Gait Analysis for Walking Paths Determination and Human Identification

Meng-Fen Ho , Chung-Lin Huang

Abstract

In this paper, we propose a gait analysis method to extract the dynamic and static information from the input video for walking path determination and human identification. Based on the periodicity of swing distances, we may estimate the gait period of each walking video sequence. For each gait cycle, we depict the dynamic information by analyzing the distribution of motion vectors, and then describe the static information by using Fourier descriptors. The extracted dynamic and static information is transformed into lower dimensional embedding space for human identity recognition. To solve the difference of walking velocity between the test and training human objects, a hybrid human ID recognition algorithm is developed to choose the effective feature. Given a test feature vector, the nearest neighbor classifier is applied for walking paths determination and human identification. The proposed algorithm is evaluated on the CASIA gait database, and the experimental results demonstrate a highly acceptable recognition rate, for example, 98% for normal walking dataset.

Keywords: Gait analysis, human identification.

Meng-Fen Ho, Associate Professor, Department of Electronic Engineering, HUST.
Chung-Lin Huang, Professor, Dept. of Electrical Engineering, National Tsing Hua University.
Received 16 December 2010; accepted 18 January 2011

步伐姿態分析運用於人物步行路徑與身分識別

何孟芬、黃仲陵

摘 要

在這篇論文中，我們擷取動態與靜態資訊來進行人物步行路徑和身分的識別。首先利用步伐擺幅距離的週期特性，可以估測每段影像的步伐週期，針對每段步伐週期，分析其運動向量的分布，即可獲得所需的動態資訊，同時利用傅立葉描述取得靜態資訊，接著將這兩種資訊轉換至低維度空間以便進行人物識別。為了解決訓練影片與測試影片中，人物走路速度可能不同的問題，我們提出了一個混合人物識別演算法來選定最有效的特徵。每當測試特徵向量進來時，利用最鄰近分類法則進行路徑確認與人物識別。本系統利用 CASIA 步伐姿態資料庫進行評估，實驗結果證明的確獲得極高的辨識率，以正常走路的資料組而言，可達到約 98%。

關鍵詞：步伐姿態分析、人物辨識。

何孟芬：修平科技大學電子工程系副教授
黃仲陵：清華大學電機工程系教授
投稿日期：99 年 12 月 16 日　接受刊登日期：100 年 1 月 18 日

I. Introduction

Intelligent video surveillance system has been widely developed of which the human identification is one of the most important functionalities. Biometric features are regularly applied for human identification to express the unique property of human object. Common biometric features include iris, face, speech, fingerprints, hand geometry, voice, and gait. Here, we choose the gait posture as our main feature for human identification. Comparing with other biometrics, gait analysis has the advantage of non-contact and can generate the perceivable biometric feature for human identification at distance.

However, gait analysis also has some disadvantages [1]. In the internal factors, gait posture has a little change accompanying the mood or physical injury of the walking people. In the external factors, gait can be affected by clothing, shoes, walking surface, or other handbag-carrying conditions. Therefore, it will induce a large gait variation of the same person and reduce system discriminating ability. Excluding the internal factors, we attempt to construct a new system that can identify the human object based on the gait postures caused only by some external factors.

A. Related Works

Current approaches of gait analysis can be divided into two categories: appearance-based and model-based approaches [1-16]. The former deals directly with the image statistics, whereas the latter models the image data and then analyzes the variation of its parameters. The majority of current approaches are the appearance-based, because they are simple and fast.

Su et al. [2] propose a method which combines both the appearance-based approach and the model-based approach to analyze and extract human gait. The static features include body height, width, etc. Instead of modeling the human body, the limb angle information is extracted by analyzing the variation of silhouette width to represent the kinematic information of gait. Then the multi-class support vector machines are used to identify human object. In [3], a new feature extraction process is proposed for gait representation. Each gait sequence is described by using a low-dimensional feature vector consisting of selected Radon template coefficients. Identification is done by using linear discriminant analysis (LDA). A different approach called component-wise comparison is proposed by [4], which calculates the distances between silhouettes

on a component by component basis, and then combines the component-wise distances into a common distance metric for the evaluation of similarity between two silhouettes.

