Check for updates

Analysis of Histogram of Oriented Gradients on Gait Recognition

Chirawat Wattanapanich[1], Hong Wei[1](✉), and Wei Xu[2]

[1] University of Reading, Reading RG6 6AY, UK
wchirawat@wu.ac.th, h.wei@reading.ac.uk
[2] University of Surrey, Guildford GU2 7XH, UK
w.xu@surrey.ac.uk

Abstract. This study investigates the impact of Histogram of Oriented Gradients (HoG) on gait recognition. HoG is applied to four basic gait representations, *i.e.* Gait Energy Image (GEI), Gait Entropy Image (GEnI), Gait Gaussian Image (GGI), and newly developed Gait Gaussian Entropy Image (GGEnI). Hence their corresponding secondary gait representations, Gradient Histogram Gait Images (GHGI), are generated. Due to the nature of HoGs, the secondary gait representations contain rich information of images from different scales and orientations. The optimized HoG parameters are investigated to establish appropriate parameter settings in the HoG operations. Evaluations are conducted by using Support Vector Machines (SVM) as classifier swith CASIA dataset B. Experimental results have shown that HoG associated secondary representations are superior to the originally basic representations in gait recognition, especially in case of coping with appearance changes, such as walking with bag and walking with coat when using normal walking samples in training. GHGIs have increased the gait recognition rate approximately of 17% to GEI, 12% to GEnI, 24% to GGI and 20% to GGEnI.

Keywords: Gait representation · Histogram of oriented gradients · Model-free gait recognition

1 Introduction

Gait recognition, which is non-intrusive in identifying individuals in distance, is still challenging in biometric research. The main advantages of gait recognition are that it allows low-resolution images in recognition, can be used for long-distance detection, and has non-interference with target activities. Moreover, gait which is the personal walking characteristic is hardly hidden and spoofed.

Gait representation which presents personal gait information is one of the most important parts in gait recognition research. There are two main stages in conventional gait recognition, gait feature extraction and classification. Gait features which represent the walking characteristic can be extracted from both gait model and gait image sequence. In a model free approach, gait features are usually extracted from gait representation called compact image which is generated from a complete gait cycle. The basic compact

© Springer Nature Switzerland AG 2021
C. Djeddi et al. (Eds.): MedPRAI 2020, CCIS 1322, pp. 86–97, 2021.
https://doi.org/10.1007/978-3-030-71804-6_7

image, called Gait Energy Image (GEI) [1] or Average Silhouette [2] can be generated by an average function. GEI has been commonly used in the model free research because of its simplicity and low-cost in computation. Based on GEI, other gait compact images have been consequently established to fulfill recognition efficiency, such as Gait Entropy Image (GEnI), Gait Gaussian Image (GGI) [3], Flow Histogram Energy Image [4], Gradient Histogram Gaussian Image [5] and Gait Information Image [6]. This study has chosen GEI, GEnI, GGI and GGEnI (a Gaussian variant of GEnI) as four basic gait representations that are used to generate Gradient Histogram Gait Image (GHGI) by the Histogram of Oriented Gradients (HoG) method.

Various feature extraction methods have been used with compact images in gait recognition, such as Principal Component Analysis (PCA) [7, 8], Linear Discriminant Analysis (LDA) [9, 10] and Convolutional Neural Network (CNN) [11–14]. PCA, as a fundamental method, is chosen for dimension reduction or feature extraction in this study.

In the classification stage,various classifiers are adapted in gait recognition, such as Nearest Neighbor (NN) [3, 15, 16], Support Vector Machine (SVM) [17–19] and CNN. This study has chosen one-against-all multi-class SVM as the classifier.

GEI, GEnI, GGI and GGEnI are used as the basic gait representation image, from which Gradient Histogram Gait Images (GHGI) are generated by applying HoGs to them, followed by gait feature extraction by PCA. The classification performance is tested by multi-class SVMs on CASIA dataset B [20], that contains three types of appearance and eleven camera view angles. It aims to investigate how parameters of HoGs impact on gait recognition performance with GHGI gait representations. The rest of the paper is organized as follows. Section 2 presents the methodology for the gait recognition system. Section 3 discusses experiments and results. The conclusion is given in Sect. 4.

