RESEARCH **Open Access**

Robust gait-based gender classification using depth cameras

Laura Igual[1,3]*, Àgata Lapedriza[2,3] and Ricard Borràs[3]

Abstract

This article presents a new approach for gait-based gender recognition using depth cameras, that can run in real time. The main contribution of this study is a new fast feature extraction strategy that uses the 3D point cloud obtained from the frames in a gait cycle. For each frame, these points are aligned according to their centroid and grouped. After that, they are projected into their PCA plane, obtaining a representation of the cycle particularly robust against view changes. Then, final discriminative features are computed by first making a histogram of the projected points and then using linear discriminant analysis. To test the method we have used the DGait database, which is currently the only publicly available database for gait analysis that includes depth information. We have performed experiments on manually labeled cycles and over whole video sequences, and the results show that our method improves the accuracy significantly, compared with state-of-the-art systems which do not use depth information. Furthermore, our approach is insensitive to illumination changes, given that it discards the RGB information. That makes the method especially suitable for real applications, as illustrated in the last part of the experiments section.

1 Introduction

Recently, human gait recognition from a medium distance has been attracting more attention. The automatic processing of this type of information has multiple applications, including medical ones [1] or its use as a biometric [2]. Beyond other biometrics such as face, iris, or finger-prints, gait patterns give information about other characteristics (for instance, gender or age) making its analysis interesting for multiple applications. Some enterprises are developing methods for automatically collecting population statistics in railway stations, airports, or shopping malls [3], while marketing departments are working on the development of interactive and personalized advertising. In these contexts, and for these goals, gait is an information source that is particularly appropriate, given that it can be acquired in a non-intrusive way and is accessible even at low resolutions [4]. This is illustrated in Figure 1, where we can see two images captured by a camera located at the entrance of a shopping mall. In this situation, it is very difficult to perform an accurate face

analysis to extract characteristics of the subjects visiting the mall, because of the view angle, the low resolution of the faces, and the illumination conditions. However, some interesting information about the subjects can be extracted by analyzing the body figure and the walking patterns.

In this article, we focus our attention on the problem of gender classification. Almost any gait classification task can benefit from a previous robust gender recognition phase. However, current systems for gender recognition are still beyond human abilities. In short, there are three main drawbacks in the automatic gait classification problems: (i) the human figure segmentation, which is usually highly computationally demanding, (ii) the changes of viewpoint, and (iii) the partial occlusions. This study deals particularly with the first two drawbacks, and presents some future lines of research regarding the third one.

In order to improve the gait-based gender classification methods, we propose to use depth cameras. More concretely, we present a gait feature extraction system that uses just depth information. In particular, we used Microsoft's Kinect, which is an affordable device provided with an RGB camera and a depth sensor. It records RGBD videos at 30 frames per second at a resolution of 640×480. This device has rapidly attracted interest among

*Correspondence: ligual@ub.edu
[1]Department of Apply Mathematics and Analysis, Universitat de Barcelona, 08007 Barcelona, Spain
[3]Computer Vision Center, O Building, UAB Campus, 08193 Bellaterra, Barcelona, Spain
Full list of author information is available at the end of the article

Figure 1 Examples of Real images captured by a camera located at the entrance of a shopping mall.

the computer vision community. For example, Shotton et al. [5] won the best paper award at CVPR 2011 for their work on human pose recognition using Kinect, while ICCV 2011 included a workshop on depth cameras, especially focused on the use of Kinect for computer vision applications.

In the recent literature, we can find some papers on the use of Kinect for human detection [6], body figure segmentation [7], or pose estimation [8-10]. However, there are few works on gait analysis that use this device, although this topic can benefit from the depth information. This study is reviewed in the next section. Notice that the use of depth cameras simplifies the human figure segmentation stage, making it possible to process this information considerably faster than before. The depth information offers the possibility of extracting gait features that are more robust against view changes.

In this study, we present a new feature extraction system for gait-based gender classification that uses the 3D point cloud of the subject per frame as a source. These point clouds are aligned according to their centroid and projected into their PCA plane. Then, a 2D histogram is computed in this plane, and it is divided into five parts, to compute the final discriminative features. We use support vector machine (SVM) during the classification stage. The proposed method is detailed in Section 3.

To test our approach we used the DGait database [11]. This database is currently the only publicly available database for gait analysis that includes depth information. It has been acquired with Kinect in an indoor semi-controlled environment, and contains videos of 53 subjects walking in different directions. Subject, gender and age labels are also provided. Furthermore, a cycle per direction and subject has manually been labeled as well. We perform different experiments with the DGait database and compared our results with the state-of-the art method for gait-based gender recognition proposed by Li et al. [12]. Our system shows higher performance across all the tests and higher robustness against changes in viewpoint. Moreover, we show results with a test

performed in real environment data, where we deal with partial occlusions.

The rest of the article is organized as follows. The next section offers a brief overview of the recent literature on gait classification. Then, Section 3 details the proposed feature extraction method. In Section 4, we present the database and the results of our experiments. Finally, the last section concludes the study and proposes a line of further research.

2 Related work

In general, there are two main approaches for gait analysis, termed model-based and model-free. The first one encodes the gait information using body and motion models, and extracts features from the parameters of the models. In the second one, no prior knowledge of the human figure or walking process is assumed. While model-based methods have shown interesting robustness against view changes or occlusions [13,14], they are usually high demanding computationally. That makes them less suitable for real-time applications than model-free approaches that do not have to understand the constraints of the walking movements.

Among the model-free approaches, one of the most successful representations for gait is the Gait Energy Image (GEI) proposed by Han and Bhanu [15]. The GEI is a 2D single image that summarizes the spatio-temporal information of the gait. It is obtained from centered binary silhouette images of gait corresponding to an entire cycle, involving a step by each leg. Figure 2 shows an example of the silhouette images of a cycle, while in Figure 3 the corresponding GEI is plotted. The experiments performed by Han and Bhanu showed GEI to be an effective and efficient gait representation, and it has been used as a baseline for recent gait-based gender recognition methods. For instance, Wang et al. [16] proposed the Chromo-Gait Image for subject recognition. This representation can be seen as an extension of the GEI that encodes the temporal information with more precision via a color mapping. On the other hand, Li et al. [12] proposed a partition of the GEI and analyzed the discriminability power of each

Figure 2 Binary silhouettes of a cycle.

part for both gender and subject classifications. In the case of gender classification, Yu et al. [17] proposed a five-part partition of the GEI with the contribution of human observers, and achieved promising results in their experiments, although they just test the method using side view sequences.

