Gait Energy Response Function for Clothing-invariant Gait Recognition

Xiang Li[1,2], Yasushi Makihara[2], Chi Xu[1,2], Daigo Muramatsu[2], Yasushi Yagi[2], and Mingwu Ren[1]

[1] School of Computer Science and Engineering,
Nanjing University of Science and Technology, Nanjing, China
{lixiangmzlx, xuchisherry}@gmail.com, renmingwu@mail.njust.edu.cn
[2] Institute of Scientific and Industrial Research, Osaka University, Osaka, Japan
{makihara, muramatsu, yagi}@am.sanken.osaka-u.ac.jp

Abstract. This paper describes a method of clothing-invariant gait recognition by modifying intensity response function of a silhouette-based gait feature. While a silhouette-based representation such as gait energy image (GEI) has been popular in gait recognition community due to its simple yet effective property, it is also well known that such a representation is susceptible to clothes variations since it significantly changes silhouettes (e.g., down jacket, long skirt). We therefore propose a gait energy response function (GERF) which transforms an original gait energy into another one in a nonlinear way, which increases discrimination capability under clothes variation. More specifically, the GERF is represented as a vector of components of a lookup table from an original gait energy to another one and its optimization process is formulated as a generalized eigenvalue problem considering discrimination capability as well as regularization on the GERF. In addition, we apply Gabor filters to the GEI transformed by the GERF and further apply a spatial metric learning method for better performance. In experiments, the OU-ISIR Treadmill dataset B with the largest clothing variation was used to measure the performance both in verification and identification scenarios. The experimental results show that the proposed method achieved state-of-the-art performance in verification scenarios and competitive performance in identification scenarios.

1 Introduction

Gait recognition [1] is one of behavioral biometrics and advantageous over the other biometrics (*e.g.*, face, iris, finger vein) because it can be used even at a distance from a camera since it does not require a high image resolution. In addition, gait is usually captured as an unconscious behavior, and hence it does not require subject cooperation in general. Due to these characteristics, it can be applied to many areas (*e.g.*, surveillance, forensics, criminal investigation[2–4]).

Gait recognition approaches can be divided into two main groups: model-based approaches [5–10] and model-free (appearance-based) approaches [11–16]. The model-based approaches have greater invariant properties and are better at

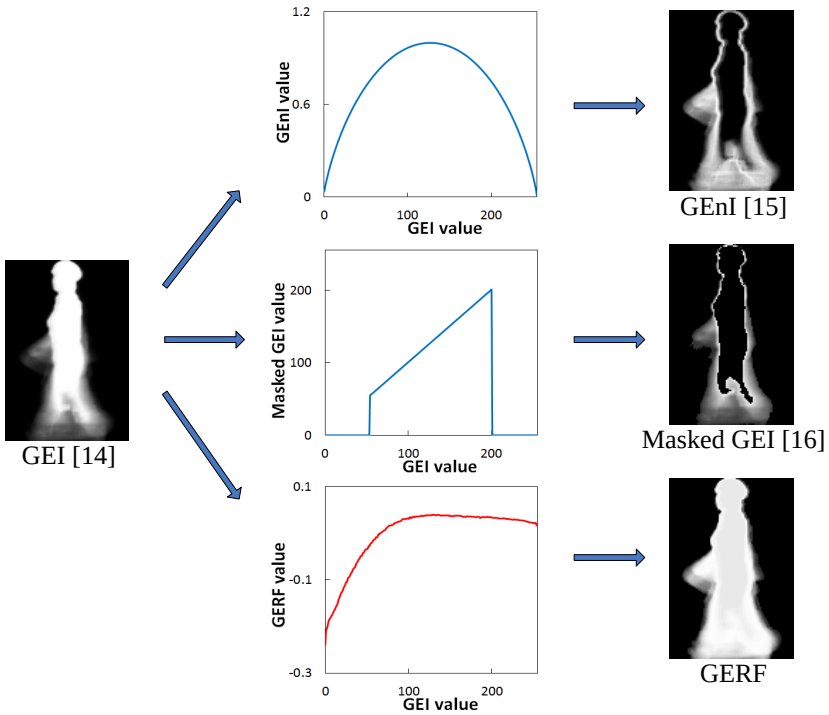


Fig. 1. Concept of the proposed GERF compared with existing intensity transformation-based approaches.

handling occlusion, noise, scale, and rotation. These approaches, however, require higher resolution images for model fitting and have relatively high computational cost.

The appearance-based approaches directly use input or silhouette images in a holistic way to extract gait features without modeling, and hence they generally work well even for relatively low-resolution images. In particular, silhouette-based representations such as gait energy image (GEI) [14], frequency-domain feature [17], chrono-gait image [18], Gabor GEI [19], are dominant in gait recognition community due to its simple yet effective property. The appearance-based approaches, however, often suffer from many covariates (*e.g.*, clothing, view, speed, and carrying status) since the appearance-based features of individuals are significantly affected by them, which causes a rapid decline in recognition rate. Among these covariates, clothing is one of the most challenge covariates [20, 11, 21–23].

There are two major categories to address the clothing-invariant problem in appearance based approaches: (1) spatial metric learning-based approaches and (2) intensity transformation-based approaches. In addition, the spatial metric learning-approaches further fall into two families: whole-based approaches [14–16, 24] and part-based approaches [25, 21–23, 26].

While the whole-based approaches usually apply discriminative projections to the holistic appearance-based features (*e.g.*, linear discriminant analysis (LDA) [14, 27, 18] in conjunction with principal component analysis (PCA), discriminant analysis with tensor representation (DATER) [28, 29], random subspace method (RSM) [23, 26] to gain robustness to the clothes variation, the part-based approaches firstly divide the whole body into multiple body parts and then exploit the body parts which are not so much affected by the clothes variation by adaptively assigning weights for individual body parts [21] or finding the effective body parts [22], which mitigates the effect of clothes variation.

