Human Gait Analysis

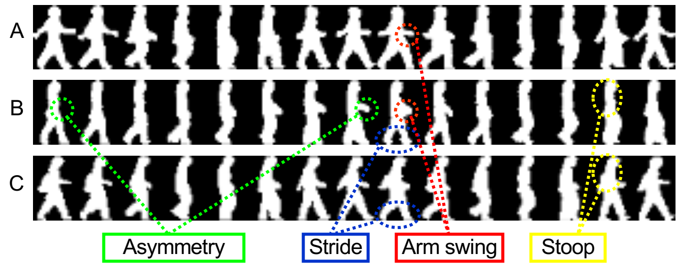
Yasushi Yagi
Executive Vice President
Osaka University, Osaka, Japan

Acknowledgements
-Yagi lab., Osaka University-

- Staff
  - Prof. Yasushi Makihara
  - Prof. Tomio Echigo
  - Prof. Yasuhiro Mukaigawa
  - Prof. Ikuhisa Mitsugami
  - Prof. Daigo Muramatsu
  - Prof. Fumio Okura
  - Dr. Ryusuke Sagawa
  - Dr. Junqiu Wang
  - Dr. Md Altab Hossain
  - Dr. Chunsheng Hua
  - Dr. Al Mansur
  - Dr. Haruyuki Iwama
  - Dr. Rasyid Aqmar
- Students
  - Mr. Kazushige Sugiura
  - Mr. Akira Tsuji
  - Dr. Hidetoshi Mannami
  - Ms. Koko Cho
  - Mr. Atsushi Mori
  - Ms. Mayu Okumura
  - Mr. Akira Shiraishi
  - Mr. Naoki Akae
  - Mr. Yusuke Fujihara
  - Ms. Betria Silvana Rossa
  - Mr. Ryo Kawai
  - Mr. Yuma Higashiyama
  - Mr. Ruochen Liao
  - Mr. Takuya Tanoue
  - Mr. Takuhiro Kimura
  - Mr. Jaemin Son
  - Mr. Xiang Li
  - Ms. Chi Xu

Human gait -Personality- Identity -

Human gait -Personality-



Gait recognition: Person authentication from gait personalities

Example of practical use (1)

- Gait recognition on burglar on CCTVs
  - Admitted as evidence in UK court[1]

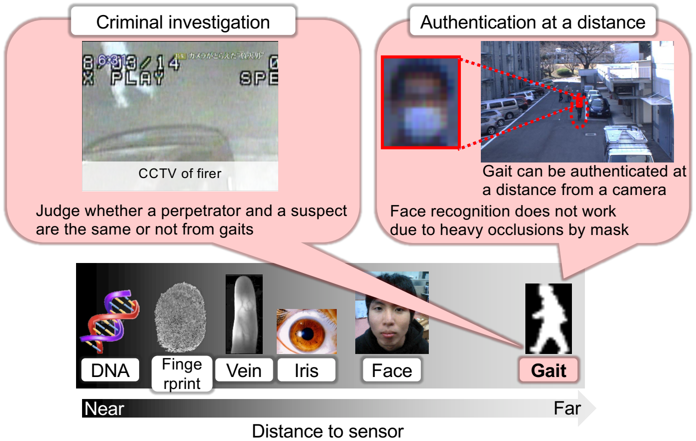
[1] http://news.bbc.co.uk/2/hi/programmes/click online/7702065.stm, "How biometrics could change security," BBC News, 31 Oct. 2008.

Example of practical use (2)

- Gait recognition on firer in Japan[2]

[2] 2009年2月20日 毎日放送 VOICE「指紋は不要？放火犯を追った驚きの科学捜査とは！-歩き方で捕まった放火男」
Mainichi Broadcast VOICE (2009/2/20)

Advantage of gait recognition



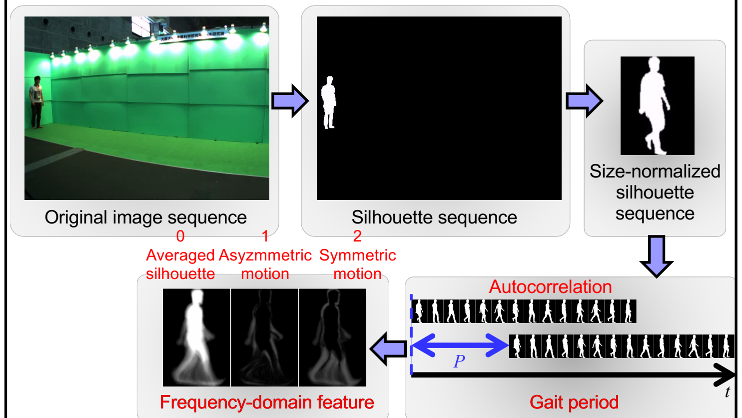
Today's topics

- Gait recognition
- What is the difficulty for applying gait recognition to wide-area surveillance ?
  - The difference of the observation direction
  - Speed change
  - The difference of clothes
  - Low sampling rate
  - Occlusion in crowd scene
- Age estimation
- Gait analysis for innovative entertainment

8

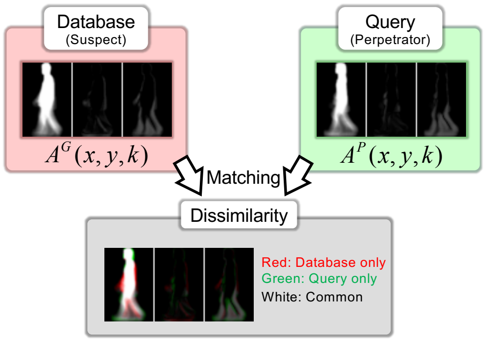
Gait identification & gait verification

Gait feature extraction



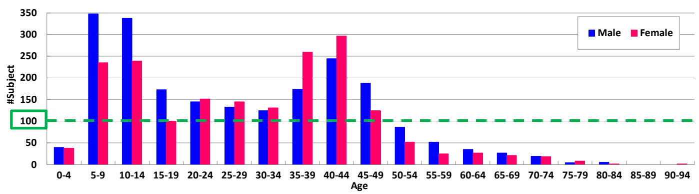
Y. Makihara, R. Sagawa, Y. Mukaigawa, T. Echigo, and Y. Yagi, ``Gait Recognition Using a View Transformation Model in the Frequency Domain,'' 9th European Conf. on Computer Vision, Vol. 3, pp. 151-163, 2006.

Dissimilarity: Single feature



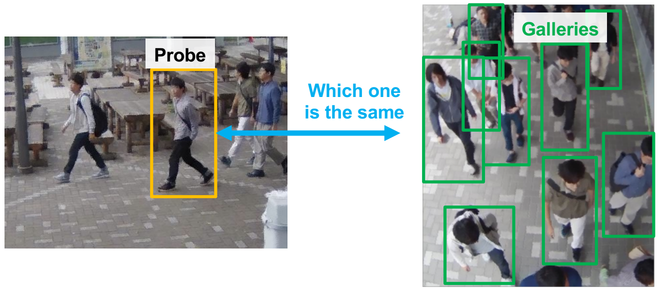
$$t = \sqrt{\sum_{x,y,k} \left(A^G(x,y,k) - A^P(x,y,k)\right)^2}$$

Database: OU-LP



The world largest database with 4,007 subjects (Male: 2,135, Female: 1,872)

Identification (1:N matching)