Actually, most of the existing methods for gait classification just deal with the case of side views. Nevertheless, we can find some approaches dealing with the multi-view problem. For example, Makihara et al. [18] proposed a spatio-temporal silhouette volume of a walking person to encode the gait features, and then applied a view transformation model using singular value decomposition to obtain a more view-invariant feature vector. More recently, a further study on the view dependency in gait recognition was presented later by Makihara et al. [19]. On the other hand, Yu et al. [20] present a set of experiments to evaluate the effect of view angle on gait

Figure 3 Example of GEI.

recognition, using the GEI images as features and classifying with Nearest Neighbors. Finally, Kusakunniran et al. [21] proposed a view transformation model (VTM), which adopts a multi-layer perceptron as a regression tool. The method estimates gait features from one view using selected regions of interest from another view. With this strategy they obtained normalized gait features of different views into the same view, before gait similarity is measured. They tested their model using several large gait databases. Their results show a significantly improved performance for both cross-view and multi-view gait recognitions, in comparison with other typical VTM methods. As previously stated, we can find some work on gait analysis that uses depth information. For instance, Ioannidis et al. [22] presented another work on gait recognition using depth images. More recently, Sivapalan et al. [23] presented a new approach for people recognition based on frontal gait energy volumes. In their study, they use a gait database that includes depth information, but the authors have not published it yet. Moreover, this dataset just includes 15 different subjects in frontal view. For this reason, this database cannot be used for this study, which aims to analyze gait from different points of view.

3 Method for gait-based gender recognition

We present a feature extraction method which is partially based on the study of Li et al. [12]. We selected this algorithm as a baseline for its tradeoff of simplicity and robustness, given that our goal is to deal with real-time applications. Moreover, it does not properly use RGB information, given that the features are extracted from the binary body silhouettes. That makes this method insensitive to illumination changes.

The steps for 2D feature extraction are the following.

1. *Data preprocessing.* For all the images in a cycle, we perform preprocessing.

   (a) We resize images to $d_x \times d_y$, to ensure that all silhouettes have the same height.
   (b) We center the upper half of the silhouette with respect to its horizontal centroid, to ensure that the torso is aligned in all the sequences.
   (c) We segment human silhouettes using depth map. The segmentation algorithm is proprietary software included in the
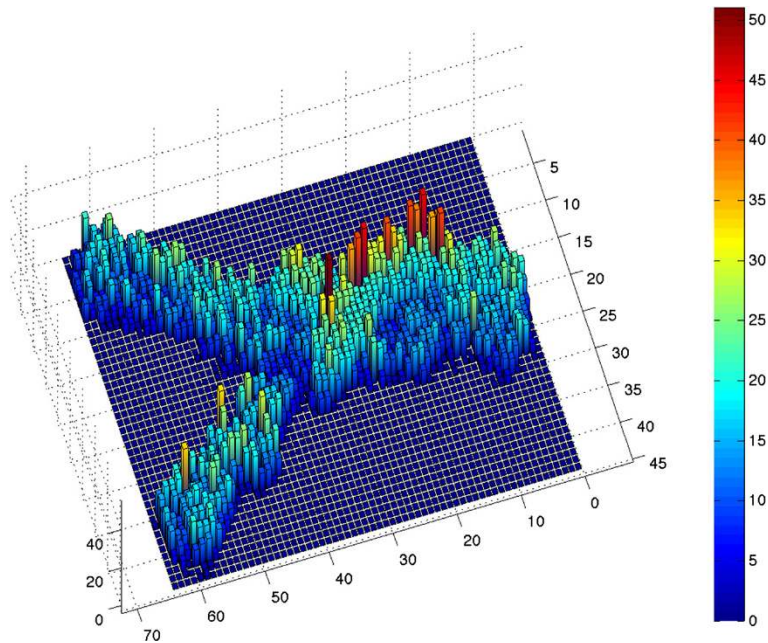
Figure 4 Projection of the point cloud into their PCA plane.

communication libraries of the Kinect (OpenNi middleware from [24]).

2. *Compute GEI.* GEI is defined as the average of silhouettes in a gait cycle (composed of T frames):

$$F(i,j) = \frac{1}{T} \sum_{t=1}^{T} I(i,j,t), \tag{1}$$

where i and j are the image coordinates, and $I(\cdot,\cdot,t)$ is the binary silhouette image obtained from the tth frame.

3. *Parts definition.* We divide the GEI into five parts corresponding to head and hair, chest, back, waist and buttocks, and legs, as defined in [17].

4. *PCA.* For every part, we compute PCA (keeping principal components that preserve the 98% of the data variance).

5. *FLD.* We compute one single feature per part using linear discriminant analysis, obtaining thus a final feature vector with five components.

The steps for 3D feature extraction are the following.

1. *3D points alignment.* For all the images of a cycle, silhouettes are segmented based on a depth map.