Whereas the above mentioned approaches mainly focus on the metric learning aspect, the intensity transformation-based approaches more focus on gait representation aspect. Since the clothes variation affects more on static parts (*e.g.*, torso and limb shapes) than on dynamic parts (*e.g.*, leg and arm motion), gait entropy image (GEnI) [15] extracts the dynamic parts from GEI by computing its Shannon entropy, where gait energy for each pixel is regarded as a foreground probability. For example, the pixels with large and small gray values (*e.g.*, 255 and 0) in GEI become small in GEnI, while the pixels with middle values (*e.g.*, 127) become large (see Fig. 1, the top row). The static parts (*i.e.*, complete foreground and background), however, still have discrimination capability to some extent even under clothes variation, and hence GEnI discards such useful information. Moreover, GEnI treats two different gait energies which are symmetric with just the middle value (*i.e.*, 127.5), as the same value, and hence it loses discrimination capability (*e.g.*, gait energies $64(= 127.5 - 63.5)$ and $191(= 127.5 + 63.5)$ returns the same value in GEnI).

In order to solve the latter problem, masked GEI [16] is proposed, where gait energies whose corresponding gait entropy is smaller than a certain threshold (*i.e.*, more static parts) are masked out and are set to zero, while the other gait energies are kept as their original values (see Fig. 1, the middle row). Masked GEI is, however, dependent on choice of the threshold to mask out and also still discards useful static information.

Because both GEnI and masked GEI are generated from GEI, we can regard this as a sort of gait energy transformation process via a gait energy response function (GERF). While both GEnI and masked GEI employ handcrafted GERFs to focus on the dynamic parts, we may generate more discriminative features under clothes variation by designing the GERF in a more general and data-driven way.

We therefore propose to introduce the GERF to transform GEI into more discriminative feature and show its effectiveness on gait recognition under clothes variation. The contributions of this work are three-fold.

1. A data-driven approach to intensity transformation

While the existing intensity transformation-based methods such as GEnI and masked GEI are designed in a handcrafted way, the proposed method learn the GERF in a data driven way. More specifically, we train the GERF so as to maximizing the discrimination capability using the training set including clothes

variation. This enables us to realize a good tradeoff between static and dynamic parts, unlike the existing method discard the static parts.

2. A closed-form solution to optimize the GERF

We train the GERF so as to maximize dissimilarity for different subjects' pairs while to minimize dissimilarities for the same subjects' pairs, and consequently formulate its optimization process as a generalized eigenvalue problem. We therefore obtain an analytic solution in a closed form without any iterations and hence avoids troublesome convergence problems which is inseparable from a nonlinear optimization framework.

3. State-of-the-art performance on clothing-invariant gait recognition

We achieved the state-of-the-art performance on clothing-invariant gait recognition using publicly available gait database containing the largest clothes variations up to 32 types, in conjunction with Gabor filtering and spatial metric learning.

2 Gait Recognition using GERF

2.1 Representation of GERF

In this section, we introduce the GERF for the most widely used gait feature, *i.e.*, GEI. For this purpose, we briefly describe the GEI at first. The GEI [14] a.k.a. averaged silhouette [30] is a size-normalized and registered silhouette averaged over one gait period (cycle) T defined as

$$I(x,y) = \frac{1}{T}\sum_{t=1}^{T} B(x,y,t), \tag{1}$$

where $B(x,y,t)$ is a size-normalized and registered binary silhouette value (0 and I_{max}[1] for background and foreground, respectively) at the position (x,y) at the n-th frame, and $I(x,y)$ is a gait energy (averaged silhouette) at the position (x,y). While the domain of the gait energy is real number, *i.e.*, $I(x,y) \in \mathbb{R}$, we approximate it as an integer number, *i.e.*, $I(x,y) \in \{0,1,\ldots,I_{max}\}$ for simplicity.

A transformation from an original gait energy $I(x,y)$ to another one $I'(x,y)$ is then defined via the GERF f as

$$I'(x,y) = f(I(x,y))\ \forall(x,y). \tag{2}$$

Since the original gait energy takes one of $(I_{max}+1)$ integer numbers from 0 to I_{max}, the GERF is represented as a lookup table $\boldsymbol{f} = [f_0,...,f_{I_{max}}]^T \in \mathbb{R}^{I_{max}+1}$, where f_i represent a transformed gait energy from an original gait energy i.

Next, we consider a dissimilarity measure between a pair of GEIs transformed from original GEIs I_1 and I_2. We simply adopt Euclidean distance between them and define its squared distance $d^2_{I_1,I_2}$ and further formulate it in a quadratic form of $\boldsymbol{f}$ as

$$d^2_{I_1,I_2} = \sum_{x,y}(f_{I_1(x,y)} - f_{I_2(x,y)})^2 = \boldsymbol{f}^T A_{I_1,I_2}\boldsymbol{f}\ , \tag{3}$$

[1] I_{max} is usually 255 for 8-bit depth.

where $A_{I_1,I_2} \in \mathbb{R}^{(I_{max}+1)\times(I_{max}+1)}$ is a coefficient matrix for quadratic-form representation and its (l,m) component is obtained using the Kronecker delta $\delta_{i,j}$ as

$$\begin{aligned}(A_{I_1,I_2})_{l,m} = \sum_{x,y} (&\delta_{I_1(x,y),l}\delta_{I_1(x,y),m} + \delta_{I_2(x,y),l}\delta_{I_2(x,y),m} \\ &- \delta_{I_1(x,y),l}\delta_{I_2(x,y),m} - \delta_{I_2(x,y),l}\delta_{I_1(x,y),m}) \ .\end{aligned} \tag{4}$$

