Gait Analysis of Gender and Age Using a Large-Scale Multi-View Gait Database

Yasushi Makihara, Hidetoshi Mannami, and Yasushi Yagi

Osaka University, 8-1 Mihogaoka, Ibaraki, Osaka, 567-0047, Japan

Abstract. This paper describes video-based gait feature analysis for gender and age classification using a large-scale multi-view gait database. First, we constructed a large-scale multi-view gait database in terms of the number of subjects (168 people), the diversity of gender and age (88 males and 80 females between 4 and 75 years old), and the number of observed views (25 views) using a multi-view synchronous gait capturing system. Next, classification experiments with four classes, namely children, adult males, adult females, and the elderly were conducted to clarify view impact on classification performance. Finally, we analyzed the uniqueness of the gait features for each class for several typical views to acquire insight into gait differences among genders and age classes from a computer-vision point of view. In addition to insights consistent with previous works, we also obtained novel insights into view-dependent gait feature differences among gender and age classes as a result of the analysis.

1 Introduction

Gait is a basic and well-trained human behavior that includes both static (e.g., height, body shape) and dynamic (e.g., stride, arm swing, walking posture) components and has attracted much interest in terms of its application in, for example, person identification [1] [2] [3] [4] [5] [6], prognosis of surgery [7], and diagnosis of brain disorders [8]. In addition, the classification of gender and age has raised expectations of the potential application of gait to areas such as detection of wandering elderly in a care center and collection of customer information at shopping malls.

Some studies have challenged gender classification [9] [10] [11] and estimation of age [12] [13]. Although these studies reached reasonable conclusions, the databases used considered only one aspect, either gender or age, and included images captured from only a few views. To deal with the challenges involved, a multi-view analysis of a database that includes a wide range of genders, ages, and views is required. Using such a database, analysis of both gender and age, which has not been undertaken in previous works, would also be possible.

In this paper, we address video-based multi-view classification and analysis of gender and age classes using a large-scale multi-view gait database. Our main contributions are as follows.

The large-scale multi-view gait database: A large population of subjects (168 people), including both genders (88 males and 80 females) and a wide range of ages (between 4 and 75 years old), enables us to study and evaluate applications that require a diverse group of subjects. Moreover, with images from 25 different views (2 heights $\times$ 12 azimuth directions, plus an overhead view), multi-view applications can be supported.

Feature analysis for gender and age classes: The uniqueness of gait features for four typical classes, namely children, adult males, adult females, and the elderly, is analyzed with respect to view, and static and dynamic components. Frequency-domain features [14], representing average silhouette, and asymmetric and symmetric motion are exploited for analysis from a computer-vision point of view, since these capture effectively the 2D spatial information of the dynamic, as well as static components of gait in a more compact representation than a raw silhouette sequence [5]. We show that some insights from our feature analysis of gender and age classes are consistent with those from other gait research studies, and that some new insights are also obtained as a result of our large-scale and multi-view feature analysis.

The remainder of this paper is organized as follows. Section 2 summarizes related works. Section 3 describes the gait capturing system and the constructed gait database. Gait features used in this analysis are explained in Section 4. Section 5 presents the classification experiments for gender and age classes. Differences in the classes are analyzed in Section 6, while Section 7 compares these differences to the results of the experiments. Finally, Section 8 concludes the paper.

2 Related Work

Gait-based gender and age classification: In gender classification, Kozlowski et al. [15] introduced the possibility of automatic gender classification using subjective tests. Li et al. [10] confirmed the possibility of gender classification through experiments with the HumanID dataset [5], the subjects of which were mainly in their 20's. Various studies have also investigated the possibility of age classification. Davis [13] presented a classification of children (3-5 years old) and adults (30-52 years old), while Begg et al. [12] presented a similar classification of adults (avg. 28.4, SD 6.4) and the elderly (avg. 69.2, SD 5.1). These two studies used comparatively small samples of gender and age, but more importantly, they considered only one aspect, either gender or age.