Han et al. [5] propose a spatial-temporal gait representation, called Gait Energy Image (GEI), to characterize human walking posture. They combine Principal Component Analysis (PCA) and Multiple Discriminant Analysis (MDA) to transform datasets to a low-dimensional space and then separate the datasets to different classes. In [6], they improve temporal templates called Gait History Image (GHI). GEI only represents the static part and the dynamic part on moving subject, but GHI increases the temporal variation and provides the better performance. Equivalently, the other authors choose gait moment image (GMI) as features to emphasize the dynamic information of human gait in [7]. GMI is the gait probability image at each key moment in the gait period.

In [8], they propose an eigen-gait method to model human motion directly, and encode the dynamic feature of gait in pair-wise image similarities of gait images. Moreover, the authors employ PCA to reduce the self-similarity plot (SSP) and use K-nearest neighbor rule to recognize the human identification. Cheng et al. [9] propose a novel algorithm for both automatic viewpoint and person identification by using only the silhouette sequence of the gait. First, the gait silhouettes are nonlinearly transformed into low-dimensional embedding by Gaussian process latent variable model (GPLVM). Then the temporal dynamics of the gait sequences are modeled by hidden Markov models (HMMs). In [10], a gait is represented by a vector of affine invariant moments obtained from the binary silhouettes. With a combination of histograms of individual silhouette and contextual silhouette, a gait appearance model is represented by a shape descriptor and gait images plane [11]. The similarity of gait appearance models are measured by Jeffrey divergence and dynamic time warping.

B. System Overview

Generally speaking, the procedure of gait recognition includes human object segmentation, feature extraction, and classification. In this paper, we propose a method which is different from the previous works in using new gait characteristics for human identification without knowing the walking direction of the person. The novel system will identify the walking path of human object, and then recognize his/her identification. The

overview of our proposed system is shown in Fig. 1.

This system is composed of four main stages: human silhouettes segmentation, static and dynamic feature extraction, dimension reduction of feature vectors, and the recognition of walking path and human identification. The first stage is developed to separate the human object from the background in binary image sequence, which includes foreground compensation, horizontal and vertical alignment, and size normalization. Then, gait period is estimated by utilizing the periodicity of swing distances, and then binary image sequence is divided into sub-cycles of silhouette.

The second stage is developed to extract the static and dynamic features in each sub-cycle of silhouette. The intersecting operation is applied to every silhouette in each sub-cycle to obtain the static region within each gait period. The contour of the static region is regarded as the static feature. Then, we analyze two continuous image frames in each sub-cycle to obtain the optical flow measurement as the motion vector field. By analyzing the statistics of the motion vector field, we generate the two-dimensional motion histogram of the magnitude and the direction of the motion vectors, which are regarded as the dynamic features.

In the third stage, we use PCA and MDA to transform the high-dimensional feature vectors to the low-dimensional subspace. In the fourth stage, the discriminant functions and the nearest neighbor classifier are applied to recognize the walking path and human identification respectively.

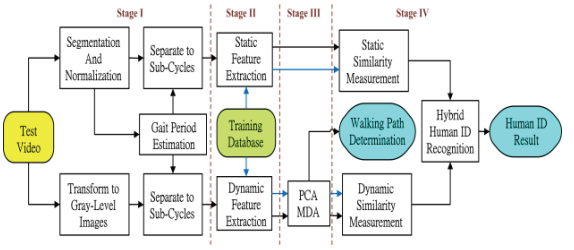


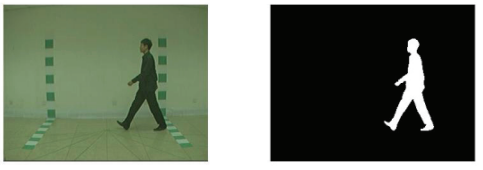
Fig. 1. System diagram of our main system.

II. Human Object Segmentation And Gait Period Estimation

A. Human Object Segmentation

In our system, we assume the camera is stationary and there is only one human object walking through the scene. We use the foreground video sequences from CASIA database. For each segmented human silhouette in each frame, we compute the centroid (xc, yc) of the human silhouette. Then, we compute the width W and the height H of the human silhouette. We use the centroid of the human silhouette to calibrate along the horizontal direction, and the vertical direction of the

frames of size 320x240. Then we use the centroid to calibrate and then normalize the human silhouette within the bounding box of size 161x101 as shown in Fig. 2.



(a) (b)

(c)

Fig. 2. Human segmentation results of the view angle 90 degree. (a) The original image. (b) The extracted silhouette. (c) The human silhouette after calibration and normalization.

