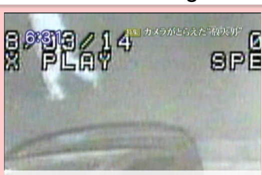
The IAPR/IEEE Winter School on Biometrics 2021 26 January, 2021

Gait Analysis

Yasushi Yagi
Osaka University, Osaka, Japan

Advantage of gait recognition

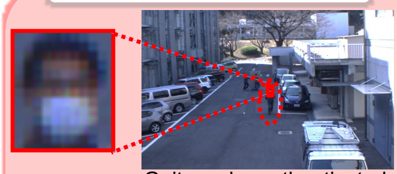
Criminal investigation



CCTV of firer

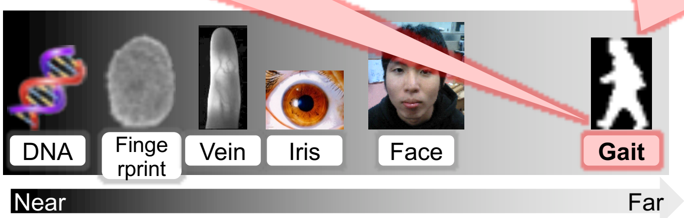
Judge whether a perpetrator and a suspect are the same or not from gaits

Authentication at a distance



Gait can be authenticated at a distance from a camera

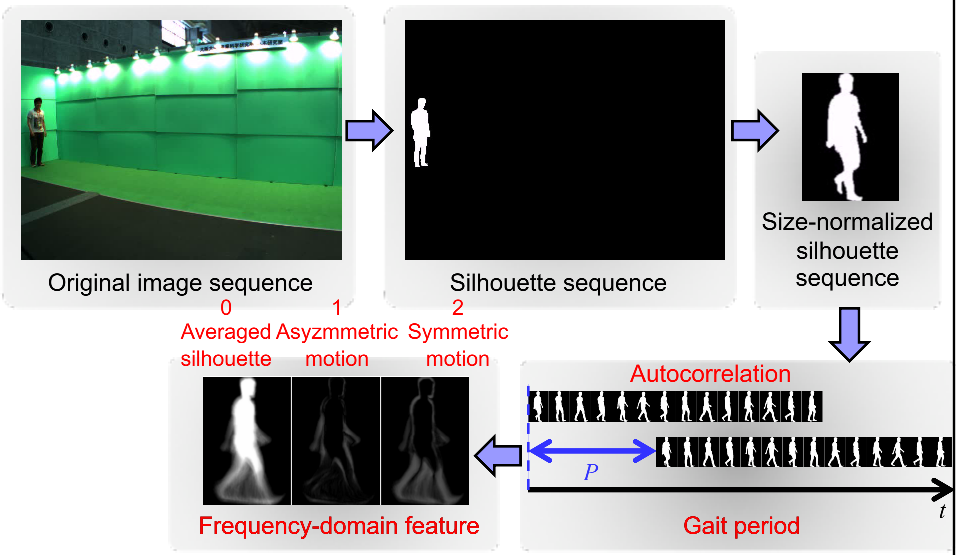
Face recognition does not work due to heavy occlusions by mask



Distance to sensor

2

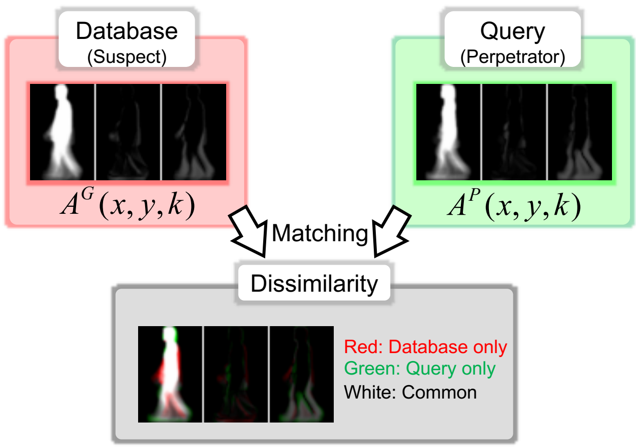
Standard approach for gait feature extraction



Y. Makihara, R. Sagawa, Y. Mukaigawa, T. Echigo, and Y. Yagi, ``Gait Recognition Using a View Transformation Model in the Frequency Domain,'' 9th European Conf. on Computer Vision, Vol. 3, pp. 151-

3

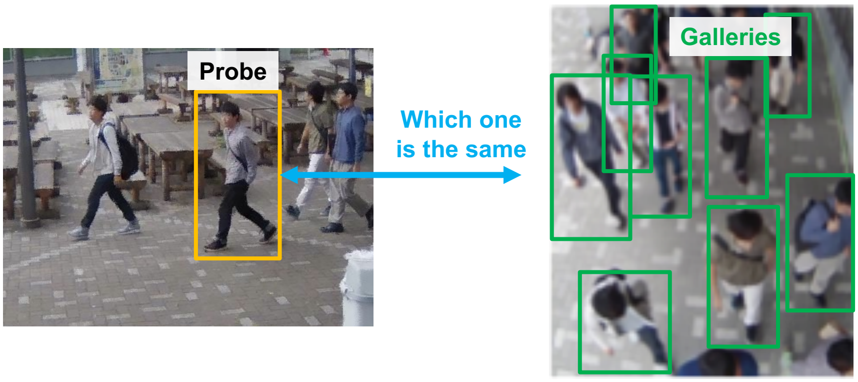
Dissimilarity: Single feature



$$t = \sqrt{\sum_{x,y,k} \left(A^G(x,y,k) - A^P(x,y,k) \right)^2}$$

4

Identification (1:N matching)

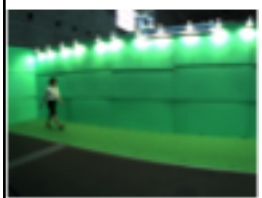


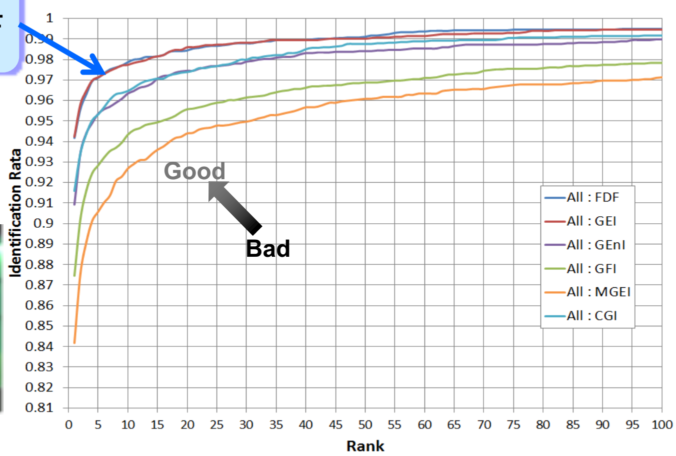
Applications
- Person re-identification
- ID-less access control

5

Performance evaluation: identification

[Iwama et al. IFS 2012]



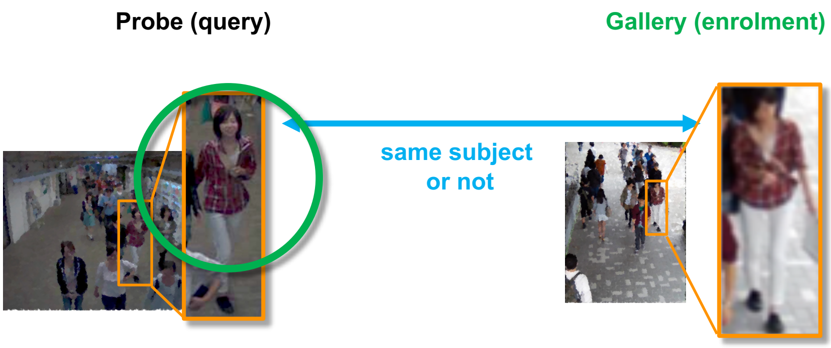


Cumulative matching characteristics (CMC) curve

94% rank-1 identification rate (N = 3,141)

