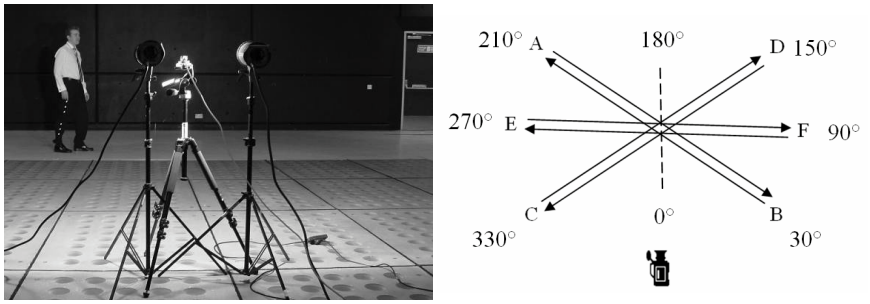

Fig. 3. Experimental setup

Because of the special reflective structure of the markers with respect to the background and to the subject clothes, the marker extraction can be easily and robustly achieved by applying an RGB threshold. Robustness to noise and lighting changes has been achieved by comparing the roundness of the objects obtained from the color segmentation. The centroid of the remaining objects is then estimated and a vertical separation of these points is used to classify them into two subsets representing the thigh and the shank (lower leg) points. The individual thigh and shank points are then matched between consecutive frames on the assumption that the same point in the next frame has not moved more than the still closest point, in 2D Euclidian norm, to the point in the present frame. This is of course dependent on the frame rate of the camera system. The frame rate of 30 fps that is used here is enough to make this classification stable for the type of movement captured.

After applying the method for extracting the markers trajectories to the video sequences, the limbs pose has been estimated frame by frame. Moreover, the hip and knee angles, θ and ϕ respectively, have been evaluated for each walking direction. Fig. 4 shows an example of markers extraction and the corresponding stick body model.

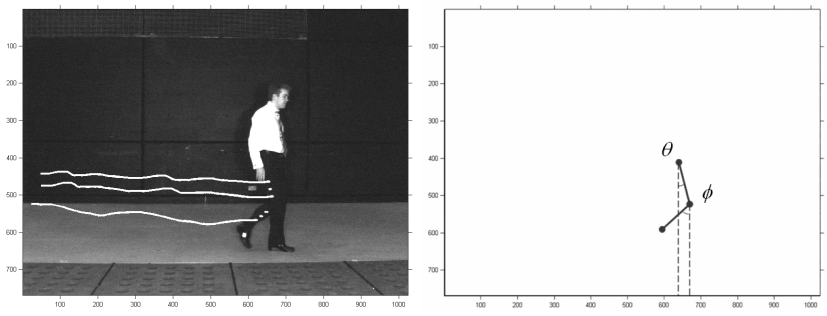

Fig. 4. Marker extraction and stick model

3.2 Gait Reconstruction

From the upper traces in Fig. 5 it is obvious that the mere extraction of limbs position from the 2D images produces angle trends that cannot be used directly for biometric identification. For this reason, the proposed method allows correction of these gait patterns so that they have trends similar to the one that can be achieved by extracting them on the anterior-posterior plane as showed in the lower part of Fig. 5. By inspection the correction achieved by our new approach has aligned the sequences such that they can now be used for identification purposes: the rectification process has led to series which are now closely aligned for the same subject. There is some slight variation between the resulting traces (in the lower part of Fig. 5) consistent with intra-subject variation between the recordings.

Data obtained by analysing the 30 video sequences has been collected and in order to quantify the angle trends matching after the proposed correction, the Mean Correlation Coefficient (MCC) along the i ($i=1,..N$) different directions has been achieved. Let S be the number of subjects, the MCC for the angle θ is defined in the following way:

$$MCC_{\theta}(i) = \frac{\sum_{k=1}^{S} CC_k^{\theta}(i)}{S} \tag{26}$$

where $CC_k^{\theta}(i)$ is the mean correlation coefficient of subject k along the walking direction i:

$$CC_k^{\theta}(i) = \frac{\sum_{j=1}^{N} R_{i,j}^{\theta}}{N} \tag{27}$$

and $R_{i,j}^{\theta}$ is the off-diagonal elements of the correlation coefficients matrix between the directions i and j and similarly for ϕ.

