Gait Recognition Based on Modified Gait Energy Image

Israel Raul Tiñini Alvarez
Centro de Investigación, Desarrollo e Innovación en Ingeniería Mecatrónica
Universidad Católica Boliviana "San Pablo"
La Paz, Bolivia
ir.tinini@acad.ucb.edu.bo

Guillermo Sahonero-Alvarez
Centro de Investigación, Desarrollo e Innovación en Ingeniería Mecatrónica
Universidad Católica Boliviana "San Pablo"
La Paz, Bolivia
g.sahonero@acad.ucb.edu.bo

***Abstract*— Biometric systems allow us to identify individuals from distinctive biological traits. Gait recognition is a biometric technique used to recognize humans based on the style of their walk. However, model-free based gait recognition performance is often degraded by the presence of some covariate factors such as view, clothing and carrying variations. From these, it has been shown that the change in appearance is the covariant that most affect the recognition performance. To address such issues, we propose to use a feature representation that takes both dynamic and static regions of silhouettes. This way, more robustness against covariates and better discriminative performance are expected. The proposed method is evaluated on one of the largest datasets available under the variations of clothing and carrying conditions: CASIA gait database B. Results show that the proposed method achieves correct classification rate up to 90% and outperformed state-of-the-art methods.**

Keywords—biometric systems, gait recognition, model free, modified gait energy image.

I. Introduction

The interest in developing automatic systems for automatic human identification has increased recently due to the growth in the number of crimes and attacks that have occurred in the past years [1], [2]. As a result, the community of researchers and companies of security systems have begun to focus on expanding and improving surveillance systems [3]. Automatic surveillance systems can be classified according to their task: pre-detection, detection/tracking and classification/identification [4]. Among these, the most attractive for intelligent surveillance is the last one, which would allow not only identify possible risks in the monitored spaces but also, access control, personal verification [5].

Currently, individual recognition can be achieved using biometrics characteristics such as palm print, iris, face, and fingerprints, etc. These has been popular due to the amount of data available and also their unicity, universality, and permanence [1], [6]. However, these biometrics technologies have two disadvantages: 1) their performance decreases in low-resolution images and 2) user cooperation is required to achieve good recognition [7]–[11]. To overcome these limitations, new biometrics characteristics such as gait have been proposed. Its advantages have attracted significant attention in the biometric research community as it is seen as a valid and even a better alternative [12], [13].

Gait can be defined as the way of walk, combined with our posture. There is a large number of works in gait recognition that have shown that gait is feasible in human identification despite the distance [14]. Gait recognition techniques can be classified into two main categories: model-based and model-free approaches. Model-based methods model the person body structure, but this process gets computationally expensive, error-prone and requires high-resolution images [15]. On the other hand, model-free or appearance-based approaches build signatures from silhouettes extracted from video sequences for the recognition. Hence, it requires less computation and its performance does not reduce when using low-resolution images; this means less memory usage [15]. Due to these advantages, the current trend on gait recognition seems to focus on model-free approaches. However, the major challenge of these methods is to cope with various covariates, such as view, clothing, carrying, shoe, surface and even time conditions [13].

This paper proposes a better way of recognition by using a modified gait representation of gait sequence. The proposed method achieves dimensionality reduction by applying principal component analysis and classification employing linear discriminant analysis. The model was tested using the CASIA-B benchmarked database and the performance of the model was measured using the rate of correct classified examples.

The rest of this paper is organized as follows: Section 2 summarizes previous works. Section 3 gives the description of the proposed method. Experiments and results are presented in Section 4. The last section, Section 5, shows our conclusions.

II. Related Works

There is a considerable amount of work focused on the use of appearance for gait recognition. Gait energy image (GEI), suggested by Han and Bhanu [3], is one of the most used representation since it considers an effective balance between computational cost and recognition performance [8]. GEI is a spatio-temporal representation of the gait that can be obtained averaging the silhouettes over a gait cycle. Although this representation is one of the most popular gait representation used, it has been found that clothing and carrying conditions influence the recognition performance [9]. Bashir et al. [15] introduced a new representation called gait entropy image (GEnI) which is constructed computing the Shannon entropy for each pixel over a gait cycle. This means that pixels with high value on the GEnI frame correspond to dynamic parts that are more robust against covariates.

In [7], three new representations based on Gabor functions (GaborD, GaborS, and GaborSd) are proposed, where linear discriminant analysis LDA is used for the stage of classification, and the performance of the model is evaluated in the HumanID dataset. Dupuis et al. [16] represent the gait sequences using a modified GEI where the feature ranking is

978-1-5386-8375-0/18/$31.00 ©2018 IEEE

carried out using tree random forest (RF), and canonical discriminant analysis (CDA) is employed for classification.

