

2D Markerless Gait Analysis

Michela Goffredo¹, John N. Carter¹ and Mark S. Nixon¹

¹ ECS, University of Southampton (United Kingdom)

Abstract— We present a 2D gait analysis system which is completely markerless and extracts kinematic information by analyzing video sequences obtained from an RGB video camera. These properties make the proposed approach particularly suitable in medical contexts where visual gait observation is still a recognised procedure or the invasiveness and high costs of marker-based systems can not be afforded. Markerless motion estimation literature for medical gait analysis is generally 2D oriented, since the majority of joints dysfunctions related to gait occur in the sagittal plane. Most of the approaches are based on time consuming human body models or need human-intervention. Conversely, the method we present this contribution is silhouette-based, completely automatic and uses information on the human body anthropometric proportions for the estimation of the lower limbs' pose in the sagittal plane with good accuracy and low computational cost. Tests on a large number of synthetic and real video sequences with normal gait have been performed. Different frame rates, image resolutions and noises have been considered. The obtained results, in terms of sagittal joint angles, have been compared with the typical trends found in biomechanical studies. The performance of the proposed method is particularly encouraging for its appliance in the real medical context.

Keywords— Markerless, Gait Analysis, Image Processing, 2D Human Motion Analysis.

I. INTRODUCTION

Clinical gait analysis aims at learning about a condition affecting different groups of patients or the effect of an intervention or a rehabilitation protocol [1][2]. To this purpose, quantitative kinematic measurements such as joint angular rotations, are generally accomplished by means of devices (i.e. markers or electrogoniometries) applied on the subject during the movement execution [3][4][5]. These systems are particularly accurate and widely accepted by the medical community.

In clinical practice, however, qualitative visual gait observation is still widely utilised because it is cost efficient, quick, and easy to use in comparison to device-based gait analysis [6][7]. Moreover, the device-free approach does not affect the function it is measuring: the walking performed with the presence of devices applied on the patients' skin, is not necessarily representative of their

natural walking.

Computer vision has been recently involved into the design of specific systems for the device-free kinematic gait analysis by analysing image sequences from video cameras. Since the video processing approach does not rely on the use of markers, this research area has been addressed as “markerless” [8]. Some approaches found in literature are based on the visual hull reconstruction and an a priori 3D human model by using a large number of video cameras [9][10]. However, in clinical gait analysis, since the majority of gait pathological knee, ankle and foot dysfunctions occur in the sagittal plane [11], markerless lower limbs' motion estimation literature is generally 2D oriented. The 2D approaches can be grouped into model-based [12][13] and semi-automatic features-based [14][15].

We present a novel markerless method for a markerless gait analysis in the sagittal plane without human intervention. Our approach is silhouette-based and uses information on the human body anthropometric proportions for the estimation of the lower limbs' pose with good accuracy and low computational cost.

II. THE NEW METHOD

Our approach automatically processes video sequences gathered with a digital camera and extract the 2D joints' location and their angle trend over time.

Let $\mathbf{I}_{\text{RGB}}(\mathbf{x}, \mathbf{y}, t_\varphi)$ be the RGB frame (of size $R \times C$ pixels) at time $\mathbf{T}=[t_1, t_2, \dots, t_\varphi, \dots, t_F]$, where (x_l, y_l) is the top-left corner of the image. By applying a background subtraction method based on the threshold of the 3 components of the color space YUV [16], the binary image $\mathbf{S}(\mathbf{x}, \mathbf{y}, t_\varphi)$ has been extracted. The pixels $(\mathbf{x}_{sil}, \mathbf{y}_{sil})$ containing the human silhouette have been selected as the object with maximum area [17].

The lower limbs pose estimation algorithm is based on the proportions of the human body segments, following the results of anatomical studies [18]. Particularly, the vertical positions of hips, knees and ankles with respect to the silhouette height H are assumed in first approximation to be

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

$$y'_{ankle} = \min(\mathbf{y}_{sil}) + 0.90 \cdot H$$

where H is the silhouette's height.

