

# A New Representation for Human Gait Recognition: Motion Silhouettes Image (MSI)

Toby H.W. Lam and Raymond S.T. Lee

Department of Computing, The Hong Kong Polytechnic University,  
Hung Hom, Kowloon, Hong Kong  
{cshwlam, csstlee}@comp.polyu.edu.hk

**Abstract.** Recently, gait recognition for human identification has received substantial attention from biometrics researchers. Compared with other biometrics, it is more difficult to disguise. In addition, gait can be captured in a distance by using low-resolution capturing devices. In this paper, we proposed a new representation for human gait recognition which is called Motion Silhouettes Image (MSI). MSI is a grey-level image which embeds the critical spatio-temporal information. Experiments showed that MSI has a high discriminative power for gait recognition. The recognition rate is around 87% in SOTON dataset by using MSI for recognition. The recognition rate is quite promising. In addition, MSI can also reduce the storage size of the dataset. After using MSI, the storage size of SOTON has reduced to 4.2MB.

## 1 Introduction

In this paper, we proposed a new representation for gait recognition, Motion Silhouettes Image (MSI). The idea of MSI was inspired by the Motion History Image (MHI) which is developed by Bobick and Davis [2]. Bobick and Davis used MHI for motion recognition and applied Hu moments for dimension reduction. Experiments had been done in a test set which contains 18 aerobics exercises and the recognition is above 80%. Our proposed MSI is similar to MHI. However, MSI is simpler and easier to implement. MSI is a gray-level image which embeds the spatial and temporal information. Experiments showed that MSI has a high discriminative power. Besides, it greatly reduced the computational cost and the storage size. In our proposed algorithm, we applied Principal Component Analysis (PCA) on MSIs for reducing the dimensionality of the input space and optimizing the class separability of different MSIs. We use the SOTON dataset [3] to demonstrate the efficacy of the proposed algorithm.

The rest of this paper is organized as follows. We show the related research about gait recognition in Section 2. In Section 3, we reveal and explain the detail about Motion Silhouettes Image (MSI) and the proposed recognition algorithm. The experimental results are shown in Section 4. Conclusion appears in Section 5.

## 2 Related Work

Murase and Sakai [4] proposed a parametric eigenspace representation for moving object recognition. Eigenspace representation is formerly used in face recognition [5].

Murase and Sakai applied eigenspace representation in gait recognition and lip reading. In their proposed algorithm, the extracted silhouettes were projected to the eigenspace by using Principle Components Analysis (PCA). The sequence of movement forms a trajectory in the eigenspace which called parametric eigenspace representation. During recognition, the input image sequence of movement was preprocessed to form a sequence of binary silhouette and these formed a trajectory in the eigenspace. The best match was the one which gets the smallest distance between the input trajectory and the reference sequence. Huang, Harris and Nixon [6] applied the similar approach for gait recognition. Instead of using PCA, they used Linear Discriminating Analysis (LDA), or namely canonical analysis, for transformation. Wang and Tan proposed another new representation for gait recognition [7]. Wang and Tan generated a distance signal by unwrapping the human silhouette. The time-varying distance signals were applied to eigenspace transformation based on PCA. The performance of our proposed algorithm is compared with the performance of Wang's algorithm. For more details, please refer to Section 4.

### 3 Recognition Algorithm

The proposed gait recognition algorithm could be divided into four steps: (i) Image sequences pre-processing, (ii) Motion Silhouettes Image (MSI) generation, (iii) Principle components generation and (iv) Classification. Fig. 1 shows a flow diagram of our proposed algorithm.

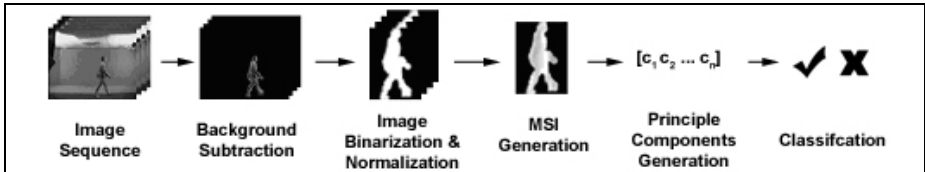


Fig. 1. Flow diagram of gait recognition by using MSI

#### 3.1 Preprocessing

In our proposed gait recognition algorithm, silhouettes are the basis of the feature for recognition. By using background subtraction and thresholding, the silhouettes are extracted from the image sequences [5]. To further eliminate the scaling effect, the silhouettes are extracted according to the size of bounding box and resize to a standard size (128 x 88 pixels). Fig. 2 shows some examples of normalized silhouettes.



