



Viewpoint-independent gait recognition through morphological descriptions of 3D human reconstructions[☆]



D. López-Fernández^{1,*}, F.J. Madrid-Cuevas¹, A. Carmona-Poyato¹, M.J. Marín-Jiménez¹,
R. Muñoz-Salinas¹, R. Medina-Carnicer¹

Department of Computing and Numerical Analysis, Maimnides Institute for Biomedical Research (IMIBIC), University of Córdoba, Córdoba, Spain

ARTICLE INFO

Article history:

Received 23 October 2014
Received in revised form 9 January 2016
Accepted 19 January 2016
Available online 23 February 2016

Keywords:

Gait recognition
Morphology
View-independent
Appearance-based
3D reconstruction
Histogram

ABSTRACT

Many studies have confirmed gait as a robust biometric feature for identification of individuals. However, direction changes cause difficulties for most of the gait recognition systems, due to appearance changes. This study presents an efficient multi-view gait recognition method that allows curved trajectories on unconstrained paths in indoor environments. The recognition is based on volumetric analysis of the human gait, to exploit most of the 3D information enclosed in it. Appearance-based gait descriptors are extracted from 3D gait volumes and temporal patterns of them are classified using a Support Vector Machine with a sliding temporal window for majority voting. The proposed approach is experimentally validated on the “AVA Multi-View Dataset (AVAMVG)” and on the “Kyushu University 4D Gait Database (KY4D)”. The results show that this new approach is able to identify people walking on curved paths.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Research on human gait as a biometric feature for identification has received a lot of attention due to the apparent advantage that it can be applied discreetly on the observed individual without needing the active participation of the individual.

Previous studies on gait recognition have been classified into two categories: model-based and model-free approaches. The model-based methods extract gait features by fitting a model to input images, whereas model-free approaches characterize the human gait pattern by a compact representation, without having to develop any articulated model for feature extraction and having practical application even with low quality images where the color and texture information is lost.

In addition, regarding viewing angle, the previous work can be categorized into two approaches: view-dependent and view-independent approaches. The view-dependent approaches assume that the appearance will not change during walking. In such methods, a change in the appearance, caused by a view change, will adversely affect performance [1]. Fig. 1 shows the influence of a curved path on the silhouette appearance. As one of the advantages of gait as biometric is that it does not need the cooperation of the individual, the trajectory of motion cannot be restricted to just straight paths.

On the other hand, the use of volumetric information allows more information to be analyzed in contrast to methods which only compute gait descriptors from silhouettes or 2D images. This paper presents an efficient view-independent method to recognize people walking along unconstrained (curved and straight) trajectories. This approach focuses on capturing 3D morphological and structural information from volumetric reconstructions of walking humans, which are previously aligned along the way. The main contribution is that our method allows direction changes, achieving a good recognition rate on unconstrained paths.

Some potential applications of this work are access control in special or restricted areas (e.g. military bases, governmental facilities and laboratories) or smart video surveillance (e.g. bank offices) [2].

This article is organized as follows. Section 2 describes works related to the topic of gait recognition. Section 3 explains the details

[☆] This paper has been recommended for acceptance by Mark S. Nixon.

* Corresponding author.

E-mail addresses: i52lofed@uco.es (D. López-Fernández), fjmadrid@uco.es (F. Madrid-Cuevas), ma1capoa@uco.es (A. Carmona-Poyato), mjmarin@uco.es (M. Marín-Jiménez), rmsalinas@uco.es (R. Muñoz-Salinas), rmedina@uco.es (R. Medina-Carnicer).

¹ Computing and Numerical Analysis Department, Edificio Einstein, Campus de Rabanales, Córdoba University, 14071 Córdoba, Spain. Tel.: +34 957212255.

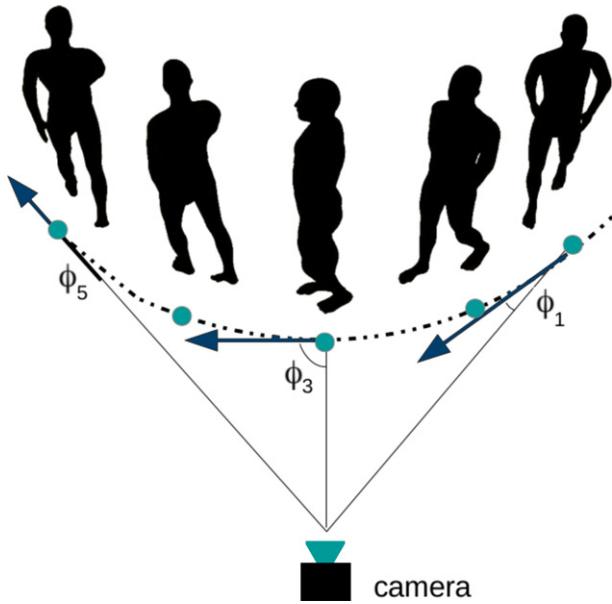


Fig. 1. In a curved path, the observation angle between the walking direction of the subject and optical axis of the camera is gradually changed, which affects the silhouette appearance.

of the proposed algorithm and describes three new descriptors which obtain information from 3D occupancy volumes. An analysis of the proposed method and the performance is given in Section 4. Finally, we conclude this paper in Section 5.

2. Related work

The previous work can be categorized into two approaches: view-dependent and view-independent approaches. In the following we describe works related to both categories.

2.1. View-dependent approaches

One of the earliest model-free and view-dependent approaches can be seen in [3], where the width of the outer contour of the binarized silhouette from a side view is used to build a descriptor which contains both structural features and dynamic aspects of gait. Feature vectors derived from binary silhouettes have been also used to train Hidden Markov Models [4]. The contours of silhouettes have been used directly [5,6], and through their Fourier descriptors [7,8].

In addition, the authors of Ref. [9] present a gait recognition method which analyzes the shape of the silhouette using Procrustes Shape Analysis and Elliptic Fourier Descriptors. The Gait Energy Image (GEI) descriptor is introduced in [10], which is the average of all silhouette images for a single gait cycle. Silhouette images are also used by Lam et al. [11] to generate the gait flow image (GFI).

Based on the idea of GEI, Depth Energy Image (DEI) was defined in [12], which is simply the average of the depth silhouettes taken along a gait cycle, over the front view. GEI is also extended in [13] to consider depth information from the side view, by means of a new feature called Depth Gradient Histogram Energy Image (DGHEI). Depth information is also used by Chattopadhyay et al. [14] to address the problem of occlusion in frontal gait recognition.

The Gait Energy Volume (GEV), a binary voxel-discretized volume which is spatially aligned and averaged over a gait cycle, is presented in [15]. The authors apply GEV on partial reconstructions

obtained with depth sensors from the front view of the individual. An extended work from GEV [15] that combines the frontal-view depth gait image and side-view 2D gait silhouette by means of a back-filling technique is presented in [16]. A spatio-temporal representation based on point clouds in a spherical coordinate space was proposed in [17], where frontal 3D point clouds of humans obtained with stereo cameras are used.

A work closely related to our proposed approach in terms of analysis by morphological size distributions was proposed in [18]. In this work, video cameras are placed in hallways to capture longer sequences from the front view of walkers rather than the side view, which results in more gait cycles per gait sequence. Despite the high recognition rate, the main drawback of this model-free approach is the dependence with respect to the viewpoint. To obtain a gait representation directly from silhouettes, the authors proposed the use of a morphological descriptor, called Cover by Rectangles, which is defined as the union of all the largest rectangles that can fit inside a silhouette.

In [19], a 3D approximation of a Visual Hull (VH) [20] is used to design a multi-modal recognition approach. Although a VH model is computed, a gait recognition scheme based on silhouette analysis is applied, which restricts a large amount of discriminant information because the recognition is based on single view silhouette analysis, instead of analyze 3D information. Seely et al. [21] use 3D volumetric data to synthesize silhouettes from a fixed viewpoint relative to the subject. The resulting silhouettes are then passed to a standard 2D gait analysis technique, such as the average silhouette. The sequences are collected from a multi-biometric tunnel, where the subjects just walk straight.

Ariyanto and Nixon [22] propose a model-fitting algorithm, correlation filters and dynamic programming to extract gait kinematics features. They use a structural model including articulated cylinders with 3D Degrees of Freedom (DoF) which are fitted to a visual hull shape to model the human lower legs. In [23], 3D data collected from a projector-camera system is used to fit 3D body models and reconstruct synthetic poses in a gait cycle.

2.2. View-independent approaches

Appearance changes due to viewing angle changes cause difficulties for most of the model-free gait recognition methods. This situation cannot be easily avoided in practical applications. There are three major approach categories to sort out this problem, namely: (1) approaches that construct 3D gait information through multiple calibrated cameras; (2) approaches that extract gait features which are invariant to viewing angle changes; (3) approaches whose performance relies on learning mapping/projection relationship of gaits under various viewing angles [24].

Approaches of the first category are represented in [25,26]. Bodor et al. [25] apply image-based rendering on a 3D VH model to reconstruct gait features under a required viewing angle. This approach tries to classify the motion of a human in a view-independent way, but it has two drawbacks. On the one hand it considers only straight paths to estimate the position and orientation of a virtual camera. Tests were performed only on straight path motions. On the other hand, not all the 3D information available in the VH is used, because feature images are extracted from 2D images rendered only from a single view.

