A Gait Classification System using Optical Flow Features

CHIH-CHANG YU[1], CHIEN-HUNG CHENG[2] AND KUO-CHIN FAN[2]
[1]*Department of Computer Science and Information Engineering*
Vanung University
Zhongli, 32061 R.O.C
[2] *Department of Computer Science and Information Engineering*
National Central University
Zhongli, 32001, R.O.C

Gait classification is an effective and non-intrusive method for human identification. This paper proposes a system to recognize human identity using optical flow features. The distinguishing characteristic of the proposed system is that we only adopt optical flow information and do not consider shape features or other information. The moving object is detected and located from the flow field using a gaussian model. Afterwards, each subject is identified via the established histogram using optical flow features. The proposed system applies and compares three different kinds of optical flow extraction algorithms. Various experiments with two different databases analyzed and discussed the feasibility of the aprroach. This work demonstrates that optical flow information is useful for gait classification even for unstable optical flows.

Keywords: Gait Classification, optical flow, histogram matching, Principle Component Analysis, Linear Discriminant Analysis

1. INTRODUCTION

Human identification is a popular research in the field of pattern recognition. This kind of applications employed biometric features such as faces, fingerprints/palm-prints, signatures, etc. The disadvantage of these applications is that subjects are required to use some specific devices to obtain features for identifing who they are. Unlike the applications mentioned above, gait classification is an effective and non-intrusive way to recognize different subjects. Human gaits are obtained using a single camera without being percieved. Most researchers proposed silhouette-based methods to solve the gait classification problem. Obviously the performance of the system more or less depends on the quality of the extracted silhouettes. In other words, high recognition rate can be achieved as long as the silhouettes are acceptable. Different from the previous works, this work proposes a non-silhouette based framework by employing optical flow features.

Optical flow is widely used in research topics such as object tracking and action recognition because it is a popular way to describe the motion of moving objects. This inspires us to use optical flow information on a more complicated problem: gait classification. Few researchers discussed the feasibility of adopting optical flow information to solve this problem. In terms of gait classification, most researches integrated optical flow information with contour-based method. For example, previous works made by Little and Boyd [1], Xu *et al.* [2] employed optical flow information and foreground silhouettes as features for gait classification. Meyer [3] proposed a gait classification approach using Hidden Markov Models. However, experiments conducted

in their work were action recognition, not gait classification. This work proposes a new framework which only uses optical flow information for gait classification. Experiments demonstrate that the optical flow features provide sufficient information to recognize different subjects.

2. RELATED WORKS

Generally, human actions recognition / behavior analysis applications are categorized into two kinds of approaches: model-based and model-free approaches [4, 6]. In model-based approaches, the system segments human body into several parts then compares these parts with the pre-trained models [5]. Ju *et al* [7] proposed a cardboard model, where they use patches to represent the segmented body parts. Fujiyoshi *et al* [8] proposed a star skeletionization model to represent the whole human body. The star skeletionization model considers the contour extremities as the features and tries to distinguish between running and walking actions. Wren *et al* [9] proposed a Pfinder real time human tracking system, the system created a multi-class statistical model which considered color and silhouette information to locate, track, and describe human's head and hands. Lakany [10] apllied Continuous Wavelet Transform (CWT) to transform the extracted spatio-temporal features and then employed Kohonen clustering algorithm (SOM) to classify the frequency distribution of the transformed features. Gaits are classified according to their diagnostics or ages. The method is useful on clinical gait diagnosis for patients. Cheng *et al* [11] proposed a manifold learning method from silhouettes in a video sequence. They applied Isometric Feature Mapping (ISOMAP) algorithm for nonlinear dimension reduction and solve the ambiguity on left/right arms and legs in a walking cycle. Then they employed Gaussian Process Latent Variable Model (GP-LVM) to choose discriminative features as the active set from the latent space.

Most human actions recognition / behavior analysis applications are categorized as silhouette-based approach. More specifically, human silhouettes are extracted from video sequences. In this paper, we proposed a framework which extracts optical flow information to classify human gaits. Optical flow algorithms are widely used on motion estimation to detect or track moving objects. Denman *et al* [12] proposed an automatic feedback mechanism to improve the accuracy of optical flow information. Jia *et al* [13] applied optical flow algorithm on autonomous vehicle and people tracking. Nejadasl *et al* [14] use optical flow to track multiple vehicles under low resolution. Their method solved the problems of occluded cars and cars with similar colors. Bhandarkar and Luo [15] proposed a multi-scale elastic matching-based optical flow method which is able to track multiple faces with sudden or gradual changes in illumination, scale and viewpoint.

