

The AVA Multi-View Dataset for Gait Recognition

D. López-Fernández, F.J. Madrid-Cuevas, A. Carmona-Poyato, M.J. Marín-Jiménez, and R. Muñoz-Salinas

Computing and Numerical Analysis Department. Campus de Rabanales. Maimónides Institute for Biomedical Research (IMIBIC). University of Córdoba, 14071, Córdoba, Spain. Email: {i52lofed, fjmadrid, malcapoa, mjmarin, rmsalinas}@uco.es

Abstract. In this paper, we introduce a new multi-view dataset for gait recognition. The dataset was recorded in an indoor scenario, using six convergent cameras setup to produce multi-view videos, where each video depicts a walking human. Each sequence contains at least 3 complete gait cycles. The dataset contains videos of 20 walking persons with a large variety of body size, who walk along straight and curved paths. The multi-view videos have been processed to produce foreground silhouettes. To validate our dataset, we have extended some appearance-based 2D gait recognition methods to work with 3D data, obtaining very encouraging results. The dataset, as well as camera calibration information, is freely available for research purposes.

1 Introduction

Research on human gait as a biometric for identification has received a lot of attention due to the apparent advantage that it can be applied discreetly on the observed individuals without needing their cooperation. Because of this, the automation of video surveillance is one of the most active topics in Computer Vision. Some of the interesting applications are, among others, access control, human-machine interface, crowd flux statistics, or detection of anomalous behaviours [1].

Most current gait recognition methods require gait sequences captured from a single view, namely, from the side view or from the front view of a walking person [2–9]. Hence, there are many existing databases which capture the gait sequences from a single view. However, new challenges in the topic of gait recognition, such as achieving the independence from the camera point of view, usually require multi-view datasets. In fact, articles related to multi-view and cross-view gait recognition have been increasingly published [10–16].

First, some of the existing multi-view current datasets were recorded in controlled conditions, and in some cases, they made use of a treadmill [17–19]. An inherent problem associated with walking on a treadmill is that the human gait is not as natural as it should be, the gait speed is usually constant, and the subjects cannot turn right or left. They are not representative of human gait in a real world. Secondly, for other multi-view datasets, calibration information

is not provided, e.g. [20]. Some of the gait recognition methodologies require camera calibration to deal with 3D information.

In addition, there are not many multi-view datasets specifically designed for gait. Some of them are designed for action recognition, and therefore they do not contain gait sequences of enough length as to contain several gait cycles, because gait is a subset of them.

For this reason, we have created a new indoor dataset to test gait recognition algorithms. This dataset can be applied in workspaces where subjects cannot show the face or use the fingerprint, and even they have to wear special clothing, e.g. a laboratory. Furthermore, in this dataset people appear walking along both straight and curved paths, which makes this dataset suitable to test methods like [10]. The cameras have been calibrated and the methods based on 3D information can use this dataset to test. The dataset is free only for research purposes.

This paper is organized as follows. Section 2 describes current datasets for gait recognition. Section 3 describes the AVA Multi-View Dataset for Gait Recognition (AVAMVG). Section 4 shows several application examples carried out to validate our database. Finally, we conclude this paper in Section 5.

2 Current datasets for gait recognition

The chronologic order of appearance of the different human action video datasets runs parallel to the challenges that the scientific community has been considering to face the problem of automatic and visual gait recognition. From this point of view, datasets can be divided into two groups: datasets for single-view gait recognition and datasets for multi-view gait recognition. Besides, the datasets can be divided in two subcategories: indoor and outdoor datasets.

Regarding single view datasets, indoor gait sequences are provided by the OU-ISIR Biometric Database [17]. OU-ISIR database is composed by four treadmill datasets, called A, B, C and D. Dataset A is composed of gait sequences of 34 subjects from side view with speed variation. The dataset B is composed of gait sequences of 68 subjects from side view with clothes variation up to 32 combinations. OU-ISIR gait dataset D contains 370 gait sequences of 185 subjects observed from the lateral view. The dataset D focuses on the gait fluctuations over a number of periods. The OU-ISIR gait dataset C is currently under preparation, and as far as we know, information about this has not been released yet.

