

Towards View-Invariant Gait modeling: Computing View-Normalized Body Part Trajectories

Frédéric Jean ^{a,*} Alexandra Branzan Albu ^b Robert Bergevin ^a

^a*Computer Vision and Systems Laboratory, Dept. of Electrical and Computer Engineering, Laval University, Québec, QC, Canada, G1K 7P4*

^b*Laboratory for Applied Computer Vision Algorithms, Dept. of Electrical and Computer Engineering, University of Victoria, Victoria, BC, Canada, V8W 3P6*

Abstract

This paper proposes an approach to compute view-normalized body part trajectories of pedestrians walking on potentially non-linear paths. The proposed approach finds applications in gait modeling, gait biometrics, and in medical gait analysis. Our approach uses the 2D trajectories of both feet and of the head extracted from the tracked silhouettes. On that basis, it computes the apparent walking (sagittal) planes for each detected gait half-cycle. A homography transformation is then computed for each walking plane to make it appear as if walking was observed from a fronto-parallel view. Finally, each homography is applied to head and feet trajectories over each corresponding gait half-cycle. View normalization makes head and feet trajectories appear as if seen from a fronto-parallel viewpoint, which is assumed to be optimal for gait modeling purposes. The proposed approach is fully automatic as it requires neither manual initialization nor camera calibration. An extensive experimental evaluation of the proposed approach confirms the validity of the normalization process.

Key words: gait, view-invariance, normalization, body part trajectories

* Corresponding author. Tel.: +1 418-656-2131 ext. 4786; fax: +1 418-656-3594.
Email addresses: fjean@gel.ulaval.ca (Frédéric Jean), aalbu@ece.uvic.ca (Alexandra Branzan Albu), bergevin@gel.ulaval.ca (Robert Bergevin).

1 Introduction

Vision-based technologies for gait analysis offer a promising alternative to their conventional marker-based counterparts. Markerless gait analysis is key for emerging application fields, such as for instance biometrics. Since humans can visually recognize the gender and identity of walkers using only motion cues [1], computer vision technologies for gait-based person identification are aiming to replicate and enhance this human capability. Other applications of vision-based gait analysis can provide information about physical and psychological conditions of monitored patients. The need for such tools is confirmed by medical studies which identify specific visually observable gait changes as early signs of depression [2], dementia [3] or as fall predictors in elderly subjects [4]. Abnormal gait detection techniques for medical applications using computer vision technologies have been proposed in [5,6].

The dynamics of gait have been extensively studied in the medical community; a seminal medical study [7] has modeled human gait as a complex articulated motion composed of twenty significant components. One may note that such a detailed model is not retrievable from visual information only. However, gait modeling techniques play a central role in computer vision-based gait analysis, regardless of the area of application. Most of the recently proposed models were developed for the purpose of gait-based person identification. A brief overview of gait modeling techniques follows below.

Gait models can be classified into two main categories: shape-based and kinematic. According to [8], these two modeling approaches are complementary. Shape-based models belong to the category of state-space methods [9] which view the gait motion as completely defined by a sequence of static body poses. Shape-based models exploit therefore silhouette information. A comparative analysis of gait-relevant information captured by average silhouettes versus differential silhouettes is provided in [10]. Average silhouettes are also used in [11]. The work in [12] uses a population Hidden Markov Model (pHMM) where states represent gait stances over one gait cycle and the observations are the silhouettes of the corresponding gait stances. The second class of gait models encode information about gait dynamics. Their origins can be found in early work on cues that humans use for gait recognition [13], which suggests that motion cannot be reduced to a collection of static forms, but represents a dynamic invariant that influences event perception. Kinematic gait models are used for representing the human body as an articulated kinematic chain, which can be either fit onto the walker’s silhouette [14] or used implicitly in order to recover gait parameters such as stride and height [15], stride width and elevation at both knee and ankle levels [16] or spatiotemporal trajectories [17,18].

The proposed work is based on the implicit use of a kinematic gait model that is similar to [17] and [18]. Specifically, we extract spatiotemporal trajectories of body parts (head and feet) for modeling gait. Our method for trajectory generation improves upon previous work by solving the manual initialization issue in [17] and by extracting the spatiotemporal trajectories in real-time from video data instead of obtaining them directly from a marker-based motion tracker [18]. More importantly, our focus is not on gait recognition, but on view-invariant gait modeling using a non-calibrated camera. The view-invariant gait model can be further used for gait recognition, gait-based biometrics, or medical gait analysis purposes. Our main contribution consists therefore in a novel technique for view normalization, which makes obtained head and feet trajectories always appear as if seen from a fronto-parallel viewpoint. Related work on view-invariant gait modeling is reviewed below.

Many gait modeling approaches are either applicable only to the fronto-parallel viewpoint [16,19], or at least view-dependent [20]. Height and stride length are estimated in a view-invariant way in [21], but the method necessitates the camera to be calibrated with respect to the ground. The method proposed in [22] uses the structure of articulated body part motion to recover the parameters of the projective transformation under which a subject is observed. The projective transformation is then used to generate a canonical fronto-parallel view. That method uses markers to obtain precise positions of the ankles, knees and hip, which are difficult to retrieve automatically from computer vision algorithms. Synthesis of a canonical side view from an arbitrary view is performed in [23] via two methods, namely perspective projection and optical flow-based structure-from-motion. However, the synthesis of a side view is only feasible from a limited number of initial views. The method in [24] involves a scaling process, for each known view, on silhouette parameters such as height and distance between head and pelvis. In [25], a method for tilt correction of silhouettes is proposed, but it requires camera calibration. Estimation of a person’s 3D trajectory from a monocular calibrated camera is discussed in [26]. The 3D trajectories are used in order to recover the walking speed. Walking directions of people are computed in [27] using a camera with known focal length and the weak-perspective projection model. The walking direction is then used to recover view-invariant lengths at different parts of the silhouettes.

Other methods integrate the information from multiple views in order to recover a canonical view or to extract features that are view-invariant. This is the case of the method presented in [28] where the desired view of a moving object is reconstructed using multiple simultaneous views. The method proposed in [29] achieves view-invariance by learning gait parameters from multiple views, and people identification is performed by providing only a single view. In [30], a bilinear model is fitted on multiple views. View-invariant identification is achieved by decoupling the identity of the person and the viewpoint from

which he is observed. A view transformation model is used in [25] in order to transform already observed gait features into the same walking direction as newly observed features. The view transformation model is learned from multiple views of walking subjects.

Some studies were performed to analyze the effect of viewpoint variation on gait identification performance. In [31], a geometrical and an analytical model is developed for that purpose. It has been found that identification is robust between two views that are close to the fronto-parallel view, which seems not to be the case for other reference views. The problem of determining the best reference view has been addressed in [32]. Using a view transformation model, it has been found that the 45° view leads to better results for a single reference view. A combination of orthogonal views, such as 90° and 0° , leads to better results for two reference views.

The approach proposed in this paper generates body part trajectories that are invariant to the walking trajectory and viewpoint. We relax the assumption generally adopted in gait analysis about linear walking trajectories. This assumption does not hold when walkers are tracked in extended premises, where they need to change their walking direction in order to follow paths, or to avoid obstacles/other walkers.

Body part trajectories are assumed to contain sufficient information about the gait of a person for view-invariant modeling. The trajectory of a body part (foot, head, hand) is defined as a sequence of the successive 2D positions it takes in the frames of a video sequence. On a frame-by-frame basis, each body part is represented by one point respectively. View normalization consists in making body part trajectories appear as if seen from the same fronto-parallel viewpoint. The proposed approach to view normalization features automatic initialization, no camera calibration, as well as a low computational complexity. Since the proposed method is calibration-free, the images from the cameras are not corrected for radial distortion. The latter is assumed to be negligible compared to the distortion resulting from perspective projection. This would not be the case for a camera with a wide field of view. Few gait analysis methods currently consider wide fields of views.

This paper represents an extension of the work presented in [33]. A new method to determine foot labels (left or right) is presented. A new evaluation method is also proposed to assess the performance of the normalization algorithm for straight line walks as well as walks with changes in direction. An extensive evaluation is performed on more than 120 walk sequences.

The remainder of the paper is organized as follows. The proposed approach is detailed in Section 2. Experimental results are presented in Section 3. Finally, Section 4 draws conclusions and outlines the main directions of future work.

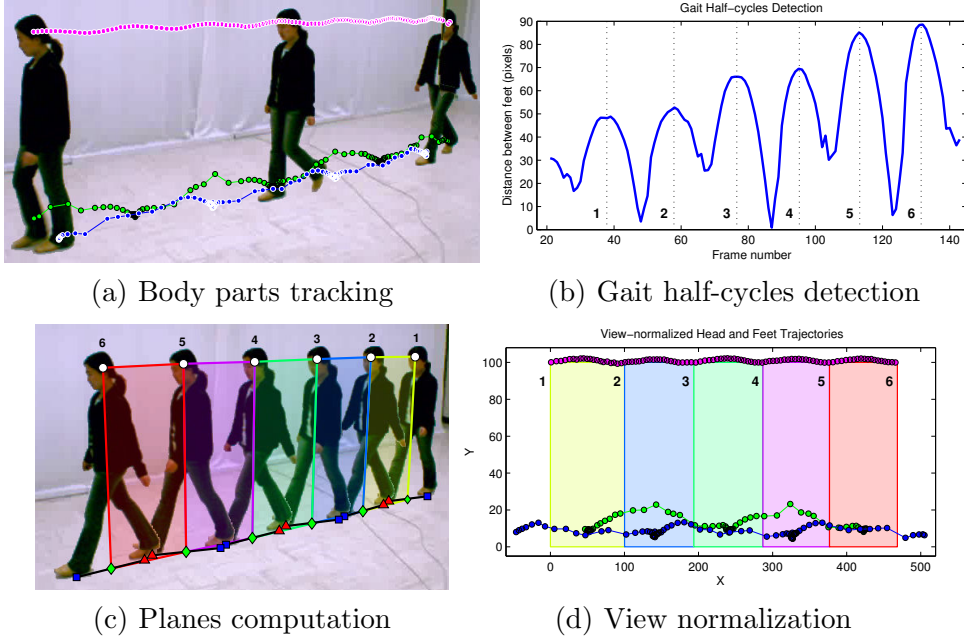


Fig. 1. Overview of the proposed method.

