

A probabilistic image-weighting scheme for robust silhouette-based gait recognition

Heesung Lee · Jeonghyun Baek · Euntai Kim

Published online: 10 July 2012
© Springer Science+Business Media, LLC 2012

Abstract Many gait recognition methods use silhouettes as a feature due to their simplicity and effectiveness. However, silhouette-based gait recognition algorithms have the drawback of performance degradation when the silhouette images are corrupted. To solve this problem, this paper proposes a new gait representation method by emphasizing the noise-free silhouettes while suppressing the corrupted ones. The probabilistic support vector machine (PSVM) is employed to weigh the silhouette images according to quality and to construct a new gait representation for robust recognition. Experiments are conducted with the CASIA and SOTON databases, and the proposed method makes silhouette-based gait recognition as reliable biometrics.

Keywords Gait recognition · Silhouette-based gait recognition · PSVM · CASIA database · SOTON database · Biometrics

1 Introduction

Gait recognition is the identification of individuals based on personal walking motions [6]. The theoretic foundation of gait recognition is the 1964 work by Murray et al. [13], where it was reported that every person has a unique gait. In general, a human's gait can be perceived at a distance [12], and gait recognition has the advantage of being noninvasive and a noncontact process. Furthermore, gait recognition can be used even when the human subject occupies too few image pixels for other biometrics to be perceivable. For this reason, the gait recognition system is an attractive and interesting subject for researchers in the field of computer vision and biometrics. Gait recognition can be roughly classified into two types: model-based and silhouette-based approaches [20].

Model-based methods [2–4] represent the human body or motion by employing explicit models which describe gait dynamics, such as stride dimensions and the kinematics of joint angles. However, model-based approaches are limited by imperfect vision techniques in the body structure/motion modeling and parameter recovery from a walking image sequence. Furthermore, the precision of the model makes the model-based approaches computationally expensive [11].

H. Lee · J. Baek · E. Kim (✉)
School of Electrical and Electronic Engineering, Yonsei University, Sinchon-dong, Seodaemun-gu,
Seoul 120-749, South Korea
e-mail: etkim@yonsei.ac.kr

In contrast, the silhouette-based methods [1, 8, 9, 16, 19] characterize body movement using statistics of walking patterns, which capture both the static and dynamic properties of body shape [10]. In these approaches, the representation method for human gait obviously plays a critical part. Gait energy image (GEI) [8], head-torso-thigh image (HTI) [16], motion silhouette image (MSI) [9], active energy image (AEI) [19], and gait entropy image (GEnI) [1] are general and popular gait representations for silhouette-based methods. GEI is the most popular representation method and is robust against segmental error [8]. HTI is actually a part of GEI and is more robust against variations in walking speed than GEI [16]. MSI is a gray level image which describes the spatiotemporal gait information with high discriminating power [9]. AEI is constructed by accumulating only the active regions of silhouette images. It has the advantage of high quality human silhouettes and sufficient dynamic characteristics [19]. GEnI selects the relevant gait features from each pair of gallery and probe gait sequences by measuring the entropy to perform good recognition [1].

Unfortunately, however, when silhouettes are defective or corrupted, the gait representations are deformed, and the recognition performance is seriously degraded. To solve this problem, the probabilistic support vector machine (PSVM) [5] is employed. PSVM can evaluate the quality of the silhouette images since it yields probabilities of comparable quality to the other methods while still retaining the sparseness of the SVM [14]. The PSVM is trained to distinguish good silhouettes from bad ones and assigns the degree of goodness to each silhouette. Based on the outputs of the PSVM, the robust gait representation image is constructed by emphasizing the noise-free silhouettes while suppressing the corrupted images.

This paper is organized as follows: In Section 2, we provide some background about various gait representations and PSVM. In Section 3, the PSVM-based image-weighting method is presented. In Section 4, the proposed method is applied to the CASIA and SOTON databases to illustrate its performance. Conclusions are drawn in Section 5.

2 Background

2.1 Gait Energy Image (GEI)

GEI is an effective representation scheme with good discriminating power and robustness against segmental errors [8]. Given the preprocessed binary gait silhouette images $f_t(x, y)$ at time t in a sequence, GEI is computed by

$$GEI(x, y) = \frac{1}{N} \sum_{t=1}^N f_t(x, y), \quad (1)$$

where N is the number of frames in the complete gait sequence, and x and y are values in the image coordinates. Figure 1 shows some examples of GEI. In comparison with the binary silhouette sequence, GEI saves both storage space and computation time for recognition and is less sensitive to noise in individual silhouette images. When a silhouette is deformed, however, even GEI exhibits a degraded performance.

2.2 Head-Torso Image (HTI)

HTI [16], proposed by Tan et al., is a gait representation containing only the head and torso part of GEI in order to focus on a more stable region during walking in which there is less

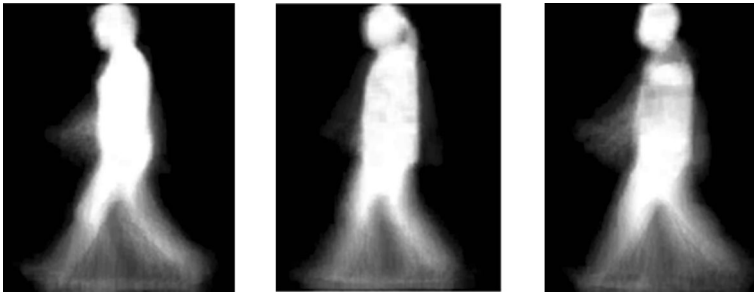


Fig. 1 Examples of GEI

movement. Furthermore, we often identify people by the profile of the head and torso. From these observations, Tan et al. [16] use only the head-torso-thigh parts of human silhouettes to represent the human gait and refer to it as HTI.

HTI can be viewed to some extent as a first-order-statistic-based description of human gait from the structural perspective. HTI is defined as

$$HTI(x, y) = \frac{1}{N} \sum_{t=1}^N H_t(x, y), \quad (2)$$

where $H_t(x, y)$ is obtained by removing part of the crura from $f_t(x, y)$, and N is the number of frames in one sequence. Figure 2 shows some examples of HTI.

2.3 Motion Silhouette Image (MSI)

MSI [9] is a gray level image which retains the critical spatiotemporal information. The intensity of each pixel in an MSI is a function of its temporal history. MSI has high discriminating power and remains the critical information source for gait recognition. MSI is obtained by the following:

$$MSI_t(x, y) = \begin{cases} 255 & \text{if } f_t(x, y) = 1 \\ \max(0, MSI_{t-1}(x, y) - 1) & \text{otherwise.} \end{cases} \quad (3)$$

Figure 3 shows examples of MSI.

