Kejun WANG, Liangliang LIU, Xinnan DING, Kaiqiang YU, Gang HU, 2021. A partition approach for robust gait recognition based on gait template fusion. *Frontiers of Information Technology & Electronic Engineering*, 22(5):709-719. https://doi.org/10.1631/FITEE.2000377

A partition approach for robust gait recognition based on gait template fusion

Key words: Gait recognition; Partition algorithms; Gait templates; Gait analysis; Gait energy image; Deep convolutional neural networks; Biometrics recognition; Pattern recognition

Corresponding author: Liangliang LIU
E-mail: liuliangliang@hrbeu.edu.cn
ORCID: https://orcid.org/0000-0002-7262-7502

Motivation

1. Gait recognition has significant potential for remote human identification, but it is easily influenced by identity-unrelated factors such as clothing, carrying conditions, and view angles.
2. The performance of projection-based methods (canonical correlation analysis (CCA) and coupled metric learning (CML)) is affected by factors that are unrelated to identity.
3. A gait template (such as GEI) and a set of images may be affected by contour changes or loss of temporal information concerning gait sequences.

Main idea

1. A gait template fusion method based on multi-channel mapping and a novel temporal approach called FEI is proposed to effectively represent more original gait information.
2. The FEI is separated into three parts: head, trunk, and leg. The features of each part of the body are learned from each block, and the low-dimensional features of the GEI are obtained after fusion.
3. A CNN-based fusion method is presented for gait feature extraction and recognition. Bottom layers with three routes in parallel are used to learn the gait features, and then identity recognition is performed through the fusion layer.

Method

This paper presents a novel fusion-based gait recognition method to take advantage of more gait information and solve the problem posed by silhouettes in different view angles or dresses.

