Markerless View Independent Gait Analysis with Self-camera Calibration

Michela Goffredo, Richard D. Seely, John N. Carter and Mark S. Nixon
ECS, University of Southampton (UK)
{mg2,rds06r,jnc,msn}@ecs.soton.ac.uk

Abstract

We present a new method for viewpoint independent markerless gait analysis. The system uses a single camera, does not require camera calibration and works with a wide range of directions of walking. These properties make the proposed method particularly suitable for identification by gait, where the advantages of completely unobtrusiveness, remoteness and covertness of the biometric system preclude the availability of camera information and use of marker based technology. Tests on more than 200 video sequences with subjects walking freely along different walking directions have been performed. The obtained results show that markerless gait analysis can be achieved without any knowledge of internal or external camera parameters and that the obtained data that can be used for gait biometrics purposes. The performance of the proposed method is particularly encouraging for its appliance in surveillance scenarios.

1. Introduction

Gait biometrics, which concerns recognizing individuals by the way they walk, is a relatively new research field where the advantages of a completely unobtrusive system without any subject's cooperation or contact for data acquisition make it particularly attractive. There is a rich literature of various gait recognition techniques that can be divided into 3D [3, 12] and 2D approaches [9, 18]. In the first group, identification relies on parameters extracted from the 3D body deformation during walking. These methods use several digital cameras and the 3D reconstruction is achieved after a camera calibration process. On the other hand, the majority of 2D gait biometric approaches use a single camera positioned perpendicular to the subject walking direction (sagittal plane). However, in real surveillance scenarios, we need a system that operates in an unconstrained environment where maybe there is no information regarding the single camera and where the subject walks freely. 2D view independent gait identification is a recent research area and the approaches found in literature extract some gait parameters which are independent from the human pose by analysing the silhouette shape over time or aim at synthetising the sagittal view of the human body from any other arbitrary views [8, 11]. The latter choice is justified by the fact that the lateral view has proven the recognition capability in a great number of works [9, 15, 18]. However, they need information on the camera parameters [11] or rely on the use of reflective markers on the body [8].

The aim of this paper is to present a novel 2D markerless view independent gait analysis algorithm which does not need camera calibration or pre-knowledge of subject pose so that it can be used for gait identification in scenarios where only a single remote camera view is available. Human motion estimation literature is rich of different approaches, as detailed in [7, 13], and the methods oriented to gait tracking are based on models [4, 17], optical flow [10], non-linear regression of joint's angles [2] or present off-line training phases for generating shape models [16].
In our approach, information on the human silhouette and on the human body anthropometric proportions [6] are used for the estimation of the lower limbs' pose in the image reference system with low computational cost. After the markeless motion estimation, the obtained angles' trends are corrected by the view point independent gait reconstruction algorithm proposed by the authors in [8], that allows to reconstruct the limbs' pose in the sagittal plane. This work is part of a wider project where a reliable and robust silhouette extraction will be achieved in a completely unconstrained environment by analysing the images at different wavelength from an hyperspectral video camera.

2. Theory

The proposed gait analysis is based on two consecutive steps: a motion estimation method which extracts the limb's orientations with respect to the image reference system and a view-point independent gait reconstruction algorithm that normalises and corrects the limbs inclinations in the lateral reference system [8].

2.1. Markerless gait analysis

Let $\mathbf{S}(\mathbf{x}, \mathbf{y}, t_\varphi)$ be the RGB frame (of size RxC pixels) at time $\mathbf{T} = [t_1, t_2, \ldots, t_\varphi, \ldots, t_F]$ where (x_1, y_1) is the top-left corner of the image. We use a standard background subtraction algorithm for RGB images whilst the hyperspectral version is not yet available. Therefore, the binary image $\overline{\mathbf{S}}(\mathbf{x}, \mathbf{y}, t_\varphi)$, containing the human silhouette in $(\mathbf{x}_{sil}, \mathbf{y}_{sil})$, has been extracted as the object with maximum area. The lower limbs pose estimation algorithm is based on the proportions of the human body segments, following the medical results of anatomical studies [6]:

$$
\begin{aligned}
y'_{hip} &= \min(\mathbf{y}_{sil}) + 0.5 \cdot H \\
y'_{knee} &= \min(\mathbf{y}_{sil}) + 0.75 \cdot H \\
y'_{ankle} &= \min(\mathbf{y}_{sil}) + 0.90 \cdot H
\end{aligned}
\tag{1}
$$

where H is the silhouette's height.