B. Gait Period Estimation

The human walking is treated as a periodic activity which corresponds to the total average pixel width of the normalized foreground. For example, double-support stance corresponds to the local maximum average width and legs-together stance corresponds to the local minimum width. Therefore, swing distance [6] that computes the total average foreground pixels' distance from the normalized foreground center is applied to detect the started and end frames of the gait period. Swing distance is described as

$$sw = \sum_{y=y_b}^{y_t - \frac{(y_t - y_c)}{2}} \sum_{x=x_l}^{x_r} \left| (x - x_c) \times \frac{I(x,y)}{255} \right| , \quad (1)$$

where (xc, yc) is the centroid of the binary silhouette, xl and xr denote the horizontal positions of the left-most and the right-most boundary pixels of the silhouette respectively, yb and yt denote the vertical positions of the bottom-most and the top-most boundary pixels of the silhouette respectively, and I(x, y) represents the pixel intensity at (x, y).

The periodicity of gait is implied by the variation of swing distance. The gait period defined here starts from the legs-together stance and passes through two double-support stances and one legs-together stance and back to the same legs-together stance. Local minimum swing distance is detected and used to separate gait periods from the video since it is less affected by background noise than local maximum. The experimental results for measuring the periodicity of binary gate sequence are shown in Fig. 3. Finally, the binary gait sequence is separated to a number of sub-cycles.

III. Gait Feature Extraction

In order to get the higher recognition rate, we must find the discriminative

characteristics for representing the walking style of each person. Using the motion vectors as the dynamic features is not effective when the walking velocities of the training video sequence and the test video sequence are different. It is necessary to utilize the static features which are independent of the walking velocity. Here, we use the Fourier descriptors to represent the contour of static region as the static features.

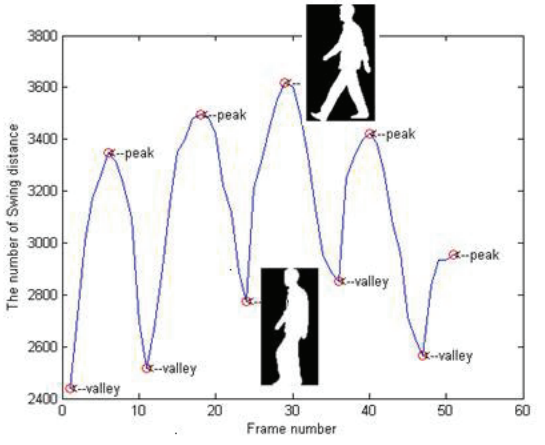


Fig. 3. The Swing distance corresponds to frame numbers. Local minimum of swing distance represents the legs-together stance, while, local maximum represents the double-support stance.

A. Dynamic Feature Extraction

After using the motion estimation to find the optical flow field of the video sequences, we evaluate the histogram distribution of the optical flow vector field, which are treated as the dynamic feature. From the experimental results, we find some useless motion vectors due to the variations of lighting condition, reflecting condition, shadow, and other reasons. By using the thresholding, we select only the meaningful and slightly larger motion vectors. We partition the magnitude and the direction of the motion vectors into 10 and 13 intervals respectively as shown in Table 1 and 2.

Table 1. Intervals of the magnitude of the optical flow vector, where mvm denotes the magnitude of every optical flow vector.

Level	**1**	**2**	**3**	**4**	**5**
MV_m	1~2	2~3	3~4	4~5	5~6
Level	**6**	**7**	**8**	**9**	**10**
MV_m	6~7	7~8	8~9	9~10	>10

Table 2. Intervals of the direction of the optical flow vector, where MVθ denotes the direction of every optical flow vector.

Interval	**1**	**2**	**3**	**4**	**5**
MV_θ	$\pi \pm \frac{\pi}{12}$	$\frac{5\pi}{6} \pm \frac{\pi}{12}$	$\frac{4\pi}{6} \pm \frac{\pi}{12}$	$\frac{3\pi}{6} \pm \frac{\pi}{12}$	$\frac{2\pi}{6} \pm \frac{\pi}{12}$
Interval	**6**	**7**	**8**	**9**	**10**
MV_θ	$\frac{\pi}{6} \pm \frac{\pi}{12}$	$0 \pm \frac{\pi}{12}$	$-\frac{\pi}{6} \pm \frac{\pi}{12}$	$-\frac{2\pi}{6} \pm \frac{\pi}{12}$	$-\frac{3\pi}{6} \pm \frac{\pi}{12}$
Interval	**11**	**12**	**13**		
MVθ	$-\frac{4\pi}{6} \pm \frac{\pi}{12}$	$-\frac{5\pi}{6} \pm \frac{\pi}{12}$	$-\pi \pm \frac{\pi}{12}$		

