Gait Recognition Using Self-Adaptive Hidden Markov Models

Xiuhui Wang

China Jiliang University, Hangzhou 310018 China

Wei Qi Yan

Auckland University of Technology, 1010 New Zealand

Abstract: Human gait recognition has intensive applications in multiple fields, such as video-based surveillance, digital security and forensics, etc. In order to enhance the recognition rate, this paper studies a gait recognition scheme based on self-adaptive Hidden Markov Models (SAHMM). Firstly, we present a novel feature extraction algorithm to calculate the Local Gait Energy Image (LGEI) and construct observation vector sets. Then, a SAHMM-based gait recognition method which adopts a parameter adaptation process to optimize the parameters of gait models is provided. Finally, the proposed method is evaluated extensively on the CASIA Dataset B for cross-walking-condition and cross-view gait recognition, and further evaluated on the OU-ISIR Large Population Dataset to verify its generalization ability with large data. Both experimental results show that the proposed method exhibits superior performance in comparison with those existing methods, and show great potential for practical applications.

Keywords: gait recognition, hidden Markov model, biometrics, video-based surveillance

1 Introduction

Gait Recognition is one of the most promising biometric recognition technologies，which achieves human identification through our walking habits and dynamic characteristics. Compared to other biometrics, gait has a myriad of advantages such as non-offensive, low-resolution and easy data-collection [1]. In recent years, video-based gait recognition has gained tremendous attention from multiple fields, and a number of gait recognition algorithms have been proposed [2]-[6]. There are many challenges in video-based gait recognition, including different camera angles, dressing and carrying conditions, different walking speed, etc. And till now, existing gait recognition methods mainly focused on different view angles, also named cross-view problems, which are widely existing in practical applications. However, currently the cross-view problems are still far from completely resolved, because the gait appearance of one person can be dramatically altered when the view angles changed.

On the other hand, Hidden Markov Models (HMM) [7] has been successfully applied to model temporal information on applications such as natural language recognition [8] and face recognition [9], and achieved remarkable results. Deng and Bao [8] propose a sparse hidden Markov model

(HMM)-based single-channel speech enhancement method that models the speech and noise gains accurately in non- stationary noise environments. In [9], adaptive HMM was used to perform video-based face recognition. Although HMM and its variations have been applied to human behavior understanding, natural language recognition, face recognition and other fields extensively, few of them are dealing with video-based gait recognition.

In this paper, to improve the recognition rate of gait recognition, we propose a novel gait recognition solution based on self-adaptive Hidden Markov Model (SAHMM). To the best our knowledge, this is the first time that this approach is proposed. There are three-fold contributions of the novel method: first, a well-designed feature extraction algorithm is presented based on Local Gait Energy Image (LGEI). Second, a novel gait recognition method based on SAHMM is proposed for classification which takes use of a parameter adaptation process to optimize the parameters of each gait model. Third, we greatly advance the record scores on the CASIA Dataset B and OU-ISIR Large Population Dataset, demonstrating that our method can work well under the condition of gait databases with large capacity.

The remaining sections of this paper are organized as follows. In Section 2, more related works on gait recognition will be introduced. Then, in Section 3, a SAHMM-based gait recognition method will be presented and demonstrated in detail, including gait feature extraction, parameter estimation and adaptation of gait models. Our experiments are carried out based on CASIA gait dataset to evaluate the proposed algorithms in Section 4, the conclusion and future work of this paper will be finally addressed in Section 5.

2 Related work

Most of the existing approaches to video-based gait recognition can be broadly classified into two categories, i.e., those that use GEI-based holistic features [2], [3], [10]-[12] and others that use detailed features[4]-[6], [13]-[15]. Gait Energy Image (GEI) [10] is a spatial-temporal gait feature representation, which is widely used to characterize human walking properties for gait recognition and gait classification. The first kinds of methods are based on the GEI and its varieties, which use the holistic gait information and generally do not need to extract specific gait features or construct 3D models. In [2], by using the GEI as original gait features, the problem of cross-view gait based human identification was investigated via deep convolutional neural networks. Connie et al. [3] combined multi-view matrix representation and a randomized kernel extreme learning machine, and proposed an end-to-end solution for view variation problems under Grassmann manifold treatment. To reduce the deteriorate effect of covariate factors in gait recognition and classification, Guan et al. [11] presented a class ensemble method, which was sensitive to locations of corrupted features and could generalize well to most covariate conditions. More recently, Islam et al. [12] designed a frequency domain gait representation, which was a variety of the GEI, and proposed a wavelet-based feature extraction method.

The second kind of methods use detail gait information, mainly including the structural features and the geometry features from 3D models. Luo et al. [4] utilized a clothes-independent 3D parametric gait model to deal with the variation in speed, inclined plane and clothing faced by gait recognition process. In [5], a covariate conscious approach for gait recognition was proposed, which addressed the variations in clothing and carrying conditions that could have negative impact on recognition performance. Tang et al. [6] assumed a 3D object shares common view surfaces in significantly different view, and addressed the arbitrary-view gait recognition problem by gait partial similarity matching. They estimated each gait pose via a level set energy cost function from silhouettes, and achieved shape deformation by Laplacian deformation energy function. Huang et al. [13] combined the gait structural profile and the shifted energy image to improve the robustness of gait recognition. In [14], a gait recognition system based on HMMs and dual discriminative observations for sub-dynamics analysis was presented. This method deployed both the gait dynamics information and the model-based features to improve the discriminatory capacity of their system. To deal with the large perspective distortion in gait recognition, Abdulsattar et al. [15] proposed an identification technique by reconstructing 3D models of the walking subjects. This technique used the Generic Fourier Descriptors as gait features, and could handle truncated gait cycles of different length.

3 The proposed methodology

Gait feature extraction and classifier design are the key processes in gait recognition problem. In this paper, a well-designed gait feature, named LGEI, and the related extraction algorithm is presented. Besides, HMM-based gait models are constructed, and a parameter adaptation method is proposed to refine them. The proposed method mainly includes four steps: 1) LGEI feature extraction which extracts gait features from the training set, the adapting set and the test set; 2) HMM gait models training. In this step, HMM gait models are constructed and trained using the training set; 3) parameter adaptation for HMM gait models. HMM gait models are refined by a small amount of gait data in adapting set; 4) SAHMM-based gait classification, which takes use of the HMM gait models with refined parameters to classify the gaits in the test set. The system diagram of gait recognition using the method mentioned above is shown in Fig. 1.