In [26], an observation angle at each frame of a gait sequence is estimated from the walking direction, by fitting a 2D polynomial curve to the foot points. Virtual images are synthesized from 3D reconstructions, so that the observation angle of a synthesized image is the same that the observation angle for the real image of the subject, which is identified by using affine moment invariants extracted from images as gait features. The advantage of this method is that the setup assumes multiple cameras for training, but only one camera for

testing. However, as in the above two works, despite 3D volumes are used, descriptors are extracted from 2D images, so that, the amount of used information is restricted. On the other hand, shadows on the floor complicate the estimation of the foot points in silhouette images.

Approaches of the second category extract gait features which are invariant to viewing angle change. A method to generate a canonical view of gait from any arbitrary view is described in [27]. The main disadvantage of this method is that the synthesis of a canonical view is only feasible from a limited number of initial views. The performance is significantly dropped down when the angle between image plane and sagittal plane is large.

In [28], a method based on homography to compute view-normalized trajectories of body parts obtained from monocular video sequences was proposed. But this method only works properly for a limited range of views. Planar homography has also been used to reduce the dependency between the motion direction and the camera optical axis [29], however this method seems not to be applicable when the person is walking nearly parallel to the optical axis. In [30] view-invariant features are extracted from GEI. Only parts of gait sequences that overlap between views are selected for gait matching, but this approach cannot cope with large view angle changes under which gait sequences of different views can have little overlap.

A self-calibrating view-independent gait recognition based on model-based gait features is proposed in [31]. The poses of the lower limbs are estimated based on markerless motion estimation. Then, they are reconstructed in the sagittal plane using viewpoint rectification. This method has two main drawbacks that are worth mentioning: 1) the estimation of the poses of the limbs is not robust from markerless motion; 2) it is not applicable for frontal view because the poses of the limbs become untraceable; and 3) this method assume that subjects walk along a straight line segment.

Zhao et al. [32] present a multi-camera approach for gait tracking and recognition. The video sequences are used as input, and then a human 3D model is set up. The lengths of key segments are extracted as static parameters, and the motion trajectories of lower limbs are used as dynamic features. A skeletal 3D model is also used in the work of Kastaniotis et al. [33], which presents a framework for pose-based gait recognition and identification, as well as gender recognition.

The approaches of the third category rely on learning mapping/projection relationship of gaits under various viewing angles. The trained relationship may normalize gait features from different viewing angles into shared feature spaces. An example from this category can be read in [34], where LDA-subspaces are learned to extract discriminative information from gait features under each viewing angle.

A View Transformation Model (VTM) was introduced in [35] to transform gait features from different views into the same view. The method of Makihara et al. [35] creates a VTM based on frequency-domain gait features, obtained through Fourier Transformation. To improve the performance of this method, Kusakunniran et al. [36] created a VTM based on GEI optimized by linear discriminant analysis. A sparse-regression-based VTM for gait recognition under various views is also proposed in [24]. However, this method cannot deal with changes in the direction of motion.

Although methods of the third category have better ability to cope with large view angle changes compared to other works, some common challenges are the following [24]: (1) performance of gait recognition decreases as the viewing angle increases; (2) since the methods rely on supervised learning, it will be difficult for recognizing gait under untrained/unknown viewing angles, (3) these methods implicitly assume that people walk along straight paths and that their walking direction does not change during a single gait cycle (i.e., that people do not walk along curved trajectories). However,

people often walk on curved trajectories in order to turn a corner or to avoid an obstacle.

3. Proposed method

We propose a model-free approach to recognize walking humans independently of the viewpoint and regardless direction changes. Our approach focuses on capturing 3D morphological and structural information from the gait through volumetric reconstructions of the walking humans.

The use of volumetric information allows more information to be analyzed in contrast to other related works, which only compute gait descriptors from silhouettes, discarding an important part of the dynamical and structural information of the gait. So that, our method extends the input domain for the morphological gait descriptors used in [18], from 2D silhouettes to 3D reconstructions of the individuals, aligned along the way.

Our approach relies on morphological analysis of series of 3D occupation volumes which are generated from the multi-view video sequences. For each time of the gait sequence, a 3D occupation volume is obtained by combining information from multiple silhouette images, from several points of view. Then, this gait volume is aligned and centered with respect to a global reference system. Next, our gait descriptor is computed from each 3D gait volume in order to provide information about their 3D appearance.

A gait signature is built by aggregating descriptors. The gait signature is a temporal pattern of gait, a sample that feeds a classifier producing a class label corresponding to the identity of a particular person. The proposed recognition algorithm is shown in Fig. 2, where the identity of a walking human is predicted at each time t . The algorithm consists of five steps which are exposed in detail in this section:

1. Silhouette extraction of each camera's view by a background subtraction technique [37].
2. 3D reconstruction from silhouettes captured from several viewpoints, by a Shape from Silhouette algorithm (SfS) [38].
3. Person detection and gait alignment.
4. Gait descriptor generation, which is used to update the gait signature.
5. Classification of gait signature by a machine learning algorithm.

The first three steps of the algorithm generate a 3D volume with occupancy information of the person at time t , whereas the last three steps perform the feature extraction, signature generation and gait classification.

3.1. 3D reconstruction, detection and alignment

We start by computing a 3D reconstruction of the individual from silhouettes extracted from several viewpoints. This procedure requires calibration parameters, such as the camera matrix, distortion coefficients (intrinsic parameters), pose and orientation (extrinsic parameters) of each camera.

After the 3D reconstruction, the individual is detected and the 3D volumes corresponding to a gait sequence are aligned and centered with respect to a global reference system, so that the generation of the descriptors can be made as if the person had walked on a treadmill in a certain direction.

3.1.1. 3D reconstruction

Since our method generates gait descriptors from 3D occupation volumes or VH, a 3D reconstruction procedure, such a Shape from Silhouette algorithm [38] is required. Fig. 3 shows a 3D reconstructed gait sequence.

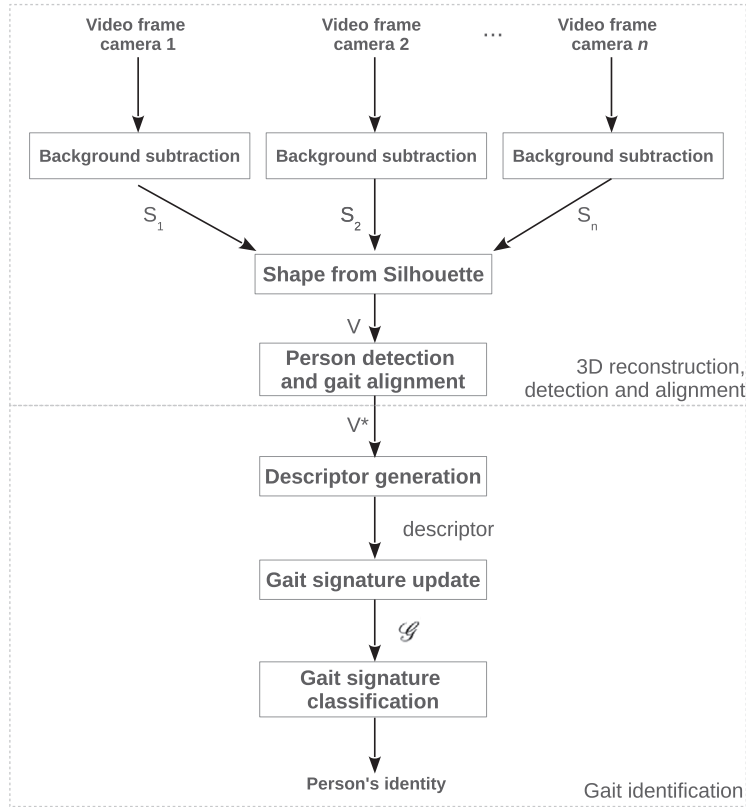


Fig. 2. Steps of our gait-recognition algorithm at time t .

3.1.2. Volume detection and alignment

Given a reconstructed volume V_t of a person at each instant t along the way, it required a mechanism of detection and alignment to achieve the independence which refers to the viewpoint. This process will allow the individual to walk freely in the scene, without adversely affecting the subsequent generation of gait descriptors.

We assume that although there is only one individual in the scene, reconstructed shadows as well as noise can coexist, due to poor segmentation. By obtaining the ground marginal distribution

of occupied voxels (ground projection of the volume), we detect the volume belonging to a person as that which has a greater volume than a certain threshold ϕ , and its volume has fully entered into the workspace. The value of ϕ is tuned up accordingly to the average corporal volume for humans and the resolution of the 3D reconstructions. This is described in Section 4.2.

When the volume belonging to a person has been detected, the centroid p of the ground projection is calculated. Then, the volume is moved into a bounding-box of average adult human's size, so that the workspace where the descriptor will be computed is bounded. This process is illustrated in Fig. 4. The estimation of the direction of path is determined by the displacement vector, defined as:

$$\vec{v}_t = p_t - p_{t-1}, \quad (1)$$

where t is the current time, p_t is the centroid's current position, and p_{t-1} is the last known position of the centroid.

The angle of the displacement vector is calculated using the expression:

$$\alpha_t = \arctan \frac{v_{ty}}{v_{tx}}. \quad (2)$$

An example of extraction of the displacement vector angle can be seen in Fig. 5, where top projections of the individual can be seen in several moments of the gait, and the principal axis (perpendicular to displacement vector) is represented. The reconstructed volume is rotated about the body vertical axis.