6

Verification (1:1 matching)

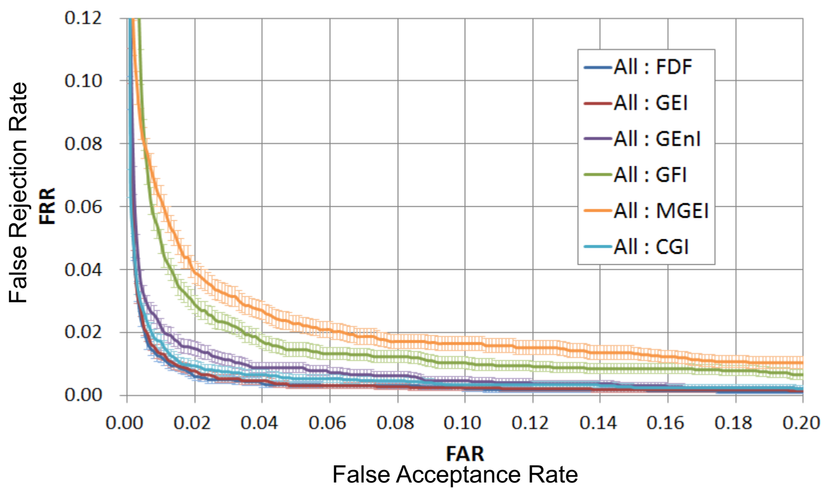


Applications

- Matching a perpetrator and suspect for a criminal investigation.
- Detecting a specific person at border control.

7

Performance evaluation: verification



Equal Error Rate (EER): 1.15%

8

World first packaged gait verification system for criminal investigation (2013)

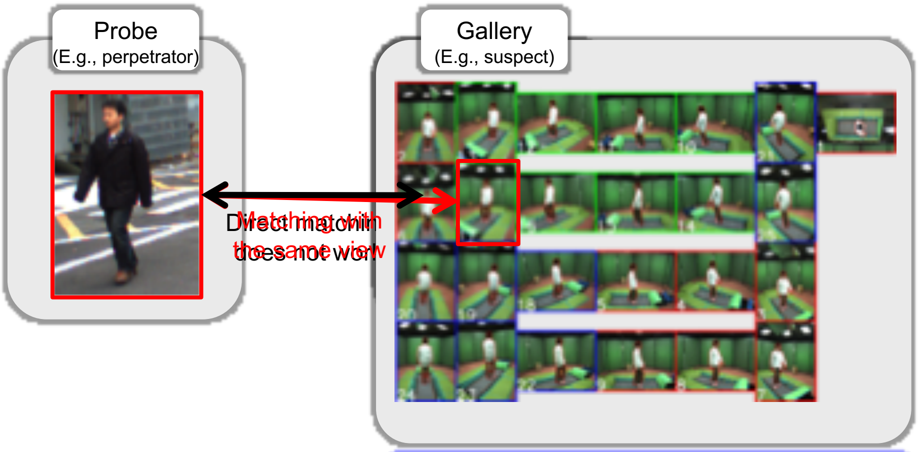
9

What is the difficulty for applying gait to wide-area surveillance ?

- The difference of the observation direction
- Occlusion in crowd scene
- Speed difference
- The difference of clothes & carrying objects
- The difference of shoes & ground plane
- The difference of health condition
- The difference of intention
- Low sampling rate

10

Challenge -View differences-

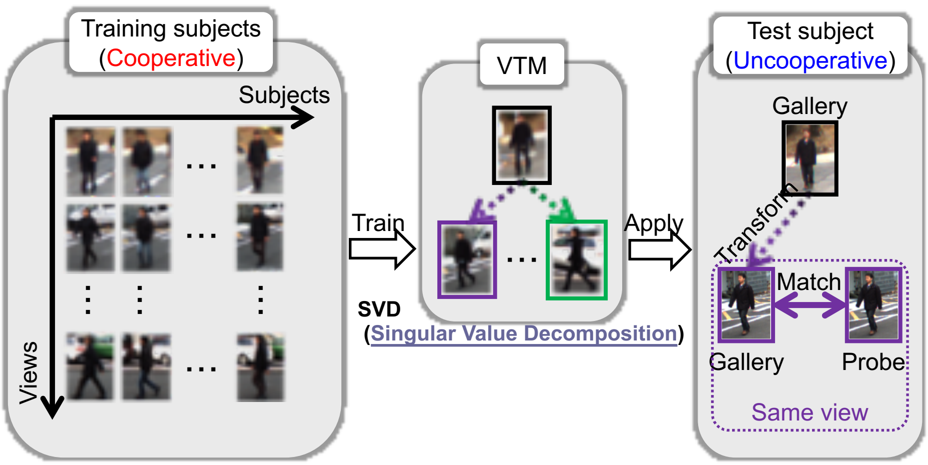


Difficult to collect multi-view gait features for uncooperative subject

11

View transformation model (VTM)

[Makihara et al. ECCV 2006]



12

Formulation of VTM in frequency domain

- Decompose training data matrix of gait features into individuals and views by SVD

$$\begin{array}{cc} & \text{individual} \\ \text{view} & \begin{bmatrix} \boldsymbol{a}_{\theta_1}^1 & \boldsymbol{a}_{\theta_1}^2 & \cdots & \boldsymbol{a}_{\theta_1}^M \\ \boldsymbol{a}_{\theta_2}^1 & \boldsymbol{a}_{\theta_2}^2 & \cdots & \boldsymbol{a}_{\theta_2}^M \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{a}_{\theta_K}^1 & \boldsymbol{a}_{\theta_K}^2 & \cdots & \boldsymbol{a}_{\theta_K}^M \end{bmatrix} \end{array} = USV^T = \begin{bmatrix} P_{\theta_1} \\ P_{\theta_2} \\ \vdots \\ P_{\theta_K} \end{bmatrix} \begin{bmatrix} \boldsymbol{v}^1 & \boldsymbol{v}^2 & \cdots & \boldsymbol{v}^M \end{bmatrix}$$

Training data matrix

Transformation matrix to each view

View-independent individual vector

- Gait feature for mth subject from θ_i view

$$\boldsymbol{a}_{\theta_i}^m = P_{\theta_i} \boldsymbol{v}^m$$

13

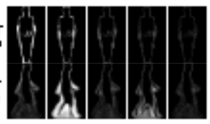
View transformation

- From a single reference θ_j to θ_i

$$\left.\begin{aligned} \boldsymbol{a}_{\theta_i}^m &= P_{\theta_i} \boldsymbol{v}^m \\ \boldsymbol{a}_{\theta_j}^m &= P_{\theta_j} \boldsymbol{v}^m \end{aligned}\right\} \Rightarrow \boldsymbol{a}_{\theta_i}^m = P_{\theta_i} P_{\theta_j}^{+} \boldsymbol{a}_{\theta_j}^m$$

0 deg

90 deg



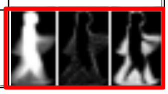
  - □ Orthogonal motion to reference θ_j is degenerated

- From multiple references $\{\theta_j(1), \ldots, \theta_j(k)\}$ to θ_i

$$\left.\begin{aligned} \boldsymbol{a}_{\theta_i}^m &= P_{\theta_i} \boldsymbol{v}^m \\ \boldsymbol{a}_{\theta_j(1)}^m &= P_{\theta_j(1)} \boldsymbol{v}^m \\ &\vdots \\ \boldsymbol{a}_{\theta_j(k)}^m &= P_{\theta_j(k)} \boldsymbol{v}^m \end{aligned}\right\} \Rightarrow \boldsymbol{a}_{\theta_i}^m = P_{\theta_i} \begin{bmatrix} P_{\theta_j(1)} \\ \vdots \\ P_{\theta_j(k)} \end{bmatrix}^{+} \begin{bmatrix} \boldsymbol{a}_{\theta_j(1)}^m \\ \vdots \\ \boldsymbol{a}_{\theta_j(k)}^m \end{bmatrix}$$

