

Repositório ISCTE-IUL

Deposited in *Repositório ISCTE-IUL*:

2018-07-12

Deposited version:

Post-print

Peer-review status of attached file:

Peer-reviewed

Citation for published item:

Verlekar, T. T., Soares, L. D. & Correia, P. L. (2018). Gait recognition in the wild using shadow silhouettes. *Image and Vision Computing*. 76, 1-13

Further information on publisher's website:

10.1016/j.imavis.2018.05.002

Publisher's copyright statement:

This is the peer reviewed version of the following article: Verlekar, T. T., Soares, L. D. & Correia, P. L. (2018). Gait recognition in the wild using shadow silhouettes. *Image and Vision Computing*. 76, 1-13, which has been published in final form at <https://dx.doi.org/10.1016/j.imavis.2018.05.002>. This article may be used for non-commercial purposes in accordance with the Publisher's Terms and Conditions for self-archiving.

Use policy

Creative Commons CC BY 4.0

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a link is made to the metadata record in the Repository
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Accepted Manuscript

Gait recognition in the wild using shadow silhouettes

Tanmay Tulsidas Verlekar, Luís Ducla Soares, Paulo Lobato
Correia



PII: S0262-8856(18)30067-2
DOI: [doi:10.1016/j.imavis.2018.05.002](https://doi.org/10.1016/j.imavis.2018.05.002)
Reference: IMAVIS 3691

To appear in:

Received date: 13 September 2017
Revised date: 31 January 2018
Accepted date: 10 May 2018

Please cite this article as: Tanmay Tulsidas Verlekar, Luís Ducla Soares, Paulo Lobato Correia , Gait recognition in the wild using shadow silhouettes. The address for the corresponding author was captured as affiliation for all authors. Please check if appropriate. *Imavis*(2018), doi:[10.1016/j.imavis.2018.05.002](https://doi.org/10.1016/j.imavis.2018.05.002)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Gait Recognition in the Wild using Shadow Silhouettes

Tanmay Tulsidas Verlekar^{1*}, Luís Ducla Soares², Paulo Lobato Correia¹

¹Instituto de Telecomunicações, Instituto Superior Técnico, Universidade de Lisboa, Portugal

²Instituto de Telecomunicações, Instituto Universitário de Lisboa (ISCTE-IUL), Portugal

*ttv@lx.it.pt

Abstract: *Gait recognition systems allow identification of users relying on features acquired from their body movement while walking. This paper discusses the main factors affecting the gait features that can be acquired from a 2D video sequence, proposing a taxonomy to classify them across four dimensions. It also explores the possibility of obtaining users' gait features from the shadow silhouettes by proposing a novel gait recognition system. The system includes novel methods for: (i) shadow segmentation, (ii) walking direction identification, and (iii) shadow silhouette rectification. The shadow segmentation is performed by fitting a line through the feet positions of the user obtained from the gait texture image (GTI). The direction of the fitted line is then used to identify the walking direction of the user. Finally, the shadow silhouettes thus obtained are rectified to compensate for the distortions and deformations resulting from the acquisition setup, using the proposed four-point correspondence method. The paper additionally presents a new database, consisting of 21 users moving along two walking directions, to test the proposed gait recognition system. Results show that the performance of the proposed system is equivalent to that of the state-of-the-art in a constrained setting, but performing equivalently well in the wild, where most state-of-the-art methods fail. The results also highlight the advantages of using rectified shadow silhouettes over body silhouettes under certain conditions.*

1. Introduction

Biometric traits such as iris, fingerprint or palmprint are widely used for user recognition as they provide a higher level of security when compared to passwords or key cards. However, these traits are mostly used in controlled environments, since they require active user cooperation. To employ biometric recognition in the wild, i.e., in less constrained or even unconstrained conditions, the used traits should:

- not require active user cooperation; and
- be collectable from a distance.

Among the currently used biometric traits satisfying the requirements above, gait, representing the static and dynamic aspects of a user's motion, is unique to a user, and collectable without active user cooperation, from a distance, even using low resolution images [1].

Gait recognition can be performed from data acquired using a wide range of devices, including body worn sensors, force plates on the floor, depth sensing cameras, and also conventional 2D video cameras. For operation in the wild, it can be difficult to setup complicated sensors on the user, and depth sensing cameras typically have a limited range of operation, thus making 2D cameras the more viable choice [2].

Most state-of-the-art image based gait recognition systems first employ a background subtraction algorithm [3] to separate the walking user from the background. The resulting body silhouette (foreground) is used to obtain features for recognition, such as gait energy image (GEI) [4], gait probability image [5] or silhouette contour [6]. In some conditions, also the silhouette of the shadows cast by the walking user on the ground can be used to characterize that users' gait and perform recognition [7].

1.1. Motivation

The exploitation of silhouettes obtained for gait recognition in the wild can be a challenging task, as the silhouettes used for representing the user's gait can be affected by several factors, notably related to the user, the camera characteristics, the light source and the environment. The state-of-the-art [16-43] addresses some of the related problems, such as the changes in the observed view of the user with respect to the camera, or changes in the user's appearance, for instance due to clothing. However, many other limitations of gait recognition, related to the factors listed above and combinations thereof can be identified. Therefore, this paper proposes a taxonomy discussing the factors affecting gait features' quality and how they can impair gait recognition. This taxonomy considers that gait recognition can exploit the user body silhouettes, but also silhouettes corresponding to the shadows cast by the user.

In addition, the paper proposes a novel gait recognition system, operating on shadow silhouettes. The use of shadow silhouettes can be advantageous in several scenarios, such as:

- When the video camera is mounted on an elevated position (e.g., placed on a lamp post, or carried by a drone, or placed near the ceiling inside a building), capturing an overhead view of the scene. Under such conditions, the user's body silhouette can be self-occluded, as illustrated in Fig. 1.a, b, while the shadow cast by the user is similar to the body silhouette and possibly less affected by occlusions.
- When gait features are acquired from the body silhouettes at different parts of a video sequence, as illustrated in Fig. 1.c, they can appear significantly different due to the problem of view change. However, under the same conditions, the features acquired from the shadow silhouettes often appear similar to each

other, allowing the use of the entire video sequence for gait recognition using shadows.

- When the video camera is mounted in a side view position, capturing both the body and the shadow cast by the user, as illustrated in Fig. 1.c. Under such conditions, features can be acquired from both the user's body and the shadow, and treated as two sources of information, in a multimodal system, which can improve recognition results.

To perform gait recognition using shadow silhouettes, the proposed system includes three main contributions: (i) a method to perform shadow segmentation, separating the shadow from the user body; (ii) a method to identify the user's walking direction, so that it can be successfully matched with a gait database which is sorted with respect to the walking directions; and (iii) a method to rectify the shadow silhouettes to compensate the distortions and deformations present in them. The method can be applied, before the matching step, to improve recognition results.

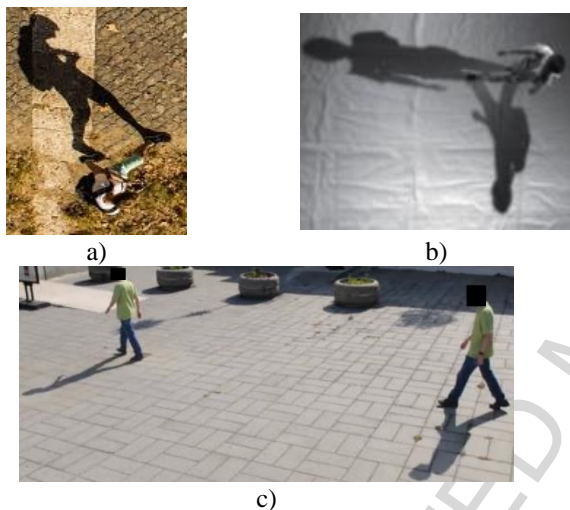


Fig. 1. Example of image captured by: a), b) overhead camera (from pixabay.com, [43]), c) side view camera.

1.2. State-Of-The-Art

A considerable amount of work has been reported on gait recognition, suggesting improvements on various components of the recognition system, such as feature representation, matching and decision, but also towards its operation in the wild, notably concerning the robustness to changes in view and/or appearance [1]. The methods employed for gait recognition can be broadly classified into: (i) **model based**, or (ii) **appearance based** methods.

Most model based methods rely on a 3D model describing the user's anatomy and/or kinematics. Recognition is performed using features obtained from the model, providing robustness to changes in view. Examples include methods that construct a 3D user model using static and dynamic features of a user's body obtained with multiple 2D cameras [8], or depth capturing cameras [3]. Multiple 2D cameras are also used in [9] to obtain visual hulls of a user to construct a 3D model that transforms the gallery gait features to match the probe view. A drawback of these methods is that they require multiple cameras or a depth camera, which typically has a limited range of operation. Also, model based methods assume that the view of the probe sequence is known. Although some methods

such as [10], [11] do identify the probe view using the feet positions of a user along time and synthesize silhouettes for the identified view using the user's 4D (3D + time) gait model, they require information of the scene being observed along with camera parameters, making them ineffective for recognition in the wild.

