EXPRESS PAPER

Directional characteristics evaluation of silhouette-based gait recognition

Yui Shigeki[1†], Fumio Okura[1*†], Ikuhisa Mitsugami[2], Kenichi Hayashi[3] and Yasushi Yagi[1]

Abstract

Gait is an important biometric trait for identifying individuals. The use of inputs from multiple or moving cameras offers a promising extension of gait recognition methods. Personal authentication systems at building entrances, for example, can utilize multiple cameras installed at appropriate positions to increase their authentication accuracy. In such cases, it is important to identify effective camera positions to maximize gait recognition performance, but it is not yet clear how different viewpoints affect recognition performance. This study determines the relationship between viewpoint and gait recognition performance to construct standards for selecting an appropriate view for gait recognition using multiple or moving cameras. We evaluate the gait features generated from 3D pedestrian shapes to visualize the directional characteristics of recognition performance.

Keywords: Silhouette-based gait recognition, Performance evaluation, Multi-camera gait recognition

1 Introduction

Individual identification is a core problem in the computer vision and biometrics fields. A particularly active area of study is the identification of individuals for security purposes using vision sensors placed in distant positions (e.g., security cameras). Face recognition [1] is a typical instance that achieves great success due to its practical recognition performance [2]. However, face recognition is notably difficult when the quality of facial features is insufficient, for example, due to the person looking down or to low-resolution input images.

Gait recognition [3], i.e., the identification of individuals from their walking styles, is another technique that offers promising solutions [4] for applications that use security cameras, which capture pedestrians from distant positions. In particular, silhouette-based gait features [5, 6], which use silhouettes of people walking, achieve state-of-the-art performance when using distant views (i.e., low-resolution images) as the input, compared with model-based approaches [7, 8] that fit human shape models. Because of its suitability for security camera applications, silhouette-based gait recognition is occasionally used for

forensics [9]. To improve performance, gait recognition can be combined with other biometric features [10] such as face images [11, 12].

One promising extension of gait recognition is to use inputs from multiple or moving cameras. Personal authentication systems at building entrances, for example, can use multiple cameras installed in appropriate positions to increase their authentication accuracy (such as "biometric tunnels" [13]). Similarly, moving cameras (e.g., cameras installed on drones) are expected to provide a novel type of security that can actively detect people behaving suspiciously by capturing pedestrians from different viewpoints, leveraging the recent developments in vision-based tracking [14]. In both cases, designating effective camera positions for gait recognition is required to maximize performance.

To date, a few studies have attempted to study the most effective method of selecting camera positions by comparing gait recognition performance from multiple viewpoints [15–18]. However, these studies used a limited number and variety of viewpoints (e.g., only horizontal views [15, 16, 18]), and it remains unclear how different viewpoints affect recognition performance.

This study aims to reveal the relationships among viewing direction, distance, and silhouette-based gait recognition performance using systematic experiments to construct standards for the selection of views for multi- or

*Correspondence: okura@am.sanken.osaka-u.ac.jp
†Yui Shigeki and Fumio Okura contributed equally to this work.
[1]The Institute of Scientific and Industrial Research, Osaka University, 8-1, Mihogaoka, Ibaraki, Osaka, Japan
Full list of author information is available at the end of the article

moving-camera for silhouette-based gait recognition. In particular, we focus gait energy image (GEI) [5], which is known as a simple yet effective feature for gait recognition. GEI is often used as a baseline in performance evaluations of silhouette-based gait recognition [4], and most appearance-based gait features (e.g. [6, 19, 20]) are designed as extensions of GEI. Because it is difficult to capture pedestrians from every direction and distance simultaneously, we used multiple cameras to reconstruct the 3D shapes of people walking to synthetically generate their gait features from various directions and distances. We then visualized the directional characteristics of the recognition performance by evaluating the generated gait features.

Given the directional characteristics, we can estimate the gait recognition performance from a given position; thus, these characteristics can be used for general purposes that involve the selection of camera positions for silhouette-based gait recognition. To demonstrate the practical application of the performance characteristics, we introduce a simple yet effective approach for person recognition that combines gait features observed from multiple viewpoints.

Contributions: The contribution of this paper is to investigate the effect of direction and distance on the performance of silhouette-based gait recognition through systematic experiments. Our study constructs a standard for selecting the view for multi- or moving-camera gait recognition using GEIs.

2   Multi-camera gait recognition

The directional characteristics of gait recognition proposed in this paper are intended to be used for practical gait recognition scenarios using multiple or moving cameras. We therefore first introduce a particular multi-camera gait recognition scenario.

Multi-camera gait recognition is intended to be used in practical scenarios such as the authentication of people at building entrances. The cameras are installed at multiple locations to simultaneously capture pedestrian video sequences. Similar to the traditional gait recognition techniques using security cameras, by registering gait features from similar viewpoints beforehand, the gait features extracted from multi-camera images can be used for authentication tasks. Moving-camera gait recognition is an asynchronous version of multi-camera recognition. A moving platform such as a drone captures multiple video sequences from multiple viewpoints while hovering and changing the position of the camera.

Open problems still need to be solved to achieve gait recognition in a practical environment, where pedestrians are often placed against complex backgrounds and occlusions, and display a large variation in walking styles. Pedestrian detection, tracking, and segmentation are common fundamental problems in silhouette-based gait recognition. To apply our directional characteristics, the direction of the pedestrian has to be estimated in a preprocessing stage. In addition, pedestrians' walking speed, baggage, clothes, and other features change their silhouettes.