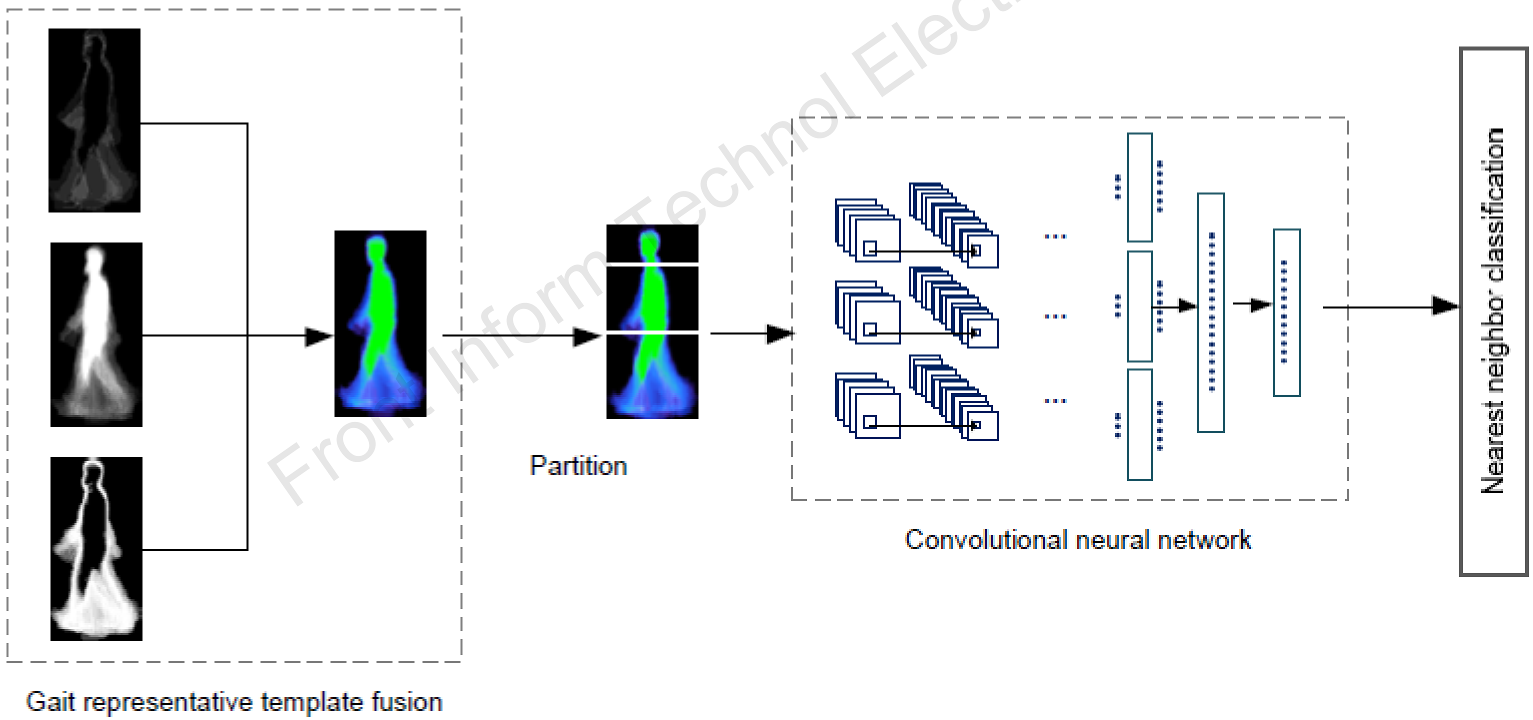

Fig. 1 Process flow of the proposed method

1) Gait representative template fusion method based on multi-channel mapping

The multi-channel mapping method is used to complete the fusion of multiple templates. Gait templates are marked with color and mapped to the RGB space in red, green, and blue channels. The **FEI** (*x*, *y*) is defined as follows:

$$\mathbf{FEI}(x,y)=$$

$$\begin{pmatrix} T_1(x,y)W_1^{\mathrm{R}}+T_2(x,y)W_2^{\mathrm{R}}+\cdots+T_m(x,y)W_m^{\mathrm{R}} \\ T_1(x,y)W_1^{\mathrm{G}}+T_2(x,y)W_2^{\mathrm{G}}+\cdots+T_m(x,y)W_m^{\mathrm{G}} \\ T_1(x,y)W_1^{\mathrm{B}}+T_2(x,y)W_2^{\mathrm{B}}+\cdots+T_m(x,y)W_m^{\mathrm{B}} \end{pmatrix}=\begin{pmatrix} E_{\mathrm{AEI}}(x,y) \\ E_{\mathrm{GEI}}(x,y) \\ E_{\mathrm{GEnI}}(x,y) \end{pmatrix}$$

$$i_m=(m-1)/(M-1),\qquad W_m^{\mathrm{G}}=\begin{cases}0, & 0\le i_m\le 1/2,\\ 2i_m-1, & 1/2<i_m<1,\end{cases}$$

$$W_m^{\mathrm{R}}=\begin{cases}1-2i_m, & 0\le i_m\le 1/2,\\ 0, & 1/2<i_m<1,\end{cases}\qquad W_m^{\mathrm{B}}=\begin{cases}2i_m, & 0\le i_m\le 1/2,\\ 2-2i_m, & 1/2<i_m<1,\end{cases}$$

where $T_m(x, y)$ is the image of the m^{th} template. In this study, we choose GEI, AEI, and GEnI to compose the FEI. i_m is the normalized number of the m^{th} template, and W_m^{R}, W_m^{G}, and W_m^{B} are the weights mapped to the RGB space.

1) Gait representative template fusion method based on multichannel mapping (Cont'd)

GEI has an efficient gait characterization template, but has no connection between adjacent frames. AEI can extract the active regions by calculating the frame difference. Note that GEnI is not sensitive to changes in static information about carrying conditions or clothing. The FEI composed of these three templates contains more gait information and the respective shortcomings are offset by each other.

$$E_{\mathrm{GEI}}(x,y)=\frac{1}{N}\sum_{n=1}^{N}B(x,y,n),\quad E_{\mathrm{AEI}}(x,y)=\frac{1}{N}\sum_{n=1}^{N}D_n(x,y,n),$$

$$D_n(x,y,n)=\left|B(x,y,n+1)-B(x,y,n)\right|,$$

$$E_{\mathrm{GEnI}}(x,y)=-E_{\mathrm{GEI}}(x,y)\log_2 E_{\mathrm{GEI}}(x,y)\\ -(1-E_{\mathrm{GEI}}(x,y))\log_2(1-E_{\mathrm{GEI}}(x,y)).$$

Here, N is the number of frames in a gait cycle and $B(x,y,n)$ is the binary and normalized silhouette of the n^{th} frame. $D_n(x,y,n)$ is the difference between the next frame and the current frame of the silhouettes.