Let \mathbf{X}^* be the set of horizontal coordinates where $\mathbf{S}(\mathbf{X}^*, y'_{knee}, t_\varphi) = 1$, then the subset of frames $\mathbf{T}^* = [t_1^*, t_2^*, \dots, t_\delta^*, \dots, t_{D < F}^*]$ where the legs do not overlap are those where \mathbf{X}^* forms 2 single connected components larger than $0.02H$ pixels. Consequently, for each frame t_δ^* , the shin extraction algorithm, based on the linear approximation of the skeleton [19] of the portion of image that includes the lower leg, is applied on $\mathbf{S}(\mathbf{x}, \mathbf{y}, t_\delta^*)$. The sub-images containing the shins (labelled as \mathbf{S}_{shin1} and \mathbf{S}_{shin2}) are extracted from

$$\mathbf{S}_{lower\ leg}(\mathbf{x}, \mathbf{y}, t_\delta^*) = \begin{cases} \mathbf{S}(\mathbf{x}, y_s, t_\delta^*) & \text{if } y_s \in \mathbf{y}_{shin} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

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

$$\mathbf{y}_{shin} = [y_{knee}^*, y_{knee}^* + 1, \dots, \min(y_{ankle}^*, y_{overlap} - 1)] \quad (3)$$

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 right ($\ell=1$) and left ($\ell=2$) shins skeletons are defined by the pixels with coordinates

$$\begin{cases} \mathbf{x}_{shin\ell}^{skel}(t_\delta^*) = [x_{1\ell}, x_{2\ell}, \dots, x_{s\ell}, \dots, x_{s\ell}] \\ \mathbf{y}_{shin\ell}^{skel}(t_\delta^*) = \mathbf{y}_{shin} \end{cases} \quad (4)$$

$$\ell = \{1, 2\}$$

where

$$x_{s\ell} = \frac{\sum_{j=1}^C x_j \cdot \mathbf{S}_{shin\ell}(x_j, y_s, t_\delta^*)}{\sum_{j=1}^C \mathbf{S}_{shin\ell}(x_j, y_s, t_\delta^*)} \quad (5)$$

Then, the shins are linearly approximated by the straight line

$$\mathbf{p}_\ell(\mathbf{x}_{shin\ell}^{skel}, t_\delta^*) = p_{\ell 0}(t_\delta^*) + p_{\ell 1}(t_\delta^*) \cdot \mathbf{x}_{shin\ell}^{skel}(t_\delta^*) \quad (6)$$

The knee angle at frame t_δ^* is $\alpha_\ell(t_\delta^*) = \pi - \arctan[p_{\ell 1}(t_\delta^*)]$. The procedure is applied for each frame in the subset $\mathbf{T}^* = [t_1^*, t_2^*, \dots, t_\delta^*, \dots, t_{D < F}^*]$.

In order to also be able to estimate α_1 and α_2 when the legs overlap, the extracted shins are

labelled as “left” or “right” according to the following rationale without losing generalities: in t_1 the shin labelled 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 resolved during the whole video sequence. This choice has been determined experimentally and considering the spectral characteristics of human gait [20].

The upper legs orientation is also extracted at every frame $\mathbf{T} = [t_1, t_2, \dots, t_\varphi, \dots, t_F]$ with a coarse to fine hips estimation procedure where at first, the hips position is achieved with coordinates $[x'_{hip}, y'_{hip}] = [x_1 + (x_P - x_1)/2, y'_{hip}]$. Subsequently, the finer hips pose estimation is obtained with a linear approximation of the thighs by the straight line

$$q_{\ell 0}(t_\varphi) = \frac{y_{shin\ell}^m \cdot x'_{hip\ell} - y'_{hip\ell} \cdot x_{shin\ell}^m}{x'_{hip\ell} + x_{shin\ell}^m} \quad (7)$$

$$q_{\ell 0}(t_\varphi) = \frac{1}{2} \cdot \left(a'_\ell + \frac{y'_{hip\ell} - q_{\ell 0}(t_\varphi)}{x'_{hip\ell}} \right)$$

where $y_{shin\ell}^m = \min(\mathbf{y}_{shin})$ and $x_{shin\ell}^m$ is the mean value of the correspondent horizontal coordinates. a'_1 and a'_2 are the slope of the straight lines that approximate the right and left edges of the positions of silhouette belonging to the thighs respectively.