2.2 Training of GERF

In order to make the transformed GEI discriminative under clothes variation, we optimize the GERF using a training set including the clothes variation. The whole training set is composed of two subsets $\mathcal{S}$ and $\mathcal{D}$, where the subset $\mathcal{S}$ is a set of GEI pairs of the same subject, while the subset $\mathcal{D}$ is a set of GEI pairs of different subjects. For better discrimination, it is preferable to make it larger the sum of squared distances $D_{\mathcal{S}}$ for the same subject pairs $\mathcal{S}$ while make it smaller the sum of squared distances $D_{\mathcal{D}}$ for the different subject pairs. Here, $D_{\mathcal{S}}$ and $D_{\mathcal{D}}$ are calculated as

$$\begin{aligned} D_{\mathcal{S}} &= \sum_{(I_1,I_2)\in\mathcal{S}} d^2_{I_1,I_2} &= \boldsymbol{f}^T S_{\mathcal{S}} \boldsymbol{f} \\ D_{\mathcal{D}} &= \sum_{(I_1,I_2)\in\mathcal{D}} d^2_{I_1,l_2} &= \boldsymbol{f}^T S_{\mathcal{D}} \boldsymbol{f} \ , \end{aligned} \tag{5}$$

where $S_{\mathcal{S}} \in \mathbb{R}^{(I_{max}+1)\times(I_{max}+1)}$ and $S_{\mathcal{D}} \in \mathbb{R}^{(I_{max}+1)\times(I_{max}+1)}$ are computed as $S_{\mathcal{S}} = \sum_{(I_1,I_2)\in\mathcal{S}} A_{I_1,I_2}$ and $S_{\mathcal{D}} = \sum_{(I_1,I_2)\in\mathcal{D}} A_{I_1,I_2}$, respectively.

Moreover, in order to make the GERF smoother, we also introduce a regularizer D_R, which is defined as

$$\begin{aligned} D_R &= w_1 \sum_{i=1}^{I_{max}} (f_i - f_{i-1})^2 + w_2 \sum_{i=1}^{I_{max}-1} (f_{i+1} - 2f_i + f_{i-1})^2 \\ &= \boldsymbol{f}^T (w_1 S_{R_1} + w_2 S_{R_2}) \boldsymbol{f} \\ &= \boldsymbol{f}^T S_R \boldsymbol{f} \ , \end{aligned} \tag{6}$$

where w_1 and w_2 are weighting parameters for the first-order and second-order smoothness, and $S_{R_1} \in \mathbb{R}^{(I_{max}+1)\times(I_{max}+1)}$ and $S_{R_2} \in \mathbb{R}^{(I_{max}+1)\times(I_{max}+1)}$ are coefficients matrices for the first-order and the second-order smoothness, which

are defined as

$$S_{R_1} = \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -1 & 2 & -1 \\ 0 & \cdots & 0 & -1 & 0 \end{bmatrix}, \quad S_{R_2} = \begin{bmatrix} 1 & -2 & 0 & \cdots & \cdots & \cdots & 0 \\ -2 & 5 & -4 & \ddots & & & \vdots \\ 0 & -4 & 6 & -4 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & -4 & 6 & -4 & 0 \\ \vdots & & & \ddots & -4 & 5 & -2 \\ 0 & \cdots & \cdots & \cdots & 0 & -2 & 1 \end{bmatrix}. \quad (7)$$

Finally, the GERF is optimized so as to maximize the ratio between the sum of squared distances $D_{\mathcal{D}}$ for the different subject pairs and those $D_{\mathcal{S}}$ for the same subject pairs plus the regularizer D_R under an L_2 norm constraint on $\boldsymbol{f}$ as

$$\boldsymbol{f}^* = \arg\max_{\boldsymbol{f}} \frac{\boldsymbol{f}^T S_{\mathcal{D}} \boldsymbol{f}}{\boldsymbol{f}^T (S_{\mathcal{S}} + S_R) \boldsymbol{f}} \quad \text{s.t.} \quad \|\boldsymbol{f}\| = 1. \quad (8)$$

In an analogous fashion to well-known LDA formulation, we can formulate this optimization problem as the following generalized eigenvalue problem

$$S_{\mathcal{D}} \boldsymbol{f} = \lambda (S_{\mathcal{S}} + S_R) \boldsymbol{f} \quad \text{s.t.} \quad \|\boldsymbol{f}\| = 1, \quad (9)$$

where λ is an eigenvalue, and $\boldsymbol{f}$ is regarded as a corresponding eigenvector. We therefore analytically obtain the optimal GERF $\boldsymbol{f}^*$ in a closed-form solution by assigning the eigenvector corresponding to the largest eigenvalue.

2.3 Gabor Filtering

In order to further improve the performance, we introduce two sequential processes after obtaining a GEI transformed with the optimal GERF (call it GEI-GERF later), since the proposed GERF can be jointly used with other filtering and spatial metric learning techniques.

The first one is Gabor filtering, which has been successfully employed in gait recognition because of its effectiveness [19, 31]. In a similar way to [19], we will briefly describe Gabor functions. The Gabor function are defined by multiplying an elliptical Gaussian envelope function with a complex oscillation, defined as

$$\psi_{s,d}(\boldsymbol{p}) = \frac{|k_{s,d}|^2}{\delta^2} \exp\left\{ -\frac{|k_{s,d}|^2 \|\boldsymbol{p}\|^2}{2\delta^2} \right\} \left[\exp\left(\boldsymbol{k}(j k_{s,d}) \cdot \boldsymbol{p}\right) - \exp\left(-\frac{\delta^2}{2}\right) \right], \quad (10)$$

where a vector $\boldsymbol{p} = [x, y]^T$ is the spatial location in Gabor kernel window, a complex number $k_{s,d} = \theta_s e^{j\phi_d}$ determines the scale ($s = 0, \ldots, N_s - 1$) and orientation ($d = 0, \ldots, N_d - 1$) of the Gabor kernel function, j is an imaginary unit, and $\boldsymbol{k}(\cdot)$ is a function to convert a complex number to a two-dimensional real vector. Specifically, $\theta_s = 2^{-s}(\pi/2)$ is the scale item, and $\phi_d = \pi d / N_d$ is the