State-of-the-art human detection approaches [21, 22], which include processes for estimating the person's position, direction, and pose, can be applied in the preprocessing step of gait recognition in a practical environment. Approaches for estimating human body shapes [23] and semantic segmentation [24] can facilitate the automatic segmentation of pedestrians for appearance-based gait recognition using pedestrian silhouettes. Meanwhile, we can employ depth cameras to acquire accurate segmentation for some cases. In addition, recent studies on gait recognition, which tackle practical problems such as speed transition [25], clothing [26], and baggage [27], show promising performance when pedestrian silhouettes are available.

On these state-of-the-art techniques, this study assumes that the preprocessing stage can be performed properly to acquire pedestrian silhouettes. Once the pedestrian silhouettes and their direction have been obtained, the directional characteristics constructed in this study provide a useful measure for designing camera positions.

3   Experimental settings

3.1   Gait feature generation

We generate gait features viewed from various directions and distances to investigate the directional characteristics of silhouette-based gait recognition. This is a challenging task when using physical cameras; thus, we apply a semi-synthetic approach using the 3D shapes of pedestrians. The 3D shapes are reconstructed as visual hulls using videos captured by 24 synchronized cameras installed around a treadmill by Muramatsu et al. [28] and converted to surface models by Ikeda et al. [29].

Silhouette sequences of pedestrians are computed by projecting the 3D shapes onto virtual cameras located at various viewpoints around the 3D shapes. The position of the virtual cameras is described by a vertical angle θ, a horizontal angle α, and a distance d (see Fig. 1). The view direction is designated so that the optical axis faces the pedestrian. GEIs are then computed as an averaged image of the height-normalized silhouette sequences created from 3D shapes during one walking period.

3.2   Omnidirectional gait feature dataset

We used the 3D pedestrian shapes of 97 subjects collected in [28]. The sequences from three walking periods are available for each subject (291 periods in total). A GEI is
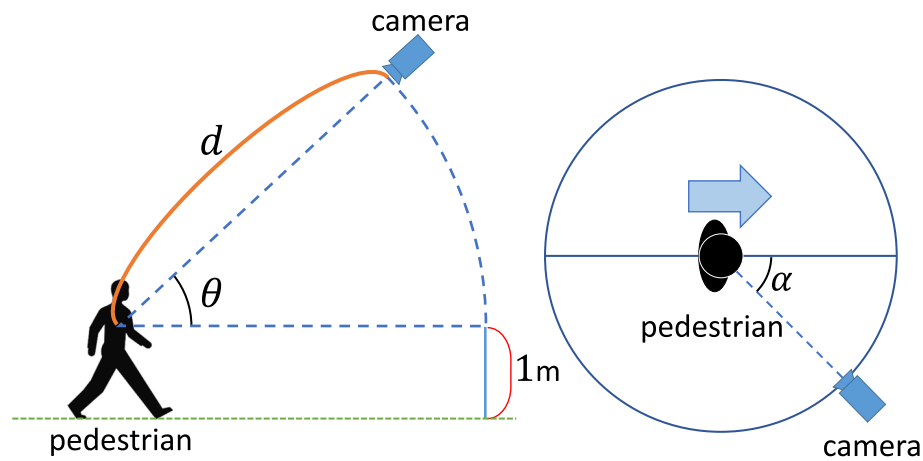
Fig. 1 Definition of virtual camera position. θ, α, and d denote vertical angles, horizontal angle, and distance, respectively

generated for each walking period from each subject by changing the position of the virtual cameras as follows:

- Vertical angle θ from 0° to 90° at 10° intervals.
- Horizontal angle α from 0° to 350° at 10° intervals.
- Distance d of 5, 10, 20, and 40 m, where the resolution of the virtual cameras is 400×300 pixels and the vertical field of view (FOV) is 53°.

A total of 1440 virtual cameras are generated from 36 horizontal, 10 vertical, and 4 distance variations. We assume a weak central projection at a distance of $d = 5$ m. Virtual views at distances of $d = 10, 20, 40$ m are therefore obtained by lowering the resolution of the silhouette sequences at $d = 5$ m. Figure 2 illustrates the relationship between the distance d and the silhouette resolution when the vertical angle is 0°.

	GEI	Silhouette height at $\theta = 0°$ [px]
$d = 5$ m		98
$d = 10$ m		49
$d = 20$ m		24
$d = 40$ m		12

Fig. 2 Distance d and silhouette resolution. Relationship between the distance and the silhouette resolution when the vertical angle is 0°