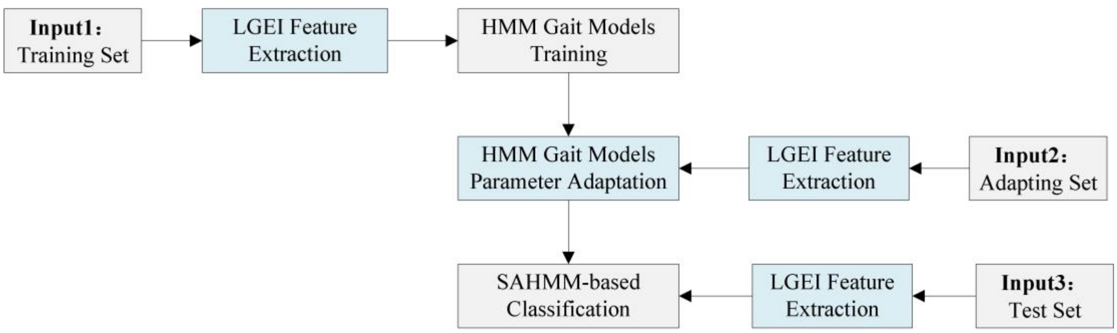


Fig. 1. System diagram of gait recognition using the proposed method

3.1 Gait Feature Extraction

Feature extraction is a crucial step in the course of gait recognition. It is usually necessary to detect and segment a moving object from a given video stream or motion pictures before gait feature extraction. At present, the algorithms of moving object detection and segmentation are well developed [16, 17], therefore, the focus of this paper will be on gait feature extraction and gait recognition. Amid gait feature extraction and HMM-based gait models training as well as gait recognition, we take gait contour into consideration directly.

According to characteristics of human legs, the gait of normal walking has three states: in the state of Legs Together(S_1)when the two legs close together, the two legs are in the same plane with a human body, including lifting up left foot to the side of right leg, lifting up the right foot to the side of left leg and normal standing; the state Left Front Right Back(S_2) means that the left foot is on front of the right one; the state S_3(Left Back Right Front) refers that the right foot is on front of the left one, as shown in Fig 2. Thus, a complete natural gait cycle is defined as $S_1 \rightarrow S_2 \rightarrow S_1 \rightarrow S_3 \rightarrow S_1$, alternatively, $S_1 \rightarrow S_3 \rightarrow S_1 \rightarrow S_2 \rightarrow S_1$ according to our walking habits of individuals. It should be noted that the gait cycle segmentation method is related to the direction of camera pointing in, when the view angle is suitable enough, the segmentation works very well.

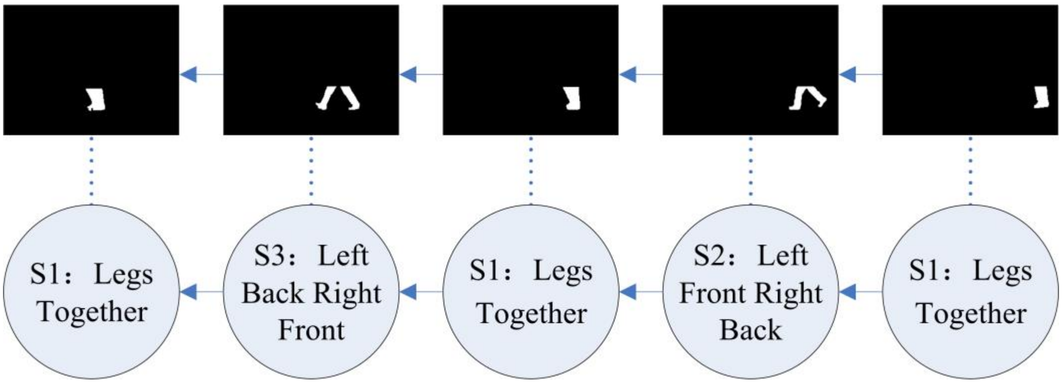
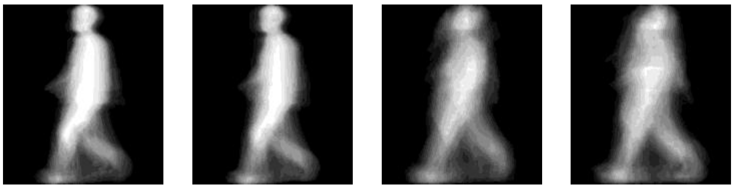


Fig.2 A natural gait cycle

After the completion of gait cycle segmentation, we can extract the gait features based on LGEI. As the first gait representation [10], given the binary gait silhouette images $B_t(x, y)$ at time $t$in a gait cycle, the Gait Energy Image (GEI) is defined as:

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

where N is the number of video frames in a complete cycle, t is the frame index, $x$and y arethe coordinates of an image pixel. Figure 3 shows examples of GEI in a complete cycle, where (a) and (b) are from the same person, but with different angles of view; (c) and (d) are from another person with different view angles also.



(a) (b) (c) (d)

Fig.3 Examples of GEI in a compete cycle

Generally, GEI reflects major shapes of gait silhouettes related to variations over a complete gait cycle. GEI has three basic characteristics [18]: (1) all the silhouette images are the space-normalized energy images of the target individual; (2) GEI is seen as the time-normalized accumulative energy image of the target individual in a complete gait cycle; (3) what pixels with bigger values in GEI means that actions occur more frequently at these points. In this paper, we refine the GEI by splitting each gait cycle into several segments and calculate relevant LGEIs on each segments, as shown in Figure 4. Compared with traditional GEI representation based on a complete gait cycle, LGEI utilizes more important details of gait changes in each gait cycle, and reduces sensitivity to variations in the view angle and other factors. The proposed feature extraction algorithm is given in Algorithm 1 by summarizing the above results.

Algorithm 1.LGEI-based gait feature extraction

Input: Periodic gait data of the person *X*

Output: Feature vector set of the person *X*

Step 1.*Find the key frames*. Divide each gait cycles in the periodic gait dataset **P**={V_i, 1≤i≤N_C}, and get a sequence of key frames I={I_j,1≤j≤N_F}, where N_F≥5, and the value of N_F= 5 has been adopted in our experiments.

Step 2.*Calculate LGEI for each key frame***.** The calculation of LGEI within the small neighborhood of each key frame is shown as Figure 4. It is easy to see that the use of GEI within a small neighborhood to replace a single key frame can reduce the key information loss caused by simple hard segmentation；

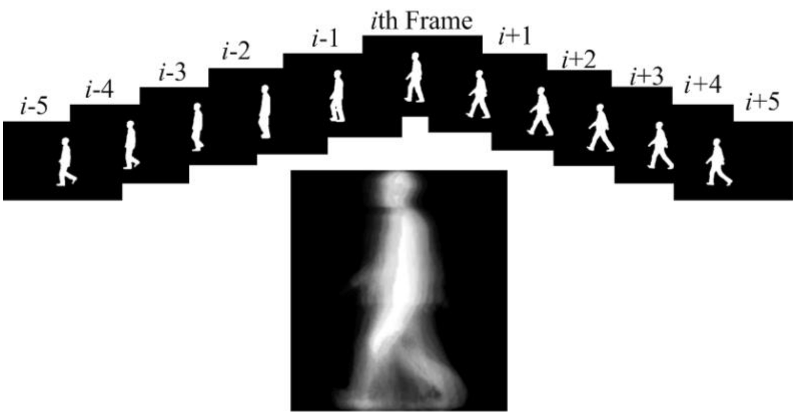


Fig. 4 Diagram of LGEI calculation

Step 3. *Construct feature vector set*. If W and H denote the width and height of each LGEI respectively, then the related original feature vector set **S** is given by,

$$\mathbf{S} = \left[\mathbf{s}1, \mathbf{s}2, ..., \mathbf{s}_W\right]^T \tag{2}$$

where $\mathbf{s}_k$ (k=1, 2,…,W) are H-dimensional feature vectors.