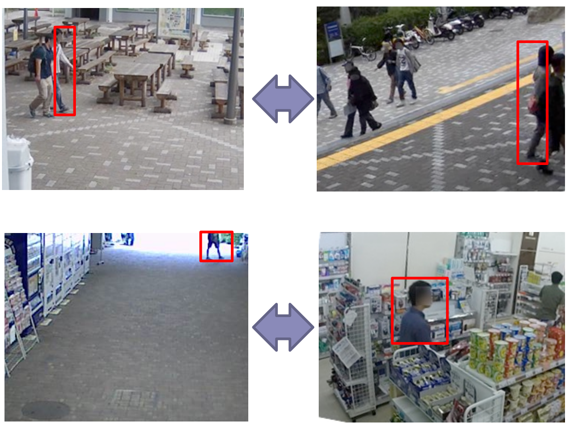
14

Transformation results

	Gallery		
0 deg			
15 deg			
30 deg			
45 deg			
60 deg			
75 deg			
90 deg			

15

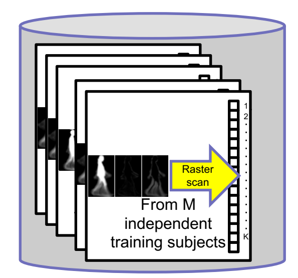
Actual situation of observed gait in surveillance

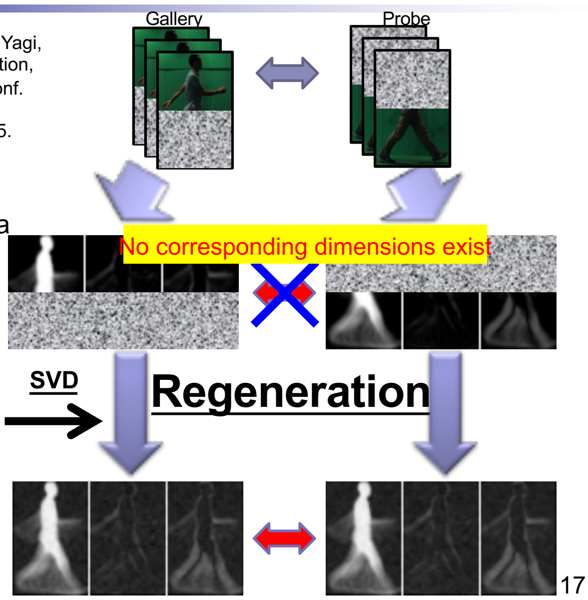


16

Gait regeneration for recognition

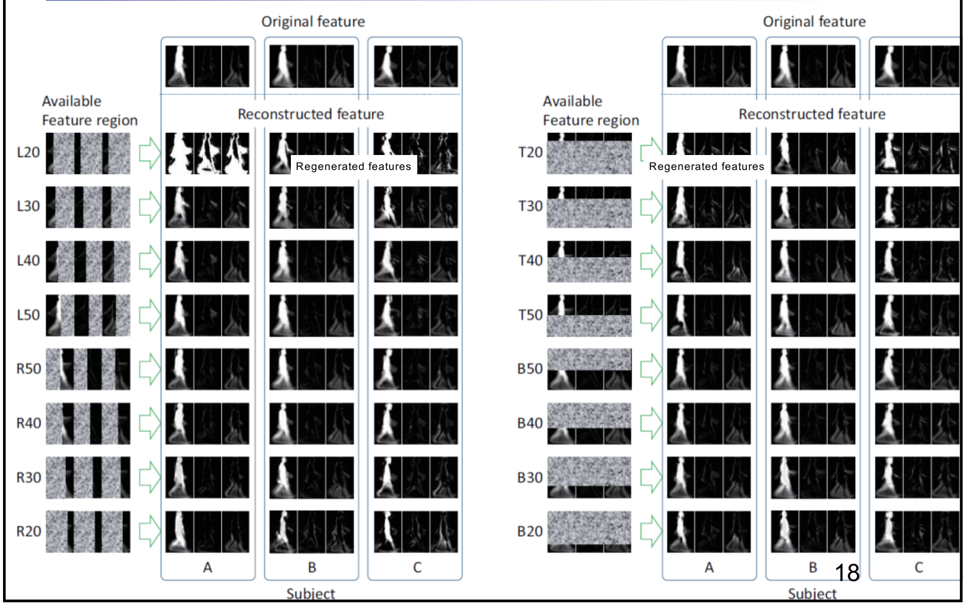
D. Muramatsu, Y. Makihara, Y. Yagi,
Gait Regeneration for Recognition,
In Proc. of the 7th IAPR Int. Conf.
on Biometrics (ICB 2015), No.
Paper ID 60, pp. 1-8, May 2015.





17

Experiment : Regenerated gait



18

Challenge -Speed difference-

2 km/h

4 km/h

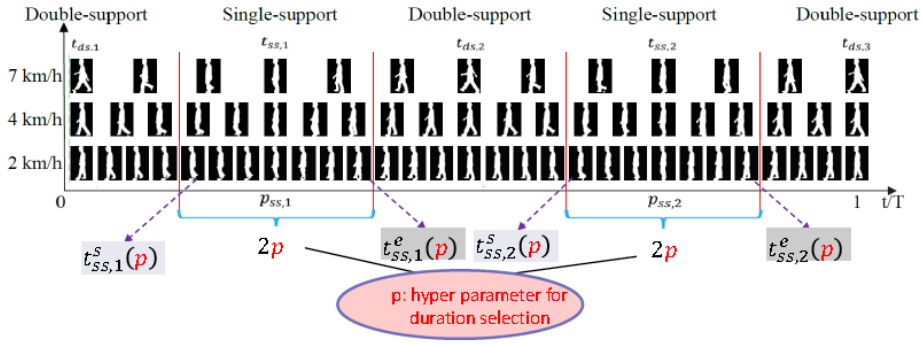
6 km/h

8 km/h

10 km/h

Single-Support GEI (SSGEI)

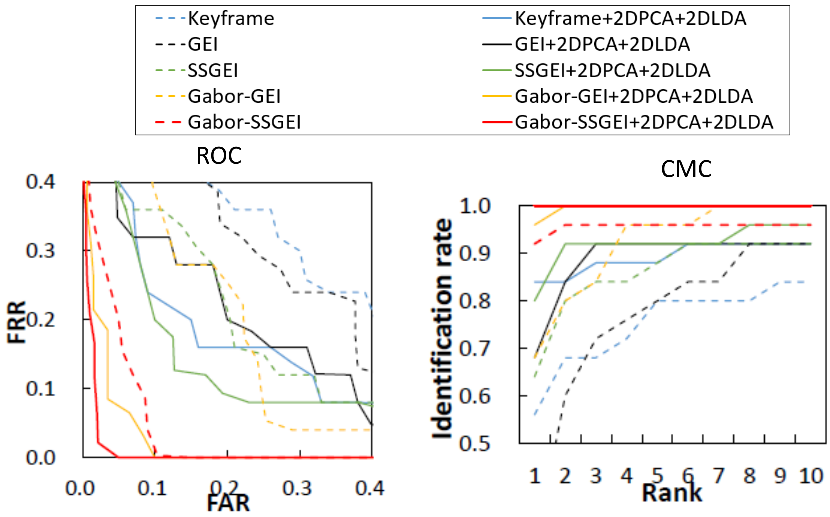
- Aggregate multiple frames of optimal duration around single-support phase.