Gait features: With respect to gait features, Kozlowski et al. [15] and Davis [13] used light points attached to a subject's body, while Begg [12] proposed the Minimum Foot Clearance, defined as the distance between the ground surface and the foot. Since all these methods require subjects to wear various devices, they are not suitable for real applications. On the other hand, various methods using gait images as contact-less information, have also been proposed. These can roughly be divided into model-based and appearance-based methods. In model-based methods, body parts are represented as various shapes (e.g., the ellipse

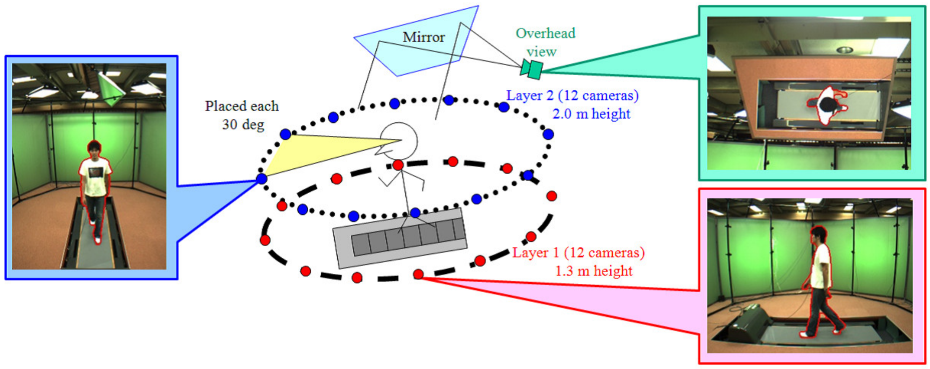


Fig. 1: Overview of multi-view synchronous gait capturing system.

model [16] and stick model [17]) and gait features are extracted as parameters of these shapes. These model-based methods often run into difficulties with model fitting, and their computational costs are comparatively high. Of the appearance-based methods, the width vector [3], frieze pattern [18], and self similarity plot [1] discard 2D spatial information, while the averaged silhouette [19] [20] [10] degrades dynamic information. On the other hand, a raw silhouette sequence [5] contains both 2D spatial and dynamic information, but lacks compactness of representation. Although most of the appearance-based methods treat only side-view or near side-view gait images, gait features captured from other views are also informative (e.g., gait-based person identification [21] and action recognition [22]). Huang et al. [9] proposed a multi-view gender classification, but the number of views was insufficient for an extensive analysis of multi-view effects.
Gait image databases: With respect to gait image databases, there are three large databases (the HumanID dataset [5], the Soton database [23], and the CASIA dataset [24]) that include over 100 subjects each. These are widely used in studies on gait-based person identification and gender classification. Each database, however, has its own particular limitations making it unsuitable for our purposes. The HumanID dataset [5] has relatively small view changes (about 30 deg.) and includes neither children nor the elderly. The Soton database [25] contains only single view images, while most of the subjects included in the CASIA dataset [24] are in their 20's or 30's.

3 The Gait Database

3.1 Multi-view Synchronous Gait Capturing System

Our image capturing system consists mainly of a treadmill, 25 synchronous cameras (2 layers of 12 encircling cameras and an overhead camera with a mirror), and six screens surrounding the treadmill, as shown in Fig. 1. The frame rate