Step 4. Reduce dimension of each original feature vector using Principal Component Analysis (PCA) and obtain a new low-dimension observational state set **O**, which is the final feature vector set,

$$\mathbf{O} = \left\{o_{i,j,k} \middle| \ i \in \left[1,\ N_F\right],\ j \in \left[1, N_C\right],\ k \in \left[1,\ N_V\right]\right\} \tag{3}$$

where N_F is the number of key frames extracted from each gait cycle, N_C is the number of total gait cycles segmented from each gait image sequence, N_V is the number of gait image sequences for the tracking person. The feature vector set **O** will be partitioned into a training set and a test set, which will be applied to parameter estimation and gait model test in Section 3.2 and Section 3.3.

3.2 HMM Gait Model Training

Suppose the state set of gait observation of the given person m is $S_m\left\{s_k \middle| k \in \left[1,\ N_S\right]\right\}$, where N_s is the number of implicit states, we construct a SAHMM gait model $\lambda_m\left(\mathbf{A}, \mathbf{B}, \boldsymbol{\pi}\right)$. **A** is the state transition matrix which is defined as:

$$\mathbf{A} = \left\{a_{ij} = P\left(S_j \middle| S_i\right) \middle| 1 \le i,\ j \le 5\right\} \tag{4}$$

where a_{ij} is the probability that the implicit state is S_i at time t, the implicit state is S_j at time$t$+1. **B** is the confusion matrix defined as,

$$\mathbf{B} = \left\{b_{ij} = P\left(O_j \middle| S_i\right) \middle| 1 \le i \le 5, 1 \le j \le N_F\right\} \tag{5}$$

Where b_{ij} is the probability at time t the implicit state is S_i and the observed state is O_j; N_F stands for the number of observable states. Considering a person's walk normally is periodic and the gait cycles are relatively stable, we extract five key gaits as the implicit states of SAHMM model, i.e. S_1(Two Feet Together), S_2(Right Foot Through the Left Side), S_3 (Left Foot Through the Right Side), S_4(Left Front Right Back), S_5(Right Front Left Back) as shown in Figure 5.

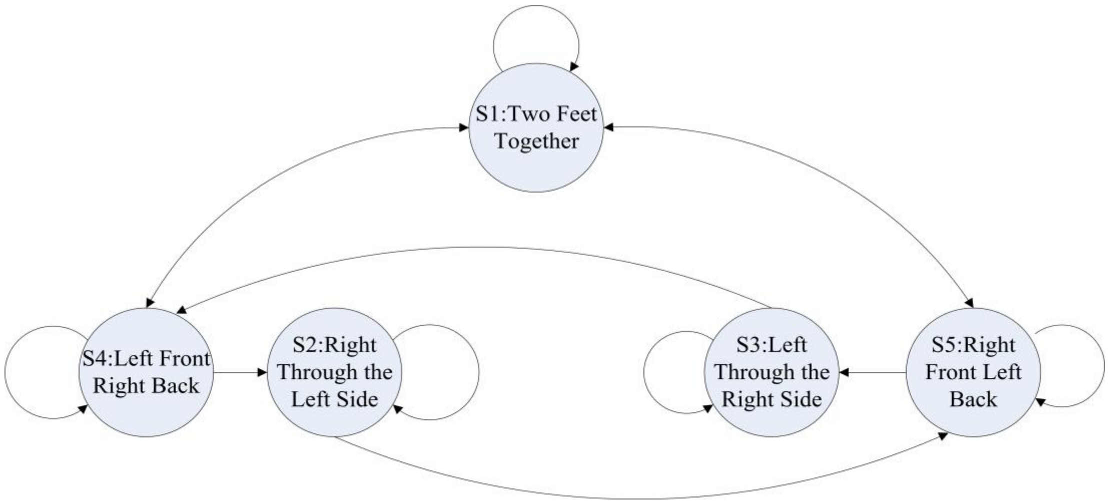


Fig.5 The hidden state transition diagram of the proposed SAHMM

By default, the gait cycle in our experiments starts from the state of two legs standing still, in most cases the gait begins from the state S_1 and seldom from S_4 and S_5. Thus, the state probability $\boldsymbol{\pi}$ is initialized as,

$$\begin{cases} \pi_1 = 0.4 \\ \pi_2 = \pi_3 = 0.3 \\ \pi_4 = \pi_5 = 0 \end{cases} \tag{6}$$

Combined with the actual situation of normal walking, we set the initial values of state transition probability **A** and matrix **B** as,

$$\mathbf{A} = \begin{bmatrix} 0.0\ 0.0\ 0.0\ 0.5\ 0.5 \\ 0.0\ 0.0\ 0.0\ 0.0\ 1.0 \\ 0.0\ 0.0\ 0.0\ 1.0\ 0.0 \\ 0.0\ 1.0\ 0.0\ 0.0\ 0.0 \\ 0.0\ 0.0\ 1.0\ 0.0\ 0.0 \end{bmatrix} \tag{7}$$

$$\mathbf{B} = \begin{bmatrix} 1.0\ 0.0\ 0.0\ 0.0\ 0.0 \\ 0.0\ 1.0\ 0.0\ 0.0\ 0.0 \\ 0.0\ 0.0\ 1.0\ 0.0\ 0.0 \\ 0.0\ 0.0\ 0.0\ 1.0\ 0.0 \\ 0.0\ 0.0\ 0.0\ 0.0\ 1.0 \end{bmatrix} \tag{8}$$