   (a) We keep for each frame the 3D point cloud of the subject contour.
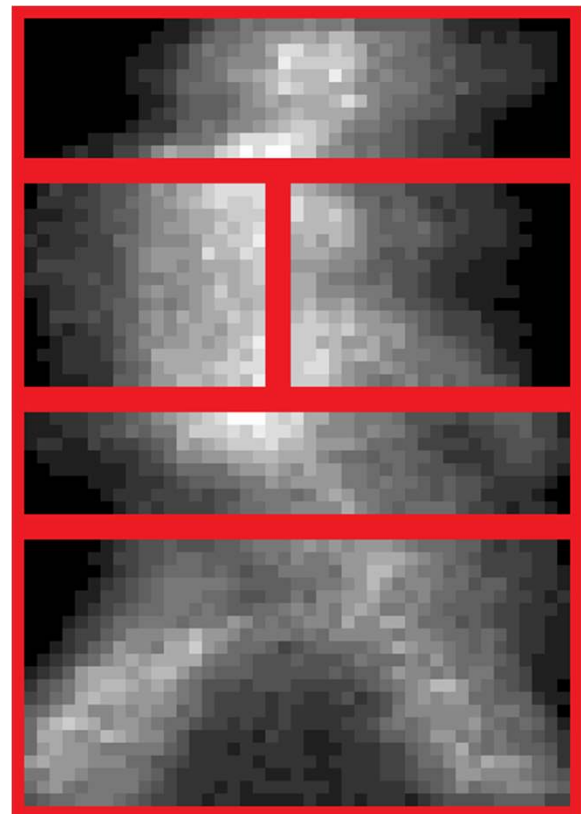   (b) We compute the 3D point cloud centroid, and use it to align the points.

Figure 5 Histogram image with the 5 body parts delimited, which are used to compute the final discriminative features.

(c) We accumulate all the centered 3D points, obtaining a single 3D point cloud that summarizes the entire cycle.

2. *PCA plane computation.* To ensure orientation invariance in the 3D-GEI features, we compute the

PCA plane of the accumulated point cloud and use it to represent the 3D information. We rotate the plane so that the *y*-axis points up and the *x*-axis points to the right.

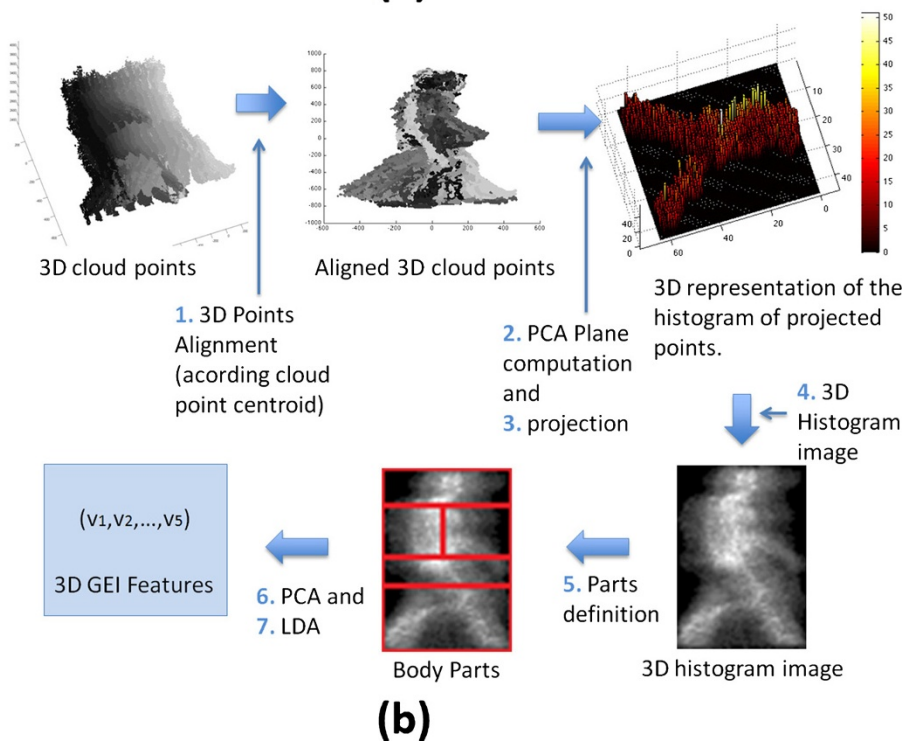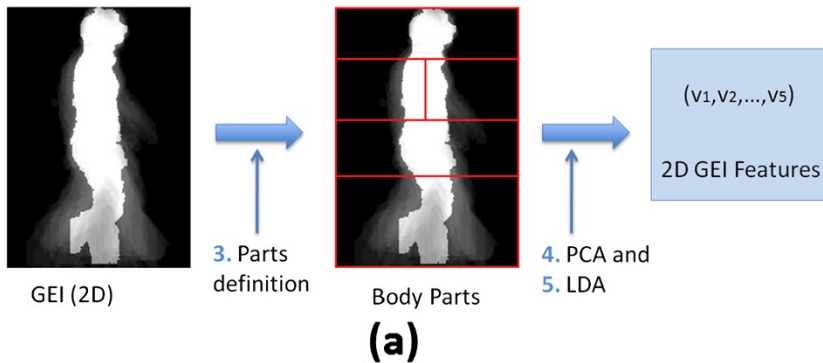3. *Point projection.* We project all the points into the PCA plane. The PCA plane contains the main

Figure 6 Flowchart of the 2D method (a) and 3D (b) feature extraction. In the 3D cloud point, the grey level represents the frame number in the cycle.

orientation of the person in the 3D frame, and projecting data into this plane allows us to capture orientation invariant shapes.

4. *3D histogram definition.* We consider the smallest window that contains all the projected points, divide it into a grid of $m_x \times m_y$ bins, and compute a histogram image, as illustrated in Figure 4. More concretely, each cell of the grid represents the number of points whose projection belongs to the cell.

5. *Parts definition.* We divide the 3D histogram into five parts corresponding to body parts as it is done for 2D-GEI images. In Figure 5, we plotted an example of the histogram image with the corresponding parts.

6. *PCA.* For every part, we compute PCA (keeping principal components that explain the 98% of data variance).

7. *FLD.* We compute one single feature per part using linear discriminant analysis, obtaining thus a final feature vector with five components.

Figure 6 shows a flowchart of the process of 2D feature extraction and 3D feature extraction.

4 Experiments
In order to test the proposed method we perform different experiments with the DGait database. This database contains DRGB gait video sequences from different points of view. The next section briefly describes this dataset.