Then we combine the magnitude and

direction histograms to a 2-D histogram. Supposing there are N gait cycles in the sequence, after evaluating the overall N 2-D combinational histograms, we utilize the Bhattacharyya distance to measure the similarity between these 2-D combinational histograms. For two discrete probability distributions p(x) and q(x), over the same domain X, the Bhattacharyya distance dbhatt(p,q) is defined as

$$d_{bhatt}(p,q) = \sum_{x \in X} \sqrt{p(x)q(x)} \quad . \tag{2}$$

Let MVC(i) denote the 2-D combinational histogram belonging to the ith gait period interval. The 2-D combinational histograms, MVC(i), i=1,…,N whose dbhatt(p, q) is lower than the threshold are truncated as MVC(i), i= 1,…,Nt, where 1< Nt <N. Then, we take the average of the spare 2-D combinational histograms as

$$r = \sum_{i=1}^{N^t} \frac{MV_C(i)}{N^t} \quad . \tag{3}$$

Finally, we transform the average 2-D combinational histogram to a column vector of size 130×1. The column vector expresses the dynamic features in the video sequence. Figure 4(a) shows the silhouette sequence of one human object, and the corresponding average 2-D combinational histogram is demonstrated in Fig. 4(b).

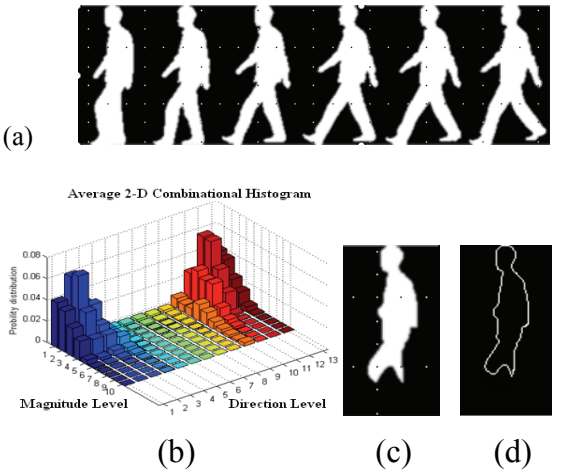


Fig. 4. (a) The silhouette sequence. (b) The average 2-D combinational histogram. (c) The static part of (a). (d) The extracted contour.

B. Static Feature Extraction

In this paper, we use Fourier descriptors (FDs) [17] to describe the contour of the static region within half cycle as the static feature. The static region of the human body represents the common region in each frame within one half cycle. Assuming S(x, y) represents the common region in each frame of one half cycle, where non-moving pixels are highlighted in binary image S, as shown in Fig. 4(c). For binary silhouette sequence BI, S can be obtained by utilizing the intersecting operation as

$$S(x,y) = \bigcap_{t=1}^{\tau} BI(x,y,t) \quad , \tag{4}$$

where BI(x,y,t) is the binary value of the pixel (x,y) in the tth frame, and τ is the time duration of the motion sequence.

FDs are invariant to translation, rotation and scaling of the object. Let the complex array p0, p1,…, pM-1 represent the boundary belonging to the region of static part, the kth Fourier transform coefficient can be calculated as

$$z_k = \sum_{n=0}^{M-1} p_n \, e^{-\frac{2\pi i k n}{M}}, \quad (5)$$

zk describes the frequency contents of the shape. Lower frequency components depict the overall shape, whereas higher frequency components describe the details of the shape. The FDs are defined as

$$c_k = |z_{k+2}| / |z_1|, \; k = 0, \cdots, M-3. \quad (6)$$

Finally, we truncate the number of evaluated FDs to 30. Assuming N is the number of total gait cycles in the video sequence, the FDs of N static regions are combined as a matrix of size N×30 to represent the static features.

IV Human ID Recognition in Multiple Paths

Given the test video sequence, the purpose of the system is to recognize the human walking path and identify the human ID as well. In the learning procedure, we adopt the dimension reduction method to map the training feature vectors onto its embedding feature space. Then we continue the walking path classification and the human ID similarity measurement in the embedding feature space.