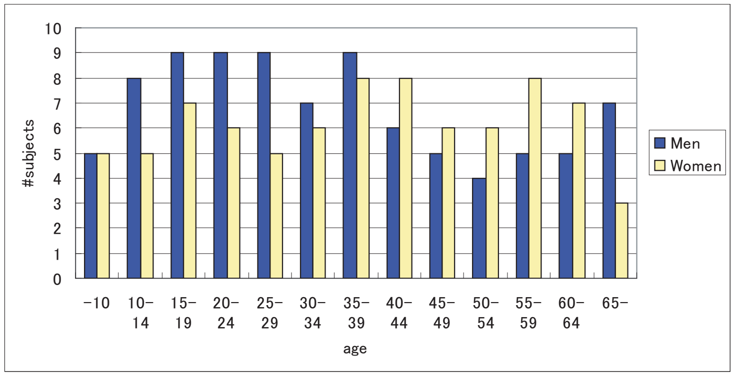


Fig. 2: Distribution of subjects' gender and age.

and resolution of each camera are set to 60 fps and VGA, respectively. The surrounding screens are used as a chroma-key background. Sample images captured in the system are also shown in Fig. 1.

3.2 Large-scale Gait Database for Gender and Age

For analyzing gender and age effects on gait, the database used must include subjects with a wide range of ages for each gender. We aimed to collect at least 10 subjects of each gender for every 10 year age group. Subjects were obtained by open recruitment and signed a statement of consent regarding the use of their images for research purposes. As a result, our database includes 168 subjects (88 male and 80 female), with ages ranging from 4 to 75 years old, (see Fig. 2 for the age distribution). A portion of the database has already been made available to the public [26]. The number of subjects exceeds that of the previous largest database and is still growing in the light of further research. Due to the reported difference between walking on the ground and on a treadmill [27], we attempted to minimize any potential difference by providing enough time for each subject to practice walking on the treadmill. After the practice sessions, subjects were asked to walk at 4 km/h or slower if necessary for children and the elderly. Subjects wore standard clothing (long-sleeved shirts and long pants, or their own casual clothes).

4 Gait Feature Extraction

In this section, we describe the gait feature extraction from the captured images. Here we define gait as a cycle of walking that includes a pair of left and right heel-strikes.

Since our purpose is the analysis of visual appearance, it is preferable to use a gait feature that captures 2D spatial information for both static and dynamic

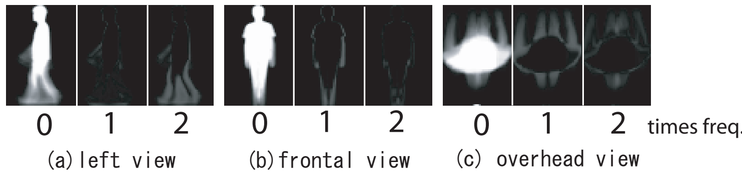


Fig. 3: Samples of frequency-domain features captured by each camera. In this paper, frequencies between 0 and 2 are used, where each frequency in side view, for example, denotes the following: 0-times represents the average silhouette, 1-time represents the asymmetry of the left and right motion, and 2-times represents the symmetry thereof.

features. Consequently, we selected a frequency-domain feature [14] because, 1) it contains both 2D spatial and dynamic information, and 2) it can be expressed in a more compact form than a raw silhouette sequence, as shown in the three frequency-channel images in Fig.3.

The procedure for frequency-domain feature extraction comprises four steps:

1. Construction of the Gait Silhouette Volume (GSV) by size normalization and registration of silhouette images (64 by 44 pixels in this paper).
2. Gait period detection by maximizing normalized autocorrelation of the GSV for the temporal axis.
3. Amplitude spectra calculation through a one-dimensional Discrete Fourier Transformation (DFT) of the temporal axis, with the detected gait period used as the base period for the DFT.
4. Application of Singular Value Decomposition (SVD) for dimension reduction of the frequency-domain feature.

A more detailed description of this process can be found in [14].

While the actual height information is obviously important for discriminating children from other age groups, the estimation process of the actual height needs camera calibration in advance. Because not all CCTV cameras in the world are calibrated, the actual height acquisition may be an unreasonable assumption. We assume the more challenging problem of classifying gender and age using an uncalibrated camera, with height-normalized gait features.