Taken the gait feature vector set $\mathbf{V}_m$ of given person m as training data, the purpose of parameter estimation for the SAHMM model λ_m is to optimize three parameters **A**, **B** and $\boldsymbol{\pi}$. In gait recognition based on image sequences, the state transition matrix **A** and the confusion matrix **B** cannot be obtained directly, hence we use the forward-backward algorithm [19] to obtain the local optimal solution. The forward-backward algorithm firstly initializes the parameters of the SAHMM gait model and takes use of the given training data to evaluate the reliability of the initialization parameters, finally adjusts the initialization parameters by minimizing the errors they cause. Specifically, given gait observation sequences $\mathbf{V}_m$, the SAHMM model λ_m and its initialization parameters, we get,

$$\gamma_t(\mathrm{i}) = P\left(q_t = s_i \left| V_m, \lambda_m \right.\right) \tag{9}$$

$$\xi_t(i, j) = P\left(q_t = s_i, q_{t+1} = s_j \left| V_m, \lambda_m \right.\right) \tag{10}$$

where $\gamma_{\mathrm{t}}(\mathrm{i})$ is the local probability of the implicit state S_i at time t, $\xi_t(i, j)$ is the local probability of the implicit state converting from S_i at time t to S_j at time t+1.Finally, we use the eq. (11)~eq.(13) to iteratively refine the initialization parameters of λ_m and obtain a set of locally optimal parameters for (**A**, **B**, $\boldsymbol{\pi}$).

$$\hat{\pi}_i = \gamma_t(i) \tag{11}$$

$$\hat{a}_{ij} = \sum_{t=1}^{T-1} \xi_t(i, j) \Big/ \sum_{t=1}^{T-1} \gamma(i) \tag{12}$$

$$\hat{b}_{ij} = \sum_{t=1}^{T} \gamma_t(i) \Big/ \sum_{t=1}^{T} \gamma_t(i) \tag{13}$$

3.3 Parameter Adaptation for Gait Models

In practical applications of gait recognition, due to external factors that affect the view angle of the camera, clothing and weight bearing, there often is a mismatch between the training data and the practical test data, which will lead to decline gait recognition performance. By parameter adaptation which adopts data from a real environment to adjust the parameters of gait models, it is therefore possible to solve this problem. HMM parameter adaption methods usually are getting well with the maximum a posteriori (MAP) estimation [20], which combines prior knowledge and the knowledge obtained from adaptive data, then perform linear interpolation between initial parameter and adaptive data to obtain final mean vector. The advantage of this method is that it makes use of prior knowledge of the model parameters, thus it has good consistent and incremental features, especially when the adaptive data is large enough. However, the MAP-based methods are sensitive to the amount of adaptation data. When the volume of adaptation data is little, the mean after adaption will depend on the initial one.

In order to overcome the disadvantages of parameter adaption methods using MAP, we propose an incremental adaptive method based on Cox regression analysis. By utilizing the relationship between HMM parameters, the proposed method employs a small amount of adaption data to adjust the gait model parameters effectively. Suppose the output of SAHMM gait models is mixed with Gaussian distribution, and the output distributions before and after the adaptive processing are,

$$\lambda_m = \left(\mu_{ij}, \Sigma_{ij} \right) \tag{14}$$

$$\tilde{\lambda}_m = \left(\tilde{\mu}_{ij}, \tilde{\Sigma}_{ij} \right) \tag{15}$$

where μ_{ij} and $\tilde{\mu}_{ij}$ are the *j*-th means of Gaussian mixed distribution of the *i*-th state before and after the adaptive processing respectively; Σ_{ij} and $\tilde{\Sigma}_{ij}$ are the covariance matrices before and after the adaptive processing separately.

Given an adaption data set $\mathbf{V}_{Am} = \left\{ \mathbf{v}_i \middle| i \in \left[1,\ N_A\right] \right\}$, the corresponding Cox regression model is defined as:

$$\tilde{\mu}_{ij} = \mathbf{K}\mu_{ij} + \mathbf{k}_0 + \delta_{ij} \tag{16}$$

where δ_{ij} is the residual error, $\mathbf{K}$ and $\mathbf{k}_0$ are the regression matrix and translation vector respectively. To solve K and $\mathbf{k}_0$, we define the objective function as,

$$f(\mathbf{K}, \mathbf{k}_0) = \sum_{p \in N} w_p \delta_p^T \cdot \delta_p = \sum_{p \in N} w_p \left(\tilde{\mu}_{ij} - \mathbf{K}\mu_{ij} - \mathbf{k}_0 \right)^T \cdot \left(\tilde{\mu}_{ij} - \mathbf{K}\mu_{ij} - \mathbf{k}_0 \right) \tag{17}$$

where *N* is the nearest neighbor set of λ_m before the adaptive processing; w_p is the weight factor that is defined as,

$$w_p = \exp(-d_p^2) \tag{18}$$

$$d_p^2 = \sum_{k=1}^{K} \left(\mu_p^k - \mu_{ij}^k \right)^2 \Big/ \beta_{ij}^k \tag{19}$$

where β_{ij}^k is the *k*-th diagonal element of covariance matrix Σ_{ij}.

From the definition, when the distance is smaller, the corresponding weight factor w_p will be larger, and thus it has a great impact on the regression model. Meanwhile, if there is no corresponding adaption data for a mean vector, we only need to find the regression class and utilize the corresponding transformation matrix to conduct the adaption transformation. For a small capacity of gait recognition based on SAHMM, one leaf node in its regression tree represents a single component, and each node on the high layer stands for a set of components with similar distance, while the root node contains all the mixture components. When a SAHMM has multiple mixed components, the leaf nodes come along with the basic classes based on the initial clustering, and each of the basic class comprehends to a set of components with similar distance. The experimental results show that the proposed method is effective for gait recognition with only a small amount of adaption data.

3.4 SAHMM-Based Gait Classification

Given a gait observation vector set $\mathbf{V} = \left\{ v_k \middle| k \in \left[1, N\right] \right\}$, the observation sequence $\mathbf{O}_k = o_{k_1} o_{k_2} \ldots o_{k_T}$ corresponding to each $\mathbf{v}_k$, and theset of all the samples in gait database, a gait recognition problem is converted to a SAHMM evaluation problem. Thus we only need to calculate the average probability,

$$\bar{P}_m = \tfrac{1}{N}\sum_{k=1}^{N} P\left(O_k \mid \lambda_m\right) \tag{20}$$

where each SAHMM gait model outputs the results { $\bar{P}_m$ } given the observation sequence. We take use of the forward algorithm to calculate the probability that each SAHMM model generates the given gait observation sequence **V**. The complete gait recognition process based on SAHMM is described as Algorithm 2.

Algorithm 2. Gait recognition using SAHMM with parameter adaptation

Input: Feature vectors from the training set and test set.

Output: The model number *n*.

Step1. Construct a SAHMM gait model $\lambda_m\left(\mathbf{A},\mathbf{B},\boldsymbol{\pi}\right)$ for each person in current gait database, and initialize parameters of all gait models. The initial values of **A**, **B** and $\boldsymbol{\pi}$ are shown in eq. (6) ~ eq. (8).

Step2. Use the forward-backward algorithm to obtain local optimal solution for parameters of all gait models. In this step, the feature vectors from the training set are adopted as eq. (11) ~ eq. (13).

Step3. Refine the parameters of each gait model by the proposed incremental adaptive method. This step encompasses parameter optimization operations of SAHMM gait models, which are relatively complex and are able to be completed offline only.

