Depth Information in Human Gait Analysis: An Experimental Study on Gender Recognition

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

[1] Centre de Visió per Computador
Edifici O Campus UAB, Bellaterra (Cerdanyola)
08193 Barcelona, Spain
[2] Universitat Oberta de Catalunya
[3] Universitat de Barcelona
{ricard.borras,agata,ligual}@cvc.uab.es
www.cvc.uab.es

Abstract. This work presents DGait, a new gait database acquired with a depth camera. This database contains videos from 53 subjects walking in different directions. The intent of this database is to provide a public set to explore whether the depth can be used as an additional information source for gait classification purposes. Each video is labelled according to subject, gender and age. Furthermore, for each subject and view point, we provide initial and final frames of an entire walk cycle. On the other hand, we perform gait-based gender classification experiments with DGait database, in order to illustrate the usefulness of depth information for this purpose. In our experiments, we extract $2D$ and $3D$ gait features based on shape descriptors, and compare the performance of these features for gender identification, using a Kernel SVM. The obtained results show that depth can be an information source of great relevance for gait classification problems.

Keywords: Biometrics, Gender Recognition, Human Gait features.

1 Introduction

This paper presents and describes a new database for human gait analysis, called *DGait*. It can be viewed and downloaded at the following address: *http://www.cvc.uab.es/DGaitDB/*.

The main advantage of this new data set with respect to existing ones is the incorporation of depth information. The database is composed of a set of relatively unconstrained gait videos. In particular, the DGait is composed of gait RGBD video sequences from 53 subjects at different view points. The database was acquired in an indoor environment, using a Microsoft's Kinect [4], which is provided with an RGB camera and a depth sensor. This device is receiving increasing attention among the computer vision research community. For example, Shotton et al. received the best paper award on CVPR 2011 for their work on real-time human pose recognition [11] using Kinect. In this paper authors

A. Campilho and M. Kamel (Eds.): ICIAR 2012, Part II, LNCS 7325, pp. 98–105, 2012.

propose a new method to quickly and accurately predict 3D positions of body joints from depth images.

The motivation of this database is the increasing interest for human gait and human figure analysis. These characteristics are accepted as biometrics, and have notable advantages beyond others, such as face, iris or finger- prints. Notice, for instance, that gait analysis can be used for uniquely recognizing humans, but also for estimating their gender or age. Moreover, gait analysis can be applied in a non-intrusive way and it is accessible at large distance or at low resolutions. These advantages of gait in front of other biometrics have raised expectations of their potential use in real applications. For example, gait can be used for automatically collecting population statistics (age, gender, ethnicity...) in train stations, airports or shopping malls. Another possible application of this biometric is in marketing, where it can be used for interactive and personalized advertisement.

We can find in the literature several approaches for gait classification [3,5,15,6,9]. In general, these methods suffer from two common drawbacks: (*i*) the difficulty of the human figure segmentation, and (*ii*) the changes on the view point. Depth information leads to realtime human figure segmentation algorithms, as the one provided in the OpenNi middleware [7]. However, it is not that evident whether the addition of depth measures can contribute to mitigate the second drawback. The analysis of this second point is the main intent of the DGait database.

In this paper, we also perform some gait-based gender recognition experiments using the DGait database. The experiments are detailed in Section 4. We extract $2D$ and $3D$ gait features, which are based on $2D$ silhouette descriptors and $3D$ body shape descriptors, respectively. Concretely, we use $3D$ body silhouette shape descriptors as proposed in [8], and their corresponding simplification to the $2D$ body silhouette shape. We compare the performance of these features, as well as their combination, for gender recognition, using a Kernel SVM classifier. The obtained results illustrate the relevance of depth information for gait classification.

The paper is organized as follows. Next section gives a brief overview of current databases for gait classification. In Section 3, we properly present the DGait database. Section 4 shows some gait-based gender recognition experiments, in order to illustrate the potential advantages of using depth information in gait classification. Finally, Section 5 gives an overview of the paper and concludes the work.