- Representation:

$$S(x,y;p) = \frac{1}{2}\sum_{k=1}^{2}\frac{1}{t^e_{ss,k}(p) - t^s_{ss,k}(p) + 1}\sum_{t=t^s_{ss,k}(p)}^{t^e_{ss,k}(p)} I(x,y,t), \qquad (0 < p \leq 1/4).$$

Experiments: Gallery 4 km/h vs. probe 7 km/h

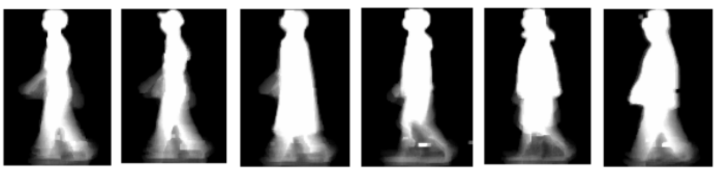


The propose method achieves the best accuracy.

21

Difference of Clothing

- Gait recognition
  - Pros:
    - Availability at a distance for an uncooperative subject (c.f. face, iris)
  - Cons:
    - Accuracy drop due to many covariates (e.g., clothing, view, speed)

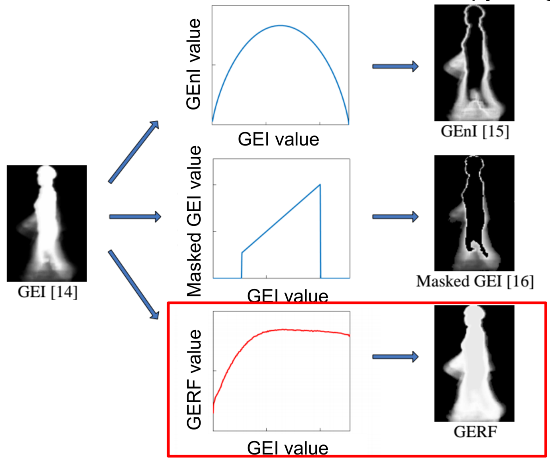


GEI [Han and Bhanu 2006] under clothing variations

22

Gait energy response function for clothes-invariant gait recognition (ACCV2016)

- Transform GEI into more discriminative feature under clothes variation

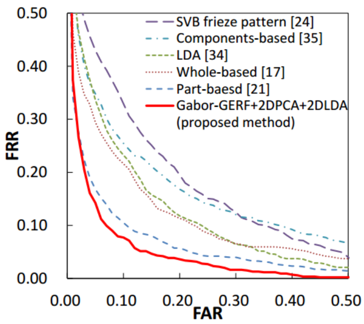


Gait Energy Response Function

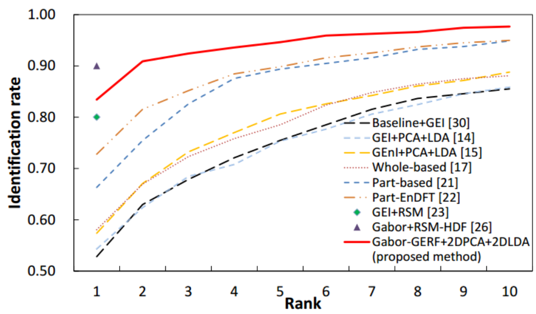
A sort of gait energy transformation process via a gait energy response function (GERF)

Comparison with state-of-the-arts methods

- Compare with the state-of-the-arts methods



ROC curves



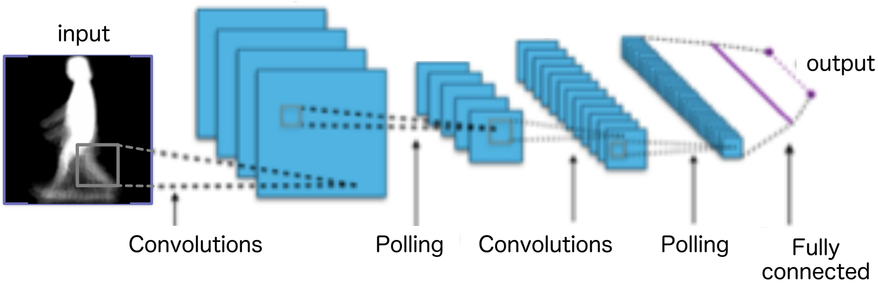
CMC curves

Gait Features

- Frequency Domain Feature (FDF)
  - Y. Makihara, R. Sagawa, Y. Mukaigawa, T. Echigo, Y. Yagi, ``Gait Recognition Using a View Transformation Model in the Frequency Domain,'' In 9th European Conf. on Computer Vision (ECCV), Vol. 3, pp. 151--163, May 2006.
- Single-Support Gait Energy Image (SSGEI)
  - C. Xu, Y. Makihara, X. Li, Y. Yagi, J. Lu, ``Speed-Invariant Gait Recognition Using Single-Support Gait Energy Image," Multimedia Tools and Applications, Vol. 78, No. 18, pp. 26509-26536, Sep. 2019 (Published online: 11 Jun. 2019).
- Gait Energy Response Function (GERF)
  - X. Li, Y. Makihara, C. Xu, D. Muramatsu, Y. Yagi, M. Ren, ``Gait Energy Response Functions for Gait Recognition against Various Clothing and Carrying Status," Applied Science, Vol. 8, No. 1380, pp. 1-22, Aug. 2018.
- Frequency-domain gait entropy (EnDFT)
  - Md. Rokanujjaman, Md. S. Islam, Md. A. Hossain, Md. R. Islam, Y. Makihara, Y. Yagi, ``Effective Part-Based Gait Identification using Frequency-Domain Gait Entrophy Features," Multimedia Tools and Applications, Vol. 74, No. 9, pp. 3099-3120, May 2015 (Published online: 22 Nov. 2013).
- Normal Distance Maps
  - H. El-Alfy, I. Mitsugami, Y. Yagi, ``Gait Recognition Based on Normal Distance Maps," IEEE Trans. on Cybernetics, Vol. 48, No. 5, pp. 1526 - 1539, May 2018 (Published online: 05 Jun. 2017).
- Two-Point Gait
  - S. Lombardi, K. Nishino, Y. Makihara, Y. Yagi, ``Two-Point Gait: Decoupling Gait from Body Shape,'' In 14th IEEE Int. Conf. on Computer Vision (ICCV 2013), pp. 1041-1048, Dec. 2013.

25

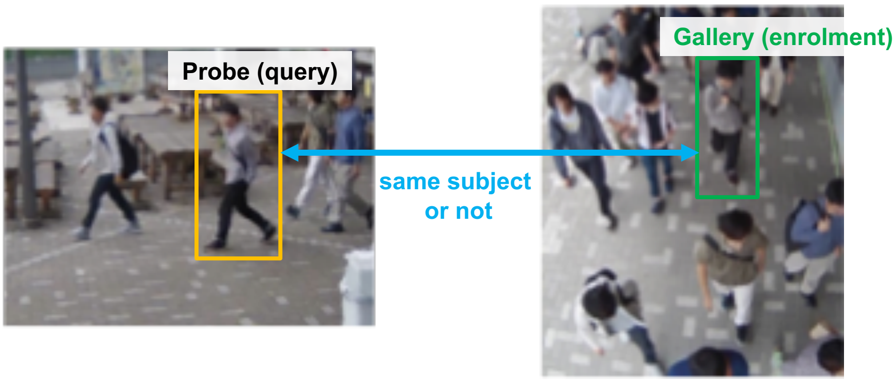
Convolutional Neural Network (CNN) -based gait recognition



- CNN-based methods have achieved state-of-the-art performance.
- Network architectures can be designed flexibly.