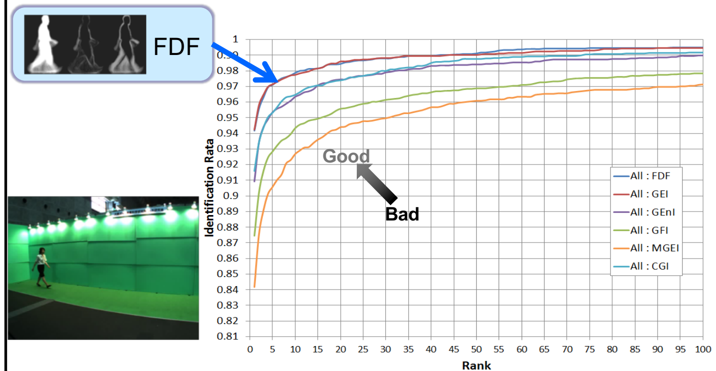
Applications

- Person re-identification
- ID-less access control

8

Performance evaluation: identification

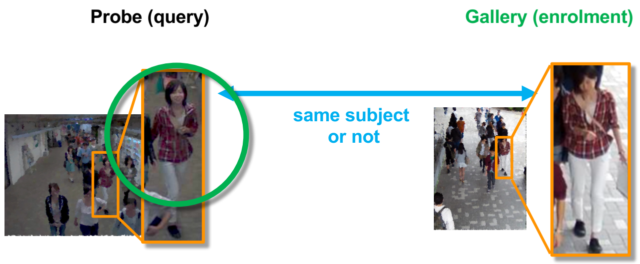
[Iwama et al. IFS 2012]



Cumulative matching characteristics (CMC) curve

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

Verification (1:1 matching)

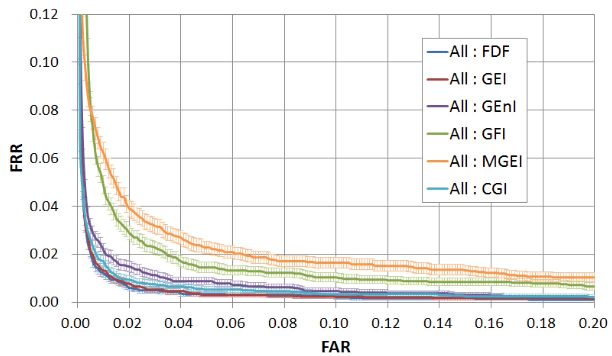


Applications

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

6

Performance evaluation: verification

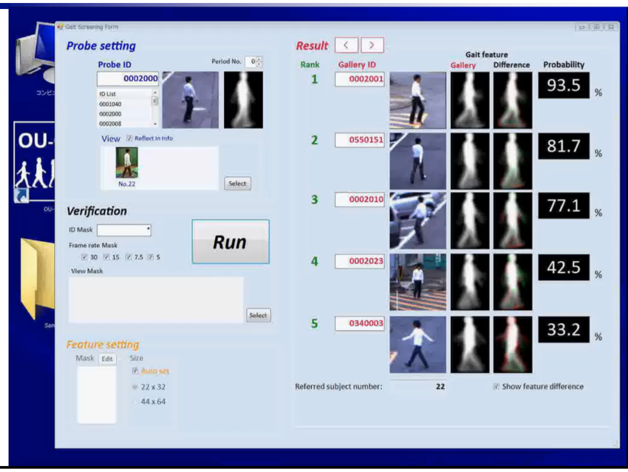


EER: 1.15%

16

World first packaged gait verification system for criminal investigation

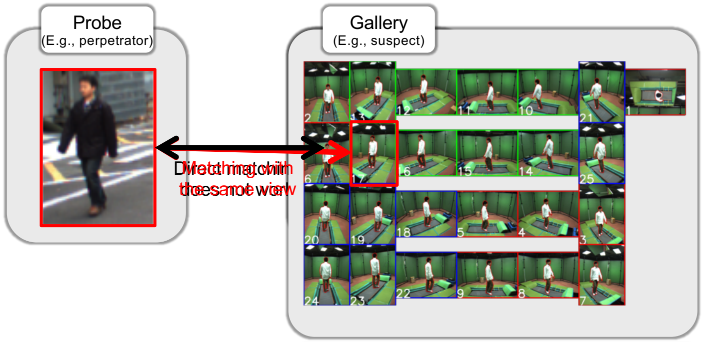
Example of batch verification



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

- The difference of the observation direction
  - ECCV2006
- Speed change CVPR2010
- The difference of clothes
  - Pattern Recognition 2010
- The difference of shoes
- Low sampling rate
  - ACCV2010, IJCB2011,
  - CVPR2012
- Occlusion in crowd scene

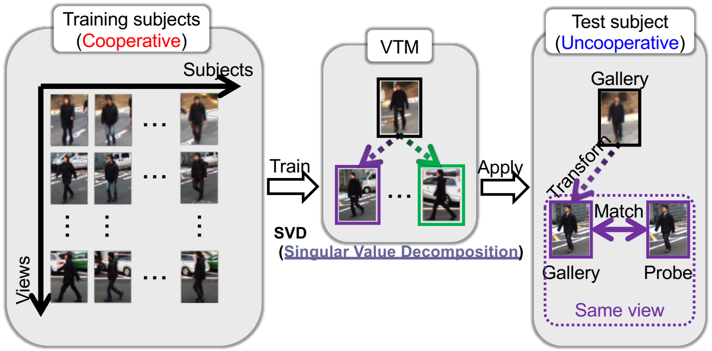
Challenge -View differences-



Difficult to collect multi-view gait features for uncooperative subject

View transformation model (VTM)

[Makihara et al. ECCV 2006]



Formulation of VTM in frequency domain

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

$$\begin{array}{c} \text{individual} \\ \left[\begin{array}{cccc} a_{\theta_1}^1 & a_{\theta_1}^2 & \cdots & a_{\theta_1}^M \\ a_{\theta_2}^1 & a_{\theta_2}^2 & \cdots & a_{\theta_2}^M \\ \vdots & \vdots & \ddots & \vdots \\ a_{\theta_K}^1 & a_{\theta_K}^2 & \cdots & a_{\theta_K}^M \end{array}\right] \end{array} \text{view} = USV^T = \begin{bmatrix} P_{\theta_1} \\ P_{\theta_2} \\ \vdots \\ P_{\theta_K} \end{bmatrix} \begin{bmatrix} v^1 & v^2 & \cdots & v^M \end{bmatrix}$$

Training data matrix

Transformation matrix to each view

View-independent individual vector

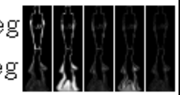
- Gait feature for mth subject from θ_i view

$$a_{\theta_i}^m = P_{\theta_i} v^m$$

View transformation

- From a single reference θ_j to θ_i

$$\left.\begin{array}{l} a_{\theta_i}^m = P_{\theta_i} v^m \\ a_{\theta_j}^m = P_{\theta_j} v^m \end{array}\right\} \Rightarrow a_{\theta_i}^m = P_{\theta_i} P_{\theta_j}^{+} a_{\theta_j}^m$$

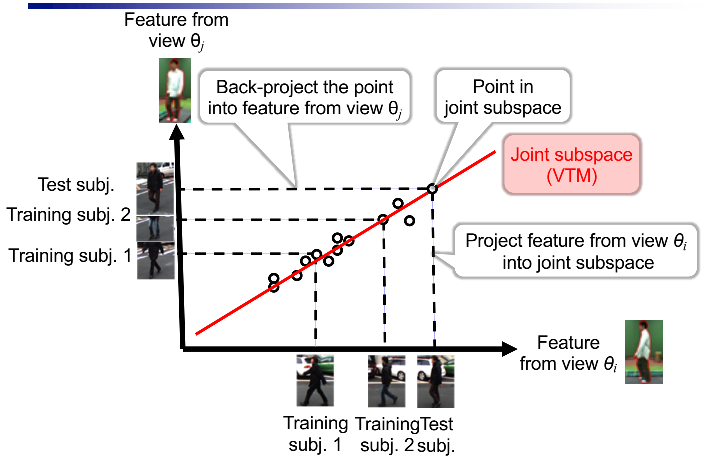


  - Orthogonal motion to reference θ_j is degenerated

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

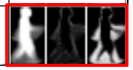
$$\left.\begin{array}{l} a_{\theta_i}^m = P_{\theta_i} v^m \\ a_{\theta_j(1)}^m = P_{\theta_j(1)} v^m \\ \vdots \\ a_{\theta_j(k)}^m = P_{\theta_j(k)} v^m \end{array}\right\} \Rightarrow a_{\theta_i}^m = P_{\theta_i} \begin{bmatrix} P_{\theta_j(1)} \\ \vdots \\ P_{\theta_j(k)} \end{bmatrix}^{+} \begin{bmatrix} a_{\theta_j(1)}^m \\ \vdots \\ a_{\theta_j(k)}^m \end{bmatrix}$$

How does it work?