2 Related Databases

In this section, we give an overview of the existing databases for gait classification. These databases have been acquired using an RGB camera. The USF HumanID is composed of 1870 video sequences of 122 persons walking in elliptical paths. Images were acquired outdoor and, for each subject, there are up to five covariates: viewpoints (left/right), two different shoe types, surface types (concrete/grass), carrying conditions (with/without a briefcaseon), and change in elapsed time (May/November). Moreover, the public domain contains the normalized and aligned data. The CASIA Gait Database [2] is composed of three datasets: Dataset A, Dataset B (multiview dataset) and Dataset C (infrared dataset). The most complete of these sets, at a computer vision level, is Dataset B. It contains sequences from 124 subjects (93 male and 31 female) and gait data was captured from 11 views. There are three covariates in this database: view angle, clothing and carrying condition changes, which are separately considered. Besides, the human silhouettes extracted from video files are also provided. The IRIP Gait Database [14] was captured in an indoor laboratory scenario, where simple background was used to simplify silhouette segmentation. Eight cameras were placed at different angles during the acquisition to capture the movements of persons from different points of view. The database include recordings of 60 subjects, 32 male and 28 female. The Large-Scale Multi-View Gait Database [6] includes sequences of 168 subjects (88 males and 80 females) from 25 different views. Currently, just a subset of this database is public. Other examples of datasets for gait analysis are the Soton Large Database [12], which only includes side views of normal walking; and the Georgia Tech Database [13], with 20 different subjects from 3 different view points.

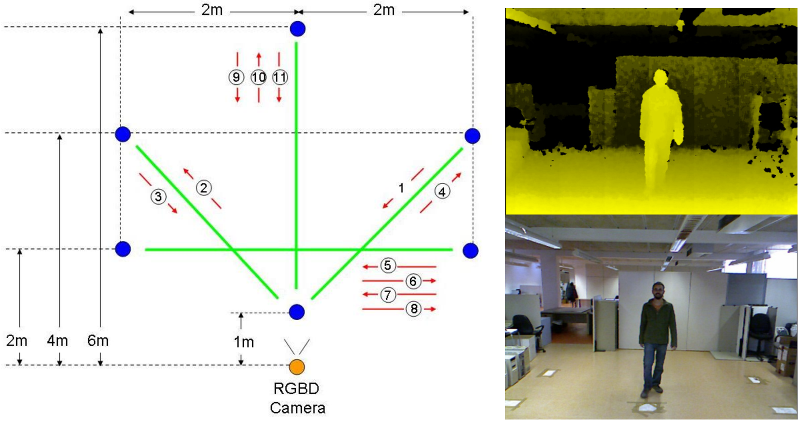
3 The *DGait* Database

The DGait database was acquired in an indoor environment, using a Microsoft's Kinect [4]. This device gives a 640×480 image at 30 frames per second with depth resolution of a few centimeters. The acquisition strategy was to simulate the scenario of a real statistical gait-based study on a shopping mall, a railway station or any public building. The dataset contains sequences from 53 subjects, 36 male (67.9%) and 17 female (32.0%), most of them Caucasian.

For the acquisition the Microsoft's Kinect was placed two meters above the ground. Subjects were asked to walk in a predefined circuit performing different trajectories. Figure 1 (a) summarizes the map of trajectories recorded, which are marked in green color. For each trajectory there are different sequences, denoted by red arrows. Thus, per each subject we have an RGBD video with a total number of 11 sequences. 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 sequences subjects are looking forward.

Figure 1 (b) shows an example of a frame in the database. On the top image we can see the depth maps, while the bottom image shows the RGB image.

The database contains one video per subject with all the sequences. The labels provided with the database are: subject, gender and age. The name of the original videos acquired with the Kinect device summarize this information. More concretely, names of files are *NsubjectGA.oni*, where *Nsubject* is integer id of the subject, G is a character denoting the gender (m for male and f for female), and A is an integer denoting the age of the subject.