Although we assume a constant walking speed, an individual could vary moderately the walking speed in a certain moment of the

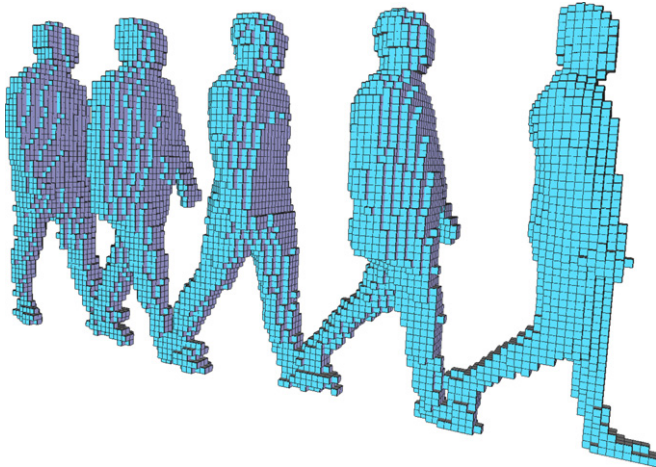


Fig. 3. Example of reconstructed gait sequence, sampled at 2Hz, where each cube represents a voxel.

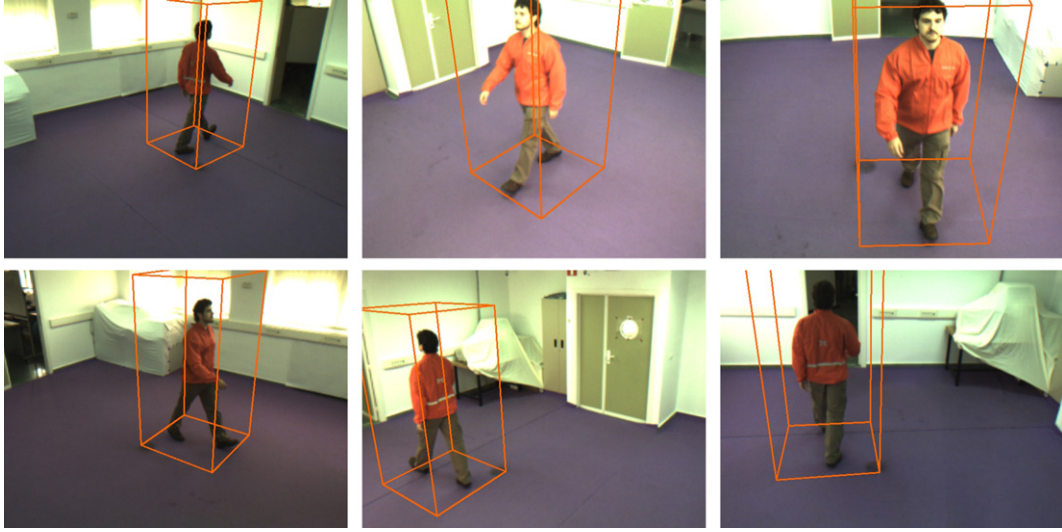


Fig. 4. Bounding-box with size of an average adult human, where the descriptor will be computed on the 3D reconstructed data at time t .

gait. It could happen, for example, when the individual is describing a curved closed path.

If the walking speed is very low at time t , $|\vec{v}_t|$ will be too small, which could result in a noisy estimation of the angle α_t . To attenuate this noise in the α_t estimation and smooth the path, we propose a weighted average of the displacement vector angle as follows:

$$\bar{\alpha}_t = \alpha_t \cdot \beta + \bar{\alpha}_{t-1} \cdot (1 - \beta), \quad (3)$$

where

$$\beta = \frac{|\vec{v}_t|}{\max_{i=0 \dots t} \{|\vec{v}_i|\}}. \quad (4)$$

The aim of β is to reduce this noise in the estimation of the alignment angle, giving more or less weight to the current estimation

depending on the walking speed. For example, if the walking speed is decreasing and it becomes too small, the magnitude of the displacement vector may not be large enough, causing oscillations in the estimation of the angle. In this case, it would be right to give less importance to the current angle estimation. However, if the walking speed is increasing, it would be more appropriate to give more weight to the current angle estimation. A method to decrease over time the denominator in Eq. (4) should be applied if the gait sequence were too large. Thus, the whole gait sequence can be centered and aligned along the path as it is illustrated in Fig. 6.

3.2. Gait identification

The algorithm steps that handle up the gait identification are described below.

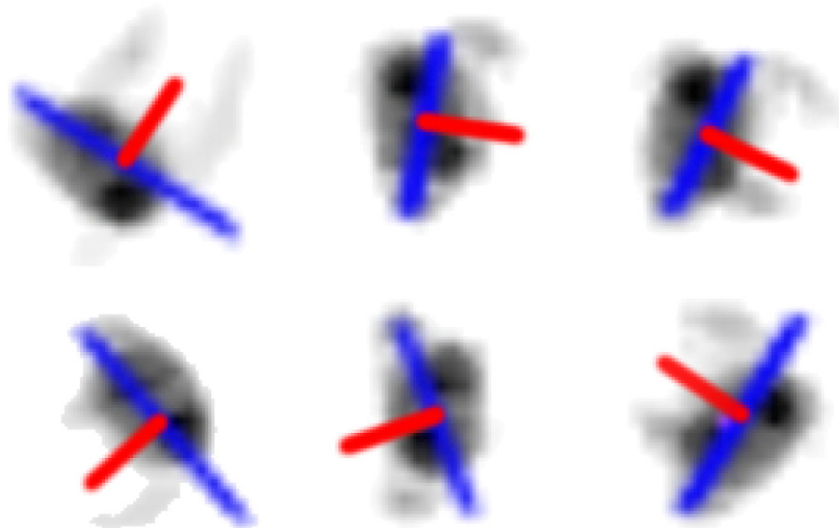


Fig. 5. Displacement vector (red line) of the individual is computed at each time. The principal axis (blue line) is perpendicular to the displacement vector.

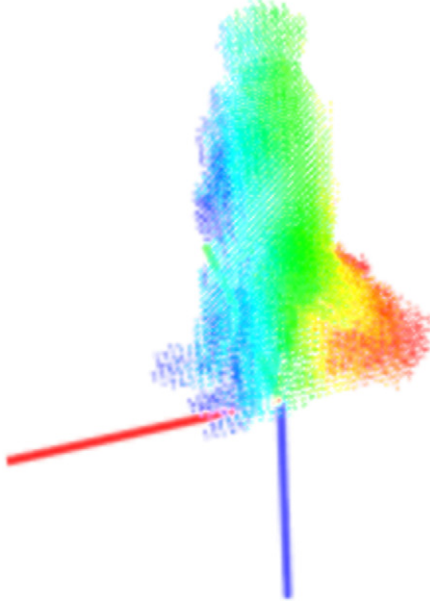


Fig. 6. Union of all aligned 3D volumes over the gait sequence. Different colors represent the distance to the center on the X axis.

3.2.1. Descriptor generation

The first step is the generation of the gait descriptor on the V_t volume. We propose the three following candidate descriptors.

- **Cover by Rectangles from frontal volume projection:**

The Cover by Rectangles descriptor, denoted here as $CR(S)$, was proposed by Barnich and Van Droogenbroeck [18]. It is defined as the union of all the largest rectangles that can fit inside of a silhouette S . In Barnich's method, video cameras are placed in hallways to capture longer sequences from the front view of walkers rather than the side view, which results in more gait cycles per gait sequence. The main drawback of this method is the dependence on the viewpoint.

Each silhouette is then converted to an intra-frame histogram, which compacts the width and height distributions of the set of all the rectangles that can be wedged inside the silhouette. In order to build the histogram, the widths and heights of the rectangles are discretized into M and N bins respectively.

As the occupation volumes have been aligned to achieve the view independence, a virtual camera can be placed in front of the volumes to obtain projections on which the $CR(S)$ can be computed, as can be seen in Fig. 7 (b).

- **Cover by Cubes:**

We propose a new gait descriptor defined as the union of all the cubes with the largest size that can fit into a volume belonging to the person. It is called *Cover by Cubes*.

Considering V as the 3D volume of a person in a moment of the gait, $CC(V)$ is the union of all cubes of maximum size that can fit in it. The new descriptor deals with three-dimensional domain spatial information, and like *Cover by Rectangles*, it has the following useful properties:

- The elements of the set overlap each other, introducing redundancy (i.e. robustness).
- Each element (cube) of $CC(V)$ covers at least one voxel that belongs to no other cube.

- The union of all cubes reconstructs the volume V so that no information is ever lost.

Let $\alpha = \#CC(V)$ be the cardinality of the set $CC(V)$. The cubes of $CC(V)$ are indexed with a parameter e , so that C_e ($e = 1, \dots, \alpha$) are the cubes of $CC(V)$. The width, height and depth of C_e are, respectively, denoted by w_e , h_e and d_e ; and they are upper-bounded by w^{max} , h^{max} and d^{max} , so $\forall e$, $w_e \leq w^{max}$, $h_e \leq h^{max}$ and $d_e \leq d^{max}$.

In order to build histograms, the widths, heights and depths of the cubes R_e are discretized into M bins $B_W(i)$, N bins $B_H(j)$ and D bins $B^D(k)$

$$B^W(i) = \left[i \frac{w^{max}}{M}, (i+1) \frac{w^{max}}{M} \right), \quad (5)$$

$$B^H(j) = \left[j \frac{h^{max}}{N}, (j+1) \frac{h^{max}}{N} \right), \quad (6)$$

$$B^D(k) = \left[k \frac{d^{max}}{D}, (k+1) \frac{d^{max}}{D} \right) \quad (7)$$

where $i = 0, \dots, M-1$; $j = 0, \dots, N-1$ and $k = 0, \dots, D-1$.