Fig. 2. Normalized silhouettes

### 3.2 Motion Silhouettes Image

Motion Silhouettes Image (MSI) is a gray-level image where the pixel intensity is a function of the temporal history of motion of that pixel. The intensity of the MSI represents motion information. MSI embeds the critical spatial and temporal information and it could be formulated by (eqn. 1) . Fig. 3 shows some examples of MSI.

$$MSI(x, y, t) = \begin{cases} 255 & \text{if } I(x, y, t) = 1 \\ \max(0, MSI(x, y, t-1) - 1) & \text{otherwise} \end{cases} \quad (1)$$

where  $I$  is the silhouette image,  $t$  is the current time,  $x$  and  $y$  are the horizontal and vertical coordinates of the image respectively.

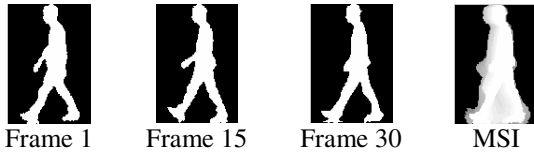


Fig. 3. Examples of Motion Silhouette Image

### 3.3 Training and Transformation

PCA is used for capturing the principle components of the input space. The purpose of using PCA is reducing the feature space to a subspace which maximizing the variance of classes. Suppose there are  $C$  classes for training, each class  $c \in C$  has  $N_c$  of  $q$ -dimensional MSI  $m_{c,i}$ , where  $i$  is the instance label. The total number of training samples is  $N_{total} = N_1 + N_2 + \dots + N_c$ . The average MSI of all samples  $\mu \in \mathfrak{R}^q$  is defined as

$$\mu = \frac{1}{N_{total}} \sum_{c \in C} \sum_{j=1}^{N_c} m_{c,j} \quad (2)$$

and the covariance matrix  $\Sigma_{MSI}$  is defined as

$$\Sigma_{MSI} = \frac{1}{N_{total}} \sum_{c \in C} \sum_{j=1}^{N_c} (m_{c,j} - \mu)(m_{c,j} - \mu)^T \quad (3)$$

A transformation matrix  $W_{pca} = [w_1, w_2, \dots, w_p]$  is obtained where  $w_1, w_2, \dots, w_p$  are the eigenvectors of the samples covariance matrix  $\Sigma_{MSI}$  corresponding to  $p$  ( $p < q$ ) largest eigenvalues. These eigenvectors are orthonormal to each other and span the eigenspace. The original MSI  $m_{c,i}$  is projected to points  $p_{c,i}$  in  $p$ -dimensional eigenspace by using the transformation matrix  $W_{pca}$ .

$$p_{c,i} = W_{pca}^T m_{c,i} = [w_1, w_2, \dots, w_p]^T m_{c,i} \quad (4)$$

where  $c \in C$  and  $i=1, 2, \dots, N_c$

### 3.4 Recognition

Suppose an input image sequence be  $r(t)$ , where  $t = 1, 2 \dots T$ . MSI  $m_r$  is generated by using eqn. (1). The MSI is projected into the eigenspace by

$$z_r = [w_1, w_2, \dots, w_p]^T m_r \quad (5)$$

The Euclidean distance  $E$  between the projected testing feature vector and the projected training feature vector is calculated (see eqn. 6).

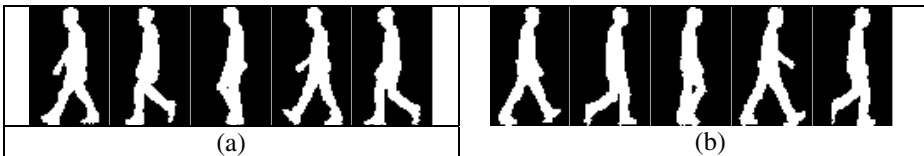
$$E(z_r, z_{c,i}) = \sqrt{(z_r - z_{c,i})^T (z_r - z_{c,i})} \quad (6)$$

where  $z_r$  is the projected feature vector of testing MSI  $m_r$  and  $z_{c,i}$  is the projected feature vector of training MSI  $m_{c,i}$ . The calculated Euclidean distance  $E$  is used for classification by using the nearest neighbor (NN) classifier. The image sequence  $r(t)$  is classified as class  $c$  if the Euclidean distance  $E$  is the minimum among other training samples.

## 4 Experiments and Results

We used SOTON dataset for evaluating the proposed algorithm. SONON dataset was developed by University of University of Southampton [3]. SOTON dataset contains 115 subjects and total 2,128 walking sequences. The walking sequences were filmed in a controlled laboratory environment and captured with a digital video cam at 25 frames per second. The sequence could be divided into two types: (I) subject walking from right to left and (II) subject walking from left to right. For our purposes, the silhouettes is extracted and normalized to a standard size (128 x 88 pixels). We adopted the scheme of FERET [8] for the evaluation and measure the recognition rate by cumulative match score. All experiments are implemented by Matlab and run in a PC computer with P4 2.26GHz and 512MB memory. Fig. 4 shows some walking sequences in SOTON dataset.