We use the manually labeled cycles to perform a first evaluation, and then we test the method without these labels on the entire trajectories of each subject. In these tests, we compare our results with the 2D feature extraction method described in [12]. We denote this method by 2D-FE, while our method is denoted by 3D-FE. Notice that both methods extract a final feature vector of five components. In all the experiments, we used the OpenNi middleware from [24] to segment silhouettes in the scene. On the other hand, we classify with SVM. Concretely, we used the OSU-SVM toolbox for Matlab [25].

In the last part of this section, we show results computed in real-time in a video acquired in a non-controlled environment using our 3D-FE method.

4.1 The DGait database
The DGait database was acquired in an indoor environment, using Microsoft's Kinect [26]. The dataset contains DRGB video sequences from 53 subjects, 36 male (67.9%) and 17 female (32.0%), most of them Caucasian. This database can be viewed and downloaded at the following address: www.cvc.uab.es/DGaitDB.

The Kinect was placed 2 m above the ground to acquire the images. Subjects were asked to walk in a predefined circuit performing different trajectories. Figure 7 summarizes the map of trajectories recorded, which are marked in purple. For each trajectory, there are different sequences, denoted by red arrows. Thus, we have an RGBD video per subject with a total of 11 sequences.
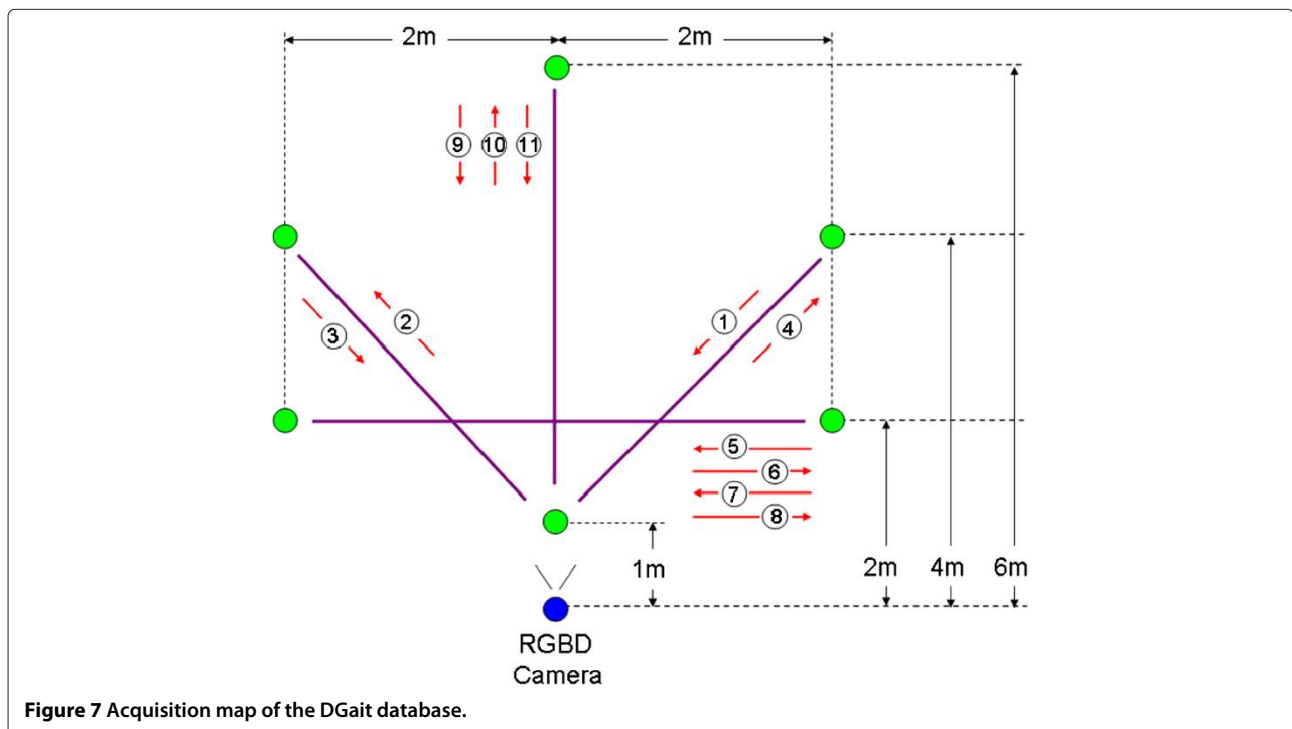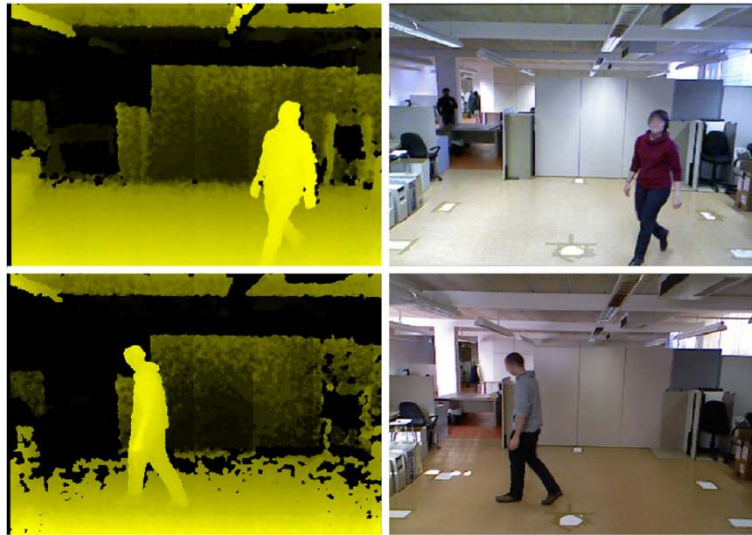

Figure 7 Acquisition map of the DGait database.

Figure 8 Example of depth map and RGB image of a frame in the DGait database.

In particular, the views from the camera are right diagonal (1,3), left diagonal (2,4), side (5,6,7,8), and frontal (9,10,11).

In the case of side views, subjects were asked to look at the camera in sequences 7,8, while in the rest of the sequences subjects are looking forward. Figure 8 shows two frames of the database. On the left we can see the depth maps, and the right images show the RGB.