2.4 Active Energy Image (AEI)

Existing gait representation methods usually have defects related to both the low quality of human silhouettes and insufficient dynamic characteristics. AEI [19] is constructed by



Fig. 2 Examples of HTI



Fig. 3 Examples of MSI

accumulating only the active regions of silhouette images in order to make up for those defects. Given a binary gait silhouette image $f_t(x, y)$, we can calculate the difference image between two adjacent silhouettes as follows:

$$D_t(x, y) = \begin{cases} f_t(x, y) & t = 1 \\ \|f_t(x, y) - f_{t-1}(x, y)\| & t > 1 \end{cases} \tag{4}$$

where $\|\cdot\|$ is the Euclidean norm, and $D_t(x, y)$ is the difference between $f_t(x, y)$ and $f_{t-1}(x, y)$, i.e., $D_t(x, y)$ is the active region at time t , and it is desirable to use the difference image to extract the dynamic parts of the moving body. Accumulating the difference images, we can obtain AEI as

$$AEI(x, y) = \frac{1}{N} \sum_{t=1}^N D_t(x, y). \tag{5}$$

Figure 4 shows some examples of AEI.

2.5 Gait Entropy Image (GEnI)

GEnI [1] is proposed to distinguish the dynamic and static areas of a GEI by measuring the Shannon entropy at each pixel location. The gait cycle consists of a sequence of human silhouettes, and the intensity value of the silhouettes at a fixed pixel location can be considered as a discrete random variable. The Shannon entropy measures the uncertainty

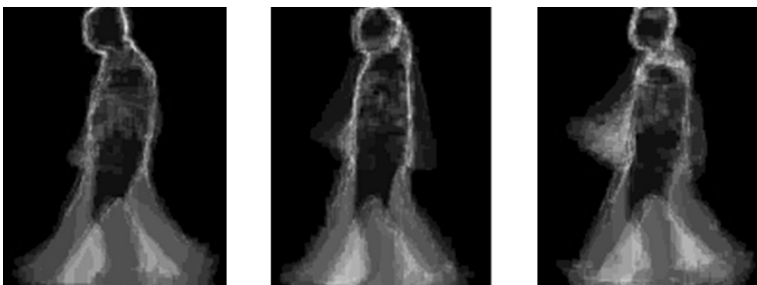


Fig. 4 Examples of MSI

associated with the random variable over a complete gait cycle. To obtain the information content of the gait sequence proportional to the entropy value, GENI is defined as

$$\text{GENI}(x, y) = - \sum_{k=0}^K p_k(x, y) \log_2 p_k(x, y), \quad (6)$$

where $p_k(x, y)$ is the probability that a pixel (x, y) has the value k . Since $f_i(x, y)$ is a binary gait silhouette image with either 0 or 1, we set $K=1$. The probability that the pixel has a value of 1 is $p_1(x, y) = \frac{1}{T} \sum_{t=1}^T f_t(x, y)$ (i.e., the GEI), and the probability that the pixel has a value of 0 is $p_0(x, y) = 1 - p_1(x, y)$. Figure 5 shows some examples of GENI.

2.6 Probabilistic Support Vector Machine (PSVM)

A posterior probability is very important in practical classification problems. SVM [17], however, cannot produce posterior probability but only a binary decision value. PSVM [5] maps the output SVM to the interval $[0, 1]$ using a sigmoid function

$$P(\omega = \{+1, -1\} | f) = \frac{1}{1 + \exp(-d(f))} \quad (7)$$

and determines the probability at which a data point belongs to a class where f is a data point; ω denotes its binary class with $\{+1, -1\}$, and $d(f)$ is the unthresholded output of SVM defined as

$$d(f) = w^T \cdot f + b, \quad (8)$$

where w and b denote a weight for f and the bias, respectively, in the SVM.

3 Probabilistic image weighting scheme

In this section, we describe a new image-weighting scheme for robust silhouette-based gait recognition. An SVM classifier is trained to distinguish noise-free images from corrupted ones. The logistic sigmoid function is then employed to map the output of the SVM classifier to a probability interval $[0, 1]$, assigning the probability to each test input image as a weight. Finally, a new gait representation is constructed using the previous methods such as GEI, HTI, MSI, etc., with the only exception being that the weights computed by the PSVM are used to represent the “goodness” of the silhouettes. For this reason, our methods can be combined



Fig. 5 Examples of GENI

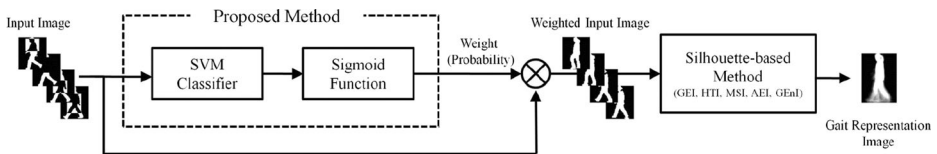


Fig. 6 Overview of the proposed method

with various silhouette-based gait-recognition systems to suppress the silhouette noise and improve the quality of the gait features. Figure 6 shows an overview of the proposed method.

3.1 Construction of the decision boundary

Both noise-free and corrupted images are used to train the SVM. For corrupted images, two kinds of rectangular noise are used: one is a black box noise, as shown in Fig. 7(b), and the other is a complement box noise which inverts the original color, as shown in Fig. 7(c). SVM is trained to separate the noise-free gait silhouettes from the corrupted ones. After training, when a new silhouette f is presented, the probability that the silhouette is noise-free is computed by

$$P(\omega = \{\text{noise-free}\} | f) = \frac{1}{1 + \exp(-d(f))} = \frac{1}{1 + \exp(-w^T \cdot f - b)}, \tag{9}$$

where w and b denote the weight for f and the bias in the SVM, respectively.

3.2 Probabilistic image weighting

SVM is an empirically optimal discriminant model which provides a deterministic binary decision on an input data point. In this paper, a scaled logistic sigmoid is employed to map the distance between the image and the decision boundary to the goodness of the image. Figure 8 shows the decision boundary of the SVM classifier and the normal or noisy images. In Fig. 8, $d(f)$ ¹ denotes the distance between the image and decision boundary and is defined in (8). In addition, ξ , is a small positive value used to define the range of the uncertain region in the SVM learning.