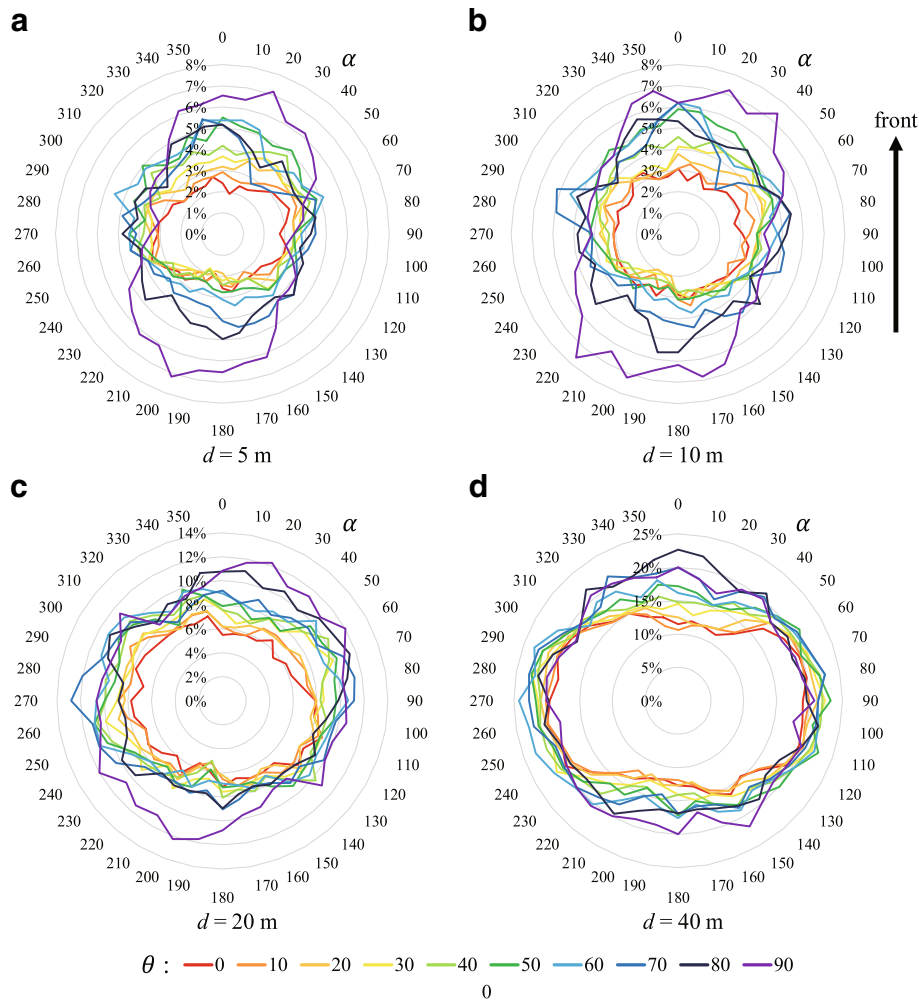**Fig. 3** Relationships among EER (%), viewing direction (α and θ), and distance d. The scale of the legend changes among the charts because the performance varies widely depending on the distance. **a** d = 5 m. **b** d = 10 m. **c** d = 20 m. **d** d = 40 m

4   Directional characteristics of gait recognition

Using gait features viewed from various viewpoints, we evaluated the recognition performance to ascertain the direction- and distance-related characteristics.

4.1   Evaluation method

According to the convention, we computed the equal error rate (EER) for visual recognition performance. The EER is obtained from the false rejection rate (FRR) and the false acceptance rate (FAR), which are calculated from a dissimilarity matrix among gait features. For each viewpoint, we created a dissimilarity matrix based on the L2 norm distance between pairs among 291 GEIs (3 walking periods for every 97 subjects). The EERs for all 1440 viewpoints were obtained from the distance matrices.

4.2   Results

Figure 3 visualizes the change in the EER with various view directions at distances of d = 5, 10, 20, and 40 m, where a smaller EER indicates better performance. Table 1 summarizes the EERs at three vertical angles θ = 0°, 50°, and

Table 1 Summary of the directional characteristics of recognition performance (distance d = 5 m)

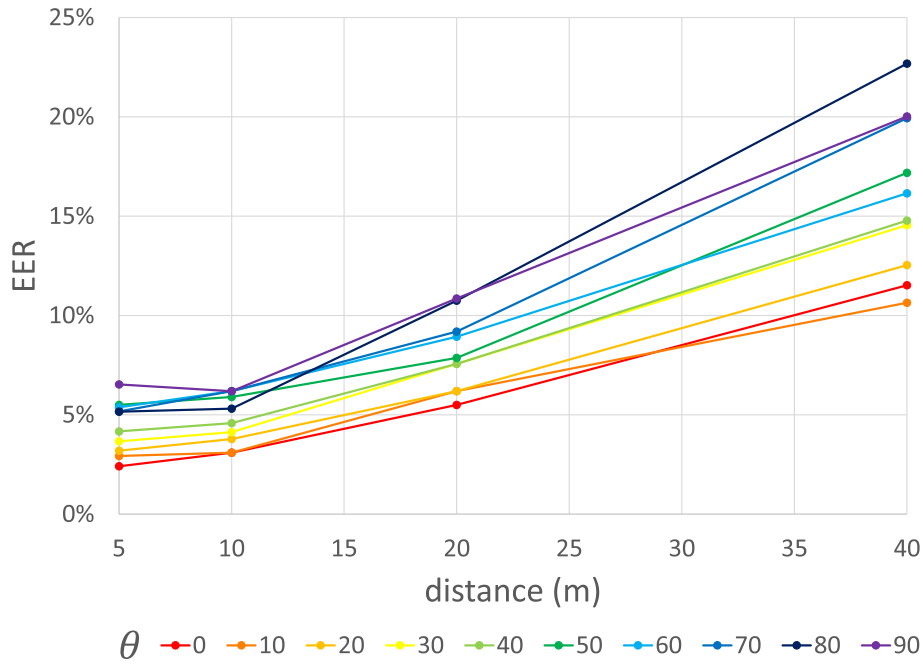| Vertical angle	Horizontal angle		
	α = 0°	α = 90°	α = 180°
	(front)	(side)	(back)
θ = 0° (horizontal view)	2.40	2.75	2.56
θ = 50°	5.50	3.34	2.75
θ = 90° (top view)	6.53	3.44	6.52

Fig. 4 Relationship between EER (%) and distance. Horizontal angle is fixed at $\alpha = 0°$ (front view)

90° at $d = 5$ m. The visualizations illustrate the following tendencies:

- Small vertical angles θ (e.g., 0°): better performance is achieved when the horizontal angle α is approximately 0° or 180° (i.e., front or back view).
- Medium θ (e.g., 50°): better performance is achieved when α is approximately 180° (i.e., captured from the back).
- Large θ (e.g. 90°): better performance is achieved when α is approximately 90°, despite a deterioration in overall performance.