In [10], GEI is used to represent the gait sequence, and a motion based vector is proposed to reduce the dimensionality of the data, and Group Lasso algorithm is used to select the most important parts of the GEI representations. Finally, CDA is used for the classification task. In the same context [9] also use the motion based vector (MBV) as features of the gait, however, the two main parts of the body are used, instead of just the most important. A different approach by the same author in [11] where the most important features are selected using Statistical Dependency (SD) with Globality-Locality Preserving Projections (GLPP) and K-nearest neighbor (KNN) is used for classification.

Recently, there is a trend of employing deep learning techniques to account the issue of appearance and view changes [1]. In [17] a sequential set of 2D stick figures is used to represent the gait cycle, and a back-propagation neural network is used as a classifier. The experiments were carried out in the SOTON database. Similarly, in [18] a new algorithm to create a 2D spatiotemporal gait template suitable for real-time surveillance application is proposed, where a neural network of five layers is used for classification. The same author in a similar research [19] proposes a specialized deep convolutional neural network (CNN), where GEI representations are used as gait feature descriptors and input to the CNN. The architecture of the network consisted of eight layers: four convolutional and four pooling layers allowed to obtain relatively good results in the CASIA-B dataset. A similar idea is presented in [20] where GEI grayscale images of 88x128x1 are used to feed a CNN of 6 layers, 3 convolutional and 3 fully connected layers. The experimental validation has been carried out on the OU-ISIR-B dataset, experimental results showed that the proposed method outperformed the state-of-the-art methods.

Even more sophisticated preprocessing algorithms are employed in [14], where a deep model based on auto-encoders is designed to overcome the issue of view variations. Specifically, Stacked Progressive AutoEncoders (SPAE) of 5 layers is used. The advantage of this model is that it can extract view variant feature from any view using only one model, therefore, another algorithm for view estimation is not needed. In [12] the same idea of SPAE is applied to deal the problem of view, clothing and carrying condition variation. Although the results obtained did not outperform the state-of-the-art methods, employing autoencoders to address the problem of covariates, is a promising idea that already showed improvements in the field of face recognition.

III. Proposed Method

In this paper, we focus on solving the problem of change in appearance caused by the change in clothing. Effects of clothing were studied by Matovski et al. [21]. Their work shows that performance is significantly reduced by clothing covariate; elapsed time, footwear, speed and size of images are not that impactful.

The proposed model is inspired by the one in [9] where a motion vector based on GEI is used to determine which parts of the body are more robust against covariates. Fig. 1 shows our framework that contains four main modules. These are described in the following sections.

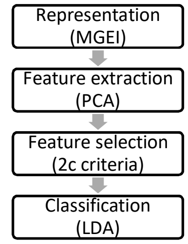

Fig. 1. The proposed framework for gait recognition

A. Representation

GEI is a spatio-temporal representation of gait cycles, which is produced averaging the silhouettes extracted over a complete gait cycle, it is nowadays the best signature for gait for its robustness to noise and easy to compute [3]. Given a size-normalized and center aligned binary gait silhouettes images $B_{t\ (x,y)}$ the GEI can be computed using the following equation:

$$G_{(x,y)} = \frac{1}{N}\sum_{t=1}^{N} B_{t\ (x,y)} \quad (1)$$

where N is the number of frames within a gait cycle, x and y are values in the 2D image coordinate and B_t is a silhouette image of frame t. Fig. 2 shows the sample silhouettes images from one person and the right most image is the corresponding GEI.

Gait energy image has two main regions: the static parts (high-intensity pixels) and the dynamic areas (low-intensity pixels) [4]. However, it is vulnerable to appearance changes of the human silhouette [15], as we can see in Fig. 3 (top row), the representation of the gait changes in function on the circumstances. The image on the left corresponds to a normal condition, the one in the center to a carrying-bag condition and the one on the right to a wearing-coat condition. Moreover, to improve the representation, [16] proposes to use only certain parts of the GEI. The parts of the body are ranked using RF. Results showed that the head and feet contain the most important features. As illustrated in Fig. 3, the modified representations look very similar even under different conditions.

B. Feature extraction and selection

1) Feature extraction

Since gait sequences are considered high-dimensional data, it is necessary to perform a dimensionality reduction stage which allows us not only to select those characteristics which contain more information with minimum redundancy but also to improve the performance of our classifier [22]. For this stage, we chose PCA for the reduction of dimensionality and its good performance as feature extraction and reduction of gait data [23], [24].