(a) (b)

Fig. 1. Map of trajectories designed for the database acquisition and an example of video frame in the DGait database (depth maps and the RGB)

On the other hand, we also provide per each video initial and final frames of an entire gait cycle per each of the 11 sequences. A gait cycle is considered as the gait unity and can be useful in the gait characterization of a subject.

These cycles have been manually labeled. This information can be found in the *xml* file *cycles.xml*, as well as in the *Matlab* file *cycles.mat*.

3.1 Intended Uses and Challenges

As previously mentioned, this database is aimed at studying the contribution of depth information in gait classification problems. The data set is also provided with the RGB information. Thus, current methods that just use this information source can be easily compared with new methods that include also depth information.

Most of the methods found in the literature dealing with gait classification are focused on subject identification. However, the extraction of other characteristics from gait, such as gender or age, can be also approached with this database.

Moreover, the DGait database can be used to analyze the robustness of the gait classifications methods in presence of view changes, or even to directly classify the gait orientation.

Finally, the resolution of the images is high enough to explore the combination of face and gait analysis in the case of frontal, diagonal or profile sequences [10].

4 Application to Gender Classification

In this section, we perform gait-based gender classification experiments using DGait database. We propose a feature extraction method to define two different types of features: 2D-GF, which stands for 2D Gait-based features, and is based only in 2D body shape information (silhouettes) and 3D-GF, which stands for 3D Gait-based features, and uses the whole 3D body shape information obtained from depth maps. In short, these features are based on shape distributions, which correspond to the generalization of geometric histograms of distances between random points belonging to the silhouettes. Our objective is to show that adding 3D-GF, which includes depth information, the gait-based gender classification can be improved. Although DGait database provides RGB images together with depth maps, the feature extraction method described here is based just on depth maps. Discarding the RGB images, the method is invariant to illumination changes, which is an important issue in real applications.

4.1 Feature Extraction

We describe here the steps of the feature extraction method and the differences between 2D-GF and 3D-GF features. Figure 2 shows the flowchart of the method with the next steps:

1. **Silhouette Segmentation:** Middleware from [7] is used to segment silhouettes in the scene.
2. **Body Shape Feature Extraction:** We describe silhouettes in frames using shape distributions, as it is proposed in [8] for general 3D shapes.
   In the 2D-GF case, we represent the shape signature for a silhouette as a probability distribution sampled from a shape function. The shape function measures the Euclidean distances between pairs of randomly selected points on the border of the 2D silhouette. Then, we construct a histogram of B bins with a fixed size. From the histogram, we reconstruct a piecewise linear function with V ($\leq B$) equally spaced vertices, which forms our representation for the shape distribution. We store the shape vector of V integers in $\mathbf{v}_i$, where i represent the frame index. In the 3D-GF case, we consider the shape function defined as the Euclidean distances between pairs of randomly selected points on the surface of the 3D silhouette.
3. **Cycle Feature Extraction:** We consider cycles of a subject as the gait unity. In order to represent a cycle we compute the mean shape distribution for all frames on the cycle: $\bar{\mathbf{v}}, = \frac{1}{c}\sum_{i=1}^{c}\mathbf{v}_i$. where c is the number of frames in the cycle.
4. **Dimensionality Reduction:** We perform PCA (keeping 98% of the variance) followed by LDA.

4.2 Gender Classification

Gender classification experiments are carried out on the DGait database using 2D-GF and 3D-GF features and SVM classifier with kernel basis function. In all the experiments, we have considered a set of $R = 10000$ random points for both 2D-GF and 3D-GF methods, histograms of size $B = 2048$ and number of vertices $V = 8$. The size of the resulting shape distribution vector $\bar{\mathbf{v}}$ is 256. To reduce the dimensionality, we used PCA, keeping 98% of the variance, followed by LDA