We also visualize the relationship between the EER and distance in Fig. 4. The change in performance is small when the viewpoints are closer than $d = 10$ m. Beyond $d = 20$ m, the deterioration in performance is approximately linear to the distance.

4.3 Discussion
4.3.1 Causes of directional characteristics
We further investigate the directional characteristics of recognition performance, which show different tendencies among viewing directions. Example GEIs at small ($\theta = 0°$) and medium ($\theta = 50°$) vertical angles are shown in Figs. 5 and 6. The differences between the front ($\alpha = 0°$) and back ($\alpha = 180°$) views are small when the vertical angle is small (cf. Fig. 5). At a medium vertical angle, however, the GEIs from the back view (Fig. 6b) show the head shape more clearly than the front view (Fig. 6a). This is considered to be caused by a characteristic of the
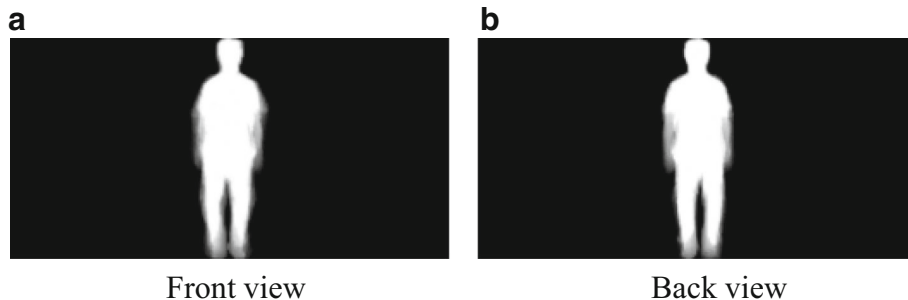

Fig. 5 Differences between GEIs from front and back views ($\theta = 0°$). The differences between the front ($\alpha = 0°$) and back ($\alpha = 180°$) views are small when the vertical angle is small. **a** Front view. **b** Back view
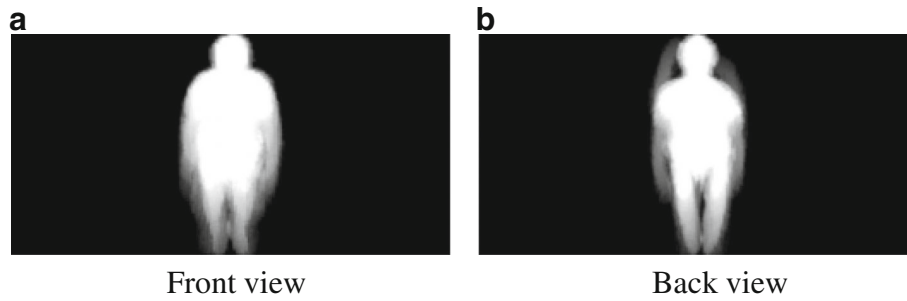
Fig. 6 Differences between GEIs from front and back views ($\theta = 50°$). The GEIs from the back view (**b**) show the head shape more clearly than the front view (**a**)

human shape, which means that some people walk with a stoop. In this case, better performance is achieved using a back view at a medium vertical angle.

The recognition performance in top-view ($\theta = 90°$) situations varies due to the height normalization of the silhouette sequences during GEI computation. A GEI is generated by averaging silhouette sequences, after preliminary normalization by the height of the silhouette to the alignment. However, this process has unexpected results when using top views, as shown in Fig. 7. When the horizontal angle α is 0° (see Fig. 7a), the silhouettes of the arms and legs move up and down in the image plane. The unstable GEIs caused by the changes in the height of the silhouette while walking result in worse performance when α is 0° and 180°.

Resolution: State-of-the-art security cameras are equipped with high-resolution cameras. Although our experiment used virtual cameras with a 400 × 300-pixel resolution and a 50° FOV, we can interpret our experimental results as equivalent to situations using high-resolution security cameras, where we assume a wide-angle camera with 4K resolution and a 90° vertical

FOV. According to Fig. 4, when $d \leq 10$ m, the EER is less than 5% if we select the capturing altitude so that $\theta \leq 40°$. Assuming a weak central projection, the gait features obtained by the virtual cameras at a distance of $d = 10$ m are equivalent to those obtained from approximately 20 m by the high-resolution cameras. Similarly, when $d = 20$ m (equivalent to 40 m distance in the high-resolution camera), the EER is less than 8% if $\theta \leq 50°$.