Transformation results

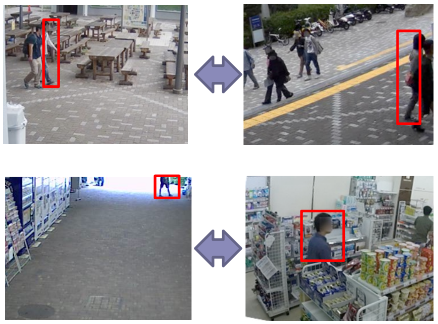
Gallery

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

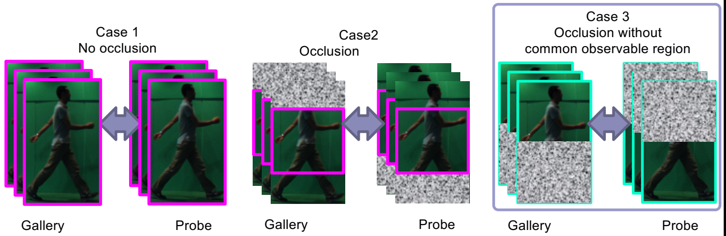
What is the difficulty for applying gait recognition to wide-area surveillance ?

- The difference of the observation direction
  - ECCV2006
- Speed change CVPR2010
- The difference of clothes
  - Pattern Recognition 2010
- Low sampling rate
  - ACCV2010, IJCB2011,
  - CVPR2012
- Occlusion in crowd scene
  - ICB2015

Actual situation of observed gait in surveillance



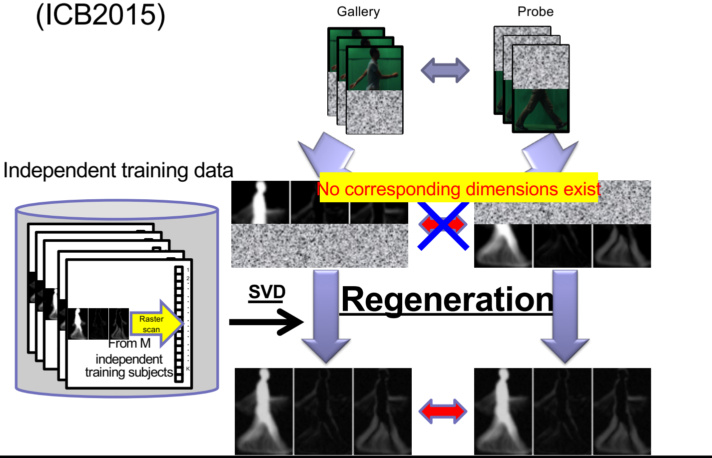
Challenge: Serious occlusion



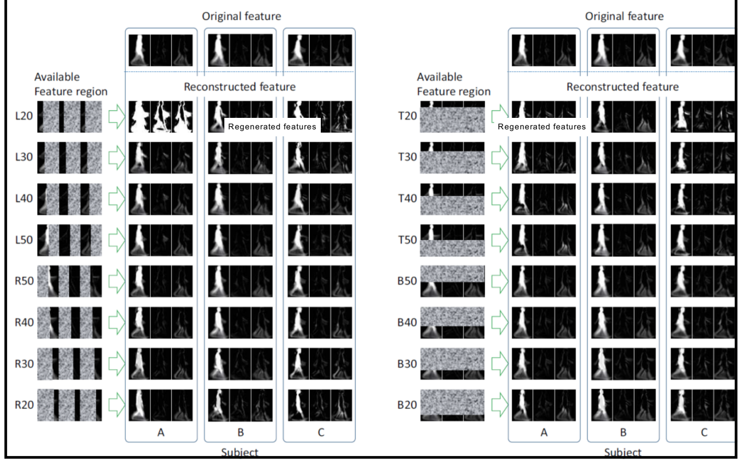
- Common observable regions (COR) are used for recognition
- Direct comparisons are impossible in case 3
  because any common region cannot be observed

Gait regeneration for recognition

(ICB2015)

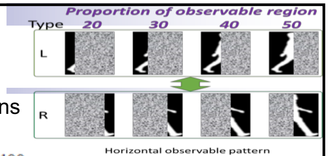


Experiment : Regenerated gait



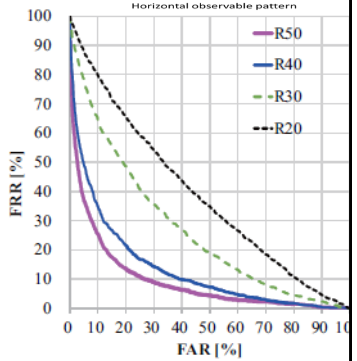
Experiment

Results with horizontal observable patterns



View Angle	Gallery / Probe	L50	L40	L30	L20
85	R50	16.4	23.6	39.3	47
	R40	20.6	25.6	38.9	48
	R30	32.9	34.6	42.2	48
	R20	41.8	41.4	44.4	48

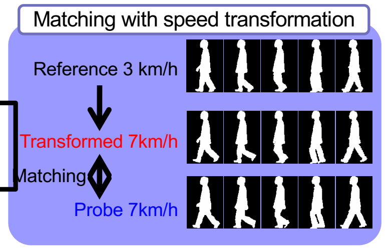
Equal Error Rate



ROC curves of propose method against gait features with view 85 deg where L50 is used for the gallery

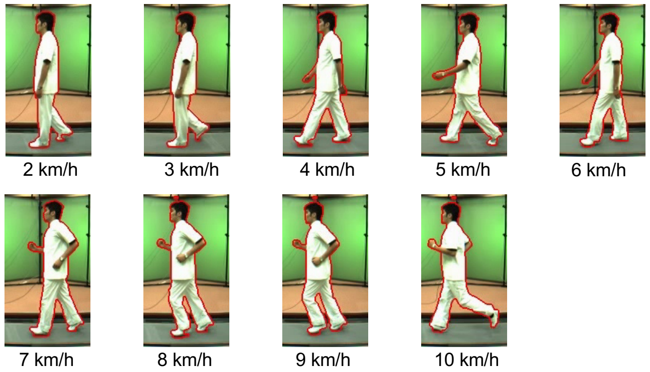
What is the difficulty for applying gait recognition to wide-area surveillance ?

- The difference of the observation direction
  - ECCV2006
- Speed change CVPR2010 ACCV2016
- The difference of clothes
  - Pattern Recognition 2010
- Low sampling rate
  - ACCV2010, IJCB2011,CVPR2012
- Occlusion in crowd scene
  - ICB2015



Speed transformation model

Challenge -Speed difference-

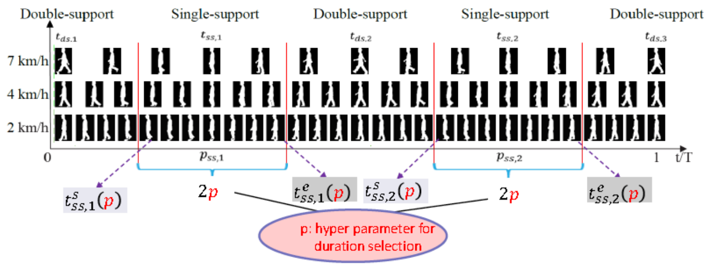


Speed Invariance vs. Stability: Cross-Speed Gait Recognition using Single-Support Gait Energy Image

C. Xu, Y. Makihara, X. Li, Y. Yagi, J. Lu, "Speed Invariance vs. Stability: Cross-Speed Gait Recognition Using Single-Support Gait Energy Image", In *Proc. of the 13th Asian Conf. on Computer Vision (ACCV 2016)*,

Single-Support GEI (SSGEI)

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

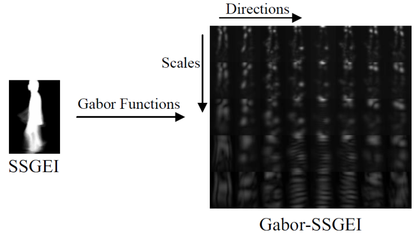


- Representation:

$$S(x, y; p) = \frac{1}{2}\sum_{k=1}^{2} \frac{1}{t^e_{ss,k}(p) - t^s_{ss,k}(p) + 1} \sum_{t=t^s_{ss,k}(p)}^{t^e_{ss,k}(p)} I(x, y, t), \qquad (0 < p \le {}^1/_4).$$

Post-process

- Gabor filtering [Tao et al. 2007]
  - Gabor kernel functions of 8 directions and 5 scales.