Three histograms are defined, $H^W(i)$, $H^H(j)$ and $H^D(k)$:

$$H^W(i) = \frac{1}{\alpha} \# \{C_e | w_e \in B^W(i)\}, \quad (8)$$

$$H^H(j) = \frac{1}{\alpha} \# \{C_e | h_e \in B^H(j)\}, \quad (9)$$

$$H^D(k) = \frac{1}{\alpha} \# \{C_e | d_e \in B^D(k)\}, \quad (10)$$

and the three-dimensional histogram $H^{W \times H \times D}$ as:

$$H^{W \times H \times D}(i, j, k) = \frac{1}{\alpha} \# \{C_e | w_e \in B^W(i), h_e \in B^H(j), d_e \in B^D(k)\}. \quad (11)$$

All these histograms are normalized taking into account the number of cubes of maximum size of the volume.

From the four histograms, $H^W(i)$, $H^H(j)$, $H^D(k)$ and $H^{W \times H \times D}(i, j, k)$, the latter is the one that better describes V . However, its dimensionality is proportional to the product of the numbers of bins ($M \times N \times D$), which might be too high, e.g., for embedded systems. To deal with such situation, a composite histogram is proposed, $H^{W+H+D}(l)$, with $l = 0, \dots, M+N+D-1$ defined as the concatenation of $H^W(i)$, $H^H(j)$ and $H^D(k)$ (marginal distributions).

An example of Cover by Cubes histograms is shown in Fig. 8.

In this figure, the joint distribution and the marginal distributions of the histograms can be seen, corresponding to Eqs. (11), (8), (9), and (10).

- **Cover by Rectangles from top, side, and frontal volume projections:**

The availability of 3D gait volumes leads us to that think we can use several projections of the gait volumes, instead of just using the frontal projection of them, in order to exploit the 3D information of gait.

So following the idea about the use of virtual cameras, we also propose a new descriptor based on computing the CR descriptor on the front, side, and top projections of the volume V , as it is shown in Fig. 7. Its concatenation can be denoted as $CRP(V)$.

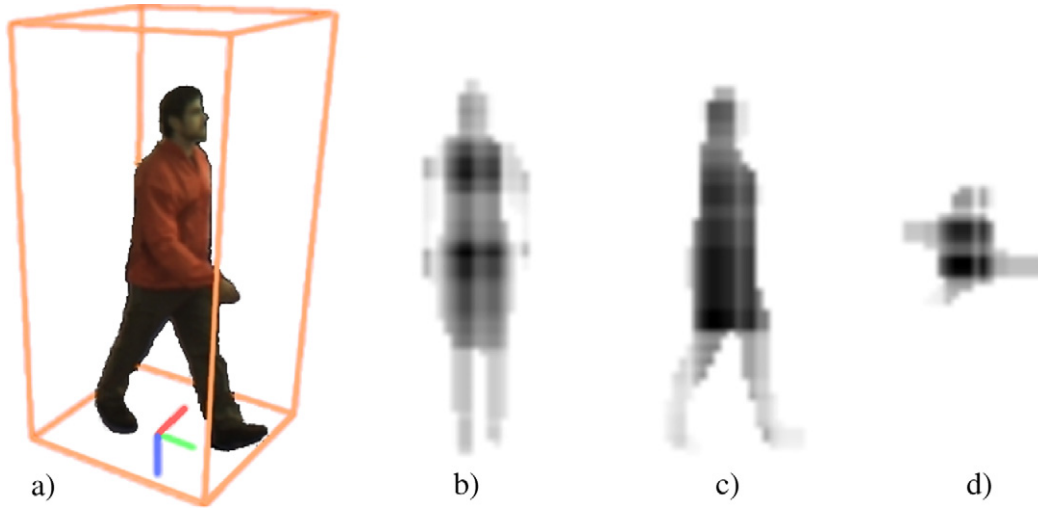


Fig. 7. Cover by Rectangles descriptors computed on the front (b), side (c), and top (d) view projections of a 3D volume reconstruction (a). A gray level of pixel in images (b) and (d) displays the density of rectangles that contain that pixel.

$$CRP(V) = CR(side), CR(front), CR(top), \quad (12)$$

where *side*, *front* and *top* are lateral, frontal and top rendered projections of *V*, respectively.

3.2.2. Gait signature update

The action immediately prior to the classification by a machine learning algorithm is the generation of the sample or vector of features. A sample is generated at every moment of the walking, which enables synchronous classification. This is known as \mathcal{G} or gait signature, and represents a temporal pattern of movement of the person.

For the CR descriptor, obtained on the front projection of the volume, the signature \mathcal{G} can be obtained by combining a number L of successive histograms into a single spatio-temporal (inter-frame) gait signature, as it was described in [18]. The signature can be made by the combination of the marginal or joint distribution of the histograms and it needs to be updated in every time t .

However, with the new proposed gait descriptors (CC and CRP) which get information from 3D occupancy volumes instead of getting

it from silhouettes, it is necessary to reformulate the procedure to construct the signature \mathcal{G} . The way in which this signature is built depends on the descriptor.

With regard to CC descriptor, the gait signature \mathcal{G} relies on temporal series of descriptors obtained from the 3D volumes of a person's gait sequence. So if t refers to the current time, and $H(i, j, k, t)$ is the Cover by Cubes descriptor obtained from the volume V_t , we have two possible signatures as follows:

$$\mathcal{G}^{W \times H \times D}(i, j, k, t) = H^{W \times H \times D}(i, j, k, t - (L - 1)), \dots, H^{W \times H \times D}(i, j, k, t) \quad (13)$$

which consist of n -uples of L consecutive histograms, and a shortened version as:

$$\mathcal{G}^{W+H+D}(o, t) = H^{W+H+D}(o, t - (L - 1)), \dots, H^{W+H+D}(o, t) \quad (14)$$

where $o = 0, \dots, M + N + D - 1$.

Similarly, for CRP descriptors, the gait signature \mathcal{G} can be also composed by aggregating, on a sliding window, L CRP descriptors (joint or marginal distribution of the histograms computed on rendered side, top and front projections of V).

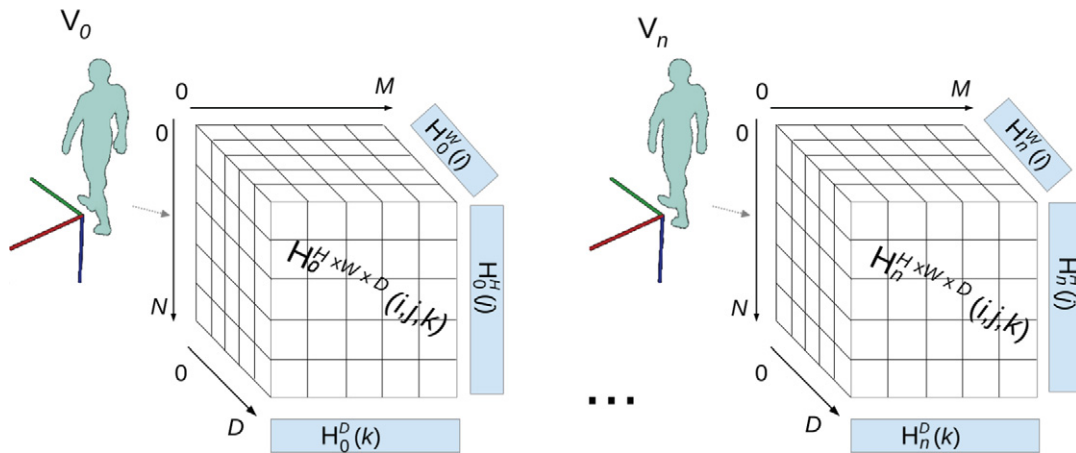


Fig. 8. Cover by Cubes histograms, where n refers to the length of the gait sequence. $M \times N \times D$ is the number of bins of the histogram $H^{W \times H \times D}$, and the blue rectangles represent marginal distributions.

3.2.3. Classification

The gait signature obtained at time t is the feature vector used for recognition. Each of these feature vectors is assigned to a class label that corresponds to one of the individuals in the database. This idea is well known as multi-class classification system. We adopt a Support Vector Machine (SVM) [39] for training and classification.

Our recognition algorithm provides the identity of the person as soon as possible, without having to split the gait sequence into gait cycles. This makes our method less restrictive compared to other techniques from the literature. A possibly different class label can be produced for each new gait signature, on the basis of L previous volumes.

To smooth and reinforce the results over time, a majority vote policy over a sliding temporal window of size W is used. As the gait signature information is computed on L previous volumes, the use of this window causes a delay of $L + W$ frames in obtaining the identity. Fig. 9 shows an example of majority voting system over a sliding temporal window, with $L = 5$ and $W = 3$.

4. Experiments and discussion

In this section we start by describing the used datasets, and then we present the experimental results.

4.1. Datasets description

To perform a 3D reconstruction by the SFS algorithm, the dataset must be multiview and calibration information have to be provided. Two datasets have been used to carry out the experiments, the “AVA Multi-View Dataset for Gait Recognition (AVAMVG)”² [40] and the “Kyushu University 4D Gait Database (KY4D)”³ [26].

In AVAMVG, 20 subjects perform 9 walking trajectories in an indoor environment. Each trajectory is recorded by 6 color cameras placed around a room that is crossed by the subjects during the performance, according to the distribution shown in the diagram of Fig. 10.