Accuracy of 3D pedestrian models: This study used 3D pedestrian models, which were originally created in [28], to model the view-dependent transformation of gait features and to generate gait features from various viewpoints. Although it is difficult to evaluate the accuracy of 3D shapes directly because of the difficulty of acquiring the ground truth person shapes, we evaluated the error metrics related to the reliability of the 3D shapes. The 3D models are visual hulls created by the volume intersection method, in which the shape errors occur in concave areas. The concave areas vary in size according to the human pose, e.g., concave areas are large when the person is standing straight and small when the person is spreading his arms. We investigated the effect of the concave
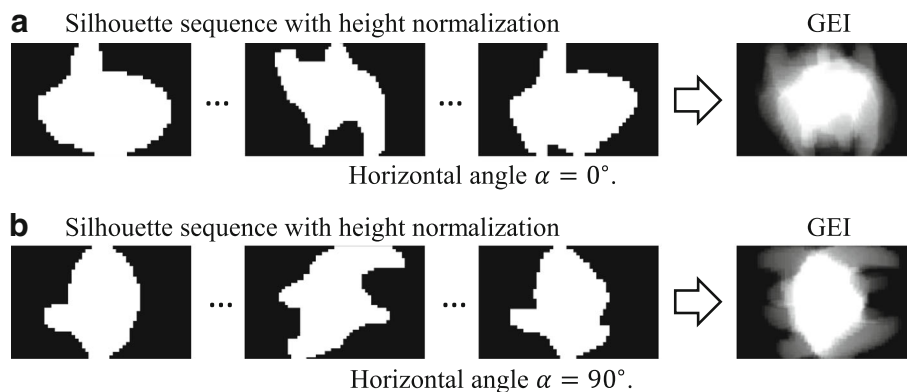

Fig. 7 Effect of height normalization. The effect of height normalization generating GEIs from top views (i.e., vertical angle $\theta = 90°$) sequences. **a** Horizontal angle $\alpha = 0°$. **b** Horizontal angle $\alpha = 90°$

Shigeki *et al. IPSJ Transactions on Computer Vision and Applications* (2018) 10:10

Page 7 of 10

areas by measuring the volume sizes in the frames in each 3D video. The average standard deviation of the 3D volume size (i.e., the number of occupied volumes) for each sequence was 1.77% over the average volume size. Assuming that the error is uniformly distributed, it leads to 0.59% error in the length of a dimension such as the height, width, and depth. Because this is less than one pixel in the images captured from 5 m in our experiment, the 3D models we used have good reliability.

Comparison with real dataset: The experiment was carried out using semi-synthetic dataset. Therefore, there is a possibility that our experiment does not fit the actual environment. Although they do not cover the entire direction, several datasets (e.g., [18]) provide multi-view videos of pedestrians. We, therefore, performed a comparison of our experimental result with an actual dataset (OU-MVLP dataset [18]) to evaluate the validity of our experiment. Figure 8 shows the EER from OU-MVLP dataset and our result from a similar view direction. While the absolute values of EER are different because the number of subjects included in the datasets is different, the tendency is similar (i.e., the accuracy is better around $\alpha = 30°$ and worse around $\alpha = 60°$). Not only when using the same distance metric to our experiment (L2-norm between GEIs), the tendency is consistent when using state-of-the-art distance metric optimized by deep neural networks [30]. This result indicates the validity of our experiment based on simulation dataset.

4.4 Multi-camera gait recognition
Designating an effective layout for multiple cameras for gait recognition is straightforward using the directional performance characteristics. Personal authentication systems at building entrances, for example, can use multiple cameras installed at fixed positions to increase their

Fig. 8 Comparison with a real dataset [18]. The recognition accuracy in our experiment shows a similar tendency with an existing dataset capturing actual pedestrians

authentication accuracy. Similarly, moving cameras such as drones can obtain gait features from multiple viewpoints as they move around.

As a proof of concept, we performed a brief multi-view gait recognition experiment to demonstrate the performance improvement. The dissimilarities D_i between the gait features captured from multiple positions are calculated for each position, where i denotes the i-th viewpoint. Assuming a simple case, we obtain an overall dissimilarity value by summing the multiple dissimilarities $D = \sum_i D_i$.

Figure 9 illustrates the performance improvement at two vertical angles $\theta = 0°$ and $70°$ and two distances $d = 20$ and 5, after combining the GEIs obtained from two horizontal angles, denoted as α_1 and α_2. We calculate the EER of the two-view recognition (denoted as EER_c) by summing the dissimilarity matrices of the two horizontal angles. Denoting the EERs for single-view recognition at α_1 and α_2 as EER_1 and EER_2, respectively, we calculate the improvement in Δ_{EER} as:

$$\Delta_{\text{EER}} = \min(\text{EER}_1, \text{EER}_2) - \text{EER}_c. \tag{1}$$

When the vertical angle is small (cf. Fig. 9a), the performance improvement is large when two gait features are captured from perpendicular directions (dashed lines in the figure). Conversely, when the vertical angle is large (cf. Fig. 9b), combined front and back views achieve a notable improvement (the circle in the figure).

4.5 Summary and application scenarios
In this section, we summarize the insights yielded by our experiment and introduce practical application scenarios. Table 2 summarizes the best single- and two-view gait recognition performance for each vertical angle.

For single-camera gait recognition, the suggested camera locations are as follows.

- Small vertical angle (e.g., $\theta = 0°$): capture from the front or back. The best performance was EER = 1.92% at $\theta = 0°$ znd $\alpha = 190°$, while the performance at $\alpha = 10°$ was almost equivalent (EER = 1.93%).
- Middle vertical angle (e.g., $\theta = 40°$): capture from the back. The best performance was EER = 1.82% at $\theta = 40°$ and $\alpha = 200°$.
- Large vertical angle (e.g., $\theta = 70°$): capture diagonally from the front. The best performance was EER = 2.75% at $\theta = 70°$ and $\alpha = 30°$.
- Top view (e.g., $\theta = 90°$): capture from the side if height normalization is applied for feature calculation. The best performance was EER = 3.09% at $\theta = 90°$ and $\alpha = 100°$.