26

Verification (1:1 matching)

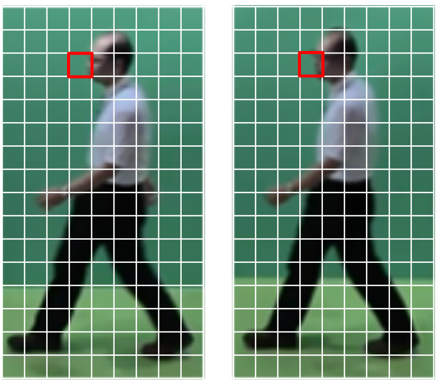


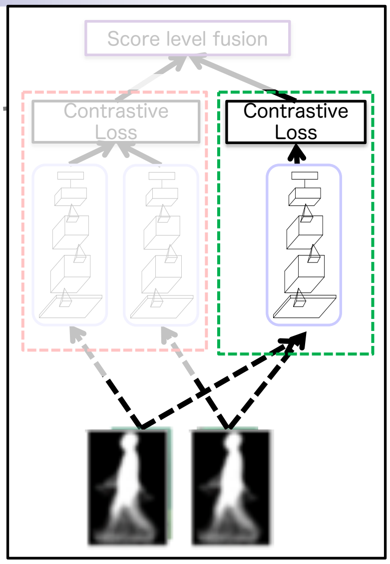
Applications

- Matching a perpetrator and suspect for a criminal investigation.
- Detecting a specific person at border control.

27

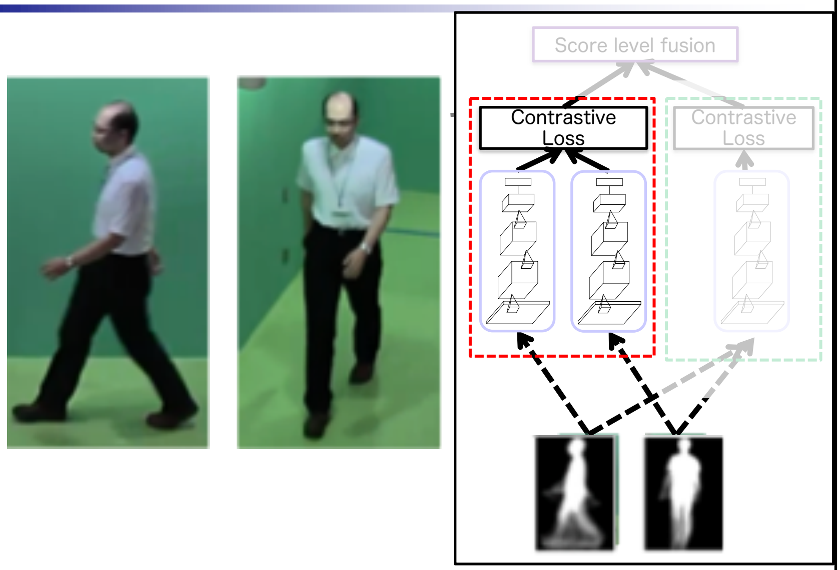
In case of small angular difference





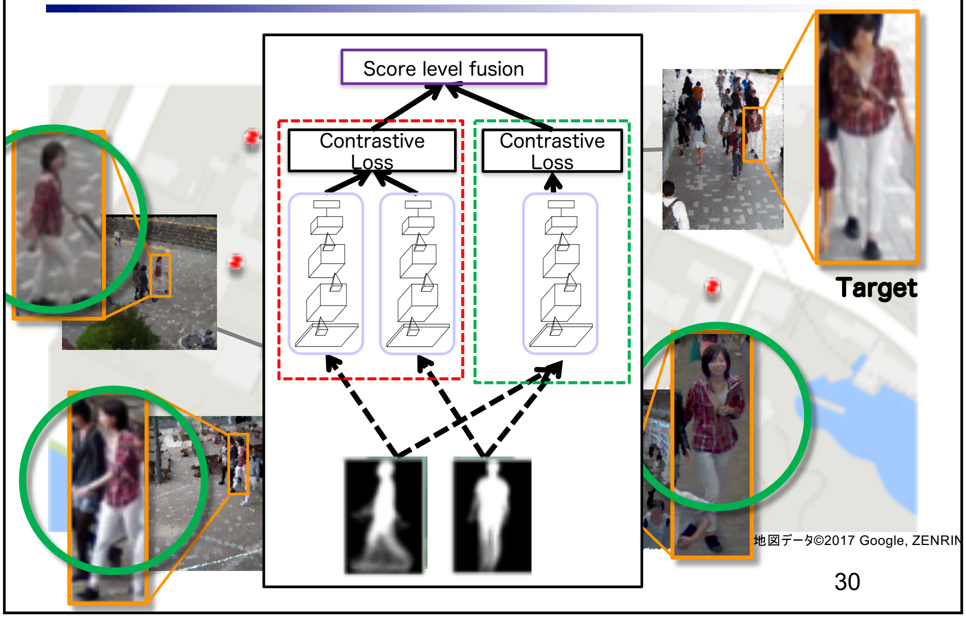
28

In case of large angular difference



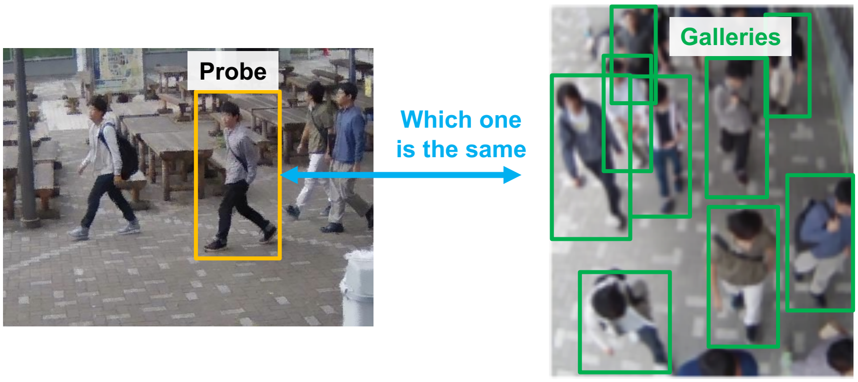
29

CNN based Cross View Approach



30

Identification (1:N matching)



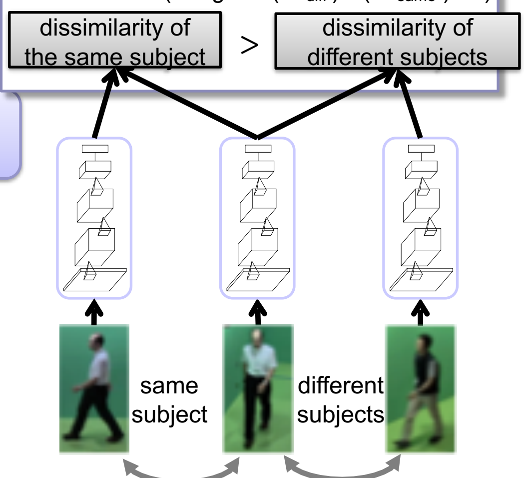
Applications

- Person re-identification
- ID-less access control

Network architecture for identification

Loss : Triplet ranking loss

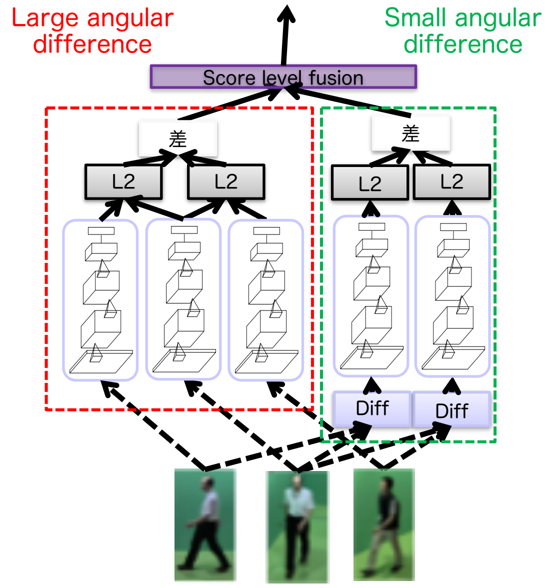
Loss = max($margin$ – ($L2_{diff}$)^2+($L2_{same}$)^2, 0)



Relative similarity scores are important !