In contrast with OU-ISIR, the first available outdoor single-view database was from the Visual Computing Group of the UCSD (University of California, San Diego) [21]. The UCSD gait database includes six subjects with seven image sequences of each, from the side view. In addition to UCSD gait database, one of the most used outdoor datasets for single-view gait recognition is the USF HumanID database [22]. This database consists of 122 persons walking in elliptical paths in front of the camera.

Other outdoor walking sequences are provided in CASIA database, from the Center for Biometrics and Security Research of the Institute of Automation of

the Chinese Academy of Sciences. CASIA Gait Database is composed by three datasets, one indoor and the other two outdoor. The indoor dataset can be also considered as a multi-view dataset, and therefore it will be discussed later. The outdoors datasets are named as Dataset A and Dataset C (infrared dataset), and they are described below.

Dataset A [23] includes 20 persons. Each person has 12 image sequences, 4 sequences for each of the three directions, i.e. parallel, 45 degrees and 90 degrees with respect to the image plane. The length of each sequence is not identical due to the variation of the walker's speed. Dataset C [24] was collected by an infrared (thermal) camera. It contains 153 subjects and takes into account four walking conditions: normal walking, slow walking, fast walking, and normal walking with a bag. The videos were all captured at night.

More outdoor gait sequences are also found in the HID-UMD database [25], from University of Maryland. This database contains walking sequences of 25 people in 4 different poses (frontal view/walking-toward, frontal view/walking-away, frontal-parallel view/toward left, frontal-parallel view/toward right).

A database containing both indoor and outdoor sequences is the Southampton Human ID gait database (SOTON Database) [18]. This database consists of a large population, which is intended to address whether gait is individual across a significant number of people in normal conditions, and a small population database, which is intended to investigate the robustness of biometric techniques to imagery of the same subject in various common conditions.

Currently, in real problems, more complex situations are managed. Thus, for example, outdoor scenarios may be appropriate to deal with real surveillance situations, where occlusions occur frequently. To address the challenge of occlusions, the TUM-IITKGP Gait Database is presented in [26].

Other indoor and outdoor datasets have been specifically designed for action recognition. However, it is possible to extract a subset of gait sequences from them. Examples of this are Weizmann [27], KTH [28], Etiseo, Visor [29] and UIUC [30]. The Weizmann database contains the walking action among other 10 human actions, each action performed by nine people. KTH dataset contains six types of human actions performed several times by 25 people in four different scenarios. ETISEO and Visor were created to be applied in video surveillance algorithms. UIUC (from University of Illinois) consists of 532 high resolution sequences of 14 activities (including walking) performed by eight actors.

With the new gait recognition approaches that deal with 3D information, new gait datasets for multi-view recognition have emerged. One of the first multi-view published dataset was CASIA Dataset B [20]. Dataset B is a large multi-view gait database. There are 124 subjects, and the gait data was captured from 11 viewpoints. Neither the camera position nor the camera orientation are provided.

As can be seen in OU-ISIR and SOTON databases among others, treadmills are widely used, nonetheless, an inherent problem associated with walking on a treadmill is that the human gait is not as natural as it should be. An example of using of the treadmill in a multiview dataset is presented in the CMU Motion of Body (MoBo) Database [19], which contains videos of 25 subjects walking on

Table 1: **Multi-view action datasets that include walking as an action.** This table shows some of the most popular multiview action datasets, which contain walking sequences among other activities.

Database	Actions	Subjects	Source	Views	Path	Year
i3DPost [31]	13	8	Indoor	Eight views	Straight	2009
MuHAVi [32]	17	14	Indoor	Eight views	Straight	2010
IXMAS [33]	11	10	Indoor	Five views	Closed curve	2006

a treadmill from multiple views. A summary of the current freely available gait datasets is shown in Table 2.