The video sequences of AVAMVG have a resolution of 640×480 pixels, and were recorded at a rate of 25 frames per second. For each actor, 9 gait sequences are captured in several trajectories as described in the figure by $\{t_1, \dots, t_9\}$. Of these trajectories, 3 are straight ($\{t_1, \dots, t_3\}$) and 6 are curved ($\{t_4, \dots, t_9\}$). An example of this dataset is shown in Fig. 11, in which several subjects are walking along different paths, from multiple viewpoints. Calibration parameters for the cameras of AVAMVG have been obtained with Aruco library [41].

With respect to KY4D Gait Database, it is composed of sequential 3D reconstructions and image sequences of 42 subjects walking along four straight and two curved trajectories. The sequences were recorded by 16 cameras, at a resolution of 1032×776 pixels. Although the KY4D Gait Database also provide sequential 3D reconstructions of subjects, we have reconstructed them with the same SFS method and resolution parameters used for reconstructing the 3D AVAMVG models.

As far as we know, there are other well-known multi-camera databases, as CMU Motion of Body (MoBo) Database [42] and CASIA Dataset B [43]. However, since these databases do not include information on camera parameters, 3D reconstructions of walking people cannot be obtained. Therefore, we did not use these databases in the experiments of the present study.

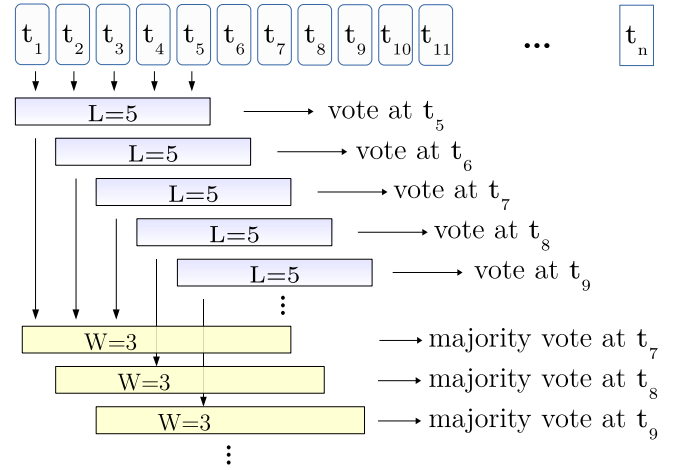


Fig. 9. Majority vote policy over a sliding temporal window. In this example, the size of the signature is set to $L = 5$, and the size of the voting window is set to $W = 3$.

4.2. Experimental results

In this section, we present the results of multiple experiments run on both gait datasets. First of all, we need to determine the value of several parameters of our method. Thus, considering the 3D reconstruction stage, the first relevant parameter is the voxel size. We consider that a voxel size of $0.27 \times 10^{-4} \text{m}^3$ is enough to get detailed 3D human reconstructions.

The average corporal volume for humans is $66.4L = 0.6640 \times 10^{-1} \text{m}^3$ measured by the water displacement method in 521 people aged 17 – 51 years [44]. Using a voxel size of $0.27 \times 10^{-4} \text{m}^3$, the number of voxels belonging to a person in the 3D volume should be

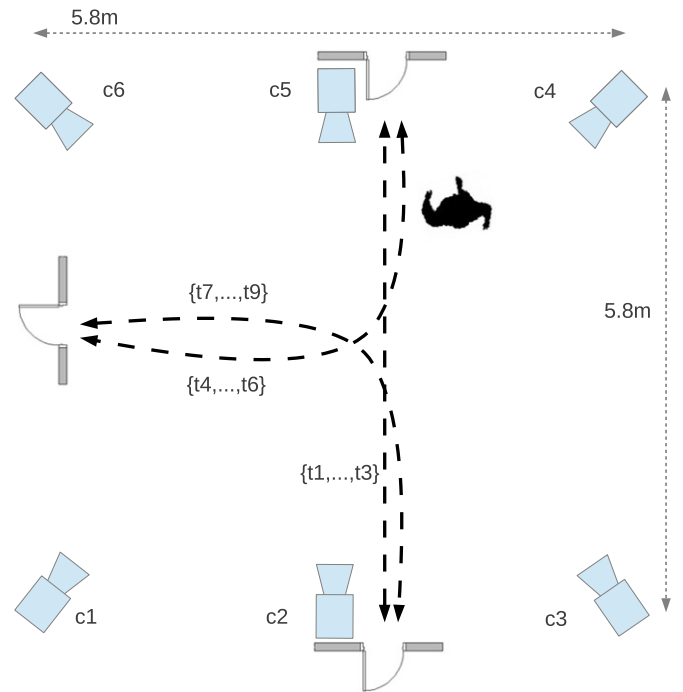


Fig. 10. Workspace setup used by AVAMVG Dataset, where $\{c_1, \dots, c_6\}$ represent the set of cameras of the multiview dataset and $\{t_1, \dots, t_9\}$ represent the different trajectories followed by each actor of the dataset.

² Publicly available at: <http://www.uco.es/investiga/grupos/ava/node/41>.

³ Publicly available at: <http://robotics.ait.kyushu-u.ac.jp/research-e.php?content=db>.



Fig. 11. Example of AVAMVG multiview dataset. People walking in different directions, from multiple points of view.

about 2459. With a value of $1 \times 10^3 < \phi < 2459$ (see Section 3.1.2) the system should be able to detect an adult human.

Regarding the volume alignment, Table 1 shows the mean error in the estimation angle for each trajectory of the KY4D dataset. We also report the 95% confidence interval on the mean, assuming that the data are normally distributed.

As it was proved in [18], the number of silhouettes (volumes in our case) aggregated in a single gait signature can be set in $L = 20$, because we have a rate of 25 volumes per second, and $L = 20$ responds to a signature of about 1 s which matches the length of a gait cycle.

The decision to build the gait signatures with joint or marginal distributions of histograms depends on the amount of training data and memory available for the classification process. If all the other parameters are kept unchanged, the use of joint distributions should lead to get better results. However, the dimensionality of the corresponding feature space using joint distributions is $M \times N \times L$ for CR [18], $M \times N \times 3 \times L$ for CRP, or $M \times N \times D \times L$ for CC, which is too expensive to compute. Alternatively, the dimensionality of the feature space using marginal distributions of histograms is $(M+N) \times L$ for CR, $(M+N) \times 3 \times L$ for CRP and $(M+N+D) \times L$ for CC. For example, for a value of $M = N = D = 25$ and $L = 20$, the feature space of the CC descriptor has a dimensionality of 312500 using joint distributions, compared to 1500 features using marginal distributions.

In several previous experiments, we noted that if we use joint distributions and SVM, the accuracy is lower than if we use marginal distributions. Maybe it can occur because the statistical significance of the joint distribution of histograms is much lower than the statistical significance of the marginal distributions of them. Moreover, may be impracticable to compute the joint distribution of histograms such as CC or CRP because of the high dimensionality. Therefore, we focus our experimentation on the use of marginal distributions of the histograms.

In order to determine M , N and D (the number of bins), we tested values ranging from 5 to 25, with step 5 on both AVAMVG and KY4D

Table 1
Mean error in the angle estimation (degrees) of the volume alignment step, for each trajectory of KY4D dataset. The 95% confidence interval on the mean is also shown.

Straight paths				Curved paths	
Tr. 1	Tr. 2	Tr. 3	Tr. 4	Tr. 5	Tr. 6
0.73 ± 0.04	0.76 ± 0.06	0.74 ± 0.05	0.76 ± 0.04	2.16 ± 0.34	2.15 ± 0.30

gait databases. It was observed that large values of M , N or D generally lead to better performance. However, the performance saturates with 20 bins and above, depending on the descriptor.

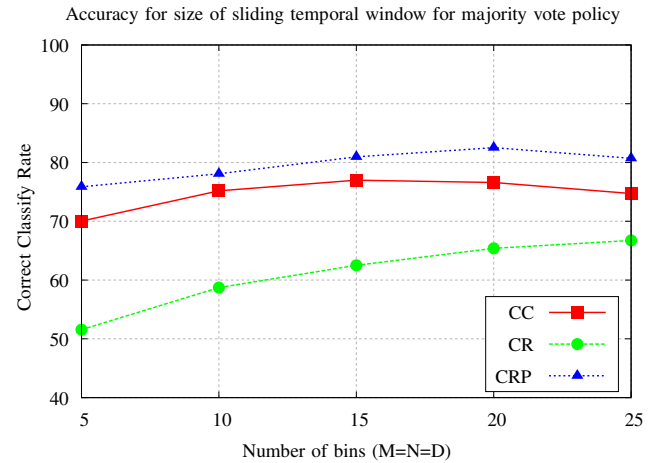


Fig. 12. Performance of each descriptor for different histogram sizes on AVAMVG gait database. The sliding temporal window for majority voting policy is disabled ($W = 1$).

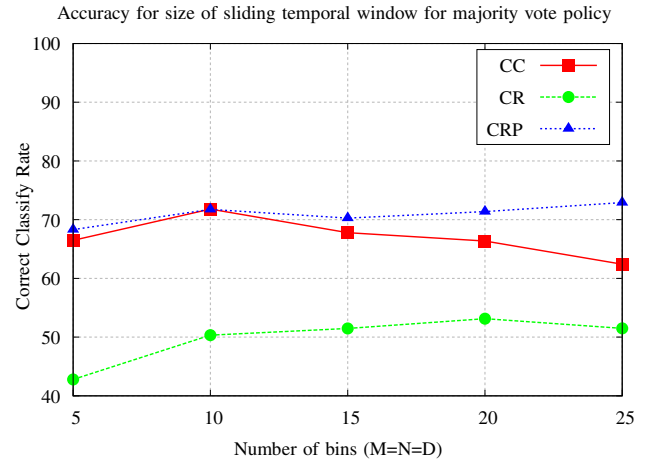


Fig. 13. Performance of each descriptor for different histogram sizes on KY4D gait database. The sliding temporal window for majority voting policy is disabled ($W = 1$).