Input: triplet GEIs including the same subject and different subjects

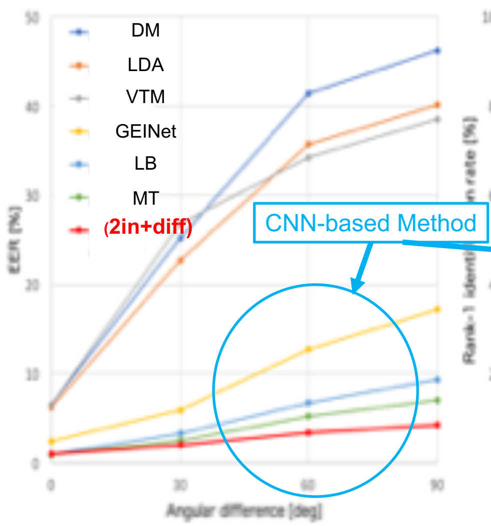
CNN based gait identification



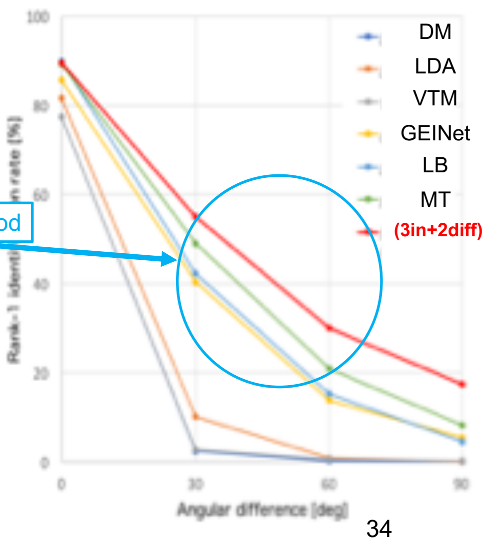
33

Comparison with benchmarks

Verification



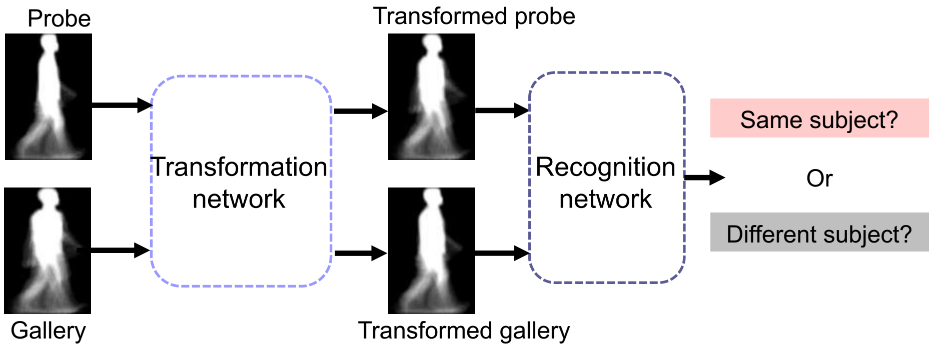
Identification



34

Cross-View Gait Recognition using Pair-wise Spatial Transformer Networks

- An end-to-end CNN-based framework integrating spatial registration and discrimination learning.

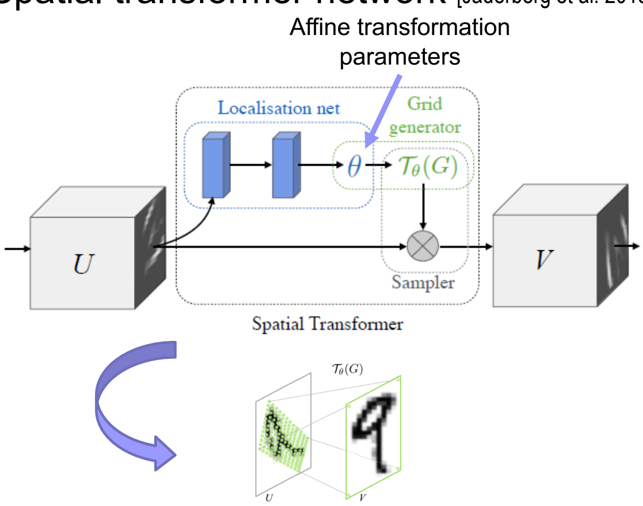


C. Xu, Y. Makihara, X. Li, Y. Yagi, J. Lu, ``Cross-View Gait Recognition using Pairwise Spatial Transformer Networks,'' IEEE Trans. on Circuits and Systems for Video Technology, 2020 (Published online: 21 Feb. 2020).

35

Inspiration

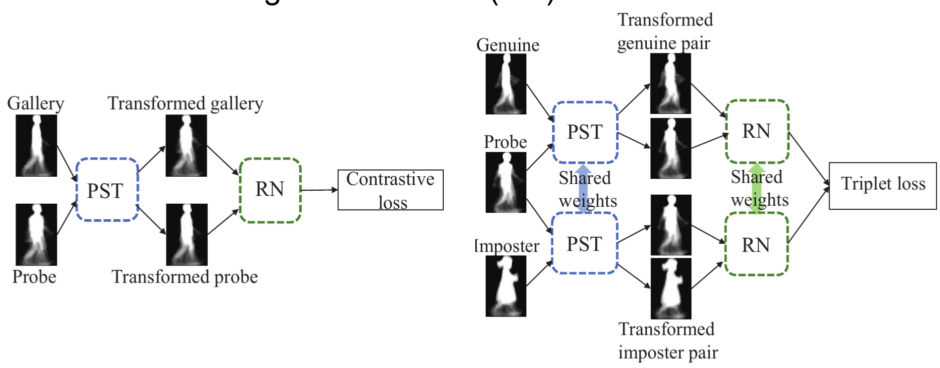
- Spatial transformer network [Jaderberg et al. 2015]



36

Gait recognition using Pair-wise Spatial Transformer Networks (PSTN)

- Two parts: Pair-wise Spatial Transformer (PST) and recognition network (RN).



Verification scenario

Identification scenario

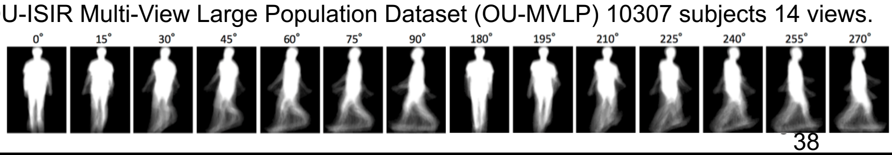
37

Comparison on OU-MVLP

- EER [%] of each angular difference.

Angular difference	0°	30°	60°	90°	Mean
DM [Takemura et al. 2018]	6.5	25.2	41.4	46.2	27.2
LDA [Otsu 1982]	6.2	22.7	35.7	40.1	24.0
VTM [Makihara et al .2006]	6.5	26.8	34.2	38.5	25.0
GEINet [Shiraga et al. 2016]	2.4	5.9	12.7	17.2	8.1
Original LB [Wu et al. 2017]	1.0	3.3	6.7	9.3	4.3
Original MT [Wu et al. 2017]	***0.9***	2.5	5.2	7.0	3.3
diff + 2in [Takemura et al. 2018]	1.0	2.0	3.4	4.2	***2.4***
PST-LB* + PST-2in (proposed)	**0.6**	**1.5**	**2.8**	**3.7**	**1.9**

OU-ISIR Multi-View Large Population Dataset (OU-MVLP) 10307 subjects 14 views.



38

Comparison on OU-MVLP

- Rank-1 identification rate [%] of each angular difference.