Other multiview datasets, specifically designed for action recognition rather than gait recognition, are described then. A summary of them can be seen in Table 1. The i3DPost Multi-View Dataset [31] was recorded using a convergent eight camera setup to produce high definition multi-view videos, where each video depicts one of eight people performing one of twelve different human actions. A subset for gait recognition can be obtained from this dataset. The actors enter the scene from different entry points, which seems to be suitable to test invariant view gait recognition algorithms. However, the main drawback of this subset is the short length of the gait sequences, extracted from a bigger collection of actions.

The Faculty of Science, Engineering and Computing of Kingston University collected in 2010 a large body of human action video data named MuHAVi (Multicamera Human Action Video dataset) [32]. It provides a realistic challenge to objectively compare action recognition algorithms. There are 17 action classes (including walk and turn back) performed by 14 actors. A total of eight non-synchronized cameras are used. The main weakness of MuHAVi is that the walking activity is carried out in an unique predefined trajectory. Due to this, this dataset is not very suitable to compare invariant-view gait recognition algorithms. This dataset was specifically designed to test action recognition algorithms, and it does not contain gait sequences of enough length.

The INRIA Xmas Motion Acquisition Sequences (IXMAS) database, reported in [33], contains five-view video and 3D body model sequences for eleven actions and ten persons. A subset for gait recognition challenges can be obtained from the INRIA IXMAS database. However, humans appear walking in very closed circle paths. Consequently, the dataset does not provide very realistic gait sequences.

Table 2: **Summary of existing gait datasets.** The table below shows some features of the existing gait datasets. Some of the current databases are divided into other subsets, to deal with specific challenges, as clothes variation, carrying conditions, or multiple view recognition. The column sequences shows the number of available sequences per subject.

Database	Subset	Type of problem	Subjects	Sequences	Source	Treadmill	Views	Path	Year
UCSD [21]	N.A	Shaded scenes	6	7	Outdoor	No	Side	Circular	1998
HID-UMD [25]	N.A	Undetermined	25	1	Outdoor	No	Front, side	Straight	2001
MoBo [19]	N.A	Multi-view recognition	25	4	Indoor	Yes	Six views	Straight	2001
SOTON [18]	Large	Multiple purposes	100	6	In-outdoor	Some seq.	0, 45, 90	Straight	2002
	Small	Diff. walk. cond.	12	15	Indoor	No	0, 45, 90	Straight	2002
CASIA	A [23]	Undetermined	20	12	Outdoor	No	0, 45, 90	Straight	2001
	B [20]	Multi-view recognition and diff. carrying cond.	124	10	Indoor	No	11 views	Straight	2005
	C [24]	Diff. walk. cond.	153	10	Outdoor	No	Side	Straight	2005
USF Human ID [22]	N.A	Covariate conditions	122	Up to 5	Outdoor	No	Side	Elliptical	2005
	TUM-IITKGP [26]	N.A	Occlusions	35	1	Indoor	No	Straight	2011
OU-ISIR [17]	A	Speed variation	34	68	Indoor	Yes	Side	Straight	2012
	B	Clothes variation	68	Up to 32	Indoor	Yes	Side	Straight	
	D	Gait fluctuation	370	185	Indoor	Yes	Side	Straight	
AVA	N.A	Multi-view recognition	20	10	Indoor	No	Six views	Curved and straight	2013

3 AVA Multi-View Dataset for Gait Recognition

In this section we briefly describe the camera setup, the database content, and the preprocessing steps carried out in order to further increase the applicability of the database.

3.1 Studio environment and camera setup

Six convergent IEEE-1394 FireFly MV FFMV-03M2C cameras are equipped in the studio where the dataset was recorded, spaced in a square of $5.8m$ of side at a height of $2.3m$ above the studio floor. The cameras provide 360° coverage of a capture volume of $5m \times 5m \times 2.2m$.

A natural ambient illumination is provided by four windows through which natural light enters into the scene. Video gait sequences were recorded at different times of day and the cameras were positioned above the capture volume and were directed downward.

Instead of using a screen backdrop of a specific color, as in [31], the background of the scene is the white wall of the studio. However, to facilitate foreground segmentation, the actors wear clothes of different color than the background scene.