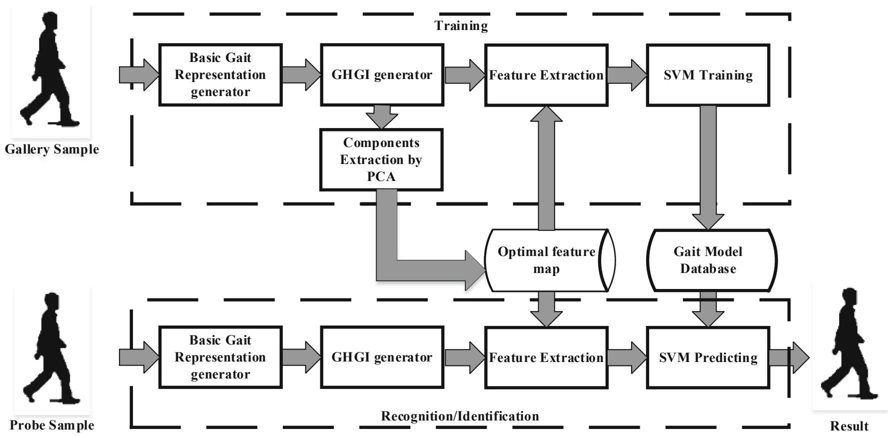
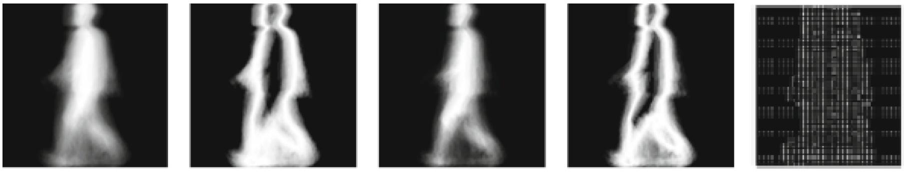


Fig. 1. General gait recognition system

2 Methodology

The gait recognition framework is shown in Fig. 1. It is a model-free gait recognition system that normally has two modes, training and testing. The training mode has four processes involved, sequence image preparation, gait representation generation, gait feature extraction, and personal model training. The testing mode comprises the similar processes as those in the training mode except that the last process is classifier's prediction. In this study, CASIA dataset B provides both training and testing images which are already processed by foreground extraction.

The study is focused on the investigation of the impact of gait representations on gait recognition performance. Four basic gait representations, GEI, GEnI, GGI and GGEnI, along with the corresponding secondary gait representations, Gradient Histogram Gait Image (GHGI) are investigated. Exemplar gait representation images are shown in Fig. 2. These gait representations are briefly described in the Sects. 2.1 to 2.5.



(a) GEI (b) GEnI (c) GGI (d) GGEnI (e) GHGI of (a)

Fig. 2. Examples of gait representations

2.1 Gait Energy Image (GEI)

Gait Energy Image (GEI) is generated by averaging all binary images in walking sequence with the same view angle, as expressed in Eq. (1).

$$G(x, y) = \frac{1}{N}\sum_{t=1}^{N} B_t(x, y) \tag{1}$$

where N is the number of silhouette frames in a complete gait sequence, t is the frame number in the gait sequence, $B_t(x, y)$ is the binary image at frame t and (x, y) is the pixel coordinate in a frame.

2.2 Gait Entropy Image (GEnI)

Gait Entropy Image (GEnI) [21] which aims to reduce unnecessary information with Shannon entropy theory is implemented as in Eq. (2).

$$GEnI = H(x, y) = \sum_{k=1}^{K} p_k(x, y) \log_2 p_k(x, y) \tag{2}$$

where (x, y) is a pixel coordinate and $p_k(x, y)$ is the k^{th} probability function which has $k = 2$ because input images are a binary image. This study follows the basic concept in [21] so that $p_2(x, y) = G(x, y)$ in Eq. (1) and $p_1(x, y) = 1 - p_2(x, y)$.

2.3 Gait Gaussian Image (GGI)

GGI is similar to GEI, however, it is produced by a Gaussian function instead of the average function. This reduces the noise effect from a individual frame in the interested gait cycle. The Gaussian function is defined in Eq. (3):

$$u_i(x) = e^{-\frac{(x_i - \bar{x})^2}{2\sigma^2}} \tag{3}$$

where u_i is Gaussian membership, x_i is the respective pixel of i^{th} frame, $\bar{x}$ is the mean of a respective pixel over all frames, and σ is the variance of the pixel vector.