2 Proposed Approach

An overview of the proposed method is illustrated in Figure 1. The three raw body part trajectories in Figure 1(a) are obtained from silhouettes extracted at each frame of a monocular video sequence. In this paper, only the head and the feet trajectories are considered. Other body part trajectories could be view-normalized in the same way. The temporal variation of the inter-feet distance is analyzed in order to determine moments at which the feet are furthest apart (maximum stance), as shown in Figure 1(b). Here, there are six such moments. Time intervals between two consecutive moments define a gait half-cycle. A view normalization is to be performed in each of these half-cycles. Figure 1(c) shows the estimated walking planes in each half-cycle, with the extracted silhouettes overlaid at maximum stance moments. If a person were observed from a fronto-parallel view, those planes would appear as rectangles; thus, the “normalized” version of a walking plane from an arbitrary view has right angles, as shown in Figure 1(d). A homography between an observed plane and its normalized version can then be computed and used in order to “normalize” the body part trajectories within the corresponding gait half-cycle.

The following sections present a detailed description of each step of the proposed method.

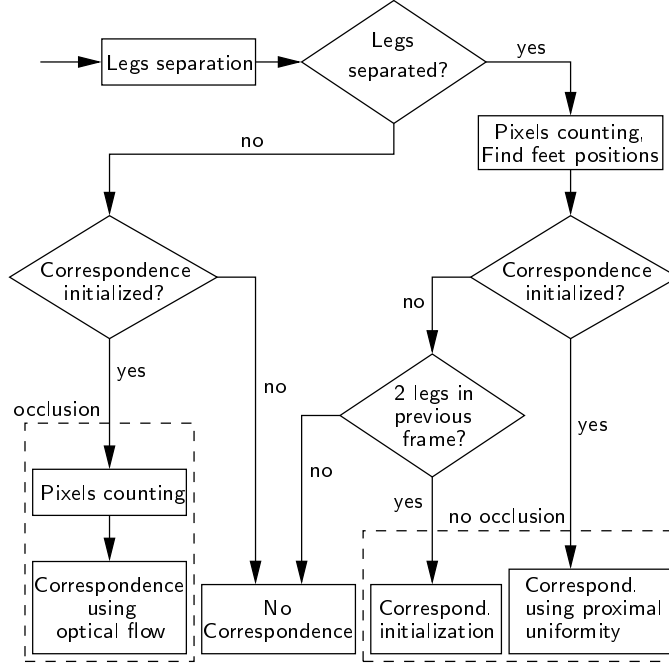


Fig. 2. Overview of the feet tracking algorithm applied at each frame.

2.1 Generation of “Raw” Body Part Trajectories

The proposed viewpoint normalization approach receives as input “raw” (i.e. view-dependent) body part trajectories. One should note that the normalization approach stays the same regardless on which (or how many) body parts are collected from the human walk. We present results for feet and head only, as we consider that the motion of these body parts encodes core, irreducible information about walking. The “raw” feet and head trajectories are generated via an algorithm customized for human walking. This algorithm is summarized below; its detailed presentation can be found in [34].

The head trajectory consists of the sequence of locations of the center of mass of the head extracted on a frame by frame basis from a walking sequence. The generation of this trajectory is straightforward, as the head is always the highest part of the human silhouette and does not suffer any occlusions during walk. This is however not the case with feet in monocular sequences. Feet occlude themselves periodically in every viewpoint except one, where the subject walks along the optical axis of the camera (frontal axis). This self-occlusion needs to be addressed in order to obtain a correct feet correspondence (i.e. left-to-left and right-to-right) across every pair of adjacent frames in the sequence. The feet correspondence problem is solved by the tracking algorithm shown in Figure 2. This algorithm handles all possible correspondence cases, as follows. Feet are first detected as regions in the image, and then each represented by one point on a frame by frame basis. In cases where legs are separable, these

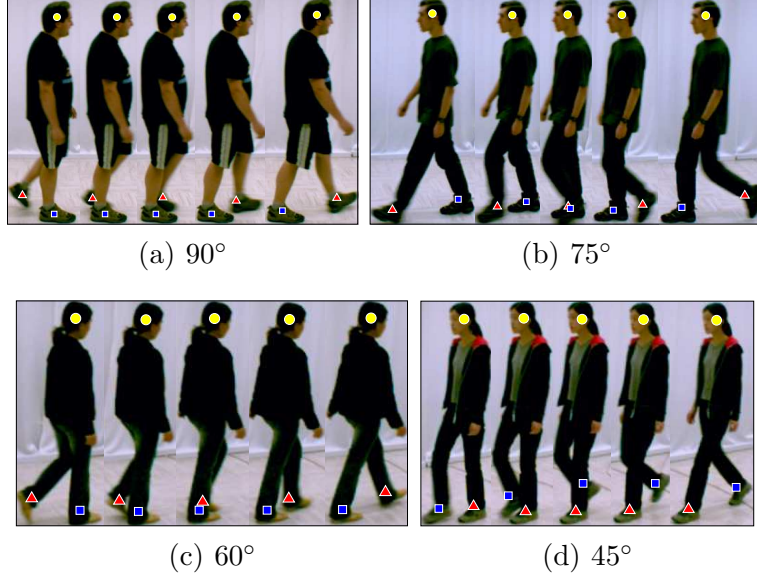


Fig. 3. Body part tracking from different views

representative points are estimates of the centers of mass of the foot regions.

First, correspondence is initialized using an adjacency criterion. This criterion states that the “right” foot in Frame i must be spatially closer to the “right” foot than to the “left” foot in Frame $(i+1)$; a similar reasoning is applied for the “left” foot correspondence. In ambiguous cases where distances are quasi-equal or where feet are in self-occlusion, the initialization is simply delayed to the next pair of adjacent frames where the ambiguity disappears. This way, the feet tracking algorithm can start at any phase of the walking cycle. One should note that “right” and “left” have arbitrary meanings here, since the algorithm does not provide a method to distinguish right from left. An algorithm for that purpose is presented in Section 2.5.

Once correspondence is initialized, the tracking algorithm propagates feet correspondence for every subsequent pair of adjacent frames. In doing so, it needs to evaluate whether self-occlusion is present or not. In the absence of self-occlusion, the motion correspondence algorithm of Rangarajan and Shah [35] is applied. The choice of this algorithm is based on the fact that it is suitable for smooth and relatively small frame-to-frame feature displacements, as it is the case with the motion of the legs’ representative points during walking.

In case self-occlusion is present, the legs’ region is not separable into two regions (one per leg). In this case, the legs’ representative points are not retrievable as centers of mass and thus need to be estimated using motion information. One may note that, in human gait, feet self-occlusions have the interesting particularity that there is only one visible foot moving, while the other is grounded as support foot. Therefore, we retrieve the representative point of the moving foot using optical flow, while the point representing the stationary

foot is the same as in the previous frame. It is important to note that the foot correspondence is implicit in the occlusion case, since the stationary and the moving foot are tracked differently. Temporary no correspondence cases are present when self-occlusion lasts more than a certain number of frames; linear extrapolation might be used during occlusion if there is not enough motion detected from optical flow.

Figure 3 shows snapshots of head and feet positions obtained for four pedestrians observed from different viewpoints. Squares and triangles represent feet positions while disks represent head positions. One can notice that correspondence is properly achieved by looking at symbols on the feet.

2.2 Segmentation of gait half-cycles

The normal gait is a periodic and symmetric motion. Symmetry arises from the fact that feet play interchanging roles during one cycle of gait: the swing of the left foot is supported by a stationary right foot and vice versa. Symmetry is important because it enables us to decompose the gait cycle into half-cycles. A gait half-cycle is defined as the interval between two consecutive stance postures (double support, where both feet are on the floor, and the inter-feet distance is maximum). The frame numbers associated with maximum inter-feet distance are detected from a maxima search on the function $d(t)$ representing the distance between the feet mass centers over time (see blue curve in Figure 1(b)). $d(t)$ is defined as follows:

$$d(t) = G(t) * \|\mathbf{p}_{f1}(t) - \mathbf{p}_{f2}(t)\|, \quad (1)$$

where: t is the frame number; $\|\cdot\|$ is the Euclidean norm; the operator $*$ denotes the convolution; $G(t)$ is a gaussian kernel used for smoothing, with $\sigma = 1$ and a kernel width of 5; $\mathbf{p}_{f1}(t)$, $\mathbf{p}_{f2}(t)$ are the positions occupied by the mass centers of feet 1 and 2 at frame t .

The maxima search is performed iteratively within a sliding window of predefined length W . This length represents the minimum duration (in frames) of a gait half-cycle, given the frame rate of the video acquisition process f_s (in frames per second), and the maximum gait cadence f_g (in Hz). Thus, the length W is mathematically defined as $W = \lfloor f_s/2f_g \rfloor + 1$ if $\lfloor f_s/2f_g \rfloor$ is even, or $W = \lfloor f_s/2f_g \rfloor$ otherwise. In this paper, $f_s = 30$ fps and f_g is set to 2 Hz, which allows for detecting the maximum distance moments for typical walking velocities.

The window W slides one frame forward per iteration. A maximum is detected at time t^* if the window W is centered on the maximum. This condition is

expressed by the following equation:

$$t^* = \arg \max_t \{d(t)\}; t \in \left[t^* - \lfloor \frac{W}{2} \rfloor, t^* + \lfloor \frac{W}{2} \rfloor \right], \quad (2)$$

where W denotes the size of a window centered at t^* .

The output of the maxima search is a series of temporal moments (expressed in frame numbers) t_i^* , $i = 1 \dots N$ who all correspond to maximum inter-feet distance.