Other methods generate view invariant features using a canonical view model. For instance, hip, knee and ankle positions [12], [13] or head and feet positions [14], are obtained from a random view and converted to the canonical view using a rectification step, before attempting recognition. View rectification is performed over entire silhouettes using a perspective projection model and a calibrated camera in [15]. Transformation of gait features into a canonical view is effective within a small range of viewing angles, thus limiting the methods' applicability. Also, since these methods rely on identifying key joints or body parts, occlusions affect their performance. As a conclusion, model based methods typically lack robustness against appearance changes, which together with the other listed limitations makes them difficult to apply for recognition in the wild.

Appearance based methods rely on spatiotemporal information obtained from the input gait sequences. Most of these methods perform recognition without requiring additional information, such as camera parameters, position or other environment settings. Recognition often relies on features representing the walking user silhouette evolution [16], such as the GEI, averaging the silhouettes over a gait cycle [17], or the gait entropy image (GENI), computing the entropy over a GEI [18].

Appearance based methods can tackle view changes by either view transformation or view tagging. View transformation methods (VTM) reported in the literature [19-25] apply singular value decomposition (SVD) to a matrix containing gallery examples in different views, generating a set of gait feature vectors and a transformation matrix that can transform probe feature vectors from one view to another. The performance of these methods can be improved by using different feature representations, such as GEIs [19], or Radon transform-based energy images [20], as well as improved classification techniques, such as support vector machines [21], or multi-layer perceptron [22]. However, a limitation of these methods is that they tackle the problem of view change assuming that the probe view is known. This limitation can be dealt with using a view transformation model to generate feature vectors for virtual views, which are then projected onto a Grassmann manifold subspace [23], [24]. Alternatively, a given silhouette, independently of the view, can be transformed into a canonical view using a transformation matrix obtained by low-rank optimization of a GTI [25]. Nonetheless, in the current literature, methods that rely on correlation between views to recognize a user usually do not consider appearance changes.

View tagging methods, on the other hand, are robust to both changes in view and appearance, following a two-step approach. The first step performs view identification, typically by analysing the user's leg region [26], for instance using features computed using GENI [27], [28], perceptual hash [29], or the user's feet position [30]. The gallery is sorted according to the available views, so that in the second step recognition can be performed with respect to the

identified view. Robustness to appearance changes can be achieved for instance using features that highlight the silhouette areas unaltered by appearance changes, as the multiscale gait image (MGI) [27], or decomposing GEIs into sections and using a weighted score for recognition [31]. Other methods include using a binary mask [32], a weighting scheme [33] or a clothing model [34] to reduce the influence of appearance changes. Since these methods can tackle both view and appearance changes, they are good candidates for gait recognition in the wild.

The appearance based methods discussed above rely on body silhouettes to obtain gait features. However, gait features can also be obtained from shadow silhouettes under certain conditions [7]. Examples of such features include: harmonic coefficients obtained from gait stripes in [35], [36], the contour of the shadow silhouette [37], or both [38], as well as affine moment invariants (AMI) obtained from the shadow GEI [39]. View change is tackled in such conditions using a transformation method [40] that uses information such as the relative positions of the light source, user and the camera. Alternatively, silhouettes can be transformed into a canonical view by applying a transformation obtained by low-rank optimization of a shadow GTI, using robust principal component analysis (RPCA) – this approach is also called transform invariant low-rank textures (TILT) [41]. It can also be tackled using a 3D model [42], where a virtual shadow is synthesized using information from the scene, such as the position of the light source and the user. Robustness to changes in appearance is dealt with in [43] by using the method proposed in [33] over GEIs computed from the shadow silhouettes, considering sequences recorded in a setup with two infrared light sources to cast two shadows, perpendicular to each other, and captured using an overhead camera.

These methods, although effective, face several limitations often needing information about the acquisition setup, such as the position of the light source, and sometimes requiring manual gait period estimation and/or manual shadow segmentation, consider a fixed camera view, a fixed walking direction, and/or a fixed position of the light source. These limitations provide opportunities to further improve these methods to operate in the wild.

1.3. Contribution

The current literature lacks an encompassing discussion of the factors affecting gait recognition systems. Thus, a taxonomy discussing the factors affecting the features obtained to perform gait recognition is proposed in this paper.

A second contribution is a novel gait recognition system that acquires gait features from shadow silhouettes. The proposed system addresses several limitations of the

state-of-the-art, such as performing shadow segmentation in the wild, rectifying the shadow silhouettes and tackling view change. The proposed system performs shadow segmentation, separating the user and the shadow, by fitting a line through the feet positions of the user, which are computed using a GTI. Then shadow silhouettes are rectified using the proposed four-point correspondence method, to compensate for the distortions and deformations present in the shadow silhouettes. The rectification allows the features acquired along a video sequence, as illustrated in Fig.1.c, to be successfully matched with the database. According to the proposed taxonomy, view change happens when the position of the user with respect to the light source and the camera changes. The proposed system tackles the problem of view change by sorting the database with respect to the observed view. It then identifies the view of the probe using the walking direction of the user, which is estimated by computing the angle of the line used to perform shadow segmentation. So, the extracted features can be matched with respect to the identified view in the database. Thus, the proposed system contains three novel contributions:

- A method to perform automatic shadow segmentation;
- A method to perform walking direction identification;
- A method to transform user's shadow silhouettes into a canonical view.

This paper also presents a new database to test the performance of the proposed gait recognition system. Moreover, it allows comparison of results obtained from the body and the shadow silhouettes.

The rest of the paper is organized as follows. Section 2 discusses the factors affecting gait recognition. Section 3 presents the proposed gait shadow recognition system and section 4 discusses the experimental results, including a description of the new database. Section 5 provides conclusions and suggests directions for future work.

2. Factors Affecting Gait Recognition

The performance of an image-based gait recognition system can be affected by a large number of factors, which can, for instance, be perceived as view changes or appearance changes. If the influence of such factors is not appropriately understood and tackled, it may lead to poor recognition results. Thus, to better understand the relevant factors affecting gait recognition, this paper proposes a taxonomy, shown in Fig.2, which groups these factors into four main dimensions:

- User related factors;
- Camera related factors;
- Light source related factors; and
- Environment related factors.

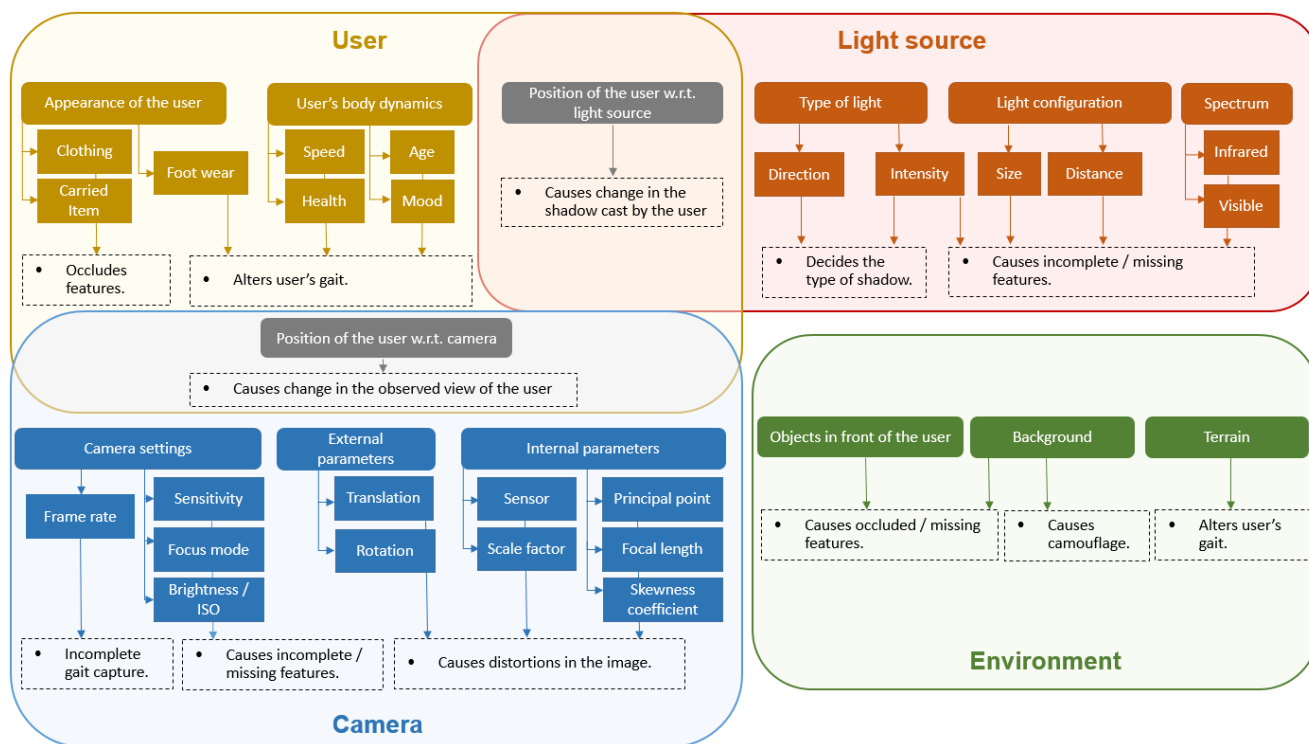


Fig. 2. Factors affecting image-based gait recognition.