Step4. Calculate the average probability of all gait models. The local probability $\gamma_t\left(j\right)$ of each hidden state is obtained by recursive computations using eq. (21).The probability of generating an observation vector set **V** by each HMM gait model equals to the sum of all the local probabilities at time *T*, as shown in eq. (21) and eq.(22).

$$\gamma_t\left(j\right) = \begin{cases} \pi\left(j\right) b_{jk_t} & t=1 \\ \sum_{i}^{5}\left(\gamma_{t-1}\left(i\right) a_{ij}\right) b_{jk_{t-1}} & t \in [2,T] \end{cases} \tag{21}$$

$$\bar{P}_m = \tfrac{1}{N}\sum_{k=1}^{N} P\left(O_k \mid \lambda_m\right) = \tfrac{1}{N}\sum_{k=1}^{N}\sum_{j=1}^{5}\gamma_T\left(j\right). \tag{22}$$

The operations from this step have moderate computational complexity, and are able to be processed in real-time.

Step5. Sort the average probabilities $\left\{\bar{P}_m\right\}$ by each gait model and get the model λ_n with the maximum average probability.

Step6. Output the model number *n*.

4 Experimental Results

Based on the CASIA Dataset B [21], and OU-ISIR Large Population Dataset [22], we designed and implemented three experiments for evaluating gait recognition. Experiment I and Experiment II is to evaluate the cross-view and cross-walking condition gait recognition on CASIA Dataset B which contains gait videos of 124 individuals including 11 views evenly distributed from 0 to 180 degrees under three kinds of walking conditions i.e. normal conditions, wearing coats, and carrying bags; Experiment III is to verify its generalization ability with large data, and takes use of the OU-ISIR Large Population Dataset which contains gait data of 20 individuals, including 12 image sequences, 4 sequences for each of the three directions, i.e. parallel, 45 degrees and 90 degrees to the image plane.

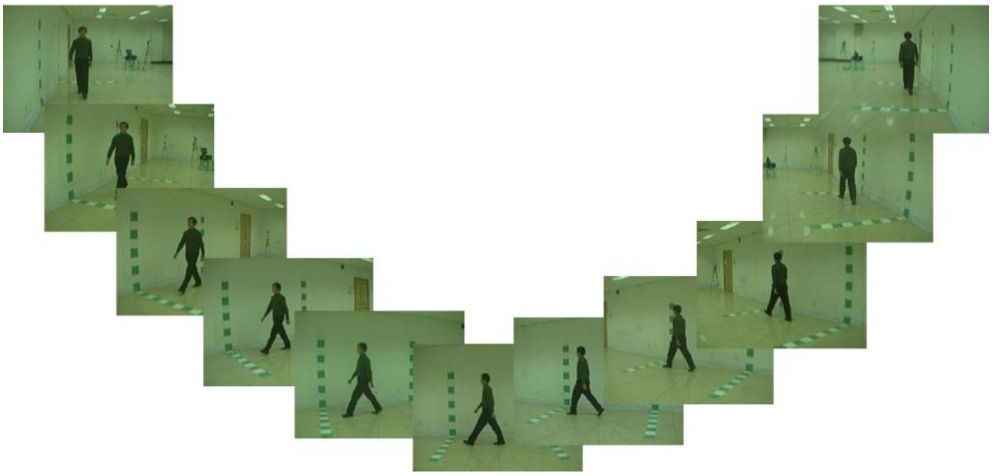
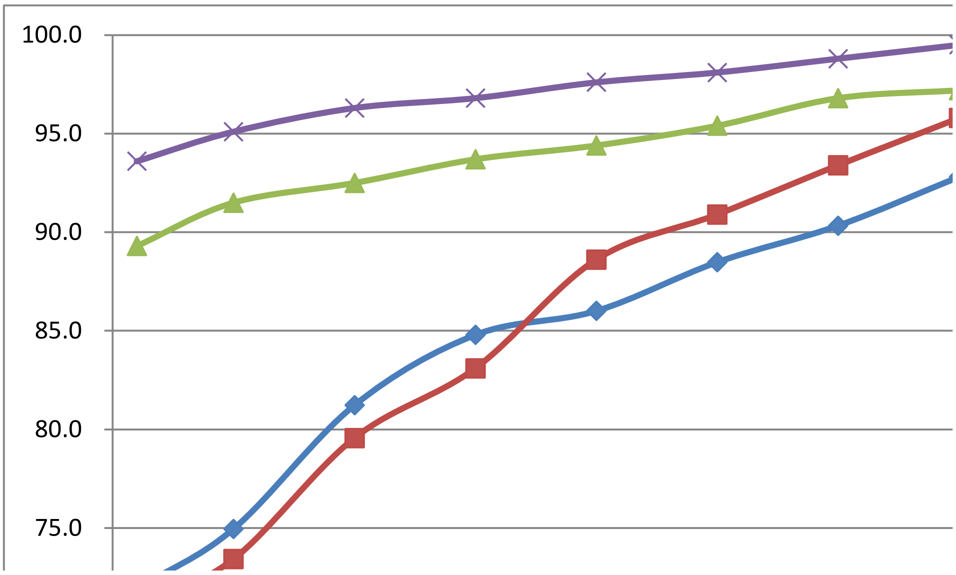


Fig. 6 11 views examples in CASIA Dataset B

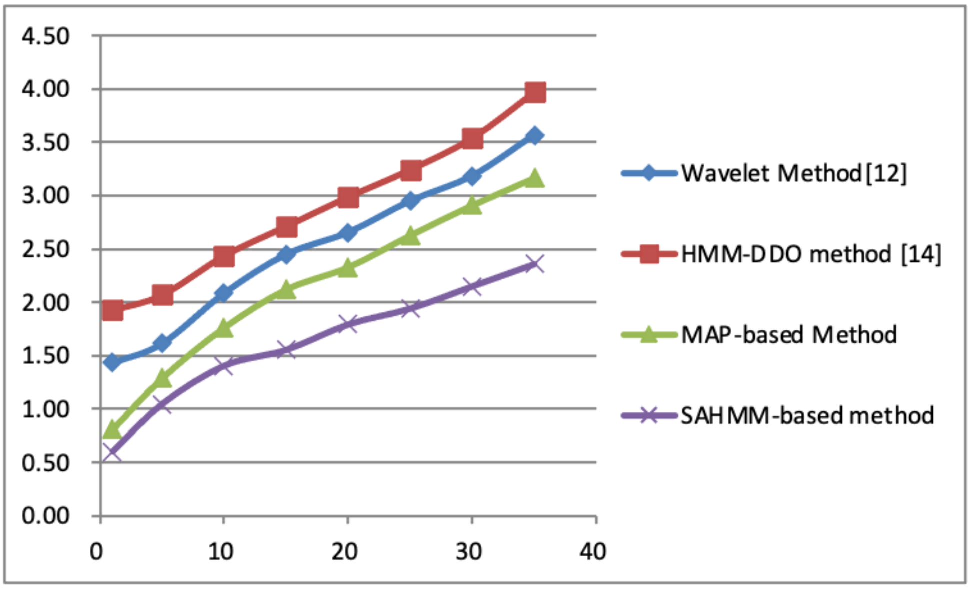
Average recognition rate and recognition rate standard deviation are taken into consideration in the evaluation indexes, which reflect the correct recognition rate and stability of related gait recognition methods. These experiments utilize the subsets of CASIA gait database to compare the proposed method with the other three methods. The experimental data were obtained by random sampling and repeated experiments (500 times).