Then the output pixel a_j is calculated from the average of the multiplied result between corresponding pixel and Gaussian membership, as shown in Eq. (4).

$$a_j = \frac{1}{N}\sum\nolimits_{i=1}^{N} a_i u_i \tag{4}$$

where j is the pixel position, i is the frame number, a_i is the pixel value of the i^{th} frame and N is the number of frame.

2.4 Gait Gaussian Entropy Image (GGEnI)

The aim of this newly purposed gait representation is for improving robustness against appearance changes in GGI, thus, the GEnI concept is combined with GGI in this representation. GGEnI is calculated with Eq. (2), but the probability function changes to a Gaussian membership function. GGEnI is defined as:

$$GGEnI = \sum\nolimits_{k=1}^{K} p_k(x, y) log_2 p_k(x, y) \tag{5}$$

$$u_i(x, y) = e^{-\frac{\left(a_i(x,y) - \overline{a(x,y)}\right)^2}{2\sigma^2}} \tag{6}$$

$$p_2(x, y) = \frac{1}{N}\sum\nolimits_{i=1}^{N} a_i(x, y) u_i(x, y) \tag{7}$$

$$p_1(x, y) = 1 - p_2(x, y) \tag{8}$$

where (x, y) is a pixel coordinate, and $p_k(x, y)$ is the k^{th} probability, $u_i(x, y)$ is Gaussian membership of the i^{th} frame, $a_i(x, y)$ is the pixel value of the i^{th} frame at (x, y), $\overline{a(x, y)}$ is the mean of pixel value at (x, y) for all frames, and σ is the variance of the pixel vector.

2.5 Gradient Histogram Gait Image (GHGI)

GHEI is obtained by applying histogram of oriented gradients (HoG) to each input original image then all output frames are averaged to generate the gait representation image [22]. Differently in this study, HoG is applied to the four basic gait images to generate GHGI. GHGI is computed with the following steps.

Step 1: Compute horizontal and vertical gradients I_x and I_y.
Step 2: Compute magnitude r and orientation θ

$$r = \sqrt{I_x^2 + I_y^2} \tag{9}$$

$$\theta(x, y) = \operatorname{atan}\left(\frac{I_x}{I_y}\right) \tag{10}$$

Step 3: Calculate cell histogram from each pixel in a cell which is a non-overlapping square region. Each cell is typically presented by 9 bin histograms.

$$\hat{\theta} = \frac{9.\theta(x, y)}{2\pi} \tag{11}$$

Cells are grouped to a block which has typically overlapped with neighbor blocks. Each block, containing 4 cells, represents a feature vector of length 36 after each cell is normalized by L1 norm.
Step 4: Combine feature vectors of all blocks which are normalized by lower-style clipped L2 norm [23].

3 Evaluation

All experiments were conducted on CASIA gait dataset B which contains videos from 124 persons. Each person had been captured in eleven view angles from 0 to 180° and three appearance variations including normal walking, walking with bag, and walking with coat. In each view angle for an individual, there are ten videos, six for normal walk, two for walking with bag, and other two for walking with coat. The datasets provide sequence images which are already processed by background subtraction, associated with the original videos. This study used only 116 persons who had complete sequence images in eleven view angles and three appearance variations.

Linear SVMs which are implemented with libSVM was chosen as classifiers for all experiments. Number of components from PCA was set as the maximum components which depend on the number of training datasets. Each personal model was trained with random error and results were averaged from five experiments.

3.1 HoG Parameters

Dalal and Triggs research [23] which suggested the optimized HoG parameters, for example number of orientation histogram bins should set to 9, is usually used as the reference for HoG parameters in Human detection research. Some research [5, 22] and some scientific program, for example MATLAB, adapt these parameters as default settings in HoGs. However, our GHGI, generated by the method described in Sect. 2.5, does not reach the maximum potential with these default settings. The initial experiments aim to find the optimized parameters for the HoG method. There are three interesting parameters, cell size, block size, and number of bins. All the experiments in this section use GEI as a basic gait representation. Personal model was trained by normal walk appearance while the recognition rate was tested by all three appearances.

The first experiment was cell size testing in which block size and number of bins were fixed as two and nine, respectively. Number of training samples or gait images per person which were used to train personal models was also considered in this experiment. Among the 6 normal walk samples of each person, we took 1, 2, 3 or 4 samples for training. Six samples per person which were two samples from each appearance were used as testing samples. The results of recognition rate are shown in Table 1.