Fig. 6 shows how the MCCs vary with respect to the walking direction i (where $i=0,\ldots,360$ as reported in Fig. 1). The results, with a mean value of 0.996, are

particularly encouraging and the peaks corresponding at the two fronto-parallel paths (at 90 and 270 degrees) confirm that the reconstructed angles along different directions are correlated with the canonical view. The front and rear views (at 0 and 180 degrees) show the lowest performance, but this is only where the MCC is still very close to the mean anyway. The errors in the estimates for the shank (lower leg) inclination are slightly greater and the MCC is slightly lower, but still at minimum 0.992. The larger error in the lower leg is due to the greater freedom in movement, and since the shank moves much faster than the thigh reducing effective resolution.

Furthermore, the information regarding the limbs' pose along different directions has allowed estimation of the root mean square (RMS) distance between the detected marker points and the projected ones.

The mean RMS error for the 5 subjects is 0.2% of the image resolution. This result is particularly encouraging especially compared with the 0.1% RMS obtained using Zhang calibration algorithm [22].

Moreover, the video sequences and the marker trajectories have been modified in order to test the method under different circumstances.

Firstly, the sensitivity of the method with respect to the camera characteristics has been evaluated by changing the image resolution and the frame rate.

Table 1 shows the mean MCC and RMS values obtained with the image spatial resolution varying from 320x240 to 640×480 to 1024×768 pixels. The frames have been resized using nearest-neighbor interpolation.

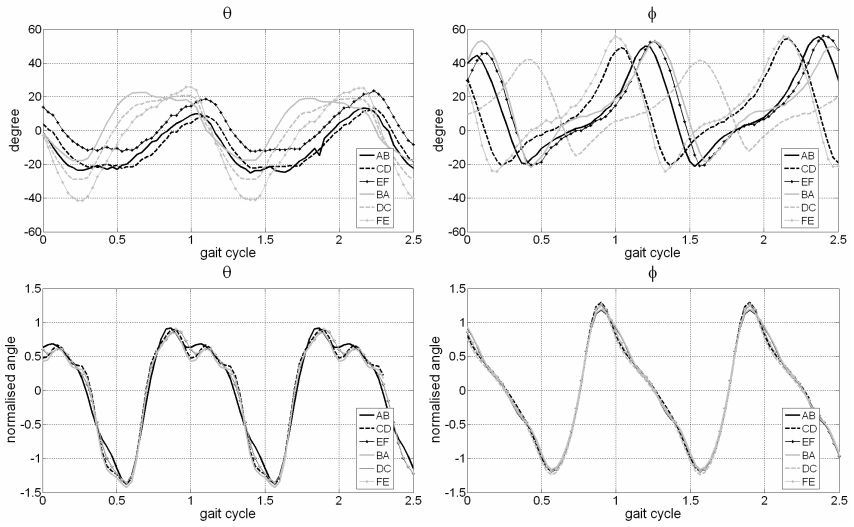

Fig. 5. Hip and knee angles in different walking directions, unprocessed (above) and corrected (below)

It is notable how the image resolution remarkably influences the RMS and how, on the other hand, the mean MCC remains higher than 0.9. The sensitivity of the proposed method to the camera frame rate is reported in Fig. 7, where the video

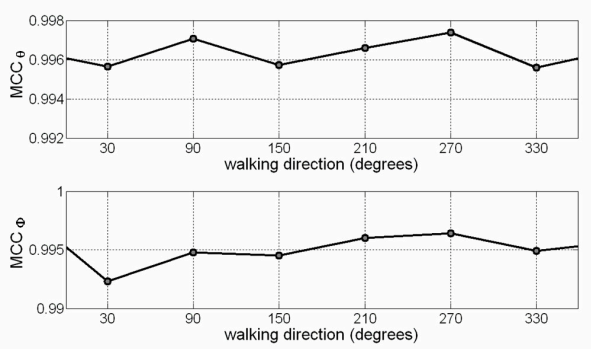

Fig. 6. Mean Correlation Coefficient (MCC) along the different walking directions

sequences have been subsampled from 20 to 30 fps. The mean MCC and RMS show that with a frame rate higher than 25 fps the performance has a linear trend.