A. Learning Procedure by PCA and MDA

We combine PCA and MDA [5] to achieve the best trade-off between the data representation and the class separability. PCA is optimal in the sense that it minimizes the mean square error between the n d-dimensional dynamic features {x1, x2,…,xn} and their approximations {y1, y2,…, yn},

$$\mathbf{y}_k = [\mathbf{u}_1, \ldots, \mathbf{u}_{d'}]^T \mathbf{x}_k, \; k = 1, \ldots, n, \quad (7)$$

where yk denotes the d′-dimensional feature vector, d′ << d, and u1, u2,…,ud′ represent the eigenvectors corresponding to the d′ largest eigenvalues.

Assuming that the principal component vectors {y1, y2,…, yn} belong to L classes (C1 to CL), MDA selects a transformation matrix W so that the ratio of the between-class scatter and the within-class scatter is maximized. Let the between-class scatter matrix be defined as

$$\mathbf{S}_b = \sum_{i=1}^{L} N_i (\overline{\mathbf{y}_{C_i}} - \overline{\mathbf{y}})(\overline{\mathbf{y}_{C_i}} - \overline{\mathbf{y}})^T , \quad (8)$$

and the within-class scatter matrix be defined as

$$\mathbf{S}_w = \sum_{i=1}^{L} \sum_{\mathbf{y}_{C_i}} (\mathbf{y}_{C_i} - \overline{\mathbf{y}_{C_i}})(\mathbf{y}_{C_i} - \overline{\mathbf{y}_{C_i}})^T \quad (9)$$

where Ni is the number of feature vectors in class Ci, yCi and $\overline{\mathbf{y}_{C_i}}$ denote the feature vectors and the mean vector belonging to class Ci respectively, and $\overline{\mathbf{y}}$ is the mean for all feature vectors. The optimal projection Wopt is chosen as the matrix with the orthonormal columns which are the generalized eigenvectors corresponding to the m largest eigenvalues,

$$\mathbf{W}_{opt} = \arg\max_{\mathbf{W}} \frac{\left|\mathbf{W}^T \mathbf{S}_b \mathbf{W}\right|}{\left|\mathbf{W}^T \mathbf{S}_W \mathbf{W}\right|} = [\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_{L-1}] . \quad (10)$$

There are at most L-1 nonzero eigenvalues and the corresponding eigenvectors. Finally, each training dynamic feature vector can be represented as

$$\begin{aligned} \mathbf{v}_k &= [\mathbf{w}_1, \ldots, \mathbf{w}_{L-1}]^T [\mathbf{u}_1, \ldots, \mathbf{u}_{d'}]^T \mathbf{x}_k \\ &= \mathbf{T}\mathbf{x}_k \ , k = 1, \ldots, n \end{aligned} \quad (11)$$

In our experiments, the training dataset must be classified to three different groups according to three different walking paths of the video sequence. In testing procedure, given a testing dynamic feature vector, it can be projected into (L-1)-dimensional space.

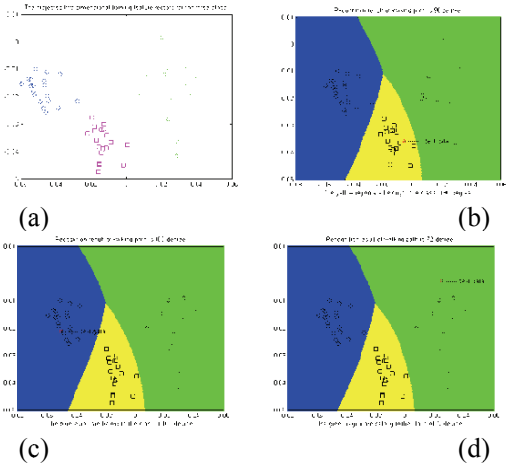
B. Walking Path Determination

We apply Bayesian classifier to divide the feature space into three decision regions corresponding to three different walking paths. The feature vector v belonging to path i if

$$g_i(\mathbf{v}) > g_j(\mathbf{v}) \quad for\ all\ j \neq i, \ \ i = 1,2,3 , \quad (12)$$

where gi(v) denotes a set of discriminant function. For multivariate normal densities $p(x|w_i) \sim N(\mu_i, \Sigma_i)$, the discriminant function can be evaluated and simplified as

$$\begin{aligned} g_i(\mathbf{v}) = &-\frac{1}{2}\mathbf{v}^t \Sigma_i^{-1} \mathbf{v} + (\Sigma_i^{-1}\mu_i)^t \mathbf{v} + \ln p(w_i) \\ &-\frac{1}{2}\ln|\Sigma_i| - \frac{1}{2}\mu_i^t \Sigma_i^{-1} \mu_i \ , \ \ i = 1,2,3. \end{aligned} \quad (13)$$



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

Fig. 5. The classification results of test feature vectors. (a) The training feature vectors from three different classes are projected onto the two-dimensional feature space.