The database contains one video per subject, containing all the sequences. The labels provided with the database are subject, gender, and age. Also the initial and final frames of an entire gait cycle per direction and subject are provided. Some baseline results of gait-based gender classification using this database are shown in [11].

In all the experiments, we considered images of dimensions $d_x = 256$, $d_y = 180$, and 3D Histograms of size $m_x = 64$, $m_y = 45$.

4.2 Experiments on the labeled cycles
In these experiments, we considered for each subject the manually labeled cycle per sequence. This is a total of 11 cycles per subject, and we group them into three categories: diagonal (denoted by D), side (denoted by S), and frontal (denoted by F).

First, we performed leave-one-subject-out validation on the set of 53 subjects of the DGait database. In each run, we trained a classifier with the cycles of all the subjects except one and estimated the gender of each of the 11 cycles of the test subject separately. The RBF parameter is learned in the leave-one-subject-out validation process and is set to $\sigma = 0.007$ for 2D-FE, and $\sigma = 0.02$ for 3D-FE.

Table 1 summarizes the results of this experiment. The measures Accuracy, F-recall, and M-recall are defined as follows.

$$\text{Accuracy} = \frac{\text{TF} + \text{TM}}{\text{TF+FF+TM+FM}}, \tag{2}$$

$$\text{F} - \text{recall} = \frac{\text{TF}}{\text{TF} + \text{FM}}, \tag{3}$$

$$\text{M-recall} = \frac{\text{TM}}{\text{TM} + \text{FF}}, \tag{4}$$

where TF and TM denote true female and true male, respectively, while FF and FM denote false female and false male, respectively.

Notice that the 3D-FE improves the 2D-FE in all the measures. In particular, the higher improvement of the 3D-FE relies on the F-recall. On the other hand, observe that M-recall is higher than F-recall in all the cases. This

Table 1 Results of the leave-one-subject-out validation on labeled cycles

Measure	2D-FE (%)	3D-FE (%)
Accuracy	93.90	98.57
F-recall	75.60	94.10
M-recall	98.52	99.64

Table 2 Results of the leave-one-orientation-out on labeled cycles

Test	2D-FE (%)	3D-FE (%)
D	92.50	98.50
S	97.20	98.08
F	90.73	99.04

Table 3 Results of the leave-one-subject-out on frames

Measure	2D-FE (%)	3D-FE (%)
Accuracy	64.19	74.70
F-recall	20.43	13.11
M-recall	44.74	57.14

is because the training data are unbalanced, since the database include fewer females than males. However, the 3D-FE can represent more accurately the gait patterns even in the case of unbalanced data.

Second, we performed an experiment using the labeled cycles corresponding to different orientation trajectories. We call this experiment "leave-one-orientation-out". The goal of this test is to show the robustness of our method against view changes. For each of the three view categories, we use the following testing protocol: (i) per subject, discard the two orientations that do not have to be tested, (ii) train a gender classifier using all the cycles belonging to the two orientations that do not have to be tested, and (iii) test with the cycles of the subject belonging to the testing orientation. Thus, for instance, in the *D* test, just *S* and *F* views are used to train. The results of these tests can be seen in Table 2. Notice that the accuracies achieved by the 3D-FE are higher in the three cases than the ones obtained with the 2D-FE, suggesting that the 3D-FE method extracts features that are more invariant to view changes than the 2D-FE method.

4.3 Experiments on the video sequences

Using the DGait database, we evaluated the performance of the leave-one-subject-out experiment for each frame, making no use of the labeled cycles of the testing subject. Concretely, we trained a classifier with the cycles of all the subjects except the subject to test, and estimated the gender of the subject at each frame of the whole trajectory, separately. Thus, given a specific frame, we compute the 2D-FE and 3D-FE features using a sliding window on the frame sequence of size 20. This size is the mean length of the labeled cycles in the DGait database.

Table 3 summarizes the percentage of frames correctly detected (accuracy), percentage of female frames correctly detected (F-recall), and percentage of male frames correctly detected (M-recall). As can be seen in the table, 3D-FE improved the results obtained by 2D-FE. Note that, as in the first experiment, the lower number of female samples leads to a lower F-recall and higher M-recall.

Figure 9 compares the percentage of correct classified frames per subject for 2D-FE and 3D-FE. It can be seen that the 3D-FE accuracy is, in general, higher than the 2D-FE accuracy. Only in 4 out of 53 of the cases do the 2D-FE attain better percentage than the 3D-FE, showing very robust results.