As shown in Fig. 8, images near the decision boundary are difficult to classify as normal or noisy, while images at a distance greater than ξ from the decision boundary can be classified. We design the weight scheme using a sigmoid function to efficiently use barely classified images in the construction of a silhouette-based image. The image weight using the sigmoid function is represented by

$$w(f) = \frac{1}{1 + \exp(-(5/\xi) \cdot d(f))}. \tag{10}$$

Figure 9 shows the weight function. When $d(f) > \xi$, the silhouette is noise-free, and the full image is used in the gait representation method. Similarly, when $d(f) < -\xi$, the silhouette is completely corrupted and should be excluded. When $\xi \geq d(f) \geq -\xi$, the silhouette is moderately corrupted and should be weighted according to the degree of corruption. The weighting scheme can then be expressed as

$$\hat{f}(x, y) = w(f) \cdot f(x, y) \tag{11}$$

¹ We assign the classes of clear and noisy images to “+1” and “-1” respectively. Therefore the positive direction of the normal vector in the decision boundary points to the set of noise-free silhouettes.

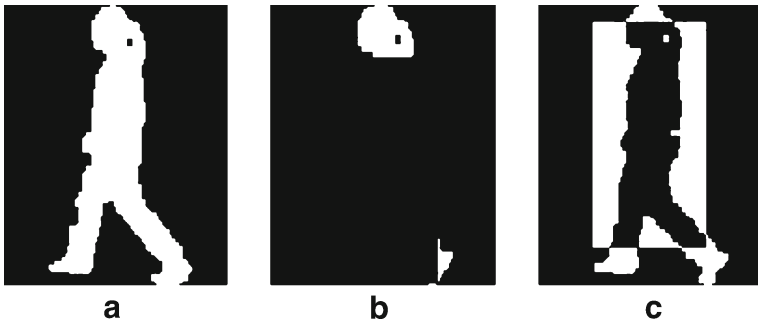


Fig. 7 Normal and noisy images: **a** normal image, **b** black box noise, **c** complement box noise

where $f(x, y)$ is the original silhouette image, $w(f)$ is the output value of the PSVM, and $\hat{f}(x, y)$ is the weighted silhouette image. Finally, the gait representation image is represented by $\hat{f}(x, y)$. Table 1 shows how the new gait representations are computed using $\hat{f}(x, y)$

3.3 Consideration

In this subsection, the performance of the proposed method is analyzed based on a simplified gait silhouette model. The model used herein is similar to the one in [8]. For the sake of simplicity and easy understanding, a single pixel model is used as a gait silhouette. That is,

$$b = f + \eta. \tag{12}$$

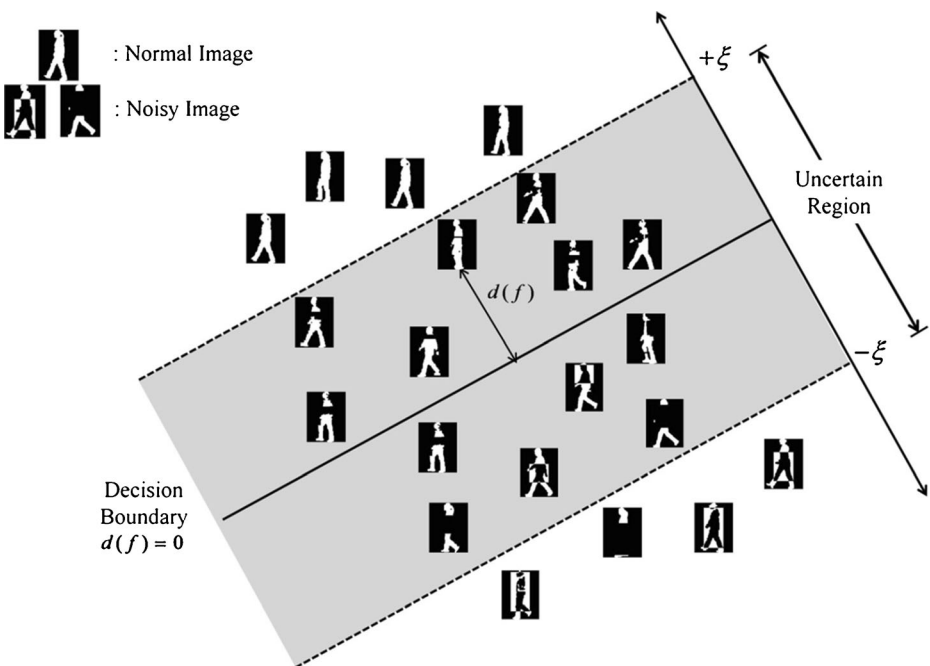
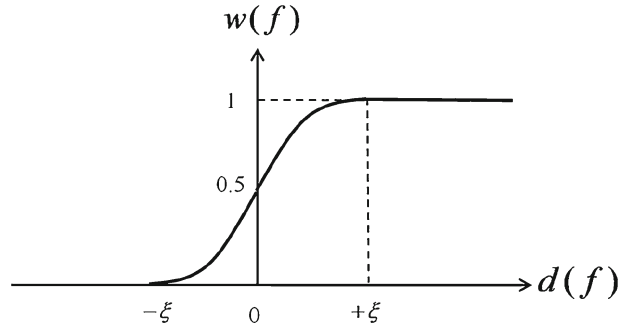


Fig. 8 Decision boundary and normal or noisy images

Fig. 9 Sigmoid weight function



where $f \in \{0, 1\}$ and $\eta \in \{-1, 0, 1\}$ denote the noise-free silhouette and noise, respectively, and $b \in \{0, 1\}$ denotes the noisy silhouette. Even though many features such as GEI, HTI, MSI, AEI and GENI were considered, only GEI is used for mathematical consideration. First of all, let us assume that the probabilities of silhouette image f being 0 is and 1 are q and $(1 - q)$, respectively:

$$P(f) = \begin{cases} P(f = 0) = q \\ P(f = 1) = 1 - q. \end{cases} \tag{13}$$

We further assume that the probability of silhouette image f being flipped by noise η is silhouette image p , thus

$$P(\eta|f) = \begin{cases} P(\eta = -1|f = 1) = p \\ P(\eta = 0|f = 1) = 1 - p \\ P(\eta = 1|f = 0) = p \\ P(\eta = 0|f = 0) = 1 - p. \end{cases} \tag{14}$$

Using our weighting scheme, (12) can be represented as

$$wb = wf + w\eta \tag{15}$$

where $w = w(f)$. For simplification, if the silhouette image is considered as a normal image (positive), then assign a positive number u for w and otherwise, we assign $(1 - u)$ for w as follows:

$$w = \begin{cases} u & \text{if } d(f) > 0 \\ 1 - u & \text{otherwise} \end{cases} \tag{16}$$

where $d(f)$ is the unthresholded output of SVM. In this case, w can be regarded as random variable according to SVM output, and it satisfies the following distribution:

$$P(w|\eta, f) = \begin{cases} P(w = u|\eta = 0, f = 1) = \alpha \\ P(w = 1 - u|\eta = 0, f = 1) = 1 - \alpha \\ P(w = u|\eta = -1, f = 1) = 1 - \beta \\ P(w = 1 - u|\eta = -1, f = 1) = \beta \\ P(w = u|\eta = 0, f = 0) = \alpha \\ P(w = 1 - u|\eta = 0, f = 0) = 1 - \alpha \\ P(w = u|\eta = 1, f = 0) = 1 - \beta \\ P(w = 1 - u|\eta = 1, f = 0) = \beta \end{cases} \tag{17}$$