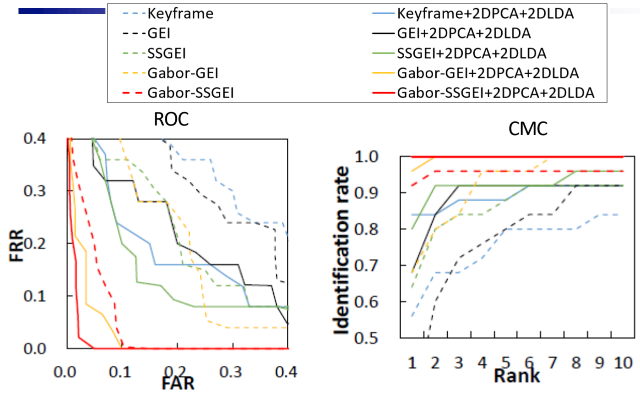
- Metric Learning
  - 2DPCA [Yang et al. 2004]: Reduce feature dimension.
  - 2DLDA [Li et al. 2005]: Achieve optimal discrimination capability.

Experiments: dataset 1

- **OU-ISIR Treadmill Dataset A** [Makihara et al. 2012]
  - Speed variation: 2 km/h ~ 7 km/h (walking)
  - Training set: 9 subjects, testing set: 25 subjects

Contains the largest speed variations.

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



The propose method achieves the best accuracy.

Experiments: Comparison with state-of-the-arts

- Rank-1 identification rate [%] in case of small and large speed changes.

Speed change	HMM [Liu et al. 2006]	SN [Tanawongsuwan and Bobick. 2004]	STM [Tsuji et al. 2010]	DCM [Kusakunniran et al. 2012]	RSM	Proposed method
Small (3 km/h and 4 km/h)	84	-	90	98	**100**	**100**
Large (2 km/h and 6 km/h)	-	35	58	82	***95***	**98**

- Averaged rank-1 identification rates [%] over 36 combinations of walking speeds of DCM, RSM and proposed method.

Algorithms	Rank-1 identification rate
DCM	92.44
RSM	***98.07***
Proposed method	**99.33**

The proposed method clearly outperforms the other algorithms, in particular in case of large speed changes.

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

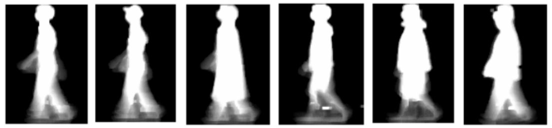
- The difference of the observation direction
  - ECCV2006
- Speed change CVPR2010
- The difference of clothes
  - Pattern Recognition 2010, ACCV2016
- Low sampling rate
  - ACCV2010, IJCB2011,
  - CVPR2012
- Occlusion in crowd scene
  - ICB2015

Gait energy response function for clothes-invariant gait recognition

X. Li, Y. Makihara, C. Xu, D. Muramatsu, Y. Yagi, M. Ren, "Gait Energy Response Function for Clothing-invariant Gait Recognition", In *Proc. of the 13th Asian Conf. on Computer Vision (ACCV 2016)*

Background

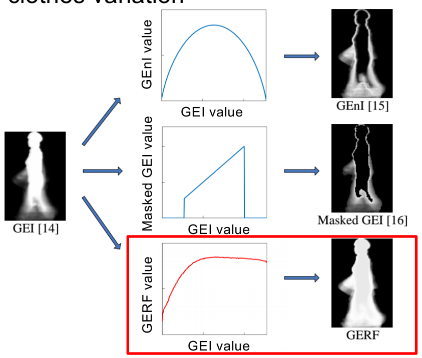
- Gait recognition
  - Pros:
    - Availability at a distance for an uncooperative subject (c.f. face, iris)
  - Cons:
    - Accuracy drop due to many covariates (e.g., clothing, view, speed)



Gait energy images [Han and Bhanu 2006] under clothing variations

Objective

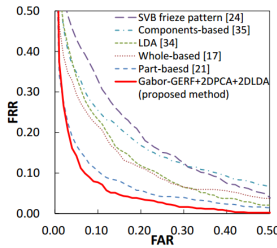
- Transform GEI into more discriminative feature under clothes variation



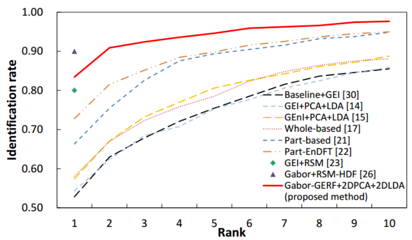
A sort of gait energy transformation process via a gait energy response function (GERF)

Comparison with state-of-the-arts methods

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



ROC curves

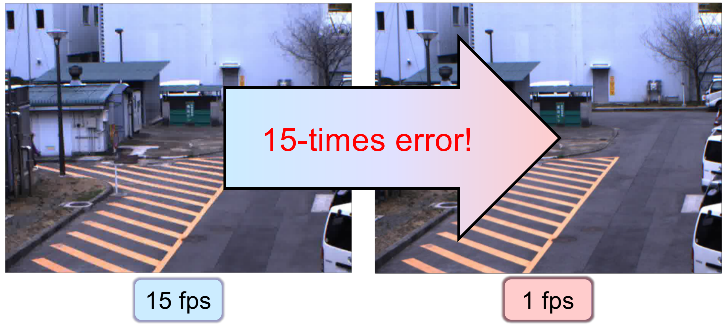


CMC curves

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

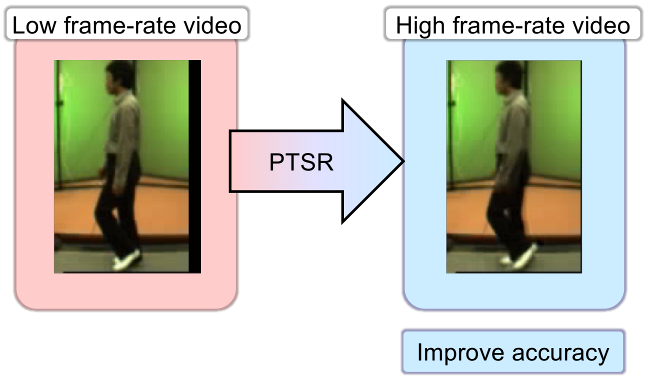
- The difference of the observation direction
  - ECCV2006
- Speed change CVPR2010
- The difference of clothes
  - Pattern Recognition 2010
- The difference of shoes
- Low sampling rate
  - ACCV2010, IJCB2011,
  - CVPR2012
- Occlusion in crowd scene
  - ICB2015

Challenge -Low frame-rate-



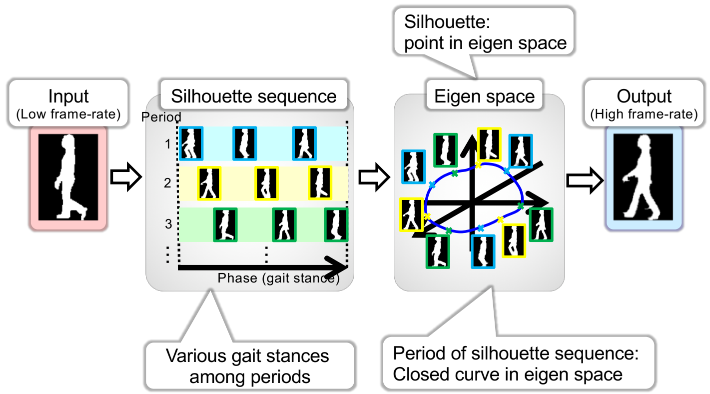
Solution

- Periodic Temporal Super Resolution (PTSR)