1) Gait representative template fusion method based on multichannel mapping (Cont'd)

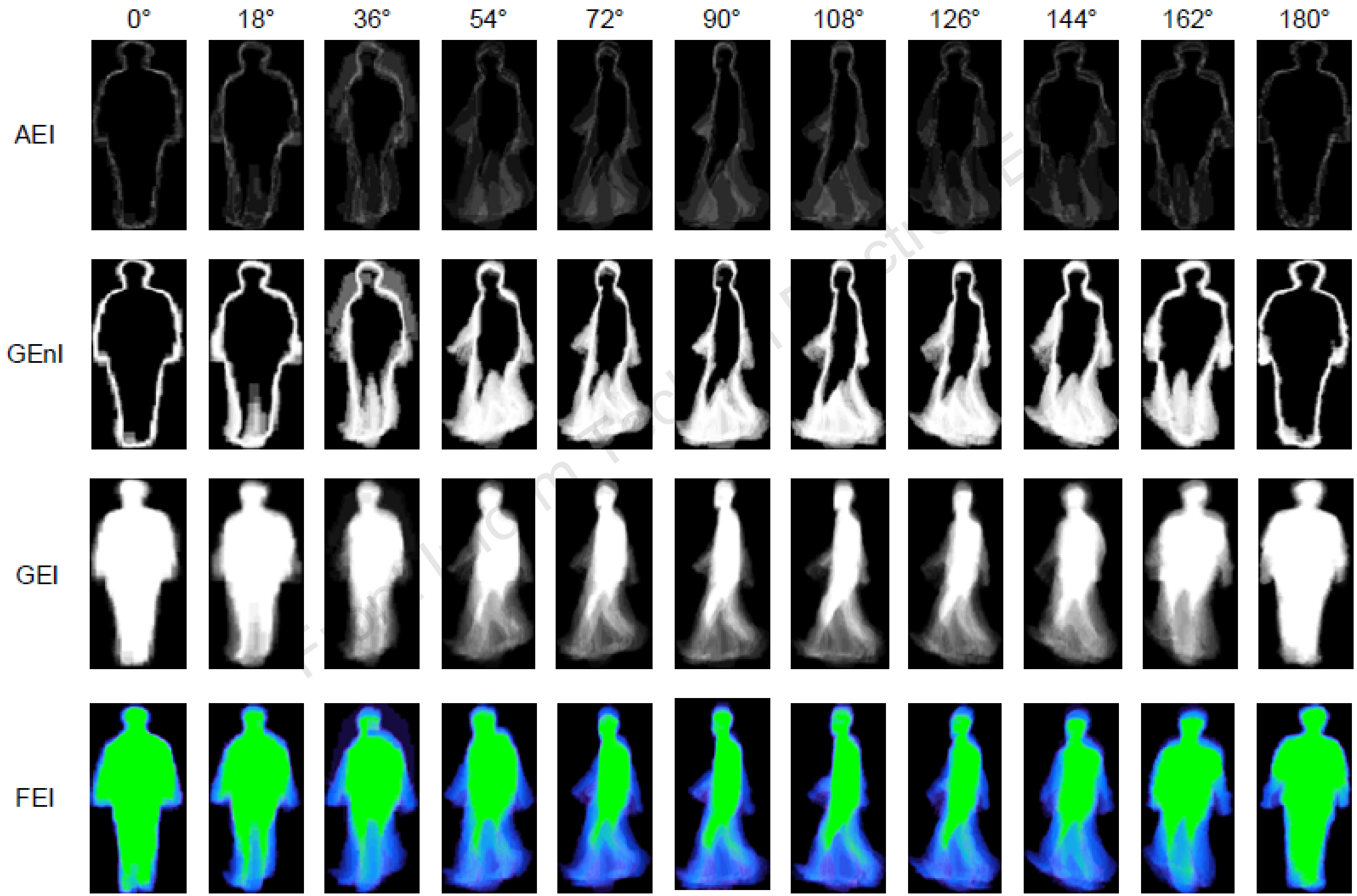

Fig. 2 AEIs, GEnIs, GEIs, and FEIs at different view angles

2) A robust partitioning approach for covariate factors

As shown in Fig. 3, the changing of silhouettes caused by the clothing or backpacks is concentrated mainly on the human body's torso, and less on the head and legs. Therefore, we present a solution to the covariate factors based on partitioning. FEI is cropped into the head, trunk, and leg regions as shown in Fig. 3. Due to normalization in the FEI generation process, it is easy to apply this method to all the subjects. In this way, attention is not solely focused on the whole silhouette, but also dispersed to each sub-FEI to reduce the effect of covariate factors.