Human gait is captured in 4:3 format with 640×480 pixels at 25Hz. Synchronized videos from all six cameras were recorded uncompressed directly to disk with a dedicated PC capture box. All cameras were calibrated to extract their intrinsic (focal length, centre of projection, and distortion coefficients) and extrinsic (pose, orientation) parameters.

To get the intrinsics of each camera, we used a classical black-white chess-board based technique [34] (OpenCV), while for the extrinsics we used the Aruco library [35] whose detection of boards (several markers arranged in a grid) have two main advantages. First, since there is more than one marker, it is less likely to lose them all at the same time. Second, the more markers detected, the more points available for computing the camera extrinsics. An example of extrinsics calibration based on Aruco library is shown in Figure 1. Calibration of the studio multi-camera system can be done in less than 10 minutes using the above referenced techniques.

3.2 Database description

Using the camera setup described above, twenty humans (4 females and 16 males), participated in ten recording sessions each. Consequently, the database contains 200 multi-view videos or 1200 (6×200) single view videos. In the following paragraphs we describe the walking activity carried out by each actor of the database.

Ten gait sequences were designed before the recording sessions. All actor depict three straight walking sequences ($\{t1, \dots, t3\}$), and six curved gait sequences ($\{t4, \dots, t9\}$), as if they had to turn a corner. The curved paths are composed by a first section in straight line, then a slight turn, and finally a final straight

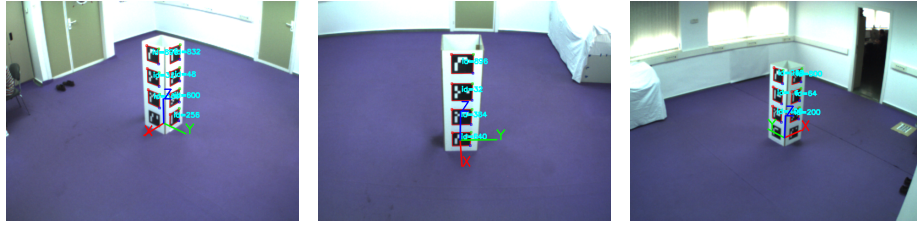


Fig. 1: **3D artifact with Aruco [35] board of markers**, used for getting the pose and orientation of each camera.

segment. These paths are graphically described in Figure 2. In the last sequence actors describe a figure-eight path (t_{10}).

3.3 Multi-view video preprocessing

The raw video sequences were preprocessed to further increase the applicability of the database. To obtain the silhouettes of actors, we have used the Horprasert's algorithm [36]. This algorithm is able to detect moving objects in a

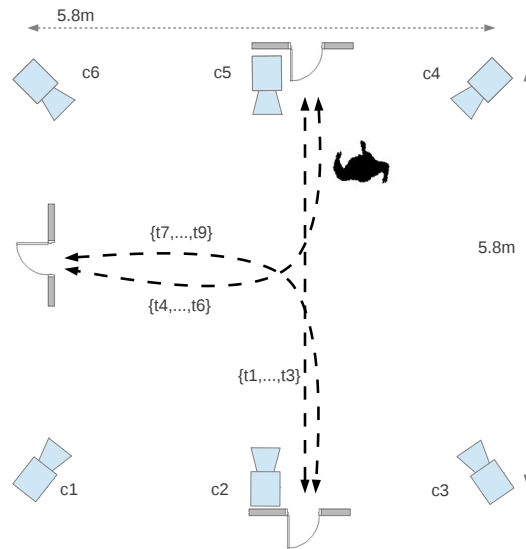


Fig. 2: **Workspace setup for dataset recording**, where $\{c1, \dots, c6\}$ represent the set of cameras of the multiview dataset and $\{t1, \dots, t9\}$ represent the different trajectories followed by each actor of the dataset.



Fig. 3: **Example of our multiview dataset.** People walking in different directions, from multiple points of view.

static background scene that contains shadows on color images, and it is also able to deal with local and global perturbations such as illumination changes, casted shadows and lightening.

