

# Clothing-invariant human gait recognition using an adaptive outlier detection method

A. Ghebleh<sup>1</sup>  · M. Ebrahimi Moghaddam<sup>1</sup>

Received: 29 September 2016 / Revised: 23 February 2017 / Accepted: 12 April 2017 /

Published online: 9 May 2017

© Springer Science+Business Media New York 2017

**Abstract** Human gait as a behavioral biometric identifier has received much attention in recent years. But there are some challenges which hinder using this biometric in real applications. One of these challenges is clothing variations which complicates the recognition process. In this paper, we propose an adaptive outlier detection method to remove the effect of clothing on silhouettes. The proposed method detects the most similar parts of probe and each gallery sample independently and uses these parts to obtain a similarity measure. Towards this end, the distances of the probe and a gallery sample are calculated row by row which are then used to obtain an adaptive threshold to determine valid and invalid rows. The average distance per valid rows is then considered as dissimilarity measure of samples. Experimental results on OU-ISIR Gait Database, the Treadmill Dataset B and CASIA Gait Database, Dataset B, show that this method efficiently detects and removes the clothing effect on silhouettes and reaches about 82 and 84% successful recognition respectively.

**Keywords** Biometrics · Clothing-invariant · Gait recognition · Outlier detection

## 1 Introduction

Biometric identification systems and security cameras are widely used in public places like banks, airports, surveillance environments, etc. A variety of biometric identifiers are used in security applications. These include signature, fingerprint, palm vein, face, iris and retina scans. These biometrics are not convenient for crowded environments since they all need the subject's cooperation and high quality images from close distances.

---

✉ A. Ghebleh  
a\_ghebleh@sbu.ac.ir

M. Ebrahimi Moghaddam  
m\_moghaddam@sbu.ac.ir

<sup>1</sup> Faculty of Computer Science & Engineering, Shahid Beheshti University: G, C, Tehran, Iran

Human gait as a behavioral biometric identifier has gained significant attention in recent years. This is due to its unique characteristics such as unobtrusiveness, recognition from distance and no need of high quality video. In addition, since existing security cameras can be used for gait recognition, it is affordable and easy to set up. However, there are difficulties in human gait recognition which impact the accuracy of the recognition process. One of these challenges is the type of clothing which confounds human gait recognition approaches. Moreover, since human identification is usually a real-time application, any solution for these challenges should not impose much computational overhead on the system.

In this work, we propose an adaptive outlier detection method to mitigate the impact of clothing in human gait recognition. Although this method uses thresholding, it does not need any manual threshold estimation. Unlike other methods, we do not use any global information like the typical body shape or any static splitting of the silhouette. Instead, our similarity measure of two samples is obtained from most similar areas of them. Towards this end, the distance of two samples is obtained row by row and the average of these distances is used as threshold to detect outliers, i.e., the rows which are affected by clothing. Finally, the average of the remaining rows is used as dissimilarity measure. This approach effectively detects and eliminates clothing effects on silhouettes.

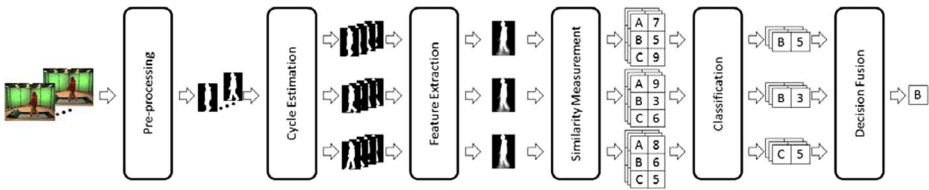
The rest of paper has been organized as follows: section 2 reviews some existing approaches briefly, in section 3 the motivation and intuition of the proposed method is explained while section 4 covers it in detail. Evaluation results are illustrated in section 5 and finally we conclude the paper in section 6.

## 2 Related works and background

Human gait recognition has attracted tremendous interest from researchers. Existing approaches can be roughly divided into two categories: model-based and model-free. Model-based approaches consider a model of human body and try to fit input frames of a walking person to that model. The parameters of the model such as joint angles are then used as features. Some examples of model-based approaches are [4–6, 8, 22, 28]. On the other hand, model-free approaches like [2, 3, 9, 11, 13, 15, 23, 25, 29] focus on the shape of the silhouette rather than fitting it to a chosen model. The desired feature is computed from the silhouettes obtained from input frames.

Theoretically, model-based approaches are more robust than model-free ones. In particular, they are mostly view and scale invariant. But these approaches suffer from model fitting errors which affect their performance. On the other hand, model-free approaches are less sensitive to the quality of input frames and have the benefit of low computational cost compared to model-based approaches. On the other hand, model-free approaches are sensitive to factors such as clothing variations which influence the shape of the silhouette.

Researchers have proposed methods to remedy the dependence of model-free approaches on clothing variations. We can consider a general framework for these methods similar to typical pattern recognition systems as depicted in Fig. 1. It is composed of pre-processing, cycle estimation, feature extraction, similarity measurement, classification, and decision fusion modules. The pre-processing module removes the background and extracts the silhouettes. In cycle estimation, complete cycles of walking are detected and then desired features are extracted using a feature extraction module. Then the similarity between the probe and gallery samples are calculated and using the classification module, one or multiple candidates are suggested. If there are more than one candidate, the decision fusion module selects one as final result.



**Fig. 1** The structure of proposed method (A, B and C denote typical gallery samples)

Considering the problem of clothing variations, different approaches focus on one or multiple modules of this framework. While pre-processing and cycle estimation modules have nothing to do with clothing variations, there are few works which try to introduce clothing-invariant features. Yogarajah et al. [26] propose a new feature called  $P_{RW}GEI$  based on Poisson Random Walks, that effectively removes the impact on silhouettes of the subject carrying a bag. However, this feature is not much effective in situations where the subject wears a coat.

Most of the approaches proposed so far to mitigate the clothing variations problem concentrate on a similarity measurement module. These approaches try to find the most significant and robust parts of the silhouette or to detect and remove disrupted parts of it. In one of the earliest of these works, Bashir et al. [1] suggest supervised and unsupervised methods to extract the most relevant and informative features from the gait energy image (GEI) [9] which is simply the average of silhouettes. Li et al. [14] segment the silhouette into seven components and study the effectiveness of each component for identification and in particular, gender recognition. Hossain et al. [10] segment the silhouette into six regions with overlaps and inspect the impact of clothing variation in each region. They propose to use the weighted sum of similarity in different regions as the overall similarity. A technique is proposed in [19–21] for defining more effective parts for clothing. In this technique, the whole gait feature like GEI is constructed incrementally row by row. Adding each row, the recognition rate is calculated and the positive or negative effect of the row at hand is obtained. The results determine three most effective and two less effective parts. Finally, in [24] the affected areas in the GEI are detected and after removing them, the recognition is performed. Towards this end, [24] introduces Typical GEI (TGEI) which is the average of all GEIs with normal clothing in the gallery and shows the form of body with no covariate factors. The difference between the probe GEI and TGEI shows the areas which are disrupted by clothing and all rows of that area are removed.

The method proposed by Guan et al. [7] is one of the most important works in classifier module which regards clothing-invariant human gait recognition. They model the effect of covariates as an unknown partial feature corruption problem and propose a classifier ensemble method based on random subspace method to reduce the effect of covariate factors in gait recognition. They extend their method by proposing local enhancement and hybrid decision-level fusion strategies that leads to good performance. In decision fusion module the most used method is majority voting which selects the most frequent candidate as the result and is used in some works such as [7].