In gait classification or action recognition application, some previous works also employed optical flow features. Fathi and Mori [16] proposed a mid-level feature representation for action recognition. In their system, low-level optical flow information was extracted first. Then the system applied AdaBoost classifier to select distinguishable features in several subregions. These selected motion features were called mid-level motion features. Finally, the system employed Hamming decoding technique to recognize multiple actions. The recognition results are impressive on actions with significant dif-

ference. Little and Boyd [1] proposed a system to classify human gaits. They applied principle component analysis (PCA) on the binarized optical flow field. They collected 13 types of features and analyzed the effectiveness of these features. The authors proposed a frequency-based method to extract the gait cycle and features. They also adopted Analysis of Variance (ANOVA) to analyze the distinguishability of these features. The authors tried different combination of these features and concluded the importance of them. Ahmad and Lee [17] proposed a method for human action recognition from multi-view image sequences. They combined motion flow and shape flow information. A combined local-global (CLG) optical flow is used to extract motion flow features. The shape flow features were extracted by computing invariant image moments from each video frame. One action is represented as a set of multi-dimensional CLG optical flow and shape flow feature vectors in the spatial and temporal action boundary. Each action in each view is trained by a multi-dimensional HMM. In our work, we adopt a histogram-based and nearest neighbor method to classify human gaits. Experiments validate that the proposed method is simple but effective.

The rest of the paper is organized as follows. Details of the proposed framework are described in section 3, including how to extract optical flow information, moving object detection, and feature selection using PCA and LDA. The experimental results are shown in section 4 followed by conclusions in section 5.

3. Proposed Framework

In this section, we will introduce the proposed framework for human gait classification. In subsection 3.1, we describe the overview of our system. How to extract the moving object from the optical flow field will be illustrated in subsection 3.2. Methods used to extract the histogram-based features will be described in subsection 3.3. We utilize the proposed feature histogram for obtaining human gait cycle and gait classification. Methods will be illustrated in subsection 3.4.

3.1 System Overview

Fig. 1 shows the overview of our system. Firstly, the system converts the input color images to gray level images. Then we employ optical flow algorithm on the image sequence to get the flow field. We employ the Gaussian Model (GM) on the flow field to find the location of the foreground object. The flow histograms are constructed by considering the polar component of flows inside the bounding box of the foreground object. Since the length of the videos may be different due to their walking speed, we apply a histogram matching method to find the gait cycle for each subject to obtain unbiased feature. Afterwards, PCA and LDA are employed to find the best projected feature space. Finally, we use k-Nearest Neighbor (k-NN) classifier to get the recognition results.

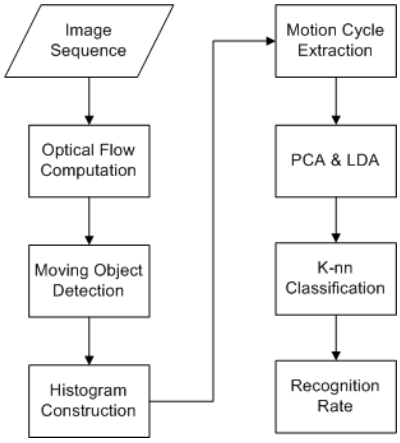


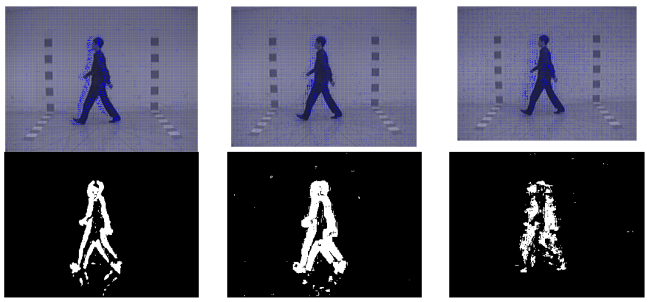
Fig. 1. System Overview.

3.2 Moving Object Detection

In our work, we assume that strong optical flow magnitude mainly occurs around the moving object. Thus, we threshold the flow field to create a binary image using the following equation:

$$I(i,j)=\begin{cases} 0 & \text{if } \sqrt{u_{ij}^2+v_{ij}^2}<1 \\ 1 & \text{otherwise} \end{cases} \tag{1}$$

where i and j are the indices of position in the image, u and v are the horizontal and vertical flows respectively. In our work, three types of optical flow algorithms are employed: Lucas-Kanade [18, 19], Horn-Schunk [20], and Multi-frame [21]. The results of the flow field and corresponding binary images are shown in Fig. 2.