Subsequently, the shins and thighs have been analysed separately. Let $\hat{\mathbf{X}}$ be the set of horizontal coordinates where $\overline{\mathbf{S}}(\hat{\mathbf{X}}, y'_{knee}, \hat{t}_\delta) = 1$, then the subset of frames $\hat{\mathbf{T}} = [\hat{t}_1, \hat{t}_2, \ldots, \hat{t}_\delta, \ldots, \hat{t}_{D \leq F}]$ where the legs do not overlap are those where $\hat{\mathbf{X}}$ forms 2 single connected components larger than $0.02H$ pixels.

Consequently, for each frame $\hat{t}_\delta$, the shin extraction algorithm, based on the linear approximation of the skeleton of the portion of image that includes the lower leg, is applied on $\overline{\mathbf{S}}(\mathbf{x}, \mathbf{y}, \hat{t}_\delta)$. The sub-images containing the shins (labeled as $\mathbf{S}_{shin1}$ and $\mathbf{S}_{shin2}$) are extracted from

$$
\mathbf{S}_{lower\ leg}(\mathbf{x}, \mathbf{y}, \hat{t}_\delta) = \begin{cases} \overline{\mathbf{S}}(\mathbf{x}, y_s, \hat{t}_\delta) & \text{if } y_s \in \mathbf{y}_{shin} \\ 0 & \text{otherwise} \end{cases}
\tag{2}
$$

where the vertical coordinates belonging to the two shins $\mathbf{y}_{shin} = [y_1, y_2, \ldots, y_s, \ldots, y_S]$ are defined in the following way:

$$
\mathbf{y}_{shin} = \left[y'_{knee}, y'_{knee} + 1, \cdots, \min(y'_{ankle}, y_{overlap} - 1) \right]
\tag{3}
$$

which includes the possibility of feet overlapping with $y_{overlap}$: vertical coordinate where the set of correspondent horizontal coordinates $\mathbf{x}_{overlap}$ belong to a single object larger than $0.02H$ pixels.

Therefore, the left and right shins skeletons are defined by the pixels with coordinates

$$
\begin{cases} \mathbf{x}^{skel}_{shin\ell}(\hat{t}_\delta) = [x_{1\ell}, x_{2\ell}, \ldots, x_{s\ell}, \ldots, x_{S\ell}] \\ \mathbf{y}^{skel}_{shin\ell}(\hat{t}_\delta) = \mathbf{y}_{shin} \end{cases} \quad \ell = \{1, 2\}
\tag{4}
$$

where

$$
x_{s\ell} = \sum_{j=1}^{C} x_j \cdot \mathbf{S}_{shin\ell}(x_j, y_s, \hat{t}_\delta) \Big/ \sum_{j=1}^{C} \mathbf{S}_{shin\ell}(x_j, y_s, \hat{t}_\delta)
\tag{5}
$$

Then, the shins can be linearly approximated by the first order polynomial with coefficients

$$
\mathbf{p}_\ell\left(\mathbf{x}^{skel}_{shin\ell}, \hat{t}_\delta\right) = p_{\ell 0}(\hat{t}_\delta) + p_{\ell 1}(\hat{t}_\delta) \cdot \mathbf{x}^{skel}_{shin\ell}(\hat{t}_\delta)
\tag{6}
$$

Therefore, the angle between the shins and the vertical axis (called knee angle) at frame $\hat{t}_\delta$ is $\alpha_\ell(\hat{t}_\delta) = \pi - \arctan\left[p_{\ell 1}(\hat{t}_\delta)\right]$. The procedure is applied for each frame in the subset $\hat{\mathbf{T}} = [\hat{t}_1, \hat{t}_2, \ldots, \hat{t}_\delta, \ldots, \hat{t}_D]$.