5 Preliminary Experiments on Gender and Age Classification

5.1 Method

We carried out a classification experiment of four typical gender and age classes, namely children (younger than 15 years old), adult male and adult female (male and females between 15 and 65 years old, respectively), and the elderly (aged 65 years and older). Classification performance is given for each view. For each classification experiment, 20 subjects of each class were randomly chosen from both the gallery and test (probe) images. The k Nearest-Neighbors (k-NN) algorithm was used as the classifier in LDA space [28], with the parameter of k-NN set to $k = 3$ based on preliminary experiments. Finally, the Correct Classification Rate (CCR) was calculated from 20 combinations of trials and the averaged values used in the evaluation.

5.2 Classification of the four classes

Classification performance for the four classes was evaluated independently for each view. According to the results shown in Fig. 4(a), the performance of different elevation angles shows little difference, while the performance of the overhead view is similar to the average performance of other views. In cases where weak perspective can be assumed, opposite-view cameras give horizontally inverted silhouette images [29], resulting in the same CCR. On the other hand, this assumption is not true for our gait measurement system, as the distance from the subject to the camera is not long enough compared to the width of the subject. Therefore, opposite-view cameras do not always achieve the same CCR as shown in Fig. 4(a).

5.3 Classification of Specific Combinations of Classes

The classification performance of each view feature was evaluated with respect to specific combinations of classes, namely children and adults (C-A), adult males and adult females (AM-AF), and adults and the elderly (A-E). Here the number of gallery and probe images was set as 10 for each. Figure 4(b) shows the classification results for each view. We can see that view effects on the performance are significant and therefore, multi-view analyses of gait features are meaningful.

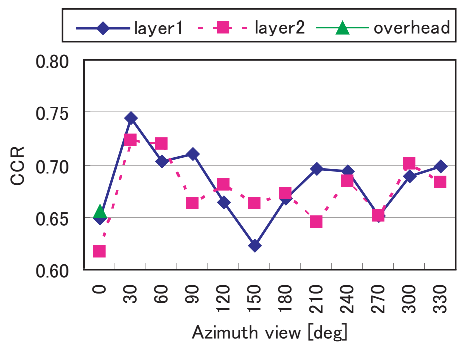
6 Feature Analysis of Gender and Age Classes

In this section, we analyze the uniqueness of gait for each class. Figure 5 illustrates the average gait features for each class for typical views (side, front, right-back, and overhead), where views from layer 1 cameras only are shown, since the elevation angle causes little difference as discussed in the previous section. Figure 6 shows a comparison of these features for combinations of classes used in the classification experiments depicted in Fig. 4(b) for the same views as in Fig. 5.

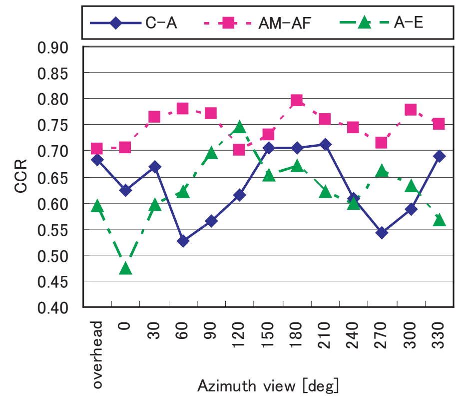
According to the above, the following trends are observed for each class.

Children

Posture, arm swing, and relative size of the head are specific features for children. Regarding posture, children tend to walk with their bodies tilted forwards, because they are often looking down at the ground, and this trend is distinctive in the side and overhead views. The captured area is larger in