Experiment I. Comparison of recognition rates of proposed methods on CASIA Dataset B

This experiment only takes use of the gait data under normal conditions in CASIA Dataset B. There are 8 training sets, which consist of all the data with 90 degrees of view and the randomly selected data from the other views with the proportion of 1%, 5%, 10%, 15%, 20%, 25%, and 35% separately. The experimental results are shown in Figure 7, where the horizontal axis is the percentage of the total adaption data, the vertical axis in Figure 7 (a) is the average recognition rate, and the vertical axis in Figure 7 (b)is the standard deviation of recognition rate.



(a) Average recognition rate



(b) The standard deviation of recognition rate

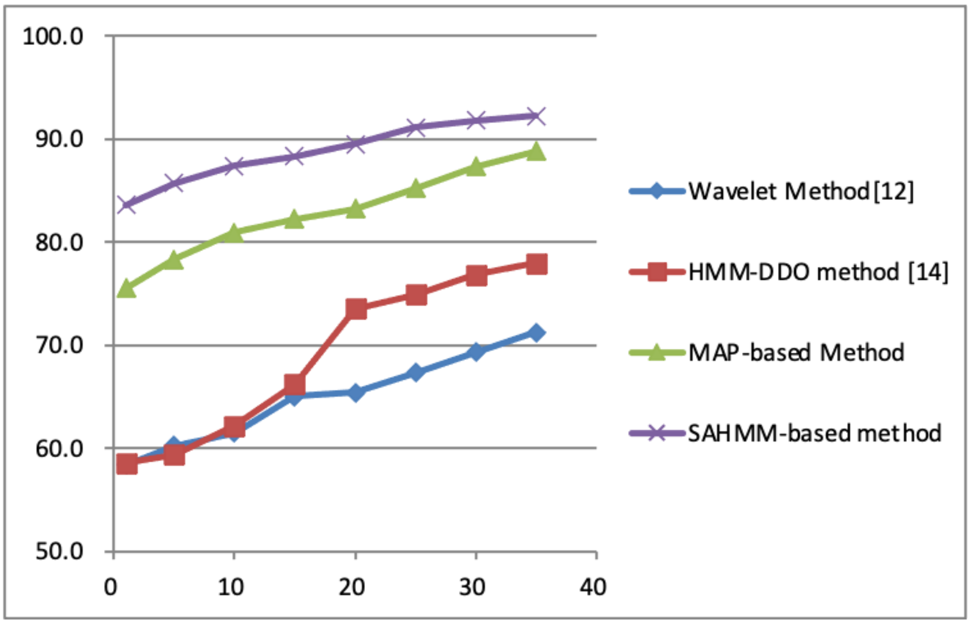
Fig.7 The recognition rate curves of different methods in Experiment I

Figure 7 shows the comparisons of the wavelet-based method [12], the HMM-DDO method [14], the MAP-based adaptive method and the SAHMM-based method. The proposed method in this paper have higher average recognition rate and less standard deviation. This mainly owes to the use of adaptive parameter optimization, which easily leads to the mismatch between the training data and the test data. Furthermore, Figure 7 also shows that with the increase of the proportion of adaption data in the total test data, the average recognition rates of the four methods have all been improved. However, the methods have a wide range of sensitivity to the adaption data, and the standard deviations of recognition rate of the methods are also different. The wavelet-based and HMM-DDO methods have not adaptive process, when the size of adaption data is small, training data with 90 degree plays a decisive role in the process of model parameter estimation and the recognition rate is very low; but when the adaption data is increased to more than 20%, which is very close to the normal amount of training data, the recognition rate will be increased sharply. Regarding the methods after an adaptive parameter optimization process, the recognition rate is already higher in the case of little adaption data, and with the increase of adopted training data, the recognition rate is relatively smooth.

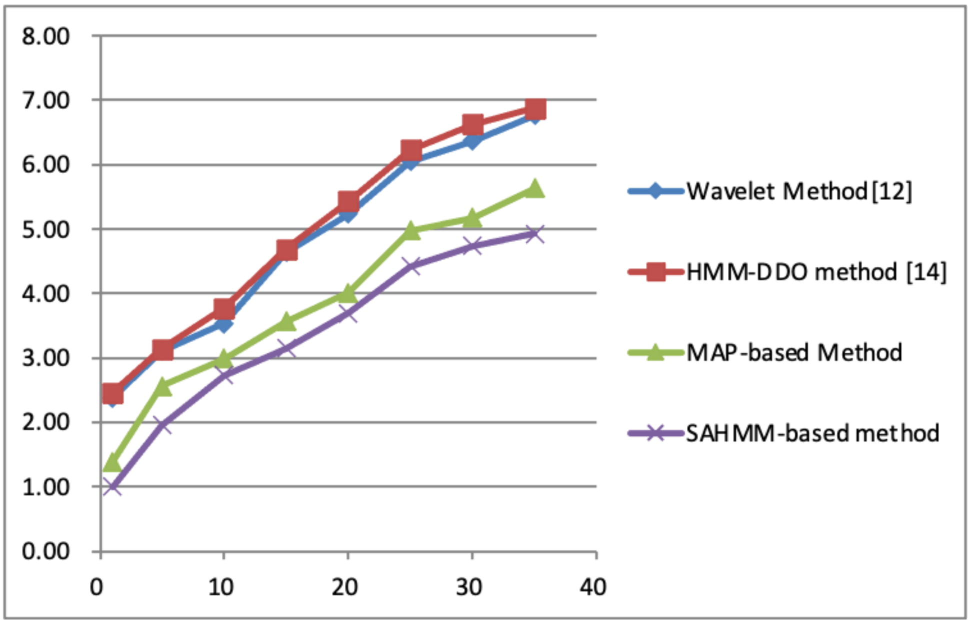
Experiment II. Comparisons of recognition rate of several methods on CASIA Dataset B

This experiment adopted all the gait data in CASIA Dataset B including the gait data under walking conditions of wearing coats and carrying bags. Similar to Experiment I, we construct 8 types of training set from all gait data with 90-degree view and randomly selected gait data from the other views with the proportion of 1%, 5%, 10%, 15%, 20%, 25%, and 35%. Figure 8 shows the experimental results with the horizontal axis representing the percentage of total adaption data, the vertical axis in Figure 8 (a) representing the average recognition rate as well as the vertical axis in Figure 8 (b) standing for the standard deviation of recognition rate.