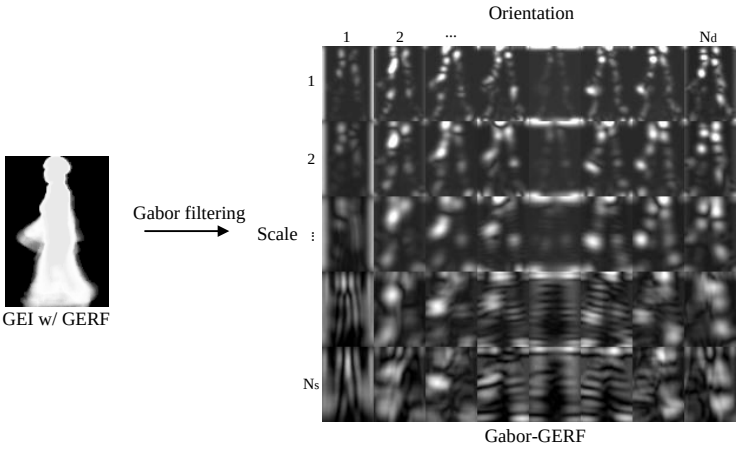


Fig. 2. An example of Gabor-GERF. The rows show different scales and the columns show different orientations. In this figure, $N_s = 5$ and $N_d = 8$.

direction item. Since we have N_s scales and N_d orientations, we then get a total of $N_s N_d$ Gabor functions.

Given a GEI-GERF whose width and height are W and H, respectively, it is convolved with all the Gabor kernel functions and further down-sampled into half size (*i.e.*, W/2 by H/2) for computational efficiency in the same way as [31]. We then concatenate all $N_s N_d$ downsampled Gabor-filtered images into a single image, where scale and orientation components are concatenated along row and column directions, respectively. As a result, we obtain a concatenated image whose width and height are $W' = N_d W/2$ and $H' = N_s H/2$, respectively. In this paper, we call it Gabor-GERF later and show an example of the Gabor-GERF in Fig. 2.

2.4 Spatial Metric Learning

Once we obtain the Gabor-GERF, we introduce a spatial metric learning, *i.e.*, two-dimensional LDA (2DLDA) in conjunction with preceding dimension reduction by two-dimensional PCA (2DPCA) [32]. Unlike PCA and LDA handle one-dimensional vector unfolded from an image matrix, a covariance matrix for 2DPCA and within-class and between-class matrices for 2DLDA are directly constructed using the original image matrices and result in smaller size of covariance/within-class/between-class matrices, which ensures lower time complexity and less singularity than PCA and LDA, respectively. We therefore adopt a combination of 2DPCA and 2DLDA (call it 2DPCA + 2DLDA later) for spatial metric learning.

Suppose that we have M samples of Gabor-GERFs $\{X_i \in \mathbb{R}^{H' \times W'}\}(i = 1, \ldots, M)$ in the training set, and its mean is denoted by $\bar{X}$. The covariance

matrix $S_T \in \mathbb{R}^{W' \times H'}$ for 2DPCA (projection for column direction) is

$$S_T = \frac{1}{M} \sum_{i=1}^{M} (X_i - \bar{X})^T (X_i - \bar{X}). \tag{11}$$

We then obtain a projection matrix $P \in \mathbb{R}^{W' \times W''}$ composed of a set of W'' eigenvectors of the covariance matrix S_T. In this paper, we set the reduced dimension W'' so as to keep more than 99% variance (*i.e.*, less than 1% information loss).

After applying 2DPCA to the Gabor-GERF and obtaining projected matrices $Y_i = (X_i - \bar{X})P$, $(i = 1, \ldots, M)$, we subsequently calculate the within-class scatter matrix $S_w \in \mathbb{R}^{H' \times H'}$ and between-class scatter matrix $S_b \in \mathbb{R}^{H' \times H'}$ as

$$S_w = \sum_{i=1}^{M} (Y_i - \bar{Y}_{l_i})(Y_i - \bar{Y}_{l_i})^T \tag{12}$$

$$S_b = \sum_{c=1}^{N_c} M_c (\bar{Y}_c - \bar{Y})(\bar{Y}_c - \bar{Y})^T, \tag{13}$$

where l_i is the class label (subject ID) for the i-th sample, $\bar{Y}_c$ is a mean for the c-th class, $\bar{Y}$ is a total mean, N_c is the number of classes, and M_c is the number of samples for the c-th class. Finally, the optimal projection $\boldsymbol{w}^*$ for 2DLDA is obtained as

$$\boldsymbol{w}^* = \arg\max_{\boldsymbol{w}} \frac{\boldsymbol{w}^T S_b \boldsymbol{w}}{\boldsymbol{w}^T S_w \boldsymbol{w}}. \tag{14}$$

We then reformulate Eq. (14) as a generalized eigenvalue problem and obtain a projection matrix $R \in \mathbb{R}^{H' \times H''}$ composed of a set of eigenvectors corresponding to the H'' largest eigenvalues.

Once we obtain the projection matrices P and R, we project the Gabor-GERF X_i into dimension reduced matrix Z_i in the 2DPCA + 2DLDA space as

$$Z_i = R^T (X_i - \bar{X}) P. \tag{15}$$

Finally, matching for a pair of Gabor-GERFs is done based on Euclidean distance in the 2DPCA + 2DLDA space.

3 Experiments

3.1 Data Set

We used the OU-ISIR Gait Database, Treadmill Dataset B [33] for our experiments, since it has the largest clothing variations. It includes 68 subjects with at most 32 combinations of different clothing. The whole dataset is divided into three subsets: training set, gallery set, and probe set. In the training set, there are 446 sequences of 20 subjects with the range of 15 to 28 different combinations of clothing. The gallery set and probe set form the testing set composed of 48 subjects, which were disjoint from the 20 subjects in the training set. The gallery contains only standard clothing type (*i.e.*, regular pant and full shirt), while the probe set includes 856 sequences of other remaining clothing types.