In figure 3, several walking subjects of the AVA Multi-View Dataset for Gait Recognition are shown.

4 Database application examples

In this section, we carry out several experiments to validate our database. First, we use a Shape from Silhouette algorithm [37] to get 3D reconstructed human volumes along the gait sequence. The whole gait sequences can be reconstructed, as shown in Figure 4. Then, these gait volumes are aligned and centred respect to a global reference system. After this, we can get rendered projections of these volumes to test 2D-based gait recognition algorithms. By this way, we can test view-dependent gait recognition algorithms on any kind of path, either curved or straight.

4.1 Gait Recognition Based on rendered Gait Entropy Images

One of the most cited silhouette-based representations is the Gait Energy Image (GEI) [9], which represents gait using a single grey scale image obtained by averaging the silhouettes extracted over a complete gait cycle. In addition to GEI, a gait representation called Gait Entropy Image (GEnI) is proposed in [7]. GEnI encodes in a single image the randomness of pixel values in the silhouette images over a complete gait cycle. Thus it is a compact representation which is an ideal starting point for feature selection. In fact, GEnI was proposed to measure the relevance of gait features extracted from the GEI.

As we have aligned gait volumes, we can use rendered side projections of the aligned volumes to compute the GEI. In addition, the GEnI can be computed



Fig. 4: **3D reconstructed gait sequences.** Example of reconstructed gait sequences, sampled at 2Hz, where each point represents the center of a squared voxel.

by calculating Shannon entropy for each pixel of the silhouette images, rendered from side projections of the aligned volumes. In this way, the methods proposed in [9] and [7] can be tested in a view invariant way.

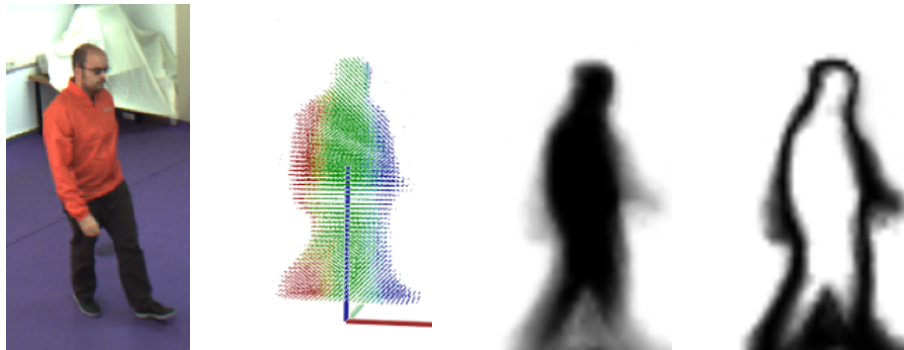


Fig. 5: **GEI and GEnI**. The leftmost image shows a walking subject. The reconstructed volumes are aligned along the gait sequence, as can be seen in the second image. The two last images show the GEI and GEnI computed over rendered images of the aligned sequence, respectively.

We designed a hold-out experiment where the gallery set is composed by the 1st, 2nd, 4th, 5th, 7th and 8th sequences and probe set is formed by 3rd, 6th, and 9th sequences of the AVA Multi-View dataset. The recognition rate obtained with the application of the gait descriptors proposed in [9] and [7] are shown in Table 3.

Table 3: **Results of the algorithm proposed in [7] on the AVA Multi-View dataset**, based on silhouettes. We report the recognition rate in %, comparing GEI with the GEnI by direct template matching, using the AVA Multi-view Dataset.

Database (probe)	GEI	GEnI
AVA Multiview Dataset (AVA B)	94.6	98.1

4.2 Front view gait recognition by intra- and inter-frame rectangle size distribution

In [8], 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. To obtain a gait representation, a morphological descriptor, called Cover by Rectangles (CR), is defined as the union of all the largest rectangles that can fit inside a silhouette. Despite of the high recognition rate, a drawback of this approach is the dependence with respect to the angle of the camera.