In Figure 8 (a), compared to Experiment I, the average recognition rates of the four methods all have a significant decline. This is due to the increase of experimental data which includes complex interference factors such as wearing coats and carrying bags. But on the whole, our algorithm has a higher average recognition rate, especially when the adaption data is less. From Figure 8 (b) we see that with the increase of adaption data, the standard deviations of recognition rate of the four methods have been significantly increased, but the standard deviations with an adaptive parameter optimization process are relatively low. The reason is that the increase of abnormal gait data, such as wearing coats and carrying bags, results in the decrease of recognition rate. On the other hand, the methods with a process of adaptive parameter optimization reduce the imbalance between the training data and the test data, the corresponding standard deviations of recognition rate are relatively small.



(a) Average recognition rate



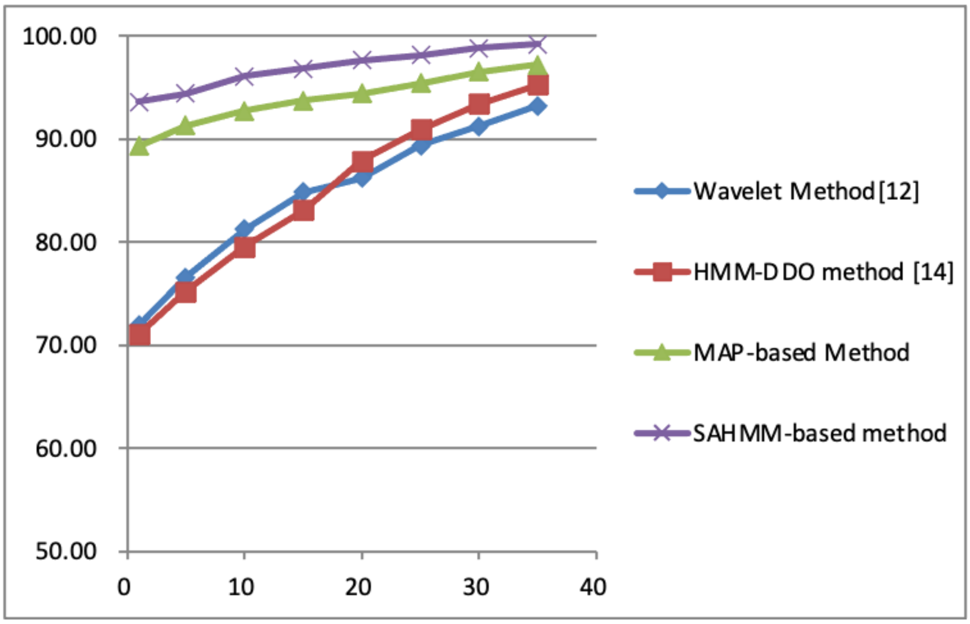
(b) The standard deviation of recognition rate

Fig.8 The recognition rate curves of different methods in Experiment II.

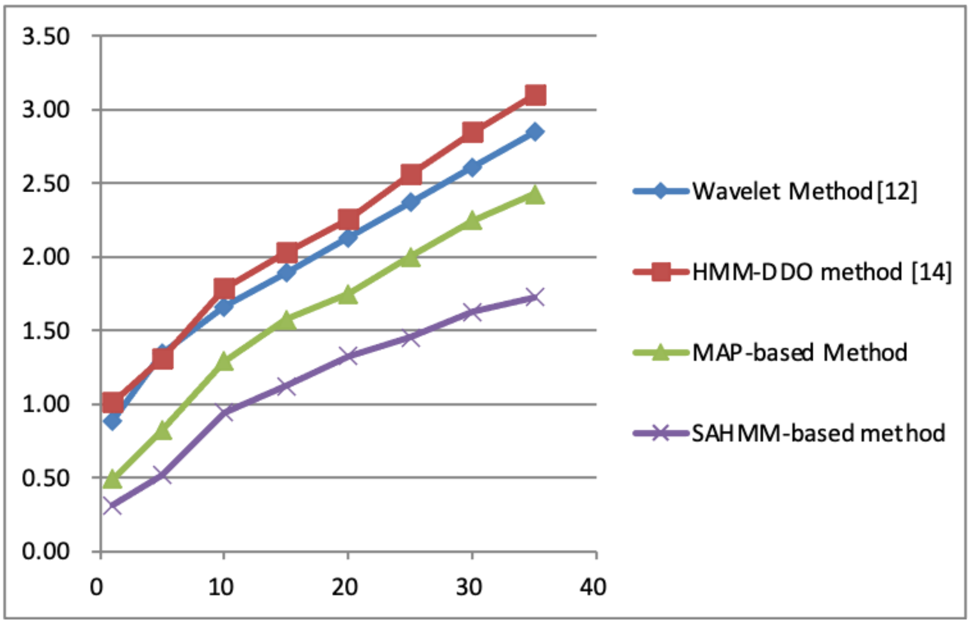
Experiment III. Recognition rates comparison of several methods on OU-ISIR Large Population Dataset

In this experiment, the effectiveness of the proposed method is assessed against the OU-ISIR Large Population Dataset, named OULP-C1V2. This dataset comprises two main subsets, A and B. A is a set of two sequences per subject, and B is a set of one sequences per subject. In addition, each of the main subsets is further divided into 5 subsets based on the observation angles, 55-degree, 65-degree, 75-degree, 85-degree, and including all four angles. The data set consists of over 4,000persons walking on the ground surrounded by the 2 cameras at 30 fps, 640 by 480 pixels.

We utilize 8 different training sets including all data in subset A and select data with other views by 1%, 5%, 10%, 15%, 20%, 25%, and 35% randomly from subset B. The related experimental results are shown in Figure 9, where the horizontal axis represents the percentage of the total adaption data, the vertical axis in Figure 9 (a) represents the average recognition rate, and the vertical axis in Figure 9(b) shows the standard deviation of recognition rate as well.



(a) Average recognition rate



(b) The standard deviation of recognition rate

Fig.9 The recognition rate curves of different methods in Experiment III

Figure 9 (a) shows the MAP-based adaptive method and the proposed method in this paper have obvious advantages in terms of average recognition rate, compared with the wavelet-based and HMM-DDO methods. The reason is that the process of adaptive parameter optimization partially solves the mismatch problem between the training data and the actual test data. Shown as Figure 9 (b), multiple adaptive methods have different sensitivity to the proportion of adaption data in the test data set, so the stability of the recognition rate is various in large repeated experiments. On the whole, the recognition rate of methods with an adaptive parameter optimization process is much stable, thus the corresponding standard deviation is rather small.