In order to also be able to estimate α_1 and α_2 when the legs overlap, the extracted shins are labeled as 'left' and 'right' according to the following rationale without losing generalities: the subject is walking for at least two gait cycles; in t_1 the shin labeled as $shin1$ belongs to the right leg; the labels swap after the legs overlap.

Subsequently, a 3rd order polynomial interpolation of α_1 and α_2 has been applied and the knees angles are achieved during the whole video sequence. This choice has been determined experimentally.

The upper legs orientation is extracted at every frame $\mathbf{T} = [t_1, t_2, \ldots, t_\varphi, \ldots, t_F]$ with a coarse to fine hips estimation procedure where at first, the hips position is achieved with

$$
\begin{cases} x'_{hip\ell} = \frac{1}{P} \cdot \sum_{j=1}^{P} \widetilde{x}_j + (2\ell - 3) \cdot H \cdot \mu \cdot 10^{-3} \\ y'_{hip\ell} = y'_{hip} \cdot (2\ell - 3) \cdot \left(\frac{\widetilde{x}_P - \widetilde{x}_1}{2}\right) \cdot \sin(0.3 \cdot \mu) \end{cases}
\tag{7}
$$

where $\widetilde{\mathbf{X}} = [\widetilde{x}_1, \widetilde{x}_2, \ldots, \widetilde{x}_j, \ldots, \widetilde{x}_P]$ is the subset of P ($P \leq C$) horizontal coordinates such as $\overline{\mathbf{S}}(\widetilde{\mathbf{X}}, y'_{hip}, t_i) = 1$.

Equations 7 puts in relationship the horizontal hips position and the walking direction μ, calculated with respect to the horizontal axes of the image reference system. These relationships have been obtained with a regressive analysis of the 3D Georgia Tech motion capture data [1] by considering different camera positions.

The angle μ is estimated as the angle of inclination of the straight line which approximates the heel strikes points with coordinates:

$$
(\mathbf{x}_{str}, \mathbf{y}_{str}) = (\mathbf{x}_{feet}, \mathbf{y}_{feet}) \mid \sum_{i=1}^{P} SK(\mathbf{x}_{feet}, \mathbf{y}_{feet}, t_i) > \tau
\tag{8}
$$

where SK is the skeleton of the image as in [14], $\mathbf{x}_{feet}$ and $\mathbf{y}_{feet}$ are the coordinates corresponding to the portions of silhouette belonging to the feet and the threshold τ is automatically set so that at least 4 heels strikes are extracted (in accordance with the hypothesis listed above).

Subsequently, a finer hips pose estimation is obtained with a linear approximation of the thighs by the first order polynomial with coefficients

$$
\begin{aligned}
q_{\ell 0}(t_\varphi) &= \frac{y^m_{shin\ell} \cdot x'_{hip\ell} - y'_{hip\ell} \cdot x^m_{shin\ell}}{x'_{hip\ell} + x^m_{shin\ell}} \\
q_{\ell 1}(t_\varphi) &= \frac{1}{2} \cdot \left(a'_\ell + \frac{y'_{hip\ell} - q_{\ell 0}}{x'_{hip\ell}}\right)
\end{aligned}
\tag{9}
$$