This series is further refined to moments with sub-frame accuracy. Sub-frame accuracy is needed for reducing the variance in the half-cycles duration, which is more significant for low acquisition frame rates. The refinement process uses parabolic interpolation as follows:

$$\gamma_i = t_i^* + \frac{d(t_i^* - 1) - d(t_i^* + 1)}{2 \{d(t_i^* + 1) - 2d(t_i^*) + d(t_i^* - 1)\}}. \quad (3)$$

The refined moments γ_i are used to delineate the gait half-cycles as time intervals $\mathcal{C}_i = [\gamma_i, \gamma_{i+1}]$, for $i = 1 \dots N - 1$. Figure 1(b) shows the detected moments as vertical dotted lines.

2.3 Estimation of Walking Planes

The work described in this paper addresses viewpoint variability in gait modeling. A fixed camera can observe subjects walking along a range of linear paths, as well as subjects changing direction during their walk. We represent information about the gait direction by using the concept of walking planes. To define such planes, we relax the assumption of walking along linear directions. Instead, we simply assume that the walking direction is invariant during each half-cycle of walk. This assumption is consistent with the fact that humans typically change direction by pivoting on their feet of support during stance phase, rather than by incorporating a rotation during swing phase.

A walking plane can be most accurately estimated for each half-cycle using the raw body part trajectories extracted with the algorithm in 2.1. One may note that feet and head are not exactly located in the same plane during walk, but on parallel planes separated by distances that are considered negligible with respect to the distance from the camera. Hence, determining the walking plane using all locations of body parts during the considered half-cycle needs a least-squares approximation which would significantly raise the computational

complexity of the technique. Therefore, we propose a simplified approach for estimating the walking plane.

Assuming a horizontal floor plane, the walking plane is orthogonal to the floor plane. We can therefore define the walking plane by using two orthogonal projections of the center of mass of the head to the floor plane. These projections are calculated at the beginning and end of the walking cycle respectively. Since this work uses a non-calibrated camera, the floor-plane location is not a priori known. Therefore, the projections of the head onto the plane floor are approximated by the middle points between the feet positions at the beginning and end frames of the half-cycle, which both contain feet in stance position. The middle points are defined by $\mathbf{p}_m(t) = \{\mathbf{p}_{c1}(t) + \mathbf{p}_{c2}(t)\} / 2$ and are shown in Figure 1(c). A walking plane Π_i is then defined as the set of four points corresponding to the head and feet projection locations at the beginning γ_i and the end γ_{i+1} times of the considered half-cycle:

$$\Pi_i : \{\mathbf{p}_h(\gamma_i), \mathbf{p}_m(\gamma_i), \mathbf{p}_h(\gamma_{i+1}), \mathbf{p}_m(\gamma_{i+1})\}. \quad (4)$$

Since γ_i and γ_{i+1} are non-integer values, the locations of head and head projections at these times are computed by linear interpolation as follows:

$$\mathbf{p}(\gamma) = \mathbf{p}(\lfloor \gamma \rfloor) + (\gamma - \lfloor \gamma \rfloor) \{\mathbf{p}(\lceil \gamma \rceil) - \mathbf{p}(\lfloor \gamma \rfloor)\}. \quad (5)$$

One may note that the proposed estimation of walking planes ensures the independence of results obtained for each half-cycle and does not allow errors to propagate from one half-cycle to the next.

2.4 View Normalization of Body Part Trajectories

The view normalization process aims to generate body part trajectories similar to those that would be acquired from a fronto-parallel viewpoint (side view). This process is to project all view-dependent trajectory information onto one reference plane, which is orthogonal to the camera's optical axis.

To obtain the normalized body part trajectories, a correspondence between walking planes and the reference plane needs to be established first. This correspondence can be expressed in terms of a homography.

Normalized walking planes are rectangles having the same height and the width proportional to the length of the walking cycle that they represent as shown in Figure 1(d). The height of the rectangles corresponds to the height of

the tracked subject and is set to an arbitrary value, since the video acquisition is performed with an uncalibrated camera.

The normalized walking plane corresponding to the half-cycle occurring in the time interval $\mathcal{C}_i = [\gamma_i, \gamma_{i+1}]$ is defined by the following set of points:

$$\tilde{\Pi}_i : \left\{ \begin{bmatrix} \alpha_{i-1} \\ \beta \end{bmatrix}, \begin{bmatrix} \alpha_{i-1} \\ 0 \end{bmatrix}, \begin{bmatrix} \alpha_i \\ \beta \end{bmatrix}, \begin{bmatrix} \alpha_i \\ 0 \end{bmatrix} \right\}, \quad (6)$$

where β is an arbitrary height (here $\beta = 100$ pixels). The horizontal positions at the beginning and the end of the corresponding half-cycle are defined as

$$\alpha_i = \tau \sum_{j=1}^i (\gamma_{j+1} - \gamma_j), \quad (7)$$

with initial value $\alpha_0 = 0$. This initialization results in spatially connected consecutive normalized planes, which in turn ensures the continuity of normalized trajectories across half-cycles.

The width of the normalized walking planes is a temporal quantity; thus, it is set to be proportional to the duration of their corresponding gait half-cycle. We assume in this paper that the velocity of the performed walk is constant, therefore, all normalized planes will have a quasi-equal width which is determined by the parameter τ , expressed in units per frame and set to an arbitrary value (here $\tau = 5$).

The homography \mathbf{H}_i that maps a walking plane Π_i to its normalized version $\tilde{\Pi}_i$ is then computed by constructing an 8-equation linear system using the correspondences between the corners of the original and normalized walking planes. Homography matrices are computed using the *normalized DLT* (Direct Linear Transform) algorithm for 2D homographies described in [36].

This homography is used on the obtained body part trajectories $\mathbf{p}_b(t)$, with $b \in \{h, f1, f2\}$. A normalized version $\tilde{\mathbf{p}}_b(t)$ within half-cycle \mathcal{C}_i is then obtained as follows:

$$\lambda \begin{bmatrix} \tilde{\mathbf{p}}_b(t) \\ 1 \end{bmatrix} = \mathbf{H}_i \begin{bmatrix} \mathbf{p}_b(t) \\ 1 \end{bmatrix}; \quad t \in [\lceil \gamma_i \rceil, \lfloor \gamma_{i+1} \rfloor], \quad (8)$$

where λ is a scale factor, and $i = 1 \dots N - 1$. One can see that each body part trajectory is view-normalized independently within each gait half-cycle using the corresponding homography.

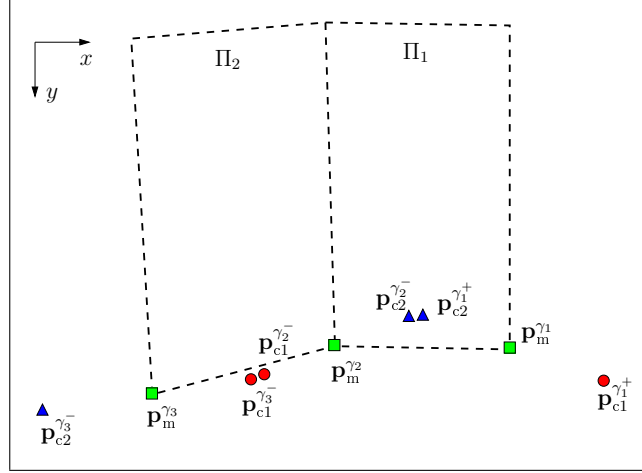


Fig. 4. Illustration of points used for foot labeling. Here, points for two gait half-cycles are shown. The points \mathbf{p}_m (squares) represent the planes Π_1 and Π_2 bottom corners which are defined at maximum stance moments γ_1 , γ_2 , and γ_3 . The points \mathbf{p}_{c1} (circles) and \mathbf{p}_{c2} (triangles) represent the feet contact points at integer moments defined as $\gamma^- = \lfloor \gamma \rfloor$ and $\gamma^+ = \lceil \gamma \rceil$.

2.5 Automatic Foot Labeling

Foot labeling is required in the experimental evaluation of the proposed normalization algorithm. Automatic foot labeling makes it possible to perform validating experiments on a large number of trajectories. In Sections 3.1 and 3.2, feet trajectories obtained from different views are compared. The comparison process should then compare a person’s right foot trajectory obtained from one view with the person’s right foot trajectory obtained from another view, and the same for the left foot. The tracking algorithm presented in Section 2.1 provides feet trajectories labeled as “foot1” and “foot2”; however, “foot1” in one sequence might not correspond to the “foot1” in another sequence. The labeling algorithm described in this section ensures that the “foot1” trajectory in each sequence corresponds always to the right foot.

The foot labeling algorithm relies on the observation that surveillance cameras are usually positioned near the ceiling and have a non-zero tilt, that is, are looking down to the floor. In a fronto-parallel view, this would make one of the foot contact points at the maximum stance moments (double support) appear a little bit higher in the image than the other one, because that foot is further than the other from the camera. Considering the direction of the walk in the image, one could deduce that the right foot is the one with the higher position if the walk is from left to right, or the lower one if the walk is from right to left. Computing feet positions at each maximum stance moment would lead to votes for “foot1” or “foot2” being the “right” foot. This whole process can also be applied for non fronto-parallel views. This is shown in Figure 1(c), where all contact points belonging to the right foot are always

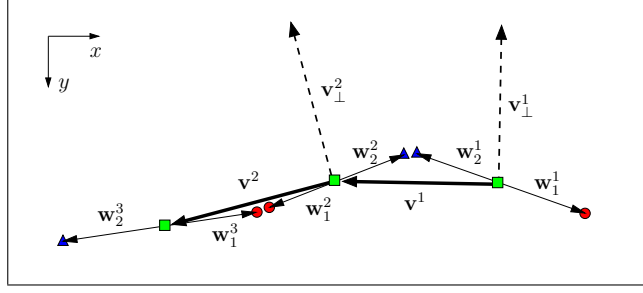


Fig. 5. Illustration of the vectors computed from point shown in Figure 4. Vectors \mathbf{v} represents the planes bottom lines, with the arrows showing the direction of the walk. Vectors \mathbf{v}_\perp are the corresponding perpendicular vector obtained from a clockwise rotation. Vectors \mathbf{w} are defined with respect to the feet contact points and the planes corner points. If the dot product $\mathbf{w} \cdot \mathbf{v}_\perp$ is greater than 0, then the foot represented by \mathbf{w} is the right foot, with respect to \mathbf{v} .

on the same side of the bottom line of the walking plane (above this line). A similar reasoning holds for the left foot. It is to be noted that no adverse effect on the normalization results have been observed due to the tilt of the camera.