Some of the considered factors affect the user's physical gait, while others affect the observation of gait and therefore the possibility to capture good features. Each of these dimensions, as well as some combinations of factors from different dimensions, are discussed in the following subsections. Section 2.1 includes a discussion on user related factors, clarifying that the problems in the observed gait caused by wearing a coat differ from those caused by wearing different footwear and, thus, should not be considered as a common problem of appearance change. Section 2.2 discusses how cameras cause distortion and deformation in the images captured by them. Section 2.3 discusses how the light source can affect the shadow silhouettes and the conditions needed to obtain sharp shadows, which are usable for gait recognition. Section 2.4 discusses other factors that can affect gait recognition, such as foreground occlusions. Finally, section 2.5 discusses problems affected by several factors, such as the perceived view change resulting from the change of user position with respect to the camera, clarifying the difference between existing works, such as those presented in [44] and [52]. It also addresses how the user's change of position with respect to the light source can affect shadow silhouettes, and discusses the advantages of using shadow silhouettes in an image-based gait recognition system.

2.1. User Related Factors

Gait recognition relies on features describing the user appearance and the corresponding motion pattern while walking. As such, any factors affecting the observable features or the user's gait should be taken into account.

One set of factors is related to the user's appearance. Some gait recognition methods assume that users' appearance will be similar along time [3-7]. However, appearance can be altered by clothing changes, e.g. when wearing a coat, a hat, or a long skirt, as well as by carrying

items, such as a bag or backpack. This can result in (partial) occlusion of the gait features useful for recognition purposes, as can be seen by comparing Fig.3.a and Fig.3.b. The problem of appearance changes has been addressed in the state-of-the-art [26-34], with some publicly available databases including test sequences allowing to test some of these conditions [44]. Appearance change can be considered as a problem of occlusion, where the gait of the user is often unaffected but some features are occluded by external objects worn or carried by the user.

A second set of factors affect the user's body dynamics, changing the gait itself. It includes factors such as speed, health, age and mood of the user. When the user increases the walking speed there is an increase in the arms swing and the stride length, as well as torso angle, as illustrated in Fig.3.c. Speed change is a covariate that has been studied in the literature, and some recognition solutions use features that remain unaltered with changes in speed [45], or use an ensemble of weak classifiers to solve the problem [46]. Some databases allow testing algorithms on a (limited) range of speeds [44] [47]. Larger speed variations and their effect on recognition requires further work, one possibility being to recognize the speed and comparing against a gallery with sequences of people walking at the identified speed.

User's health also affects the body dynamics. Injuries or other health problems may cause shorter or irregular strides, bending of the spine, or restricted limb movement, altering the user's gait, as illustrated in Fig.3.d. The health of the user also includes fitness, and weight variations. Age is another factor affecting gait; comparisons made between features acquired over longer periods of time are prone to be affected by changes in the user's height, weight, and muscular and bone development. The literature reports methods able to detect certain health issues, such as leg

injuries or a hunchback [48] [49], however, user recognition under such conditions is yet to be fully explored.

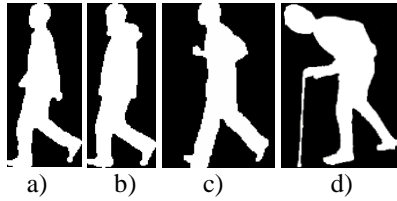


Fig. 3. Examples of silhouettes for: a) normal gait, b) partly occluded by a coat, c) altered by speed and d) affected by health.

Two other factors that can affect gait include the type of footwear, and the user's mood. The impact of using different footwear has been addressed in the literature, for instance devising feature representations that remain stable in different conditions, like MGI [27]. The influence of mood on gait has been discussed in medical literature, such as [50], however its effect on gait recognition is yet to be explored.

2.2. Camera Related Factors

This paper is concerned with gait recognition systems that use a video camera to capture gait information. Therefore, any image distortions and deformations resulting from the camera setup can be, typically described by its external and internal parameters [51], can affect the computation of gait features. Camera external parameters include rotation and translation matrices, denoting the transformations from the 3D world coordinates to the 3D camera coordinates. Camera internal parameters include focal length, image sensor format, principal point, scale factor and skewness coefficient which define the transformation of 3D camera coordinates into a 2D image (following a pinhole camera model). Thus, the combination of internal and external parameters describes the transformation of the 3D world coordinates into a 2D image. They also control the field of view, scale, skewness and resolution of the resulting image, with any distortions or deformations in the image affecting the gait features, as illustrated in Fig.4, and eventually leading to poor recognition results [41] [52].



Fig. 4. Examples of shadow silhouette deformations caused due to field of view, skewness and scale changes, as presented in [41].

Gait recognition relies on the successful capture of key instances of a gait cycle, such as heel strike, loading response, mid stance, etc. – see Fig.5. Since a normal walking gait cycle lasts for only 1 to 2 seconds, the camera must capture images at a sufficiently high frame rate to capture these key instances. Therefore, the acquisition frame rate of the camera is a determinant factor for the quality of gait features obtained [53].

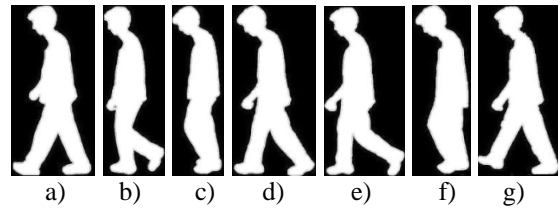


Fig. 5. Key instances of a gait cycle: a) heel strike, b) loading response, c) mid stance, d) terminal stance, e) pre swing, f) mid swing, g) terminal swing.

Other camera related factors are the sensitivity, focus mode and brightness/white control settings, which can help to better distinguish the user from the background, thus affecting the quality of silhouettes from which features are often extracted. Poor camera settings may lead to incomplete or missing features. As discussed in [54], changes in silhouette quality (i.e., incomplete features) can affect recognition results.

2.3. Light Source Related Factors

The third taxonomy dimension is related to the light source illuminating the scene. Light intensity is the main factor determining whether it is possible for the camera to “see” and therefore to allow identification of a walking user, or if the scene is poorly illuminated preventing the acquisition of usable features. Another important factor that characterizes the light source is its spectrum range. For instance, the works reported in [49] and [43] discuss the merits of using infrared light to perform gait recognition.

The scene illumination also determines whether there will be a shadow cast by the user which can be used for gait recognition purposes. The quality of the shadow silhouette depends primarily on the direction of the light rays emitted by the light source. When light rays travel in the same, well-defined direction, towards the user, i.e. under a collimate source of light, a sharp shadow is cast, displaying features similar to those that can be computed from the user's body silhouette – see Fig.6.a. If the light rays have many different directions, the resulting shadow will be diffused, appearing as a blob around the user with no distinguishing gait features – see Fig.6.b. Thus, a gait recognition system can only operate on sharp shadow silhouettes.

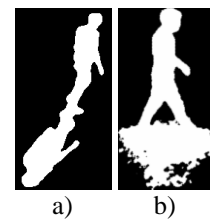


Fig. 6. Examples of a) sharp and b) diffused shadows cast by the walking user.

Other factors affecting the cast shadows are the distance and size of the light source. If the user is too close to the light source, an incomplete shadow will be produced, only including the body parts illuminated by it. While, if the user is too far away, sufficient intensity of light may not reach the user to cast a usable shadow. Finally, the illumination source size should be sufficiently large to uniformly illuminate the entire scene. However, if the light source size is large but not sufficiently far away then light rays from multiple directions might cause multiple

overlapping shadow contributions, resulting in a diffused shadow from which gait is not identifiable.