In addition, to simulate the limb tracking imprecision, zero-mean Gaussian noise has been added to the markers' trajectories. The standard deviation of the added noise varies from 0 (original data) to 10 pixels. For each noise level, 6 different trials have been conducted and added to each marker trajectory. Obviously the error increases with the level of noise added to the trajectories but the decrease of MCC is particularly interesting because it is higher than 0.9 even at high noise levels.

Table 1. Sensitivity of MCC and RMS with respect to image resolution

	320x240	**640x480**	**1024x768**
Mean MCC	0.989	0.992	0.996
RMS (%)	0.990	0.549	0.278

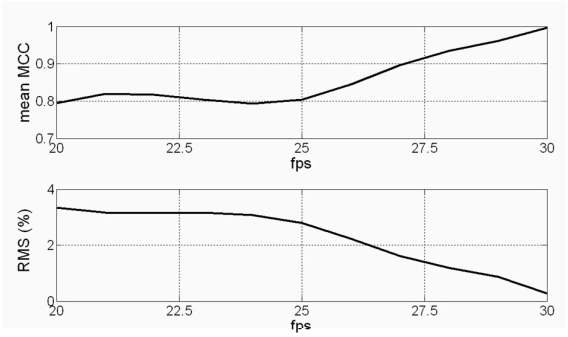

Fig. 7. Sensitivity of MCC and RMS with respect to camera frame rate

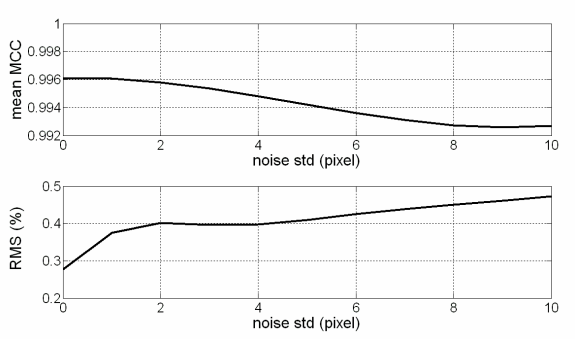

Fig. 8. Sensitivity of MCC and RMS with respect to Gaussian noise in the markers points

Moreover, the noise study allows to understand how the proposed method depends on an accurate limb position extraction and thus to extend the approach in a markerless context.

4 Conclusion

It is widely recognized that gait identification has unique advantages, such as the unobtrusiveness, respect to other biometric techniques. Recently a lot of research has been conducted into the recognition of fronto-parallel gait or using more views after the calibration process. However, considering biometrics aims and the criminal investigation context these approaches appear limited, especially because usually there is no access to the cameras and the subjects do not walk along the canonical direction. For these reasons a system is required which does not rely on the subject's pose or on the calibration of the camera.

In this context this paper has introduced a novel method for a pose invariant gait analysis based on the assumptions that the nature of human gait is cyclic, the distances between the bone joints are invariant and the articulated leg motion is approximately planar.

Experimental tests have been conducted on 5 subjects walking freely along 6 different directions in a $7x7m^2$ area. Different conditions concerning image resolution, frame rate and noisy limb trajectories, have been compared. Moreover the influence of the walking direction on the angle's estimation has been analysed. This quantity of data and the amount is sufficient to demonstrate that the proposed new approach can achieve view-point independent gait reconstruction as described.

The obtained results are particularly encouraging for future system development concerning the markerless extraction of limbs' pose, and for its appliance in real surveillance scenarios.

Acknowledgments. The study was supported by DTC through General Dynamics.