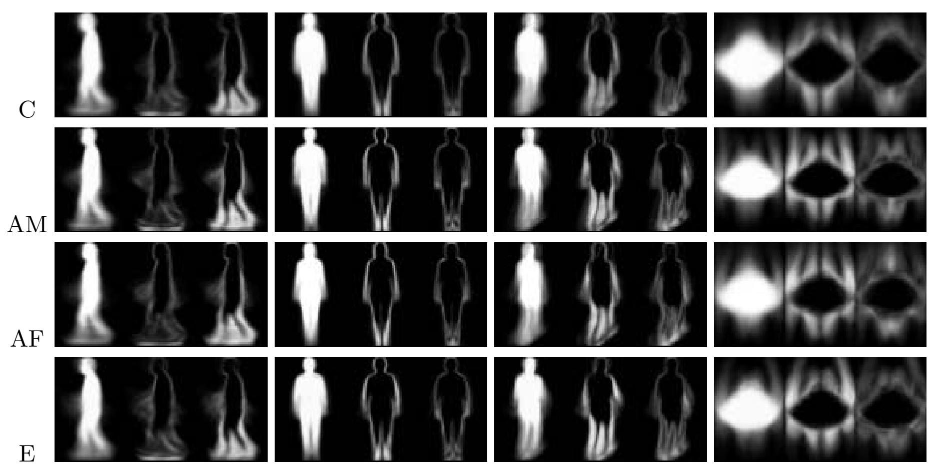


(a) CCR for each camera for the four classes



(b) CCR for each camera for specific combinations of classes

Fig. 4: Performance of gender and age classification for each view. The horizontal axis denotes the azimuth view from the subject, where front corresponds to 0 deg., right side to 90 deg., and so on. The vertical axis denotes CCR. In (a), the separate graphs depict the cameras in layer 1, layer 2, and the overhead camera. In (b), results of layer 1 and the overhead view are shown.



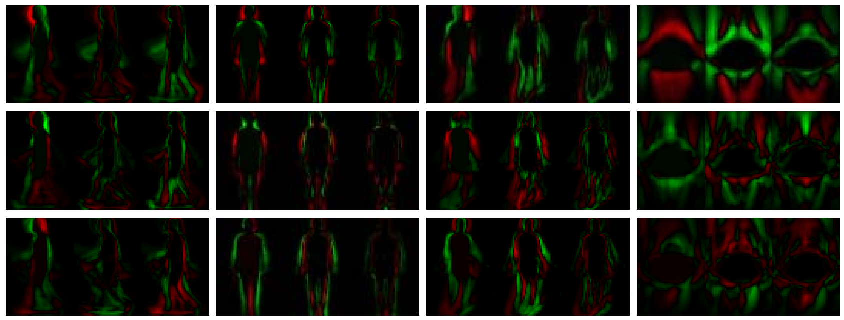
(a)left (b)front (c)right-back (d)overhead

Fig. 5: Average gait features for each class. The features are shown with their 1- and 2-times frequency multiplied 3 times for highlighting.

C-A

AM-AF

A-E



(a) Left (b) Front (c) Right-back (d) Overhead

Fig. 6: Differences in average features. Color is used to denote which class's feature appears more strongly. Red indicates that the feature of the leftmost class (e.g. C of C-A) appears more strongly, while green depicts the opposite. The features are shown with their 1- and 2-times frequency multiplied 3 times for highlighting.

the overhead view compared with that of a straight posture, while the tilted angle can easily be observed in the side view. The arm swings of children tend to be smaller than those of adults since walking in children is less mature. This feature is observed in the side and overhead views. Since the relative head size of children is larger than that of adults, children have larger heads and smaller bodies as a volume percentage against the whole body. These features are observed especially in the frontal views; arms appear lower, the size of the head area is larger and the body shape is wider.

Adult males and females

The main differences between the genders are observed in the body frame and stride. Regarding body frame, males have wider shoulders while females have more rounded bodies. These features are particularly noticeable in the frontal and side views. Differences are observed in the shoulder area in the frontal view, and in the breast and hair area in side views. Since males are taller than females on average, the walking stride of a male appears narrower in a side view when the images showing individuals walking at constant speed are normalized with respect to height. Related to physical strength, while the walking motion of adults is well balanced and more symmetrical, that of children and the elderly tends to be more asymmetrical. Therefore, adults have much lower values in the 1-time frequency in the side view, which represents asymmetry of gait, compared with children or the elderly. This asymmetry plays an important role in age classification.