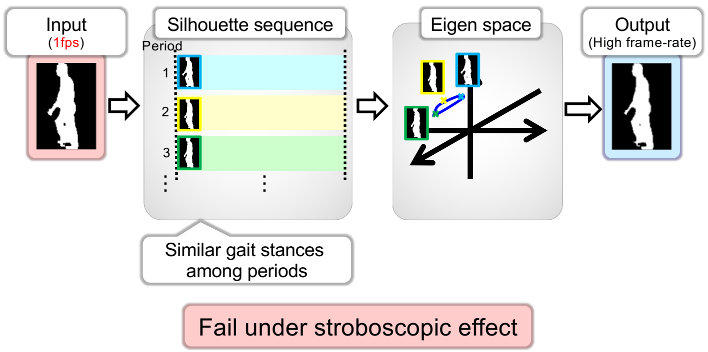


Reconstruction-based PTSR

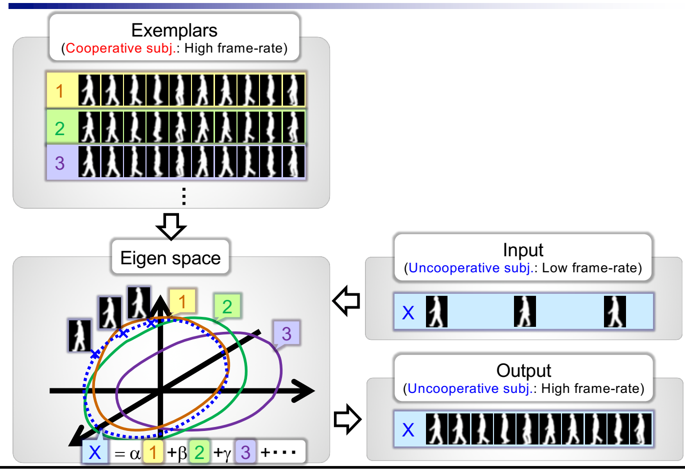
-Overview- [Makihara et al. ACCV 2010]



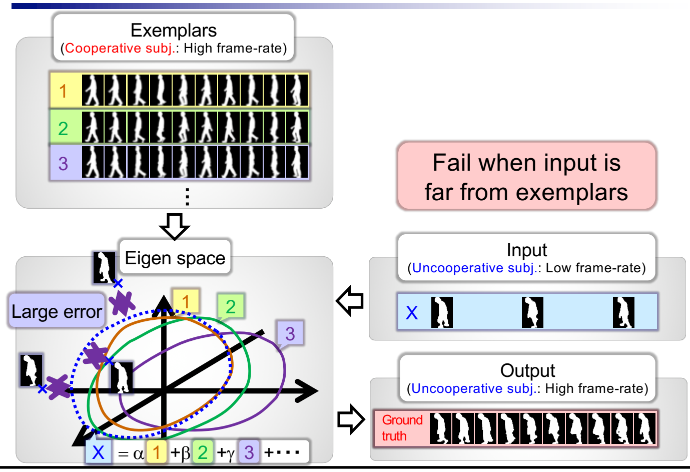
Reconstruction-based PTSR -Failure mode-



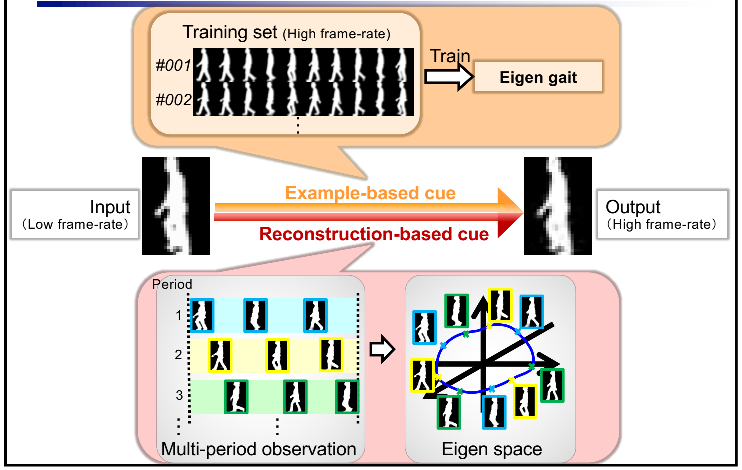
Example-based PTSR -Overview-



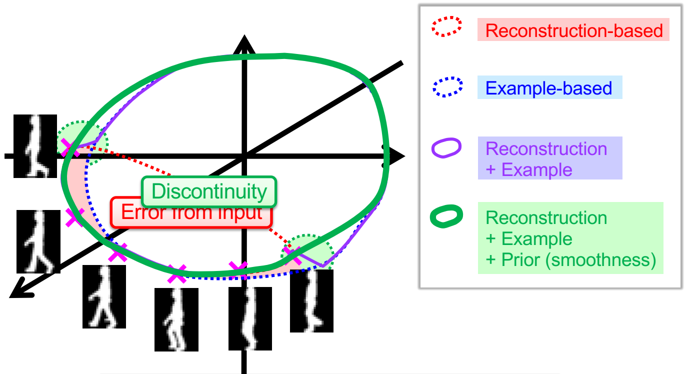
Example-based PTSR -Failure mode-



Unified approach to PTSR [Akae et al. CVPR 2012]

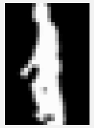
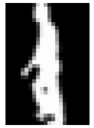
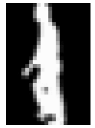
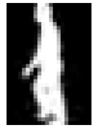


How does it work?

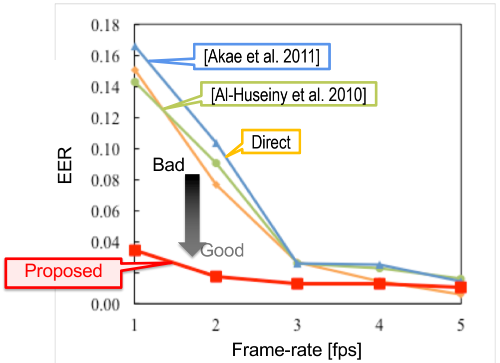


Solve by energy minimization framework

PTSR results -1 fps-

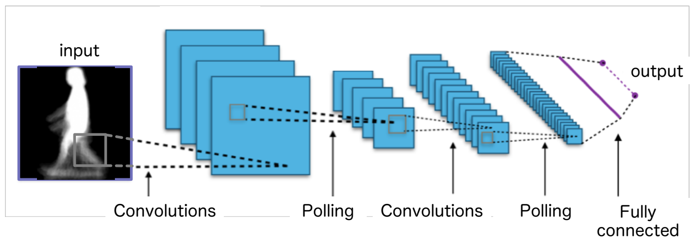
Frame-rate of input	Input	[Al-Huseiny et al. 2010]	[Akae et al. 2011]	Proposed
1 fps				