Angular difference	0°	30°	60°	90°	Mean
DM [Takemura et al. 2018]	77.4	2.4	0.2	0.0	20.3
LDA [Otsu 1982]	81.6	10.1	0.8	0.1	24.4
VTM [Makihara et al .2006]	77.4	2.7	0.6	0.2	20.5
GEINet [Shiraga et al. 2016]	85.7	40.3	13.8	5.4	40.7
Original LB [Wu et al. 2017]	89.9	42.2	15.2	4.5	42.6
Original MT [Wu et al. 2017]	89.3	49.0	20.9	8.2	46.9
2diff + 3in [Takemura et al. 2018]	89.5	55.0	30.0	17.3	52.7
PST-2LB* + PST-4in (proposed)	**93.9**	**69.2**	**41.9**	**25.9**	**63.1**

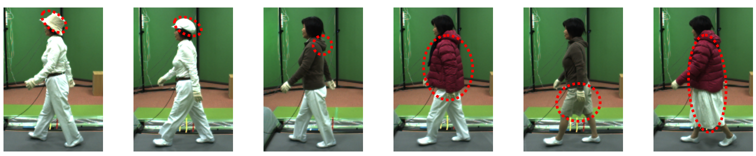
39

Joint Intensity Transformer Network for Gait Recognition against Clothing and Carrying status

- Carrying status and clothing appear in various spatial positions

Carrying status

Clothing



40

Joint intensity and spatial metric learning [CVPR 2017]

Joint Intensity and Spatial Metric Learning for Robust Gait Recognition

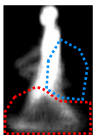
- General dissimilarity measure

$$D(\mathbf{v}^P, \mathbf{v}^G) = \sum_{i=1}^{N_S} w_{S,i}\,\eta(v_i^P, v_i^G)$$

N_S: #Pixels

i: Position index

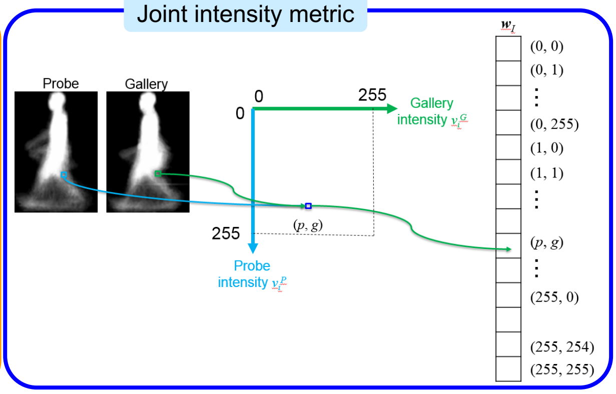
Spatial metric



Want to enhance weight: $w_{S,i} > 1$

Want to reduce weight: $w_{S,i} < 1$

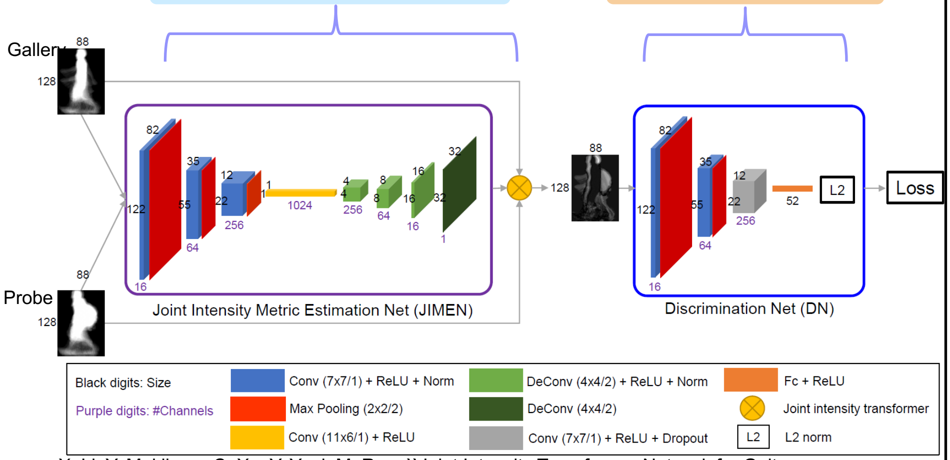
Joint intensity metric



41

Proposed method

- Joint intensity transformer network (JITN)

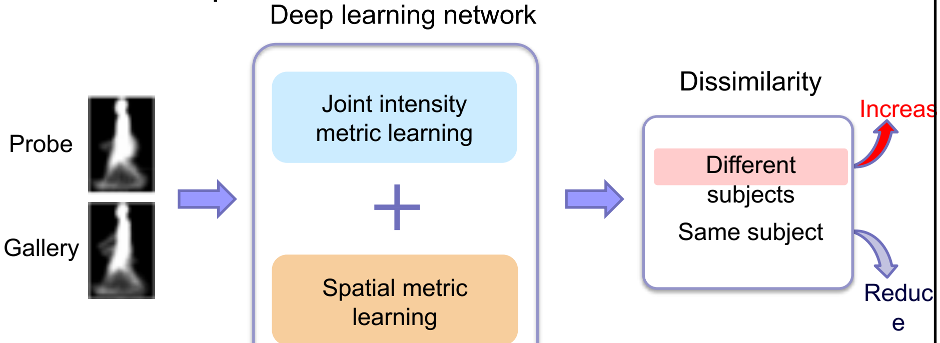


X. Li, Y. Makihara, C. Xu, Y. Yagi, M. Ren, ``Joint Intensity Transformer Network for Gait Recognition Robust against Clothing and Carrying status," IEEE Transactions on Information Forensics and Security, Vol. 14, No. 12, pp. 3102-3115, Dec. 2019

42

Joint Intensity Transformer Network

- Extend joint intensity and spatial metric learning method using deep learning-based techniques



X. Li, Y. Makihara, C. Xu, Y. Yagi, M. Ren, ``Joint Intensity Transformer Network for Gait Recognition Robust against Clothing and Carrying status,'' IEEE Transactions on Information Forensics and Security, Vol. 14, No. 12, pp. 3102-3115, Dec. 2019. [open access]

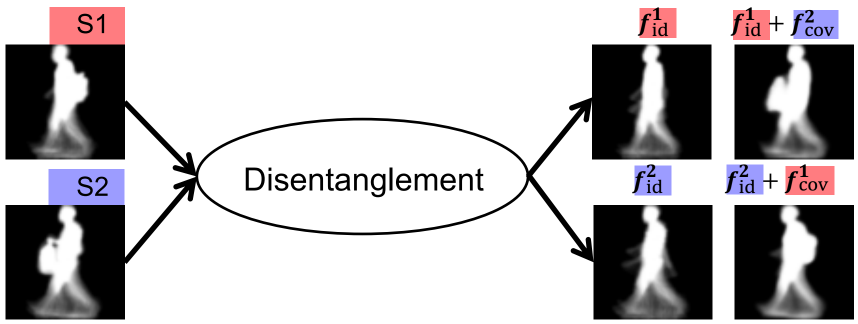
43

Gait Recognition via Semi-supervised Disentangled Representation Learning to Identity and Covariate Features

X. Li, Y. Makihara, C. Xu, Y. Yagi, M. Ren

CVPR SEATTLE WASHINGTON JUNE 16-18 2020

- Use DRL to directly separate identity and covariate features ($\boldsymbol{f}_{\rm id}, \boldsymbol{f}_{\rm cov}$) from silhouette-based gait representations



✓ Purified identity features for better recognition
✓ Transfer covariate features from one subject to another (GEI editing)

44

Comparison on OU-LP-Bag

	Cooperative setting		Uncooperative setting	
Methods	Rank1 [%]	EER [%]	Rank1 [%]	EER [%]
DM	17.74	18.46	15.90	29.89
GEI w/ LDA	40.79	7.35	31.44	14.40
GEI w/ RSVM	24.66	9.58	18.28	14.69
GERF	38.48	7.97	31.24	11.35
GEINet	22.26	11.29	18.52	14.68
SIAME	49.80	2.17	50.27	2.22
LB	74.39	1.68	70.53	1.66
2diff / diff	73.14	1.36	72.75	1.35
Proposed method	74.44	1.25	74.03	1.25

Recent Studies against Carrying status

- X. Li, Y. Makihara, C. Xu, Y. Yagi, M. Ren, ``Gait recognition invariant to carried objects using alpha blending generative adversarial networks," Pattern Recognition, 2020 (Published online: 28 Apr. 2020).