Table 1 Preparation of a new gait representation image

<p>input:</p> <p>$f_1(x, y), \dots, f_N(x, y)$: noisy gait silhouette images</p> <p>ξ: range of the uncertain region</p> <p>$d(f_1), \dots, d(f_N)$: unthresholded SVM output value of noisy gait silhouette images.</p> <p>initialize: $n = 0$.</p> <p>output:</p> <p>New silhouette-based representation images: GEI, HTI, MSI, AEI, and GENI.</p> <p>1. Construct the weighted gait silhouette image sequence: $\hat{f}_1(x, y), \dots, \hat{f}_N(x, y)$.</p> <p>for $t \leftarrow 1$ to N</p> <p style="padding-left: 20px;">Calculate weight $w(f_t)$ from the sigmoid weight function:</p> $w(f_t) = \frac{1}{1 + \exp(-(5/\xi) \cdot d(f_t))}$ <p style="padding-left: 20px;">if $w(f_t) \neq 0$ then</p> <p style="padding-left: 40px;">1) Increase n by 1: $n = n + 1$.</p> <p style="padding-left: 40px;">2) Multiply the weight by the noisy gait silhouette image: $\hat{f}_n(x, y) = w(f_t) \cdot f_t(x, y)$.</p> <p>2. Set the number of weighted gait silhouette images: $K = n$.</p> <p>3. Prepare new silhouette-based gait representation images (GEI, HTI, MSI, AEI, and GENI) with the weighted gait sequence: $\hat{f}_1(x, y), \dots, \hat{f}_K(x, y)$.</p>

where α and β are the performance of SVM. Using true positive (TP), true negative (TN), false positive (FP), and false negative (FN), we can obtain

$$\alpha = \frac{TP}{TP + FN} \quad \left(1 - \alpha = \frac{FN}{TP + FN} \right). \quad (18)$$

Table 1 (continued)

<p>-GEI</p> <p>1) Compute the average image of the weighted gait sequence $\hat{f}_1(x, y), \dots, \hat{f}_K(x, y)$:</p> $GEI(x, y) = \frac{1}{K} \sum_{k=1}^K \hat{f}_k(x, y).$ <p>-HTI</p> <p>1) Obtain $\hat{H}_1(x, y), \dots, \hat{H}_K(x, y)$ in which the crura is removed from $\hat{f}_1(x, y), \dots, \hat{f}_K(x, y)$.</p> <p>2) Compute the average image of the gait sequence $\hat{H}_1(x, y), \dots, \hat{H}_K(x, y)$:</p> $HTI(x, y) = \frac{1}{K} \sum_{k=1}^K \hat{H}_k(x, y).$ <p>-MSI</p> <p>1) Compute MSI with the weighted gait silhouette sequence: $\hat{f}_1(x, y), \dots, \hat{f}_K(x, y)$.</p> <pre style="margin-left: 20px;"> for $k \leftarrow 1$ to K if $k = 1$ then $MSI_1(x, y) = \hat{f}_1(x, y)$ else for $x \leftarrow 1$ to width of $\hat{f}_k(x, y)$ for $y \leftarrow 1$ to height of $\hat{f}_k(x, y)$ if $\hat{f}_k(x, y) = 1$ then $MSI_k(x, y) = 255$ else $MSI_k(x, y) = \max(0, MSI_{k-1}(x, y) - 1)$ </pre>

and

$$\beta = \frac{TN}{TN + FP} \left(1 - \beta = \frac{FP}{TN + FP} \right). \tag{19}$$

Table 1 (continued)

<p>-AEI</p> <p>1) Compute the difference gait silhouette image $D_1(x, y), \dots, D_K(x, y)$.</p> <p style="margin-left: 20px;">for $k \leftarrow 1$ to K</p> <div style="margin-left: 40px;"> <p>if $k = 1$ then</p> <div style="margin-left: 20px;"> $D_k(x, y) = f_k(x, y)$ </div> <p>else</p> <div style="margin-left: 20px;"> $D_k(x, y) = \ f_k(x, y) - f_{k-1}(x, y)\$ </div> </div> <p>2) Compute the average image of the difference gait silhouette images $D_1(x, y), \dots, D_K(x, y)$:</p> $AEI(x, y) = \frac{1}{K} \sum_{k=1}^K D_k(x, y).$ <p>-GENI</p> <p>1) Compute the probability that each pixel of $\hat{f}_k(x, y)$ has a value of 1 ($p_1(x, y)$) or 0 ($p_0(x, y)$):</p> $p_1(x, y) = \frac{1}{K} \sum_{k=1}^K \hat{f}_k(x, y), \quad p_0(x, y) = 1 - p_1(x, y).$ <p>2) Compute the GENI with two probabilities, $p_0(x, y)$ and $p_1(x, y)$:</p> $GENI(x, y) = - \sum_{i=0}^1 p_i(x, y) \log_2 p_i(x, y)$
--

Therefore, we have

$$P(wf) = \begin{cases} P(wf = u) = \alpha(1 - p)(1 - q) & \text{when } \eta = 0, f = 1 \\ P(wf = 1 - u) = (1 - \alpha)(1 - p)(1 - q) & \text{when } \eta = 0, f = 1 \\ P(wf = u) = (1 - \beta)p(1 - q) & \text{when } \eta = -1, f = 1 \\ P(wf = 1 - u) = \beta p(1 - q) & \text{when } \eta = -1, f = 1 \\ P(wf = 0) = \alpha(1 - p)q & \text{when } \eta = 0, f = 0 \\ P(wf = 0) = (1 - \alpha)(1 - p)q & \text{when } \eta = 0, f = 0 \\ P(wf = 0) = (1 - \beta)pq & \text{when } \eta = 1, f = 0 \\ P(wf = 0) = \beta pq & \text{when } \eta = 1, f = 0 \end{cases} \quad (20)$$