Figure 4 illustrates the feet labeling problem for two gait half-cycles. The notation \mathbf{p}^t is used to represent a position at time t . The planes' bottom corners \mathbf{p}_m will be used to compute the apparent walking direction in the image. Three such points are shown in the image, namely one for each maximum stance moment γ_i , $i = 1 \dots N$ (here $N = 3$). For the first maximum stance moment ($i = 1$), the considered feet contact positions \mathbf{p}_{c1} and \mathbf{p}_{c2} are the ones at integer moment $\gamma_1^+ = \lceil \gamma_1 \rceil$. For $i = 2 \dots N$, feet contact positions at moments $\gamma_i^- = \lfloor \gamma_i \rfloor$ are considered. These integer moments were chosen in order to use positions that are within the intervals defined between γ_i values.

Using these points, it is possible to compute a vector representing the direction of the walk, and the position of the feet contact with respect to the walk direction. Those vectors are illustrated in Figure 5. The vector \mathbf{v}^i , representing the direction of the walk for the gait half-cycle \mathcal{C}_i , is computed as

$$\mathbf{v}^i = \mathbf{p}_m^{\gamma_{i+1}} - \mathbf{p}_m^{\gamma_i}, \quad (9)$$

for $i = 1 \dots N-1$. The perpendicular vectors \mathbf{v}_\perp^i are computed in the image by performing a 90 degrees clockwise rotation on the vector \mathbf{v}^i . One should note that the vector \mathbf{v}_\perp^i always indicates the side of the vector \mathbf{v}^i where the “right” foot position on the floor should be. The vector representing the position of the foot with respect to the walk direction, \mathbf{p}_{c1} and \mathbf{p}_{c2} , is computed as follows:

$$\mathbf{w}^i = \begin{cases} \mathbf{p}_c^{\gamma_i^+} - \mathbf{p}_m^{\gamma_i} & \text{if } i = 1 \\ \mathbf{p}_c^{\gamma_i^-} - \mathbf{p}_m^{\gamma_i} & \text{otherwise} \end{cases} \quad (10)$$

Algorithm 1 Voting process

```
1:  $V \leftarrow 0$ 
2:  $A_1 \leftarrow 0, A_2 \leftarrow 0$ 
3: for  $i = 1 \dots N$  do
4:   if  $i = 1$  then {special case}
5:     if  $\mathbf{w}_f^1 \cdot \mathbf{v}_\perp^1 > 0$  and  $\mathbf{w}_{\bar{f}}^1 \cdot \mathbf{v}_\perp^1 < 0$  then
6:        $A_f \leftarrow A_f + 1, V \leftarrow V + 1$ 
7:     end if
8:   else if  $i = N$  then {special case}
9:     if  $\mathbf{w}_f^N \cdot \mathbf{v}_\perp^{N-1} > 0$  and  $\mathbf{w}_{\bar{f}}^N \cdot \mathbf{v}_\perp^{N-1} < 0$  then
10:       $A_f \leftarrow A_f + 1, V \leftarrow V + 1$ 
11:    end if
12:   else {general case}
13:     if  $\mathbf{w}_f^i \cdot \mathbf{v}_\perp^{i-1} > 0$  and  $\mathbf{w}_{\bar{f}}^i \cdot \mathbf{v}_\perp^{i-1} < 0$  then
14:        $A_f \leftarrow A_f + 1, V \leftarrow V + 1$ 
15:     end if
16:     if  $\mathbf{w}_f^i \cdot \mathbf{v}_\perp^i > 0$  and  $\mathbf{w}_{\bar{f}}^i \cdot \mathbf{v}_\perp^i < 0$  then
17:        $A_f \leftarrow A_f + 1, V \leftarrow V + 1$ 
18:     end if
19:   end if
20: end for
```

for $i = 1 \dots N$. A foot contact position \mathbf{p}_c can be associated with the “right” foot if the corresponding vector \mathbf{w} shows that the point is at the right of a vector \mathbf{v} , which is indicated by the vector \mathbf{v}_\perp . That is, the foot contact position is associated with the right foot if the dot product between \mathbf{w} and \mathbf{v}_\perp is greater than 0.

Foot contact positions at a given time provide one vote to determine which one of the foot trajectories is the “right” foot trajectory. The votes are accumulated in accumulators A_f , $f \in \{1, 2\}$, for the “foot1” and the “foot2” trajectories respectively. The voting process is performed as stated in Algorithm 1. The notation \bar{f} is used to denote the “other” foot, that is, if $f = 1$, then $\bar{f} = 2$, and vice versa. This voting process consists in three parts, where the first two deal with the special cases, that is, the foot contact position at the beginning of the first gait half-cycle and at the end of the last gait half-cycle. For those two cases, the dot product can only be computed with one walk direction. In the general case, the dot product is computed with two walk directions. One may notice that the algorithm takes care of the contradictions that may arise during the voting process. Indeed, foot contact positions at a given time must lead to a foot being the “right” foot and the other one as the “left” foot. Otherwise, the vote is not considered. The number of votes that are considered is denoted by V .

The number of votes accumulated in A_1 and A_2 are used to compute a probability of being the “right” foot for each foot trajectory “foot1” and “foot2”.

Those probabilities are computed as $P_1 = A_1/V$ and $P_2 = A_2/V$, with $P_1 + P_2 = 1$. If $P_2 > P_1$, then the foot trajectories are swapped such as the “right” foot trajectory is the “foot1” trajectory: $\mathbf{p}_{f1} \leftrightarrow \mathbf{p}_{f2}$. Otherwise, the “right” foot trajectory is already the “foot1” trajectory, or $P_1 = P_2$, in which case no decision can be made.

The foot labeling process can be performed right after the walking planes have been determined. It does not require the foot trajectories to be normalized. As it was said previously, this algorithm is used during the feet trajectories comparison. However, it could also be used by a gait modelization algorithm that would need the knowledge of a “right” and a “left” foot.

3 Experimental Evaluation

Two experiments have been performed in order to test the proposed algorithms. The first experiment implies subjects walking on a straight line path, while the second experiment implies subjects walking on paths with changes in the walk direction. The methodology and the obtained experimental results are detailed for both experiments in the following sections.

3.1 *Experiment 1: Evaluating the similarity of normalized and fronto-parallel trajectories for single direction walks*

The purpose of the normalization process is to obtain body part trajectories that appear to have been obtained from a fronto-parallel viewpoint. Thus, the effectiveness of the normalization algorithm can be assessed by comparing the body part trajectories obtained from different views of the same single direction walk to the reference view (fronto-parallel view). By performing the comparison on both non-normalized (raw) and normalized body part trajectories, one could appreciate the improvements that can be achieved by normalizing the trajectories with the proposed method. In the ideal case, normalized trajectories obtained from a non-fronto-parallel view would be identical to the ones obtained from the fronto-parallel view. However, this will not be the case in a real system because of noise induced by the tracking process.

3.1.1 *Evaluation Methodology*

One convenient way of comparing two 2D trajectories is to align (or fit) one trajectory on the other one, in a least squares sense, in order to compute a deviation measure. This distance measure would indicate how similar these

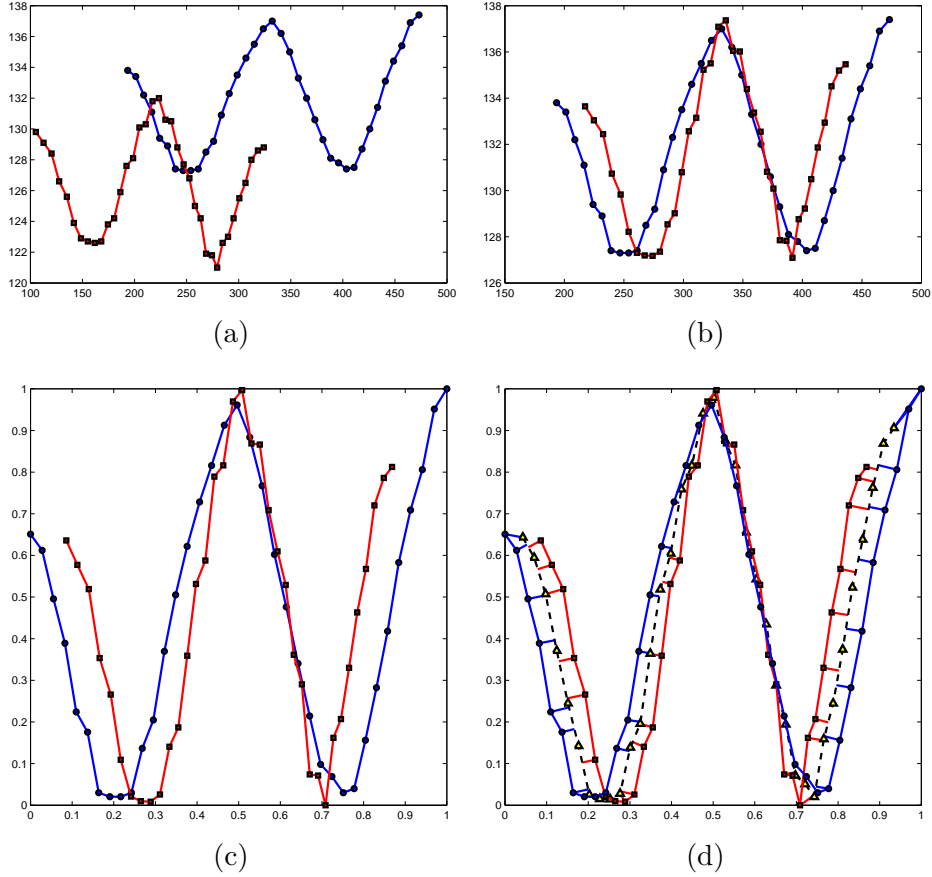


Fig. 6. Example for the trajectory alignment process. In a) The original head trajectories from a 90° view (squares) and a 45° view, as observed in the camera. In b), optimal rotation and translation of the two trajectories. In c), the trajectories are scaled in a 1×1 box. In d), an average trajectory is computed (triangles), and distance to this average trajectory is computed for all points of the two trajectories.

trajectories are. For instance, a distance of zero would mean that the trajectories are identical.