where $y^m_{shin\ell} = \min(\mathbf{y}_{shin})$ and $x^m_{shin\ell}$ is the mean value of the correspondent horizontal coordinates. a'_1 and a'_2 are the
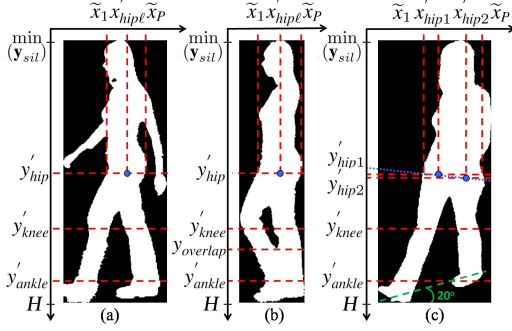
Figure 1. Markerless gait analysis: $\mu = 0^o$ view with non-overlapped (a) and overlapped (b) feet; $\mu = 20^o$ view (c)

slope of the straight lines that approximate respectively the left and right edges of the positions of silhouette belonging to the thighs.

Therefore, the angle between the thighs and the vertical axis at every frame t_φ is $\gamma_\ell(t_\varphi) = \pi - \arctan[q_{\ell 1}(t_\varphi)]$.

Ultimately, the thighs are labeled as 'left' or 'right'in accordance with the corresponding shin labeling. Figure 1 shows the shins and thighs extraction procedure.

The angles α_1, α_2, γ_1 and γ_2 are then corrected by projecting the limbs trajectories on the lateral plane, as explained in the following paragraph.

2.2. View-point reconstruction

The method is based on four main assumptions: the nature of human gait is cyclic; subjects walk along a straight line; the distances between the bone joints are constant; and the articulated leg motion is approximately planar. Therefore, multiple periods of linear gait motion appear analogous to a single period viewed from many cameras related by linear translation and the positions of the legs joints lie in an auto-epipolar configuration.

If $\mathbf{j}_i^\ell$ is 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 $\mathbf{j}_i^\ell \times \mathbf{P}_i \cdot \mathbf{J}^\ell = 0$, where $\mathbf{P}_i = \left[\mathbf{R}_\mathbf{e}^T, -i\mathbf{e}_0\right]$ and $\mathbf{R}_\mathbf{e}^T$ is the rotation matrix for aligning the epipolar vector $\mathbf{e}_0$ with the horizontal axis X. Then,

$$\mathbf{j}_i^\ell = \mathbf{P}_i \begin{pmatrix} 1 & 0 \\ 0 & \mathbf{H}_\mathbf{V}^{-1} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \mathbf{H}_\mathbf{V} \end{pmatrix} = \mathbf{H} \cdot \mathbf{J}^\ell \qquad (10)$$

having expressed the limb plane transformation matrix with $\mathbf{H}_\mathbf{V}$ so that the two cross section plane lines are centred and normalised respect to Y and Z axes and parallel with Y. By assuming the lengths of the articulated limbs $\mathbf{D}_\ell^2 = \Delta\mathbf{j}_i^{\ell T}\Delta\mathbf{j}_i^\ell$ 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

$$\Delta\mathbf{j}_i^{\ell T} \cdot \mathbf{H}^T \cdot \mathbf{H} \cdot \Delta\mathbf{j}_i^\ell = \Delta\mathbf{j}_{i+1}^{\ell T} \cdot \mathbf{H}^T \cdot \mathbf{H} \cdot \Delta\mathbf{j}_{i+1}^\ell \qquad (11)$$

Figure 3. Limb's pose estimation with bad background subtraction

After recovering the fronto-parallel structure of subject gait, the representation of the leg joints function $\left[\mathbf{J}_x^\ell(t), \mathbf{J}_y^\ell(t)\right]$ is found by fitting a modified Fourier series to the data with fixed fundamental frequency f_0 and period T:

$$\mathbf{J}_x^\ell(t) = v_x t + \sum_{k=1}^n A_k \cos\left(2\pi k f_0 \left(t + \frac{(\ell-1)T}{2}\right) + \phi_k\right) + \mathbf{J}_{x0}^\ell \qquad (12)$$

analogously for $\mathbf{J}_y^\ell(t)$. Thus, the projection of the leg joints on the lateral plane is obtained with an optimized procedure in the following way

$$\breve{\mathbf{J}}^\ell(t) = \begin{bmatrix} h_1 & h_2 & h_3 \end{bmatrix} g\left(t + \frac{(\ell-1)T}{2} : f_0, \mathbf{D}_\ell, v_x, v_y, F\right) \qquad (13)$$

where $g(t)$ is the bilateral Fourier series function with coefficients F and h are the values of the inverse normalization transform matrix. Therefore, starting from a video sequence from a single camera and without any calibration, the proposed markerless system, in junction with [8], estimates the gait parameters projected on the lateral plane.