(circle: 108°, square: 90°, diamond: 72°) (b)-(d) The testing feature vector is correctly classified to its class.

Here, we suppose an equal prior probability for each class, i.e. p(w1) = p(w2)= p(w3)=1/3. Figure 5(a) shows the distribution for training data from three different walking paths.

Then using the following decision rule to divide the feature space into three decision regions, $\Re_1$, $\Re_2$, and $\Re_3$, which are referred as three different walking paths,

$$\mathbf{v} \in \Re_i, \ \text{if } g_i(\mathbf{v}) > g_j(\mathbf{v}) \quad \text{for all } j \neq i, \ i = 1,2,3 \quad (14)$$

According to the decision rule, we can classify the feature vector v to the ith class. Figures 5(b) - 5(d) shows the classification results for the test data of three different walking paths.

C. Hybrid Human ID Recognition

Once the human walking path has been identified, the human identity can be recognized. Given a testing dynamic feature vector r, we transform it onto a low-dimensional feature space, i.e., $r' = Tr$. The similarity measurement for dynamic feature vectors is described as

$$D_i(\mathbf{r}', \mathbf{v}_i) = n(n-1)\left\|\mathbf{r}', \mathbf{v}_i\right\| \Big/ \sum_{i=1}^{n}\sum_{j=1, j\neq i}^{n}\left\|\mathbf{v}_i, \mathbf{v}_j\right\| \quad (15)$$

where $\frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1, j\neq i}^{n}\left\|\mathbf{v}_i, \mathbf{v}_j\right\|$ is the average distance between training feature vectors of every two classes, and n is the number of human objects in the database. Hence r' is classified to the kth-human object if and only if $D_k(\mathbf{r}', \mathbf{v}_k) = \min_i D_i(\mathbf{r}', \mathbf{v}_i), \ i = 1, \cdots, n$.

For the similarity measure of the static feature, assuming Ct =[c1t,…, cNtt]T is the testing static feature, and Ci =[c1i,…, cNii]T is the training static feature belonging to the ith-human object, where Nt denotes the number of gait cycles and ckt denotes the Fourier descriptors of the testing sequence. Ni represents the average gait cycles, and cki represents the evaluated Fourier descriptors for the ith-human object. Using Euclidean metric to measure the distance in the frequency domain as

$$D_i(C^t, C^i) = \left\|C^t - C^i\right\| = \frac{\sum_{k=1}^{\min(N_i, N_t)}\left\|c_k^t - c_k^i\right\|}{\min(N_i, N_t)} \quad (16)$$

Finally, Ct is assigned to the kth-human object if and only if $D_k(C^t, C^k) = \min_i D_i(C^t, C^i), \ i = 1, \cdots, n$.

Here we propose a hybrid human ID recognition based on the walking velocity

of human object. If the walking cycle period of the testing sequence is similar to the training sequence, the dynamic ranking is effective; otherwise, the static ranking is considered. The dynamic ranking identifies the human ID based on the similarity of the dynamic features, i.e., Eq. (15), whereas the static ranking recognizes the human ID based on the similarity of the static features, i.e., Eq. (16). The hybrid human identification process consists of the following steps: (1) The difference of gait period between the testing video sequence (Ptest) and the highest rank of human ID from the database ($\hat{P}_{train}$) is calculated. (2) If the difference is lower than threshold TP, we will recognize human ID by dynamic ranking. (3) If the difference of the periods is higher than threshold TP, we calculate the average similarity distance of static feature and compare with threshold TD to determine which ranking is better. (4) When the average similarity distance is lower than threshold TD, we will recognize the human ID by static ranking; otherwise, the dynamic ranking is chosen. The flowchart of the hybrid human ID recognition is described in Fig. 6.

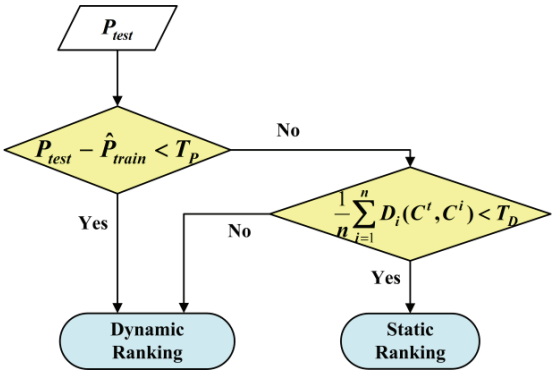


Fig. 6. The flowchart of the hybrid human ID recognition.