An example of the trajectory alignment process is detailed in Figure 6. This example shows the alignment process for the head trajectory obtained simultaneously from a 90° view (circles) and a 45° view (squares). Both trajectories have the same number of points T . In order to perform the alignment, the cameras must be synchronized, so each corresponding point between the trajectories represents the head position observed at the same moment. The original trajectories, $\mathbf{h}_{90}(t)$ and $\mathbf{h}_{45}(t)$, are shown in Figure 6(a). Figure 6(b) shows the first step of the alignment process, which consists in optimally aligning the trajectories using the method in [37] (optimal 2D rigid transformation). Because of the perspective projection and the noise in the tracking process, the corresponding points do not fit perfectly. In order to compute a “deviation” measure, the aligned trajectories are first scaled down to fit in a unit box, as shown in Figure 6(c). These new trajectories are denoted $\tilde{\mathbf{h}}_{90}(t)$ and

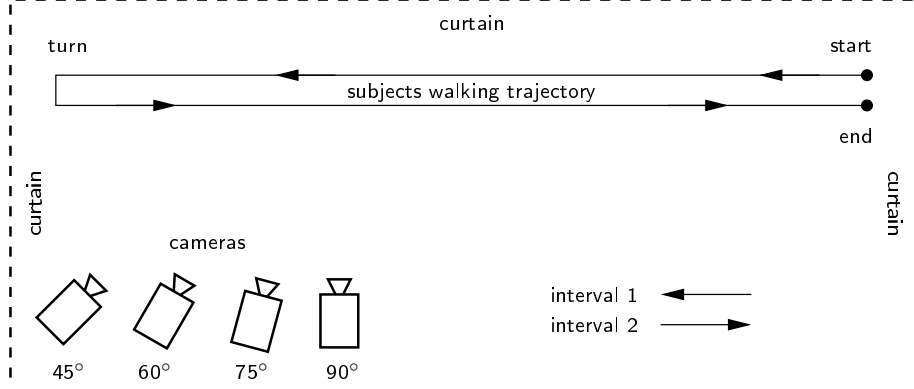


Fig. 7. Acquisition setup for the first experiment.

$\tilde{\mathbf{h}}_{45}(t)$. Next, an average trajectory $\mathbf{h}_{\text{avg}}(t)$ is computed by taking the middle point between each pair of corresponding points of the two trajectories: $\mathbf{h}_{\text{avg}}(t) = (\tilde{\mathbf{h}}_{90}(t) + \tilde{\mathbf{h}}_{45}(t))/2$. The average trajectory is shown as the triangles curve in Figure 6(d). Considering the average trajectory as a polyline curve, it is possible to find for each point of $\tilde{\mathbf{h}}_{90}(t)$ and $\tilde{\mathbf{h}}_{45}(t)$ its closest point on the polyline curve. This is presented as small segments in Figure 6(d), whose lengths are denoted $e_{90}(t)$ and $e_{45}(t)$. The length of these segments can be thought of as a deviation between the trajectories in a given point. It is then possible to compute the Root Mean Squares Deviation (RMSD) R , which provides a global deviation measure between the trajectories:

$$R = \sqrt{\frac{1}{2T} \sum_{t=1}^T [e_{90}^2(t) + e_{45}^2(t)]} \quad (11)$$

This measure is appropriate for the trajectory comparison process since it measures a deviation in the shape of the trajectory, which is mainly caused by the perspective projection.

3.1.2 Results

In the first experiment, ten volunteers are asked to walk back and forth on a straight line in front of four roughly time-synchronized color cameras (30 frames per second, resolution of 640×480 , synchronization of ± 3 frames, 6mm lenses, indoor environment). Cameras are positioned such that their optical axis intercepts the walking direction at different angles : 90° , 75° , 60° , and 45° . The minimum distance between the subjects and the cameras is about 2.5m, and the maximal distance is about 8.5m. Figure 7 shows the setup used for the acquisition process. This setup is appropriate to test the performances of the normalization algorithm since it provides four views of the same walk, including the fronto-parallel view (90°) to which other views

are to be compared. A subject first appears in right side of the image and then disappears at left side (interval 1). He reappears one or two seconds later in left side of the image and then disappears at right side (interval 2). This provides four video sequences (four views) for each subject. Preprocessing of the video sequences yields 80 head and feet trajectories ($10 \text{ subjects} \times 4 \text{ views} \times 2 \text{ intervals}$). Depending on the view and the subject, each sequence interval contains from one to three visible gait cycles.

In order to properly evaluate the normalization process, time-synchronized cameras are needed in this experiment. Imperfect time-synchronization between cameras was refined interactively by selecting key frames in each view. One should note that time synchronization is necessary for evaluation purposes only. Such time-synchronization is not needed when gait models are built and compared from trajectories extracted from different cameras at different times. In this experiment, the trajectories alignment between two synchronized views is performed for the time interval they share. That is, only the parts of the trajectories that represent body part positions that occurred at the same time in both views is considered.

Table 1 presents results of the body part trajectory alignments for three combinations of views : $90^\circ - 75^\circ$, $90^\circ - 60^\circ$, and $90^\circ - 45^\circ$. Alignment was performed for non-normalized (raw) and normalized (norm) body part trajectories for each view. RMSD values of the resulting alignment are presented for each subject, interval and body part. Table 2 presents alignment statistics for all subjects and intervals mixed. It is possible to see from these results that trajectories obtained from 75° , 60° , and 45° views are closer to the trajectories obtained from the fronto-parallel view (90°) after normalization. The importance of the trajectory normalization process becomes clearer as the difference in angle between the compared views is greater. In Figures 8 and 9, aligned trajectories of two subjects are shown. Trajectory alignment is more difficult when there is noise in the body part trajectories, as one may see in Figure 9 for the raw and normalized feet trajectories. This noise comes from the tracking algorithm whose performance may vary according to the viewpoint. The person’s silhouette may indeed be noisier in some views. Moreover, feet occlusions become more difficult to handle as the view is departing from the fronto-parallel view. These factors explain why the observed RMSD values increase as the angle between compared views is getting higher. The increase of RMSD values for the normalized trajectories is however much smaller than the increase observed for raw trajectories.

Table 3 shows the results of a paired T-Test between RMSD values of the raw and normalized trajectory alignment from Table 1. The null hypothesis states that the distance value of the raw trajectories alignment and the normalized trajectory alignment do not differ. This hypothesis is verified against a statistical significance level of 0.01. One may notice that for the closest view

Subject	Interval	Type	90deg - 75deg			90deg - 60deg			90deg - 45deg		
			Head	Foot1	Foot2	Head	Foot1	Foot2	Head	Foot1	Foot2
1	1	raw	0.96	1.10	2.76	2.09	3.11	3.30	3.20	5.81	6.62
		norm	0.95	1.14	2.52	1.03	1.94	2.91	1.06	4.26	6.00
	2	raw	0.98	2.60	1.49	2.01	3.36	2.33	4.34	6.48	4.28
2	1	norm	0.77	2.19	1.36	0.78	2.97	2.71	1.23	5.07	1.78
		raw	1.00	1.03	1.32	2.23	2.61	3.81	3.31	4.12	5.06
	2	norm	0.80	0.71	1.49	0.98	2.01	2.59	1.67	3.28	2.77
3	1	raw	1.35	1.49	2.31	2.83	2.57	3.58	3.76	4.23	4.50
		norm	0.63	1.43	1.96	1.23	1.91	2.62	1.17	2.63	2.73
	2	raw	1.36	1.95	2.28	2.32	2.13	3.28	4.54	5.12	4.81
4	1	norm	1.23	2.12	2.14	0.77	1.60	1.70	1.67	1.76	3.25
		raw	1.77	2.03	1.86	2.28	2.03	2.95	3.34	3.57	4.06
	2	norm	1.38	2.22	1.72	1.03	1.85	2.14	1.35	2.66	2.34
5	1	raw	0.90	1.34	1.78	1.81	3.02	3.33	3.69	8.25	7.50
		norm	0.58	1.27	1.53	1.02	2.83	2.17	1.51	7.93	6.02
	2	raw	1.29	1.92	3.61	2.25	2.85	2.88	3.39	3.41	3.86
6	1	norm	0.82	1.62	3.56	0.63	2.85	1.74	0.95	3.29	1.82
		raw	1.07	2.18	1.76	2.10	3.09	2.79	4.16	6.05	5.95
	2	norm	0.83	2.25	1.89	0.86	2.88	2.02	1.74	4.08	5.10
7	1	raw	1.37	2.72	3.93	2.18	5.38	5.05	2.91	4.28	4.13
		norm	1.14	2.89	3.95	1.35	5.02	4.94	1.19	2.73	2.97
	2	raw	1.36	2.86	2.84	2.67	2.89	4.78	4.00	4.32	6.94
8	1	norm	1.12	3.00	2.91	0.85	2.64	4.21	1.27	3.68	5.57
		raw	1.11	2.18	2.44	2.13	3.24	3.21	3.90	4.56	5.84
	2	norm	0.68	2.67	2.42	0.95	2.96	3.41	1.28	3.84	3.89
9	1	raw	1.38	3.31	1.71	2.57	3.30	3.36	3.55	3.96	5.21
		norm	0.95	3.01	1.19	0.95	2.74	1.83	1.14	2.89	1.93
	2	raw	1.07	1.61	2.85	2.58	2.59	3.78	3.65	3.76	5.09
10	1	norm	0.83	1.79	2.97	1.38	1.73	2.61	1.55	2.73	2.57
		raw	1.19	1.65	2.28	2.50	2.48	3.72	3.64	4.13	5.80
	2	norm	0.73	1.60	2.22	1.18	2.21	2.86	1.13	3.58	4.84
11	1	raw	0.70	2.05	1.13	2.38	3.67	3.01	3.40	4.54	4.05
		norm	0.61	2.46	1.12	0.99	3.72	1.71	1.32	4.07	1.89
	2	raw	0.93	1.40	1.72	2.16	2.44	3.15	3.09	3.37	4.71
12	1	norm	0.67	1.68	1.60	1.14	2.47	1.68	1.28	3.05	2.25
		raw	1.02	1.77	1.78	2.00	2.82	2.60	3.12	3.35	4.10
	2	norm	0.75	1.78	1.51	0.63	2.78	1.89	1.01	3.01	2.71
13	1	raw	1.06	1.86	1.70	2.08	3.21	3.01	3.24	4.88	3.57
		norm	0.96	1.97	1.75	1.42	2.23	2.40	1.33	3.10	3.13
	2	raw	1.52	1.72	1.47	2.85	3.31	2.73	3.73	4.21	3.95
14	1	norm	0.98	1.78	1.17	1.31	2.49	1.60	1.07	2.81	2.18

Table 1
RMSD values ($\times 10^2$) obtained for the three combination of views.

alignment ($90^\circ - 75^\circ$), the RMSD values are not statistically significant for the feet trajectories. However, the RMSD values for “foot 2” trajectory are statistically significant at the 0.05 significance level. This is due to the fact that the views are indeed similar, and that the feet trajectories are noisier than the head trajectory because of the occlusion. The RMSD values for the other view alignments are statistically significant since the normalization makes the trajectories appear as a fronto-parallel view whereas the raw trajectories are distorted by perspective projection.