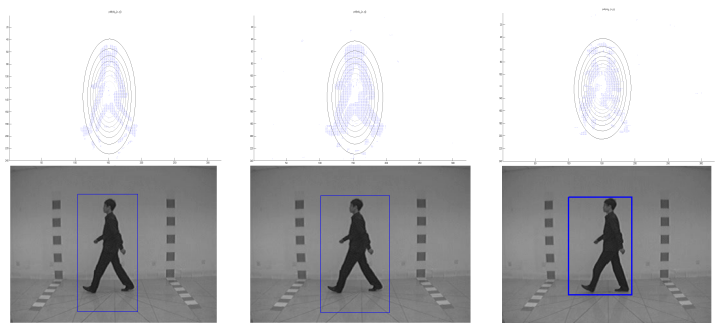
(a) (b) (c)

Fig. 2. The optical flow filed and the corresponding binary images using the (a) Horn-Schunck (b) Lucas-Kanade (c) Multi-Frame method.

After obtaining the optical flow field, we apply a Gaussian Model to find the location of the moving object. We assume that each binarized pixel is a single point in the two-dimensional spatial space. The Gaussian model is simply applied to this space to obtain the rough location of the object. The eigenvalues of the covariance matrix determine the lengths of major and minor axes of the ellipse in the two-dimensional Gaussian distribution. Hence, the width and height of the bounding box are proportional to the lengths of the axes. Since only one moving object exists in the input video, we apply a single Gaussian model for each binary image. After computing the eigenvalues D and eigenvectors V of the covariance matrix from the Gaussian model, we obtain a bounding box by using the following equation:

$$\begin{cases} x_l = \mu_i - V_i \alpha \sqrt{D_i} \\ x_r = \mu_i + V_i \alpha \sqrt{D_i} \end{cases} \tag{2}$$

where x_l and x_r are the upper left and bottom right location of the bounding box respectively. In our work, α is set as 2. The results of the bounding box that locates the moving object are shown in Fig.3. The first row is the results of Gaussian Modeling and the second row is the extracted bounding box on the input video frame.



(a) (b) (c)

Fig.3. Results of moving object detection using the (a) Horn-Schunck (b) Lucas-Kanade (c) Multi-Frame method.

After obtaining the bounding box of the moving object, we consider the optical flow information inside the bounding box to construct the flow histogram. In next subsection, we will explain how to construct the optical flow feature histogram.

3.3 Histogram Construction

The optical flow of each pixel can be represented as a vector $[u\ v]^T$ where u and v are the horizontal and vertical optical flow values respectively. In our work, we transform the optical flow values from the Cartesian coordinate system to the Polar coordinate system $[r\ \theta]^T$ by using the following equations:

$$\begin{cases} r = \sqrt{u^2 + v^2} \\ \theta = \begin{cases} 0 \quad if\ u = 0\ and\ v = 0 \\ \arcsin(\frac{v}{r}) \quad if\ u \geq 0 \\ -\arcsin(\frac{v}{r}) + \pi \quad if\ u \leq 0 \end{cases} \end{cases} \tag{3}$$

The flow histogram is constructed according to the value r and θ. Fig. 4 is an example of the flow histogram. We quantize r and θ from 1 to 10 and 0 to 360 respectively. Hence the number of bins of the flow histogram is 10x360=3600. Since the bounding boxes in different frames have different sizes which result in different number of considered flow points, it is a must to normalize these histograms. The flow histogram FH is normalized by using the following equation:

$$FH(i, j) = \frac{FH(i, j)}{\sum_{\forall i, j} FH(i, j)}, \quad 1 \leq i \leq 10, \quad 0 \leq j \leq 360 \tag{4}$$

where i and j are the indices of the flow histogram.

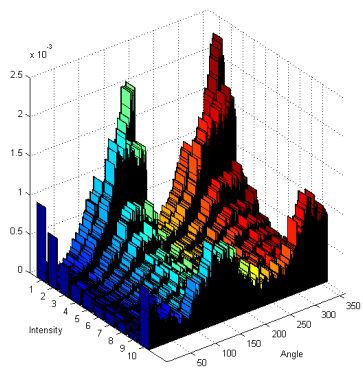


Fig. 4. Example of flow histogram

3.4 Gait Cycle Extraction

We perform histogram matching method to extract the gait cycle. Given a video sequence with T frames, we compute the gait similarity function d by using the following equation:

$$d(t) = f(FH_i, FH_r), \quad 1 \leq i \leq T \tag{5}$$

where FH_i is the flow histogram at frame i and FH_r is the flow histogram of the reference frame. f is the L_2-norm similarity measurement function. To accurately estimate the gait cycle, we take the middle frame (marked with blue rectangle) of the video sequence as the reference frame and use an average filter to smooth the function d. Fig. 5 is an example of the gait similarity function d.