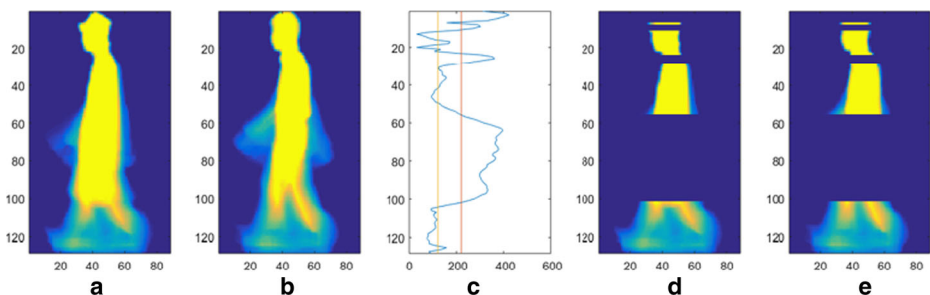
### 3 Motivation and intuition

When trying to decide whether two pictures belong to the same person, one usually looks for similarities between the two pictures and dismisses any misleading dissimilarities. On

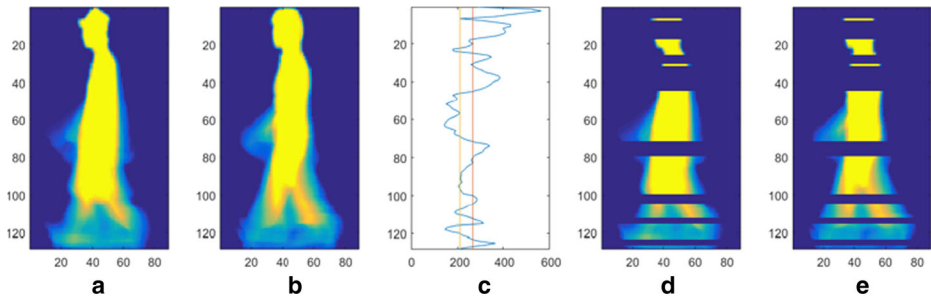
the other hand, we do not assume a general model and prejudge cases. For example, if some parts have 100% match and the nose is different we can conclude that these pictures belong to the same person but we cannot assume that the nose should always be discarded. In other words, to find the overall similarity of two pictures, we try to find the most similar parts of them.

As mentioned before, clothing effect on silhouettes troubles human gait recognition. So we should detect and eliminate its effect for successful recognition. Assume that we have two samples with same general shape, for example two rectangles. Subtracting these samples row by row after alignment and normalization leads to distances which are equal or nearly equal. But if one of them has any deficiencies, the distances in the corresponding parts are higher. Likewise, since the general shape of silhouettes is similar, the row by row distance of samples should be nearly constant. If distance value in some rows is much higher than others, we conclude that these rows are affected by clothing variations. Figure 2 shows the GEIs of same person with different cloths. As presented in this figure, the distance values in rows which are affected by cloths are higher than the distance values in other rows. So we can use thresholds to detect and eliminate these rows. The critical issue is the threshold value. It is obvious that using a strict value is not appropriate for all cases. So we use statistical measures to define the threshold value. More specifically, we use the average and standard deviation of distances. The average of distances obtained from rows represents the typical range of distances and rows which produce distances higher than this threshold are marked as outliers. We do not claim that this value is best for threshold but we demonstrate the effectiveness of this approach.

In Fig. 2c the red line represents the average distance of all rows which is used as threshold and the yellow line shows the average distance after removing outliers. As we can see this approach can successfully detect and eliminate areas which are affected by clothing variations. Note that original samples are not much similar but after removing outliers they are approximately same. On the other hand, Fig. 3 shows the case in which samples are from different persons. We can see that although some rows are eliminated but the average distance after removing them is still higher than the case illustrated in Fig. 2. Numerical analysis as presented in experimental results also verifies that using this approach is effective.



**Fig. 2** Intuition of proposed method for same person. (a) probe, (b) gallery, (c) distances in each row, (d) probe after removing outliers, (e) gallery after removing outliers



**Fig. 3** Intuition of proposed method for different persons. (a) probe, (b) gallery, (c) distances in each row, (d) probe after removing outliers, (e) gallery after removing outliers

## 4 Proposed method

The proposed method is explained in the following sections. Algorithm 1 shows the overall process of recognition.

---

### Algorithm 1 Recognition process

---

**Input:** Collection of all gallery samples

A probe sample

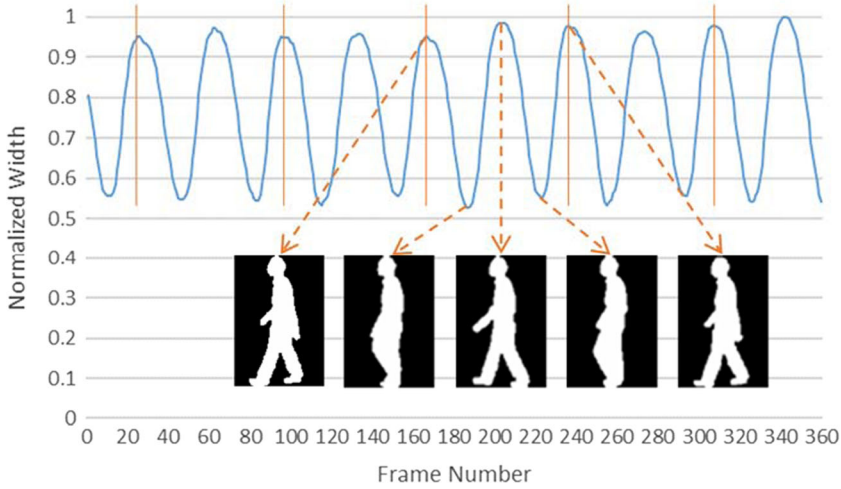
**Output:** A subject

---

- 1- Preprocess inputs to extract silhouettes
  - 2- Estimate cycles for all input samples
  - 3- Extract features from all complete cycles of all gallery samples which makes  $G = \{g^1, g^2, \dots, g^n\}$
  - 4- Extract features from all complete cycles of probe which makes  $p = \{p^1, p^2, \dots, p^m\}$
  - 5- for  $p^i \in p$
  - 6- Perform similarity measurement (Algorithm 2) using  $G$  and  $p^i$  which results in  $L^i_+$ ,  $L^i_-$  and  $L^i_*$
  - 7- Perform classification (Algorithm 3) using  $L^i_+$  and  $G$  to obtain  $Y^i$  and  $Z^i$
  - 8- end for
  - 9- Perform decision fusion (Algorithm 4 or Algorithm 5) using  $Y^i$  and  $Z^i$  for  $i = \{1, 2, \dots, m\}$
- 

### 4.1 Pre-processing

The input of this module is the frame sequence of a person's walking cycle. In each frame, the background is subtracted and using morphological operations, the silhouette of the person is obtained and scaled to the same height. The most challenging problem in this module is background detection and shadow removal. A simple yet effective method is to record the background in advance and subtract this recorded image from the input frame. The silhouette of the person is then obtained using some morphological operations. Some other methods for background detection and removal are described in [18].