The elderly

Body width, walking posture, and arm swings are distinctive features of the elderly. Due to middle-age spread, they have wider bodies than adults, which is clearly observed in the frontal view. This feature is, however, different from the similar feature in children. The elderly have wider bodies, but the size of their head is almost the same as in adults. Posture is discriminative in side views, which show a stooped posture. It is also observed that the elderly have larger arm swings. The reason seems to be that elderly subjects consciously swing their arms in the same way as brisk walkers exercising. In fact, the interviews confirmed that most of the elderly subjects do some form of daily exercise, notably walking.

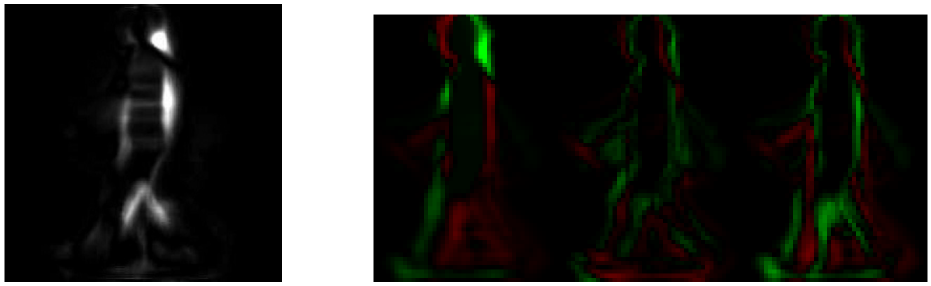
These trends also explain the classification results. According to Fig. 4(b), the performance of the overhead view for AM-AF and A-E is not as high as that of most other views. The reason is obvious from Fig. 5 in that gait features such as leg movement and posture, are not apparent in this view. However, in the children-adults classification, the relative performance of the overhead view to other views is better compared with the other classifications, as the dominant differences are the tilting posture and size of head. This indicates that the effect of the view on classification performance may differ depending on the class.

7 Discussion

Comparison with observations from previous works: Observations from previous works include results showing significant differences in the stride and breast area between genders [10], and that the difference in stride contributes to gender classification [17]. These observations are also evident in our research as shown in Fig. 6. Moreover, quite similar observations of gait feature differences between genders are reported in [11]. Yu et al. [11] analyzed the gender difference in Gait Energy Image (GEI) [20], corresponding to the 0-times component of our frequency-domain feature, and calculated the ANOVA F-statistic value between genders as shown in Fig. 7(a). The ANOVA F-statistic [30] is a measure for evaluating the discriminative capability of different features. The greater the F value, the better the discriminative capability. Because the difference at 0-times frequency (Fig. 7(b)) has similar properties to the F value of the GEI, a similar part of the image dominates in both approaches, that is, hair, back, breast area, and legs, which enhances the validity of our results. Note that our frequency-domain feature provides additional information on the dynamic component (1-time and 2-times frequency), as well as the static component, as shown in Fig. 7(b)).

In addition, we obtained further insights, such as view dependencies of classification performance and observed gait features, which previous works did not analyze. Consequently, our analyses are more comprehensive than those of previous works with respect to gender and age, dynamic components, and view.

Classification performance: The classification results show that the more diverse the subjects are in terms of gender and age, the more difficult the classification becomes. Though our classification results were 80% for AM-AF and 74% for C-A, the accuracy improves to 91% for AM-AF and 94% for C-A if the adult age range is limited to between 25 and 34 years. This improved accuracy correlates with that obtained for other state-of-the-art systems, 90% for AM-AF [9], 95.97% for AM-AF [11], and 93-95% for C-A [13], although the detailed



(a) F-statistic for GEI [11]

(b) Difference in frequency-domain features (ours)

Fig. 7: Comparison of differences between genders.

distribution and the number of subjects differ. This indicates that current methods could be improved to realize some feasible applications such as detecting lost children in airports and collecting customer statistics for stores.