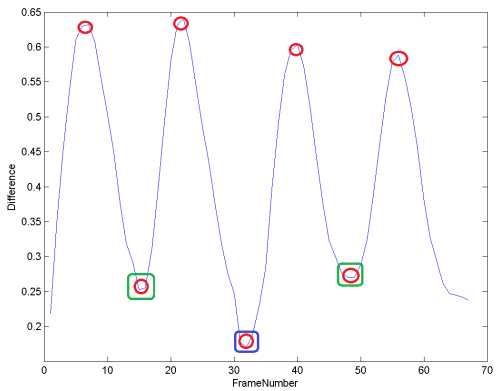


Fig. 5. Gait similarity function.

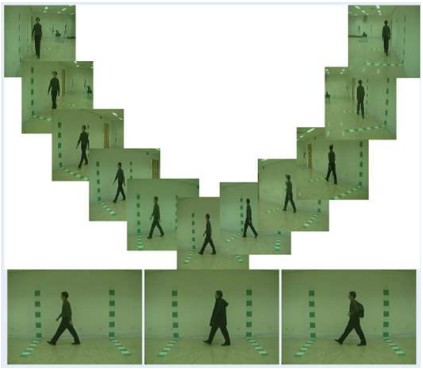
The red circles are the local maximum/minimum points which represent the best (lowest) or the worst (highest) matching results of dissimilarity. Similar gaits in different frames have small values so that we can regard a period between two peaks/valley as a gait cycle. We take the frames between two best matched valley points (marked as green boxes) as a gait cycle. Finally, we can obtain an Averaged Flow Histogram (AFH) using the following equation:

$$AFH = \frac{\sum_{i=t_1}^{t_2} FH_i}{t_2 - t_1} \tag{6}$$

where t_1 and t_2 are the beginning and ending frame indices of a gait cycle. After obtaining the AFH, we employ PCA to reduce the dimension of feature space while maintaining 99% of original information. While applying LDA, the feature dimension will be reduced to N-1 where N is the number of classes (subjects). Finally, the nearest neighbor method is employed to obtin the classification results. In the next section, we will analyze the performance of the proposed framework via some experiments.

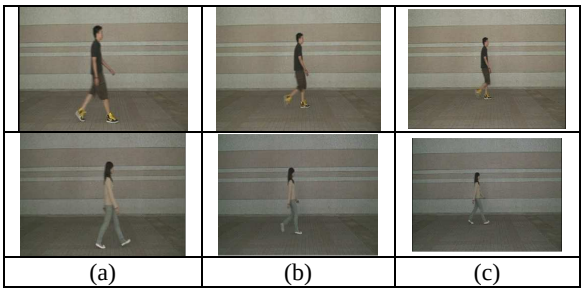
4. EXPERIMENTAL RESULTS

In this section, three experiments are conducted to verify the feasibility and effectiveness of the proposed system. The first database used to verify our framework is the CASIA database [22], which contains three different dataset A, B and C with different environmental settings. We employ the largest dataset B which contains 124 subjects and each subject has 3 different apparels under 11 different camera directions. The angle interval of each camera is 18 degrees (see Fig. 7). In our work, we only use the videos taken from 90 degree. More specifically, subjects are walking perpendicular to the camera view angle. We also compare our work with the silhouette-based methods using this database. The second and third database we used is created by ourselves. In these two databases, subjects have different distance to the camera (see Fig. 8 and Fig. 9). Three optical flow algorithms: Lucas-Kanade (LK), Horn-Schunk (HS) and Multi-frame (MF) are employed in our work. We will explain each experiment in detail in the following subsections.



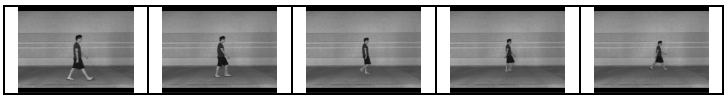
(a) (b) (c)

Fig. 7. CASIA gait database 11 different camera directions and 3 different apparels (a) normal (b) overcoat (c) backpack.



(a) (b) (c)

Fig. 8. Three different ranges in our own gait database (a) near (b) medium (c) far.