Typically, the literature addressing gait recognition using cast shadows considers recognition using a setup where all these factors are pre-determined [35-43]. To be able to address more real-life situations, the influence of variations in these factors needs to be studied.

2.4. Environment Related Factors

The environment, as observed by the camera, contains three main factors that can affect gait recognition: i) the objects in front of the user, ii) the background and iii) the terrain on which the user walks. The objects in front of the user can cause occlusions, which can lead to missing features. Background properties, such as colour and texture, if similar to the user's clothing can cause camouflage. In this case, distinguishing the user from the background becomes difficult, which can also lead to incomplete or missing features.

The terrain on which a user walks can cause changes in the user's gait. Those changes can be attributed to terrain properties, such as elevation, friction and irregularities. For example, a user has to put extra efforts to walk up a slope, when compared to a flat surface. These efforts, as well as terrain properties, alter the user's arm swing, stride length and the orientation of some body parts. Thus, the terrain can significantly affect the observed features, as discussed in [55].

2.5. Combination of Factors

The above taxonomy dimensions help to understand the factors affecting gait recognition. But, in the wild, some factors appear in combination across different dimensions, with further effects on the observed gait features. For instance, a camera captures only the part of the 3D world visible to it. Therefore, changes in the user position and walking direction relative to the camera, result in changes in the view of the user observed by the camera. Under such conditions, the features captured by the camera can change even when the physical gait of the user remains unchanged and un-occluded. For example, when a user walks towards the camera, the user's front view is observed, as illustrated in Fig.7.a, while if the user walks perpendicularly to the camera axis, the observed side view (see Fig.7.b) is significantly different, as discussed in [44]. Even under constrained conditions, when the user walks along a straight line across the (fixed) camera (see Fig.1.c), the captured view of the user at the start and at the end of the gait sequence will be significantly different, as illustrated in Fig.7.c and d, respectively, as discussed in [52]. Consequently, features computed for the same walking direction with a fixed camera may not match each other. The problem of view changes can be addressed using transformation or view tagging [19-30], but identifying users from a missing view is still a challenging task.

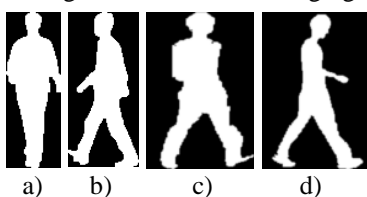


Fig. 7. Different views of the same user: a) frontal, b) lateral, c) view at the start of the sequence, d) view at the end.

Another relevant combination across taxonomy dimensions, especially when considering the recognition using shadows, is that of user and light source factors. This is especially relevant when there are considerable body self-occlusions while the cast shadow is clearly visible, as illustrated in Fig.1.a, for instance when a camera observes a user from a top view. In this case, the cast shadows depend on the combination of the user walking direction and the illumination direction, as illustrated in Fig.8.a, b.

Also, depending on the relative positions of the user, light source and camera, the user body may occlude the cast shadow, or the shadow may be cast by an overhead position of the light source, preventing acquisition of useful gait features. However, for an un-occluded shadow silhouette cast by a sideways position of the light source an interesting observation can be made. Under a collimated light source, when a user walks along a straight-line perpendicular to the camera axis, the shadows cast by the user along the trajectory will always appear similar to each other, since the user's position with respect to the light source does not change significantly. This is different from what happens for the body silhouettes (see Fig.7.c, d). As illustrated in Fig.8 the shadow silhouette Fig.8.b appears to be a skewed version of the shadow silhouette Fig.8.c. Also, since the cast shadow corresponds to an area darker than its surroundings, methods such as RPCA [3] can more easily distinguish the shadow from the background, than the user body.

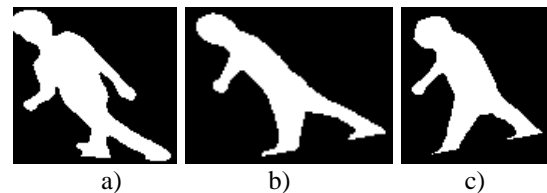


Fig. 8. Change in the shadow cast by the same user: a), b) when user's position w.r.t. the light source changes, and at the b) start and c) end of a linear walking sequence.

3. Proposed System

A number of factors can prevent successful acquisition and matching of body gait features, such as the observed view of the user, or the background properties, as discussed in section 2. Under such conditions, the shadow silhouette can be an alternative source of features, if the shadow cast by the user is sharp and un-occluded, as illustrated in Fig.6.a. Thus, several methods addressing gait recognition using shadows have been proposed in the literature, although not exactly considering operation in the wild. The problems caused by factors such as camera parameters, and changes in the cast shadow due to change in the position of the user with respect to the light source, are yet to be effectively addressed.

This paper proposes a system, suitable for operation in the wild, as it is able to tackle the problem of changes in shadows' appearance caused by the different positions of the user with respect to the light source and camera. The system also performs automatic shadow segmentation and silhouette rectification, compensating for image distortions due to the camera. Using the proposed system, features

obtained at any part of the gait sequence can be matched with the database, which otherwise can be a challenging task for body and shadow silhouettes, as discussed in section 2.

3.1. Shadow Segmentation Method

The video sequences captured by the camera contain the user, the corresponding shadow, and the background. To isolate the moving areas, a background subtraction method can be employed, such as RPCA [3], to produce foreground silhouettes containing the body and the shadow of the user, with both silhouettes connecting at the feet position. Once the foreground silhouette is available, the body and shadow silhouettes can be segmented (i.e., separated).

The shadow segmentation problem is discussed in [35], where an automatic shadow segmentation method is proposed, which performs extremely well under a constrained setting. It computes the sum of intensities along the first principal component direction over a GTI, identifying the point of separation as the one with the highest summation value. The GTI is constructed by averaging the K images containing the silhouettes $I(x, y, t)$, according to (1).

$$GTI(x, y) = \sum_{t=1}^K I(x, y, t) / K \quad (1)$$

However, the method presented in [35] is limited in its use in the wild as the principal component computed over the GTI does not always align itself with the walking direction (see Fig.9.a), and, when aligned, the highest summed intensity does not always correspond to the user's feet position (see Fig.9.b).

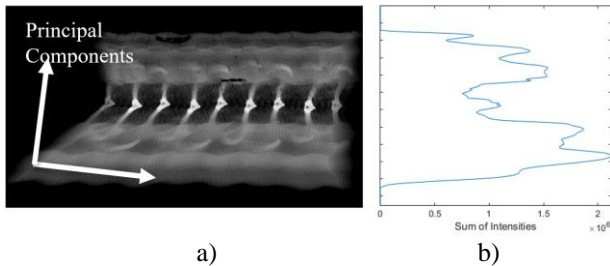


Fig. 9. Illustration of the drawbacks of the shadow segmentation method in [35]: a) principal component misalignment; b) sum of intensities along x-axis.

Nevertheless, user's feet positions in the GTI are represented by the higher intensity values, as seen in Fig.9.a. Therefore, the proposed method uses a threshold to identify those feet positions. The threshold is selected as 80% (empirically determined) of the highest intensity pixel value in the image, as it usually correspond to the feet positions of the user. The method then fits a line through the centroids of the various feet positions to separate the body (above) from the shadow silhouettes (below the line), as illustrated in Fig.10. Since under certain conditions, some of the high intensity values may correspond to the arm position, due to the overlapping silhouettes in the GTI, the proposed method applies a random sample consensus (RANSAC) line fitting algorithm [56]. RANSAC classifies the centroids as inliers and outliers, fitting the line only using the inliers, making the method robust to the falsely identified feet positions. The method starts by considering a line model using two

random centroids, followed by selecting a set of inliers for the given model. The process is repeated until a maximum number of centroids agree on the selected model.

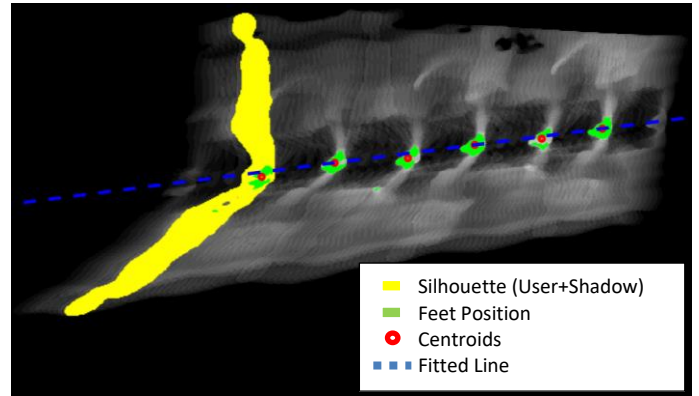


Fig. 10. Proposed shadow segmentation method.