Regarding the distance, there was no notable drop in recognition accuracy when the silhouette height of the pedestrians was around 50 pixels or more when we
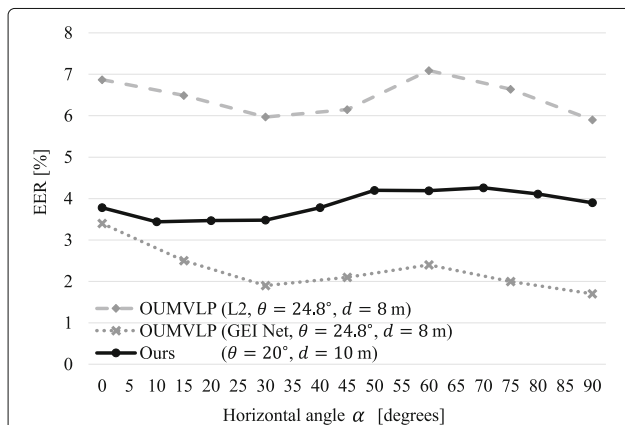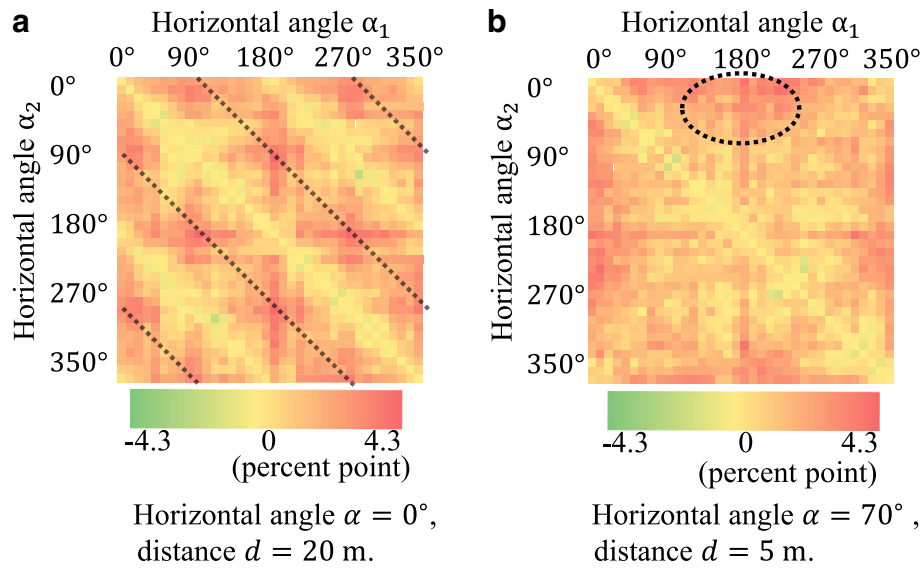
Shigeki *et al. IPSJ Transactions on Computer Vision and Applications* (2018) 10:10

Page 8 of 10

Fig. 9 Performance improvement (i.e., decrease in EER) for two-view gait recognition. Larger values indicate greater improvement (e.g., highlighted by dashed lines and a circle). **a** Horizontal angle $\alpha = 0°$, distance $d = 20$ m. **b** Horizontal angle $\alpha = 70°$, distance $d = 5$ m

used a simple (GEI + L2 distance-based) gait recognition approach.

For two-camera scenarios, we obtained the following insights.

- Low vertical angle: combine views from perpendicular directions, e.g., front/back and side. The best performance was EER = 1.58% achieved at $\theta = 0°$ by combining $\alpha = 190°$ and 280°.
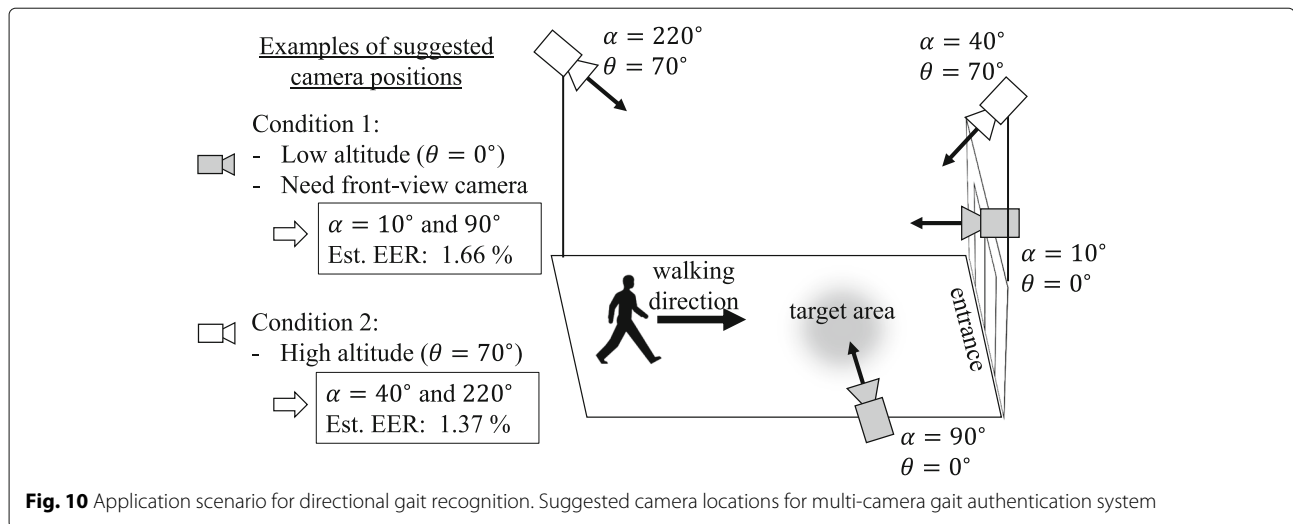- Middle vertical angle: combine two diagonal back views. The best performance was EER = 1.52% achieved at $\theta = 40°$ by combining $\alpha = 190°$ and 210°.
- Large vertical angle: combine diagonal front and back views. The best performance was EER = 1.37% achieved at $\theta = 70°$ by combining $\alpha = 40°$ and $\alpha = 220°$.
- Top views: the combination does not affect performance because the same views are acquired.