3. Experimental tests

The proposed method has been tested on 3 subjects walking freely along a straight line in a $3x5m^2$ area. The video sequences have been acquired with a non-calibrated digital camera, FLEA IEEE-1394 Digital Camera (Point Grey Research) with 4mm lens, positioned in 6 different orientation respect to the walking line: 3 different azimuth, AZ, $(0^o, 20^o, 40^o)$ and 2 different elevation, EL, $(0^o, 15^o)$ angles. Moreover different frame rates (from 5 fps to 30 fps) and frame sizes (from 170x128 pixels to 1024x768 pixels) have been considered.

3.1. Results

The new method has been applied to the video sequences, the limbs pose has been estimated frame by frame and the hip and knee angles have been extracted for each camera position and for each subject. Figure 2 shows an example of the limbs pose estimation in the 6 directions. The algorithm allows to estimate the limb's pose and appears particularly robust respect to noisy and broken silhouettes (figure 3). A quantitative validation has been obtained in a virtual environment (Poser 7®) with a humanoid walking for one gait cycle. The results in estimating the leg's angles,
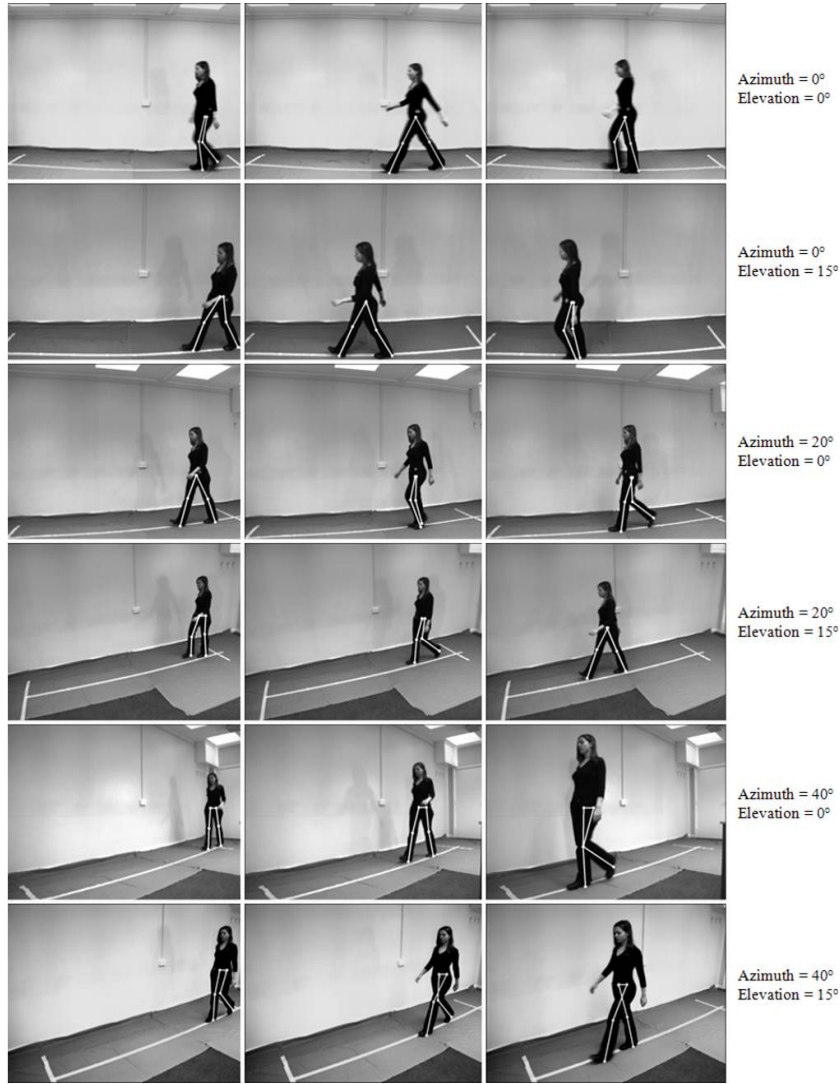
Figure 2. Markerless joints extraction in the different view points