Performance evaluation: Verification



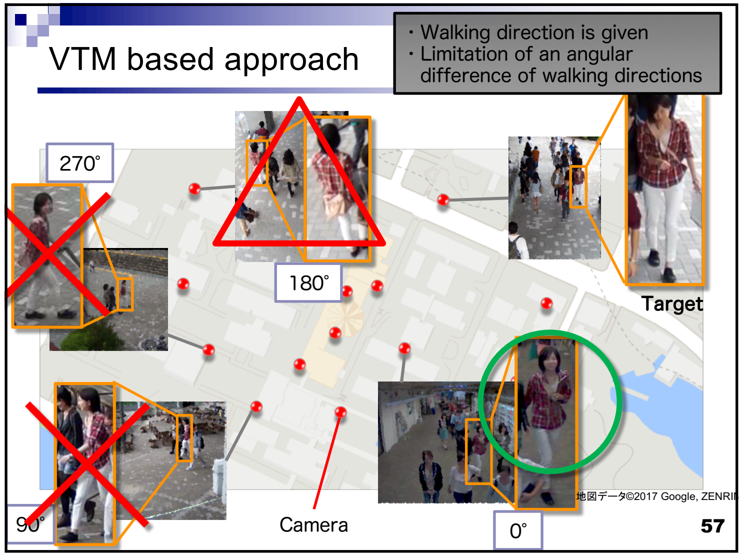
CNN-based discriminative approach

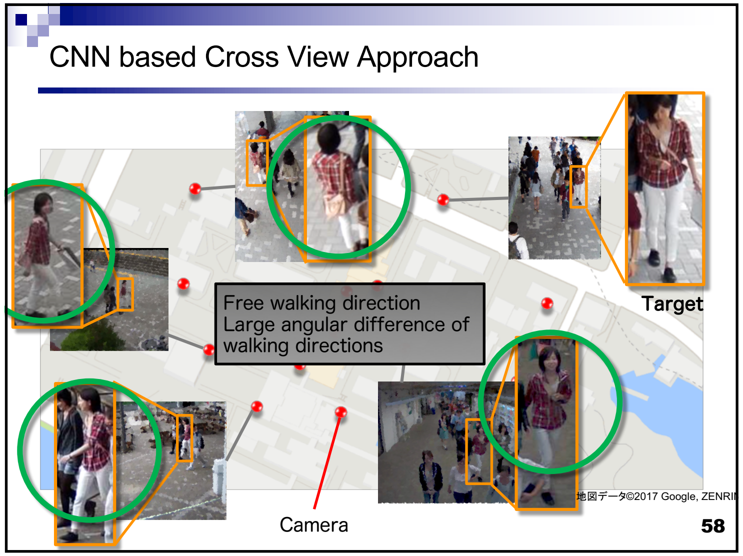
Convolutional Neural Network (CNN)-based gait recognition

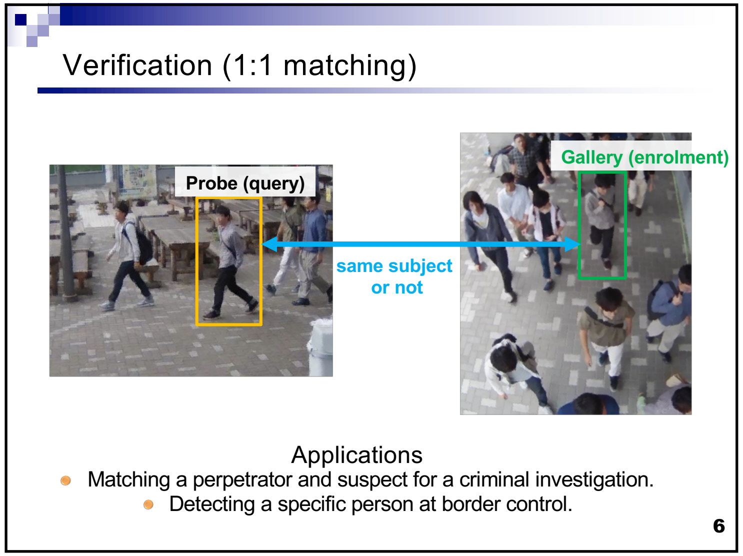


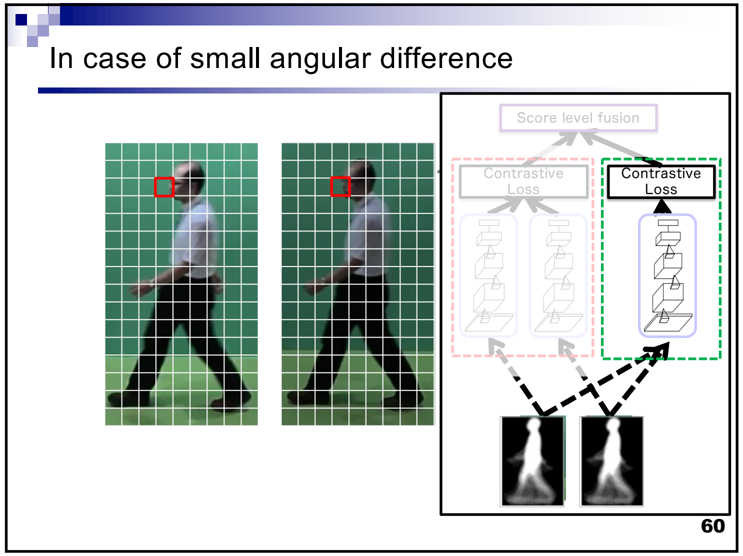
- CNN-based methods have achieved state-of-the-art performance.
- Network architectures can be designed flexibly.

56

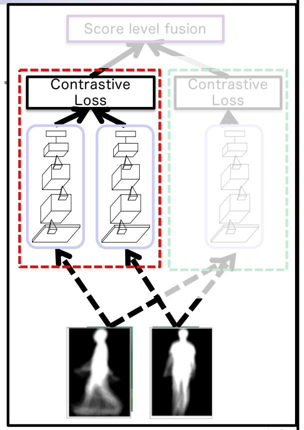

VTM based approach
・Walking direction is given
・Limitation of an angular difference of walking directions
270°
180°
Target
90°
Camera
0°
地図データ©2017 Google, ZENRIN
57


CNN based Cross View Approach
Free walking direction
Large angular difference of walking directions
Target
Camera
地図データ©2017 Google, ZENRIN
58


Verification (1:1 matching)
Probe (query)
Gallery (enrolment)
same subject or not
Applications
Matching a perpetrator and suspect for a criminal investigation.
Detecting a specific person at border control.
6

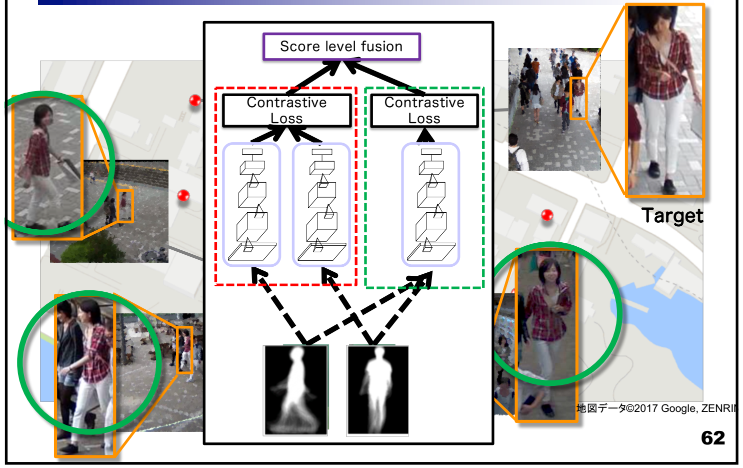

In case of small angular difference
Score level fusion
Contrastive Loss
Contrastive Loss
60

In case of large angular difference



61

CNN based Cross View Approach

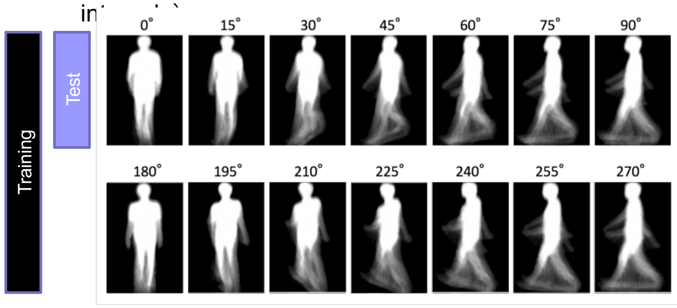


62

Gait database

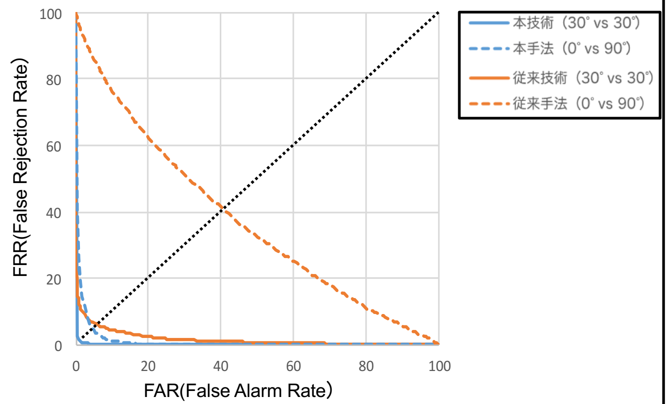
OU-MVLP（OU-ISIR Multi-View Large Population）