References

1. 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)
2. Nixon, M.S., Carter, J.N.: Automatic Recognition by Gait. Proceeding of the IEEE 94(11), 2013–2023 (2006)
3. Nixon, M.S., Carter, J.N., Cunado, D., Huang, P.S., Stevenage, S.V.: In: Jain, A., Bolle, R., Pankanti, S. (eds.) Biometrics: Personal Identification in a Networked Society, pp. 231–250. Kluwer Academic Publishing, Dordrecht (1999)
4. Cutting, J., Kozlowski, L.: Recognizing friends by their walk: gait perception without familiarity cues. Bull. Psychonom. Soc. 9, 353–356 (1977)
5. Murray, M.P., Drought, A.B., Kory, R.C.: Walking patterns of normal men. J. Bone and Joint Surgery 46-A(2), 335–360 (1964)
6. BenAbdelkader, C., Cutler, R., Davis, L. et al.: Motion-Based Recognition of People in Eigengait Space. In: Proc. Automatic Face and Gesture Recognition (2002)
7. Wagg, D.K., Nixon, M.S.: On automated model-based extraction and analysis of gait. In: Proc. Automatic Face and Gesture Recognition, pp. 11–16 (2004)
8. Collins, R., Gross, R., Shi, J.: Silhouette-Based Human Identification from Body Shape and Gait. In: Proc. Int. Conf. Automatic Face and Gesture Recognition (2002)
9. Veres, G.V., et al.: What image information is important in silhouette-based gait recognition. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition (2004)
10. Kale, A., Rajagopalan, A.N., Sundaresan, A., Cuntoor, N., RoyChowdhury, A., Kruger, V., Chellappa, R.: Identification of Humans Using Gait, Image Processing. IEEE Transactions on Publication 13(9), 1163–1173 (2004)
11. Phillips, P.J., Sarkar, S., Robledo, I., Grother, P., Bowyer, K.: The Gait Identification Challenge Problem: Data Sets and Baseline Algorithm. In: 16th ICPR, pp. 385–389 (2002)
12. Gross, R., Shi, J.: The CMU Motion of Body (MoBo) Database, tech. report CMU-RI-TR-01-18, Robotics Institute, Carnegie Mellon University (2001)
13. Shutler, J.D., Grant, M.G., Nixon, M.S., Carter, J.N.: On a Large Sequence-Based Human Gait Database. In: Proc. 4th Int. Conf. RASC, Nottingham (UK) (2002)
14. Spencer, N., And Carter, J.: Towards Pose Invariant Gait Reconstruction. In: IEEE International Conference On Image Processing, vol. 3, pp. 261–264 (2005)
15. Kale, A., Chowdhury, R., Chellappa, R.: Towards a view invariant gait recognition algorithm. In AVSS 2003, pp. 143–150 (2003)
16. Lynnerup, N.: Person Identification by Gait Analysis & Photogrammetry. In: Proc. 2nd Annual Conference Crime Solutions (2006)
17. Yamada, Y.: Advanced method for improvement of obscure video image Security Technology. In: Proceedings. IEEE 33rd Annual 1999 International Carnahan Conference on, pp. 440–445 (1999)
18. Dick, A.R., Brooks, M.J.: Issues in automated visual surveillance. In: Proc. VIIth Digital Image Comp. Tech. and App. pp.195–204, Sydney, (December 2003)
19. Kirtley, C.: Clinical Gait Analysis: Theory and Practice (June 2005) ISBN-10: 0443100098
20. Abdel-Aziz, Y.I., Arara, H.M.: Direct Linear Transformation from comparator coordinates into object space coordinates in close range photogrammetry. In: Proceedings of the Symposium on Close-Range Photogrammetry pp. 1–18 (1971)

21. Hartley, R.I., Zisserman, A.: Multiple View Geometry in Computer Vision. Cambridge University Press, Cambridge (2004)
22. Zhang, Z.: A flexible new technique for camera calibration, Pattern Analysis and Machine Intelligence. IEEE Transactions 22(11), 1330–1334 (2000)