Table 1. Cell size testing for HoG method

Cell size	Number of training samples			
	1	2	3	4
1	78.29	90.45	91.34	92.46
2	79.69	**91.31**	**91.39**	**92.46**
3	**82.52**	90.70	90.84	92.03
4	76.06	89.50	89.69	91.16
5	70.31	88.66	89.03	90.62
6	71.34	88.07	88.37	90.35
7	70.54	87.21	87.84	89.77
8	69.32	85.95	86.08	88.58

The optimized cell size for one training dataset was 3×3 while the rest was 2×2. After this optimized point, the recognition rate dropped continually. The cell size of 2×2 was set up as an initial parameter for the second experiment of block size testing. The number of bins was fixed as nine, the same as in the first experiment. Results are shown in Table 2.

The optimized block size was 3×3. When block size increases over 3×3, the recognition rate drops. Next, both cell size and block size were fixed to find the optimized number of bins. The testing result of number of bins against recognition rate was shown in Fig. 3. When one dataset was used in training of a personal model, it had the highest score at 87.42% with 12 bins. The rest had consequently highest scores as 92.47% (19

Table 2. Block size testing for HoG method

Block size	Number of training samples			
	1	2	3	4
1	83.14	90.97	91.42	92.58
2	84.47	91.31	91.37	92.46
3	**85.67**	**91.38**	**91.47**	**92.79**

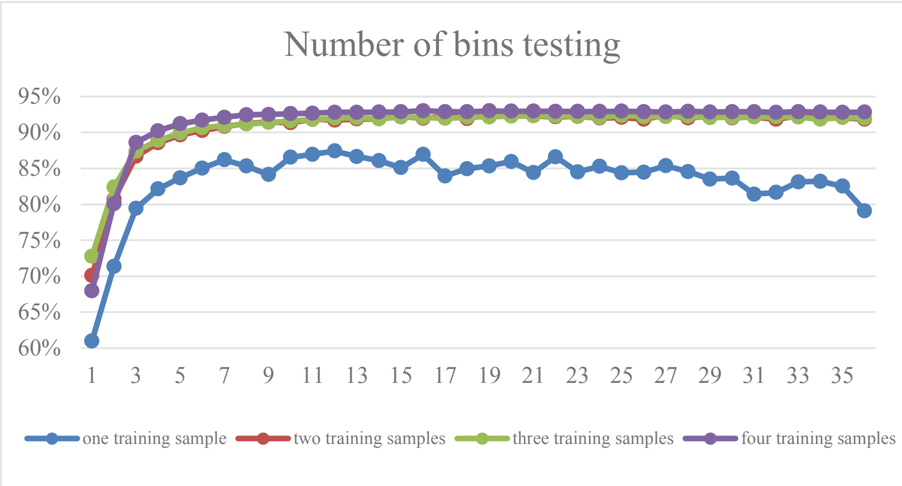


Fig. 3. Testing the number of bins for optimized HoG parameters

bins), 92.35% (25 bins) and 93.03% (16 bins) for two, three and four training samples, respectively. The optimized number of bins calculated from all testings was fifteen. As the result, the optimized parameters of the HoG method used in the experiments in this study are summarized in Table 3.

Table 3. The optimized parameter for HoG

Parameter	Number of training samples			
	1	2	3	4
Cell size	3	2	2	2
Block size	3	3	3	3
Bins	15	15	15	15

From the results demonstrated in Tables 1, 2 and 3, there is no significant difference w.r.t. recognition rate by using two or three or four training samples (only 1–2% difference). They show a similar trend whens varying the three parameters. Thus two training samples were suggested as the minimum number of training samples. In contrast, one training sample had the lowest recognition rate and it has a different pattern w.r.t. the set of best parameters, as shown in Fig. 3. It is interesting that if the parameter values were set high, the performance of GHGI was very low in terms of cell size and number of bins. The optimized parameters of HoG were applied in the other basic gait representations demonstrated in the next section.

3.2 Testing Gait Representations

Experiments involved in this section expand the HoG method to the other three basic gait representations. It shows that GHGI can improve the gait recognition rate compared to their counterparts of the basic gait representations.