V. Experimental Results

The gait video sequences in CASIA database [18] are used to evaluate the effectiveness of our proposed system. The database consists of 124 subjects captured from 11 different views simultaneously, and has 10 walking sequences for each individual. There are six normal walking sequences (i.e., Set A), two carrying-bag sequences (i.e., Set B) and two wearing-coat sequences (i.e., Set C) as shown in Fig. 7. In our experiments, we collect the first four sequences of each individual in Set A as the training set (i.e., Set A1), and use the rest two sequences in Set A (i.e., Set A2), Set B and Set C as the test set. Here, we only select three walking paths, in 72° viewing direction (i.e., Path 1), in 90° viewing direction (i.e., Path 2), and in 108° viewing direction (i.e., Path 3), to test the performance of our proposed

system. Figure 8 shows the images from three different walking paths in the perspective view.

(a) (b) (c)

Fig. 7. Three different sets from CASIA database. (a) Normal walking, (b) walking with wearing a coat, and (c) walking with carrying a bag.

(a) (b) (c)

Fig. 8. Three different walking paths. (a) 72° (Path 1), (b) 90° (Path 2), and (c) 108° (Path 3)

A. Walking Path Recognition and Human ID Recognition

We select 124 human objects from Set A2, Set B, and Set C walking in three different paths. The accuracy of our system for identifying the correct walking path direction is shown in Table 3. Once the walking path is identified, we do the experiments of testing 124 human objects from Set A2, Set B, and Set C in the three different walking paths. The average human ID recognition rates are shown in Table 4. The results show that the recognition rates for the normal walking sequences (Set A2) are much higher than the other two sets. Besides, we can further increase the recognition rate by using the proposed hybrid features.

Table 3. The Walking Path Recognition Rate using CASIA Gait Database.

Correct Path Recognition Ratio				
Input	**Path 1**	**Path 2**	**Path 3**	**Recognition Rate**
Path 1	491	8	5	98.99%
Path 2	5	476	9	95.97%
Path 3	0	12	482	97.18%
Average success rate: 97.38%				

Table 4. The experimental results for average Human ID recognition in multiple paths.

	Human ID Recognition Rate			
	Dynamic Feature Only		**Hybrid features**	
Test Sequence	**Rank1**	**Rank3**	**Rank1**	**Rank3**
Set A2	97.45%	99.33%	97.98%	99.73%
Set B	86.02%	95.70%	86.02%	95.70%
Set C	83.06%	92.20%	85.22%	93.82%
Average Rate	90.99%	96.64%	91.80%	97.24%

B. Human ID Recognition Using Only Training Dataset from Path 2

In the 2nd experiment, we use only the video sequences in 90° viewing direction

(i.e., Path 2) as our training sequence to recognize the walking people captured from three different viewing directions. This experiment aims to test the performance of different methods when the viewing direction of the test video sequence is different from the viewing direction of the training video sequence. Table 5 demonstrates the comparison of our system with other methods where "Sup" is the method of supervised feature selection [13], and "CAS" is the method of using direct GEI (Gait Energy Image) shape match [14]. The experimental results show that our proposed method is insensitive to different test paths by using only the training datasets of path 2. Compared with the supervised feature selection (Sup), which is computationally expensive, our method not only offers a similar performance for normal cases, but also performs better for the special cases of the input videos of people carrying a handbag.

Table 5. Comparison of recognition performance of our proposed method with others by using only the training dataset of path 2.

Comparison of the Recognition performance (%)									
	Set A			**Set B**			**Set C**		
Path	**CAS**	**Sup**	**Ours**	**CAS**	**Sup**	**Ours**	**CAS**	**Sup**	**Ours**
1	82.3	90.3	89.5	42.3	79.4	79.8	20.6	77.5	75.8
2	97.6	98.6	98.4	52	85.5	87.1	32.7	88.7	86.3
3	82.3	78.5	79.8	31.9	60.6	62.9	16.5	62.3	61.3

C. Comparison with Other Proposed Methods

We also do the experiments of using the input videos of 90° viewing direction for both the training and the testing datasets. Table 6 shows the comparison of the proposed method with four different existing methods. The results demonstrate that our method is efficient for human identification of different datasets under various clothing and handbag-carrying conditions.