2) Feature selection

Once the important components are extracted with PCA, the selection of the number of components to feed the classifier was determined according to the number of samples.

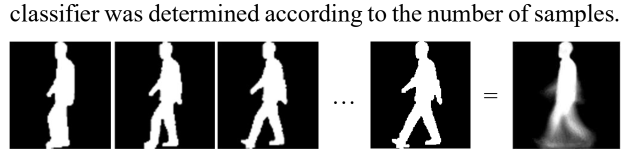

Fig. 2. Gait energy image (the right one)

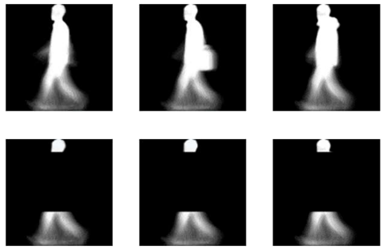

Fig. 3. Modified gait energy image (bottom row)

To determine the number of components to be used, we consider the work [3], where it is suggested to retain 2c components where c corresponds to the number of classes.

C. Classification

The classification is performed to make the decision whether the subject belongs to a class in database or not. Unlike most of the reviewed works which use KNN for classification stage [4], we used LDA given its discriminatory capacity. The performance of our method is measured by the CCR that corresponds to the ratio of the number of correctly classified samples over the total number of samples.

IV. Experiments and Results

A. Dataset description

We have used CASIA to evaluate our model. CASIA-B [25] is one of the largest datasets available for benchmarking gait recognition techniques, which was collected by The Institute of Automation Chinese Academy of Sciences. It is an indoor gait dataset and comprises 124 subjects captured from 11 views. For each subject, there are 10 walking sequences consisting of 6 normal walking (SetA), 2 carrying-bag sequences (SetB) and 2 wearing-coat sequences (SetC). Each sequence contains at least one gait cycle. Database original image size is 320x240.

B. Experiments design

As this work focuses on the effect of appearance variations, we carried out our experiments under 90° view angle, since it has been proved to be the best view because it allows to capture more dynamic information [18]. All the representation images were resized to 100x100 resolution. The first four sequences of SetA (SetA1) were used for training, the two-remaining denoted as SetA2 along with SetB and SetC were used for testing the normal, carrying, and clothing conditions.

Since GEI is used as representation, we will first compare the performance of our modified representation with respect the original representation. Next, we compare the performance of the proposed method with state-of-the-art methods.

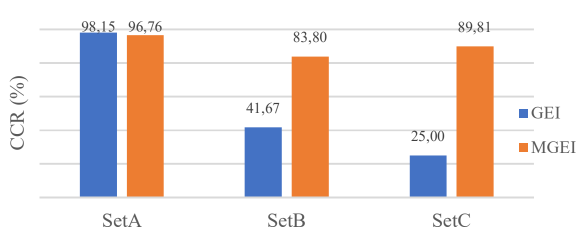

Fig. 4. Recognition rates of the proposed method

C. Results and analysis

To better illustrate the performance of our method, we decided to compare the CCR of the original representation and the modified proposed using the proposed procedure. Fig. 4 shows the comparison of the recognition rate under different covariate factors.

The CCR of MGEI based procedure is slightly lower in the normal condition but considerably higher in the carrying-bag and waring-coat conditions. This is because the representation takes in consideration both dynamic and static parts of the body that are more robust against covariates and have most discriminative power.

Figure 4 shows a lower CCR in the case of carrying-bag sequence (SetB) in comparison with normal sequences (SetA). So, we decided to repeat the training and testing process for the first 20 samples and form the confusion matrix to determine the effect of the covariant. From the confusion matrix of the first 20 sequences, corresponding to the first 20 subjects, it was observed that the gait patterns of subjects 2 and 7 were not well generalized, so the binary silhouettes and the modified GEI were analyzed. This revealed that the problem was because subjects 2 and 7, during their walk sequence, are carrying a suitcase at the knees, which modifies their silhouette in the segmentation stage. As this is not considered by the classifier, it causes errors in classification. Fig. 5 shows the GEI obtained and some binary silhouettes.

An analysis like the previous one was done for the covariant of wearing-coat (SetC). The confusion matrix was formed again to determine the effect of the covariant and it was observed that the subject 22 of the top 30 suffered a bad classification. It was determined that the change of extreme clothes (wearing a coat that covered the legs) caused the bad classification, since the coat gravely modifies the silhouette, which is consistent with [12], [14], where it is reported that some clothes, such as long overcoats, can occlude leg motion. Figure 6 shows the GEI obtained and some binary silhouettes.