According to the authors, Cover by Rectangles has the following useful properties: (1) the elements of the set overlap each other, introducing redundancy (i.e. robustness), (2) each rectangle covers at least one pixel that belongs to no other rectangle, and (3) the union of all rectangles reconstructs the silhouette so that no information is ever lost. A representation of this descriptor can be seen in Figure 6.

As we have aligned gait volumes, we can use front-rendered projections of the aligned volumes to compute the CR, and therefore the method proposed in [8] can be tested in a view invariant way.

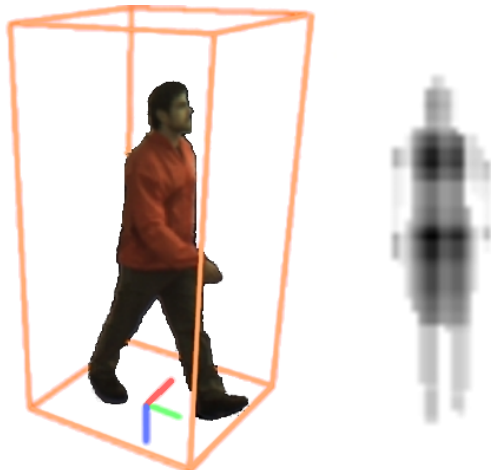


Fig. 6: **Cover by Rectangles descriptor.** Bounding box of a walking human (left), Cover by Rectangles descriptor (right). A gray level on pixel displays the density of rectangles that contains that pixel.

To test the algorithm with the AVA Multi-view Dataset, we use a leave-one-out cross-validation. Each fold is composed by a tuple formed by a set of 20 sequences (one sequence per actor) for testing, and by the remaining eight sequences of each actor for training, i.e. 160 sequences for training and 20 sequences for test. We use SVM with Radial Basis Functions, since we obtained better results than with others classifiers. To make the choice of SVM parameters independent of the sequence test data, we cross-validate the SVM parameters on the training set. For this experiment, the features vector size was set to $L = 20$, and the histogram size with which the highest classification rate is achieved is

$M = N = 25$ (see [8]). With the CR descriptor applied on the frontal volume projection, we obtain a maximum accuracy of 84.52%, as can be seen in Figure 7.

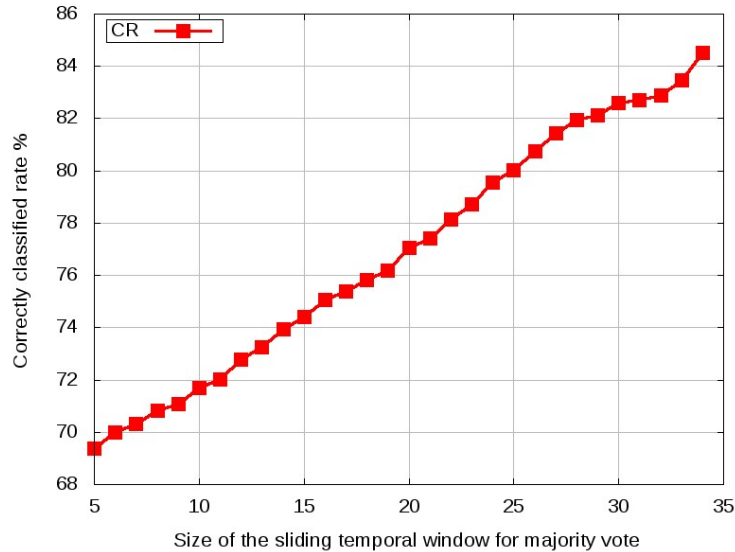


Fig. 7: **Recognition rate obtained with the application of the appearance based algorithm proposed in [8].** Since we have aligned the reconstructed volumes along the gait sequence, we can use a frontal projection of them. We show the effect on the classification rate of using a sliding temporal window for voting.