Application of view dependency: Based on our research, CCRs obtained with individual single views differ each othera whole-view feature are superior to the average CCRs obtained with an individual single view. Thus, classification performance can be improved by selecting an optimal combination of views for real applications. For example, combining the overhead and side views will ensure good performance in the classification of children, whereas the installation of overhead cameras seems to be feasible for surveillance purposes, making it a good candidate for detecting lost children.

8 Conclusion

We have described video-based gait feature analysis for gender and age classification using a large-scale multi-view gait database. We started by constructing a large-scale multi-view gait database using a multi-view synchronous gait capturing system. The novelty of the database is not only its size, but also the fact that it includes a wide range of gender and age combinations (88 males and 80 females between 4 and 75 years old), and a large number of observed views (25 views). Gait features for four classes, namely children (younger than 15 years old), adult males (15 to 65 years old), adult females (15 to 65 years old) and the elderly (65 years and older), were analyzed in terms of views, and static and dynamic components, and insights regarding view changes and classification were acquired from a computer-vision point of view. A part of our gait database have been already published on a website [26] and new gait database will be added in the future

Future directions for this research are as follows:

- Modeling the aging effect to extend the application of gait, for example, to person identification across a number of years, and the synthesis of subjects' gait for any given age.

– Optimal combination of views for better gender and age classification considering the obtained view dependency on the classification.

Acknowledgement. This work was supported by Grant-in-Aid for Scientific Research(S) 21220003.

References

1. BenAbdelkader, C., Culter, R., Nanda, H., Davis, L.: Eigengait: Motion-based recognition people using image self-similarity. In: Proc. of Int. Conf. on Audio and Video-based Person Authentication. (2001) 284–294
2. Bobick, A., Johnson, A.: Gait recognition using static activity-specific parameters. In: Proc. of IEEE Conf. on Computer Vision and Pattern Recognition. Volume 1. (2001) 423–430
3. Cuntoor, N., Kale, A., Chellappa, R.: Combining multiple evidences for gait recognition. In: Proc. of IEEE Int. Conf. on Acoustics, Speech, and Signal Processing. Volume 3. (2003) 33–36
4. Horita, Y., Ito, S., Kanade, K., Nanri, T., Shimahata, Y., Taura, K., Otake, M., Sato, T., Otsu, N.: High precision gait recognition using a large-scale pc cluster. In: Proc. of the IFIP International Conference on Network and Parallel Computing. (2006) 50–56
5. Sarkar, S., Phillips, J., Liu, Z., Vega, I., Grother, P., Bowyer, K.: The humanid gait challenge problem: Data sets, performance, and analysis. Trans. of Pattern Analysis and Machine Intelligence **27** (2005) 162–177
6. Liu, Z., Sarkar, S.: Improved gait recognition by gait dynamics normalization. IEEE Transactions on Pattern Analysis and Machine Intelligence **28** (2006) 863–876
7. Su, F., Wu, W., Cheng, Y., Chou, Y.: Fuzzy clustering of gait patterns of patients after ankle arthorodesis based on kinematic parameters. Medical Engineering and Physics **23** (2001) 83–90
8. Liston, R., Mickelborough, J., Bene, J., Tallis, R.: A new classification of higher level gait disorders in patients with cerebral multi-infarct states. Age and Ageing **32** (2003) 252–258
9. Huang, G., Wang, Y.: Gender classification based on fusion of multi-view gait sequences. In: Proc. of the 8th Asian Conf. on Computer Vision. Volume 1. (2007) 462–471
10. Li, X., Maybank, S., Yan, S., Tao, D., Xu, D.: Gait components and their application to gender recognition. Trans. on Systems, Man, and Cybernetics, Part C **38** (2008) 145–155
11. Yu, S., Tan, T., Huang, K., Jia, K., Wu, X.: A study on gait-based gender classification. IEEE Trans. on Image Processing **18** (2009) 1905–1910
12. Begg, R.: Support vector machines for automated gait classification. IEEE Trans. on Biomedical Engineering **52** (2005) 828–838
13. Davis, J.: Visual categorization of children and adult walking styles. In: Proc. of the Int. Conf. on Audio- and Video-based Biometric Person Authentication. (2001) 295–300
14. Makihara, Y., Sagawa, R., Mukaigawa, Y., Echigo, T., Yagi, Y.: Gait recognition using a view transformation model in the frequency domain. In: Proc. of the 9th European Conf. on Computer Vision. Volume 3., Graz, Austria (2006) 151–163