In Figure 10, we show the accuracy results obtained at the different areas of the circuit (displayed with different colors in the figure). These areas of the circuit are defined by a clustering over the subject centroid points. In particular, we use the *X* and *Z* coordinates of every centroid, since the height of the subject is not important, and we
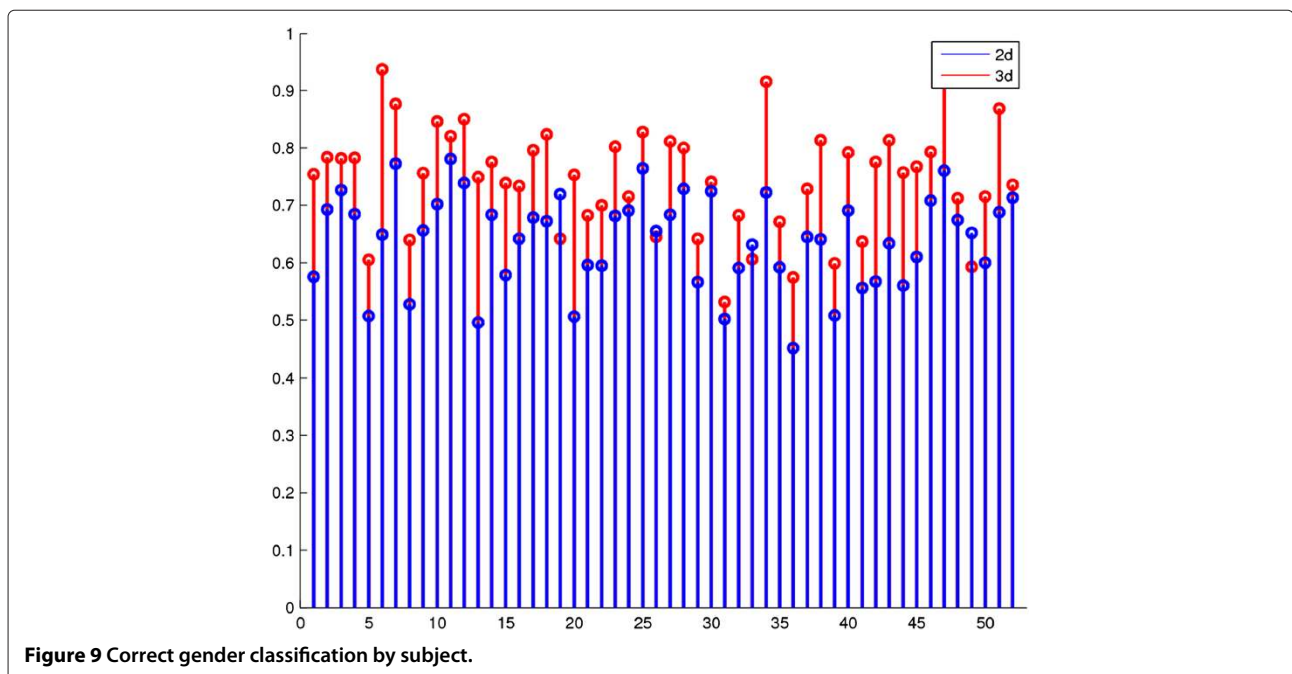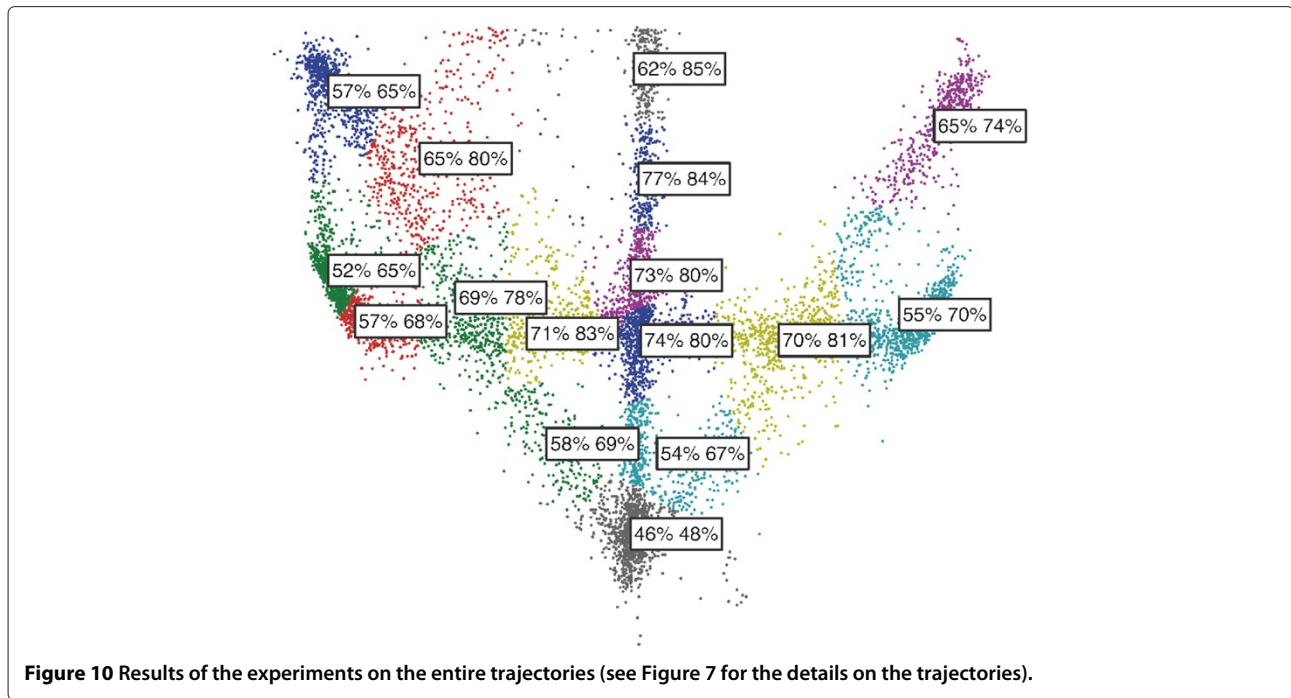

Figure 9 Correct gender classification by subject.

Figure 10 Results of the experiments on the entire trajectories (see Figure 7 for the details on the trajectories).

perform a *k*-means clustering. The value of *k* was manually set to 16 for convenience. The image orientation is the same as in the map of Figure 7. In the boxes, the accuracy values on the left correspond to the 2D-FE method, while the ones on the right were achieved by the 3D-FE method. Notice that in all the cases the 3D-FE performs better than the 2D-FE. On the other hand, the lower accuracies are obtained at the edges of the circuit for both methods, given that sometimes the human figure is partially out of the plane. Moreover, these results are evidence that 3D-FE is more robust against view changes than 2D-FE.

Finally, in this experiment, we also classified the gender of the subjects using the previous results at frame level. For that, we imposed a threshold of 60% on the percentage of classified frames. The gender of a subject will be assigned to be the gender of 60% of the frames of that subject. The obtained correct gender classification results are 65.38% for 2D-FE and 92.31% for 3D-FE. Thus, the improvement provided by 3D-FE is evident.

4.4 Evaluation on real-conditions
The last experiment performed in this study is a quantitative and qualitative evaluation of our method using a video acquired in non-controlled conditions. The Kinect was placed 2 m above the ground, and five different subjects where asked to walk around, with no specific paths. The algorithm ran in this video in real-time, and we used the SVM classifier previously learned with the DGait database for the gender recognition.

We perform the test on the whole sequence as described in Section 4.3, using a sliding window on frames of size

20, to compute the 3D-FE features. A total of 745 frames have been tested, considering each frame as many times as the number of people that appear in it. In this experiment, the OpenNi middleware was used to identify the different subjects, in order to compute the 3D-FE of each subject at each frame.