5 Conclusions

In this paper, we present a new multi-view database containing gait sequences of 20 actors that depict ten different trajectories each. The database has been specifically designed to test multi-view and 3D based gait recognition algorithms. The dataset contains videos of 20 walking persons (men and women) with a large variety of body size, who walk along straight and curved paths. The cameras have been calibrated and both calibration information and binary silhouettes are also provided.

To validate our database, we have carried out some experiments. We began with the 3D reconstruction of volumes of walking people. Then, we aligned and centred them respect to a global reference system. After this, since we have reconstructed and aligned gait sequences, we used rendered projections of these

volumes to test some appearance-based algorithms that work with silhouettes to identify an individual by his manner of walking.

This dataset can be applied in workspaces where subjects cannot show the face or use the fingerprint, and even they have to wear special clothing, e.g. a laboratory. The dataset is free only for research purposes¹.

6 Acknowledgements

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. Hu, W., Tan, T., Wang, L., Maybank, S.: A survey on visual surveillance of object motion and behaviors. *IEEE Transactions on Systems, Man, and Cybernetics* **34** (2004) 334–352
2. Lee, C.P., Tan, A.W.C., Tan, S.C.: Gait recognition via optimally interpolated deformable contours. *Pattern Recognition Letters* **34** (2013) 663–669
3. Das Choudhury, S., Tjahjadi, T.: Gait recognition based on shape and motion analysis of silhouette contours. *Computer Vision and Image Understanding* **117** (2013) 1770–1785
4. Zeng, W., Wang, C.: Human gait recognition via deterministic learning. *Neural Networks* **35** (2012) 92–102
5. Roy, A., Sural, S., Mukherjee, J.: Gait recognition using pose kinematics and pose energy image. *Signal Processing* **92** (2012) 780–792
6. Huang, X., Boulgouris, N.: Gait recognition with shifted energy image and structural feature extraction. *IEEE Transactions on Image Processing* **21** (2012) 2256–2268
7. Bashir, K., Xiang, T., Gong, S.: Gait recognition without subject cooperation. *Pattern Recognition Letters* **31** (2010) 2052 – 2060
8. Barnich, O., Van Droogenbroeck, M.: Frontal-view gait recognition by intra- and inter-frame rectangle size distribution. *Pattern Recognition Letters* **30** (2009) 893–901
9. Han, J., Bhanu, B.: Individual recognition using gait energy image. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **28** (2006) 316–322
10. Iwashita, Y., Ogawara, K., Kurazume, R.: Identification of people walking along curved trajectories. *Pattern Recognition Letters* (2014)
11. Kusakunniran, W., Wu, Q., Zhang, J., Li, H.: Gait recognition under various viewing angles based on correlated motion regression. *IEEE Transactions on Circuits and Systems for Video Technology* **22** (2012) 966–980
12. Krzeszowski, T., Kwolek, B., Michalczyk, A., Świtoński, A., Josiński, H.: View independent human gait recognition using markerless 3D human motion capture. In Bolc, L., Tadeusiewicz, R., Chmielewski, L.J., Wojciechowski, K., eds.: *Computer Vision and Graphics*. Number 7594 in *Lecture Notes in Computer Science*. Springer Berlin Heidelberg (2012) 491–500