with a mean value of 2.63 ± 2.61 deg, are particularly encouraging and present same magnitude to the ones obtained with 3D markerless systems and 2D complex model based methods [2, 5].

Figure 4(a) shows the variations of γ and α during two gait cycle for the six different camera positions in the real experimental tests. Predictably, the angles trends, extracted in the image reference system, are influenced by the subject pose respect to the camera and they cannot be used directly for biometric identification. For this reason, the view point correction algorithm is applied and the angle trends after the correction are shown in figure 4(b). The slight variations between the resulting traces in figure 4(b) are consistent with intra-subject variation between the video acquisitions.

In order to quantify the angles matching after the correction, two measures have been considered: the Mean Correlation Coefficient (MCC) of the 3 subjects along the 6 different directions; the Root Mean Square Distance (RMS) between the joints positions, estimated with the markerless algorithm, and the projected ones, after the view point correction, as in [8]. Table 1 shows the variation of the MCCs with respect to the camera AZ and EL. The results, with a mean value of 0.919, are consistent with the value of MCC obtained using reflective markers [8]. Therefore, the correlation values obtained with the angles trends estimated with the proposed method, appear particularly encouraging. The MCC peaks, corresponding to AZ less of 20° and EL 0°, confirm that the corrected angles are correlated with the lateral angle trends. Conversely, the angles obtained with the camera positioned with AZ equal to 40° present the lowest values of MCCs, due to the almost frontal view of the subject walking. Anyway, these values are still closer to the
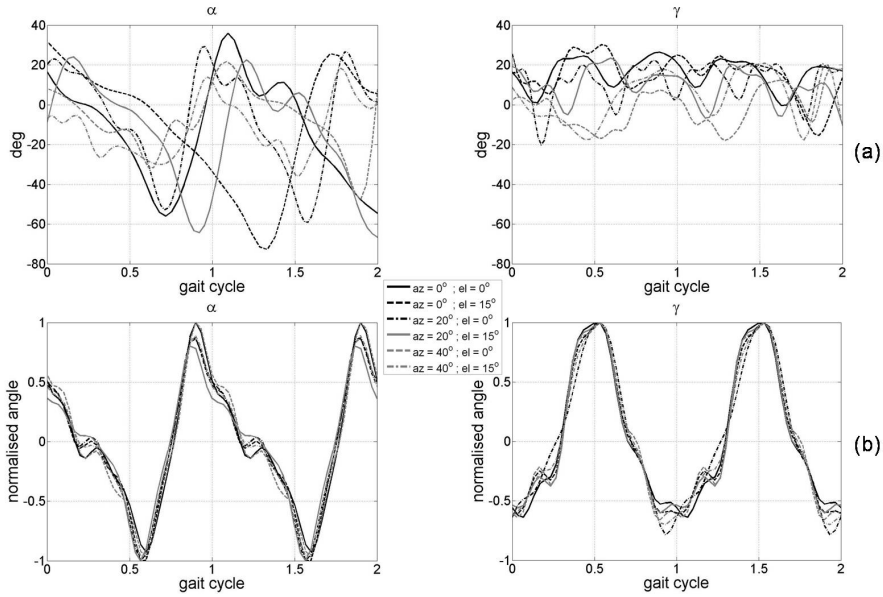
Figure 4. Hip (γ) and knee (α) angles in different camera positions (subject 1), unprocessed (a) and corrected (b)