For this experiment, foot labeling was needed in the alignment process in order to compare a foot trajectory in one view with the corresponding foot trajectory from another view. The algorithm described in Section 2.5 was used for that

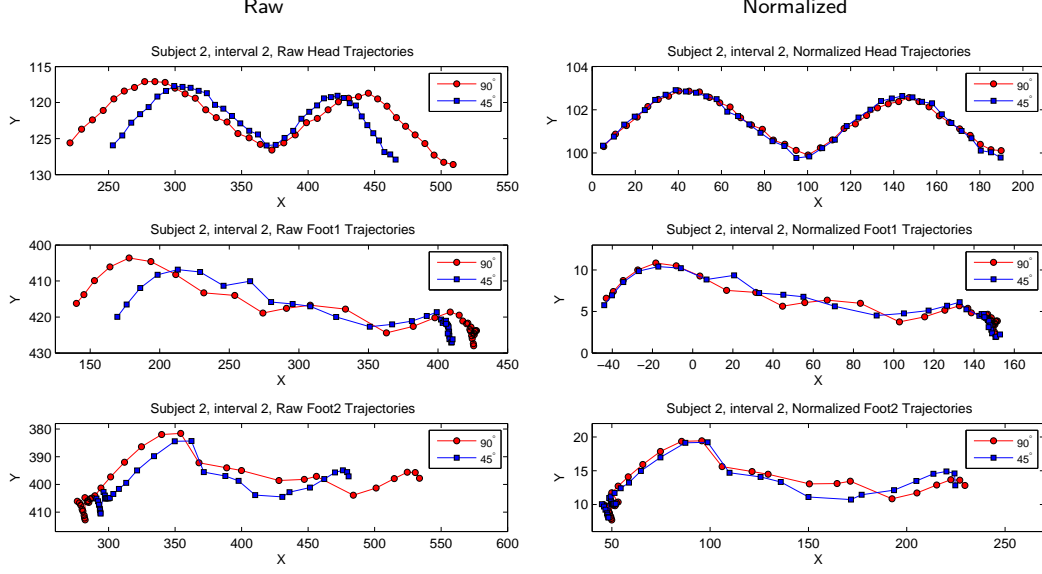


Fig. 8. Raw and normalized trajectory alignments of Subjects 2, $90^\circ - 45^\circ$

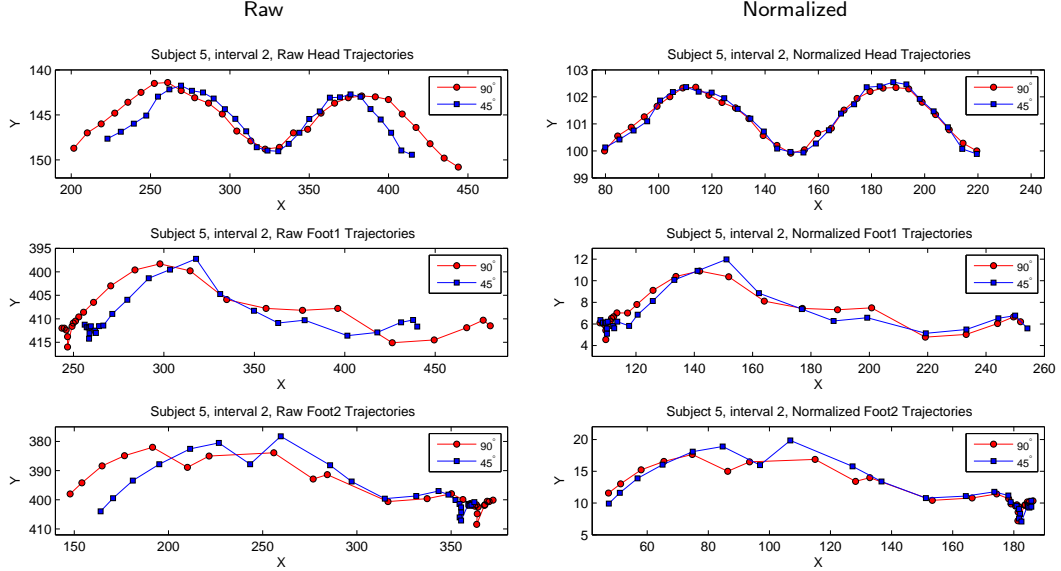


Fig. 9. Raw and normalized trajectory alignments of Subjects 5, $90^\circ - 60^\circ$

purpose. This algorithm always labels the right foot as “foot1”, and the left foot as “foot2”. The labeling results of the algorithm were compared with labeling obtained manually by looking at each color video sequence (ground-truth). The feet were correctly labeled for all 80 continuous sequences (10 subjects, 4 views, 2 intervals). Statistics on the labeling process are shown in Table 4. One may see that according to those statistics, the foot trajectory being chosen as the “right” foot was very likely correct most of the time. Moreover, most of the potential votes are used in determining the “right” foot trajectory (very few are discarded). Those results show that the foot labeling algorithm works properly and could be useful for a gait modeling algorithm

Statistic	Type	90deg - 75deg			90deg - 60deg			90deg - 45deg		
		Head	Foot1	Foot2	Head	Foot1	Foot2	Head	Foot1	Foot2
mean	raw	1.17	1.94	2.15	2.30	3.00	3.33	3.60	4.62	5.00
	norm	0.87	1.98	2.05	1.02	2.59	2.49	1.30	3.52	3.29
median	raw	1.09	1.89	1.82	2.24	2.96	3.25	3.59	4.26	4.76
	norm	0.82	1.88	1.82	1.00	2.57	2.28	1.27	3.19	2.75
std	raw	0.25	0.59	0.74	0.28	0.70	0.67	0.43	1.22	1.12
	norm	0.22	0.63	0.80	0.23	0.78	0.88	0.23	1.27	1.43
min	raw	0.70	1.03	1.13	1.81	2.03	2.33	2.91	3.35	3.57
	norm	0.58	0.71	1.12	0.63	1.60	1.60	0.95	1.76	1.78
max	raw	1.77	3.31	3.93	2.85	5.38	5.05	4.54	8.25	7.50
	norm	1.38	3.01	3.95	1.42	5.02	4.94	1.74	7.93	6.02

Table 2

Statistics on RRMSD ($\times 10^2$) from Table 1. Statistics are computed over all subjects and intervals.

T-Test	90°- 75°			90°- 60°			90°- 45°		
	Head	Foot1	Foot2	Head	Foot1	Foot2	Head	Foot1	Foot2
<i>P</i> -Value	3.0e-7	4.7e-1	2.2e-2	9.1e-14	4.1e-5	1.3e-6	2.8e-16	3.7e-6	1.6e-9
Reject <i>H</i> (0.01)?	yes	no	no	yes	yes	yes	yes	yes	yes

Table 3

Statistical significance for the comparison of normalized body part trajectories. A paired T-Test is performed on the values obtained in Table 1. Each raw value is paired with its corresponding normalized value, and the *p*-value is obtained for each body part and each view comparison. The null hypothesis *H* states that the raw and normalized trajectories comparison does not differ. Rejection of the null hypothesis is performed at the statistical significance level of 0.01.

	Mean	STD	Median	Min	Max
Probability (%)	98.63	4.97	100.00	75.00	100.00
# Considered votes	7.34	2.11	8.00	4.00	12.00
# Discarded votes	0.11	0.50	0.00	0.00	4.00

Table 4

Statistics from the results of foot labeling algorithm over Experiment 1 sequences. Statistics were computed over the 80 sequences. First row shows statistics on the certainty for the chosen foot of being the “right” foot. The two other rows show the statistics for the number of considered and discarded votes respectively.

that would require information about foot labels.