3.2. Walking Direction Identification Method

The line separating the shadow and body silhouettes highlights the walking direction of the user. This way, given a scenario where the user's view is unknown, the fitted line representing the walking direction can be used to identify the captured view, as discussed in [30].

The state-of-the-art [26-30] relies on the bottom third of the body silhouette to perform walking direction identification and hence requires a good shadow segmentation method. Since the proposed shadow segmentation also allows identifying the user's walking direction, it eliminates the need of an additional step. With the proposed method, walking direction identification can be done by computing the angle between the fitted line and the coordinate axis. The walking direction thus obtained corresponds to a unique observed view of the user. Recognition can then be performed by selecting a subset of the database, identified by the walking direction.

3.3. Shadow Silhouette Rectification Method

The images acquired by the camera contain distortions and deformations due to the camera setting, as discussed in section 2, which also affect the shadow silhouettes obtained from such images. To perform recognition, the distortions and deformations need to be compensated for, by rectifying the shadow silhouettes. The proposed method rectifies the shadow silhouettes by transforming them into a common canonical view, using a homographic projective transformation. The transformation parameters cannot be determined directly from the image, requiring the usage of a method like TILT [41], to perform a low-rank optimization of the GTI. However, TILT being an optimization method cannot be fully controlled and thus, it may retain some deformations, while also distorting the aspect ratio of the shadow silhouettes.

The proposed shadow silhouette transformation method, here denoted as "four-point correspondence method", estimates the transformation parameters using head and feet positions of shadow silhouettes at the beginning and end of the available gait sequence. Since along a gait cycle, the best estimates for the head and feet positions are obtained during the mid-stance/mid-swing, where arms and feet of the user are closer to the body, the proposed method selects shadow silhouettes from the first

mid-stance and last mid-swing phases available. The mid-stance phase is determined by analysing the normalized aspect ratio, obtained by subtracting the mean and dividing by the standard deviation of the cropped shadow silhouettes along the gait sequence. Notice that the mid-stance/mid-swing correspond to the lowest normalized aspect ratio values within each gait cycle.

Once both the first and the last shadow silhouettes in the mid-stance/mid-swing phase are identified, the proposed method applies principal component analysis (PCA) to estimate the orientation of the selected shadow silhouettes, followed by the computation of their convex hulls and centroids. Finally, the method passes a line through each selected shadow silhouettes' centroid, with their first principal component's orientation, intersecting the convex hull at the head and feet positions of the respective shadow silhouettes, as illustrated in Fig. 11. Thus, the method obtains four points corresponding to the head and feet position of the shadow silhouettes belonging to the first mid-stance and the last mid-swing phase in the observed view.

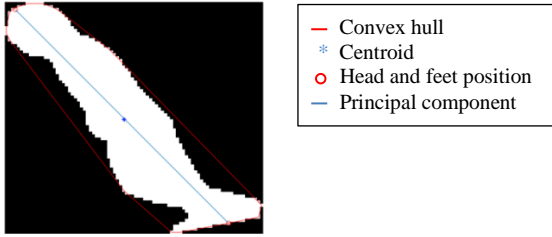


Fig. 11. Head and feet positions of the shadow silhouette.

The proposed method then estimates the location of the four corresponding points in the canonical view, which allows to establish a one-to-one correspondence between the observed and the canonical views according to equation (2), and as illustrated in Fig. 12.

$$\begin{aligned} x_l &= x_1, x_r = x_1 + \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}, \\ y_t &= (y_1 + y_2)/2, y_b = (y_3 + y_4)/2. \end{aligned} \quad (2)$$

where the points representing the head and the feet positions of the user at the first mid-stance phase are (x_1, y_1) , (x_3, y_3) and at the last mid-stance phase are (x_2, y_2) , (x_4, y_4) , respectively. Their mappings onto a canonical view are (x_l, y_t) , (x_l, y_b) , (x_r, y_t) and (x_r, y_b) , respectively.

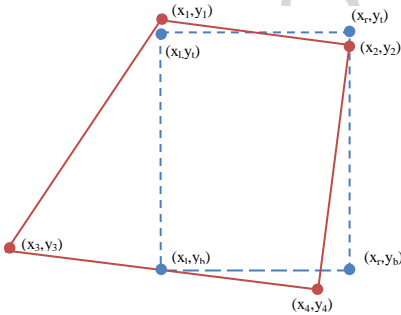


Fig. 12. The head and feet positions in the observed view (red) and their corresponding mapping in the canonical view (blue).

Using the estimated head and feet positions and their mappings in the canonical view, the proposed method estimates the parameters for the homographic projective transformation as a matrix H , consisting of three sub-

matrices T , R and P , [51], where T represents translation, R represents rotation, scaling and shear, while P represents perspective transformations, according to (3).

$$H = \begin{bmatrix} R & P \\ T & 1 \end{bmatrix} \quad (3)$$

where $R = \begin{bmatrix} r_{00} & r_{01} \\ r_{10} & r_{11} \end{bmatrix}$, $T = [t_x \ t_y]$, $P = \begin{bmatrix} p_x \\ p_y \end{bmatrix}$

The parameters of T , R and P are obtained from (4), assuming (x_l, y_t) to be the coordinate axis origin.

$$P = \begin{bmatrix} 1 & \frac{|x_1 - x_2 + x_4 - x_3 \ x_3 - x_4|}{|y_1 - y_2 + y_4 - y_3 \ y_3 - y_4|} \\ x_r - 1 & \frac{|x_2 - x_4 \ x_3 - x_4|}{|y_2 - y_4 \ y_3 - y_4|} \\ 1 & \frac{|x_2 - x_4 \ x_1 - x_2 + x_4 - x_3|}{|y_2 - y_4 \ y_1 - y_2 + y_4 - y_3|} \\ y_b - 1 & \frac{|x_2 - x_4 \ x_3 - x_4|}{|y_2 - y_4 \ y_3 - y_4|} \end{bmatrix}$$

$$T = [x_1 \ y_1] \quad (4)$$

$$R = \begin{bmatrix} \frac{x_2 - x_1}{x_r - 1} + p_x x_2 & \frac{x_3 - x_1}{y_b - 1} + p_y x_3 \\ \frac{y_2 - y_1}{x_r - 1} + p_x & \frac{y_3 - y_1}{y_b - 1} + p_y \end{bmatrix}$$

Using the estimated parameters, a given point (x_{iv}, y_{iv}) belonging to a shadow silhouette in the observed view is transformed into a point (x_{cv}, y_{cv}) in the canonical view, according to (5).

$$\begin{aligned} x_{cv} &= \frac{r_{00}x_{iv} + r_{10}y_{iv} + t_x}{p_x x_{iv} + p_y y_{iv} + 1} \\ y_{cv} &= \frac{r_{01}x_{iv} + r_{11}y_{iv} + t_y}{p_x x_{iv} + p_y y_{iv} + 1} \end{aligned} \quad (5)$$

The proposed method, transforms the available shadow silhouettes into a canonical view, correcting for skewness, scale, orientation and other distortions and deformations. This transformation allows the successful matching of shadow silhouettes captured at any part of the gait sequence, which would otherwise be a challenging task.

3.4. User Recognition

To obtain features for user recognition, the proposed system constructs a GEI from the available rectified shadow silhouettes. The GEI is obtained by averaging the N available cropped shadow silhouettes $Ic(x, y, t)$ belonging to a gait cycle, according to (6). The use of a GEI allows the proposed method to minimize the impact of residual distortions and deformations in the shadow silhouettes by highlighting the dynamic part of the user's gait.

$$GEI = \sum_{t=1}^N Ic(x, y, t)/N \quad (6)$$

The performance of a recognition system can be improved by either improving the quality of the features or by using better classification tools. To emphasize the quality of the features obtained from the rectified shadow silhouettes, the proposed system employs a simple classifier. However, before applying the classifier, the proposed

system performs dimensionality reduction and data decorrelation using PCA and LDA. PCA rearranges the data along the principal components such that the first principal component corresponds to the dimension with highest variance, thus allowing the system to select only the components with significant variation. The system then uses LDA to identify a projection matrix Φ onto a subspace that maximizes the ratio of intra- to inter-class scatter, using Fisher's criterion. Thus, given k classes with their respective centroids \bar{x}_k , recognition of a probe GEI z , can be performed by computing the Euclidean distance $d(\cdot)$ between the centroids and the probe in the transformed space according to (7).

$$\arg \min_k d(z\Phi, \bar{x}_k\Phi) \quad (7)$$

4. Experimental Results

The proposed system is tested using two different databases. The first database, published by the Kyushu University, Japan, called KU IR shadow database [43], is used to compare the proposed system with the state-of-the-art. The database contains 54 users with 6 sequences for each user. The six sequences consist of 4 normal gait sequences. The remaining 2 sequences are altered by carrying a bag, and changing clothes respectively. The database is captured in a constrained setting with a lighting setup such that the user casts two shadows, as illustrated in Fig.1.b. Sequences are captured from an overhead view, resulting in self-occluded body silhouettes, for a duration of approximately one gait cycle.