Table 4. Recognition rate of the four basic gait representations and their GHGI

Representations	Appearances	Training samples							
		Only Basic Representation				GHGI			
		1	2	3	4	1	2	3	4
GEI	Normal	95.37	98.68	**99.21**	99.14	95.24	98.11	97.86	98.63
	Bag	60.67	68.85	69.95	71.82	83.83	90.88	91.31	**92.30**
	Coat	47.53	54.59	54.62	55.24	78.15	87.19	87.01	**88.45**
	Mixed	67.86	74.04	74.59	75.40	85.74	92.06	92.06	***93.13***
GEnI	Normal	94.99	98.74	99.02	**99.15**	94.33	97.69	98.09	98.86
	Bag	68.54	76.49	78.07	80.81	81.78	90.48	90.83	**92.15**
	Coat	51.12	57.78	57.30	60.17	76.46	86.08	86.25	**87.78**
	Mixed	71.55	77.67	78.13	80.04	84.19	91.42	91.72	**92.93**
GGI	Normal	94.93	98.13	98.78	99.09	96.27	99.28	99.63	**99.73**
	Bag	38.91	43.92	46.67	48.23	67.78	82.13	84.74	**87.39**
	Coat	20.45	26.61	28.72	29.75	41.73	58.95	61.99	**64.78**
	Mixed	51.43	56.22	58.06	59.02	68.60	80.12	82.12	**83.97**
GGEnI	Normal	94.72	98.11	98.65	98.99	95.47	99.00	99.55	**99.73**
	Bag	46.65	54.71	57.33	58.61	66.82	81.25	83.60	**85.67**
	Coat	22.76	28.96	31.51	32.40	38.97	58.60	62.00	**65.13**
	Mixed	54.71	60.59	62.50	63.33	67.09	79.62	81.72	**83.51**

Table 4 shows the recognition rate of using basic gait representations and Gradient Histogram Gait Image. Results were taken from the average of the elven view angles. GEnI was the best basic representation achieving the mixed average recognition rate of 80.04% with four normal walk training samples. All basic representations had a problem with appearance changes especially GGI and GGEnI which were generated by convolving Gaussian kernels.

From Table 4, it is clear that GHGI achieved relatively higher recognition rate over the basic representations except normal walk testing in case of GEI and GEnI. The recognition rate in cases of walking with bag and coat was significantly increased by using GHGI as gait representations. This shows that when the HoG is applied to the basic representations the new secondary representations are more robust to appearance changes. In contrast, the recognition rate in terms of normal walk was slightly decreased. GEI + HoG with four training samples had the highest averaged recognition rate of 93.13%. This confirms that GEI, which was simple and had less computation cost, was the best gait representation when combined with the HoG method.

The detailed recognition rate over 11 view angles for the secondary representations (GHGI) was shown in Fig. 4, which was resulted from four normal walk training samples. If only normal walk testing was considered, GHGI to GGI and GGEnI had the best result with the same value in every view angles. The recognition rate of 100% at 0°, 18°, 54°, 126°, 144°, 168° and 180° is shown to GGI and GGEnI, while GEI and GEnI had the best recognition rate at 99.57% (18°) and 99.89% (54°), respectively.

GHGI to both GEI and GEnI had the better recognition rate in case of walking with bag and coat. Both representations had recognition rate with bag more than 95% in 72°, 90° and 108°. And they had recognition rate with coat more than 90% in 18°, 36° and 54°. If all results from every gait presentation had been calculated together in each view angle, GHGI had the highest recognition rate at 91.23% in 72° while basic representations had the highest rate at 65.01% in 180°.

Overall, GHGI had increased the recognition rate of approximately 17% to GEI, 12% to GEnI, 24% to GGI and 20% to GGEnI. This confirms that HoG scan directly apply to the basic gait representations and increase their recognition rate especially in case of appearance changes. However, this experiment was based only on CASIA dataset B which captured each person in the similar style of coat, bag and walkway. More challenging scenarios need to be explored, such as view transformation, large population dataset, as well as in a real environment. Nevertheless, to some environments where walkway (airport gate) and cloths (hospital settings) could be controlled, the proposed gait representations should work well.