Table 4 summarizes the results of this experiment. We can see that, again, the F-recall is lower than the M-recall, given that the system has been trained with unbalanced data.

In Figure 11, we show qualitative results of some examples of classified frames. The results can be seen in the bounding box labels, indicating the estimated gender for each person. Notice the presence of partial occlusions, for instance in frames (b), (d), and (e). There are also changes with regards to what people are carrying and clothes. The female in frame (e) and a male in frame (f) are carrying the same bag, and in both cases the system is able to correctly recognize the gender. The same subjects are correctly classified in frames (a) and (b), respectively. In frame (c), we can see an example of misclassification.

5 Conclusions
In this study, we presented a new approach for gait-based gender classification using Kinect, which can run in real-time. Specifically, we proposed a feature extraction

Table 4 Results of the real-conditions evaluation

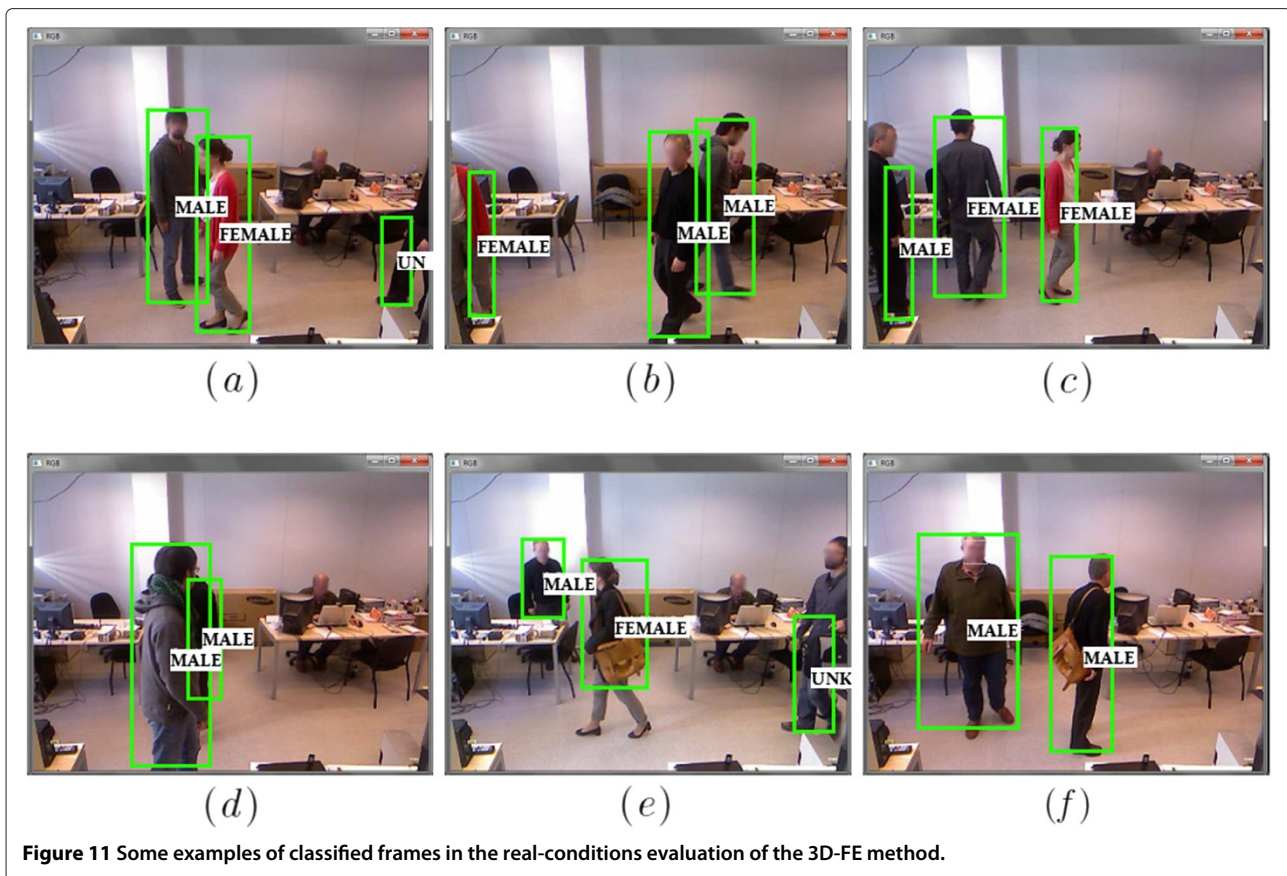| Accuracy (%)	F-recall (%)	M-recall (%)
93.90	87.20	95.70

Figure 11 Some examples of classified frames in the real-conditions evaluation of the 3D-FE method.

algorithm that takes as input the 3D point cloud of the video frames. The system does not make use of RGB information, making it insensitive to illumination changes. In short, the 3D point cloud of a cycle sequence are aligned and grouped, and then projected into their PCA plane. A 2D histogram is computed in this plane, and then final discriminative features are obtained by first dividing the histogram into parts and then using linear discriminant analysis.

To evaluate the proposed methodology we have used a DRGB database with Kinect which is the first publicly available database for gait analysis that includes both RGB and depth information. As shown in the experiments, our proposal effectively encodes the gait sequences, and is more robust against view changes than other state-of-the-art approaches that can be run without RGB information. Our method is fast and suitable for real-time and real environment applications. In the last part of our tests we show an example of its performance in this context.

In our future work, we want to focus our attention on the problem of partial occlusions. In this case, we plan to process the 3D data and 2D histogram build in a more sophisticated way, to develop a system that is more robust against missing data.

Author details
[1]Department of Apply Mathematics and Analysis, Universitat de Barcelona, 08007 Barcelona, Spain. [2]Universitat Oberta de Catalunya, Barcelona, Spain. [3]Computer Vision Center, O Building, UAB Campus, 08193 Bellaterra, Barcelona, Spain.