15. Kozlowski, L., Cutting, J.: Recognizing the sex of a walker from a dynamic point-light display. Perception and Psychophysics **21** (1977) 575–580
16. Lee, L., Grimson, W.: Gait analysis for recognition and classification. In: Proc. of the 5th IEEE Conf. on Face and Gesture Recognition. Volume 1. (2002) 155–161
17. Yoo, J., Hwang, D., Nixon, M.: Gender classification in human gait using support vector machine. In: Advanced Concepts For Intelligent Vision Systems. (2006) 138–145
18. Liu, Y., Collins, R., Tsin, Y.: Gait sequence analysis using frieze patterns. In: Proc. of the 7th European Conf. on Computer Vision. Volume 2. (2002) 657–671
19. Veres, G., Gordon, L., Carter, J., Nixon, M.: What image information is important in silhouette-based gait recognition? In: Proc. of IEEE Conf. on Computer Vision and Pattern Recognition. Volume 2. (2004) 776–782
20. Han, J., Bhanu, B.: Individual recognition using gait energy image. Trans. on Pattern Analysis and Machine Intelligence **28** (2006) 316– 322
21. Sugiura, K., Makihara, Y., Yagi, Y.: Gait identification based on multi-view observations using omnidirectional camera. In: Proc. 8th Asian Conference on Computer Vision, Tokyo, Japan (2007) 452–461 LNCS 4843.
22. Weinland, D., Ronfard, R., Boyer, E.: Automatic discovery of action taxonomies from multiple views. In: Proc. of the 2006 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition. Volume 2., New York, USA (2006) 1639–1645
23. Shutler, J., Grant, M., Nixon, M., Carter, J.: On a large sequence-based human gait database. In: Proc. of the 4th Int. Conf. on Recent Advances in Soft Computing, Nottingham, UK (2002) 66–71
24. Yu, S., Tan, D., Tan, T.: A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition. In: Proc. of the 18th Int. Conf. on Pattern Recognition. Volume 4., Hong Kong, China (2006) 441–444
25. Nixon, M., Carter, J., Shutler, J., Grant, M.: Experimental plan for automatic gait recognition. Technical report, Southampton (2001)
26. : Ou-isir gait database. (http://www.am.sanken.osaka-u.ac.jp/GaitDB/index.html)
27. Lee, S.J., Hidler, J.: Biomechanics of overground vs. treadmill walking in healthy individuals. Journal of Applied Physiology **104** (2008) 747–755
28. Otsu, N.: Optimal linear and nonlinear solutions for least-square discriminant feature extraction. In: Proc. of the 6th Int. Conf. on Pattern Recognition. (1982) 557–560
29. Makihara, Y., Sagawa, R., Mukaigawa, Y., Echigo, T., Yagi, Y.: Which reference view is effective for gait identification using a view transformation model? In: Proc. of the IEEE Computer Society Workshop on Biometrics 2006, New York, USA (2006)
30. Brown, M.B., Forsythe, A.B.: The small sample behavior of some statistics which test the equality of several means. Technometrics **16** (1974) 129–132