The proposed system addresses several limitations of the state-of-the-art such as view change and rectification of shadow silhouettes. It also explores the possible advantages of using rectified shadow silhouettes over body silhouettes. Thus, the database should contain un-occluded body and shadow silhouettes, captured along different views. The KU IR shadow database is inadequate when considering such conditions. Thus, a new database is presented in this paper, containing both body and shadow silhouettes captured in two different views. The new gait database will be made publicly available, so that other methods can be tested using it.

4.1. Shadow Gait Database

The new database called "IST gait database" includes information computed for 21 users. The database was collected outdoors, with the setup illustrated in Fig.13.a, where each user walks along two directions: BA and BC, as illustrated in Fig.13.b. Video acquisition is performed using a Nikon D300S camera, at 24fps with a spatial resolution of 640×424 pixels. The data was collected in July 2017, between 5:30 and 6:30 pm, in the campus of Instituto Superior Técnico, Lisbon, Portugal. Each user is recorded on 2 different days (sessions). Each user walks 3 times in each direction during each session, amounting to 12 gait sequences, each one including at least 3 complete gait cycles.

To test the system using the database, the following protocol is proposed. The system is first trained using the sequences belonging to the first session and tests use the sequences of the second session. Next, the system is trained using the sequences belonging to the second session and tested against the sequences of the first session. Finally, the mean of the two results is presented as the correct user

recognition rate and the standard deviation is presented as a measure of the confidence value.

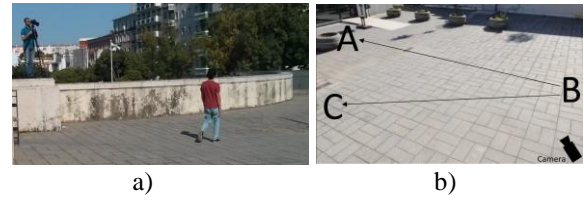


Fig. 13. Acquisition setup a) camera position b) walking directions.

4.2. User Recognition using Shadow Silhouettes

The first test is conducted using the KU IR database presented in [43]. However, since the proposed system in its current form is not robust to appearance change, only the first 4 normal gait sequences are used for testing. Also, unlike the method presented in [43] that uses both shadows and body silhouettes, the proposed system is tested only on (lateral view) shadow silhouettes. The recognition results are obtained following the 4-fold cross validation protocol presented in [43]. Following the protocol, three normal gait sequences are used for training and one normal gait sequence is used for testing. The test is repeated until all possible combinations are explored.

Among the methods presented in the state-of-the-art, the ones presented in [35], [36] perform poorly as the considered feature "gait stripe", which is computed as the maximum width of the shadow silhouette, does not seem to work well with larger databases, while the methods such as [39], [41] and [43] that use features such as entire shadow silhouettes and shadow GEIs perform extremely well. The proposed system performs equally well under such conditions with a correct recognition rate of $97 \pm 2\%$. However, it should be noted that under these testing conditions even the conventional shadow GEI provides good recognition results, as reported in Table 1. The good results can be attributed to the training sequences which contain the same type of distortions and deformations as observed in the testing sequences. The state-of-the-art methods such as [39] and [43] rely on the same assumption, as illustrated in Fig.14.a, b. They further improve the recognition results by employing sophisticated classifiers. The method presented in [41] rectifies the shadow silhouettes, compensating the distortions and deformations present in the shadow silhouettes. However, it can sometimes retain some residual deformations, as illustrated in Fig.14.c, which end up affecting its recognition results. The proposed four-point correspondence method always results in rectification of the shadow silhouettes into a canonical view, as illustrated in Fig.14.d; its performance does not depend on previous knowledge of distortions and deformations in the database, leading to good recognition results even with a simple classifier, as reported in Table 1.

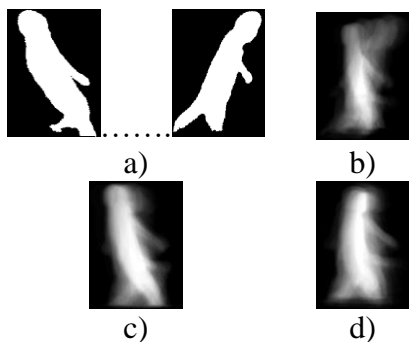


Fig. 14. Features used by the state-of-the-art: a) shadow silhouettes by method [43], b) GEI by method [39], c) rectified GEI by method [41], d) rectified GEI by the proposed method.

Table 1: Correct recognition rate with viewpoint changes

Methods	Correct Recognition Rate (%)
Proposed four-point correspondence + Shadow GEI	97
Shadow silhouettes + voting + K-divisions + AMI method [43]	97
Shadow GEI + K-divisions + AMI method [39]	94
TILT + Shadow GEI [41]	94
Shadow GEI	92

4.3. Walking Direction Identification

While performing shadow segmentation, the line separating body and shadow silhouettes also provides the walking direction of the user. Thus, to observe the performance of the proposed walking direction identification method, it is tested using all the available gait sequences from the IST gait database, as the proposed method does not require a training step. The walking direction is classified as BC if the angle obtained by the proposed method is approximately 0° , else it is classified as BA. Since the classification is binary and the separation between the two walking directions is significantly large, the proposed method performs 100% correct walking direction identification with this database.

To observe the influence of the proposed walking direction identification method on the recognition results, a test is conducted where the proposed method performs user recognition without walking direction identification, using both body and shadow silhouettes, following the proposed protocol on the IST gait database. The results of the test are reported in the first and third rows of Table 2. Next, the database is sorted with respect to the walking direction following the approach presented in [26-30]. The proposed system then identifies the walking direction of the probe and performs recognition with respect to the identified walking direction, following the proposed protocol on the IST gait database. The method thus limits the recognition process

only to a subset of the database identified by the walking direction. This improves the recognition results of both body and shadow GEIs, as reported in the second and forth rows of Table 2 respectively.

It should be noted that in the current tests, the database considers only two walking directions. However, the use of the proposed walking direction identification method improves the recognition results by 5-8%, for GEIs obtained from both body and rectified shadow silhouettes. The gain in the performance of the system can be expected to increase as the number of different walking directions registered in the database increases.

Table 2: Correct recognition rate with and without walking direction identification.

Methods	Correct Recognition Rate (%)
Body GEI	64±1
Walking direction identification + Body GEI	69±6
Proposed four-point correspondence method + Shadow GEI	81±3
Walking direction identification + Proposed four-point correspondence method + Shadow GEI	89±4

4.4. Shadow Silhouettes vs. Body Silhouettes

It can be observed from Table 2 that, in the given setup, user recognition results using body silhouettes are inferior to the ones obtained using rectified shadow silhouettes. The difference in performance can be attributed to the change in the observed view of the user which affects the body silhouettes, as discussed in section 2.5. The shadow silhouettes, under such conditions, remain relatively unaffected. However, they are affected by distortion and deformations caused by the camera. The proposed four-point correspondence method rectifies the shadow silhouettes leading to improved recognition results.

To highlight the significance of the proposed four-point correspondence method and the advantage of using shadow silhouettes over body silhouettes, a test is conducted on the IST gait database where the three gait cycles obtained along each walking direction are used to obtain three GEI groups: 1 (start), 2 (middle) and 3 (end). Recognition can then be performed across groups, following the proposed protocol.

Table 3: Correct recognition rate using body silhouettes [%].

Train/Test	Group 1		Group 2		Group 3	
	BA	BC	BA	BC	BA	BC
Group 1	60±1	62±4	27±0	39±3	05±2	19±4
Group 2	31±1	33±2	58±5	68±0	29±1	40±1
Group 3	08±7	09±3	28±3	27±2	65±9	73±0

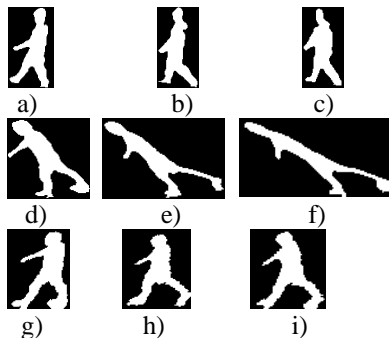
Table 4: Correct recognition rate using shadow silhouettes [%].

Train/Test	Group 1		Group 2		Group 3	
	BA	BC	BA	BC	BA	BC
Group 1	86±1	84±6	43±6	50±1	10±2	11±2
Group 2	21±6	13±1	73±6	85±1	43±1	34±4
Group 3	10±0	06±8	32±9	30±2	73±4	77±9

Table 5: Correct recognition rate using rectified shadow silhouettes using the proposed method [%].