**Fig. 4** Variations of silhouettes’ widths. Complete cycles are separated by vertical red lines and corresponding silhouettes are shown for third cycle

### 4.2 Cycle estimation

In cycle estimation module, complete cycles of input silhouettes are extracted using variations of silhouettes’ widths [12]. The gait cycle starts when one foot makes contact with the ground and ends when that same foot contacts the ground again. The variations of the silhouette’s width are shown in Fig. 4. We can see that it has a regular pattern, the maximum value is when the legs are in maximum distance, and the minimum corresponds to the situation when legs pass over each other. So three adjacent maximums denote a complete gait cycle.

### 4.3 Feature extraction

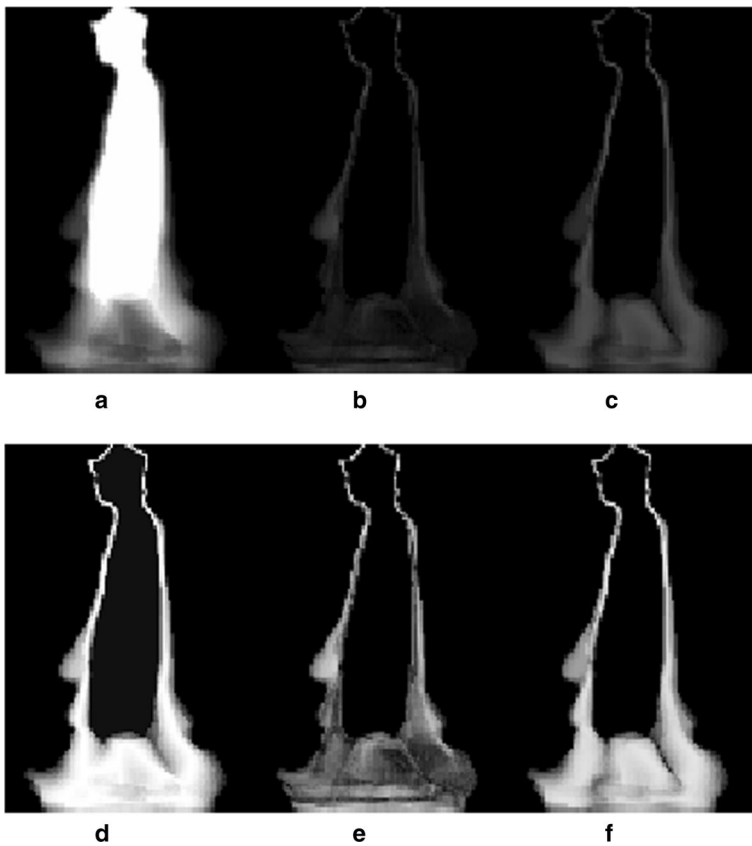
In the feature extraction module, each complete cycle is used to compute the desired feature. In this work, we have tested four different features, namely GEI, GENI [2], DFT [17] and EnDFT [21] as shown in Fig. 5 but other features could also be used. GEI is simply the average of silhouettes after normalization and alignment, and GENI is the same after applying entropy on each pixel. That is,

$$GEI(x, y) = \frac{1}{l} \sum_{i=1}^l silhouette_i(x, y) \tag{1}$$

$$GENI(x, y) = GEI(x, y) * \log(GEI(x, y)) + (1 - GEI(x, y)) * \log(1 - GEI(x, y)) \tag{2}$$

DFT is an extension of GEI and uses the Discrete Fourier Transform for feature extraction:

$$DFT_k(x, y) = \frac{1}{l} \left| \sum_{i=1}^l silhouette_i(x, y) * e^{-j2\pi ki / l} \right| \quad \text{for } k = 0, 1, 2 \tag{3}$$



**Fig. 5** DFT and EnDFT features. (a)  $DFT_0$  (b)  $DFT_1$  (c)  $DFT_2$  (d)  $EnDFT_0$  (e)  $EnDFT_1$  (f)  $EnDFT_2$

Concatenation of features  $DFT_0$ ,  $DFT_1$ ,  $DFT_2$  forms the DFT feature. Note that  $DFT_0$  equals GEI. Similarly to GENI, EnDFT is calculated as:

$$EnDFT(x, y) = DFT(x, y) * \log(DFT(x, y)) + (1 - DFT(x, y)) * \log(1 - DFT(x, y)) \quad (4)$$

#### 4.4 Similarity measurement

Assume that  $p = \{p^1, p^2, \dots, p^m\}$  represents the features computed from probe's complete cycles and  $G = \{g^1, g^2, \dots, g^n\}$  denotes gallery features, which are computed similarly to  $p$  using all gallery samples. The similarity measurement module compares each  $p^i$  to features from the gallery. The result of similarity measurement is three different lists for each  $p^i$ . Namely,  $L_i^+$ ,  $L_i^-$  and  $L_i^0$  containing the distances of  $p^i$  to each gallery feature. These lists are obtained using three different thresholds namely,  $\hat{t}$ ,  $t^+$  and  $t^-$  respectively which are explained in next subsections. The algorithm of this module is shown in Algorithm 2.

**Algorithm 2** Similarity measurement

**Input:** Gallery features

A probe feature

**Output:** Lists of distances

1. for  $g^s \in G$
2.     for  $j \in \{1, 2, \dots, h\}$
3.         Compute the distance of  $j$ -th row of  $p^i$  and  $g^s$ ,  $D_{sj}$
4.     end for
5. end for
6. Compute  $D$  the average of all  $D_{sj}$ s
7. for  $s \in \{1, 2, \dots, n\}$
8.      $t_s \leftarrow \text{mean}(\{D_{sj} \mid j = 1, 2, \dots, h\})$
9.      $r_s \leftarrow t_s / D$
10.     $\hat{t}_s \leftarrow t_s + \log(r_s) \times \text{std}(\{D_{sj} \mid j = 1, 2, \dots, h\})$
11.     $t_s^+ \leftarrow t_s + \text{std}(\{D_{sj} \mid j = 1, 2, \dots, h\})$
12.     $t_s^- \leftarrow t_s - \text{std}(\{D_{sj} \mid j = 1, 2, \dots, h\})$
13.     $L_s^{is} \leftarrow \text{mean}(\{D_{sj} \mid j = 1, 2, \dots, h \wedge D_{sj} < \hat{t}_s\})$
14.     $L_s^{is+} \leftarrow \text{mean}(\{D_{sj} \mid j = 1, 2, \dots, h \wedge D_{sj} < t_s^+\})$
15.     $L_s^{is-} \leftarrow \text{mean}(\{D_{sj} \mid j = 1, 2, \dots, h \wedge D_{sj} < t_s^-\})$
16. end for

4.4.1 Baseline threshold

As mentioned before, we propose a thresholding method to eliminate the effect of clothing. In this method, the average of data is adopted as the threshold value. Towards this end, the distance of the two samples in comparison is obtained row by row. In other words, each row of the first feature is compared to its corresponding row of the second feature. These distances reflect the similarity of two features and contain the clothing effect. Clearly, rows that are affected by clothing produce higher distances than rows which are not and can be removed by thresholding. Each row with distance higher than the threshold is removed and distances of the remaining rows are used for comparing the two features. Since the number of valid rows is not the same in all cases, the average of valid distances is used as the dissimilarity measure of the two features.