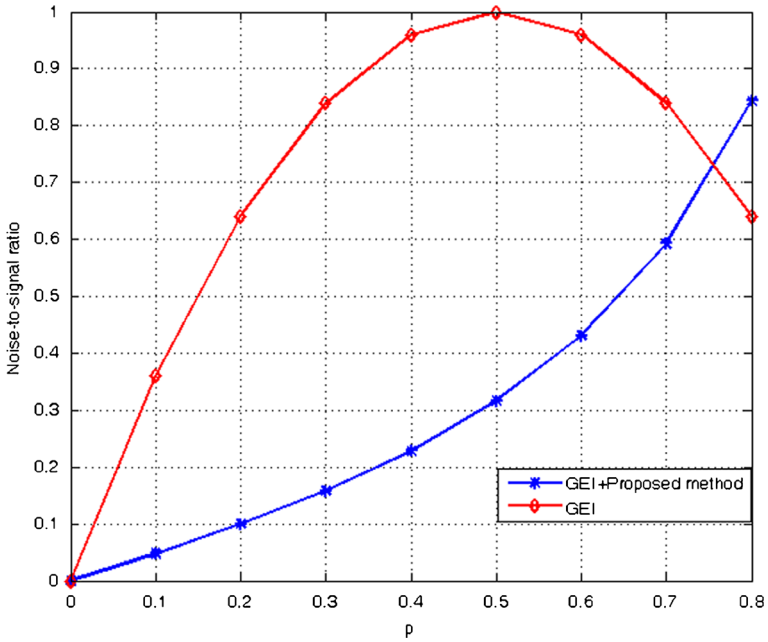


Fig. 10 NSRs of the conventional and proposed methods: $\alpha=\beta=0.9$

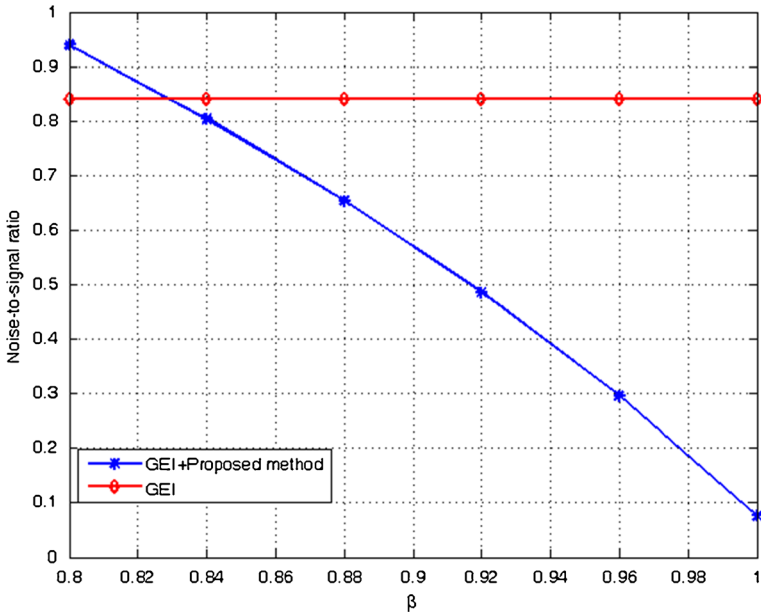


Fig. 11 NSRs of the conventional and proposed methods: $\alpha=0.9$ and $p=0.7$

Table 2 Comparison of performance: GEI

Noise rate (%)	GEI (%)	PSVM+GEI (%)	IMP (%)
0	83.33	81.67	-1.66
10	78.00	81.67	3.67
15	80.00	83.00	3.00
20	80.00	81.67	1.67
25	76.67	85.00	8.33
30	75.00	80.00	5.00
35	71.67	83.33	11.66
40	65.00	85.00	20.00
Average	76.21	82.67	6.46

and

$$P(w\eta) = \begin{cases} P(w\eta = 0) = \alpha(1-p)(1-q) & \text{when } \eta = 0, f = 1 \\ P(w\eta = 0) = (1-\alpha)(1-p)(1-q) & \text{when } \eta = 0, f = 1 \\ P(w\eta = -u) = (1-\beta)p(1-q) & \text{when } \eta = -1, f = 1 \\ P(w\eta = -(1-u)) = \beta p(1-q) & \text{when } \eta = -1, f = 1 \\ P(w\eta = 0) = \alpha(1-p)q & \text{when } \eta = 0, f = 0 \\ P(w\eta = 0) = (1-\alpha)(1-p)q & \text{when } \eta = 0, f = 0 \\ P(w\eta = u) = (1-\beta)pq & \text{when } \eta = 1, f = 0 \\ P(w\eta = (1-u)) = \beta pq & \text{when } \eta = 1, f = 0 \end{cases} \quad (21)$$

In order to see the effect of weight w , the noise-to-signal ratios (NSRs) of the conventional and proposed methods are compared. If NSR of the proposed method is lower than that of the conventional method, we can conclude that the effect of noise is relatively decreased by using our method. Given a walking cycle with N frames, NSR of GEI is defined as

$$\frac{\text{Var}\left(\frac{1}{N} \sum_{i=1}^N \eta_i\right)}{\text{Var}\left(\frac{1}{N} \sum_{i=1}^N f_i\right)} = \frac{\text{Var}(\eta)}{\text{Var}(f)}. \quad (22)$$

Table 3 Comparison of performance: HTI

Noise rate (%)	HTI (%)	PSVM+HTI (%)	IMP (%)
0	90.00	88.33	-1.67
10	90.00	90.00	0.00
15	81.67	86.67	5.00
20	86.67	90.00	3.33
25	75.00	86.67	11.67
30	78.33	88.33	10.00
35	75.00	88.33	13.33
40	63.33	86.67	23.34
Average	80.00	88.13	8.13

Table 4 Comparison of performance: MSI

Noise rate (%)	MSI (%)	PSVM+MSI (%)	IMP (%)
0	61.67	66.67	5.00
10	15.00	63.00	48.00
15	28.00	63.00	35.00
20	18.33	58.33	40.00
25	26.67	60.00	33.33
30	25.00	58.33	23.33
35	25.00	56.67	31.67
40	21.00	63.00	42.00
Average	27.58	61.13	32.29

From (13) to (14), $Var(\eta)$ and $Var(f)$ can be represented in terms of p, q as

$$Var(\eta) = p(1 - p) \tag{23}$$

$$Var(f) = q(1 - q) \tag{24}$$

Thus,

$$\frac{Var\left(\frac{1}{N} \sum_{t=1}^N \eta_t\right)}{Var\left(\frac{1}{N} \sum_{t=1}^N f_t\right)} = \frac{Var(\eta)}{Var(f)} = \frac{p(1 - p)}{q(1 - q)}. \tag{25}$$

In similar, NSR of the GEI combined by the proposed method is defined as

$$\frac{Var\left(\frac{1}{N} \sum_{t=1}^N w_t \eta_t\right)}{Var\left(\frac{1}{N} \sum_{t=1}^N w_t f_t\right)} = \frac{Var(w\eta)}{Var(wf)} = \frac{E(w^2 \eta^2) - (E(w\eta))^2}{E(w^2 f^2) - (E(wf))^2}. \tag{26}$$

