Image and Vision Computing 76 (2018) 1–13

Contents lists available at ScienceDirect

Image and Vision Computing

journal homepage: www.elsevier.com/locate/imavis

Gait recognition in the wild using shadow silhouettes☆

Tanmay Tulsidas Verlekar [a,*], Luís Ducla Soares [b], Paulo Lobato Correia [a]

[a] Instituto de Telecomunicações, Instituto Superior Técnico, Universidade de Lisboa, Portugal
[b] Instituto de Telecomunicações, Instituto Universitário de Lisboa (ISCTE-IUL), Portugal

ARTICLE INFO

Article history:
Received 13 September 2017
Received in revised form 31 January 2018
Accepted 10 May 2018
Available online 17 May 2018

Keywords:
Shadow biometrics
Gait recognition
Biometric recognition
View invariant

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 and 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

☆ This paper has been recommended for acceptance by Vitomir Štruc.

* Corresponding author.
E-mail address: ttv@lx.it.pt (T.T. Verlekar).

https://doi.org/10.1016/j.imavis.2018.05.002
0262-8856/

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. 1a, 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. 1c, 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. 1c. 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.

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.

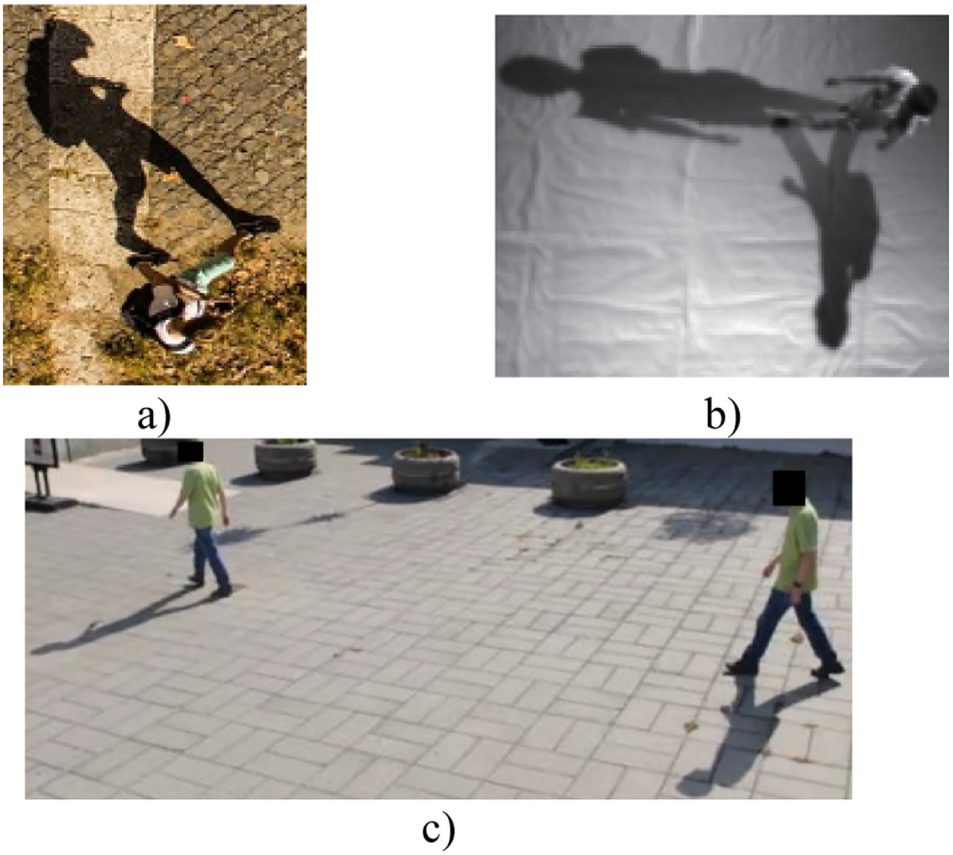

Fig. 1. Example of image captured by: a), b) overhead camera (from pixabay.com, [43]), c) side view camera.