More formally, let  $G = \{g^1, g^2, \dots, g^n\}$  be the gallery set and let  $p^i$  be a probe feature. Suppose that all gallery and probe features are of size  $h \times w$  and for each  $1 \leq s \leq n$  and  $1 \leq j \leq h$  let  $g_j^s$  and  $p_j^i$  denote the  $j$  th row of  $g^s$  and  $p^i$  respectively. The distance between corresponding rows of  $g^s$  and  $p^i$  is defined as:

$$D_{sj} = d(p_j^i, g_j^s), \tag{5}$$

where  $d$  denotes the Euclidean distance. For every gallery feature  $g^s$ , the average value of the distances  $D_{sj}$  is used as a threshold to filter out outliers in  $D_{sj}$ :

$$t_s = \frac{1}{h} \sum_{j=1}^h D_{sj} \tag{6}$$



Hence, the following set

$$I_s = \{j : D_{sj} \leq t_s\} \quad (7)$$

is the collection of (indices of) admissible rows of the probe when compared with the gallery feature  $g^s$ .

The average value of admissible row distances is used as the dissimilarity measure between  $g^s$  and  $p^i$ .

$$A_s = \frac{1}{|I_s|} \sum_{j \in I_s} D_{sj} \quad (8)$$

#### 4.4.2 Refined threshold

To further refine the results, we use the ratio of the obtained distance for each gallery feature to the average distance of all gallery features. Let  $D$  denote the average of  $D_{sj}$  for all  $1 \leq s \leq n$  and  $1 \leq j \leq h$ , then

$$r_s = \frac{t_s}{D} \quad (9)$$

And the refined threshold is defined as

$$\hat{t}_s = t_s + \sigma_s \log r_s \quad (10)$$

Where  $\sigma_s$  denotes the standard deviation of  $D_{sj}$  for all  $1 \leq j \leq h$ . In this expression, the term  $\log r_s$  fine-tunes the threshold for each gallery feature according to the similarity of the probe to all gallery features. More specifically, if  $t_s$  is less than the average distance  $D$ , the gallery feature  $g^s$  is more similar to the probe feature  $p^i$  than the average of all gallery features. In this case  $r_s < 1$  has a negative log which yields in a smaller threshold  $\hat{t}_s$  for this gallery feature. This results in the comparison between the probe feature  $p^i$  and the gallery feature  $g^s$  to be performed using fewer rows, thus producing a more exact result. Similarly, for the case  $r_s > 1$ , the opposite effect is observed. In this way, the thresholds of the most similar features decrease and so the comparison among them is more focused on rows which produce smaller distances. In other words, the comparison is focused on most similar parts of most similar cases.

#### 4.4.3 Optimistic and pessimistic thresholds

Although applying the refined threshold leads to acceptable results, on average, it eliminates half of the data regardless of clothing type. The drawback is that while in complex cases where there are more outliers than simple cases, a low threshold helps eliminate outliers, in simple cases that contain less outliers causes removal of some valid data. So the threshold in simple cases should be higher than the threshold used in complex ones. To accommodate this concern, we define two additional thresholds:

$$t_s^+ = t_s + \sigma_s \quad (11)$$

$$t_s^- = t_s - \sigma_s \quad (12)$$

We can consider these cases as optimistic and pessimistic thresholds.

### 4.5 Classification

For each list  $L_i^+$ ,  $L_i^-$  and  $L_i^0$ , we sort the gallery features according to the corresponding distances. To simplify the notation, for each threshold, we refer to these lists as  $L^i$ . Therefore, the result would be  $Y^i = \{y^1, y^2, \dots, y^n\}$  which is obtained by sorting  $L^i$  and  $Z^i = \{z^1, z^2, \dots, z^n\}$  containing corresponding subjects. With this assumption,  $z^1$  corresponds to the subject with the least distance to the probe and  $y^1$  denotes its distance to the probe. We call the  $z^{i,1}$  as candidate subject for  $p^i$ . The algorithm of this module is presented in Algorithm 3.

---

**Algorithm 3** Classification module

---

**Input:** Gallery subjects,  $G$

List of distances,  $L^i$

**Output:** Sorted distances,  $Y^i$

Sorted gallery subjects according to distances,  $Z^i$

---

1. Sort  $L^i$  ascending and permute  $G$  accordingly to get  $Y^i$  and  $Z^i$
- 

### 4.6 Decision fusion

The final result is obtained in the decision fusion module from the candidates computed by the classification module. In this work, we study two approaches for the decision fusion module. The first approach simply uses majority of candidates for all  $p^i$ s, and if majority fails, i.e., if more than one case reaches majority, the one with minimum distance is chosen. The algorithm of this module is shown in Algorithm 4.

The second approach selects the final result based on computed distances. Towards this end, a certainty measure is computed for each of the candidates, and the candidate with the highest total certainty is selected as the final result. The certainty measure of the candidate in each list is computed based on its distance to the probe by combining the distance of the next best subject from the same list. Suppose that  $x$  is the smallest index of the list  $Z^i$  such that  $z^{i,x} \neq z^{i,1}$ . Then we define the certainty measure as:

$$\text{certainty measure} = \frac{y^{i,x} - y^{i,1}}{y^{i,1}} \tag{13}$$

Note that since we have extracted one feature per each complete gait cycle, there may be more than one feature for each subject. Therefore,  $z^{i,x}$  is not necessarily the second element of the list  $Z^i$ . It is obvious that larger values of  $y^{i,x} - y^{i,1}$  and smaller values of  $y^{i,1}$  mean more confidence in the candidate  $z^{i,1}$ . Hence, the subject with higher values of this certainty measure is more probable to be the correct case. If a subject is the candidate of more than one list, we add the certainty measures of that subject. Finally, the subject with highest total of certainty measure is chosen as the final result. The algorithm of this module is shown in Algorithm 5.

---

**Algorithm 4** Decision fusion based on majority

---

**Input:** Sorted gallery subjects,  $\{Z^i \mid i = 1, 2, \dots, m\}$

**Output:** A subject

---

1.  $C \leftarrow \{Z^{i,1} \mid i = 1, 2, \dots, m\}$
  2. Select majority( $C$ ) as result
-

**Algorithm 5** Decision fusion based on certainty measure

**Input:** Sorted distances,  $Y^i$

Sorted gallery subjects according to distances,  $Z^i$

**Output:** A subject

1. for  $i \in \{1, 2, \dots, m\}$
2. Find the least  $x$  such that  $z^{i,x} \neq z^{i,1}$
3.  $mc^i \leftarrow \frac{y^{i,x} - y^{i,1}}{y^{i,1}}$
4. end for
5. Compute total of  $mc^i$  for each subject
6. Select the subject with maximum  $mc$

### 5 Evaluation

For evaluation of the proposed method we use the OU-ISIR Gait Database, the Treadmill Dataset B [16] and the CASIA Gait Database, Dataset B [27]. The first one contains walking cycles of 48 persons with up to 32 variations of clothing each as shown in Fig. 6, Tables 1 and 2. There is one sequence of type 9 for each subject that corresponds to normal clothing and is used as gallery and the other types are used as probe samples. The CASIA Gait Database, Dataset B contains walking cycles of 124 subjects with three different clothing types each (Fig. 7). There are ten sequences per subject at perpendicular viewing angle, two of which are captured with the subject carrying a bag or a backpack, two sequences with the subject wearing a coat, and six sequences captured with normal clothing. Four of the sequences with normal clothing are used as gallery and the other six sequences as probe data. Evaluation results are illustrated in following subsections.