The experimental results demonstrate the effectiveness of the proposed algorithm. This is mainly due to the well-designed feature extraction algorithm, and partially benefits from the parameter adaptation optimization process in the proposed SAHMM-based gait recognition algorithm.

5 Conclusion and Future Work

In this paper, a well-designed feature representation and the related feature extraction algorithm are presented. Then a novel gait recognition method based on SAHMM is proposed, we construct the observed state set from each gait cycle and train the SAHMM gait models. In addition, in order to improve the gait recognition rate, we design and implement an adaptive algorithm to automatically adjust the parameters of SAHMM.

To the best of our knowledge, this is the first time we work for gait recognition using SAHMM algorithms, our contributions are that we, (1)present a well-designed feature extraction algorithm based on a new gait feature, named LGEI, which utilizes important details of gait changes in each gait cycle, and release sensitivity to variations in view angle; (2) propose a novel gait recognition method based on SAHMM for gait classification which takes use of a parameter adaptation process to optimize the parameters of each gait model; (3) greatly advance the record scores on the CASIA Dataset B and OU-ISIR Large Population Dataset, demonstrating that our method can work well under the condition of gait databases with large capacity. The limitation of the proposed algorithms is that, in each gait image sequence, changes of human moving direction according to the camera should not be too intensive. The problems such as how to reduce or weaken the artificially restricted conditions, how to improve our gait recognition algorithms to match the practical needs are our future work [23, 24].

Acknowledgment

This work is supported by the NSF of China (No.61303146, 61602431).

References

[1] Sudeep Sarkar, P. Jonathon Phillips, Zongyi Liu, et al. “The HumanID Gait Challenge Problem: Data Sets, Performance, and Analysis”. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(2): 162-77

[2] Zifeng Wu, Yongzhen Huang, Liang Wang, Xiaogang Wang, Tieniu Tan. “A Comprehensive Study on Cross-View Gait Based Human Identification with Deep CNNs,” IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016

[3] Tee Connie, Michael Kah Ong Goh, Andrew Beng Jin Teoh. “A Grassmannian Approach to Address View Change Problem in Gait Recognition,” IEEE Transactions on Cybernetics, 2016

[4] Jian Luo, Jin Tang,Tardi Tjahjadi et al."Robust arbitrary view gait recognition based on parametric 3D human body reconstruction and virtual posture synthesis, " Pattern Recognition,60(2016): 361-377

[5] Himanshu Aggarwal,Dinesh K. Vishwakarma."Covariate conscious approach for Gait recognition based upon Zernike moment invariants," IEEE Transactions on Cognitive and Developmental Systems, 2017

[6] Jin Tang, Jian Luo, Tardi Tjahjadi."Robust Arbitrary-View Gait Recognition Based on 3D Partial Similarity Matching," IEEE TRANSACTIONS ON IMAGE PROCESSING,2017,26(1):7-23

[7] Chung Pau-Choo, Liu Chin-De. A daily behavior enabled hidden Markov model for human behavior understanding. PATTERN RECOGNITION, 2008,41(5):1572-1580

[8] Deng Feng, Bao Changchun; Kleijn, W. Bastiaan. Sparse Hidden Markov Models for Speech Enhancement in Non-Stationary Noise Environments. IEEE- ACM TRANSACTIONS ON AUDIO SPEECH AND LANGUAGE PROCESSING, 2015, 23(11):1973-1987

[9] Xiaoming Liu, Tsuhan Chen. Video-based Face Recognition Using Adaptive Hidden Markov Models. In proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Wisconsin, USA, 2013: 207-212

[10] Ju Han,Bir Bhanu."Individual Recognition Using Gait Energy Image," IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, 2006, 28(2):316-323

[11] Yu Guan, Chang-Tsun Li, Fabio Roli. “On Reducing the Effect of Covariate Factors in Gait Recognition: A Classifier Ensemble Method,” IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(7): 1521-1529

[12] Md. Shariful Islam, Md. Rabiul Islam, Md. Altab Hossain, Alexander Ferworn, Md. Khademul Islam Molla. “Subband entropy-based features for clothing invariant human gait recognition,” Advanced Robotics, 2017

[13] Xiaxi Huang, Nikolaos V. Boulgouris. “Gait Recognition with Shifted Energy Image and Structural Feature Extraction,” IEEE Transactions on Image Processing, 2012, 21(4): 2256-2268

[14] Nikolaos V. Boulgouris, Xiaxi Huang. “Gait Recognition Using HMMs and Dual Discriminative Observations for Sub-Dynamics Analysis,” IEEE Transactions on Image Processing, 2013, 22(9): 3636-3647

[15] Fatimah Abdulsattar, John Carter. "Performance Analysis of Gait Recognition with Large Perspective Distortion," IEEE International Conference on Identity, Security and Behavior Analysis, 2016

[16] JunseokKwon,Kyoung Mu Lee.Adaptive Visual Tracking with Minimum Uncertainty Gap Estimation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(1):18-31

[17] Dong Wang, Huchuan Lu, Ming-Hsuan Yang. Robust Visual Tracking via Least Soft-Threshold Squares. IEEE Transactions on Circuits and Systems for Video Technology,2016,26(9):1709-1721

[18] PorntepTheekhanont, WerasakKurutach, Serge Miguet. Gait Recognition Using GEI and Pattern Trace Transform. In proceedings of Information Technology in Medicine and Education 2012, Hokodate, Hokkaido, 2012:936-940

[19] A.P.Dempster, N.M.Laird, D.B.Rubin. Maximum Likelihood from Incomplete Data via the EM Algorithm. Journal of the Royal Statistical Society,1977,39(1):1-38

[20] Song Yin,Liang Weiqian. Experimental Study of Discriminative Adaptive Training and MLLR for Automatic Pronunciation Evaluation.Tsinghua Science and Technology,2011,16(2):189-193

[21] Shiqi Yu, Daoliang Tan, Tieniu Tan. A Framework for Evaluating the Effect of View Angle, Clothing and Carrying Condition on Gait Recognition. In proceedings of International Conference on Pattern Recognition． Hong Kong，China, 2006:441-444

[22] Haruyuki Iwama, Mayu Okumura, Yasushi Makihara, and Yasushi Yagi. The OU-ISIR Gait Database Comprising the Large Population Dataset and Performance Evaluation of Gait Recognition. IEEE Transaction on Information Forensics and Security, 2012, 7(5): 1511-1521.

[23] Keji Zheng, Wei Qi Yan, Parma Nand. Video dynamics detection using deep neural networks. IEEE Transactions on Emerging Topics in Computational Intelligence, 2(3), 224-234.

[24] Wei QI Yan. Introduction to Intelligent Surveillance Surveillance: Data Capture, Transmission, and Analytics Authors, 2017, Springer.