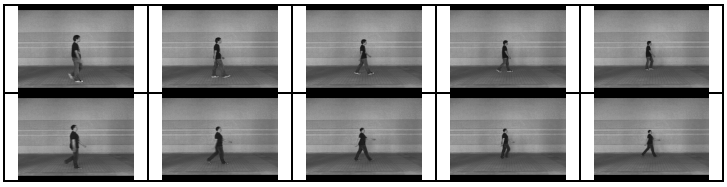


Fig. 9. A dataset contains 5 different distances from subjects to camera

4.1 Experiment 1: Regular Test

The proposed method is tested with the dataset of CASIA database. As shown in Fig. 7(a) to Fig. 7(c), the CASIA database contains three types of wearings: normal clothing (defined as "NM"), overcoat (defined as "OV") and carrying a backpack (defined as "BP") respectively. From all 124 subjects we select 100 subjects for testing. Videos that the optical flow information cannot be successfully extracted are excluded. There are 6 videos in the dataset of NM type for each subject. In the first experiment, we only consider the NM type videos and adopt the leave-one-out cross validation rule for the experiment. The recognition result of our method is shown in Table 1. Table 2 tabulates the results of the silhouette-based method proposed in [23]. Three video sequences are used for training and the rest are used for testing in their approach. Results show that our method is better than silhouette-based methods in dataset B. In this experiment, high recognition rate is achieved using three different types of optical flow algorithms. In our opinion, we think that is because each subject wears same clothing and the distance from the camera and is fixed. We think this condition is not realistic in realworld. Hence, we try to test the robustness of our framework under more challenging conditions.

Table 1. Recognition rate of proposed framework for CASIA normal type.

Normal	Rank 1		Rank 2		Rank 3	
HS	98/100	98%	99/100	99%	99/100	99%
LK	97/100	97%	98/100	98%	99/100	99%
MF	97/100	97%	98/100	98%	98/100	98%

Table 2. Recognition rate of silhouette-based approach for CASIA normal type adopted in [23].

Normal	PCA+LDA	
GEI	319/372	85.67%
EGEI	302/372	81.17%
AEI	330/372	88.71%

4.2 Experiment 2: Apparel Change

In the second experiment, we consider about whether performance of our framework is still good enough when subjects change their clothes. In this experiment, we use all NM type videos (see Fig. 7(a)) for training and use OV and BP type videos for testing. The authors [23] only take 3 NM type videos as the training data and use OV or BP data for

testing. Table 3 and Table 5 depict the classification results of our method of OV and BP data respectively. The classification results of silhouette-based methods of OV and BP data are also shown in Table 4 and Table 6 respectively. The OV data has lower recognition rate than the BP data using the proposed method. Figure 10 shows the extracted optical fields among different apparel. The gait from same object slightly changes when subject is wearing a coat or bringing a bag. If a subject's body is occluded by a backpack or an overcoat, the optical flow information will be different with the NM data. This work compares the results of the proposed method with the results of three silhouette-based methods: AEI, GEI and EGEI [23]. AEI has the best recognition rate among the four methods. The reason is that AEI computes the difference of adjacent frames so that the significance motion information is shown at lower body in AEI images. In other words, the effect of changing clothings can be remedied. We can see that AEI focuses on the information of lower body while other methods consider the overall information of the subject. That is why AEI method outperformed the proposed method in this experiment.

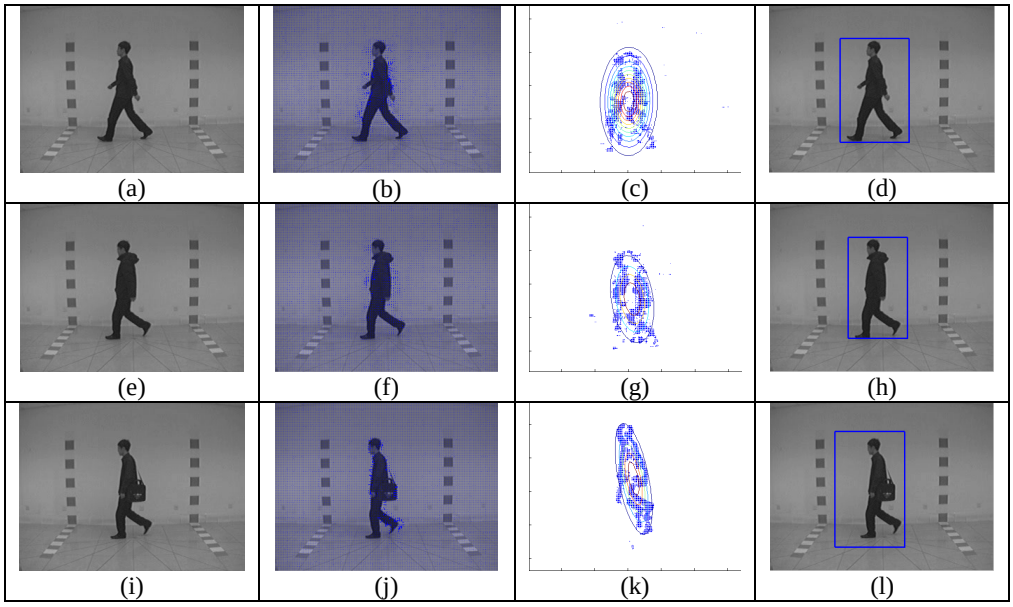


Fig. 10. Comparison of optical flow field with different apparel.