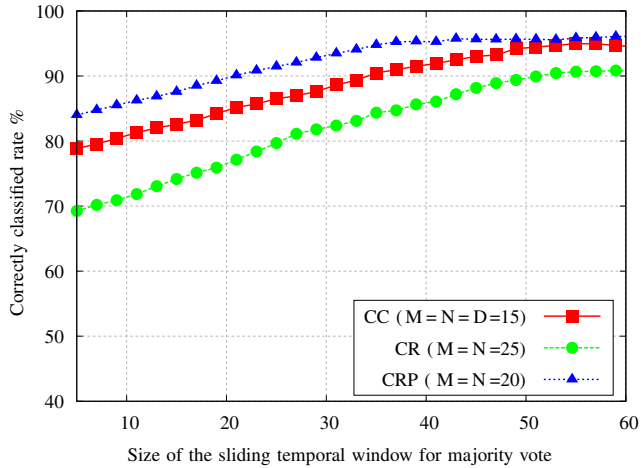


Fig. 14. Performance of each descriptor on AVAMVG database for different lengths for the majority voting window.

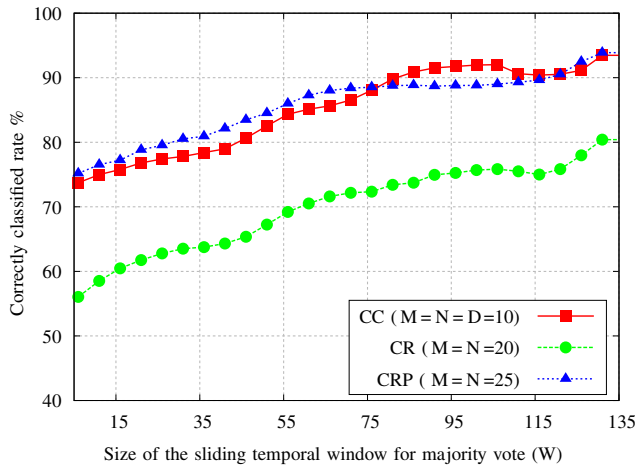


Fig. 15. Performance of each descriptor on KY4D database for different lengths for the majority voting window.

Depending on the size of the training dataset and the resolution of the 3D reconstructed volumes, the statistical significance of all the bins of the histograms needs to be taken into account. The first series of experiments consisted in determining the appropriate number of bins for each descriptor. For the sake of simplicity, we restricted to the case where $M=N$ or $M=N=D$, and disabled the majority vote on the previous W frames (or equivalently set W to 1).

We use a leave-one-out cross-validation strategy. Regarding experiments on the AVAMVG dataset, each fold is composed by a tuple formed by a set of 20 sequences (one sequence per actor) for

Table 3

Correct classification rate on KY4D. These results correspond to a cross-validation experiment. Each column corresponds to a test trajectory, using the remaining trajectories as training set. Each row corresponds to a different method. Each entry contains the percentage of correct recognition for each tuple trajectory-setup.

Method	Tr. 1	Tr. 2	Tr. 3	Tr. 4	Tr. 5	Tr. 6	Mean
CRP($M = N = 25$, $W = 135$)	92.6%	100%	100%	97.5%	84.9%	87.8%	93.8%
CC($M = N = D = 10$, $W = 135$)	97.5%	97.5%	95.1%	97.5%	82.9%	90%	93.4%
Seely et al. [21]	95.1%	100%	97.5%	100%	68.2%	72.5%	88.8%
Ariyanto and Nixon [22]	41.4%	41.4%	43.9%	53.6%	19.5%	17.5%	36.2%

testing, and by the remaining eight sequences of each actor for training, i.e. 8×20 sequences for training and 20 sequences for test. For the KY4D gait dataset, each fold is composed by 42 sequences for testing (one sequence per actor) and by the remaining five sequences of each actor (i.e. 42×5 sequences) for training.

We use a C-SVC SVM, which allows imperfect separation of classes with penalty multiplier C for outliers. Several SVM kernels were tested, and finally we selected Radial Basis Function since we obtained better results than with linear, polynomial, or sigmoid kernels. We set the same weight to all classes. To make the choice of SVM parameters independent of the sequence test data, we cross-validate the SVM parameters on the training set. We report the improvement with respect to different kernels in Table 4.

Fig. 12, shows the performance of each descriptor with $L = 20$ and different histogram sizes on the AVAMVG dataset. We use marginal distributions of histogram and we get the best results with $M = N = 25$ for the CR descriptor, $M = N = 20$ for the CRP descriptor, and $M = N = D = 15$ for the CC descriptor.

On the other hand, Fig. 13 shows the performance of each descriptor with different histogram sizes, applied on the KY4D gait dataset. We use marginal distributions, and in this case, we get the best results with $M = N = 20$ for the CR descriptor, $M = N = 25$ for the CRP descriptor, and $M = N = D = 10$ for the CC descriptor. In this experiment, for the sake of simplicity, we disabled the sliding temporal window for majority vote.

The second series of experiments that were carried out consisted in determining the optimum size of the sliding temporal window for majority voting. In Figs. 14 and 15 we show how the accuracy increases with respect to the size of the sliding temporal window for majority voting on both datasets, using the histogram sizes selected in the previous experiment.

The sliding temporal window of majority voting stage improves the performance of the method with any of the three proposed descriptors for both datasets. Nevertheless, the size of the sliding temporal window for voting is limited by the number of available gait signatures for each sequence.

For the AVAMVG, with CR descriptors applied on frontal volume projections, we obtain a maximum accuracy of 90.8%. Nevertheless, using the CC descriptor computed on the entire volume, the accuracy is about 94.5% and finally, with the CRP descriptor, the system was able to correctly identify up to 96.1% of subjects. On the other hand,

Table 2

Correct classification rate on AVAMVG. These results correspond to a cross-validation experiment. Each column corresponds to a test trajectory, using the remaining trajectories as training set. Each row corresponds to a different method. Each entry contains the percentage of correct recognition for each tuple trajectory-setup.

Method	Tr. 1	Tr. 2	Tr. 3	Tr. 4	Tr. 5	Tr. 6	Tr. 7	Tr. 8	Tr. 9	Mean
CRP($M = N = 20$, $W = 60$)	100%	88%	100%	99.3%	99.2%	97.7%	96.2%	84.8%	100%	96.1%
CC($M = N = D = 15$, $W = 60$)	100%	96%	75.5%	98.6%	87.8%	99.1%	99.5%	94%	100%	94.5%
Seely et al. [21]	90%	80%	94.7%	90%	60%	100%	80%	84.2%	90%	85.4%
Ariyanto and Nixon [22]	55%	45%	52.6%	45%	26.3%	35%	35%	31.5%	40%	40.6%

Table 4

Correct classification rate on KY4D with respect to different kernels on KY4D dataset.

Descriptor	Linear	Polynomial	Sigmoid	RBF
CR($M = N = 20, W = 135$)	78.3%	78.4%	78.0%	80.4%
CC($M = N = D = 10, W = 135$)	89.3%	91.8%	92.2%	93.4%
CRP($M = N = 25, W = 135$)	90.2%	91.5%	90.1%	93.8%

Table 5

Comparative of recognition results on AVAMVG dataset [40]. Each row corresponds to a different method. The second row indicates the training trajectory. The third and fourth columns indicate the tested trajectory. The fifth column shows the average value of the recognition rate for the curved paths.

Method	Training path	Curve t4	Curve t7	Mean
CRP($M = N = 20, W = 60$)	Straight {t1,t2,t3}	45%	81.5%	63.2%
Seely et al. [21]	Straight {t1,t2,t3}	55%	70%	62.5%
Iwashita et al. [26]	Straight {t1,t2,t3}	35.1%	37.7%	36.4%
Ariyanto and Nixon [22]	Straight {t1,t2,t3}	30%	30%	30%

Table 6

Comparative of recognition results on KY4D gait dataset [26]. Each row corresponds to a different method. The second column indicates the training trajectory. The third and fourth columns indicate the tested trajectory. The fifth column shows the average value of the recognition rate for the curved paths.

Method	Training trajectories	Curve 1	Curve 2	Mean
CRP($M = N = 25, W = 135$)	Straight {1,2,3,4}	60.4%	81.7%	71%
Iwashita et al. [26]	Straight {1,2,3,4}	61.9%	71.4%	66.6%
Seely et al. [21]	Straight {1,2,3,4}	19.5%	35%	27.2%
Ariyanto and Nixon [22]	Straight {1,2,3,4}	12.1%	15%	13.5%

on the KY4D dataset we obtain a maximum accuracy of 80.4% for CR, 93.4% for CC and 93.8% for CRP descriptors.

Tables 2 and 3 show detailed results for the leave-one-out experiment on AVAMVG and KY4D datasets respectively. We compare the accuracy of our approach for CRP and CC signatures with the accuracy of state-of-art methods such as [21,22] on AVAMVG and KY4D datasets. For the comparison with Ariyanto and Nixon we used the best kinematics features proved in [22], whereas for the case of Seely et al. [21] we have used the side-on, front-on, top-down average silhouettes. Since these methods are not designed to cope with curved trajectories, we have aligned the gait volumes along the path (see Section 3.1.2). The resolution of the reconstructed volumes was the same for all cases.