- Gait feature: GEI (Gait energy image)
- #Subjects：about 10,000 (training : testing = 1 : 1)
- View variation：14 views（0-90°，180-270°，15° - in

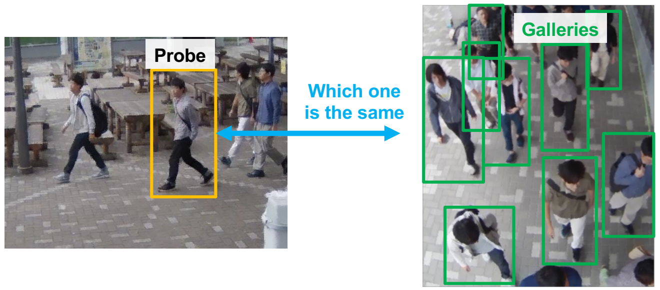


63

ROC curve of new combined CNN model

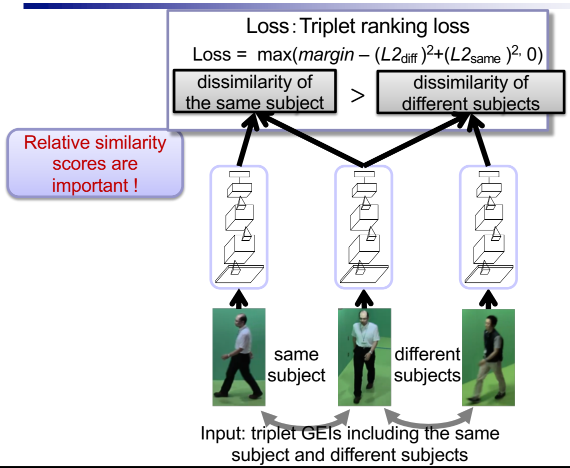


Identification (1:N matching)

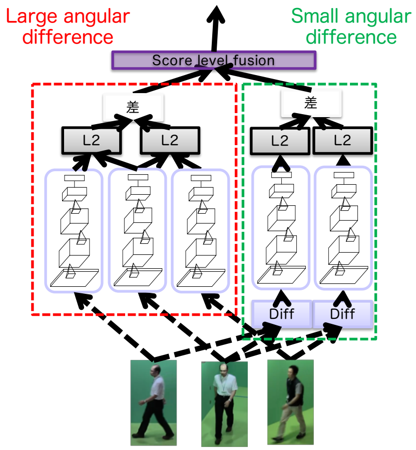


Applications
- Person re-identification
- ID-less access control

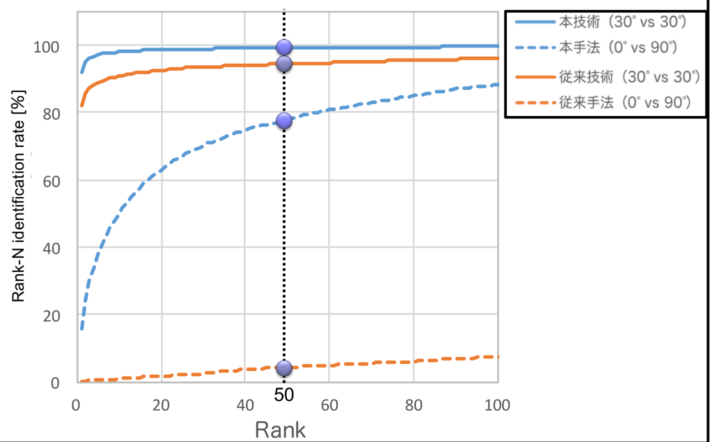
Network architecture for identification



CNN based gait identification



Rank-N identification rate（Subject about 5,000）



Benchmarks

- Direct matching
  - DM (L2 distance of two GEIs as the dissimilarity)
- Generative approach
  - VTM (View transformation model) [ECCV2006]
- Discriminative approach
  - LDA (Linear discriminant analysis) [ICB2014]
- CNN-based discriminative approach
  - GEINet [IEICE2016]
  - MT (Mid-level@Top) [TPAMI2016]
  - LB (Local@Bottom) [TPAMI2016]

16

Evaluation criteria

- For verification

  Equal error rate (EER)

  EER of false acceptance rates (FARs) and false rejection rates (FRRs) .

- For identification

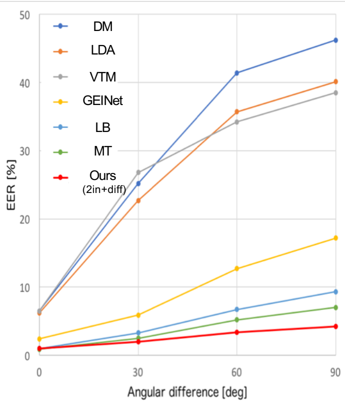
  Rank-1 identification rate

  Rate of hitting the best matching as the correct matching.

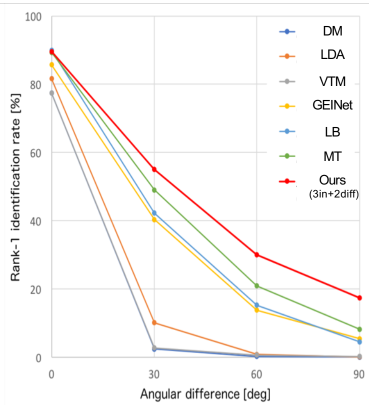
15

Comparison with benchmarks

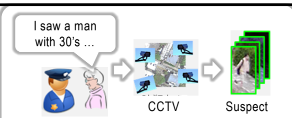
Verification



Identification



Age-dependent model for attention/intention understanding

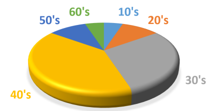


Criminal investigation (Find suspect with witness)

Gait-based age estimation (AIU2018 Dec. 3rd, 2018)

Surveillance (Find wandering elderly)

Access control (Age restriction)

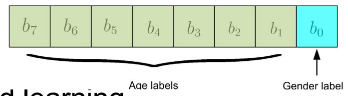


Market research (Customer counting)

72

Related work

- Regression
  - Gaussian process regression (GPR) [Makihara+ 2011, 2016]
  - Support vector regression (SVR) [Guo+ 2008]
- Classification
  - Multi-label-guided (MLG) subspace analysis [Lu+ 2010]



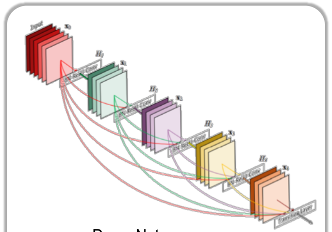
- Manifold learning
  - Ordinary preserving manifold analysis [Lu+ 2013]
  - Orthogonal locality preserving projections [Li+ 2018]

Few deep learning approaches to gait-based age estimation

73

Objective

- Validate the effectiveness of deep learning approaches to gait-based age estimation



DenseNet [Huang+ 2017]
State-of-the-art network architecture

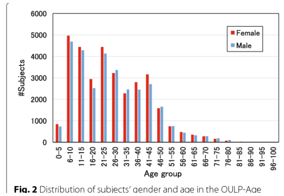


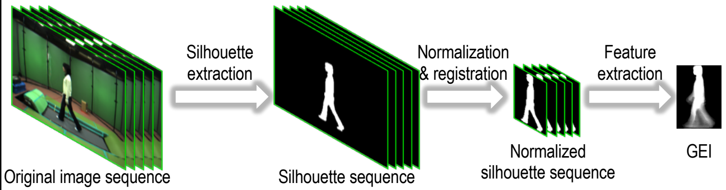
Fig. 2 Distribution of subjects' gender and age in the OULP-Age