Table 2 Best single- and two-view gait recognition performance

Vertical angle	$\theta = 0°$	$\theta = 10°$	$\theta = 20°$	$\theta = 30°$
Best EER	1.92%	1.98%	1.78%	1.72%
(Single view)	($\alpha = 190°$)	($\alpha = 190°$)	($\alpha = 190°$)	($\alpha = 200°$)
Best EER	1.58%	1.72%[†]	1.37%	1.41%
(Two views)	($\alpha = 190°, 280°$)	($\alpha = 190°, 280°$)	($\alpha = 190°, 230°$)	($\alpha = 170°, 200°$)
Vertical angle	$\theta = 40°$	$\theta = 50°$	$\theta = 60°$	
Best EER	1.82%	2.38%	2.75%	
(Single view)	($\alpha = 200°$)	($\alpha = 200°$)	($\alpha = 200°$)	
Best EER	1.52%	2.06%[†]	2.41%[†]	
(Two views)	($\alpha = 190°, 210°$)	($\alpha = 120°, 200°$)	($\alpha = 110°, 190°$)	
Vertical angle	$\theta = 70°$	$\theta = 80°$	$\theta = 90°$	
Best EER	2.75%	3.44%	3.09%	
(Single view)	($\alpha = 30°$)	($\alpha = 310°$)	($\alpha = 100°$)	
Best EER	1.37%	2.19%	3.09%[†]	
(Two views)	($\alpha = 40°, 220°$)	($\alpha = 10°, 140°$)	($\alpha = 100°, 100°$)	

Note "†" denotes that several combinations yield the same accuracy ($d = 5$ m)

Fig. 10 Application scenario for directional gait recognition. Suggested camera locations for multi-camera gait authentication system

When using three or more cameras, the recognition performance and its improvements can easily be obtained from the dissimilarity matrices.

By leveraging the directional recognition characteristics, we can systematically design the camera locations for multi-camera gait recognition applications. We provide a concrete example of an application scenario in which we design the camera positions for a gait authentication system using two cameras in front of a building entrance. In a practical situation, we need to consider several requirements for the camera setting, e.g., where can we physically install the cameras? As shown in Fig. 10, centering on the target area where the cameras shoot the pedestrian, we can observe the optimal set of camera positions while taking other factors such as physical restrictions into account. If it is necessary to install a front-view camera, our results suggest that it is better to install the second camera perpendicular to the original camera (cf. condition 1 in Fig. 10). It is naturally difficult to install cameras at the same height as pedestrians. If we have to install the cameras above $\theta = 70°$, our results suggest installing two cameras at $\alpha = 40°$ and 210° (cf. condition 2 in Fig. 10). Because we can estimate the recognition errors for a multi-camera input, we can also evaluate how many cameras are required (and where to locate them) to achieve sufficient recognition accuracy for the specific application. Once the cameras are installed, the system requires the user to be captured by the cameras in gallery sequences, similar to other biometric authentication systems. Using the gallery sequences, the system performs the authentication tasks.

5 Conclusion

This paper has described the directional performance characteristics of silhouette-based gait recognition for developing standards for GEI-based gait recognition using multiple or moving cameras. We found that the EERs of gait recognition based on GEIs varied with the horizontal view direction; moreover, performance varied notably in the vertical direction due to the characteristics of the human body. Given the performance characteristics, we proposed a view selection scheme for multi-camera gait recognition using GEIs. We plan to develop novel security applications such as drone-view gait recognition and to investigate unknown tendencies in biometrics. Together with the recent progress in human tracking and gait recognition research, we firmly believe that the directional characteristics of silhouette-based gait recognition will be helpful for future security applications.

Authors' contributions
YS played the key role in designing the experiments. FO conducted the experiments and mainly wrote and edited the paper. The first two authors contributed equally to this study. IM supported the experiments and played an important role in editing the paper. KH contributed a key idea to the research design. YY played an important role in editing the paper. All authors have read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
[1] The Institute of Scientific and Industrial Research, Osaka University, 8-1, Mihogaoka, Ibaraki, Osaka, Japan. [2] Graduate School of Information Sciences,

Shigeki *et al. IPSJ Transactions on Computer Vision and Applications* (2018) 10:10

Page 10 of 10

Hiroshima City University, 3-4-1, Ozuka-Higashi, Asa-Minami-Ku, Hiroshima, Hiroshima, Japan. [3]Qoncept, Inc., 3-8-1, Nishikanda, Chiyoda, Tokyo, Japan.