The third series of experiments that we performed consisted in testing our method, which allows completely free trajectories, with the method presented in [26], for identification of people walking along curved trajectories. Comparative of recognition results on both datasets are shown in Tables 5 and 6. In this experiment, we used the straight trajectories for training, and the curved trajectories for testing.

As we can observe in Tables 4 and 5, the performance of compared methods is dropped down when the training set does not contain curved trajectories. Moreover, we have noticed a decrease of performance of the method presented in [26] when it is trained with the straight paths and tested with the curves of AVAMVG. We think it may be due to the low number of cameras of AVAMVG and therefore to the quality of the 3D reconstructions. Besides that, in the AVAMVG dataset, depending on the viewpoint and performed trajectory, people appear at diverse scales, even showing partially occluded body parts.

We think that when the subject walks on a curved path, the gait pattern is consequently modified, as we can see in Fig. 16. For this reason, we consider that it is not entirely correct to train the classifier of our model-free approach with straight paths only.

5. Conclusions

This paper has proposed a method to recognize walking humans independently of the viewpoint and regardless direction changes. The method focuses on capturing 3D morphological and structural information from volumetric reconstructions of the gait. The main contribution is that the method achieves a good recognition rate on completely unconstrained paths, allowing direction changes, in contrast to others view-independent approaches where the view change is restricted to a few angles. In our method, the individual can walk freely in the scene without adversely affect to the recognition.

For this purpose, it was designed a mechanism of person detection and gait alignment based on 3D reconstructions. In order to extract information from the 3D volumes, three gait morphological descriptors are proposed. The first one is the Cover by Rectangles (CR) [18] applied on rendered front projections of the gait volume. The second is composed by an aggregation of three Cover by Rectangles descriptors computed on the top, side, and frontal projections of the gait volume (CRP). Lastly, the third new proposed descriptor is called Cover by Cubes (CC) and it is defined as the union of all the cubes with the largest size that can fit into a gait volume of a person.

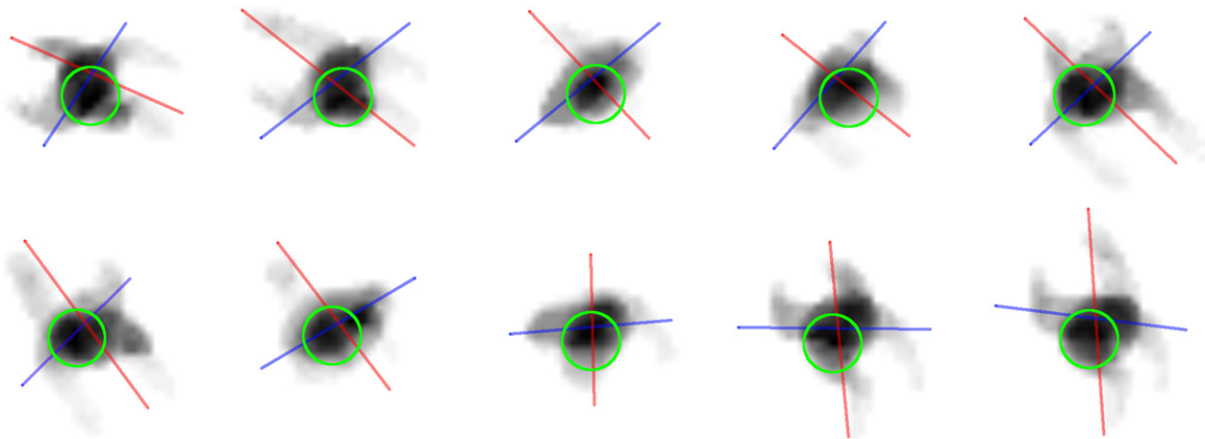


Fig. 16. Example of a curved gait cycle. We show several ground marginal distributions of occupied voxels. The velocity vector is represented by a red line, the blue line represents the torso main axis, and the position of the head is represented by a green circle. We can note that in a curved trajectory, the person rotates his/her torso and leans towards the walking direction.

The experimental results show that the CRP descriptor is the most reliable for using with our gait recognition method, providing good results in both AVAMVG and KY4D gait databases. The experimental results also show what is the optimal size for the histograms of each descriptor on each dataset. Finally, by using a majority vote policy on a sliding temporal window, the system is able to correctly identify up to 96% of the subjects of the AVAMVG gait database and nearly 94% of subjects of the KY4D dataset.

Acknowledgments

This work has been developed with the support of the Research Projects called TIN2012-32952 and BROCA both financed by Science and Technology Ministry of Spain and FEDER.