Train/Test	Group 1		Group 2		Group 3	
	BA	BC	BA	BC	BA	BC
Group 1	73±2	80±2	64±6	86±4	45±5	79±1
Group 2	71±1	77±5	81±0	89±4	70±5	91±0
Group 3	53±1	65±7	74±3	80±7	81±4	90±1

In Tables 3, 4 and 5, each entry represents the correct recognition rate along the two considered walking directions: BA and BC. In all cases, entries along the diagonal represent the best recognition results, as the testing and the training sequences are obtained from the same part of the gait sequence (start, middle or end). Most state-of-the-art methods rely on such conditions to provide their results. However, recognition performance when using either body or shadow silhouettes deteriorates when training and testing is performed using silhouettes from different parts of the gait sequence. The deterioration is caused because the body silhouettes undergo a view change as the user proceeds along a walking direction, as illustrated in Fig.15.a, b and c. In the case of shadow silhouettes, since the position of the light source (Sun) with respect to the user along a walking direction remains the same, the shadow silhouettes can be expected to remain almost unchanged, apart from the camera distortions and deformations – see Fig.15.d, e, and f. However, such distortions and deformations lead to poor recognition results, as reported in Table 4. The problem is addressed by rectifying the shadow silhouettes using the proposed four-point correspondence method – Fig.15.g, h, and i. The significance of the proposed system is highlighted in Tables 3 and 4, where the mean recognition rate using the body and shadow silhouette GEIs is 38% and 43%, respectively. Under the same conditions, the proposed four-point correspondence method performs significantly better, by rectifying the shadow silhouettes, achieving a more consistent performance across all parts of the gait sequence, with a mean recognition rate of 75%. Also, given training silhouettes belonging to Group 2 (i.e., the middle part of the gait sequence), only the proposed system performs gait recognition consistently with a mean recognition rate of 80% – see Table 5.

**Fig. 15.** Silhouettes at the start, middle and end of a gait sequence. a), b), c) body silhouettes; d), e), f) shadow silhouettes and g), h), i) rectified shadow silhouettes.

5. Conclusion and Future Work

Gait recognition systems rely on successful acquisition of gait features to perform user recognition. However, the acquisition of features can be hindered by various factors. This paper presents a novel taxonomy, grouping the identified factors across four different dimensions: user, camera, light source and environment. It discusses the influence of factors belonging to each taxonomy dimension, as well as combinations of factors across dimensions, on the acquired gait features.

The paper considers the use of shadow silhouettes to perform gait recognition, which can be advantageous under scenarios where, for example, the acquisition of body features is impaired. To effectively use shadow silhouettes under such conditions, the paper presents a novel method to segment the shadow silhouettes, by fitting a line through the user's feet positions identified in a GTI. The fitted line can also be used to identify the walking direction of the user, which allows the use of a database sorted with respect to the walking directions, thus improving recognition results. Since segmented shadow silhouettes are usually affected by distortions and deformations due to the camera, the paper also presents a novel method, the four-point correspondence method, to rectify the shadow silhouettes using a homographic projective transformation. The method estimates the transformation parameters using the head and feet positions of the user at the start and the end of the gait sequence and its mapping onto a canonical view, allowing shadow silhouettes obtained at different parts of the gait sequences to be successfully matched with the database. Thus, the proposed system performs equivalently to the state-of-the-art, in a constrained setting and outperforms them in the wild.

The proposed system currently addresses the problems caused by change in walking direction of the user with respect to the light source. But, as discussed in section 2, several other factors can affect the performance of a gait system, such as change in appearance or body dynamics of the user. However, unlike body silhouettes, the use of shadow silhouettes in recognition systems introduces additional problems, such as a possible change in the available gait features caused by changes in the position of the light source. The influence of such problems on the proposed system is yet to be studied. Thus, future work will include testing the system under a wider range of illumination conditions. Also, alternative feature representations for shadow silhouettes and sophisticated classification tools to perform gait recognition, including deep learning techniques, will be investigated. The new database presented in this paper consists of data from only 21 users walking along two directions. The database will be extended to include more users, walking along more directions, and with different positions of the light source (by capturing sequences at different hours of the day).

Acknowledgement

This work was partially supported by Instituto de Telecomunicações under Fundação para a Ciência e Tecnologia Grant UID/EEA/50008/2013.

References

- [1] Makihara, Y., Matovski, D., Nixon, S., Carter, J., Yagi, Y., "Gait recognition: Databases, representations, and applications," Wiley Encyclopedia of Electrical and Electronics Engineering, 2015.
- [2] Gafurov, D., "A survey of biometric gait recognition: Approaches, security and challenges," in Annual Norwegian computer science conference, 2007.
- [3] Rodriguez, P., Wohlberg, B., "Fast principal component pursuit via alternating minimization," in IEEE International Conference on Image Processing, 2013.
- [4] Fazenda, J., Santos, D., Correia, P., "Using Gait to Recognize People," in The International Conference on Computer as a Tool, EUROCON, Belgrade, 2005.
- [5] Choudhury, D., Tjahjadi, T., "Silhouette-Based Gait Recognition using Procrustes Shape Analysis and Elliptic Fourier Descriptors," Pattern Recognition, vol. 45, no. 9, pp. 3414–3426, 2012.
- [6] Lee, C.P., Tan, A.W., Tan, S.C., "Gait Probability Image: An Information-Theoretic Model of Gait Representation," Journal of Visual Communication and Image Representation, vol. 25, no. 6, pp. 1489–1492, 2014.
- [7] Stoica, A., "Towards Recognition of Humans and their Behaviors from Space and Airborne Platforms: Extracting the Information in the Dynamics of Human Shadows," Bio-inspired, Learning and Intelligent Systems for Security, Edinburgh, 2008.
- [8] Zhao, G., Liu, G., Li, H., Pietikainen, M., "3D Gait Recognition Using Multiple Cameras," in 7th International Conference on Automatic Face and Gesture Recognition, Southampton, 2006.
- [9] Muramatsu, D., Shiraishi, A., Makihara, Y., Yagi, Y., "Arbitrary view transformation model for gait person authentication," in Biometrics: Theory, Applications and Systems (BTAS), IEEE Fifth International Conference on, 2012.
- [10] Iwashita, Y., Kurazume, R., Baba R., Ogawara, k., "Method for Gait-Based Biometric Identification Robust to Changes in Observation Angle," in 26th International Conference of Image and Vision Computing, New Zealand, Auckland, 2011.
- [11] Iwashita, Y., Kurazume, R., Baba R., Ogawara, k., "Identification of people walking along curved trajectories," Pattern Recognition Letters, vol. 48, pp. 60-69, 2014.
- [12] Bouchrika, I., Carter, J.N., Nixon, M.S., "Towards automated visual surveillance using gait for identity recognition and tracking across multiple non-intersecting cameras," Multimedia Tools and Applications, vol. 75, no. 2, pp. 1201-1221, 2016.
- [13] Goffredo, M., Bouchrika, I., Carter, J.N., Nixon, M.S., "Self-Calibrating View-Invariant Gait Biometrics," IEEE Transactions on Systems, Man, and Cybernetics, vol. 40, no. 4, pp. 997 - 1008, 2010.
- [14] Jeana, F., Albus, A.B., Bergevin, R., "Towards View-Invariant Gait Modeling: Computing View-Normalized Body," Pattern Recognition, vol. 42, no. 11, pp. 2936-2949, 2009.
- [15] Kale, A., Chowdhury, A.R., Chellappa, R., "Towards a view invariant gait recognition algorithm," in IEEE Conference on Advanced Video and Signal Based Surveillance, 2003.
- [16] Lv, Z., Xing, X., Wang, K., Guan, D., "Class energy image analysis for video sensor-based gait recognition: a review," Sensors, vol. 15, pp. 932-964, 2015.
- [17] Man, J., Bir, B., "Individual recognition using gait energy image," IEEE transactions on pattern analysis and machine intelligence, vol. 28, no. 2, pp. 316-322, 2006.
- [18] Bashir, K., Tao, X., Shaogang, G., "Gait Recognition Using Gait Entropy Image," in 3rd International Conference on Crime Detection and Prevention (ICDP), 2009.
- [19] Kusakunniran, W., Wu, Q., Li, H., Zhang, J., "Multiple Views Gait Recognition using View Transformation Model Based on Optimized Gait Energy Image," in IEEE 12th International Conference on Computer Vision Workshops (ICCV Workshops), Kyoto, 2009.
- [20] Liu, N., Tan, Y. P., "View Invariant Gait Recognition," in IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP), Dallas, TX, 2010.
- [21] Kusakunniran, W., Wu, Q., Zhang, J. Li, H., "Support vector regression for multi-view gait recognition based on local motion feature selection," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2010.
- [22] Kusakunniran, W., Wu, Q., Zhang, J., Li, H., "Cross-view and multi-view gait recognitions based on view transformation model using multi-layer perceptron," Pattern Recognition Letters, vol. 7, no. 33, pp. 882-889, 2012.
- [23] Connie, T., Goh, M.K.O., Teoh, A.B.J., "A Grassmannian Approach to Address View Change Problem in Gait Recognition," IEEE transactions on cybernetics, 2016.
- [24] Connie, T., Goh, M.K.O., Teoh, A.B.J., "Multi-view gait recognition using a doubly-kernel approach on the Grassmann manifold," Neurocomputing, vol. 216, pp. 534-542., 2016.
- [25] Kusakunniran, W., Wu, Q., Zhang, J., Ma, Y., Li, H., "A new view-invariant feature for cross-view gait recognition," Information Forensics and Security, IEEE Transactions, vol. 8, no. 10, pp. 1642-1653, 2013.
- [26] Khalid, B., Tao, X., Shaogang, G., "Cross View Gait Recognition using Correlation strength," in British