Table 3. Recognition rate of the proposed framework for CASIA overcoat type data.

Overcoat	Rank 1		Rank 2		Rank 3	
HS	44/200	22.00%	51/200	25.50%	55/200	27.50%
LK	43/200	21.50%	56/200	28.00%	64/200	32.00%
MF	34/200	17.00%	41/200	20.50%	46/200	23.00%

Table 4. Recognition rate of the contour-based approach for CASIA overcoat type data in [23].

Overcoat	PCA+LDA	
GEI	64/248	25.81%
EGEI	53/248	21.37%
AEI	143/248	57.62%

Table 5. Recognition rate of the proposed framework for CASIA backpack type data.

Backpack	Rank 1		Rank 2		Rank 3	
HS	143/200	71.50%	165/200	82.50%	172/200	86.00%
LK	143/200	71.50%	158/200	79.00%	162/200	81.00%
MF	134/200	67.00%	144/200	72.00%	149/200	74.50%

Table 6. Recognition rate of the silhouette-based approach for CASIA backpack type data in [23].

Backpack	PCA+LDA	
GEI	85/248	34.27%
EGEI	105/248	42.34%
AEI	186/248	75.00%

4.3 Experiment 3: Variable Distance to Camera

In this experiment, we consider about whether our system is robust enough when subject has different distance to the camera. Since the CASIA does not provide such dataset, we record a new database by ourselves for testing. As shown in Fig. 8, videos with three different distances to camera (nearest, middle and farest) are recorded and the interval between each distance is about 1 meter. For simpilicity, we denote these three types of data as D_{near}, D_{middle} and D_{far} according to the distance from the subject to the camera respectively. This database contains 13 subjects and 9 videos for each person (3 videos for each distance). We try two kinds of tests in this experiment. In the first test, we select 2 videos from D_{near}, D_{middle} and D_{far} for training. Hence, the training data has 6 samples and testing data has 3. Table 7 tabulates the classification results. We can see that each optical flow algorithm performs good recognition rate. The reason is that data in all ranges have been trained. In the second test, we only select one specific distance data for testing and the rest of data for training. Table 7 to Table 9 tabulates the results of using D_{near}, D_{middle} and D_{far} as the testing data respectively.

Experiments show that Lucas-Kanade (LK) method has the lowest recognition rate. It is because our system framework is less favorable for this approach. We compute optical flow of the entire image and do not select good feature points in advance. Although we have removed most noisy flow values, the LK method still provides limited useful information. In our experiments, the LK optical flow algorithm which considers local information has less endurance to the changing of the distance from subjects to camera. Conversely, the Multi-frame method refers more than two images and indeed improves the credibility of the flow values. The HS optical flow algorithm has as good results as MF method. That is because it applied smoothness constraint that remedies the difference of optical flow values between D_{near}, D_{middle} and D_{far}.

We find that the highest recognition rate is achieved when using D_{middle} as the testing

data. We observe that D_{middle} falls between D_{near} and D_{far} in the projected feature space. We also observe that the recognition rate of using D_{near} as testing is the worst. We examine the distribution of the flow histogram from each distance and find that the distribution is not proportion to the ratio of distance. Although we have normalized the feature histogram in our framework, our method can not solve the distance problem. We assume that it is a must to train enough datasets with variable distances to the camera so that the camera distance problem can be remedied.

To verify our assumption, we record an additional dataset (shown in Fig. 9), which contains 11 subjects and 5 different distances to the camera. According to the distance from the subject to the camera, we denote the data as D_1 to D_5, which correspond to the nearest to the farthest distance repectively. Each dataset also contains 3 videos. For this dataset, we make two experiments. In the first experiment, we use D_1, D_3 and D_5 data as the training data while D_2 and D_4 are used for testing. Table 11 tabulates the recognition result. We can see that high recognition rates are achieved whether using the HS or MF method. In the second experiment, we only use D_1 and D_5 data as the training data and the rest are used for testing. Table 12 tabulates the the recognition results. The recognition rates drop significantly with all three optical flow methods. Hence, we conclude that using the data D_1, D_3 and D_5 have better representation ability of the feature space than using D_1 and D_5. Through this test, we can confirm that our assumption is correct.

Table 7. Recognition rate when use data with all three distances for training.

	Rank 1		**Rank 2**		**Rank 3**	
HS	38/39	97.44%	38/39	100.00%	39/39	100.00%
LK	37/39	94.87%	38/39	97.44%	39/39	100.00%
MF	38/39	97.44%	38/39	97.44%	38/39	97.44%

Table 8. Recognition results using the data with the nearest distance as testing.