Table 5 Comparison of performance: AEI

Noise rate (%)	AEI (%)	PSVM+AEI (%)	IMP (%)
0	78.33	81.67	3.34
10	53.33	83.33	30.00
15	36.37	81.67	45.00
20	31.67	80.00	48.33
25	18.33	83.33	65.00
30	16.67	85.00	68.33
35	23.33	71.67	48.34
40	20.00	70.00	50.00
Average	34.75	79.58	44.79

Table 6 Comparison of performance: GEnI

Noise rate (%)	GEnI (%)	PSVM+GEnI (%)	IMP (%)
0	86.67	85.00	-1.67
10	71.67	88.33	16.66
15	70.00	88.33	18.33
20	73.00	85.00	12.00
25	63.33	88.33	25.00
30	51.07	85.00	34.33
35	60.00	85.00	25.00
40	40.00	81.67	41.67
Average	64.47	85.83	21.42

From (20), $E(w\eta)$ and $E(w^2\eta^2)$ are

$$\begin{aligned}
 E(w\eta) &= -u(1-\beta)p(1-q) - (1-u)\beta p(1-q) + u(1-\beta)pq + (1-u)\beta pq \\
 &= u(1-\beta)p(2q-1) + (1-u)\beta p(2q-1) \\
 &= p(2q-1)\{u(1-\beta) + (1-u)\beta\}
 \end{aligned}
 \tag{27}$$

and

$$\begin{aligned}
 E(w^2\eta^2) &= u^2(1-\beta)p(1-q) + (1-u)^2\beta p(1-q) + u^2(1-\beta)pq + (1-u)^2\beta pq \\
 &= u^2(1-\beta)p + (1-u)^2\beta p \\
 &= p\{u^2(1-\beta) + (1-u)^2\beta\}.
 \end{aligned}
 \tag{28}$$

In the same way, $E(wf)$ and $E(w^2f^2)$ are

$$\begin{aligned}
 E(wf) &= u\alpha(1-p)(1-q) + (1-u)(1-\alpha)(1-p)(1-q) \\
 &\quad + u(1-\beta)p(1-q) + (1-u)\beta p(1-q) \\
 &= \{u\alpha + (1-u)(1-\alpha)\}(1-p)(1-q) + \{u(1-\beta) + (1-u)\beta\}p(1-q)
 \end{aligned}
 \tag{29}$$

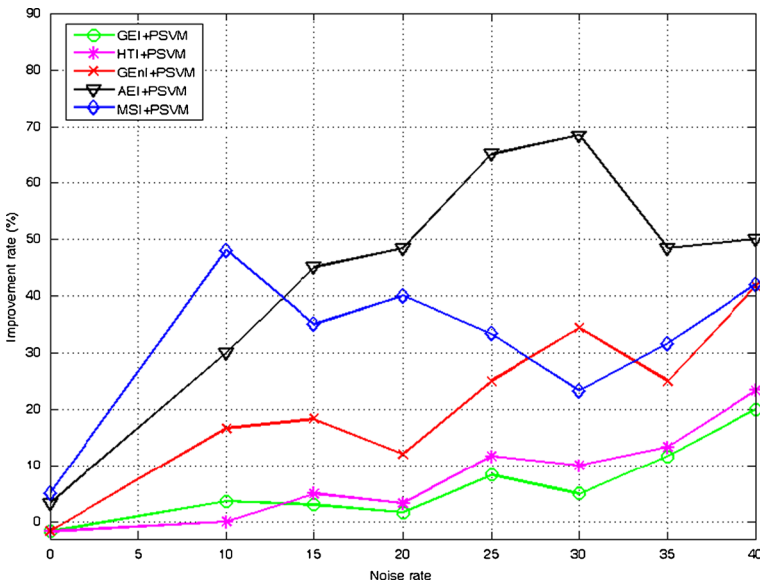


Fig. 12 Improvement rate according to noise rate

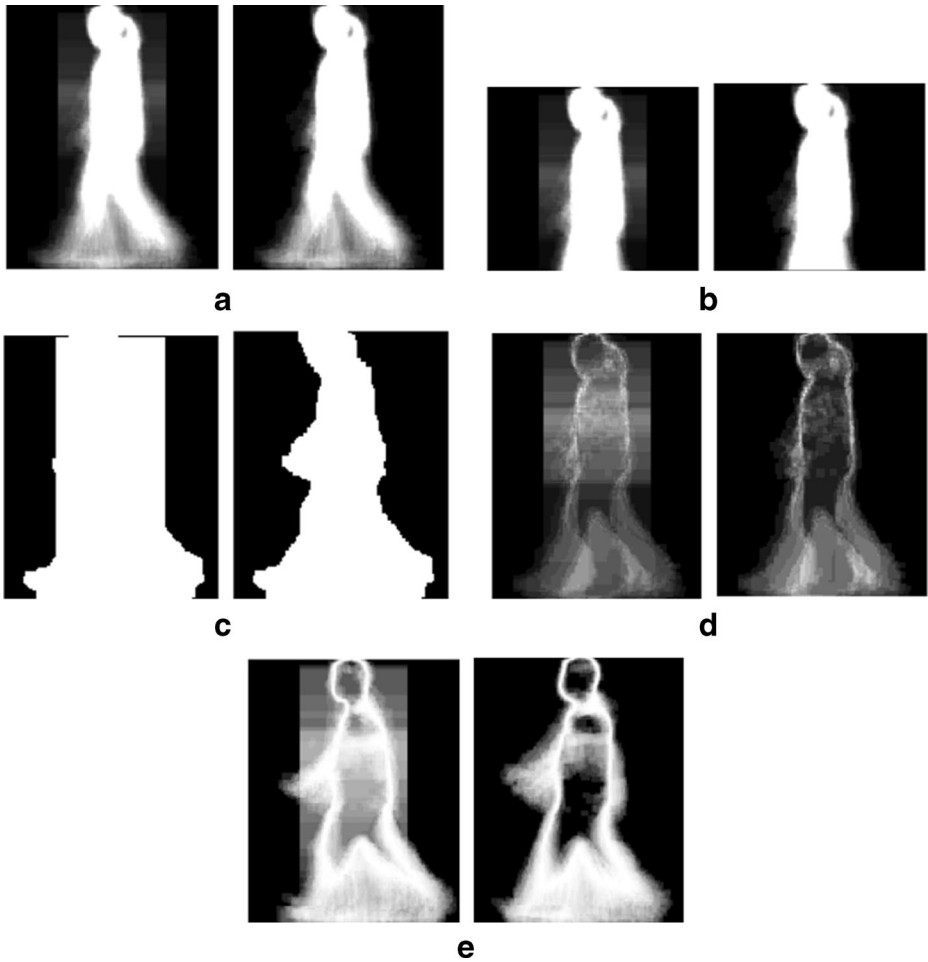


Fig. 13 Silhouette-based representation images using the corrupted dataset (noise rate: 25 %) **a** GEI, **b** HTI, **c** MSI, **d** AEI, **e** GEnI, (*left*: conventional method, *right*: proposed method)

and

$$\begin{aligned}
 E(w^2f^2) &= u^2\alpha(1-p)(1-q) + (1-u)^2(1-\alpha)(1-p)(1-q) \\
 &\quad + u^2(1-\beta)p(1-q) + (1-u)^2\beta p(1-q) \\
 &= \{u^2\alpha + (1-u)^2(1-\alpha)\}(1-p)(1-q) + \{u^2(1-\beta) + (1-u)^2\beta\}p(1-q).
 \end{aligned}
 \tag{30}$$

Finally, the NSRs of the conventional and proposed methods are compared. Let us assume that $q=0.5$ and $u=0.9$.