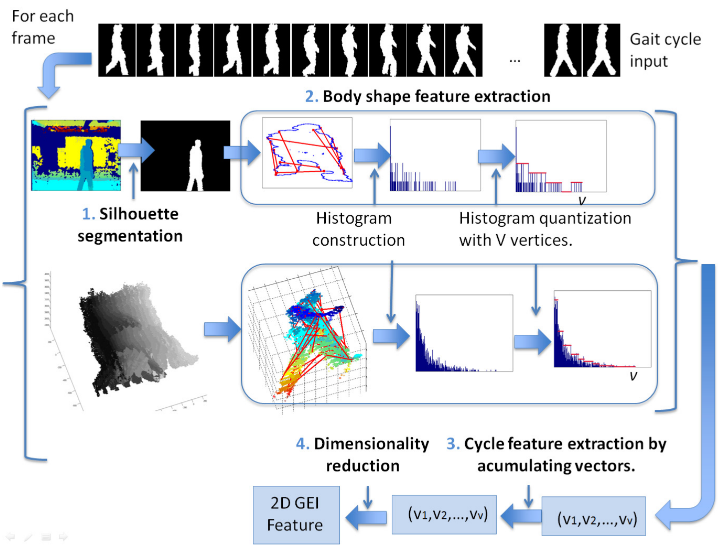


Fig. 2. Flowchart of the 2D and 3D feature extraction method

to define 2D-GF and 3D-GF features. We have used the osu-svm toolbox for Matlab [1]. The KBF parameter is learned in the leave-one-subject-out validation process and is set to $\sigma = 0.08$ for 2D-GF and $\sigma = 0.05$ for 3D-GF.

First, we estimate the gender of the DGait cycles separately. For that, we evaluated the performance using leave-one-subject-out validation on the set of 53 subjects of the DGait. The results were validated using several measures, described in terms of true female (TF), true male (TM), false female (FF) and false male (FM): Error= FF+FM, Female Sens = $\frac{\text{TF}}{\text{TF+FM}}$, Male Spec = $\frac{\text{TM}}{\text{TM+FF}}$, Female predictive ratio, F-PR = $\frac{\text{TF}}{\text{TF+FF}}$, Male predictive ratio, M-PR= $\frac{\text{TM}}{\text{TM+FM}}$, False Female Ratio, F-FR= $\frac{\text{FF}}{\text{TF+FM}}$, False Male Ratio, F-MR = $\frac{\text{FM}}{\text{TM+FF}}$. Table 1 (a) contains the results of this validation. As can be seen, combining 2D-GF and 3D-GF we can improve the results obtained from a single feature type. From 2D-GF and 3D-GF, we obtain useful complementary information. Note that the lower number of female samples leads to a lower Sensitivity and F-PR and higher F-Far than Specificity and M-PR and M-Far.

Second, we evaluated the gender classification results on 5 different orientation groups of sequence: face diagonal (sequences 1,3), back diagonal (sequences 2,4), side (sequences 5-8), face frontal (sequences 9,11) and back frontal (sequences 10). Table 1 (b) contains the mean error on these orientations and the variance obtained between them (last row). As it can be seen, the variance for 3D-GF is much smaller than for 2D-GF. This agrees with our hypothesis that 3D-GF is more robust in front of view changes.

Finally, we propose two strategies to evaluate the gait-based classification results: first computing the percentage of subjects that have been correctly classified in all the 11 sequence cycles, second computing the percentage of subjects

Table 1. Results of the gender classification per cycle (a) per orientation (b) and per subject (c)

	2D-GF	3D-GF	[2D+3D]-GF
(a)			
Mean Error	20.58	20.05	13.43
Sens	74.05	75.00	82.28
Spec	81.19	80.05	88.30
F-PR	55.11	51.14	73.86
M-PR	90.93	92.00	82.28
F-Far	60.31	71.67	29.11
M-Far	8.10	6.96	7.12
(b)			
Face diagonal	15.43	18.14	10.60
Back diagonal	19.08	17.58	12.55
Side	20.29	19.21	11.11
Face frontal	25.55	23.33	15.55
Back frontal	13.20	21.84	5.66
Variance	22.64	5.16	12.94
(c)			
% Subjects with no errors	33.96	43.40	47.17
% Subjects well classified	90.57	86.79	96.23

that are well classified with a vote strategy, taking into account the classification of their 11 sequences. Concretely, it is the percentage of subjects such that 6 or more of their sequences have been correctly classified. Table 1 (c) contains the results of the two alternative strategies. Notice that the best accuracies are always obtained with the combinations of $2D$ and $3D$ features, meaning that depth represents relevant additional information to the $2D$ features.