Therefore, the hip angles are $\gamma_\ell(t_\varphi) = \pi - \arctan[q_{\ell 1}(t_\varphi)]$ at every frame t_φ .

Ultimately, the thighs are labelled as “left” or “right” in accordance with the correspondent shin labelling. Figure 1 shows the shins and thighs extraction procedure.

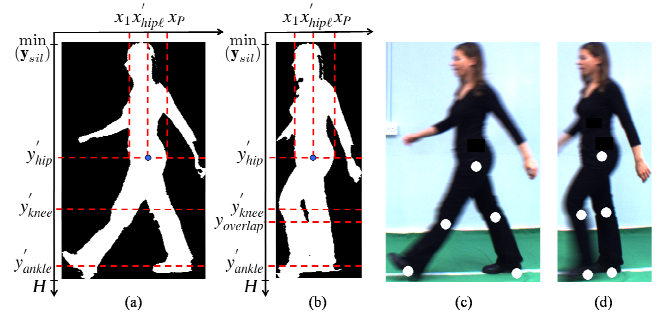


Fig. 1 Markerless gait analysis with non overlapped (a) and overlapped (b) feet and back-projected joints' position (c)(d).

III. EXPERIMENTAL TESTS

A quantitative validation of the proposed method for markerless gait analysis has been obtained in a virtual environment: synthetic video sequences (500x490 pixels, 30fps) of a humanoid walking for one gait cycle have been

created with Poser 7® (by e-frontier) in order to provide the ground truth for validation purposes.

Different spatial resolutions (250x245, 163x163 pixels) have been considered by sub-sampling the original images with factors 2 and 3. Table 1 shows the comparative results in terms of mean (m_e) and standard deviation (σ_e) of joint's angles compared to the ground truth over the gait cycle. The results, with a mean value of 2.63 ± 2.61 , are particularly encouraging and present same magnitude to the ones obtained with 3D markerless systems or 2D complex model based ones [9][12]. The image resolution and the silhouette's height reduce the algorithm performance, as predictable. On the other hand, a 50-pixels-height silhouette still maintains an acceptable result. Moreover, table 1 shows the mean time for processing one frame (Matlab7® code processed with 2GHz Intel CPU): the model-free proposed approach allows to extract the kinematic information with computational cost per silhouette's pixel lower than $2 \cdot 10^{-4}$ s.

Table 1 Angle's errors (in terms of mean and standard deviation) on synthetic gait with different spatial resolution.

Image resolution / mean sil's height	m_e (deg)	σ_e (deg)	Processing time (s/frame)
500x490/200	2.04	2.12	1.44
250x245/100	2.80	2.15	0.97
163x163/50	3.06	3.56	0.37

In addition, to simulate background subtraction imprecision, zero-mean Gaussian noise has been added to the synthetic images. The standard deviation (σ_n) of the added noise varies from 0 (original data) to 150 corresponding to PSNR [21] from ∞ to 7.01 dB. Figure 2 compares the original images with 2 noised images; the extracted silhouettes are also shown.

Results in Table 2 show m_e and σ_e with respect to the different PSNR. Obviously the error increases with the level of noise added to the frames but it is lower than 5 degrees even at high noise levels.

The noise robustness allows to understand how the proposed method depends on an accurate silhouette extraction and thus to extend the approach in a real context.

Additionally, our approach has been tested on real data where 3 young healthy subjects walk freely along a straight line in a $3 \times 5 \text{m}^2$ area. The video sequences have been acquired with a calibrated digital camera, FLEA IEEE-1394 Digital Camera (Point Grey Research) with 4mm lens, positioned perpendicular to the walking direction. The image resolution and the frame rate are 1024x768 pixels and 30fps respectively.