In our proposed algorithm, we used MSI as the basis for recognition. As we mentioned before, SOTON dataset contains two types of walking sequences: people walking from left to right and people walking from right to left. We did not separate these sequences for the experiments. In the testing and training subsets, they included sequences from different walking direction. As a result, it is more difficult for the recognition. The experiment could reveal the discriminative power of the proposed representation in gait recognition.



**Fig. 4.** Examples of walking sequence in SOTON dataset: (a) walking from right to left (b) walking from left to right

### 4.1 Recognition with Different Number of Training and Testing Samples

To evaluate the recognition effects with different number of training and testing samples, we conduct three tests: (A) 90% of the image sequences in each class is used for training and other 10% is used for testing; (B) 50% of the image sequences in each class is used for training and other 50% is used for testing; (C) 75% of the image sequences in each class is use for training and other 15% is used for testing. The recognition procedure is mentioned in section 3.2 and the result is shown in Table 1. In this experiment, 95% of the accumulated variance of eigenvalues was used. The best identification rate was 87.15% when using (C) 90% train, 10% test. Fig. 5a shows the ranking order statistics for three different tests. The best identification rate was 87.15% by using (A) 90% train and 10%. The experiment revealed that the overall recognition rate by using (A) 90% train and 10% train is higher than using (B) 75% train and 15% test and (C) 50% train and 50% test (see Fig. 5a).

**Table 1.** Recognition rates for different number of training and testing samples in SOTON

	<b>Top 1</b>	<b>Top 5</b>	<b>Top 10</b>
<b>(A) 90% train 10% test</b>	87.15%	95.18%	95.98%
<b>(B) 75% train 15% test</b>	84.85%	93.76%	95.37%
<b>(C) 50% train 50% test</b>	82.10%	92.31%	94.94%

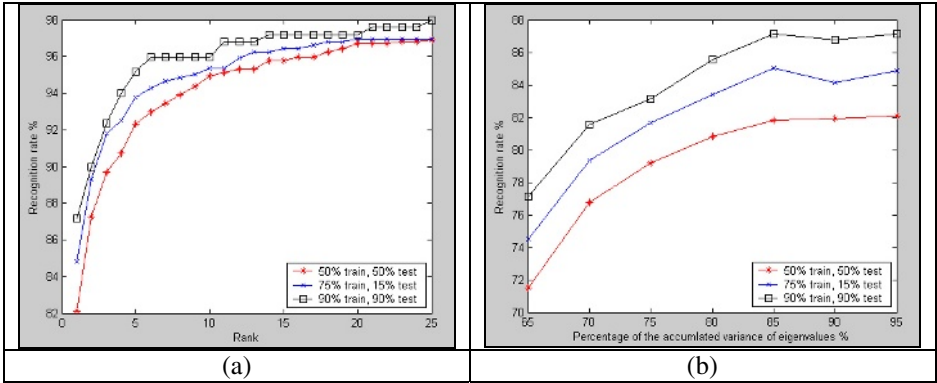
### 4.2 Recognition with Different Eigenvalues

In this experiment, we used seven different percentages of accumulated variance of eigenvalues for evaluating the influence of the eigenvalues. The result is tabulated in Table 2. The result showed that the recognition rate is directly proportional to the number of eigenvectors (see Fig. 5b). Again, the overall recognition rate by using (A) 90% train and 10% test is better than the others (B & C).

**Table 2.** Recognition rates with different eigenvalues

		<b>Percentages of accumulated variance of eigenvalues</b>						
		<b>65%</b>	<b>70%</b>	<b>75%</b>	<b>80%</b>	<b>85%</b>	<b>90%</b>	<b>95%</b>
<b>(A) 90% train, 10% test</b>	(i)	77.11%	81.53%	83.13%	85.54%	87.15%	86.75%	87.15%
	(ii)	87.55%	90.36%	92.37%	93.57%	95.58%	94.78%	95.18%
	(iii)	93.98%	94.38%	95.98%	95.98%	95.98%	95.98%	95.98%
	(iii)							
<b>(B) 75% train, 15% test</b>	(i)	74.51%	79.32%	81.64%	83.42%	85.03%	84.14%	84.85%
	(ii)	88.77%	90.37%	91.27%	92.51%	93.76%	93.58%	93.76%
	(iii)	92.51%	93.94%	94.83%	94.83%	95.01%	95.19%	95.37%
	(iii)							
<b>(C) 50% train, 50% test</b>	(i)	71.51%	76.76%	79.19%	80.79%	81.82%	81.91%	82.10%
	(ii)	87.82%	89.32%	90.91%	91.47%	92.22%	92.41%	92.31%
	(iii)	91.75%	92.88%	94.38%	94.56%	94.75%	94.56%	94.94%
	(iii)							

**(i) – Top 1, (ii) – Top 5, (iii) – Top 10**



**Fig. 5.** (a) Identification performance in terms of rank order statistics. (b) The recognition performance under different eigenvalues.