  X. Li, Y. Makihara, C. Xu, Y. Yagi, M. Ren, ``Gait Recognition via Semi-supervised Disentangled Representation Learning to Identity and Covariate Features,'' In the 33rd IEEE Conf. on Computer Vision and Pattern Recognition (CVPR 2020), Seattle, WA, USA, pp. 1-11, Jun. 2020.

ECCV'20 ONLINE
23-28 AUGUST 2020

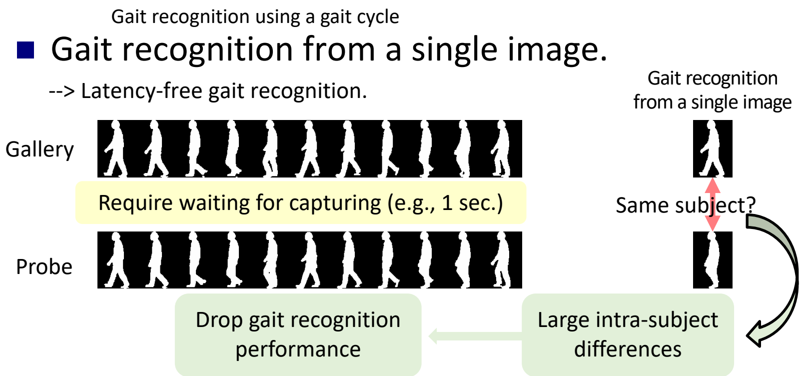
16TH EUROPEAN CONFERENCE ON
COMPUTER VISION
WWW.ECCV2020.EU

Gait Recognition from a Single Image using a Phase-Aware Gait Cycle Reconstruction Network

Chi Xu[1,2], Yasushi Makihara[2], Xiang Li[1,2], Yasushi Yagi[2], and Jianfeng Lu[1]
[1] Nanjing University of Science and Technology
[2] Osaka University

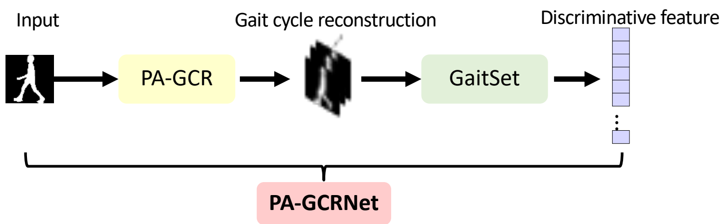
Problem setting

- Gait recognition using a gait cycle.
- Gait recognition from a single image.
  --> Latency-free gait recognition.



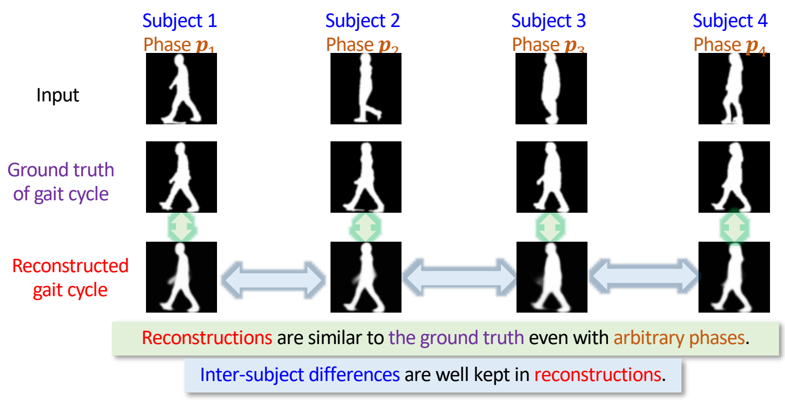
Proposed method

- **PA-GCRNet**: Phase-aware gait cycle reconstruction network
  - An end-to-end framework composed of the two components:
    - PA-GCR: Phase-aware gait cycle reconstructor
    - GaitSet [7]: Recognition network



Reconstruction examples

- Inputs from different subjects with arbitrary phases.



Recognition performance

- Comparison on OU-MVLP.
  - Rank-1 identification rate [%] and equal error rate (EER) [%].

Method	Rank-1	EER
DM	4.4	41.3
GaitSet [7]	***14.0***	***19.6***
PA-GCRNet (proposed)	**80.3**	**1.3**

5.5-times better

15-times better

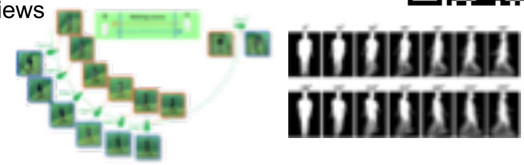
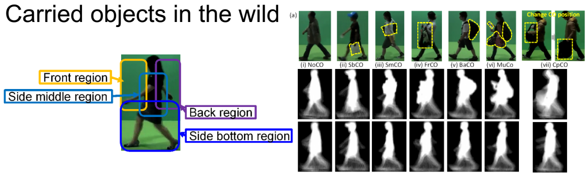
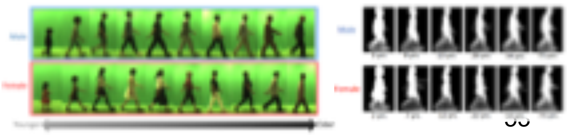
The proposed method significantly outperforms the benchmarks.

List of representative gait databases (DBs)

Name	#Subjects	#Sequences	Covariates	#Viewpoints	Indoor (I) / Outdoor (O)
CMU MoBo	25	600	✔	6	I (Treadmill)
Georgia Tech	15	268	✔	-	O
	18	20	✔	-	-
HID-UMD	25	100		1	O
	55	222	✔	2	O
SOTON Small Database	12	-	✔	3	I
SOTON Large Database	115	2,128	✔	2	I/O
SOTON Multimodal	>300	>5,000	✔	12	I
SOTON Temporal	25	2,280	✔	12	I
USF HumanID	122	1,870	✔	2	O
CASIA A	20	240	✔	3	I
CASIA B	124	13,640	✔	11	I
CASIA C	153	1,530	✔	1	O
CASIA D	88	2640	✔	1	O
OU-ISIR, Treadmill A	34	612	✔	1	I (Treadmill)
OU-ISIR, Treadmill B	68	2,764	✔	1	I (Treadmill)
OU-ISIR, Treadmill C	200	200	✔	25	I (Treadmill)
OU-ISIR, Treadmill D	185	370		1	I (Treadmill)
OU-ISIR, LP	4,007	7,842		2	I
OU-ISIR, LP-Age	63,846	63,846	✔	1	I
OU-ISIR, LP-Bag	62,528	178,018	✔	1	I
OU-ISIR, MVLP	10,307	277,358		14	I
TUM-IITKGP	35	850	✔	1	O
TUM-GAID	305	3,370	✔	1	O
WOSG	155	684	✔	8	O

52

Ad: World's largest gait database

Data set	#Subjects	Covariates
OUMVLP	10,307	14 views 
OULP-Bag	62,528	Carried objects in the wild 
OULP-Age	63,846	Wide age range 

THANKS FOR YOUR ATTENTION

Yasushi YAGI

E-mail : yagi@am.sanken.osaka-u.ac.jp
URL: http://www.am.sanken.osaka-u.ac.jp/