**Fig. 6** Clothing samples of OU-ISIR Gait Database, the Treadmill Dataset B [16]

**Table 1** List of clothes used in OU-ISIR Gait Database, the Treadmill Dataset B [16]

Abbreviation	Name	Abbreviation	Name	Abbreviation	Name
RP	Regular Pants	HS	Half Shirt	CW	Casual Wear
BP	Baggy Pants	FS	Full Shirt	RC	Rain Coat
SP	Short Pants	LC	Long Coat	Ht	Hat
Sk	Skirt	Pk	Parker	Cs	Casquette Cap
CP	Casual Pants	DJ	Down Jacket	Mf	Muffler

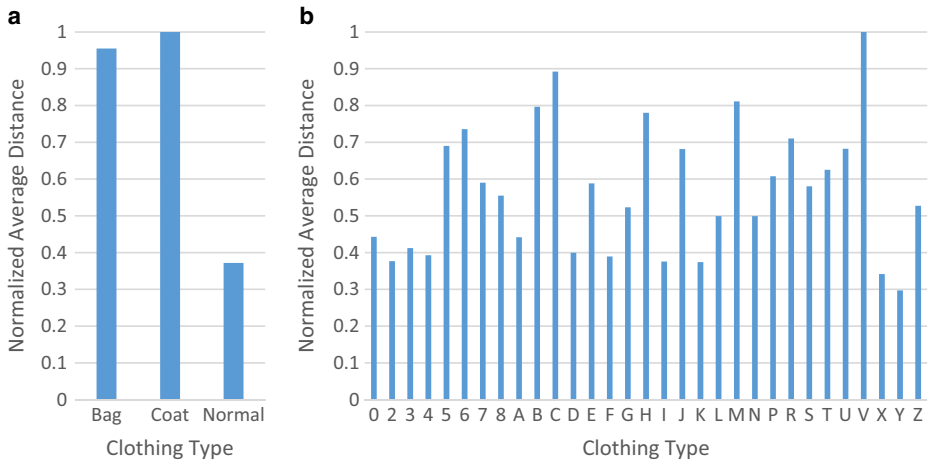
**Table 2** Different clothing combinations in OU-ISIR Gait Database, the Treadmill Dataset B [16]

#	S <sub>1</sub>	S <sub>2</sub>	S <sub>3</sub>	#	S <sub>1</sub>	S <sub>2</sub>	#	S <sub>1</sub>	S <sub>2</sub>
2	RP	HS	-	A	RP	Pk	T	Sk	FS
3	RP	HS	Ht	B	RP	DJ	U	Sk	Pk
4	RP	HS	Cs	I	BP	HS	V	Sk	DJ
9	RP	FS	-	K	BP	FS	D	CP	HS
X	RP	FS	Ht	J	BP	LC	F	CP	FS
Y	RP	FS	Cs	L	BP	Pk	E	CP	LC
5	RP	LC	-	M	BP	DJ	G	CP	DJ
6	RP	LC	Mf	N	SP	HS	H	CP	DJ
7	RP	LC	Ht	Z	SP	FS	0	CP	CW
8	RP	LC	Cs	P	SP	Pk	R	RC	-
C	RP	DJ	Mf	S	Sk	HS			

## 5.1 The effect of clothing

We begin by studying the similarity between samples belonging to the same person. The measure used here is the average distance between pairs of samples one with normal clothing and the other with any prescribed type of clothing. For ease of comparison, these average distances are normalized with respect to the maximum value found. Fig. 8a shows the values obtained for the CASIA database, where it is observed that the average distance for the sequences in which the subject carries a bag or wears a coat is at least 2.5 times as large as the average distance between samples with normal clothing. Figure 8b shows results of similar comparisons for the OU-ISIR database where type 9 is taken as normal clothing. The effect of clothing variations is also visible in these results. In particular, it can be seen that clothing types B, C, M, V have high average distance from the normal clothing which makes identification of the subject with these clothing types more challenging. On the other hand, clothing types 2, I, X, Y seem to be easier to handle.

**Fig. 7** Clothing samples of CASIA Gait Database, Dataset B (a) carrying a bag, (b) wearing a coat and (c) normal case



**Fig. 8** Normalized average distances between samples with normal clothing and other clothing types for the same person in (a) CASIA Gait Database, Dataset B (b) OU-ISIR Gait Database, the Treadmill Dataset B

### 5.2 Effectiveness of the proposed method

Since the proposed method removes outlier rows before computing the average row distance between two samples, the distance it computes for two samples is less than the distance computed from all rows. This does not depend on whether the two samples at hand belong to the same person or not. While this behavior is desired when comparing samples belonging to the same subject, it is not welcome when comparing samples from different subjects. To study the trade-off between these two sides, we define a discrimination ratio as follows:

$$\text{discrimination ratio} = \frac{\text{average intra-subject distance}}{\text{average inter-subject distance}} \tag{14}$$

Clearly, higher values of inter-subject distance and lower values of intra-subject distance leads to better discrimination. Hence, a lower discrimination ratios means better recognition of samples belonging to the same subject.

Tables 3 and 4 respectively present discrimination ratios computed for the CASIA and OU-ISIR datasets, according to various methods presented in this work. For better visibility, the smallest and the largest values in each row of these tables are displayed in bold and bold italic respectively. It is observed that with easier clothing types (as explained in 5.1) the refined method provides the best separation, while for more challenging clothing types the pessimistic method has the best performance. The first column in either table corresponds to computation of the average distance using all rows. It is seen in these tables that in most cases this is the

**Table 3** The discrimination ratio for different clothing types in CASIA Gait Database, Dataset B

	ALL	Baseline	Refined	Optimistic	Pessimistic
Bag	<b><i>0.783764</i></b>	0.721085	<b>0.664581</b>	0.752496	0.693808
Coat	<b><i>0.802069</i></b>	0.651106	<b>0.611203</b>	0.749114	0.616425
Normal	0.539385	0.548853	<b>0.462490</b>	0.546430	<b><i>0.577931</i></b>

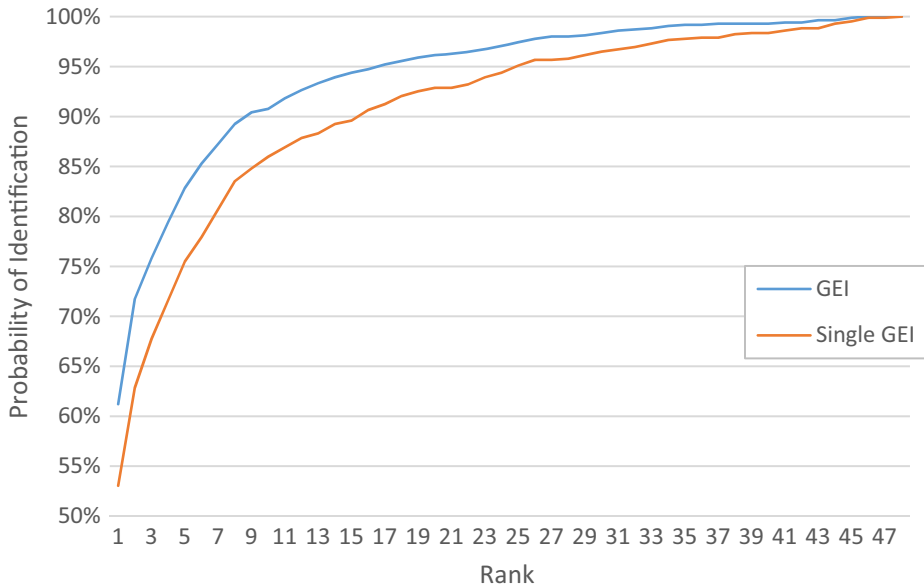
**Table 4** The discrimination ratio for different clothing types in OU-ISIR Gait Database, the Treadmill Dataset B