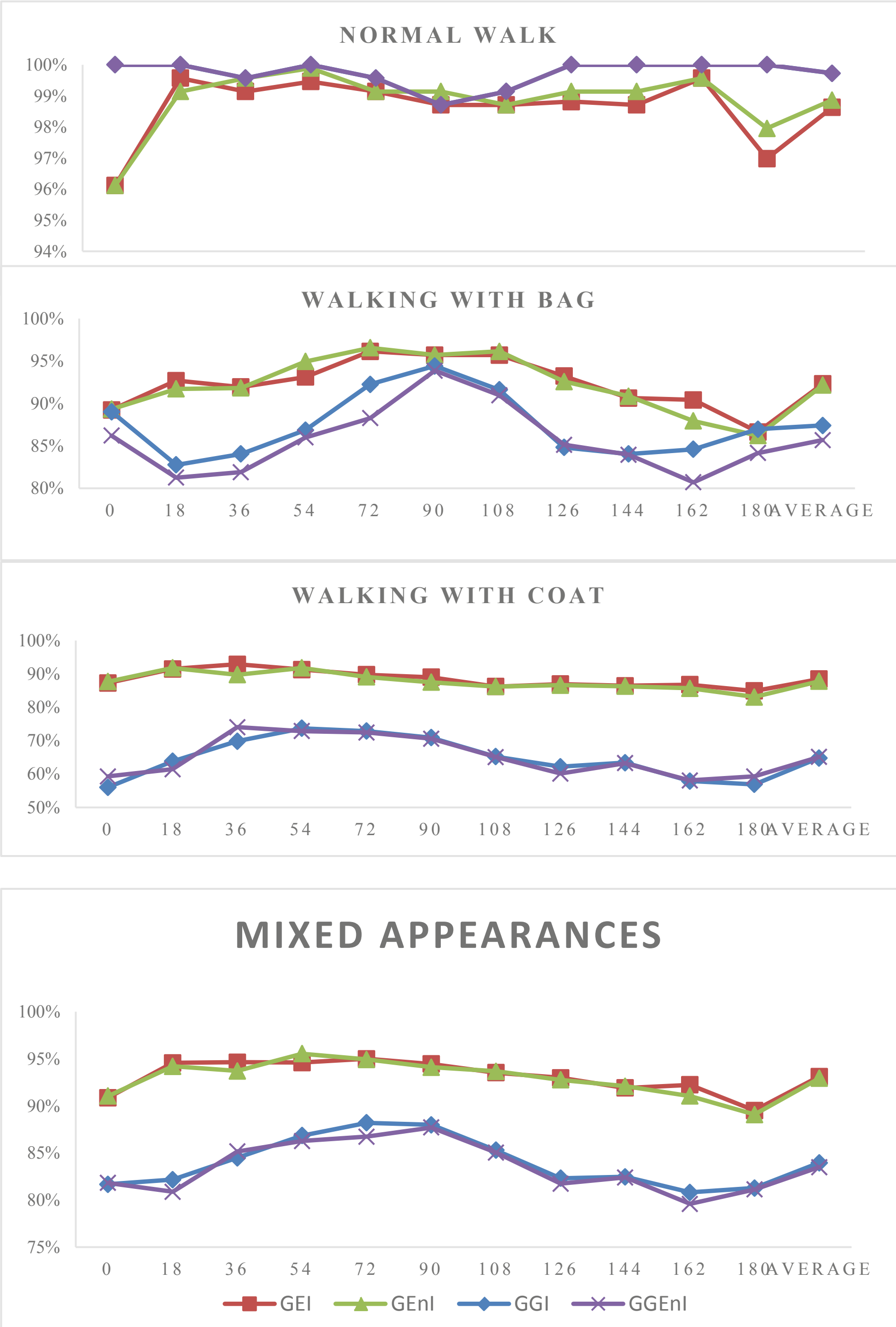


Fig. 4. Recognition rate of GHGI with different basic gait representation image

4 Conclusion

The secondary gait representations which are generated by directly applying HoGs to the basic gait representation images can improve the gait recognition rate as the experimental results shown from CASIA dataset B. Performance of these representations is dependent on parameter settings used in HoGs. When cell size and block size are decrease, small-scale details in the basic gait representations are captured. When the number of orientation histogram bins is increased, finer orientation details are captured. In contrast, they need more computational time. From our experiments, the optimized parameter settings are 2×2 cell size, 3×3 block size and 15 bins for the CASIA dataset B, except of the case of using one training sample. These secondary gait representations using the optimized HoG parameters are more robust to appearance changes as it can be seen from the results in Table 4. Among the four secondary gait representations, GHGI from GEI had the best average recognition rate at 85.74% of one training sample and 93.13% of four training samples. GHGI had increased the recognition rate approximately of 17% to GEI, 12% to GEnI, 24% to GGI and 20% to GGEnI. This study is focused on the CASIA gait dataset B. The proposed framework could be adapted for selecting optimized parameters of HoGs to other gait datasets, in which these parameters may vary.