- Machine Vision Conference, London, 2010.
- [27] Choudhury, D., Tjahjadi, T., “Robust View Invariant Multiscale Gait Recognition,” *Pattern Recognition*, vol. 48, no. 3, pp. 798-811, 2015.
- [28] Guan, Y., Li, C.T., Hu, Y., “An Adaptive System for Gait Recognition in Multi-View Environments,” in *14th ACM Multimedia and Security Workshop*, Coventry, UK, 2012.
- [29] Verlekar, T., Correia, P., “Walking Direction Identification Using Perceptual Hashing,” in *International Workshop on Biometrics and Forensics - IWBIF*, Limassol, Cyprus, 2016.
- [30] Verlekar, T., Correia, P., Soares, L., “View-Invariant Gait Recognition Exploiting Spatio-Temporal Information and a Dissimilarity Metric,” in *In Biometrics Special Interest Group (BIOSIG)*, 2016 International Conference of the IEEE, 2016.
- [31] Verlekar, T., Correia, P., Soares, L., “View-Invariant Gait Recognition System Using a Gait Energy Image Decomposition Method,” *IET Biometrics*, 2017.
- [32] Bashir, K., Tao, X., Shaogang, G., “Gait recognition without subject cooperation,” *Pattern Recognition Letters*, vol. 31, no. 13, pp. 2052-2060, 2010.
- [33] Iwashita, Y., Uchino, K., Kurazume, R., “Gait-Based Person Identification Robust to Changes in Appearance,” *Sensors - Open Access Journal*, vol. 13, no. 6, pp. 7884-7901, 2013.
- [34] Liang, Y., Li, C.T., Guan, Y., Hu, H., “Gait recognition based on the golden ratio,” *Journal on Image and Video Processing*, no. 22, 2016.
- [35] Iwashita, Y., Stoica, A., Kurazume, R., “Person Identification using Shadow Analysis,” in *British Machine Vision Conference, BMVC 2010*, Aberystwyth, 2010.
- [36] Iwashita, Y., Stoica, A., “Gait Recognition Using Shadow Analysis,” in *Symposium on Bio-inspired Learning and Intelligent Systems for Security*, Edinburgh, 2009.
- [37] Iwashita, Y., Stoica, A., Kurazume, “Finding people by their shadows: Aerial surveillance using body biometrics extracted from ground video,” in *Third International Conference on Emerging Security Technologies (EST)*, 2012.
- [38] Iwashita, Y., Stoica, A., Kurazume, “Gait identification using shadow biometrics,” *Pattern Recognition Letters*, vol. 33, no. 16, p. 2148-2155, 2012.
- [39] Iwashita, Y., Stoica, A., Kurazume Y., Stoica, A., Uchino, K., “Gait identification from invisible shadows,” in *Sensing Technologies for Global Health, Military Medicine, Disaster Response, and Environmental Monitoring II; and Biometric Technology for Human Identification IX*, Baltimore, Maryland, 2012.
- [40] Iwashita, Y., Stoica, A., Kurazume Y., Stoica, A., “People identification using shadow dynamics,” in *Image Processing (ICIP)*, 2010 17th IEEE International Conference on, Hong Kong, 2010.
- [41] Verlekar, T., Correia, P., Soares, L., “Shadows, Gait Recognition Using Normalized,” in *25th European Signal Processing Conf. - EUSIPCO*, Kos Island, Greece, 2017.
- [42] Shinzaki, M., Iwashita, Y., Kurazume, R., Ogawara, K., “Gait-Based Person Identification Method Using Shadow Biometrics for Robustness to Changes in the Walking Direction,” in *Conference on Applications of Computer Vision (WACV)*, 2015 IEEE Winter , Waikoloa, HI, 2015.
- [43] Iwashita, Y., Stoica, A., Kurazume, “Gait Identification Using Invisible Shadows: Robustness to Appearance Changes,” in *Emerging Security Technologies (EST)*, 2014 Fifth International Conference on, Alcalá de Henares, 2014.
- [44] “CASIA Gait Database,” <http://www.sinobiometrics.com>.
- [45] Xu, C., Makihara, Y., Li, X., Yagi, Y., Lu, J., “Speed Invariance vs. Stability: Cross-Speed Gait Recognition Using Single-Support Gait Energy Image,” *ACCV*, vol. 2, pp. 52-67, 2016.
- [46] Guan, Y., Li, C.T., “A robust speed-invariant gait recognition system for walker and runner identification,” in *Biometrics (ICB)*, 2013 International Conference on IEEE., 2013.
- [47] Makihara, Y., Mannami, H., Tsuji, A., Hossain, M.A., Sugiura, K., Mori, A., Yagi, Y., “The OU-ISIR Gait Database Comprising the Treadmill Dataset,” *IPSN Trans. on Computer Vision and Applications*, vol. 4, pp. 53-62, 2012.
- [48] Nieto-Hidalgo, M., Ferrández-Pastor, F.J., Valdivieso-Sarabia, R.J., Mora-Pascual, J., García-Chamizo, J.M., “A vision based proposal for classification of normal and abnormal gait using RGB camera,” *Journal of biomedical informatics*, vol. 63, pp. 82-89, 2016.
- [49] Nguyen, T.N., Huynh, H.H. and Meunier, J., “Extracting silhouette-based characteristics for human gait analysis using one camera,” *Proceedings of the Fifth Symposium on Information and Communication Technology*. ACM, 2014.
- [50] Michalak, J., Troje, N.F., Fischer, J., Vollmar, P., Heidenreich, T., Schulte, D., “Embodiment of sadness and depression—gait patterns associated with dysphoric mood,” *Psychosomatic medicine*, vol. 71, no. 5, pp. 580-587, 2009.
- [51] Hartley, R., Zisserman, A., “*Multiple View Geometry in Computer Vision*,” Cambridge University Press., 2003, p. 155-157.
- [52] Iwama, H., Okumura, M., Makihara, Y., Yagi, Y., “The OU-ISIR Gait Database Comprising the Large Population Dataset and Performance Evaluation of Gait Recognition,” *IEEE Trans. on Information Forensics and Security*, vol. 7, no. 5, 2012.
- [53] Guan, Y., Sun, Y., Li, C.T., Tistarelli, M., “Human gait identification from extremely low-quality videos: an enhanced classifier ensemble method,” *IET biometrics*, vol. 3, no. 2, pp. 84-93, 2014.
- [54] Jia, N., Sanchez, V., Li, C.T., Mansour, H., “On

Reducing the Effect of Silhouette Quality on Individual Gait Recognition: a Feature Fusion Approach,” in Biometrics Special Interest Group (BIOSIG), 2015 International Conference of the, 2015.

- [55] Sarkar, S., Phillips, P.J., Liu, Z., Vega, I.R., Grother, P., Bowyer, K.W., “The humanid gait challenge problem: Data sets, performance, and analysis,” IEEE transactions on pattern analysis and machine intelligence, vol. 27, no. 2, pp. 162-177, 2005.
- [56] Fischler, M.A., Bolles, R.C., “Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography,” Communications of the ACM, vol. 24, no. 6, pp. 381-395, 1981.

ACCEPTED MANUSCRIPT

Highlights

The main contributions of this paper are:

- A taxonomy discussing the factors affecting the features in gait recognition;
- A method to perform automatic shadow segmentation;
- A method to perform walking direction identification, and tackle view change;
- A rectification method to transform gait shadow silhouettes into a canonical view;
- A new database to test the performance of the proposed shadow gait recognition system.

ACCEPTED MANUSCRIPT