### 4.3 MSI vs. Unwrapping

We implemented Wang’s unwrapping transformation algorithm by using Matlab [7]. The performance of our proposed algorithm is compared with the performance of the Wang’s unwrapping algorithm. We also used PCA for dimension reduction before classification. We chose 95% of the accumulated variance of eigenvalues in this experiment. Table 3 shows the recognition rate of the unwrapping algorithm and our proposed algorithm. The recognition rate of our proposed algorithm is comparable with unwrapping algorithm.

**Table 3.** Recognition rates by using proposed algorithm and unwrapping algorithm in SOTON database (PCA)

Method	MSI			Unwrapping			
	Rank	Top 1	Top 5	Top 10	Top 1	Top 5	Top 10
<b>a. 50% train 50% test</b>		87.15%	95.18%	95.98%	93.16%	97.56%	98.31%
<b>b. 75% train 25% test</b>		84.85%	93.76%	95.37%	94.47%	97.33%	98.04%
<b>c. 90% train 10% test</b>		82.10%	92.31%	94.94%	94.38%	97.59%	98.39%

Remark: 95% of the accumulated variance of eigenvalues is used in this experiment

## 5 Conclusion and Future Work

In this paper, we proposed a new representation for gait recognition. We used Motion Silhouettes Image (MSI) as the feature template for the recognition. The experiments showed that MSI has a high discriminative ability. The recognition performance is quite promising. The recognition rate is over 80 percent in SOTON dataset by using

MSI. It is comparable with Wang's unwrapping algorithm [7]. The experimental results reveal that MSI has a great potential as a feature template in gait recognition. Furthermore, by using MSI, it can reduce the data storage significantly, but still retain the critical spatial-temporal information for recognition. The original size of SOTON dataset is around 1.92GB. The dataset size reduced to 4.2MB which is reduced 99.7% of the original dataset after transformed to MSIs. In this paper, we evaluated the proposed algorithm in the dataset which captured in the indoor environment. In future, we would like to evaluate the performance of our proposed algorithm under different independent variables situation such as loading, surface and footwear. In addition, we would like to explore other potential representation for gait recognition and we would like to use other dimension reduction technique such as kernel PCA to further improve the recognition rate.

## Acknowledgements

We are grateful to the partial support from the CERG grant for the iMASS Project (B-Q569) and the CRG Project (A-PF74) from the Hong Kong Polytechnic University.

## References

1. M.P. Murray, A.B. Drought and R.C. Kory, "Walking patterns of normal men," *Journal of Bone and Joint Surgery*, Vol. 46 – A, No. 2, pp. 335-60
2. Bobick, A.F.; Davis, J.W., "The recognition of human movement using temporal templates", *IEEE Trans. on PAMI*, Vol. 23, No. 3, pp. 257 – 267, March 2001
3. J. D. Shutler, M. G. Grant, M. S. Nixon, and J. N. Carter "On a Large Sequence-Based Human Gait Database", *Proc. 4th International Conference on Recent Advances in Soft Computing*, Nottingham (UK), pp 66-71, 2002
4. H. Murase, R. Sakai, "Moving object recognition in eigenspace representation: gait analysis and lip reading," *Pattern Recognition Letters*, Vol. 17, pp. 155-62, 1996
5. M. Turk and A. Pentland, "Face Recognition using Eigenfaces," in *Proceedings of the Computer Vision and Pattern Recognition*, 1991
6. P.S. Huang, C.J. Harris and M.S. Nixon, "Human Gait Recognition in Canonical Space Using Temporal Templates," *IEE Proceedings - Vision, Image and Signal Processing*, 146(2), pp. 93-100, 1999
7. L. Wang and T. Tan, "Silhouette Analysis-Based Gait Recognition for Human Identification," *IEEE Trans on PAMI*, Vol. 25 (12), 1505-1518, 2003
8. J. Phillips, H. Moon, S. Rizvi, and P. Raue, "The FERET Evaluation Methodology for Face Recognition Algorithms," *IEEE Trans. PAMI*, Vol. 22, No.8, pp. 1090-1104, Oct. 2000