¹ Full database access information: <http://www.uco.es/grupos/ava/node/41>

13. Lu, J., Tan, Y.P.: Uncorrelated discriminant simplex analysis for view-invariant gait signal computing. *Pattern Recognition Letters* **31** (2010) 382–393
14. Goffredo, M., Bouchrika, I., Carter, J., Nixon, M.: Self-calibrating view-invariant gait biometrics. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics* **40** (2010) 997–1008
15. Kusakunniran, W., Wu, Q., Li, H., Zhang, J.: Multiple views gait recognition using view transformation model based on optimized gait energy image. In: 2009 IEEE 12th International Conference on Computer Vision Workshops (ICCV Workshops). (2009) 1058–1064
16. Bodor, R., Drenner, A., Fehr, D., Masoud, O., Papanikolopoulos, N.: View-independent human motion classification using image-based reconstruction. *Image and Vision Computing* **27** (2009) 1194–1206
17. Makihara, Y., Mannami, H., Tsuji, A., Hossain, M., Sugiura, K., Mori, A., Yagi, Y.: The ou-isir gait database comprising the treadmill dataset. *IPSN Trans. on Computer Vision and Applications* **4** (2012) 53–62
18. Shutler, J., Grant, M., Nixon, M.S., Carter, J.N.: On a large sequence-based human gait database. In: *Proc RASC*, Springer Verlag (2002) 66–72
19. Gross, R., Shi, J.: The cmu motion of body (mobo) database. Technical Report CMU-RI-TR-01-18, Robotics Institute, Pittsburgh, PA (2001)
20. Yu, S., Tan, D., Tan, T.: A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition. In: *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on. Volume 4.* (2006) 441–444
21. Nixon, M.S., Tan, T.N., Chellappa, R.: Human identification based on gait. Volume 4. Springer (2006)
22. 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* **27** (2005) 162–177
23. Wang, L., Tan, T., Ning, H., Hu, W.: Silhouette analysis-based gait recognition for human identification. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **25** (2003) 1505–1518
24. Tan, D., Huang, K., Yu, S., Tan, T.: Efficient night gait recognition based on template matching. In: *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on. Volume 3.* (2006) 1000–1003
25. Chalidabhongse, T., Kruger, V., Chellappa, R.: The umd database for human identification at a distance. Technical report, University of Maryland (2001)
26. Hofmann, M., Sural, S., Rigoll, G.: Gait recognition in the presence of occlusion: A new dataset and baseline algorithm. In: *Proc. 19th Intern. Conf. on Computer Graphics, Visualization and Computer Vision (WSCG), Plzen, Czech Republic.* (2011) 31.01.-03.02.2011.
27. Blank, M., Gorelick, L., Shechtman, E., Irani, M., Basri, R.: Actions as space-time shapes. In: *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on. Volume 2.* (2005) 1395–1402 Vol. 2
28. Schuldt, C., Laptev, I., Caputo, B.: Recognizing human actions: a local svm approach. In: *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on. Volume 3.* (2004) 32–36 Vol.3
29. Vezzani, R., Cucchiara, R.: Video surveillance online repository (visor): an integrated framework. *Multimedia Tools Appl.* **50** (2010) 359–380
30. Tran, D., Sorokin, A.: Human activity recognition with metric learning. In: *Computer Vision – ECCV 2008. Volume 5302 of Lecture Notes in Computer Science.* Springer Berlin Heidelberg (2008) 548–561

31. Gkalelis, N., Kim, H., Hilton, A., Nikolaidis, N., Pitas, I.: The i3dpost multi-view and 3d human action/interaction database. In: Proceedings of the 2009 Conference for Visual Media Production. CVMP '09, Washington, DC, USA, IEEE Computer Society (2009) 159–168
32. Singh, S., Velastin, S., Ragheb, H.: Muhavi: A multicamera human action video dataset for the evaluation of action recognition methods. In: Advanced Video and Signal Based Surveillance (AVSS), 2010 Seventh IEEE International Conference on. (2010) 48–55
33. Weinland, D., Ronfard, R., Boyer, E.: Free viewpoint action recognition using motion history volumes. *Comput. Vis. Image Underst.* **104** (2006) 249–257
34. Bradski, G., Kaehler, A.: *Learning OpenCV: Computer Vision with the OpenCV Library*. O'Reilly, Cambridge, MA (2008)
35. Garrido-Jurado, S., Muñoz-Salinas, R., Madrid-Cuevas, F., Marín-Jiménez, M.: Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognition* **47** (2014) 2280 – 2292
36. Horprasert, T., Harwood, D., Davis, L.S.: A statistical approach for real-time robust background subtraction and shadow detection. In: Proc. IEEE ICCV. (1999) 1–19
37. Díaz-Más, L., Muñoz-Salinas, R., Madrid-Cuevas, F., Medina-Carnicer, R.: Shape from silhouette using dempster–shafer theory. *Pattern Recognition* **43** (2010) 2119 – 2131