D. Comparison with the state-of-the-art

To better illustrate the performance of the proposed method, we also compare our method against the reported by other state-of-the-art methods. Gait recognition techniques are compared in Table 1 based on their CCR, but also the preprocessing and the classification technique is shown.

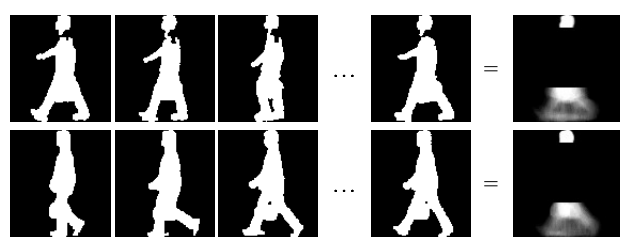

Fig. 5. Modified gait energy image created from a sequence of binary silhouettes (2 and 7 subjects respectively)

Fig. 6. Modified gait energy image created from a sequence of binary silhouettes (subject 22 of the dataset)

TABLE I. COMPARING THE STATE-OF-THE-ART GAIT RECOGNITION METHODS ON CASIA-B FOR SIDE VIEW, USING GEI REPRESENTATION

Work	Pre-processing	Classifier	SetA	SetB	SetC	Overall CCR (%)
[12]	SPAE + PCA	KNN	95,97	65,32	42,74	68,01
[11]	SD + GLPP	KNN	98,80	70,10	89,29	86,06
[9]	MBV + PCA	CDA	98,39	75,89	91,96	88,75
[10]	MBV + PCA	CDA	95,56	74,11	86,61	85,43
[19]	-	CNN	95,50	88,30	76,20	86,67
Ours	PCA	LDA	96,76	83,80	89,81	90,12

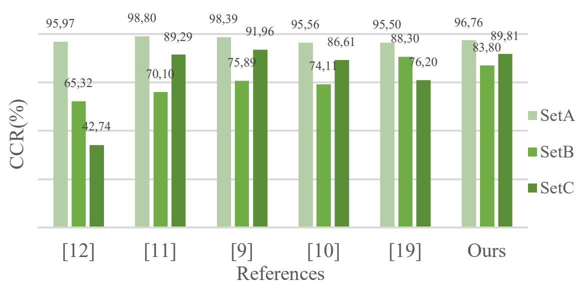

Fig. 7. Comparison of CCR from some state-of-the-art methods under different covariate factors

PCA is a common preprocessing technique used to perform dimensionality reduction. Also, simple classifiers are preferred, this is because the number of samples in CASIA-B is limited for training [12]. Fig. 7 shows the CCR comparison considering the three covariate factors.

In real life, people wear different clothes depending on days (cool or warm days) and the season (winter or summer) therefore clothing variants is always present, being this an important issue [21]. The proposed representation addresses the clothing variation problem since it significantly outperforms other approaches as is shown in Fig. 7.

V. CONCLUSION

In this paper we used a modified gait representation, performed dimensionality reduction using PCA and the classification task was solved using LDA to improve the gait recognition accuracy. The proposed framework achieves the best performance among all the approaches. This is verified by experiments, which show that the proposed method significantly outperforms the state-of-the-art methods especially in the case of carrying-bag and clothing-coat variations, which is very suitable for practical application in intelligent surveillance systems. We also analyze the effect caused by the inclusion of covariates in the gait sequences and determined that long coats and suitcases at the height of the feet occlude the leg motion.

In the future, we will improve this model to deal more challenging covariates, such as view variations, prove the model in other benchmarking large databases since the use of more subjects is important to have reliable results and improve the accuracy in the case of clothing-coat conditions.

REFERENCES

[1] P. Karampelas and T. Bourlai, *Surveillance in Action*. 2018.

[2] H. Iwama, M. Okumura, Y. Makihara, and Y. Yagi, "The OU-ISIR gait database comprising the large population dataset and performance evaluation of gait recognition," *IEEE Trans. Inf. Forensics Secur.*, vol. 7, no. 5, pp. 1511–1521, 2012.

[3] J. Han and B. Bhanu, "Individual recognition using gait energy image," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 2, pp. 316–322, 2006.

[4] M. Tariq and M. A. Shah, "Review of Model-Free Gait Recognition in Biometric Systems," 2017.