The limbs pose has been estimated frame by frame and the hip and knee angles have been extracted in the lateral plane with a processing time of 3.05s/frame. Figure 3 and 4 shows the obtained results in terms joints' position estimation and

angles' trend. Figure 3, in particular, depicts the joints' angle trends for the 3 subjects.

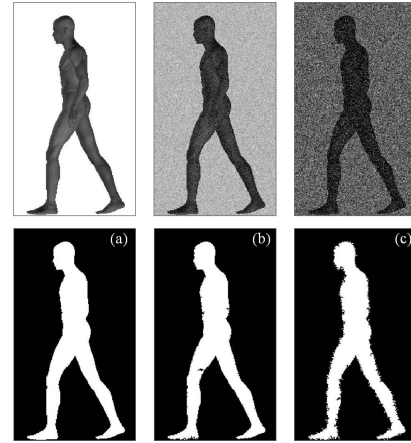


Fig. 2 Comparison between the original synthetic image (a) and the ones with zero-mean Gaussian noise with $\sigma_n=25$ (b) and $\sigma_n=100$ (c).

Table 2 Angle's errors (in terms of mean and standard deviation) on synthetic gait with different zero-mean Gaussian noises.

σ_n /PSNR(dB)	m_e (deg)	σ_e (deg)
0/ ∞	2.04	2.12
5/28.12	2.30	2.13
25/24.25	3.01	3.38
50/21.89	3.30	3.72
100/15.23	3.25	3.68
150/7.01	4.25	3.49

The results obtained via our approach are consistent with biomechanical data by Winter [22] shown in bold in figure 3. The variations between the resulting traces in figure 3 are consistent with intra-subject variability [23].

Figure 4 shows the joints' pose back-projected on the images (half gait cycle, subject 1): the proposed method extracts the limb's trajectories with good accuracy and solves efficiently the self-occlusion problems.

The obtained results encourage for future tests on a larger number of subjects and comparative analysis with marker-based systems.

IV. CONCLUSIONS

A new method for 2D markerless gait analysis has been presented. It extracts kinematic information by analyzing video sequences with a silhouette-based approach based on the human body anthropometric proportions.

Tests on synthetic and real data have shown good accuracy in estimating the limb's pose, low processing time and robustness with respect to different spatial resolution and noisy images. Comparisons with data found in

biomechanical studies are particularly encouraging for the appliance of the proposed method in the real medical context where the qualitative visual gait observation and the highly invasive and costly marker-based systems can not be adopted.

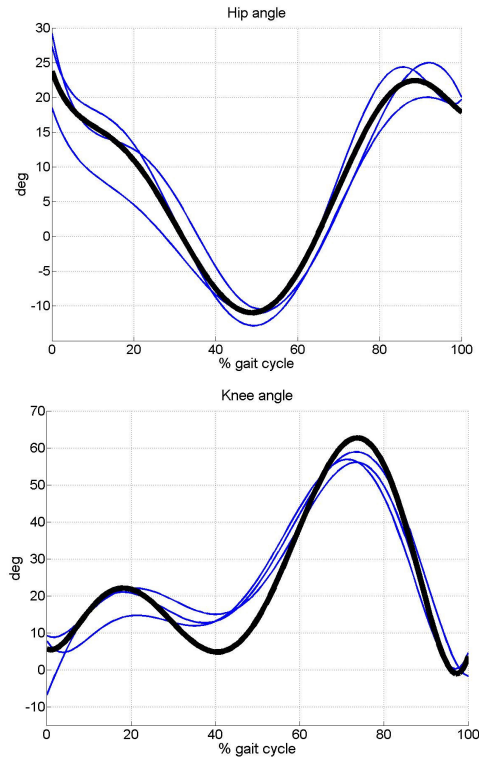


Fig. 3 Sagittal angles of hip and knee in the 3 normal subjects during one gait cycle. Comparison with biomechanical data by Winter (in bold) [22].