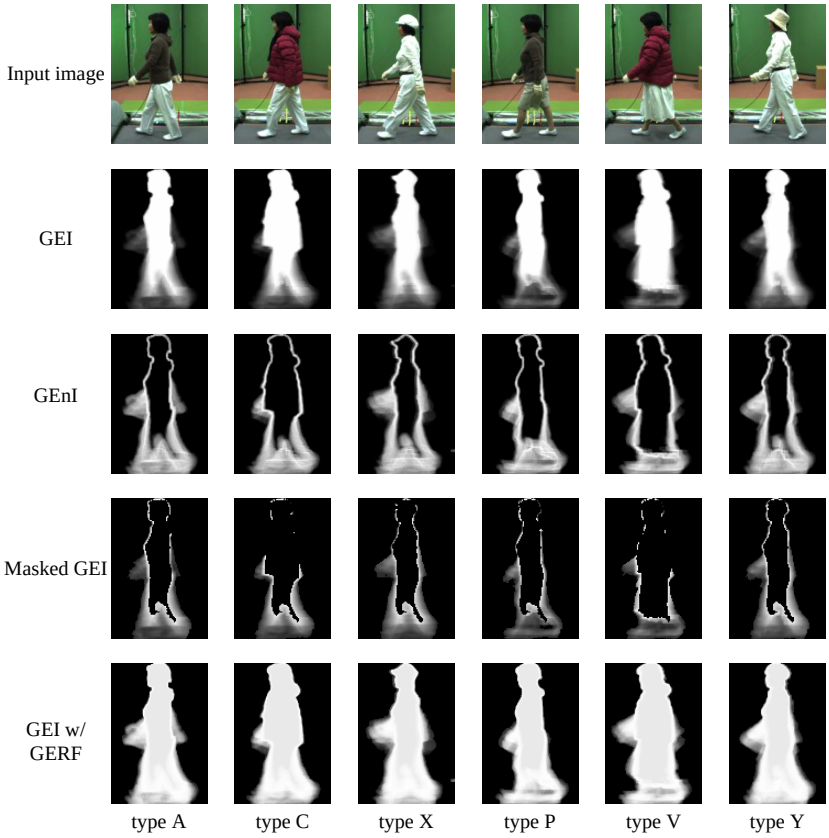


Fig. 3. Examples of extracted features for intensity transformation-based methods.

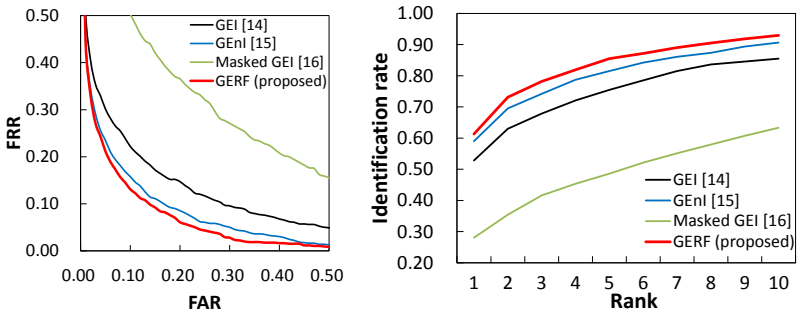
3.2 Parameter Setting

There are two main hyper parameters of our GERF in the training stage: weighting parameters of the regularizer w_1 and w_2. We experimentally set $w_1 = w_2 = 5000$. About parameters in Gabor filtering, we set Gabor kernel window size to 41×41 and set parameter δ (in Eq. (10)) to 2π. The number of scales N_s and orientations N_d are set to 5 and 8, respectively. Since the silhouette image resolution provided in the database is 128×88, and hence the resolution of the Gabor-GERF is 320×352.

3.3 Comparison with Intensity Transformation-based Methods

To investigate the effectiveness of the proposed GERF module, we firstly conducted comparison experiments with a family of intensity transformed-based methods, *i.e.*, GEnI and Masked GEI as well as GEI as a baseline.

We show examples for the four gait features, *i.e.*, GEI, GEnI, Masked GEI, and GEI w/ GERF, as well as cropped original images in Fig. 3. Note that



(a) ROC curves (b) CMC curves

Fig. 4. ROC and CMC curves for intensity transformation-based methods.

Table 1. EER [%] and rank-1 identification rate (denoted as Rank-1) [%] for intensity transformation-based methods. Bold and Italic bold fonts indicate the best and the second best, respectively, which is consistent throughout this paper.

Method	EER	Rank-1
GEI [14]	16.12	52.80
GEnI [15]	***12.81***	***59.00***
Masked GEI [16]	28.15	28.04
GERF (proposed)	**11.57**	**61.33**

the trained GERF is depicted as a red curve at the bottom row of Fig. 1. The profile of the GERF for smaller gait energy (*e.g.*, gait energy from 0 to 127) is similar to the profile of GEnI, which suggests to emphasize the difference from background to middle-level gray value. On the other hand, the profile of the GERF for larger gait energy is approximately flat and hence the complete background and foreground is still differentiated, unlike GEnI or Masked GEI confuse it. In this way, the proposed GERF can highlight differences in dynamic parts on one hand, and it keeps static information on the other hand.

In addition, performances in verification (one-to-one matching) and identification (one-to-many matching) scenarios are evaluated. In verification scenarios, we employ an receiver operating characteristics (ROC) curve which indicates the tradeoff between the false rejection rate (FRR) of the same subject and the false acceptance rate (FAR) of different subjects when an acceptance threshold changes. Moreover, an equal error rate (EER) of FAR and FRR is also evaluated. In identification scenarios, we employ cumulative matching characteristics (CMC) curve which shows the rates that the true subjects are included within each of rank.