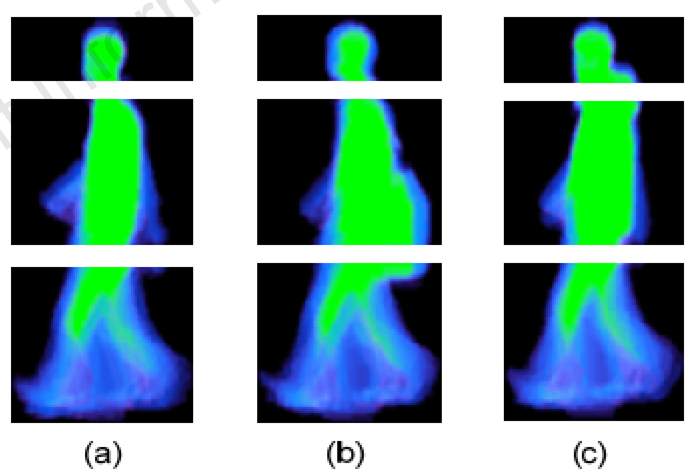

Fig. 3 Partitioned FEI when the person is in the normal condition (a), carrying a bag (b), and wearing a coat (c)

3) Feature-level information fusion-based deep convolutional networks

Three sub-FEIs are sent to the CNN to learn the gait feature, and the fusion layer combines multiple features to recognize it. Table 1 shows the structure of the network in detail.

Table 1 Implementation details of the network

Layer (head)	Layer (trunk)	Layer (leg)
Conv5-ReLU-Batch_norm (60, 28, 16)	Conv5-ReLU-Batch_norm (60, 44, 16)	Conv5-ReLU-Batch_norm (60, 44, 16)
	Conv3-ReLU (58, 42, 32)	Conv3-ReLU (58, 42, 32)
Maxpool2 (30, 14, 16)	Maxpool2 (29, 21, 32)	Maxpool2 (29, 21, 32)
Conv3-ReLU-Batch_norm (28, 12, 32)	Conv3-ReLU-Batch_norm (27, 19, 64)	Conv3-ReLU-Batch_norm (27, 19, 64)
	Conv3-ReLU (25, 17, 64)	Conv3-ReLU (25, 17, 64)
Maxpool2 (14, 6, 32)	Maxpool2 (14, 8, 64)	Maxpool2 (14, 8, 64)
FC-ReLU (512)	FC-ReLU (1024)	FC-ReLU (1024)
Concat (2560)	Concat (2560)	Concat (2560)
FC-ReLU (1024)	FC-ReLU (1024)	FC-ReLU (1024)
FC-ReLU (512)	FC-ReLU (512)	FC-ReLU (512)
FC-ReLU (62)	FC-ReLU (62)	FC-ReLU (62)

Major results

Impact of fusing gait representative templates

Table 2 Comparison of different gait representative templates by accuracy

Probe	Accuracy (%)			
	AEI	GEI	GEnI	FEI
NM	88.7	93.5	94.8	96.8
CL	35.2	32.7	34.3	35.4
BG	46.3	44.5	45.8	47.0

Gallery: NM 1–4, 90°; probe: NM 5–6, CL, BG

Major results (Cont'd)

Impact of CNN-based partition and fusion

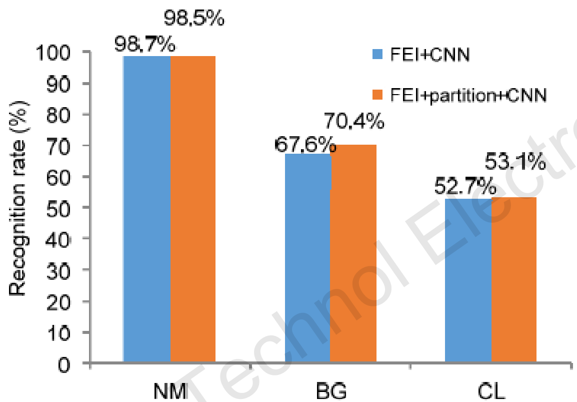

Fig. 4 Comparison of partition and fusion by accuracy (References to color refer to the online version of this figure)

Gallery: NM 1–4, 90°; probe: NM 5–6, CL, BG

Table 3 Comparison of accuracy at different numbers of partitions

Probe	Accuracy (%)			
	Partition number=1	Partition number=2	Partition number=3	Partition number=4
NM	98.7	98.5	98.5	98.0
CL	67.6	70.1	70.4	68.2
BG	52.7	53.3	53.1	53.1