OULP-Age [Xu+ 2017]
The world largest database (+60,000 subj)

74

Input data

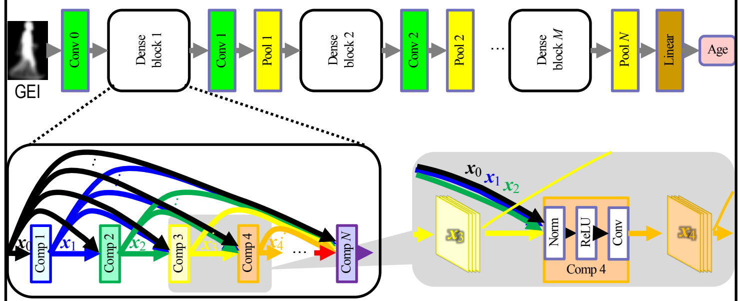
- Gait energy image (GEI) [Han+ 2006]
  - Simple yet effective gait feature
  - The most widely used in gait analysis community



75

Network architecture

- Densely connected convolutional network (DenseNet) [Huang+ 2017]
  - State-of-the-art performance in many computer vision tasks
  - Utilize skip connection from all preceding layers



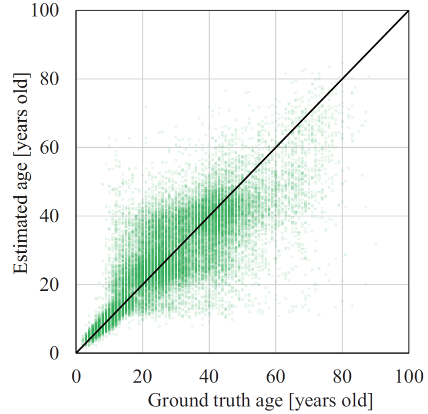
76

Experiments: Data set

- The OU-ISIR Gait Database, Large Population Dataset with Age (OULP-Age) [Xu+ 2017]
  - 63,846 subjects (31,093 males and 32,753 females)
  - Age range: 2 to 90 years old
  - Protocol
    - Training set : 31,923 subjects
    - Testing set: 31,923 subjects

77

Experiments: Scatter plots



78

Experiments: Mean absolute error

$$\mathrm{MAE} = \frac{1}{N}\sum_{k=1}^{N}\left|\hat{l}_k^t - l_k^t\right|$$

#Test samples
Estimated age
Ground truth age

		MAE [years]
Non-deep learning	MLG	10.98
	GPR (K=10)	8.83
	GPR (K=100)	7.94
	GPR (K=1000)	7.3
	SVR (linear)	8.73
	SVR (Gaussian)	7.66
	OPLDA	8.45
	OPMFA	9.08
	AGDMLR	6.78
Deep learning	GEINet	6.22
	Proposed	5.79

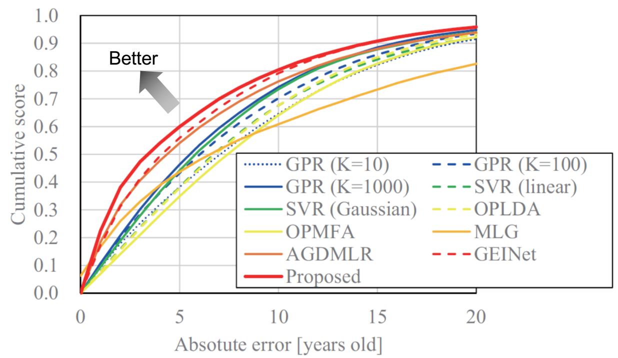
Better

79

Experiments: Cumulative score

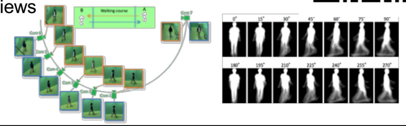
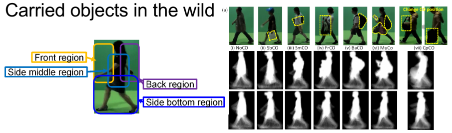
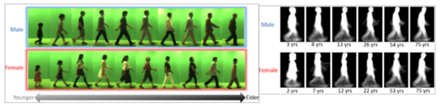
$$CS(j) = \frac{N_{e\le j}}{N}$$

#Test samples whose absolute errors are lower than j years
#Test samples



80

Ad: World's largest gait database

Data set	#Subjects	Covariates
OUMVLP [Takemura+ 2017]	10,307	14 views 
OULP-Bag [Uddin + 2018]	62,528	Carried objects in the wild 
OULP-Age [Xu+ 2018]	63,846	Wide age range 

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

OU-ISIR Gait Database

http://www.am.sanken.osaka-u.ac.jp/BiometricDB/index.html

- *The OU-ISIR Gait Database*
  - *Treadmill Dataset*
    - **dataset A -Speed variation-**
    - **dataset B -Clothes variation-**
    - **dataset C -view variation-**
    - **dataset D -Gait fluctuation-**
  - *Large Population Dataset (4,016 subjects)*
  - *Large Population Dataset with Bag* (62,528 subjects)
  - *Large Population Dataset with Age* (63,846 subjects)
  - *Large Population Multi-view Population Dataset* (*10,307* subjects)
  - *Inertial Sensor Dataset*
  - *Similar Actions Inertial Dataset*
- *The OU-ISIR Biometric Score Database*

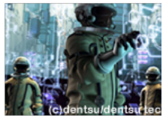
Please note that each corresponding signed release agreement is required to get the access to each dataset, i.e., required release agreements are different among datasets.

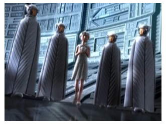
Gait Analysis for Innovative Entertainment

Dive Into the Movie

"Dive into the Movie (DIM)" is a name of project to aim to realize a world innovative entertainment system which can provide an immersion experience into the story by giving a chance to audience to share an impression with his family or friends by watching a movie in which all audience can participate in the story as movie casts.

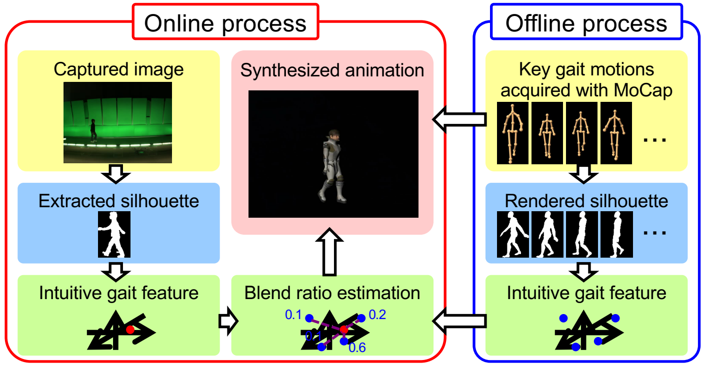
To realize this system, we are trying to model and capture the personal characteristics instantly and precisely in face, body, gait, hair and voice.





Collaborated with
Waseda University (Prof. Morishima)
Advanced Telecommunications Research Institute International (ATR). (Dr Nakamura, NAIST)

Online measurement of intuitive gait feature for digital entertainment [3][4]



[3] M. Okumura, Y. Makihara, S. Nakamura, S. Morishima, and Y. Yagi, ``The Online Gait Measurement for the Audience-Participant Digital Entertainment," Proc. of Invited Workshop on Vision Based Human Modeling and Synthesis in Motion and Expression, No. 5, pp. 1-10, Xi'an, China, Sep. 2009.
[4] Y. Makihara, M. Okumura, Y. Yagi, and S. Morishima, ``The Online Gait Measurement for Characteristic Gait Animation Synthesis," Proc. of Human Computer Interaction International 2011, Virtual and Mixed Reality - New Trends, vol. 6773, pp.325--334, Springer, Orlando, FL, USA, Jul. 2011.

Digital Entertainment

THANKS FOR YOUR ATTENTION

Yasushi YAGI

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