The ROC curves in Fig. 4(a) show that the proposed GERF outperforms other features. In Table 1, EER for the proposed GERF method is the lowest 11.57%, which represents the best verification performance. The CMC curves

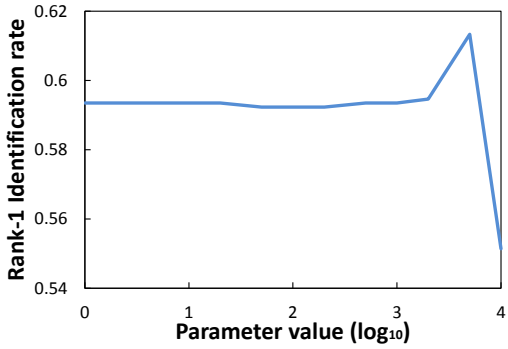
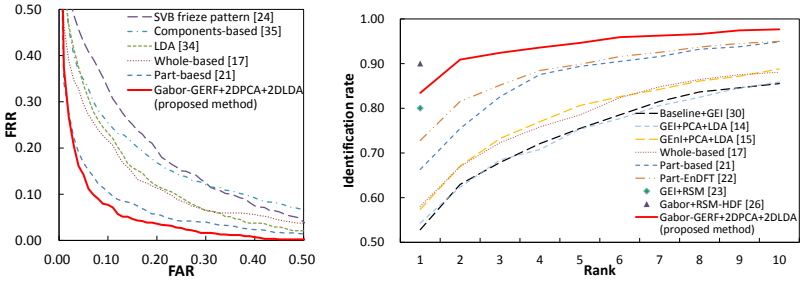


Fig. 5. Sensitivity analysis of the hyper parameters on rank-1 identification rates. The horizontal axis is shown by log-scale.



(a) ROC curves (b) CMC curves

Fig. 6. ROC and CMC curves compared with the state-of-the-arts.

in Fig. 4(b) also show that the proposed GERF yielded the best performance among the four features. In Table 1, rank-1 identification rate is the highest 61.33%, which indicates the best identification performance.

As for reference, we have investigated the sensitivity of the hyper parameters w_1 and w_2 on rank-1 identification rate of the proposed GERF. For simplicity, we set the same parameters both for w_1 and w_2, and changed it in the range from 1 to 10,000 as shown in Fig. 5. As a result, rank-1 identification rate is not so much degraded for smaller range of the hyper parameters (less than 5,000) and is still better than the second best method, *i.e.*, GEnI with 59.0% rank-1 identification rate. It is therefore turned out that the proposed method is not so insensitive to the setting of parameter w_1 and w_2 as long as we use less than 5,000.

3.4 Comparison with the State-of-the-arts Methods

In verification scenarios, we compare the proposed method with the frequency-domain feature (denoted as whole-based) [17], part-based method with adaptive weight control (denoted as part-based) [21], GEI with LDA (denoted as

Table 2. EERs [%] and rank-1 identification rates [%] compared with the state-of-the-arts.

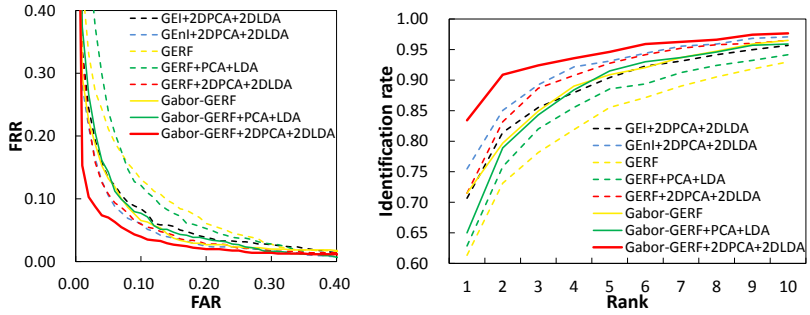
Method	EER	Rank-1
Baseline+GEI [30]	-	52.8
LDA [34]	15.48	-
SVB frieze pattern [24]	19.81	-
Components-based [35]	18.25	-
GEI+PCA+LDA [14]	-	54.3
GEnI+PCA+LDA [15]	-	57.4
Whole-based [17]	14.88	58.1
Part-based [21]	***10.26***	66.3
Part-EnDFT [22]	-	72.8
GEI+RSM [23]	Not applicable	80.4
Gabor+RSM-HDF [26]	Not applicable	**90.7**
Gabor-GERF+2DPCA+2DLDA (proposed method)	**6.19**	***83.4***

LDA) [34], SVB frieze pattern [24] and gait components-based method (denoted as components-based) [35] to confirm its effectiveness. The performance is evaluated by ROC curves in Fig. 6(a). As a result, the proposed method gets the state-of-the-art performance in contrast to the other methods.

In identification scenarios, we compare the proposed method with the averaged silhouette (denoted as baseline+GEI) [30], GEI+PCA+LDA [14], GEnI + PCA + LDA [15], whole-based [17], part-based [21], part-EnDFT [22], GEI + RSM [23], Gabor + RSM-HDF [26] to confirm its effectiveness. Note that this different list of benchmarks in identification scenario from that in verification scenario comes from the difference in the availabilities of reported results in each paper. The performance is evaluated by CMC curve s in Fig. 6(b). In addition, the Rank-1 identification rate is shown in Table 2. The proposed method gets the second best performance, lower than the Gabor+RSM-HDF [26]. However, we need to point out that the RSM framework cannot guarantee a stable accuracy because of its randomness. Moreover, the RSM framework is only applicable to identification scenarios since it relies on a framework of majority voting to all of the galleries. Considering these points, we can say that the proposed method is promising since it can be employed both in identification and verification scenarios, which indicates the widely application range of the proposed method.