References

- [1] S. Yu, D. Tan, T. Tan, Modelling the effect of view angle variation on appearance-based gait recognition, in: P.J. Narayanan, S.K. Nayar, H.-Y. Shum (Eds.), *Computer Vision - ACCV 2006, Lecture Notes in Computer Science 3851*, Springer Berlin Heidelberg 2006, pp. 807–816. ISBN 978-3-540-31219-2. http://dx.doi.org/10.1007/11612032_81.
- [2] W. Hu, T. Tan, L. Wang, S. Maybank, A survey on visual surveillance of object motion and behaviors, *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.* 34 (3) (2004) 334–352. ISSN 1094-6977. ISSN. <http://dx.doi.org/10.1109/TSMCC.2004.829274>.
- [3] A. Kale, N. Cuntoor, B. Yegnanarayana, A.N. Rajagopalan, R. Chellappa, *Gait analysis for human identification, Audio- and Video-based Biometric Person Authentication 2688*, Springer Berlin Heidelberg 2003, pp. 706–714. ISBN 978-3-540-40302-9.
- [4] M.-H. Cheng, Meng-F. Ho, C.-L. Huang, Gait analysis for human identification through manifold learning and (HMM), *Pattern Recogn.* 41 (8) (2008) 2541–2553. ISSN 0031-3203. ISSN. <http://dx.doi.org/10.1016/j.patrec.2007.11.021>.
- [5] L. Wang, T. Tan, H. Ning, W. Hu, Silhouette analysis-based gait recognition for human identification, *IEEE Trans. Pattern Anal. Mach. Intell.* 25 (12) (2003) 1505–1518. ISSN 0162-8828. ISSN. <http://dx.doi.org/10.1109/TPAMI.2003.1251144>.
- [6] S. Hong, H. Lee, E. Kim, Automatic gait recognition using width vector mean, *Industrial Electronics and Applications*, 2009. ICIEA 2009. 4th IEEE Conference on, 2009, pp. 647–650. <http://dx.doi.org/10.1109/ICIEA.2009.5138285>.
- [7] S.D. Mowbray, M.S. Nixon, Automatic gait recognition via Fourier descriptors of deformable objects, in: J. Kittler, M.S. Nixon (Eds.), *Audio- and Video-based Biometric Person Authentication, Lecture Notes in Computer Science* 2688, Springer Berlin Heidelberg 2003, pp. 566–573. ISBN 978-3-540-40302-9. http://dx.doi.org/10.1007/3-540-44887-X_67.
- [8] C.P. Lee, A.W.C. Tan, S.C. Tan, Gait recognition via optimally interpolated deformable contours, *Pattern Recogn. Lett.* 34 (6) (2013) 663–669. ISSN 0167-8655. ISSN. <http://dx.doi.org/10.1016/j.patrec.2013.01.013>.
- [9] S.D. Choudhury, T. Tjahjadi, Silhouette-based gait recognition using Procrustes shape analysis and elliptic Fourier descriptors, *Pattern Recogn.* 45 (9) (2012) 3414–3426. ISSN 0031-3203. ISSN best Papers of Iberian Conference on Pattern Recognition and Image Analysis (IbPRIA'2011). <http://dx.doi.org/10.1016/j.patrec.2012.02.032>.
- [10] J. Han, B. Bhanu, Individual recognition using gait energy image, *IEEE Trans. Pattern Anal. Mach. Intell.* 28 (2) (2006) 316–322. ISSN 0162-8828. ISSN. <http://dx.doi.org/10.1109/TPAMI.2006.38>.
- [11] T.H.W. Lam, K.H. Cheung, J.N.K. Liu, Gait flow image: a silhouette-based gait representation for human identification, *Pattern Recogn.* 44 (4) (2011) 973–987. ISSN 0031-3203. ISSN. <http://dx.doi.org/10.1016/j.patrec.2010.10.011>.
- [12] C. Rougier, E. Auvinet, J. Meunier, M. Mignotte, J.A de Guise, Depth energy image for gait symmetry quantification, *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, 2011, pp. 5136–5139. ISSN 1557-170X. ISSN. <http://dx.doi.org/10.1109/IEMBS.2011.6091272>.
- [13] M. Hofmann, S. Bachmann, G. Rigoll, 2.5D gait biometrics using the depth gradient histogram energy image, *Biometrics: Theory, Applications and Systems (BTAS)*, 2012 IEEE Fifth International Conference on, 2012, pp. 399–403. <http://dx.doi.org/10.1109/BTAS.2012.6374606>.
- [14] P. Chattopadhyay, S. Sural, J. Mukherjee, Frontal gait recognition from occluded scenes, *Pattern Recogn. Lett.* 63 (2015) 9–15. ISSN 0167-8655. ISSN. <http://dx.doi.org/10.1016/j.patrec.2015.06.004>.
- [15] S. Sivapalan, D. Chen, S. Denman, S. Sridharan, C. Fookes, Gait energy volumes and frontal gait recognition using depth images, *Biometrics (IJCB)*, 2011 International Joint Conference on, 2011, pp. 1–6. <http://dx.doi.org/10.1109/IJCB.2011.6117504>.
- [16] S. Sivapalan, D. Chen, S. Denman, S. Sridharan, C. Fookes, The backfilled GEI—a cross-capture modality gait feature for frontal and side-view gait recognition, *Digital Image Computing Techniques and Applications (DICTA)*, 2012 International Conference on, 2012, pp. 1–8. <http://dx.doi.org/10.1109/DICTA.2012.6411694>.
- [17] J. Ryu, S. Kamata, Front view gait recognition using Spherical Space Model with Human Point Clouds, *Image Processing (ICIP)*, 2011 18th IEEE International Conference on, 2011, pp. 3209–3212. ISSN 1522-4880. ISSN. <http://dx.doi.org/10.1109/ICIP.2011.6116351>.
- [18] O. Barnich, Ma. Van Droogenbroeck, Frontal-view gait recognition by intra- and inter-frame rectangle size distribution, *Pattern Recogn. Lett.* 30 (10) (2009) 893–901. ISSN 0167-8655. ISSN. <http://dx.doi.org/10.1016/j.patrec.2009.03.014>.
- [19] G. Shakhnarovich, L. Lee, T. Darrell, Integrated face and gait recognition from multiple views, *Computer Vision and Pattern Recognition*, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on, 1, 2001, pp. 439–446. ISSN 1063-6919. ISSN. <http://dx.doi.org/10.1109/CVPR.2001.990508>.
- [20] A. Laurentini, The visual hull concept for silhouette-based image understanding, *IEEE Trans. Pattern Anal. Mach. Intell.* 16 (2) (1994) 150–162. ISSN 0162-8828. ISSN. <http://dx.doi.org/10.1109/34.273735>.
- [21] R.D. Seely, S. Samangoeei, M. Lee, J.N. Carter, M.S. Nixon, The University of Southampton multi-biometric tunnel and introducing a novel 3D gait dataset, *Biometrics: Theory, Applications and Systems*, 2008. BTAS 2008. 2nd IEEE International Conference on, 2008, pp. 1–6. <http://dx.doi.org/10.1109/BTAS.2008.4699353>.
- [22] G. Ariyanto, M.S. Nixon, Model-based 3D gait biometrics, *Biometrics (IJCB)*, 2011 International Joint Conference on, 2011, pp. 1–7. <http://dx.doi.org/10.1109/IJCB.2011.6117582>.
- [23] K. Yamauchi, B. Bhanu, H. Saito, Recognition of walking humans in 3D: initial results, *Computer Vision and Pattern Recognition Workshops*, 2009. CVPR Workshops 2009. IEEE Computer Society Conference on, 2009, pp. 45–52. ISSN 2160-7508. ISSN. <http://dx.doi.org/10.1109/CVPRW.2009.5204296>.
- [24] W. Kusunniran, Q. Wu, J. Zhang, H. Li, Gait recognition under various viewing angles based on correlated motion regression, *IEEE Trans. Circuits Syst. Video Technol.* 22 (6) (2012) 966–980. ISSN 1051-8215. ISSN. <http://dx.doi.org/10.1109/TCSVT.2012.2186744>.
- [25] R. Bodor, A. Drenner, D. Fehr, O. Masoud, N. Papanikolopoulos, View-independent human motion classification using image-based reconstruction, *Image Vis. Comput.* 27 (8) (2009) 1194–1206. ISSN 0262-8856. ISSN. <http://dx.doi.org/10.1016/j.imavis.2008.11.008>.
- [26] Y. Iwashita, K. Ogawara, R. Kurazume, Identification of people walking along curved trajectories, *Pattern Recogn. Lett.* 48 (2014) 60–69. ISSN 0167-8655. ISSN celebrating the life and work of Maria Petrou. <http://dx.doi.org/10.1016/j.patrec.2014.04.004>.
- [27] A. Kale, A. Sundaresan, A.N. Rajagopalan, N.P. Cuntoor, A.K. Roy-Chowdhury, V. Kruger, R. Chellappa, Identification of humans using gait, *IEEE Trans. Image Process.* 13 (9) (2004) 1163–1173. ISSN 1057-7149. ISSN. <http://dx.doi.org/10.1109/TIP.2004.832865>.
- [28] F. Jean, A.B. Albu, R. Bergevin, Towards view-invariant gait modeling: computing view-normalized body part trajectories, *Pattern Recogn.* 42 (11) (2009) 2936–2949. ISSN 0031-3203. ISSN. <http://dx.doi.org/10.1016/j.patrec.2009.05.006>.
- [29] S. Jeong, T.-h. Kim, J. Cho, Gait recognition using description of shape synthesized by planar homography, *J. Supercomput.* 65 (1) (2013) 122–135. ISSN 0920-8542. ISSN. <http://dx.doi.org/10.1007/s11227-013-0897-8>.
- [30] J. Han, B. Bhanu, A.K. Roy-Chowdhury, A study on view-insensitive gait recognition, *Image Processing*, 2005. ICIP 2005. IEEE International Conference on, 3, 2005. (III-297-300). <http://dx.doi.org/10.1109/ICIP.2005.1530387>.
- [31] M. Goffredo, I. Bouchrika, J.N. Carter, M.S. Nixon, Self-calibrating view-invariant gait biometrics, *IEEE Trans. Syst. Man Cybern. B Cybern.* 40 (4) (2010) 997–1008. ISSN 1083-4419. ISSN. <http://dx.doi.org/10.1109/TSMCB.2009.2031091>.
- [32] G. Zhao, G. Liu, H. Li, M. Pietikainen, 3D gait recognition using multiple cameras, Automatic face and gesture recognition, 2006. FGR 2006. 7th International Conference on, 2006, pp. 529–534. <http://dx.doi.org/10.1109/FGR.2006.2>.
- [33] D. Kastaniotis, I. Theodorakopoulos, C. Theoharatos, G. Economou, S. Fotopoulos, A framework for gait-based recognition using Kinect, *Pattern Recogn. Lett.* 68, Part 2 (2015) 327–335. ISSN 0167-8655. ISSN special Issue on “Soft Biometrics”. <http://dx.doi.org/10.1016/j.patrec.2015.06.020>.
- [34] N. Liu, Y.-P. Tan, View invariant gait recognition, *Acoustics Speech and Signal Processing (ICASSP)*, 2010 IEEE International Conference on, 2010, pp. 1410–1413. ISSN 1520-6149. ISSN. <http://dx.doi.org/10.1109/ICASSP.2010.5495466>.
- [35] Y. Makiyama, R. Sagawa, Y. Mukaigawa, T. Echigo, Y. Yagi, Gait recognition using a view transformation model in the frequency domain, in: A. Leonardis, H. Bischof, A. Pinz (Eds.), *Computer Vision – ECCV 2006, Lecture Notes in Computer Science 3953*, Springer Berlin Heidelberg 2006, pp. 151–163. ISBN 978-3-540-33836-9. http://dx.doi.org/10.1007/11744078_12.
- [36] W. Kusunniran, Q. Wu, H. Li, J. Zhang, Multiple views gait recognition using View Transformation Model based on optimized Gait Energy Image, *Computer Vision Workshops (ICCV Workshops)*, 2009 IEEE 12th International Conference on, 2009, pp. 1058–1064. <http://dx.doi.org/10.1109/ICCVW.2009.5457587>.
- [37] T. Horprasert, D. Harwood, L.S. Davis, A statistical approach for real-time robust background subtraction and shadow detection, *Proc. IEEE ICCV*, 1999, pp. 1–19.
- [38] L. Dáz-Más, R. Muñoz-Salinas, F.J. Madrid-Cuevas, R. Medina-Carnicer, Shape from silhouette using Dempster-Shafer theory, *Pattern Recogn.* 43 (6) (2010) 2119–2131. ISSN 0031-3203. ISSN. <http://dx.doi.org/10.1016/j.patrec.2010.01.001>.

- [39] C.J.C. Burges, A tutorial on support vector machines for pattern recognition, *Data Min. Knowl. Disc.* 2 (2) (1998) 121–167. ISSN 1384-5810. ISSN. <http://dx.doi.org/10.1023/A:1009715923555>.
- [40] D. López-Fernández, F.J. Madrid-Cuevas, A. Carmona-Poyato, M.J. Marín-Jiménez, R. Muñoz-Salinas, The AVA multi-view dataset for gait recognition, *Activity Monitoring by Multiple Distributed Sensing* Springer International Publishing 2014, pp. 26–39. ISBN 978-3-319-13322-5. http://dx.doi.org/10.1007/978-3-319-13323-2_3.
- [41] S. Garrido-Jurado, R. Muñoz-Salinas, F.J. Madrid-Cuevas, M.J. Marín-Jiménez, Automatic generation and detection of highly reliable fiducial markers under occlusion, *Pattern Recogn.* 47 (6) (2014) 2280–2292. ISSN 0031-3203. ISSN. <http://dx.doi.org/10.1016/j.patcog.2014.01.005>.
- [42] R. Gross, J. Shi, The CMU Motion of Body (MoBo) Database, Robotics Institute, Pittsburgh, PA, 2001.
- [43] S. Zheng, J. Zhang, K. Huang, R. He, T. Tan, Robust view transformation model for gait recognition, *Image Processing (ICIP)*, 2011 18th IEEE International Conference on, 2011, pp. 2073–2076. ISSN 1522-4880. ISSN. <http://dx.doi.org/10.1109/ICIP.2011.6115889>.
- [44] D.T. Drinkwater, W.D. Ross, *Anthropometric fractionation of body mass*, *Kinanthropometry*, 2, University Park Press, Baltimore, 1980, pp. 178–189.