5 Summary and Conclusions

In this paper we have presented a new database for human gait analysis, called DGait. It has been acquired with Microsoft's Kinect, a device provided with an RGB camera and a depth sensor. Up to our knowledge, the DGait is the first publicly available data set that includes depth information.

The main intended use for the DGait is to explore the contribution of depth information in gait classification. The database includes RGBD videos from 53 subjects. Each video is composed of 11 gait sequences from different views. Identity, gender and age labels are provided. Moreover, initial and final frames of a cycle at each sequence are also available.

In this work, we also present some experiments on gait-based gender recognition using the DGait. These experiments illustrate the potential benefits of using depth information for gait classification purposes. We compared the performance of 2D gait features, 3D gait features and their combination for gender recognition. These features are based on 2D body silhouette descriptors and 3D body shape descriptors respectively. Our results show the combination of 2D and 3D features to be the most robust feature set for gender recognition. This feature combination obtains a higher accuracy in all the sequence types of the

DGait database, considering different orientations. These results show depth is useful in front of view changes and encourage the further research on the use of depth for gait classification purposes.

Acknowledgements. This work has been partially supported by Grants TIN2009-14404-C02 and CONSOLIDER-INGENIO 2010 (CSD2007-00018).

References

1. Osu svm is a support vector machine (svm) toolbox for the matlab numerical environment, `http://sourceforge.net/projects/svm/`
2. Casia: Casia gait database (2005), `www.cbsr.ia.ac.cn/english/GaitDatabases.asp`
3. Han, J., Bhanu, B.: Individual recognition using gait energy image. IEEE Transactions on Pattern Analysis and Machine Intelligence, 316–322 (2006)
4. Kinect: Microsoft corp. redmond wa. kinect for xbox 360 (2010)
5. Li, X., Maybank, S.J., Yan, S., Tao, D., Xu, D.: Gait components and their application to gender recognition. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews 38(2), 145–155 (2008)
6. Makihara, Y., Mannami, H., Yagi, Y.: Gait Analysis of Gender and Age Using a Large-Scale Multi-view Gait Database. In: Kimmel, R., Klette, R., Sugimoto, A. (eds.) ACCV 2010, Part II. LNCS, vol. 6493, pp. 440–451. Springer, Heidelberg (2011)
7. OpenNI: Openni organization, `www.openni.org`
8. Osada, R., Funkhouser, T., Chazelle, B., Dobkin, D.: Shape distributions. ACM Transactions on Graphics 21(4), 807–832 (2002)
9. Sarkar, S., Phillips, P., Liu, Z., Vega, I., Grother, P., Bowyer, K.: The humanid gait challenge problem: data sets, performance, and analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence 27(2), 162–177 (2005)
10. Shan, C., Gong, S., McOwan, P.W.: Fusing gait and face cues for human gender recognition. Neurocomputing 71(10-12), 1931–1938 (2008)
11. Shotton, J., Fitzgibbon, A., Cook, M., Sharp, T., Finocchio, M., Moore, R., Kipman, A., Blake, A.: Real-time human pose recognition in parts from single depth images. In: Proceedings of the Computer Vision and Pattern Recognition Conference (2011)
12. Soton: University of southampton. database collection of programme automatic gait recognition for human id at a distance at soton
13. Tech., G.: Georgia tech. gvu center/college of computing. data of project human identification at a distance
14. Wang, Y.: Investigating the separability of features from different views for gait based gender classification. In: 19th International Conference on Pattern Recognition, ICPR 2008., pp. 1–4. IEEE (2008)
15. Yu, S., Tan, T., Huang, K., Jia, K., Wu, X.: A study on gait-based gender classification. IEEE Transactions on Image Processing 18(8), 1905–1910 (2009)