[5] G. V. Veres, L. Gordon, J. N. Carter, and M. S. Nixon, "What image information is important in silhouette-based gait recognition?," *Proc. 2004 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognition, 2004. CVPR 2004.*, vol. 2, pp. 776–782, 2004.

[6] S. Das, U. Kumar, and S. Meher, "Improving Single View Gait Recognition Using Sparse Representation Based Classification," pp. 317–321, 2016.

[7] D. Tao, X. Li, X. Wu, and S. J. Maybank, "General tensor discriminant analysis and Gabor features for gait recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 10, pp. 1700–1715, 2007.

[8] I. Rida, S. Almaadeed, and A. Bouridane, "Gait recognition based on modified phase-only correlation," *Signal, Image Video Process.*, vol. 10, no. 3, pp. 463–470, 2016.

[9] I. Rida, X. Jiang, and G. L. Marcialis, "Human body part selection by group lasso of motion for model-free gait recognition," *IEEE Signal Process. Lett.*, vol. 23, no. 1, pp. 154–158, 2016.

[10] I. Rida, S. Al-Maadeed, and A. Bouridane, "Unsupervised Feature Selection Method for Improved Human Gait Recognition," pp. 1133–1137, 2015.

[11] I. Rida, L. Boubchir, N. Al-Maadeed, S. Al-Maadeed, and A. Bouridane, "Robust model-free gait recognition by statistical dependency feature selection and Globality-Locality Preserving Projections," *2016 39th Int. Conf. Telecommun. Signal Process. TSP 2016*, pp. 652–655, 2016.

[12] S. Yu, H. Chen, Q. Wang, L. Shen, and Y. Huang, "Invariant feature extraction for gait recognition using only one uniform model," *Neurocomputing*, vol. 239, pp. 81–93, 2017.

[13] S. Sarkar, P. J. Phillips, Z. Liu, I. R. Vega, P. Grother, and K. W. Bowyer, "The humanID gait challenge problem: Data sets, performance, and analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 2, pp. 162–177, 2005.

[14] S. Yu, Q. Wang, L. Shen, and Y. Huang, "View invariant gait recognition using only one uniform model," *Proc. - Int. Conf. Pattern Recognit.*, pp. 889–894, 2017.

[15] B. Khalid, T. Xiang, and S. Gong, "Gait recognition using gait entropy image," *3rd Int. Conf. Imaging Crime Detect. Prev. (ICDP 2009)*, pp. P2–P2, 2009.

[16] Y. Dupuis, X. Savatier, and P. Vasseur, "Feature subset selection applied to model-free gait recognition," *Image Vis. Comput.*, vol. 31, no. 8, pp. 580–591, 2013.

[17] J.-H. Yoo, D. Hwang, K.-Y. Moon, and M. S. Nixon, "Automated Human Recognition by Gait using Neural Network," *1st Work. Image Process. Theory, Tools Appl. (IPTA 2008)*, pp. 1–6, 2008.

[18] M. Alotaibi and A. Mahmood, "Automatic Real Time Gait Recognition based on Spatiotemporal Templates," 2015.

[19] M. Alotaibi and A. Mahmood, "Improved Gait recognition based on specialized deep convolutional neural networks," *2015 IEEE Appl. Imag. Pattern Recognit. Work.*, pp. 1–7, 2015.

[20] T. Yeoh, A. E. Hernan, and K. Tanaka, "Clothing-invariant Gait Recognition Using Convolutional Neural Network," *2016 Int. Symp. Intell. Signal Process. Commun. Syst.*, pp. 1–5, 2016.

[21] D. S. Matovski, M. S. Nixon, S. Mahmoodi, and J. N. Carter, "The effect of time on gait recognition performance," *IEEE Trans. Inf. Forensics Secur.*, vol. 7, no. 2, pp. 543–552, 2012.

[22] X. Shi, Z. Guo, F. Nie, L. Yang, J. You, and D. Tao, "Two-Dimensional Whitening Reconstruction for Enhancing Robustness of Principal Component Analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 10, pp. 2130–2136, 2016.

[23] T. Chau, "A review of analytical techniques for gait data. Part 1: Fuzzy, statistical and fractal methods," *Gait Posture*, vol. 13, no. 1, pp. 49–66, 2001.

[24] T. Chau, "A review of analytical techniques for gait data. Part 2: neural network and wavelet methods," *Gait Posture*, vol. 13, no. 2, pp. 102–120, 2001.

[25] S. Yu, D. Tan, and T. Tan, "A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition," *Proc. - Int. Conf. Pattern Recognit.*, vol. 4, pp. 441–444, 2006.