References

1. Han, J., Bhanu, B.: Individual recognition using gait energy image. IEEE Trans. Pattern Anal. Mach. Intell. **28**(2), 316–322 (2006)
2. Liu, Z., Sarkar, S.: Simplest representation yet for gait recognition: averaged silhouette. In: 2004 Proceedings of the 17th International Conference on Pattern Recognition, ICPR 2004 (2004)
3. Arora, P., Srivastava, S.: Gait recognition using gait Gaussian image. In: 2015 2nd International Conference on Signal Processing and Integrated Networks (SPIN) (2015)
4. Yang, Y., Tu, D., Li, G.: Gait recognition using flow histogram energy image. In: 2014 22nd International Conference on Pattern Recognition (ICPR) (2014)
5. Arora, P., et al.: Improved gait recognition using gradient histogram Gaussian image. Procedia Comput. Sci. **58**, 408–413 (2015)
6. Arora, P., Hanmandlu, M., Srivastava, S.: Gait based authentication using gait information image features. Pattern Recogn. Lett. **68, Part 2**, 336–342 (2015)
7. Zheng, S., et al.: Robust view transformation model for gait recognition. In: 2011 18th IEEE International Conference on Image Processing (ICIP) (2011)
8. Chetty, G., et al.: Multiview gait biometrics for human identity recognition. In: 2014 International Conference on Computing for Sustainable Global Development (INDIACom) (2014)
9. Xue, H., Hao, Z.: Gait recognition based on gait energy image and linear discriminant analysis. In: 2015 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC) (2015)
10. Zhihui, L., et al.: Human gait recognition via sparse discriminant projection learning. IEEE Trans. Circ. Syst. Video Technol. **24**(10), 1651–1662 (2014)
11. Alotaibi, M., Mahmood, A.: Improved gait recognition based on specialized deep convolutional neural networks. In: 2015 IEEE Applied Imagery Pattern Recognition Workshop (AIPR) (2015)

12. Shiraga, K., et al.: GEINet: view-invariant gait recognition using a convolutional neural network. In: 2016 International Conference on Biometrics (ICB) (2016)
13. Wolf, T., Babaee, M., Rigoll, G.: Multi-view gait recognition using 3D convolutional neural networks. In: 2016 IEEE International Conference on Image Processing (ICIP) (2016)
14. Yan, C., Zhang, B., Coenen, F.: Multi-attributes gait identification by convolutional neural networks. In: 2015 8th International Congress on Image and Signal Processing (CISP) (2015)
15. Lan, W., et al.: Multi-view gait recognition with incomplete training data. In: 2014 IEEE International Conference on Multimedia and Expo (ICME) (2014)
16. Luo, C., Xu, W., Zhu, C.: Robust gait recognition based on partitioning and canonical correlation analysis. In: 2015 IEEE International Conference on Imaging Systems and Techniques (IST) (2015)
17. Flora, J.B., et al.: Improved gender classification using nonpathological gait kinematics in full-motion video. IEEE Trans. Hum.-Mach. Syst. **45**(3), 304–314 (2015)
18. Gupta, A., et al.: Hybrid method for Gait recognition using SVM and Baysian Network. In: 2015 IEEE 8th International Workshop on Computational Intelligence and Applications (IWCIA) (2015)
19. Das, D.: Human gait classification using combined HMM & SVM hybrid classifier. In: 2015 International Conference on Electronic Design, Computer Networks & Automated Verification (EDCAV) (2015)
20. Yu, S., Tan, D., Tan, T.: A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition. In: 2006 18th International Conference on Pattern Recognition, ICPR 2006 (2006)
21. Bashir, K., Tao, X., Shaogang, G.: Gait recognition using gait entropy image. In: 3rd International Conference on Crime Detection and Prevention (ICDP 2009) (2009)
22. Hofmann, M., Rigoll, G.: Improved gait recognition using gradient histogram energy image. In: 2012 19th IEEE International Conference on Image Processing (2012)
23. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2005) (2005)