References
1. N Sen Köktas, N Yalabik, G Yavuzer, R Duin, A multi-classifier for grading knee osteoarthritis using gait analysis. Pattern Recognition Letters. **31**(9), 898–904 (2010)
2. K Bashir, T Xiang, S Gong, Gait recognition without subject cooperation. Pattern Recognition Letters. **31**(13), 2052–2060 (2010)
3. Trumedia, TruMedia and Dzine Introduce Joint Targeted Advertising Solution PROM Intergrated into DISplayer. http://www.trumedia.co.il/trumedia-and-dzine-introduce-joint-targeted-advertising-solution-prom-intergrated-displayer
4. DK Wagg, MS Nixon, in *Proceedings of the Sixth IEEE international conference on Automatic face and gesture recognition, FGR' 04.* On automated model-based extraction and analysis of gait (IEEE Computer Society, Seoul, Korea, 2004), pp. 11–16
5. J Shotton, A Fitzgibbon, M Cook, T Sharp, M Finocchio, R Moore, A Kipman, A Blake, in *Computer Vision and Pattern Recognition (CVPR), IEEE Conference on*. Real-time human pose recognition in parts from single

depth images (IEEE Publisher, Piscataway, New Jersey USA, 2011), pp. 1297–1304

6. L Spinello, K Arras, People detection in RGB-D data. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 3838–3843 (2011)

7. V Gulshan, V Lempitsky, A Zisserman, in *1st IEEE Workshop on Consumer Depth Cameras for Computer Vision (ICCV Workshops)*. Humanising GrabCut: Learning to segment humans using the Kinect (IEEE Publisher, Piscataway, New Jersey USA, 2011), pp. 1127–1133

8. HP Jain, A Subramanian, S Das, A Mittal, in *Proceedings of the 5th international conference on Computer vision/computer graphics collaboration techniques, MIRAGE'11*. Real-time upper-body human pose estimation using a depth camera (Springer-Verlag, Berlin, Heidelberg, 2011), pp. 227–238

9. RB Girshick, J Shotton, P Kohli, A Criminisi, AW Fitzgibbon, in *IEEE International Conference on Computer Vision (ICCV) (2011)*. Efficient regression of general-activity human poses from depth images (IEEE Publisher, Piscataway, New Jersey USA, pp. 415–422

10. A Baak, M Müller, G Bharaj, HP Seidel, C Theobalt, in *IEEE 13th International Conference on Computer Vision (ICCV), (IEEE 2011)*. A Data-Driven Approach for Real-Time Full Body Pose Reconstruction from a Depth Camera IEEE Publisher, Piscataway, New Jersey USA, pp. 1092–1099

11. R Borràs, A Lapedriza, L Igual, in *Proceedings of the International Conference on Image Analysis and Recognition*. Depth Information in Human Gait Analysis: An Experimental Study on Gender Recognition (Springer-Verlag, Berlin Heidelberg, 2012), pp. 98–105

12. X Li, S Maybank, S Yan, D Tao, D Xu, Gait components and their application to gender recognition. Systems, Man, and, Cybernetics, Part, C: Applications and Reviews. IEEE Transactions on. **38**(2), 145–155 (2008)

13. I Bouchrika, MS Nixon, in *Proceedings of the 3rd international conference on Computer vision/computer graphics collaboration techniques, MIRAGE'07*. Model-based feature extraction for gait analysis and recognition (Springer-Verlag, Berlin, Heidelberg, 2007), pp. 150–160

14. C Yam, M Nixon, Model-based Gait Recognition. Enclycopedia of Biometrics. **1**, 1082–1088 (2009)

15. J Han, B Bhanu, Individual Recognition Using Gait Energy Image. IEEE Trans. Pattern Anal. Mach. Intell. **28**, 316–322 (2006)

16. C Wang, J Zhang, J Pu, X Yuan, L Wang, in *Proceedings of the 11th European conference on Computer vision: Part I*. Chrono-gait image: a novel temporal template for gait recognition (Springer-Verlag, Berlin, Heidelberg, 2010), pp. 257–270

17. S Yu, T Tan, K Huang, K Jia, X Wu, A study on gait-based gender classification. Image Processing, IEEE Transactions on. **18**(8), 1905–1910 (2009)

18. Y Makihara, R Sagawa, Y Mukaigawa, T Echigo, Y Yagi, in *Proceedings of the 9th European conference on Computer Vision - Volume Part III*. Gait recognition using a view transformation model in the frequency domain (Springer-Verlag, Berlin, Heidelberg, 2006), pp. 151–163

19. Y Makihara, H Mannami, Y Yagi, in *Proceedings of the 10th Asian conference on Computer vision - Volume Part II , (ACCV'10)*. Gait analysis of gender and age using a large-scale multi-view gait database (Springer-Verlag, Berlin, Heidelberg, 2011), pp. 440–451

20. S Yu, D Tan, T Tan, in *18th International Conference on Pattern Recognition, (ICPR), Volume 4, (IEEE 2006)*. A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition, New Jersey USA, pp. 441–444

21. W Kusakunniran, Q Wu, J Zhang, H Li, Cross-view and multi-view gait recognitions based on view transformation model using multi-layer perceptron. Pattern Recognition Letters. **33**(7), 882–889 (2012)

22. D Ioannidis, D Tzovaras, I Damousis, S Argyropoulos, K Moustakas, Gait recognition using compact feature extraction transforms and depth information. Information Forensics and Security, IEEE Transactions on. **2**(3), 623–630 (2007)

23. S Sivapalan, D Chen, S Denman, S Sridharan, CB Fookes, in *Proc. the 1st IEEE Int. Joint Conf. on Biometrics*. Gait energy volumes and frontal gait recognition using depth images, Washington DC, USA, 2011), pp. 1–6

24. OpenNI, OpenNI Organization. www.openni.org

25. OSU-SVM, Support Vector Machine (SVM) toolbox for the MATLAB numerical environment. http://sourceforge.net/projects/svm/

26. Kinect, Microsoft Corp. Redmond WA. Kinect for Xbox 360. http://www.microsoft-careers.com/go/Kinect-for-Xbox-360-Jobs/150565/