Table 6. Comparison of recognition performance of our proposed method with other existing methods.

Comparison of the Recognition performance					
	CAS [14]	UCR [5]	Sup[13]	Un-Sup [13]	**Ours**
Set A2	97.6%	99.4%	98.6%	99.4%	98.39%
Set B	32.7%	60.2%	85.5%	79.9%	87.10%
Set C	52.0%	22.0%	88.8%	31.3%	86.29%

VI. Conclusion

In this paper, we propose a novel hybrid system using the dynamic feature extracted by optical flow estimation as well as the static feature to recognize the human walking path and human identification. The proposed method can be applied for

different walking path and walking velocity of the gait video sequence. The experimental results show the effectiveness of our system and demonstrate its ability even for various clothing and carrying conditions.

Acknowledgment

This paper uses the CASIA Gait Database collected by Institute of Automation, Chinese Academy of Sciences.

References

[1] N.V. Boulgouris, D. Hatzinakos, and K.N. Plataniotis, "Gait Recognition: A Challenging Signal Processing Technology for Biometric Identification," IEEE Signal Processing Magazine, Vol. 22, No 6, pp.78 – 90, Nov. 2005.

[2] H. Su and F. G. Huang, "Human Gait Recognition Based on Motion Analysis," in Proceedings of the Fourth International Conference on Machine Learning and Cybernetics, Guangzhou, pp. 4464 – 4468, Aug. 2005.

[3] N. V. Boulgouris and Z. X. Chi, "Gait Recognition Using Random Transform and Linear Discriminant Analysis," IEEE Transactions on Image Processing, Vol. 16, No. 3, pp. 731 – 740, Mar. 2007.

[4] N. V. Boulgouris and Z. X. Chi, "Gait Recognition Based on Human Body Components," in ICIP 2007, pp. I–353–356.

[5] J. Han and B. Bhanu, "Individual Recognition using Gait Energy Image," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 28, No. 2, pp. 316 – 322, Feb. 2006.

[6] J. Liu and N. Zheng, "Gait History Image: A Novel Temporal for Gait Recognition," in ICME 2007, pp. 663 – 666.

[7] Q. Ma, S. Wang, D. Nie and J. Qiu, "Recognizing Humans Based on Gait Moment Image," in *Eighth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing 2007 IEEE*, pp. 606 – 610.

[8] P. S. Huang, C. J. Harris, and M. S. Nixon, "Recognizing Humans by Gait via Parametric Canonical Space," *Artificial Intelligence in Eng.,* Vol. 13, pp. 359 – 366, 1999.

[9] M. H. Cheng, M. F. Ho and C. L. Huang, "Gait Analysis for Human Identification Through Manifold

Learning and HMM," *Pattern Recognition,* Vol. 41, No. 8, pp. 2541 – 2553, Aug. 2008.

[10] A. Bissacco, P. Saisan and S. Soatto, "Gait Recognition using Dynamic Affine Invariants," in *Proc. of the MTNS*, 2004.

[11] S. Chen and Y. Gao, " An Invariant Appearance Model for Gait Recognition," in *ICME* 2007, pp. 1375 – 1378.

[12] 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 Transactions on Pattern Analysis and Machine Intelli*gence, Vol. 27, No. 2, pp. 162 – 177, Feb. 2005.

[13] K. Bashir, T. Xiang, S. Gong and Q. Mary, "Feature Selection on Gait Energy Image for Human Identification," in *ICASSP* 2008, pp. 985 – 988 .

[14] S. Yu, D. Tan and T. Tan, "A Framework for Evaluating the Effect of View Angle, Clothing and Carrying Condition on Gait Recognition," in *ICPR* 2006, pp. 441 – 444.

[15] A. Kale, R. Chowdhury and R. Chellappa, "Towards a View Invariant Gait Recognition Algorithm," in *Proc. Of the IEEE Conf. on Advanced Video and Signal Based Surveillance*, 2003, pp. 143 – 150.

[16] X. Huang and N. Boulgouris, "Model-Based Human Gait Recognition Using Fusion of Features," in *ICASSP* 2009, pp. 1469 – 1472.

[17] R. C. Gonzales and R. E. Woods, *Digital Image Processing*, Prentice Hall.

[18] *CASIA Gait Database*, http://www.sinobiometrics.com, 2006.