Gallery: NM 1–4, 90°; probe: NM 5–6, CL, BG Partition number=1: the whole image without partition; Partition number=2: two partition parts involve double 64×64×3 sub-FEIs; Partition number=3: three sub-FEIs are 64×32×3, 64×48×3, and 64×48×3; Partition number=4: four partition parts involve 64×32×3, 64×24×3, 64×24×3, and 64×48×3 sub-FEIs

Major results (Cont'd)

Comparison with other methods

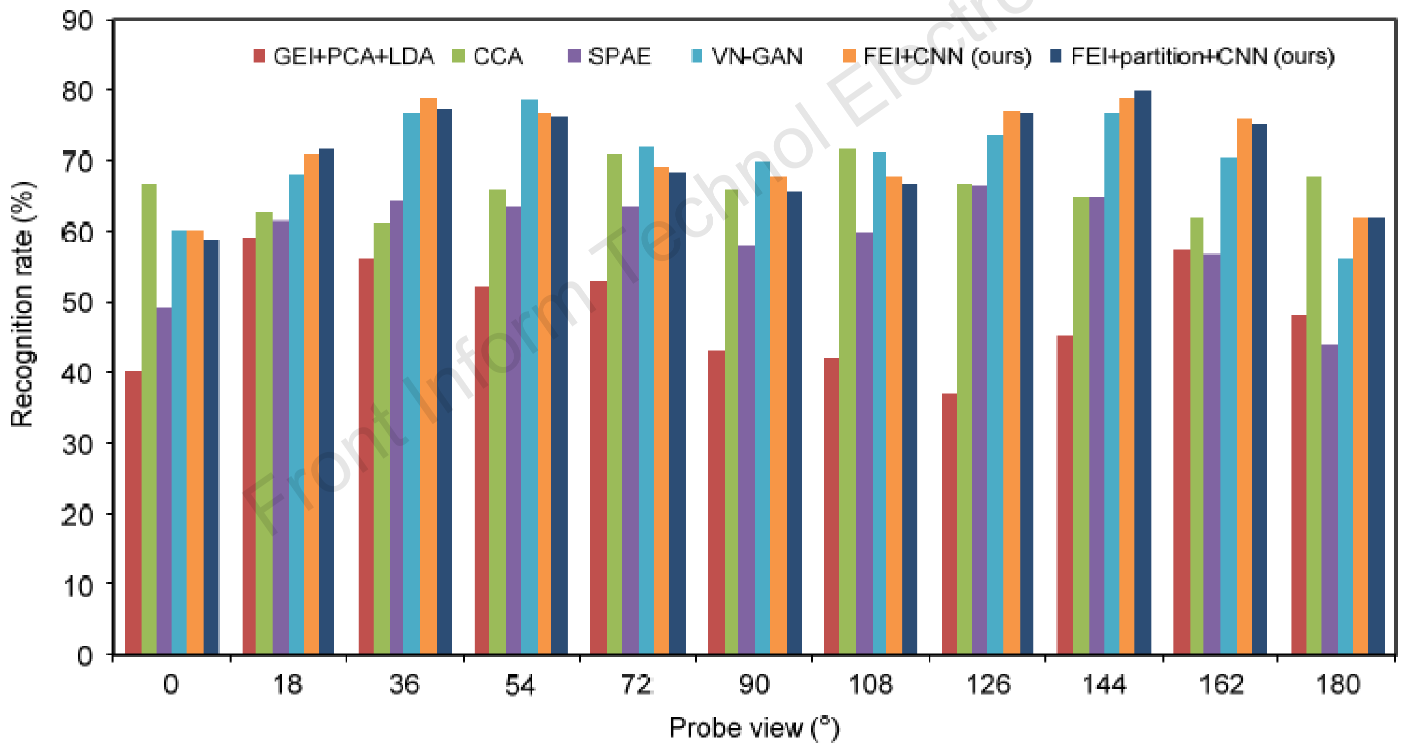

Fig. 5 Comparison of the proposed method with existing ones by accuracy under probe NM 5–6, 0°–180° (gallery: NM 1–4, 0°–180°) (References to color refer to the online version of this figure)

Major results (Cont'd)