3.5 Analysis of Individual Modules

In order to investigate the effectiveness of individual modules (GEI/GEnI/GERF, Gabor filtering and spatial metric learning methods), we compare totally eight methods: GEI+2DPCA+2DLDA, GEnI+2DPCA+2DLDA, GERF, GERF+PCA +LDA, GERF+2DPCA+2DLDA, Gabor-GERF, Gabor-GERF+PCA+LDA and Gabor-GERF+2DPCA+2DLDA. The ROC and CMC curves are reported in Figs. 7(a) and 7(b), while EER and rank-1 identification rate are reported in Table 3, respectively. If we exclude Gabor filtering and 2DPCA + 2DLDA from



(a) ROC curves (b) CMC curves

Fig. 7. ROC and CMC curves of GERF and Gabor-GERF w/ metric learning.

Table 3. EER [%] and rank-1 identification rate (denoted as Rank-1) [%] of GERF and Gabor-GERF w/ metric learning.

Method	EER	Rank-1
GEI+2DPCA+2DLDA	8.91	70.68
GEnI+2DPCA+2DLDA	***7.48***	***75.47***
GERF	11.57	61.33
GERF+PCA+LDA	10.98	62.85
GERF+2DPCA+2DLDA	7.94	71.50
Gabor-GERF	8.41	71.50
Gabor-GERF+PCA+LDA	8.53	65.00
Gabor-GERF+2DPCA+2DLDA	**6.19**	**83.41**

the full proposed method (Gabor-GERF+2DPCA+2DLDA), rank-1 identification rates drops by approximately 10%, and EER increases by approximately 2%, and hence we confirmed that Gabor filtering and spatial metric learning successfully enhance the proposed GERF framework.

3.6 Analysis of Difficulty Levels by Clothing Type

To evaluate the difficulty levels of clothes variation for the proposed method, we compute the rank-1 identification rates for all probe clothing types and list them in descending order as shown in Fig. 8. For the first 21 probe clothes types, the proposed method achieves over 85% rank-1 identification rate, which include more complex clothing type, such as clothing type B (*i.e.*, regular pants + down jacket), clothing type 6 (*i.e.*, regular pants + long coat + muffler). It is therefore validated that the proposed method effectively gains the discriminative features under a certain clothes types.

For the rest of the probe clothing types, the average rank-1 identification rate drops to approximately 67%. Specifically, the lowest two clothing types are type V (*i.e.*, skirt + down jacket) with 44% rank-1 identification rate and type

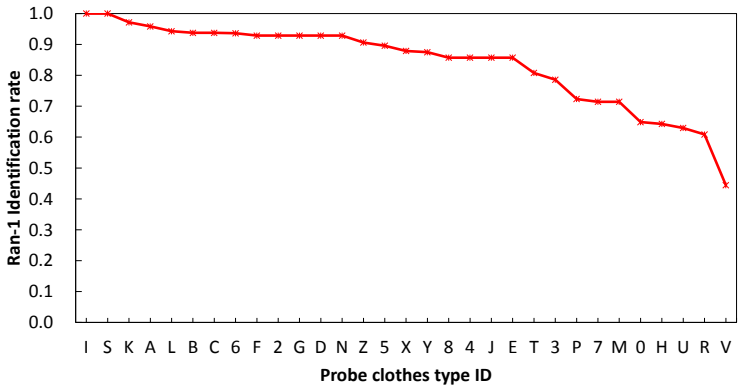


Fig. 8. Sorted clothing types according to recognition rate with the proposed method (Gabor-GERF+2DPCA+2DLDA).

R (*i.e.*, raincoat) with 61% rank-1 identification rate. In fact, the clothing type V has quite different appearance from the others as shown in Fig. 3, and even the proposed GERF suffers from large intra-subject variations. Use of a single common GERF for all the clothes type may cause this performance degradation, one of future research avenues is a clothes type-adaptive selection of a suitable GERF from multiple GERFs in future.

4 Conclusion

The paper described a data-driven framework to learn GERF for clothes-invariant gait recognition. The GERF transforms an original gait energy into another one so as to make it more discriminative under clothes variation. The GERF is represented as a look-up table vector and is optimized through efficient generalized eigenvalue problem, which enables us to obtain analytical solution in a closed form without any iterations. In addition, in order to boost the GERF performance, Gabor filtering and 2DPCA+2DLDA are employed. Through comprehensive experiments, the proposed method shows the state-of-the-art performance in verification scenarios and competitive performance in identification scenarios.

Since, we only use the eigenvector corresponding to the largest eigenvalue as the GERF, the use of multiple eigenvectors will be investigated in the future. Moreover, since we use a common GREF regardless clothes type and spatial positions, further performance improvement is expected by introducing adaptive selection of GREF in future.

Acknowledgement. This work was supported by JSPS Grants-in-Aid for Scientific Research (A) JP15H01693, the JST CREST "Behavior Understanding based on Intention-Gait Model" project and Nanjing University of Science and Technology.