	ALL	Baseline	Refined	Optimistic	Pessimistic
0	<i>0.700324</i>	0.685711	<b>0.615489</b>	0.695025	0.673148
2	<i>0.599106</i>	0.537607	<b>0.459979</b>	0.577149	0.509427
3	<i>0.643179</i>	0.582403	<b>0.514936</b>	0.617299	0.547118
4	<i>0.689949</i>	0.660771	<b>0.602481</b>	0.670835	0.643544
5	<i>0.808625</i>	0.733520	0.687803	0.770478	<b>0.678162</b>
6	<i>0.842546</i>	0.770151	0.727884	0.822232	<b>0.726161</b>
7	<i>0.765215</i>	0.712529	<b>0.647393</b>	0.729945	0.663895
8	<i>0.758080</i>	0.691473	<b>0.628152</b>	0.716475	0.629771
A	<i>0.652129</i>	0.591643	<b>0.514445</b>	0.614125	0.535187
B	<i>0.827520</i>	0.674858	0.642528	0.757889	<b>0.606125</b>
C	<i>0.851448</i>	0.686280	0.657169	0.772172	<b>0.618423</b>
D	<i>0.652015</i>	0.642225	<b>0.566734</b>	0.640346	0.624701
E	<i>0.791646</i>	0.743588	<b>0.685801</b>	0.761612	0.702432
F	<i>0.652793</i>	0.634735	<b>0.556746</b>	0.643691	0.617400
G	<i>0.741068</i>	0.706290	<b>0.647556</b>	0.706462	0.663157
H	<i>0.857668</i>	0.759751	0.723285	0.839592	<b>0.680413</b>
I	<i>0.658705</i>	0.636885	<b>0.566303</b>	0.648513	0.620536
J	<i>0.811523</i>	0.760800	<b>0.715352</b>	0.776745	0.730285
K	0.641799	<b>0.643416</b>	<b>0.564024</b>	0.639453	0.630791
L	<i>0.713951</i>	0.681261	0.604824	0.678085	<b>0.601590</b>
M	<i>0.841832</i>	0.731906	0.699899	0.791940	<b>0.659624</b>
N	<i>0.688394</i>	0.629517	<b>0.546334</b>	0.648410	0.558917
P	<i>0.770368</i>	0.722846	<b>0.663087</b>	0.734827	0.668636
R	<i>0.847382</i>	0.771242	<b>0.734751</b>	0.806700	0.741081
S	<i>0.794310</i>	0.653699	0.612260	0.760145	<b>0.597994</b>
T	<i>0.755640</i>	0.664751	<b>0.605017</b>	0.708317	0.615439
U	<i>0.818057</i>	0.744413	<b>0.698098</b>	0.794322	0.710178
V	<i>0.931668</i>	0.888333	0.858194	0.911767	<b>0.796878</b>
X	<i>0.607067</i>	0.589217	<b>0.496511</b>	0.591438	0.544549
Y	<i>0.560956</i>	0.555623	<b>0.456455</b>	0.553845	0.519151
Z	<i>0.705838</i>	0.682208	<b>0.601656</b>	0.666150	0.605869

least effective method, thus supporting the effectiveness of the proposed method. This claim is further supported by observing in Table 3 that even for normal clothing in the CASIA dataset, the refined method gives the best result. Therefore, while applying this method leads to mitigate the clothing effect and improves the recognition of these cases, has not negative effect on recognizing normal cases.

### 5.3 Cycle estimation and decision fusion

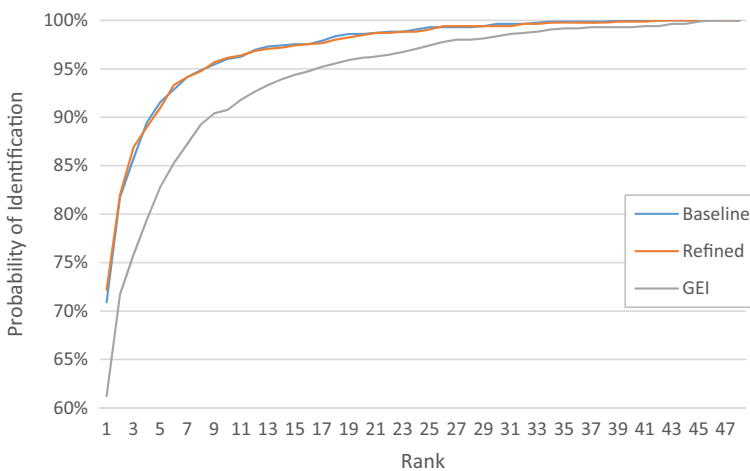
Each sample of the OU-ISIR Gait Database, the Treadmill Dataset B contains 360 frames that form multiple gait cycles. Samples in CASIA Gait Database, Dataset B have multiple gait cycles too. As described before, instead of using the most appropriate cycle or using all frames to make the desired feature, we extract complete gait cycles, perform recognition on each independently, and use decision fusion module to obtain the final result. Figure 9 compares the performance of using all frames for making single GEI with making GEI from each cycle, performing recognition on each and using majority to obtain the result. As we can see, using this approach improves the results significantly.



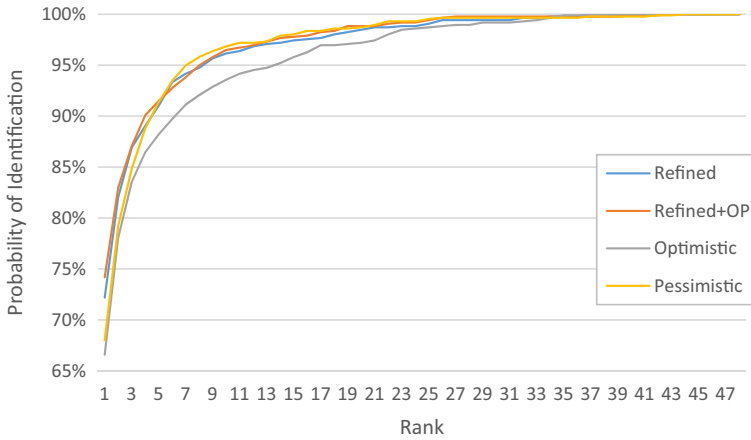
**Fig. 9** CMC curves representing the effect of cycle detection and decision fusion in OU-ISIR Gait Database, the Treadmill Dataset B

#### 5.4 Baseline and refined methods

To observe the effectiveness of the proposed baseline and refined methods the CMC curves for cases with and without using these methods is presented in Fig. 10. It is evident that using the proposed method to eliminate outliers increases the recognition rate by about 10% and using refined threshold further improves the results.



**Fig. 10** CMC curves representing the effect of the proposed baseline and refined methods comparing to pure GEI on OU-ISIR Gait Database, the Treadmill Dataset B. The rank-1 values are 70.91, 72.19 and 61.21% respectively



**Fig. 11** CMC curves representing the effect of optimistic, pessimistic, refined threshold and their combination in OU-ISIR Gait Database, the Treadmill Dataset B. The rank-1 values are 66.58, 67.99, 72.19 and 74.18% respectively

### 5.5 Optimistic and pessimistic thresholds

Figure 11 demonstrates the effect of using optimistic and pessimistic thresholds. We can see that both optimistic and pessimistic thresholds lead to lower recognition rates but in combination with refined threshold and using majority for decision fusion, the results are improved. Another important observation is that choosing average distance as threshold is a proper choice since lower and higher thresholds lead to worse results.