Comparison with other methods

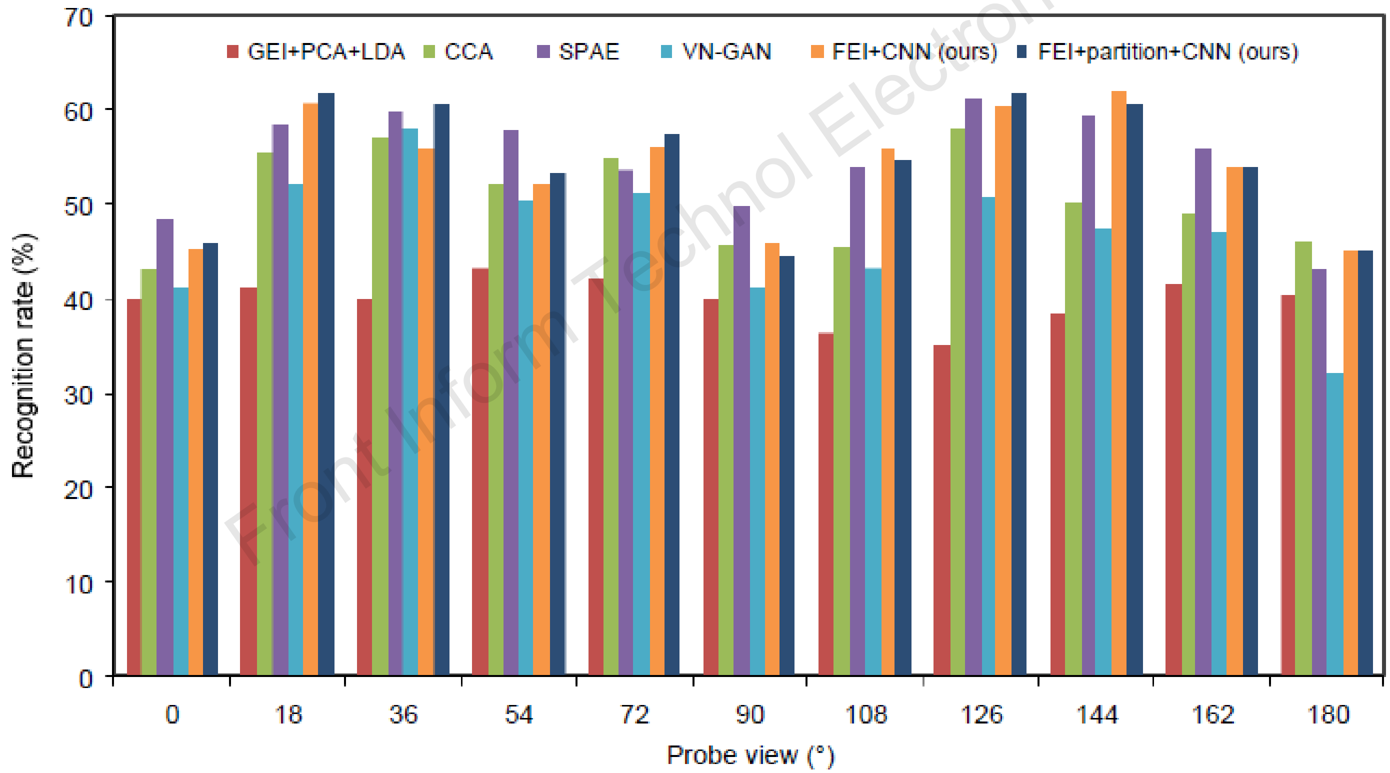

Fig. 6 Comparison of the proposed method with existing ones by accuracy under probe BG 0°–180° (gallery: NM 1–4, 0°–180°) (References to color refer to the online version of this figure)

Major results (Cont’d)