References

1. Nixon, M.S., Tan, T.N., Chellappa, R.: Human Identification Based on Gait. Int. Series on Biometrics. Springer-Verlag (2005)
2. Bouchrika, I., Goffredo, M., Carter, J., Nixon, M.: On using gait in forensic biometrics. Journal of Forensic Sciences **56** (2011) 882–889
3. Iwama, H., Muramatsu, D., Makihara, Y., Yagi, Y.: Gait verification system for criminal investigation. IPSJ Transactions on Computer Vision and Applications **5** (2013) 163–175
4. Lynnerup, N., Larsen, P.: Gait as evidence. IET Biometrics **3** (2014) 47–54
5. Bobick, A., Johnson, A.: Gait recognition using static activity-specific parameters. In: Proc. of the 14th IEEE Conference on Computer Vision and Pattern Recognition. Volume 1. (2001) 423–430
6. Cunado, D., Nixon, M., Carter, J.: Automatic extraction and description of human gait models for recognition purposes. Computer Vision and Image Understanding **90** (2003) 1–41
7. Urtasun, R., Fua, P.: 3d tracking for gait characterization and recognition. In: Proc. of the 6th IEEE International Conference on Automatic Face and Gesture Recognition. (2004) 17–22
8. Wagg, D., Nixon, M.: On automated model-based extraction and analysis of gait. In: Proc. of the 6th IEEE Int. Conf. on Automatic Face and Gesture Recognition. (2004) 11–16
9. Yam, C., Nixon, M., Carter, J.: Automated person recognition by walking and running via model-based approaches. Pattern Recognition **37** (2004) 1057–1072
10. Zhao, G., Liu, G., Li, H., Pietikainen, M.: 3d gait recognition using multiple cameras. In: 7th International Conference on Automatic Face and Gesture Recognition (FGR06). (2006) 529–534
11. Sarkar, S., Phillips, J., Liu, Z., Vega, I., ther, P.G., Bowyer, K.: The humanid gait challenge problem: Data sets, performance, and analysis. IEEE Transactions of Pattern Analysis and Machine Intelligence **27** (2005) 162–177
12. Murase, H., Sakai, R.: Moving object recognition in eigenspace representation: Gait analysis and lip reading. Pattern Recognition Letters **17** (1996) 155–162
13. Wang, L., Tan, T., Ning, H., Hu, W.: Silhouette analysis-based gait recognition for human identification. Pattern Analysis and Machine Intelligence, IEEE Transactions on **25** (2003) 1505–1518
14. Han, J., Bhanu, B.: Individual recognition using gait energy image. IEEE Transactions on Pattern Analysis and Machine Intelligence **28** (2006) 316– 322
15. Bashir, K., Xiang, T., Gong, S.: Gait recognition using gait entropy image. In: Proc. of the 3rd Int. Conf. on Imaging for Crime Detection and Prevention. (2009) 1–6
16. Bashir, K., Xiang, T., Gong, S.: Gait recognition without subject cooperation. Pattern Recognition Letters **31** (2010) 2052–2060
17. Makihara, Y., Sagawa, R., Mukaigawa, Y., Echigo, T., Yagi, Y.: Gait recognition using a view transformation model in the frequency domain. In: Proc. of the 9th European Conference on Computer Vision, Graz, Austria (2006) 151–163
18. Wang, C., Zhang, J., Wang, L., Pu, J., Yuan, X.: Human identification using temporal information preserving gait template. IEEE Transactions on Pattern Analysis and Machine Intelligence **34** (2012) 2164 –2176
19. Tao, D., Li, X., Wu, X., Maybank, S.J.: General tensor discriminant analysis and gabor features for gait recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence **29** (2007) 1700–1715

20. Matovski, D., Nixon, M., Mahmoodi, S., Carter, J.: The effect of time on gait recognition performance. Information Forensics and Security, IEEE Trans. on **7** (2012) 543 –552
21. Hossain, M.A., Makihara, Y., Wang, J., Yagi, Y.: Clothing-invariant gait identification using part-based clothing categorization and adaptive weight control. Pattern Recognition **43** (2010) 2281–2291
22. Rokanujjaman, M., Islam, M., Hossain, M., Islam, M., Makihara, Y., Yagi, Y.: Effective part-based gait identification using frequency-domain gait entrophy features. Multimedia Tools and Applications **74** (2015) 3099–3120
23. Guan, Y., Li, C.T., Hu, Y.: Robust clothing-invariant gait recognition. In: Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), 2012 Eighth International Conference on. (2012) 321–324
24. Lee, S., Liu, Y., Collins, R.: Shape variation-based frieze pattern for robust gait recognition. In: Proc. of the 2007 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, Minneapolis, USA (2007) 1–8
25. Boulgouris, N., Chi, Z.: Human gait recognition based on matching of body components. Pattern Recognition **40** (2007) 1763–1770
26. Guan, Y., Li, C.T., Roli, F.: On reducing the effect of covariate factors in gait recognition: A classifier ensemble method. IEEE Transactions on Pattern Analysis and Machine Intelligence **37** (2015) 1521–1528
27. Liu, Z., Sarkar, S.: Effect of silhouette quality on hard problems in gait recognition. Trans. of Systems, Man, and Cybernetics Part B: Cybernetics **35** (2005) 170–183
28. Yan, S., Xu, D., Yang, Q., Zhang, L., Tang, X., Zhang, H.J.: Discriminant analysis with tensor representation. In: Proc. of the IEEE Computer Society Conf. Computer Vision and Pattern Recognition. (2005) 526–532
29. Xu, D., Yan, S., Tao, D., Zhang, L., Li, X., jiang Zhang, H.: Human gait recognition with matrix representation. IEEE Trans. Circuits Syst. Video Technol **16** (2006) 896–903
30. Liu, Z., Sarkar, S.: Simplest representation yet for gait recognition: Averaged silhouette. In: Proc. of the 17th International Conference on Pattern Recognition. Volume 1. (2004) 211–214
31. Xu, D., Huang, Y., Zeng, Z., Xu, X.: Human gait recognition using patch distribution feature and locality-constrained group sparse representation. IEEE Transactions on Image Processing **21** (2012) 316–326
32. Yang, J., Zhang, D., Frangi, A.F., yu Yang, J.: Two-dimensional pca: a new approach to appearance-based face representation and recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence **26** (2004) 131–137
33. Makihara, Y., Mannami, H., Tsuji, A., Hossain, M., Sugiura, K., Mori, A., Yagi, Y.: The ou-isir gait database comprising the treadmill dataset. IPSJ Transactions on Computer Vision and Applications **4** (2012) 53–62
34. Liu, Z., Sarkar, S.: Improved gait recognition by gait dynamics normalization. IEEE Transactions on Pattern Analysis and Machine Intelligence **28** (2006) 863–876
35. Li, X., Maybank, S., Yan, S., Tao, D., Xu, D.: Gait components and their application to gender recognition. Trans. on Systems, Man, and Cybernetics, Part C **38** (2008) 145–155