The recognition rate per each type of clothing for these thresholds is shown in Fig. 12. We can see that as expected, the optimistic threshold has better performance in less complex cases like types 0, D, G while the pessimistic threshold performs much better in complex cases such as types 5, C, H, S, U, and V.

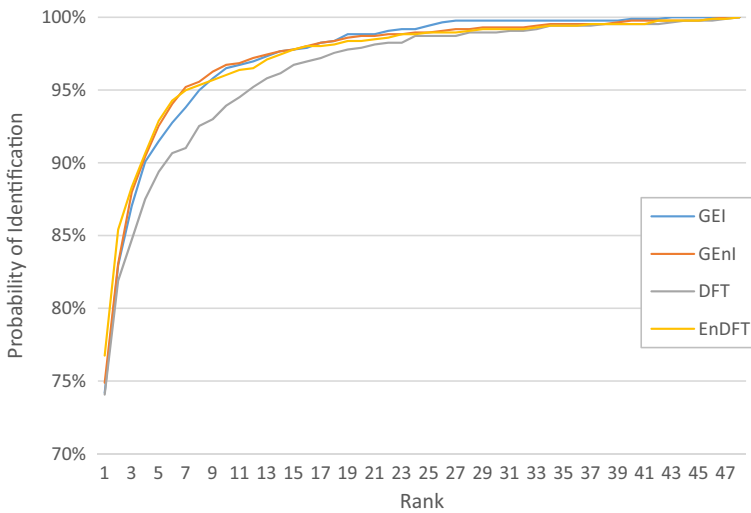
### 5.6 Different features

In Fig. 13, the performance of the proposed method using different features is illustrated. We can see that these features perform well with proposed approach. While GEI as a simple feature is used in



**Fig. 12** Rank-1 recognition rates for different thresholds per clothing type in OU-ISIR Gait Database, the Treadmill Dataset B



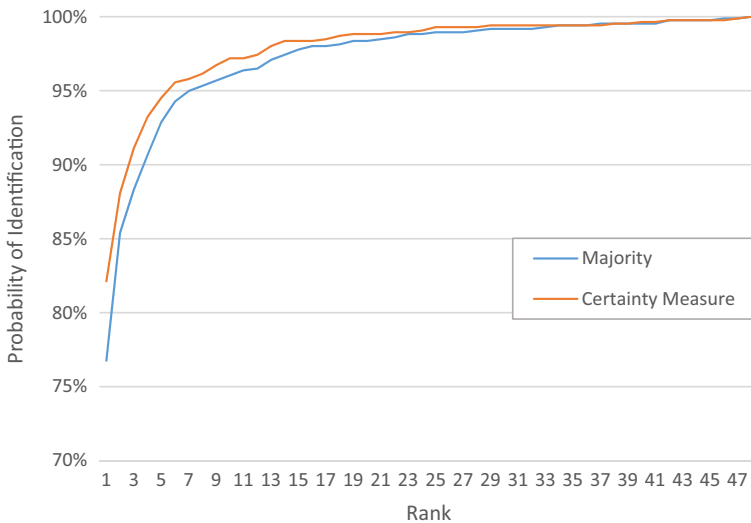


**Fig. 13** CMC curves illustrating the performance of GEI, GENI, DFT and EnDFT features using the proposed method in OU-ISIR Gait Database, the Treadmill Dataset B. The rank-1 values are 74.18, 74.88, 74.06 and 76.75% respectively

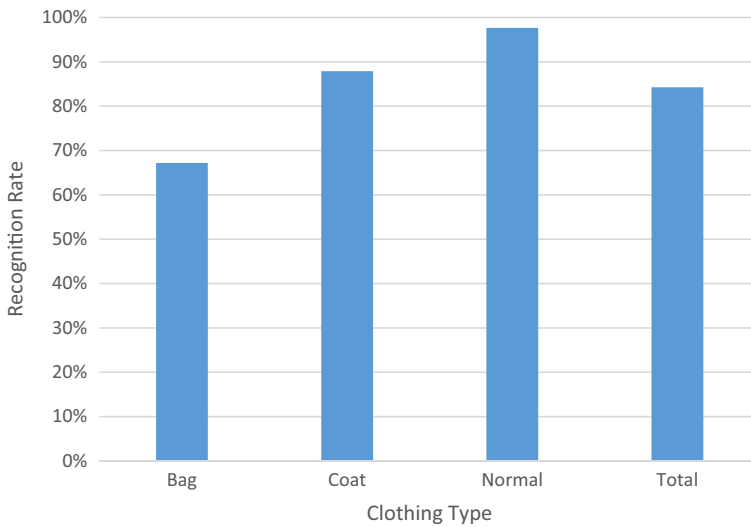
previous sections to illustrate the performance of the proposed method, we can see that GENI, DFT and EnDFT which are more sophisticated features perform well with this method.

### 5.7 Decision fusion

As described in section 4.6, we use two different methods in decision fusion module. The results obtained using these methods are presented in Fig. 14. We can see that using certainty measure improves the recognition rate about 6% that shows the effectiveness of this method.



**Fig. 14** CMC curves illustrating the performance of decision fusion methods in OU-ISIR Gait Database, the Treadmill Dataset B. The rank-1 values are 76.65 and 82.12% for majority and certainty measure respectively



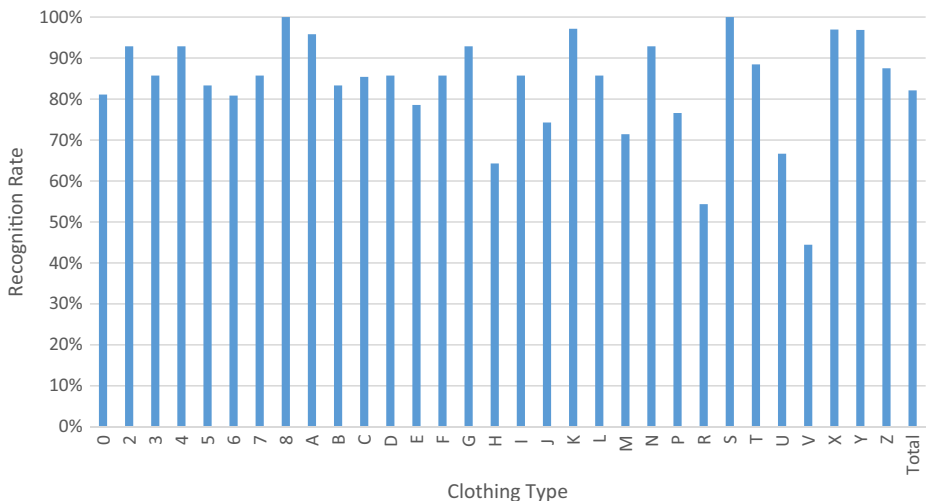
**Fig. 15** Recognition rates per clothing types on CASIA Gait Database, Dataset B

### 5.8 Recognition per clothing types

Figures 15 and 16 shows the recognition rates per each clothing type on CASIA and OUISIR databases respectively. It is obvious that the proposed method improves the recognition rates of complex clothing types and has not negative effect on normal and simple types. These results also confirms the results obtained in section 5.2 about the discrimination ratios.