3.2 Experiment 2: Evaluating the stability of normalized trajectories under walk direction changes

An important issue addressed by the proposed method is the viewpoint variation that occurs when a person changes direction during walk. In this case, the quality of the simulated fronto-parallel views obtained with normalization algorithms can be assessed by observing the stability of the spatial properties of the trajectories between each gait half-cycle. In the ideal case, the stability

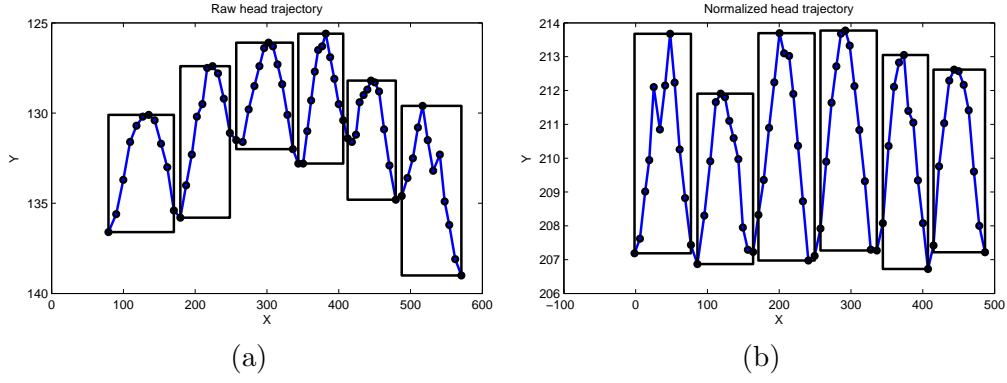


Fig. 10. Example of stability of the width and height of the head trajectory for a walk with a change in direction halfway in the walk. In a), the raw trajectory, with $c_h = 0.180$ and $c_w = 0.113$. In b), the normalized trajectory, with $c_h = 0.145$ and $c_w = 0.084$. In both cases, the bounding boxes for each gait half-cycle is shown.

of the normalized trajectories should be perfect, that is, no variation should be observed in the spatial properties of the trajectories. A comparison of the stability for both non-normalized (raw) and normalized body part trajectories should show less variation in the spatial properties in the normalized case.

3.2.1 Evaluation Methodology

The analysis of the height and the length of the body part trajectories across all gait half-cycles is a simple but efficient way to evaluate the stability of the normalized trajectories. Indeed, those two measures directly represent static gait characteristics, that is, stride length (trajectory length) and foot-floor clearance (trajectory height) in the case of the foot trajectories.

In order to evaluate the stability, the coefficients of variation of the height and the width of the body part trajectories across gait half-cycles are computed. They are computed as $c_h = \sigma_h / \mu_h$ for the height, and $c_w = \sigma_w / \mu_w$ for the width. σ and μ are respectively the standard deviation and the mean of the size of bounding boxes defined around the trajectory at each gait half-cycles. An example of the stability of the width and the height of the raw and normalized head trajectories is shown in Figure 10(a) and 10(b). In this example, the person walked on two straight line segments, with a change in direction halfway in the walk. One may notice that the amplitude (height) of the trajectory and the width have less variation across the gait half-cycle in the normalized case. The normalized trajectory has a lower coefficient of variation than the raw trajectory.

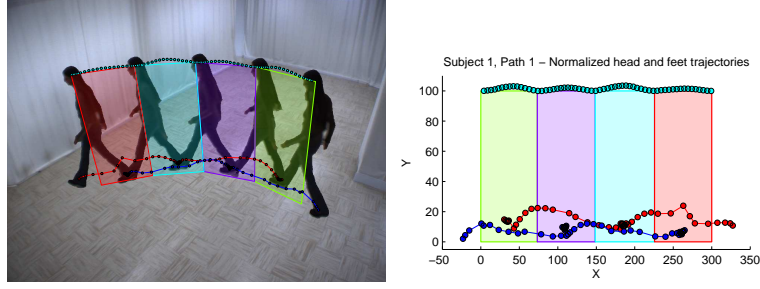


Fig. 11. Example of a subject walking on path 1, which is a walk with a smooth and continuous change in direction. Left figure shows the original body part trajectories along with computed walking planes. The silhouettes that have been observed at the maximum stance moments are shown. The figure at right shows the obtained normalized body part trajectories.

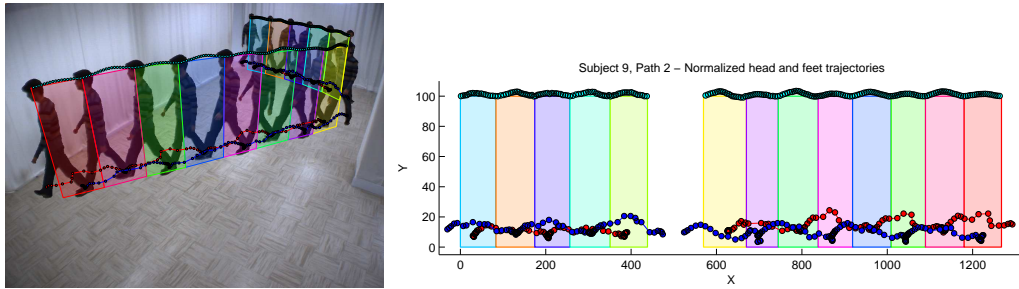


Fig. 12. Example of a subject walking on path 2, which is a walk with a sudden and strong change in the walk direction. The figure at left shows the original body part trajectories along with computed walking planes. The silhouettes that have been observed at the maximum stance moments are shown. The figure at right shows the obtained normalized body part trajectories. In this case, the body part tracking has been lost during the change in direction because of the long feet occlusion. Therefore, the sequence has been automatically split in two intervals.

3.2.2 Results

In the second experiment, 20 subjects were asked to walk on two different paths, which comprise changes in the walk direction. Forty video sequences were acquired using one camera, with frame resolution of 1024×768 , and frame rate of 30 fps. A variable focal length lens was used (3.5mm - 8mm) and it was set approximately to 4mm. The minimum distance between the subjects and the camera is about 2.5m, and the maximum distance is about 8.5m. The first path consists in a walk with a smooth and continuous change in direction. The second path consists in a walk with a sudden and strong change in the walking direction, which leads to two quasi-linear walk segments. Examples of the two paths are shown in the left image of Figures 11 and 12.

Unlike the evaluation method for the first experiment, trajectory alignment cannot be used to assess the effectiveness of the normalization algorithm since

Subject	Head				Feet			
	Height		Width		Height		Width	
	raw	norm	raw	norm	raw	norm	raw	norm
1	0.422	0.340	0.058	0.061	0.518	0.222	0.057	0.032
2	0.384	0.173	0.051	0.032	0.501	0.171	0.036	0.024
3	0.351	0.371	0.035	0.059	0.504	0.243	0.046	0.041
4	0.439	0.403	0.078	0.036	0.377	0.076	0.030	0.029
5	0.432	0.094	0.046	0.033	0.605	0.320	0.013	0.019
6	0.430	0.214	0.046	0.042	0.363	0.322	0.030	0.033
7	0.482	0.219	0.065	0.052	0.628	0.310	0.019	0.028
8	0.464	0.242	0.111	0.046	0.463	0.303	0.038	0.022
9	0.458	0.379	0.041	0.052	0.442	0.308	0.036	0.054
10	0.611	0.215	0.112	0.056	0.633	0.167	0.041	0.044
11	0.525	0.188	0.075	0.032	0.621	0.166	0.032	0.012
12	0.370	0.224	0.039	0.044	0.386	0.220	0.056	0.038
13	0.366	0.239	0.063	0.028	0.314	0.392	0.026	0.033
14	0.463	0.373	0.051	0.050	0.458	0.357	0.050	0.053
15	0.410	0.423	0.071	0.070	0.478	0.159	0.063	0.060
16	0.501	0.336	0.075	0.044	0.452	0.146	0.045	0.026
17	0.616	0.306	0.032	0.076	0.647	0.303	0.028	0.077
18	0.489	0.273	0.024	0.029	0.474	0.166	0.032	0.029
19	0.464	0.188	0.084	0.049	0.401	0.289	0.107	0.062
20	0.408	0.232	0.028	0.055	0.440	0.247	0.033	0.039
P-Value	1.34e-06		7.83e-02		1.83e-07		0.470	

Table 5

Coefficients of variation of the height and the width of the body part trajectories for all subjects in path 1. Values are shown for non-normalized (raw) and normalized trajectories (norm). The last line shows P-Values of a paired T-Test on corresponding coefficients of variation for raw and normalized body part trajectories.

there is no fronto-parallel view available for each walk direction in the second experiment. Assumptions made for the second experiment are that the subjects have walked at a constant speed, and that a subject's strides length are equal for both feet and for all gait half-cycles. The purpose of our evaluation is to show that in the case of the normalized trajectory, there is less variation across gait half-cycle for the width and the height of the body part trajectories. In the case of non-normalized trajectories, there will be more variation for those measures since the bounding boxes size changes under perspective projection. This is true unless the walk is entirely seen from a fronto-parallel view. Examples of raw and normalized trajectories from both paths are shown in Figures 11 and 12.

Table 5 and Table 6 shows the width and the height coefficient of variation of the bounding box of the body part trajectories for both paths. For the feet trajectories, only the gait half-cycles where the feet were moving have been considered. Moreover, coefficients of variation have been computed for both feet together. A paired T-Test has also been performed between coefficients of variation from raw and normalized trajectories to see if the difference is statistically significant.

One may see in those results that there is less variation for the normalized trajectories than for the non-normalized one in most case. Results from the first path (Table 5) shows that the variation is statistically significant at the 0.01 significance level for the height, but not for the width. This may be

Subject	Head				Feet			
	Height		Width		Height		Width	
	raw	norm	raw	norm	raw	norm	raw	norm
1	0.533	0.214	0.492	0.096	0.338	0.219	0.410	0.057
2	0.434	0.152	0.373	0.029	0.246	0.344	0.321	0.037
3	0.433	0.222	0.404	0.061	0.306	0.310	0.343	0.055
4	0.408	0.364	0.476	0.065	0.290	0.196	0.399	0.054
5	0.424	0.170	0.381	0.065	0.293	0.267	0.328	0.061
6	0.510	0.174	0.417	0.064	0.321	0.312	0.343	0.053
7	0.325	0.259	0.446	0.050	0.225	0.227	0.371	0.041
8	0.461	0.151	0.392	0.072	0.295	0.277	0.327	0.063
9	0.510	0.344	0.399	0.085	0.284	0.240	0.348	0.079
10	0.408	0.241	0.421	0.040	0.287	0.271	0.365	0.051
11	0.386	0.196	0.431	0.075	0.306	0.252	0.371	0.056
12	0.479	0.154	0.434	0.075	0.286	0.253	0.359	0.056
13	0.484	0.181	0.381	0.020	0.324	0.323	0.304	0.029
14	0.556	0.244	0.426	0.033	0.320	0.295	0.374	0.043
15	0.512	0.272	0.477	0.068	0.303	0.260	0.408	0.066
16	0.447	0.225	0.420	0.049	0.318	0.281	0.377	0.058
17	0.436	0.173	0.412	0.038	0.260	0.228	0.361	0.037
18	0.455	0.226	0.415	0.080	0.283	0.264	0.360	0.062
19	0.495	0.110	0.454	0.046	0.336	0.264	0.395	0.046
20	0.442	0.194	0.505	0.042	0.383	0.374	0.445	0.043
P-Value	1.22e-10		1.53e-20		1.06e-02		9.72e-20	

Table 6

Coefficients of variation of the height and the width of the body part trajectories for all subjects in path 2. Values are shown for non-normalized (raw) and normalized trajectories (norm). The last line shows P-Values of a paired T-Test on corresponding coefficients of variation for raw and normalized body part trajectories.

explained by the fact that all parts of this path are performed at almost the same distance from the camera, which makes the width of the body part trajectories similar in the raw and normalized case. The results of the second path (Table 6) are statistically significant at the 0.01 significance level, except for the feet height, which is border line. The feet tracking algorithm sometimes returns noisy points during feet occlusion when the observed subject is far away, as in the case of the second path. This noise is mostly in the y component of the feet positions, which accounts for the variation in the height of the obtained feet trajectories. However, the results from this path are statistically significant at the 0.05 significance level.