Nearest	**Rank 1**		**Rank 2**		**Rank 3**	
HS	19/39	48.72%	20/39	51.28%	23/39	58.97%
LK	10/39	25.64%	10/39	25.64%	12/39	30.77%
MF	16/39	41.03%	18/39	46.15%	21/39	53.85%

Table 9. Recognition results using the data with middle distance as testing.

Middle	**Rank 1**		**Rank 2**		**Rank 3**	
HS	35/39	89.74%	38/39	97.44%	39/39	100.00%
LK	11/39	28.21%	12/39	30.77%	18/39	46.15%
MF	36/39	92.31%	39/39	100.00%	39/39	100.00%

Table 10. Recognition results using the data with farthest distance as testing.

Farthest	**Rank 1**		**Rank 2**		**Rank 3**	
HS	25/39	64.10%	27/39	69.23%	27/39	69.23%
LK	16/39	41.03%	19/39	48.72%	19/39	48.72%
MF	30/39	76.92%	30/39	76.92%	30/39	76.92%

Table 11. Recognition results using D_2 and D_4 distance data as the test data.

	Rank 1		Rank 2		Rank 3	
HS	61/66	92.42%	62/66	93.94%	64/66	96.97%
LK	37/66	56.06%	42/66	63.64%	51/66	77.27%
MF	64/66	96.97%	65/66	98.48%	65/66	98.48%

Table 12. Recognition results using D_2, D_3 and D_4 distance data as the test data.

	Rank 1		Rank 2		Rank 3	
HS	69/99	69.70%	78/99	78.79%	82/99	82.83%
LK	25/99	25.25%	32/99	32.32%	43/99	43.43%
MF	77/99	77.78%	82/99	82.83%	85/99	85.86%

5. CONCLUSION

In this paper, we propose a gait classification system using optical flow features. We combine the intensity and orientation information of optical flow and then employ histogram matching technique to extract the gait cycle. The classification is done by using PCA and LDA with the nearest neighbor rule. Through the experiments, we find that the optical flow information of human gaits is a very effective feature to classify human's identity even though the difference of walking postures between different people is hard to observe. Although the accuracy of optical flow information is sometimes inaccurate, we validate that optical flow information is useful for gait classification problem. We only adopt optical flow information and do not consider shape features or other information. Moreover, experiments demonstrated that the framework contributes fair recognition rates in CASIA and our own database. We adopt three different optical flow algorithms in our proposed framework. Through the experiments we conclude that the Horn-Schunck method is the best choice in this framework. It has stable performance on different kinds of test and spends the shortest computation time than the other methods. Although the third method, Multi-Frame optical flow method, has the most stable recognition rate, it requires more time for computation.

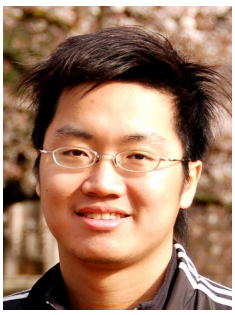
There are still two problems need to be resolved in the future. First, the proposed framework is a view-sensitive approach. Like most silhouette-based methods, it is not easy to compute the gait cycle under some view angles such as walking away from the camera. Because the flow intensity differs in each frame, it is very difficult to normalize the flow intensity of the moving object in order to compute the gait cycle. The second problem is that the system cannot classify human gaits unless the gait cycle is fully extracted. That is, our system cannot handle those videos which contain too few frames.