- 1) Let $\alpha=\beta=0.9$ and let us compare the NSRs while varying the probability p that the silhouette image is contaminated. In Fig. 10, two NSRs are compared. It is worth noting that when the probability of noise p is less than around 0.75, the proposed method exhibits lower NSR than that of the GEI. This figure explains why the proposed method outperforms the GEI in terms of recognition accuracy.

Table 7 Comparison of performance: GEI

Noise rate (%)	GEI (%)	PSVM+GEI (%)	IMP (%)
0	98.23	98.23	0
10	97.64	97.35	-0.29
15	96.75	97.64	0.89
20	96.46	97.94	1.48
25	93.51	97.24	3.73
30	93.21	97.64	4.43
35	91.44	97.64	6.20
40	90.56	95.58	5.02
Average	94.73	97.41	2.68

2) Let $\alpha=0.9$ and $p=0.7$ and let us compare the NSRs while varying the TRR (true rejection rate) β of the SVM. Two NSRs are compared in Fig. 11.

As expected, when TRR β is higher than around 0.83, noisy silhouettes are reliably rejected and the proposed method demonstrates better NSR than the plain GEI. Figures 10 and 11 explain why the proposed method outperforms the GEI in terms of recognition accuracy.

4 Experimental results

4.1 CASIA database

In this section, we apply the proposed method to the CASIA database [18] and show a significant performance improvement in a noisy environment. This database is widely used to benchmark algorithms in gait recognition, and it is also known as the NLPR gait database. The database includes 20 subjects, each of which has four sequences: two sequences for one walking direction, and two for the reverse walking direction.

For the sake of simplicity, the nearest neighbor (NN) [7] is used as the classifier in our experiment. We use one sequence as a training set for PSVM, two sequences as a training set for NN, and the other sequence as a testing set to evaluate the performance of the proposed method. The proposed weighting scheme is combined with previous silhouette-based

Table 8 Comparison of performance: HTI

Noise rate (%)	HTI (%)	PSVM+HTI (%)	IMP (%)
0	98.23	98.23	0
10	97.64	97.35	-0.29
15	96.76	97.64	0.88
20	96.46	97.94	1.48
25	93.51	97.94	4.43
30	93.22	97.64	4.42
35	91.45	97.64	6.19
40	90.56	95.58	5.02
Average	94.73	97.50	2.84

Table 9 Comparison of performance: MSI

Noise rate (%)	MSI (%)	PSVM+MSI (%)	IMP (%)
0	88.20	88.50	0.30
10	13.27	88.79	75.52
15	13.27	87.02	73.75
20	15.34	87.31	71.97
25	23.89	86.72	62.83
30	17.99	84.07	66.08
35	18.58	83.18	64.60
40	22.42	82.59	60.17
Average	26.62	86.02	59.40

approaches, and their correct classification rates (CCR) are compared with one other and the previous methods. The simulation is repeated while the probability of noise in the gait silhouette database is changed from 0 to 0.4. Three-fold validation is used, and the experimental results of each silhouette-based method are shown in Tables 2, 3, 4, 5 and 6. In the tables, IMP denotes the improved rate and is defined as

$$IMP = C_p - C_s, \tag{31}$$

where C_s denotes the CCR from only silhouette-based gait recognition in a noisy environment, while C_p represents the CCR when the proposed method is applied to silhouette-based gait recognition in a noisy environment.

Figure 12 shows the improvement rate according to noise rate for each silhouette-based representation image. From Fig. 12, we can observe that our proposed method degrades the performances of GEI by -1.66% and HTI, GENI by -1.67% at a noise rate of 0% . However, all silhouette-based methods for the general case having noisy silhouette images are improved using the proposed method. As mentioned above, the representation images are critical parts in gait recognition. The proposed method can exclude severe noisy images and decrease the influence of noisy images when constructing the gait representation image. This is the reason for such high performance. Figure 13 compares several silhouette-based representations with and without the proposed method in the corrupted CASIA database, where the probability of noise is 0.25.

Table 10 Comparison of performance: AEI

Noise rate (%)	AEI (%)	PSVM+AEI (%)	IMP (%)
0	98.82	98.82	0
10	55.16	98.53	43.37
15	46.31	99.12	52.81
20	29.20	97.44	78.24
25	24.78	97.05	72.27
30	15.63	95.58	79.95
35	15.34	95.28	79.94
40	12.68	92.33	79.65
Average	37.24	96.77	60.78

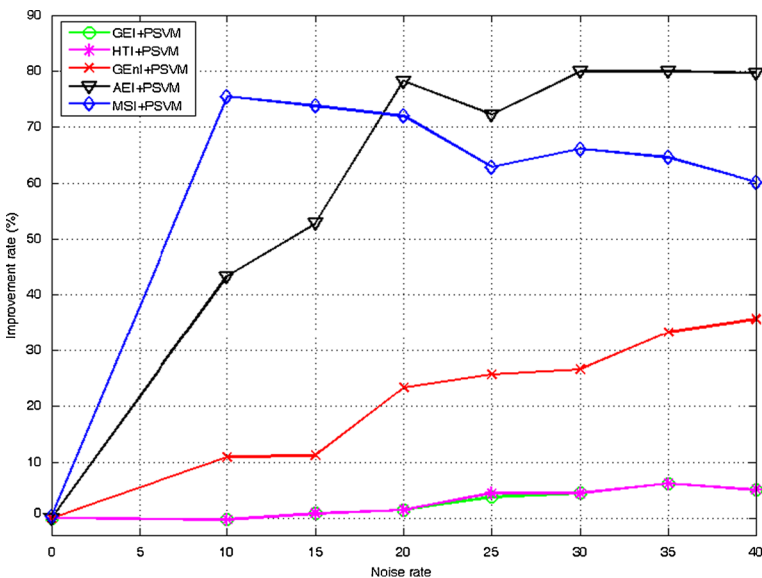
Table 11 Comparison of performance: GEnI

Noise rate (%)	GEnI (%)	PSVM+GEnI (%)	IMP (%)
0	97.05	97.05	0
10	86.14	97.05	10.91
15	83.19	94.35	11.16
20	74.04	97.35	23.31
25	71.39	97.05	25.66
30	70.50	97.05	26.55
35	63.13	96.46	33.33
40	59.29	94.99	35.70
Average	75.59	96.42	20.83

4.2 SOTON database

To make a generalized statement of our proposed method, we employ the larger SOTON database [15] for the second example. The SOTON database was created by Shutler et al. at the University of Southampton. It consists of more than 100 subjects. Similar to the experiment with the CASIA database, we use one sequence as a training set for the image weighting, two sequences as training for NN, and the other sequence as the testing set. Tables 7, 8, 9, 10 and 11 show the experimental results of each silhouette-based representation image.