References

1. Turk MA, Pentland AP (1991) Face recognition using eigenfaces. In: Proc. 1991 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR'91). IEEE. pp 586–591
2. Schroff F, Kalenichenko D, Philbin J (2015) FaceNet: A unified embedding for face recognition and clustering. In: Proc. 2015 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR'15). IEEE. pp 815–823
3. Nixon MS, Tan T, Chellappa R (2010) Human identification based on Gait. Springer Science & Business Media, NY
4. Iwama H, Okumura M, Makihara Y, Yagi Y (2012) The OU-ISIR gait database comprising the large population dataset and performance evaluation of gait recognition. IEEE Trans Inf Forensic Secur 7(5):1511–1521
5. Man J, Bhanu B (2006) Individual recognition using gait energy image. IEEE Trans Pattern Anal Mach Intell 28(2):316–322
6. Lam TH, Cheung KH, Liu JN (2011) Gait flow image: a silhouette-based gait representation for human identification. Pattern Recognit 44(4):973–987
7. Urtasun R, Fua P (2004) 3D tracking for gait characterization and recognition. In: Proc. Sixth IEEE Int'l Conf. on Automatic Face and Gesture Recognition (FG'04). IEEE. pp 17–22
8. Cunado D, Nixon MS, Carter JN (2003) Automatic extraction and description of human gait models for recognition purposes. Comp Vision Image Underst 90(1):1–41
9. Bouchrika I, Goffredo M, Carter J, Nixon M (2011) On using gait in forensic biometrics. J Forensic Sci 56(4):882–889
10. Kimura T, Makihara Y, Muramatsu D, Yagi Y (2014) Quality-dependent score-level fusion of face, gait, and the height biometrics. IPSJ Trans Comput Vis Appl 6(3):53–57
11. Jung SU, Nixon MS (2010) On using gait biometrics to enhance face pose estimation. In: Proc. Fourth IEEE Int'l Conf. on Biometrics: Theory Applications and Systems (BTAS'10). IEEE. pp 1–6
12. Shakhnarovich G, Lee L, Darrell T (2001) Integrated face and gait recognition from multiple views. In: Proc. 2001 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR'01), vol. 1. IEEE
13. Seely RD, Samangooei S, Lee M, Carter JN, Nixon MS (2008) The University of Southampton Multi-Biometric Tunnel and introducing a novel 3D gait dataset. In: Proc. Second IEEE Int'l Conf. on Biometrics: Theory, Applications and Systems (BTAS'08). IEEE. pp 1–6
14. Canals R, Roussel A, Famechon JL, Treuillet S (2002) A biprocessor-oriented vision-based target tracking system. IEEE Trans Ind Electron 49(2):500–506
15. Huang X, Boulgouris NV (2008) Human gait recognition based on multiview gait sequences. EURASIP J Adv Sig Process 2008(1):629102
16. Yu S, Tan D, Tan T (2006) Modelling the effect of view angle variation on appearance-based gait recognition. In: Proc. Seventh Asian Conf. on Computer Vision (ACCV'06). Springer. pp 807–816
17. Sarkar S, Phillips PJ, Liu Z, Vega IR, Grother P, Bowyer KW (2005) The humanID gait challenge problem: data sets, performance, and analysis. IEEE Trans Pattern Anal Mach Intell 27(2):162–177
18. Takemura N, Makihara Y, Muramatsu D, Echigo T, Yagi Y (2018) Multi-view large population gait dataset and its performance evaluation for cross-view gait recognition. IPSJ Trans Comput Vis Appl 10(1):4
19. Muramatsu D, Shiraishi A, Makihara Y, Uddin MZ, Yagi Y (2015) Gait-based person recognition using arbitrary view transformation model. IEEE Trans Image Proc 24(1):140–154
20. Muramatsu D, Makihara Y, Yagi Y (2016) View transformation model incorporating quality measures for cross-view gait recognition. IEEE Trans Cybern 46(7):1602–1615
21. Wei SE, Ramakrishna V, Kanade T, Sheikh Y (2016) Convolutional pose machines. In: Proc. 2016 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR'16). IEEE. pp 4724–4732
22. Cao Z, Simon T, Wei SE, Sheikh Y (2017) Realtime multi-person 2D pose estimation using part affinity fields. In: Proc. 2017 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR'17). IEEE
23. Bogo F, Kanazawa A, Lassner C, Gehler P, Romero J, Black MJ (2016) Keep it SMPL: Automatic estimation of 3D human pose and shape from a single image. In: Proc. 2016 European Conf. on Computer Vision (ECCV'16). Springer. pp 561–578
24. Lin G, Milan A, Shen C, Reid I (2017) RefineNet: Multi-path refinement networks with identity mappings for high-resolution semantic segmentation. In: Proc. 2017 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR'17). IEEE
25. Xu C, Makihara Y, Li X, Yagi Y, Lu J (2016) Speed invariance vs. stability: cross-speed gait recognition using single-support gait energy image. In: Proc. 2016 Asian Conf. on Computer Vision (ACCV'16). Springer. pp 52–67
26. Li X, Makihara Y, Xu C, Muramatsu D, Yagi Y, Ren M (2016) Gait energy response function for clothing-invariant gait recognition. In: Proc. 2016 Asian Conf. on Computer Vision (ACCV'16). Springer. pp 257–272
27. Makihara Y, Suzuki A, Muramatsu D, Li X, Yagi Y (2017) Joint intensity and spatial metric learning for robust gait recognition. In: Proc. 2017 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR'17). IEEE
28. Muramatsu D, Shiraishi A, Makihara Y, Yagi Y (2012) Arbitrary view transformation model for gait person authentication. In: Proc. IEEE Fifth Int'l Conf. on Biometrics: Theory, Applications and Systems (BTAS'12). IEEE. pp 85–90
29. Ikeda T, Mitsugami I, Yagi Y (2015) Depth-based gait authentication for practical sensor settings. IPSJ Trans Comput Vis Appl 7:94–98
30. Takemura N, Makihara Y, Muramatsu D, Echigo T, Yagi Y (2018) On input/output architectures for convolutional neural network-based cross-view gait recognition. IEEE Trans Circ Syst Video Technol 28(1)