Comparison with other methods

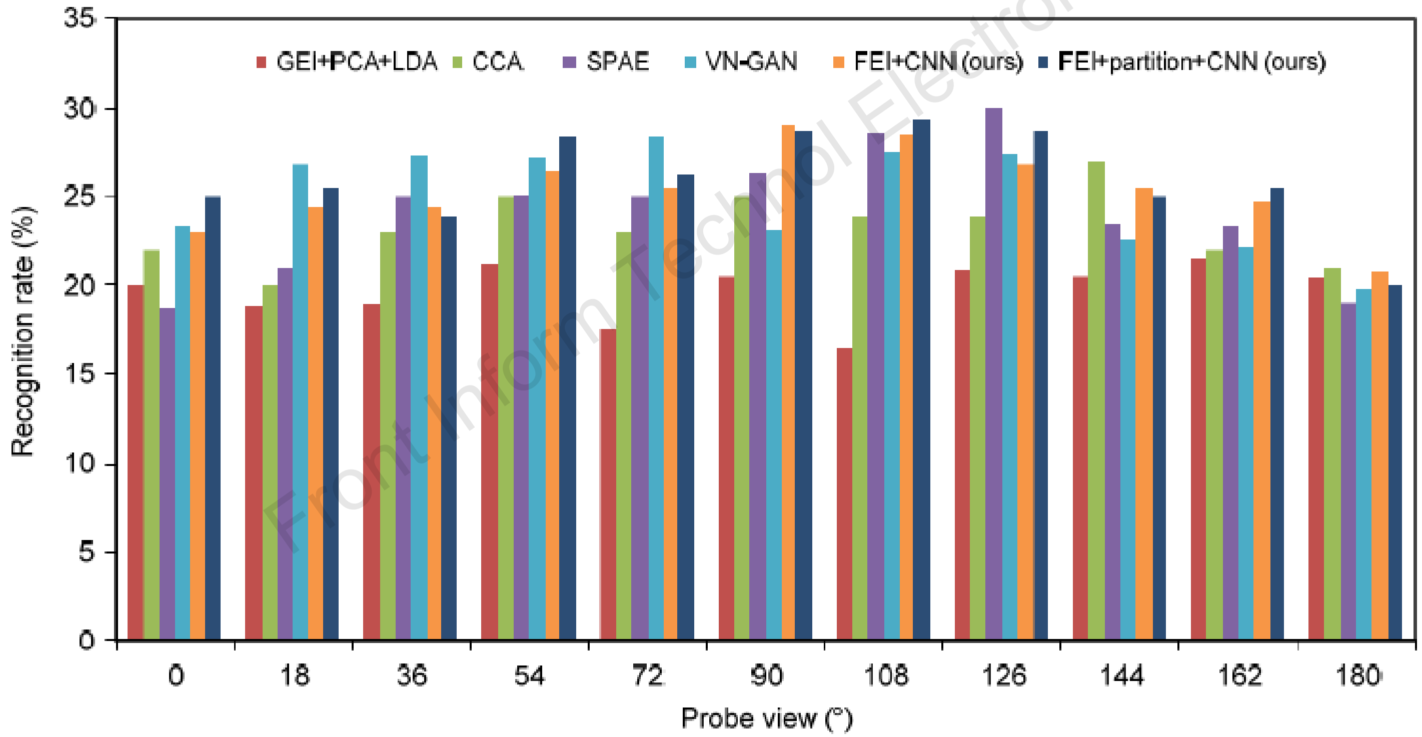

Fig. 7 Comparison of the proposed method with existing ones by accuracy under probe CL 0°–180° (gallery: NM 1–4, 0°–180°) (References to color refer to the online version of this figure)

Conclusions

1. We have employed gait template fusion using multi-channel mapping to obtain an FEI to represent more original gait features. The FEI approach also combines the advantages of the existing templates to increase the robustness of identity-unrelated factors.

2. Moreover, the partition and fusion method has been applied to a CNN, where the features of a partitioned FEI and fusion for identification have been extracted. This is effective in reducing the impact of changes in clothing and carrying conditions.

3. Experimental results confirm the proposed method's effectiveness and robustness for gait recognition, and show especially good performance in various complex environments.

Kejun WANG received the Ph.D. degree in special auxiliary devices and systems for ship and ocean engineering from Harbin Engineering University, Harbin, China, in 1995. From 1996 to 1998, he was a postdoctoral research fellow in fluid power transmission and control at Harbin Institute of Technology. He is currently a professor with College of Intelligent Systems Science and Engineering in Harbin Engineering University. His research interests cover biometrics and pattern recognition, machine learning, and neural network.

Liangliang LIU is currently a Ph.D. candidate in control theory and control engineering from Harbin Engineering University, Harbin, China. He received the M.S. degree in computer science and technology from Harbin Engineering University in 2008. His research interests cover gait recognition and machine learning.

Xinnan DING is currently a Ph.D. candidate in control theory and control engineering from Harbin Engineering University, Harbin, China. She received the M.S. degree in control engineering from Harbin Engineering University in 2018. Her research interests cover gait recognition and machine learning.