AZ (deg)	EL (deg)	α MCC	γ MCC	RMS(%)
0	0	0.974	0.972	0.995
0	15	0.975	0.970	1.242
20	0	0.982	0.960	1.441
20	15	0.977	0.968	0.799
40	0	0.959	0.927	2.669
40	15	0.960	0.941	1.754

Table 1. MCC and RMS with respect to the camera positions

F	MLL	$m_{\mathbf{e}}$	$\sigma_{\mathbf{e}}$	MCC	RMS(%)
1	225	0	0	0.919	1.154
2	112	0.789	2.39	0.917	1.289
3	75	0.945	2.645	0.896	2.681
4	56	2.320	7.809	-	-
5	45	2.485	6.941	-	-
6	38	3.148	8.342	-	-

Table 2. $m_{\mathbf{e}}$, $\sigma_{\mathbf{e}}$, MCC and RMS respect to image resolution

mean value and greater than 0.9, so they can be considered acceptable. The mean RMS error for the 3 subjects is 1.4% of the image resolution. Encouraginlgy, this result is only slightly higher than the one obtained with reflective markers by [8] and it suggests that the RMS can be considered correlated with the error of the markerless joints extraction.

The video sequences spatial and temporal characteristics have been modified in order to test the performance of the proposed method under different video setups. Firstly, the sensitivity of the method with respect to the image resolution has been evaluated. The frame size has been scaled-down with a factor $F = [1, 2, 3, 4, 5, 6]$ that corresponds to to the mean leg's length, MLL, varying from 225 to 38 pixels. The markerless algorithm has been applied to the reduced video sequences and the estimated angle trends have been compared with respect to the ones obtained from the full resolution videos. Table 2 shows how much the mean angles trends, expressed in terms of mean value ($m_{\mathbf{e}}$) and standard deviation ($\sigma_{\mathbf{e}}$), differ from those extracted with full resolution and that about 70 pixels can be considered as the
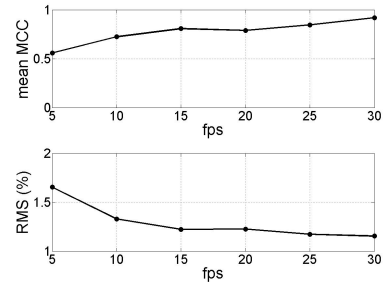

Figure 5. MCC and RMS with respect to camera frame rate

lower MLL spatial limit so that the extracted angles can be considered plausible. For this reasons table 2 also shows the mean MCC and RMS values calculated only for the first three image resolutions. It is notable how the image resolution remarkably influences both the mean MCC and RMS and how, on the other hand, the mean MCC remains higher than 0.8.

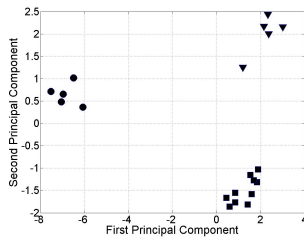Furthermore, the sensitivity of the proposed method to

Figure 6. Cluster analysis on 3 subjects

camera frame rate has been evaluated and the mean MCC and RMS (figure 5) show that a frame rate greater that 15 fps can be considered acceptable for gait identification purpose.

In order to demonstrate the applicability of the proposed approach in a biometric context and that the view-point reconstruction algorithm does not remove the inter-subject gait characterization, a cluster analysis on the PCA of the spectral characteristics of the angles has been achieved. Figure 6 shows the first and the second components and even if the number of subjects does not allow an analysis of CCR. The clustering appears particularly encouraging for future tests on a larger database and suggests that individuality is reached.