REFERENCES

1. J. Little, J. Boyd, "Recognition People by Their Gait: The Shape of Motion",

Videre : Journal of Computer Vision Research, vol. 1, no. 2, 1998
2. X. Sun, Q. Zhang, and Y. Xu, “Application of Segmentation Based on Optical Flow for Gait Recognition,” in Proc. 3rd IEEE International Conference on Advanced Computer Theory and Engineering, 2010, pp. 567-571, 2010.
3. Dorthe Meyer, “Human Gait Classification Based on Hidden Markov Models,” 3D Image Analysis and Synthesis'97, 1997
4. D.M. Gavrila, “The Visual Analysis of Human Movement: A survey”, Computer Vision Image Understanding, vol. 73, No.1, 1999, pp. 82-98.
5. C.C. Yu, Y.N. Chen, H.Y. Cheng, J.N. Hwang and K.C. Fan, "Connectivity based Human Body Modeling and Behavior Analysis from Monocular Camera", Journal of Information Science and Engineering, vol.26, no.3, 2010.
6. J.K. Aggarwal and Q. Cai, “Human Motion Analysis: a Review”, Computer Vision and Image Understanding, vol. 73, no.3, 1999, pp.428-440.
7. Shanon X. Ju, Michael J. Black, Yaser Yacoob, “Cardboard People: A parameterized Model of Articulated Image Motion”, Second IEEE International Conference on Face and Gesture Analysis, 1996, pp. 38.
8. Hironobu Fujiyoshi, Alan J. Lipton and Takeo Kanade “Real-Time Human motion Analysis by Image Skeletionization”, IEICE TRANS. INF. & SYST., vol. E87-D, no. 1, 2004
9. C. Wren, A. Azarbayejani, T. Darrell and A. Pentland, “Pfinder Real-Time Tracking of Human Body”, IEEE Transitions on Pattern Analysis and Machine Intelligence, vol. 19, No. 7, 1997, pp. 780-785.
10. Heba Lakany, “Extracting a diagtnostic gait signature”, Pattern Recognition, vol. 41, no. 5, 2008, pp. 1627-1637.
11. Min-Hsu Cheng, Meng-Fen Ho, Chung-Lin Huang, “Gait analysis for human identification through manifold learning and HMM”, Pattern Recognition, vol. 41, no. 8, 2008, pp. 2541-2553.
12. Simon Denman, Vinod Chandran and Sridha Sridharan, “An adaptive optical flow technique for person tracking systems”, Pattern Recognition Letters, vol. 28, 2007, pp. 1232-1239.
13. Zhen Jia, Arjuna Balasuriya and Subhash Challa, “Vision based data fusion for autonomous vehicles target tracking using interacting multiple dynamic models”, Computer Vision and Image Understanding, vol. 109, 2006, pp. 1-21.
14. Fatemeh Karimi Nejadasl, Ben G.H. Gorte and Serge P. Hoogendoorn, “Optical flow based vehicle tracking strengthened by statistical decisions”, ISPRS jornal of Photogrammetry & Remote Sensing, vol. 61, 2006, pp. 159-169.
15. Suchendra M. Bhandarkar and Xingzhi Luo, “Integrated detection and tracking of multiple faces using particle filtering and optical flow-based elastic matching”, Computer Vision and Image Understanding, vol. 113, 2009, pp. 708-725.
16. Alireza Fathi, Greg Mori, “Action Recognition by Learning Mid-level Motion Features”, IEEE Conference on Computer Vision and Pattern Recognition, 2008, pp. 1-8.
17. Mohiuddin Ahmad, Seong-Whan Lee, “Human action recognition using shape and CLG-motion flow from multi-view image sequences”, Pattern Recognition, vol. 41, 2008, pp. 2237-2252.
18. B.D. Lucas, “Generalized Image Matching by the Method of Differences”, PhD

Dissertation, Dept. of Computer Science, Carnegie-Mellon University, 1984.
19. B. Lucas and T. Kanade, ”An iterative image registration technique with an application to stereo vision”, Proc. DARPA IU Workshop, 1981, pp. 121-130.
20. B.K.P Horn and B.G. Schunck, “Determining optical flow”, AI, vol. 71 no. 17, 1981, pp. 185-204.
21. Chia-Ming Wang, Kuo-Chin Fan, and Cheng-Tzu Wang, “Estimating Optical Flow by Integrating Multi-Frame Information,” Journal of Information Science and Engineering, Vol. 24, No. 6, 2008, pp. 1719-1731.
22. Chinese Academy of Sciences, Institute of Automation Gait Database
23. Erhu Zhang, Yongwei Zhao, Wei Xiong, “Active energy image plus 2DLPP for gait recognition”, Signal Processing, vol. 90, 2010, pp. 2295-2302.



Chih-Chang Yu (余執彰) was born in Taipei, Taiwan, on July 30, 1979. He received the BS degree in the Department of Computer Science and Information Engineering from National Central University, Taiwan, ROC in 2002. He received the Ph.D. degree at the same university in 2009. In the same year, he joined the Department of Computer Science and Information Engineering, Vanung University as an assistant professor. His current research interests are optical character recognition, computer vision techniques on mobile applications and image/video analysis.

Chien-Hung Cheng (鄭建宏) received the BS degree in the Department of Computer Science and Information Engineering from Tunghai University, Taiwan, ROC in 2008. He received the Master degree in the Department of Computer Science and Information Engineering, National Central University in 2010. He is the research engineer in Sunplus Technology Co.

Kuo-Chin Fan (**范國清**) was born in Hsinchu, Taiwan, R.O.C., on June 21, 1959. He received the B.S. degree in electrical engineering from the National Tsing-Hua University, Hsinchu, in 1981, and the M.S. and Ph.D. degrees from the University of Florida (UF), Gainesville, in 1985 and 1989, respectively. In 1989, he joined the Institute of Computer Science and Information Engineering, National Central University, Jhongli, Taiwan, where he became a Professor in 1994. From 2008 to 2011, he was the Vice-President of Fo Guang University, I-Lan, Taiwan. His current research interest includes image analysis, optical character recognition, and document analysis.