The second experiment has shown that the proposed normalization algorithm can be used to make body part trajectories appear as if they were obtained from a fronto-parallel view, even if there are changes in the walk direction.

4 Conclusion

An approach for view-normalizing body part trajectories was presented. The normalization process consists in the computation of walking planes from head and feet positions for each gait half-cycle. A homography transformation is then computed for each walking plane in order to make it appear as if it was observed from a fronto-parallel view, that is, a rectangle. Each computed

homography transforms the body part trajectories within the corresponding gait half-cycle, which makes them appear as obtained from a fronto-parallel view.

The proposed approach is promising since it has direct applications to gait modeling. Indeed, most gait characteristics are known to be easier to extract from a fronto-parallel view. As validated experimentally, the normalized trajectories of head and feet from different views are well aligned with real fronto-parallel view trajectories. Moreover, the proposed method can even be applied on trajectories for which changes in direction occurred during the walk.

Ongoing work focuses on testing the proposed approach on trajectories of additional body parts (hands, knees etc.) involved in human walk. More tests are to be performed on trajectories with changes in walking direction. Gait modeling will be performed by extracting gait characteristics from normalized body part trajectories, which will eventually lead to view-invariant gait-based identification.

Acknowledgments

This work is supported by the Natural Sciences and Engineering Research Council of Canada, the Fonds Québécois de la Recherche sur la Nature et les Technologies, and by Precarn inc.

References

- [1] G. Johansson, Visual perception of biological motion and a model for its analysis, *Perception Psychophysics* 14 (1973) 201–211.
- [2] M. R. Lemke, N. Koethe, M. Schleidt, Segmentation of behavior and time structure of movements in depressed patients, *Psychopathology* 33 (2000) 131–136.
- [3] J. Verghese, R. B. Lipton, C. B. Hall, G. Kuslansky, M. J. Katz, H. Buschke, Abnormality of gait as a predictor of non-alzheimer’s dementia, *The New England Journal of Medicine* 347 (2002) 1761–1768.
- [4] L. Z. Rubenstein, D. H. Solomon, C. P. Roth, R. T. Young, P. G. Shekelle, J. T. Chang, C. H. MacLean, C. J. Kamberg, D. Saliba, N. S. Wenger, Detection and management of falls and instability in vulnerable elders by community physicians, *Journal of the American Geriatrics Society* 52 (2004) 1527–1531.

- [5] L. Wang, Abnormal walking gait analysis using silhouette-masked flow histograms, in: Proceedings of the 18th International Conference on Pattern Recognition, Vol. 3, 2006, pp. 473–476.
- [6] C. Bauckhage, J. K. Tsotsos, F. E. Bunn, Automatic detection of abnormal gait, *Image and Vision Computing* 27 (2009) 108–115.
- [7] M. P. Murray, Gait as a total pattern of movement, *American Journal of Physical Medicine* 13 (1967) 290–332.
- [8] A. Veeraraghavan, A. R. Chowdhury, R. Chellappa, Role of shape and kinematics in human movement analysis, in: Proceedings of the Conference on Computer Vision and Pattern Recognition, Vol. 1, 2004, pp. I-730–I-737 Vol.1.
- [9] C. BenAbdelkader, R. Cutler, H. Nanda, L. S. Davis, *EigenGait: Motion-Based Recognition of People Using Image Self-Similarity*, Vol. 2091 of LNCS, Springer, 2001, pp. 284–294.
- [10] G. V. Veres, L. Gordon, J. N. Carter, M. S. Nixon, What image information is important in silhouette-based gait recognition?, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Vol. 2, 2004, pp. 776–782.
- [11] J. Han, B. Bhanu, Statistical feature fusion for gait-based human recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Vol. 2, 2004, pp. 842–847.
- [12] Z. Liu, S. Sarkar, Improved gait recognition by gait dynamics normalization, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28 (2006) 863–876.
- [13] J. E. Cutting, D. R. Proffitt, Gait perception as an example of how we may perceived events, *Intersensory perception and sensory integration* (1981) 249–273.
- [14] J. Han, B. Bhanu, Performance prediction for individual recognition by gait, *Pattern Recognition Letters* (2005) 615–624.
- [15] C. BenAbdelkader, R. G. Cutler, L. S. Davis, Person identification using automatic height and stride estimation, in: Proceedings of the 16th International Conference on Pattern Recognition, Vol. 4, 2002, pp. 377–380.
- [16] R. Zhang, C. Vogler, D. Metaxas, Human gait recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Vol. 1, 2004, pp. 18–25.
- [17] A. Bissacco, A. Chiuso, Y. Ma, S. Soatto, Recognition of human gaits, in: Proceedings of the Conference on Computer Vision and Pattern Recognition, Vol. 2, 2001, pp. II-52–II-57 vol.2.
- [18] A. F. Bobick, A. Y. Johnson, Gait recognition using static, activity-specific parameters, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2001.

- [19] J.-H. Yoo, M. S. Nixon, Markerless human gait analysis via image sequences, in: Proceedings of International Society of Biomechanics XIXth Congress, 2003.
- [20] S. Yu, L. Wang, W. Hu, T. Tan, Gait analysis for human identification in frequency domain, in: Proceedings of the Third International Conference on Image and Graphics, 2004, pp. 282–285.
- [21] C. BenAbdelkader, R. Cutler, L. Davis, View-invariant estimation of height and stride for gait recognition, in: Proceedings of the International ECCV Workshop on Biometric Authentication, Vol. LNCS 2359, 2002, pp. 155 – 167.
- [22] N. M. Spencer, J. Carter, Towards pose invariant gait reconstruction, in: Proceedings of the IEEE International Conference on Image Processing, Vol. 3, 2005, pp. 261–264.
- [23] A. Kale, A. K. R. Chowdhury, R. Chellappa, Towards a view invariant gait recognition algorithm, in: Proceedings of IEEE Conference on Advanced Video and Signal Based Surveillance, 2003, pp. 143–150.
- [24] A. Y. Johnson, A. F. Bobick, A Multi-view Method for Gait Recognition Using Static Body Parameters, Vol. 2091 of LNCS, Springer, 2001, pp. 301–311.
- [25] Y. Makihara, R. Sagawa, Y. Mukaigawa, T. Echigo, Y. Yagi, Adaptation to walking direction changes for gait identification, in: Proceedings of the 18th International Conference on Pattern Recognition, Vol. 2, 2006, pp. 96–99.
- [26] M. Hild, Estimation of 3d motion trajectory and velocity from monocular image sequences in the context of human gait recognition, in: Proceedings of the 17th International Conference on Pattern Recognition, Vol. 4, 2004, pp. 231– 235.
- [27] X. Han, J. Liu, L. Li, Z. Wang, Gait recognition considering directions of walking, in: Proceedings of the IEEE Conference on Cybernetics and Intelligent Systems, 2006, pp. 1–5.
- [28] A. Tyagi, J. W. Davis, M. Keck, Multiview fusion for canonical view generation based on homography constraints, in: Proceedings of the 4th ACM international workshop on Video surveillance and sensor networks, ACM Press, Santa Barbara, California, USA, 2006, pp. 61–70.
- [29] C.-S. Lee, A. Elgammal, Towards Scalable View-Invariant Gait Recognition: Multilinear Analysis for Gait, Vol. 3546 of LNCS, Springer, 2005, pp. 395–405.
- [30] F. Cuzzolin, Using bilinear models for view-invariant action and identity recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Vol. 2, 2006, pp. 1701– 1708.
- [31] S. Yu, D. Tan, T. Tan, Modelling the Effect of View Angle Variation on Appearance-Based Gait Recognition, Vol. 3851 of LNCS, Springer, 2006, pp. 807–816.
- [32] Y. Makihara, R. Sagawa, Y. Mukaigawa, T. Echigo, Y. Yagi, Which reference view is effective for gait identification using a view transformation model?, in: Proceedings of the Conference on Computer Vision and Pattern Recognition Workshop, 2006, pp. 45–52.

- [33] F. Jean, R. Bergevin, A. B. Albu, Trajectories normalization for viewpoint invariant gait recognition, in: Proceedings of the 19th International Conference on Pattern Recognition, Tampa, Florida, USA, 2008.
- [34] F. Jean, R. Bergevin, A. B. Albu, Body tracking in human walk from monocular video sequences, in: Proceedings of the Second IEEE Canadian Conference on Computer and Robot Vision, Victoria, BC, Canada, 2005, pp. 144 – 151.
- [35] K. Rangarajan, M. Shah, Establishing motion correspondence, in: Proceedings of the Conference on Computer Vision and Pattern Recognition, 1991, pp. 103–108.
- [36] R. I. Hartley, A. Zisserman, Multiple View Geometry in Computer Vision, 2nd Edition, Cambridge University Press, 2004.
- [37] K. S. Arun, T. S. Huang, S. D. Blostein, Least-squares fitting of two 3-d point sets, IEEE Transactions on Pattern Analysis and Machine Intelligence 9 (1987) 698–700.