REFERENCES

- Baker R (2006) Gait analysis methods in rehabilitation. *J Neuroengineering Rehabil* 3:4
- Davis RB, DeLuca PA (1996) *Clinical Gait Analysis: Current Methods and Future Directions*. In *Human Motion Analysis*, GF Harris and PA Smith Editors, IEEE Press, Piscataway, New Jersey
- Sutherland DH (2002) The evolution of clinical gait analysis - Part II Kinematics. *Gait and Posture*, 16:2: 159-179
- Davis RI, Ounpuu S, Tyburski D, Gage J (1991) A gait data collection and reduction technique. *Hum Mov Sci*, 10:575-587
- Brady R, Pavol M, Owing T, Grabiner M (2000) Foot displacement but not velocity predicts the outcome of a slip induced in young subjects while walking. *J Biomech* 33:803-808
- The Pathokinesiology Department & The Physical Therapy Department Rancho Los Amigos Medical Center (1996) *Observational Gait Analysis Handbook*
- Lord SE, Halligan PW, Wade DT (1998) Visual gait analysis: the development of a clinical assessment and scale. *Clin Rehabil* 12:107-119
- Schrotter G (2005) Markerless tracking and surface measurements in biomechanical applications, *IASTED Proc., Int. Conf. on Robotics and Applications*, Cambridge, USA, 2005
- Corazza S, Mundermann L, Chaudhari A M, Demattio T, Cobelli C, Andriacchi T P (2006) A markerless motion capture system to study musculoskeletal biomechanics: visual hull and simulated annealing approach. *Annals of Biomedical Engineering* 34:6: 1019-29
- Zhanfeng Y, Liang Z, Chellappa R (2003) View synthesis of articulating humans using visual hull. *ICME Proc., vol. 1, Int. Conf. on Multimedia and Expo, Baltimore, USA, 2003*, pp 489-92
- Perry J (1992) *Gait Analysis: Normal and Pathological Function*. McGraw Hill, New York
- Saboune J, Charpillet F (2005) Markerless human motion capture for gait analysis. *EMBEC Proc., European Medical and Biological Engineering Conf. Advancement of Medicine and Health Care through Technology, Prague, 2005*
- Orrite-Urunuela C, Del Rincon J, Herrero-Jaraba J, Rogez G (2004) 2D silhouette and 3D skeletal models for human detection and tracking. *IEEE Proc., Int. Conf. on Pattern Recognition, vol. 4, pp 244-247, Cambridge, UK, 2004*
- Goffredo M, Schmid M, Conforto S, Carli M, Neri A, D'Alessio T (2005) Coarse-to-fine markerless gait analysis based on PCA and Gauss-Laguerre decomposition. *SPIE Proc., Medical Imaging: Image Processing, vol. 5747, pp.1076-1084, San Diego, USA, 2005*.
- Ju S, Black M, Yacoob Y (1996) Cardboard people: a parameterized model of articulated image motion. *IEEE Proc., Int. Conf. Automatic Face and Gesture*, pp 38-44, Killington, USA, 1996
- Wren CR, Azarbajejani A, Darrell T, Pentland AP (1997) Pfinder: real-time tracking of the human body. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 19:7:780-785
- Haralick RM, Shapiro LG (1992) *Computer and Robot Vision*, volume 1. Addison-Wesley
- Dempster WT, Gaughran GRL (1965) Properties of body segments based on size and weight. *American Journal of Anatomy* 120:33-54
- Nixon MS, Aguado A (2007) *Feature Extraction & Image Processing*. 2nd edition, Academic Press
- Murray MP, Drought AB, Kory RC (1964) Walking patterns of normal men. *J Bone Joint Surgery*, 46:2:335-360
- Netravali AN, Haskell BG (1995) *Digital Pictures: Representation, Compression, and Standards*. Plenum Press, New York
- Winter DA (1990) *The Biomechanics and Motor Control of Human Movement*. John Wiley & Sons
- Murray MP (1967) Gait as a Total Pattern of Movement. *Am J Physical Medicine* 46:1:290-329



Fig. 4 Markerless gait analysis on half gait cycle (subject 1)