### 5.9 Comparison with existing methods

Table 5 compares the performance of the proposed method with [21] and [7] on the OU-ISIR Gait Database, the Treadmill Dataset B. Also the comparison with [21] is presented in Table 6. We can



**Fig. 16** Recognition rates per clothing types on OU-ISIR Gait Database, the Treadmill Dataset B

**Table 5** Comparison with other methods on the OU-ISIR Gait Database, the Treadmill Dataset B

Algorithm	Rank-1	Rank-5	Rank-10
[21]	71.02%	89.71%	96.02%
[7]	80.44%	–	–
Proposed	82.13%	94.51%	97.20%

**Table 6** Comparison with other methods on the CASIA Gait Database, Dataset B

Algorithm	Rank-1	Rank-5	Rank-10
[21]	77.69%	90.07%	93.06%
Proposed	84.24%	92.71%	95.18%

see that the proposed method outperforms these works. Note that the proposed method does not use precomputed fixed parts, which makes it more reliable, realistic, and flexible in real applications. Besides, it does not use PCA or any other costly methods that makes it more suitable for using in real applications.

## 6 Conclusion

This paper proposes an adaptive outlier detection method to mitigate the problem of clothing effect in human gait recognition. The main idea is to compare most similar parts of samples and ignore rough variations. Towards this end, the distances of two features are calculated row by row and the average of these distances is used as threshold. The rows that have distances smaller than the threshold are considered as valid data and the average of these rows is used as similarity measure of features. The threshold is further refined and considering pessimistic and optimistic situations leads to some candidates. The final result is then obtained using decision fusion methods that use majority and certainty measure. It is worth to mention that the thresholds are adaptive and there is no need for manual tuning. More specifically for each pair of features, these thresholds varies dynamically to discard clothing variations and maintain valid information. The evaluation results show that this method is effective and increases the recognition rate significantly.

## References

1. Bashir K, Xiang T, Gong S (2008) Feature selection on gait energy image for human identification. in Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE Int Conf IEEE
2. Bashir K, Xiang T, Gong S (2010) Gait recognition without subject cooperation. *Pattern Recogn Lett* 31(13):2052–2060
3. Bobick AF, Davis JW (2001) The recognition of human movement using temporal templates. *Patt Anal Mach Int, IEEE Trans* 23(3):257–267
4. Bobick AF, Johnson AY (2001) Gait recognition using static, activity-specific parameters. in *Computer Vision and Pattern Recognition. CVPR 2001. Proc IEEE Comput Soc Conf IEEE*
5. Cunado D, Nixon MS, Carter JN (2003) Automatic extraction and description of human gait models for recognition purposes. *Comput Vis Image Underst* 90(1):1–41
6. Dockstader SL, Berg MJ, Tekalp AM (2003) Stochastic kinematic modeling and feature extraction for gait analysis. *Image Proc, IEEE Trans* 12(8):962–976
7. Guan Y, Li C-T, Hu Y (2012) Robust clothing-invariant gait recognition. *Int Info Hiding Mult Sig Proc (IIH-MSP) Eighth Int Conf IEEE*
8. Haiping LP, Konstantinos NV, Anastasios N (2008) A full-body layered deformable model for automatic model-based gait recognition. *EURASIP J Adv Sign Proc*
9. Han J, Bhanu B (2006) Individual recognition using gait energy image. *Patt Anal Mach Int, IEEE Trans* 28(2):316–322

10. Hossain A, Makihara Y, Wang J, Yagi Y (2010) Clothing-invariant gait identification using part-based clothing categorization and adaptive weight control. *Pattern Recogn* 43(6):2281–2291
11. Lam TH, Cheung KH, Liu JN (2011) Gait flow image: a silhouette-based gait representation for human identification. *Pattern Recogn* 44(4):973–987
12. Lee S, Liu Y, Collins R (2007) Shape variation-based frieze pattern for robust gait recognition. *Comput Vis Patt Recog CVPR'07. IEEE Conf IEEE*
13. Li X, CHEN Y (2013) Gait recognition based on structural gait energy image. *J Comput Info Syst* 9(1):121–126
14. Li X, Maybank SJ, Yan S, Tao D, Xu D (2008) Gait components and their application to gender recognition. *Syst, Man, Cyber, Part C: Appl Rev, IEEE Trans* 38(2):145–155
15. Liu J, Zheng N (2007) Gait history image: a novel temporal template for gait recognition. in *Multimedia and Expo. IEEE Int Conf IEEE*
16. Makihara Y, Mannami H, Tsuji A, Hossain MA, Sugiura K, Mori A, Yagi Y (2012) The OU-ISIR gait database comprising the treadmill dataset. *IPSI Trans Comput Vis Appl* 4:53–62
17. Makihara Y, Sagawa R, Mukaigawa Y, Echigo T, Yagi Y (2006) Gait recognition using a view transformation model in the frequency domain, in *Computer Vision—ECCV 2006. Springer* 151–163
18. Piccardi M (2004) Background subtraction techniques: a review. *Syst, Man Cyber IEEE Int conf IEEE*
19. Rokanujjaman M, Hossain MA, Islam MR (2012) Effective part selection for part-based gait identification. *Elect Comput Eng (ICECE) 2012 7th Int Conf*
20. Rokanujjaman M, Hossain M, Islam M (2013) Effective part definition for gait identification using gait entropy image. *Info, Elect Vis (ICIEV), Int Conf IEEE*
21. Rokanujjaman M, Islam MS, Hossain MA, Islam MR, Makihara Y, Yagi Y (2013) Effective part-based gait identification using frequency-domain gait entropy features. *Mult Tools Appl* 1–22
22. Wang L, Ning H, Tan T, Hu W (2004) Fusion of static and dynamic body biometrics for gait recognition. *Circ Syst Video Technol, IEEE Trans* 14(2):149–158
23. Wang L, Tan T, Ning H, Hu W (2003) Silhouette analysis-based gait recognition for human identification. *Patt Anal Mach Int, IEEE Trans* 25(12):1505–1518
24. Whytock TP, Belyaev A, Robertson NM (2013) Towards Robust Gait Recognition. *Adv Vis Comput. Springer* 523–531
25. Yang X, Zhou Y, Zhang T, Shu G, Yang J (2008) Gait recognition based on dynamic region analysis. *Signal Process* 88(9):2350–2356
26. Yogarajah P, Condell JV, Prasad G (2011) P<inf>RW</inf>GEI: Poisson random walk based gait recognition. in *Image and Signal Processing and Analysis (ISPA). 7th Int Symp IEEE*
27. Yu S, Tan D, Tan T (2006) A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition. *Patt Recog. ICPR 2006. 18th Int Conf IEEE*
28. Zhang R, Vogler C, Metaxas D (2007) Human gait recognition at sagittal plane. *Image Vis Comput* 25(3):321–330
29. Zhang E, Zhao Y, Xiong W (2010) Active energy image plus 2DLPP for gait recognition. *Signal Process* 90(7):2295–2302



**Abbas Ghebleh** is a PhD student of Software Engineering at Shahid Beheshti University, Tehran, Iran. He received his B.Sc. and M.Sc. in Software Engineering from Shahid Beheshti University, Tehran, Iran (2006 and 2010). His research interests are Digital Image Processing and Computer Vision including human gait recognition and action recognition.



**Dr. Mohsen Ebrahimi Moghaddam** is an associate professor in electrical and computer engineering department, Shahhid Beheshti university of Iran since 2006. He has got his Ph.D., M.Sc., and B.Sc. from Sharif University in Iran. His research interests are image processing and pattern recognition specially using artificial intelligence techniques such as image security, watermarking, deblurring, and biometrics. He is the image processing lab chief in his department.