Figure 14 shows the improvement rate according to noise rate. It can be observed from Tables 6, 7, 8, 9, 10 and 11 that the proposed method degrades the performances of GEI and HTI by 0.29 % at a noise rate of 10 %. Except for these two cases, however, the proposed method shows a significant improvement in performance for the corrupted SOTON database compared to the various silhouette-based gait recognition methods.

**Fig. 14** Improvement rate according to noise rate: SOTON database

5 Conclusions

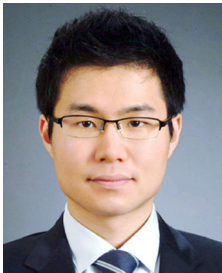
Gait representations for silhouette-based approaches are obviously important for recognition. However, corrupted and noisy silhouette images are deformed and negatively affect the performance of gait recognition systems. In this paper, we employ PSVM, which outputs an adequate weight for an individual silhouette image according to its clarity, to successfully improve the quality of gait representations. The proposed method is tested with the CASIA and SOTON databases, and shows a significant performance improvement from the general silhouette-based gait recognitions in a noisy environment, thereby increasing the reliability of the gait recognition system.

References

1. Bashir K, Xiang T, Gong S (2010) Gait recognition without subject cooperation. *Pattern Recognit Lett* 31 (13):2052–2060
2. Bazin A, Nixon M (2005) Gait verification using probabilistic methods, in *Proc. IEEE Workshop on Applications of Computer Vision*, vol. 1, Jan, pp 60–65
3. Ben Abdelkader C, Culter R, Nanda H, Davis L (2001) EigenGait: motion-based recognition of people using image self-similarity, in *Proc of International Conference on Audio- and Video-Based Biometric Person Authentication*, pp 284–294
4. Bobick A, Johnson A (2001) Gait recognition using static activity-specific parameters. *Proc IEEE Computer Vision and Pattern Recognition* 1:423–430
5. Bogdanovaa A, Nikolikb D, Curfsc L (2004) Probabilistic SVM outputs for pattern recognition using analytical geometry. *Neurocomputing* 62:293–303
6. Boulgouris NV, Chi Z (2007) Gait recognition using Radon transform and linear discriminant analysis. *IEEE Trans Image Process* 16(3):731–740
7. Cover TM, Hart PE (1967) Nearest neighbor pattern classification. *IEEE Trans Information Theory* 13 (1):21–27
8. Han J, Bhanu B (2006) Individual recognition using gait energy image. *IEEE Trans Pattern Analysis Machine Intelligence* 28:316–322
9. Lam T, Lee R (2006) A new representation for human gait recognition: Motion Silhouettes Image (MSI), in *Proc. International Conference on Biometrics*, pp 612–618
10. Lee H, Hong S, Kim E (2009) An efficient gait recognition with backpack removal, *EURASIP Journal on Advances in Signal Processing*, ID. 384384, vol. 2009
11. Lee H, Hong S, Nizami IF, Kim E (2009) A noise robust gait representation: motion energy image. *Int J Control Autom Syst* 7(4):638–643
12. Li X, Maybank S, Yan S, Tao D, Xu D (2008) Gait components and their application to gender recognition. *IEEE Trans Systems, Man and Cybernetics* 38(2):145–155
13. Murray M, Drought A, Kory R (1964) Walking patterns of normal men. *J Bone Joint Surg Br* 46-A (2):335–360
14. Platt J (1999) Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. *Advances in Large Margin Classifiers* 10:61–74
15. Shutler JD, Grant MG, Nixon MS, Cater JN (2002) On a large sequence-based human gait database. *Proc. 4th International Conference on Recent Advances in Soft Computing*, pp 66–71
16. Tan D, Huang K, Yu S, Tan T (2006) Efficient night gait recognition based on template matching. *Proc International Conference on Pattern Recognition* 3:1000–1003
17. Vapnik V (1988) *Statistical learning theory*. Wiley, NY
18. Wang L, Tan T, Ning H, Hu W (2003) Silhouette analysis-based gait recognition for human identification. *IEEE Trans Pattern Analysis Machine Intelligence* 25:1505–1518
19. Zhang E, Zhao Y, Xiong W (2010) Active energy image plus 2DLPP for gait recognition. *Signal Process* 90(7):2295–2302
20. Zhou X, Bhanu B (2007) Integrating face and gait for human recognition at a distance in video. *IEEE Trans Systems, Man and Cybernetics* 37(5):1119–1137



Heesung Lee received the BS, MS, and PhD degrees in electrical and electronic engineering from Yonsei University, Seoul, Korea, in 2003, 2005, 2010, respectively. He is currently a senior engineer of Image Computing Research Team (ICRT) in Samsung S1 Co., Ltd. His current research interests include computational intelligence, pattern recognition, biometrics, and neural networks.



Jeonghyun Baek received the BS degree in electrical and electronic engineering from Yonsei University, Seoul, Korea, in 2011. He is a graduate student of the combined master's and doctoral degree programs at Yonsei University. He has studied machine learning, biometrics and optimization.



Euntai Kim received the BS (with top honors), MS, and PhD degrees in electronic engineering from Yonsei University, Seoul, Korea, in 1992, 1994, and 1999, respectively. From 1999 to 2002, he was a full-time lecturer with the Department of Control and Instrumentation Engineering at Hankyong National University, Gyeonggi-do, Korea. Since 2002, he has been with the School of Electrical and Electronic Engineering at Yonsei University, where he is currently a professor. He was a visiting scholar with the University of Alberta, Edmonton, Canada, in 2003, and a visiting researcher with the Berkeley Initiative in Soft Computing (BISC), UC Berkeley, USA, in 2008. His current research interests include computational intelligence and machine learning and their application to intelligent service robots, unmanned vehicles, home networks, biometrics, and evolvable hardware.