4. Conclusions

The remote and covert nature of gait as a biometric has continued the interest in this field. However, most approaches to date limit their application to the fronto-parallel plane, or use camera calibration or even require the use of reflective markers attached to the subject. Considering real scenarios, a system which does not use markers and which estimates the the gait parameters without relying on the subjects pose or on camera calibration is required. Therefore, this paper has introduced a novel method for a view-independent markerless gait analysis based on the anthropometric proportions of human limbs and the characteristics of the gait task. Experimental tests have been conducted on over 200 video sequences 3 subjects viewed at 6 different camera inclinations. Different spatial resolutions and frame rates have been considered and demonstrated the robustness of the method. The results are consistent with those obtained with reflective markers and then particularly encouraging for application in real scenarios. Further, we confirm intuition in that the most discriminative camera view includes shoulders, torso and legs in contrast the the plethora of studies which concentrate on images where the plane of the camera view is normal to the direction of walking. As such, our new approach is completely viewpoint invariant, though discriminability reduces as the camera moves to view a subject walking from the front.

5. Acknowledgements

References

[1] www.cc.gatech.edu/cpl/projects/hid.

[2] A. Agarwal and B. Triggs. Recovering 3d human pose from monocular images. *IEEE TPAMI*, 28:44–58, 2006.

[3] B. Bhanu and J. Han. Human recognition on combining kinematic and stationary features. In *Proc Int Conf AVBPA*, pages 600–608, 2003.

[4] Z. Chen and H.-J. Lee. Knowledge-guided visual perception of 3-d human gait from a single image sequence. *IEEE Tran SMC*, 22(2):336–342, 1992.

[5] S. Corazza, L. Mndermann, A. M. Chaudhari, T. Demattio, C. Cobelli, and T. P. Andriacchi. A markerless motion capture system to study musculoskeletal biomechanics: Visual hull and simulated annealing approach. *ABME*, 6:1019–1029, 2006.

[6] W. T. Dempster and G. R. L. Gaughran. Properties of body segments based on size and weight. *Am J Anat*, 120:33–54, 1965.

[7] D. A. Forsyth, O. Arikan, L. Ikemoto, J. O'Brien, and D. Ramanan. Computational studies of human motion: part 1, tracking and motion synthesis. *J FTCGV*, 1:77–254, 2005.

[8] M. Goffredo, N. Spencer, D. Pearce, J. N. Carter, and M. S. Nixon. *Human Perambulation as a Self Calibrating Biometric.*, volume 4778 of *LNCS*. Springer, 2007.

[9] P. Huang, C. Harris, and M. Nixon. Recognising humans by gait via parametric canonical space. *J Artif Intell Eng*, 13(4):359–366, Nov. 1999.

[10] S. Ju, M. Black, and Y. Yacoob. Cardboard people: a parameterized model of articulated image motion. *Proc IEEE Conf AFG*, pages 38–44, 1996.

[11] A. Kale, A. Chowdhury, and R. Chellappa. Towards a view invariant gait recognition algorithm. *Proc IEEE Conf AVSS*, pages 143–150, 2003.

[12] L. Middleton, D. K. Wagg, A. I. Bazin, J. N. Carter, and M. S. Nixon. Developing a non-intrusive biometric environment. *Proc IEEE Int Conf IROS*, pages 723–728, 2006.

[13] T. B. Moeslund and E. Granum. A survey of computer vision-based human motion capture. *J CVIU*, 81(3):231–268, 2001.

[14] M. Nixon and A. Aguado. *Feature Extraction & Image Processing*. Academic Press, 2nd edition, Dec. 2007.

[15] M. S. Nixon and J. N. Carter. Automatic recognition by gait. *Proc. of the IEEE*, 94(11):2013–2024, Nov. 2006.

[16] C. Orrite-Urunuela, J. del Rincon, J. Herrero-Jaraba, and G. Rogez. 2d silhouette and 3d skeletal models for human detection and tracking. *Proc ICPR*, 4:244–247, 2004.

[17] J. M. Wang, D. J. Fleet, and A. Hertzmann. Gaussian process dynamical models for human motion. *IEEE TPAMI*, 30(2):283–298, 2008.

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