Person Re-identification:

—— Recent Challenges

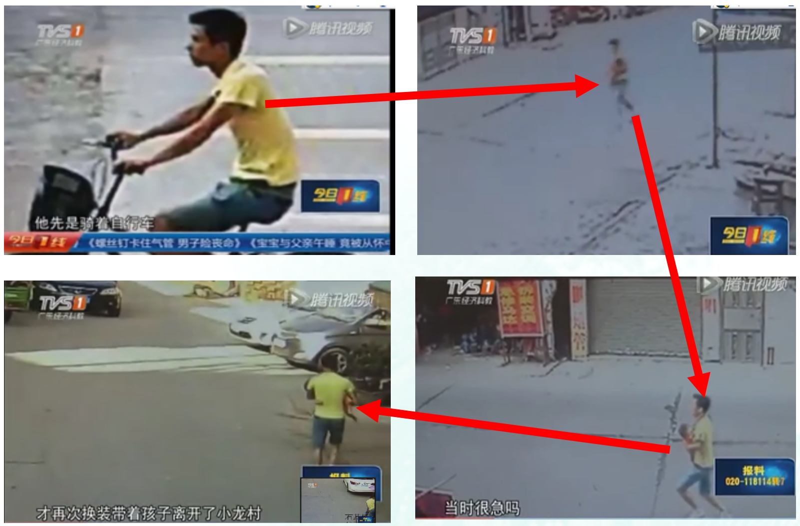
Wei-Shi Zheng (郑伟诗)

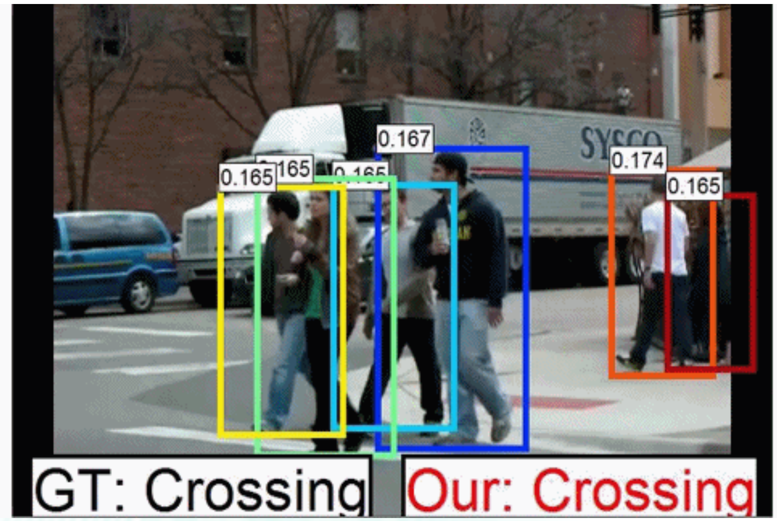
http://isee.sysu.edu.cn/~zhwshi

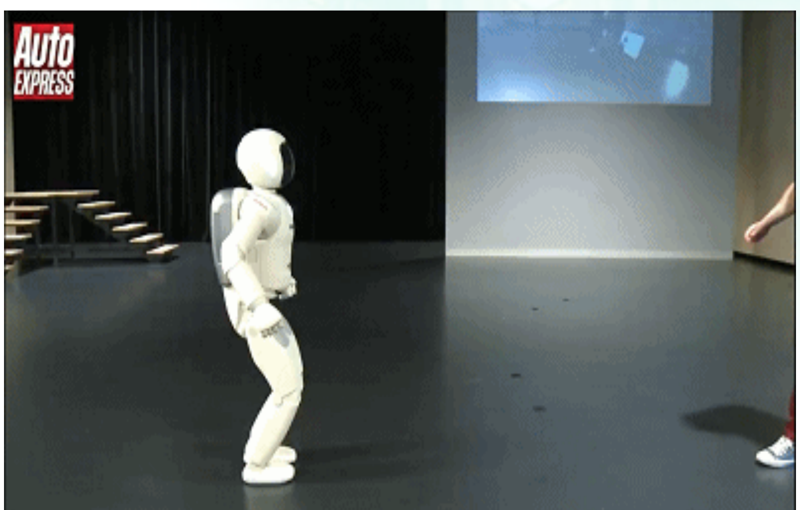
Sun Yat-sen University

机器智能与先进计算
教育部重点实验室

My Research







面向视觉对象关联的机器学习建模

Background

The whole story

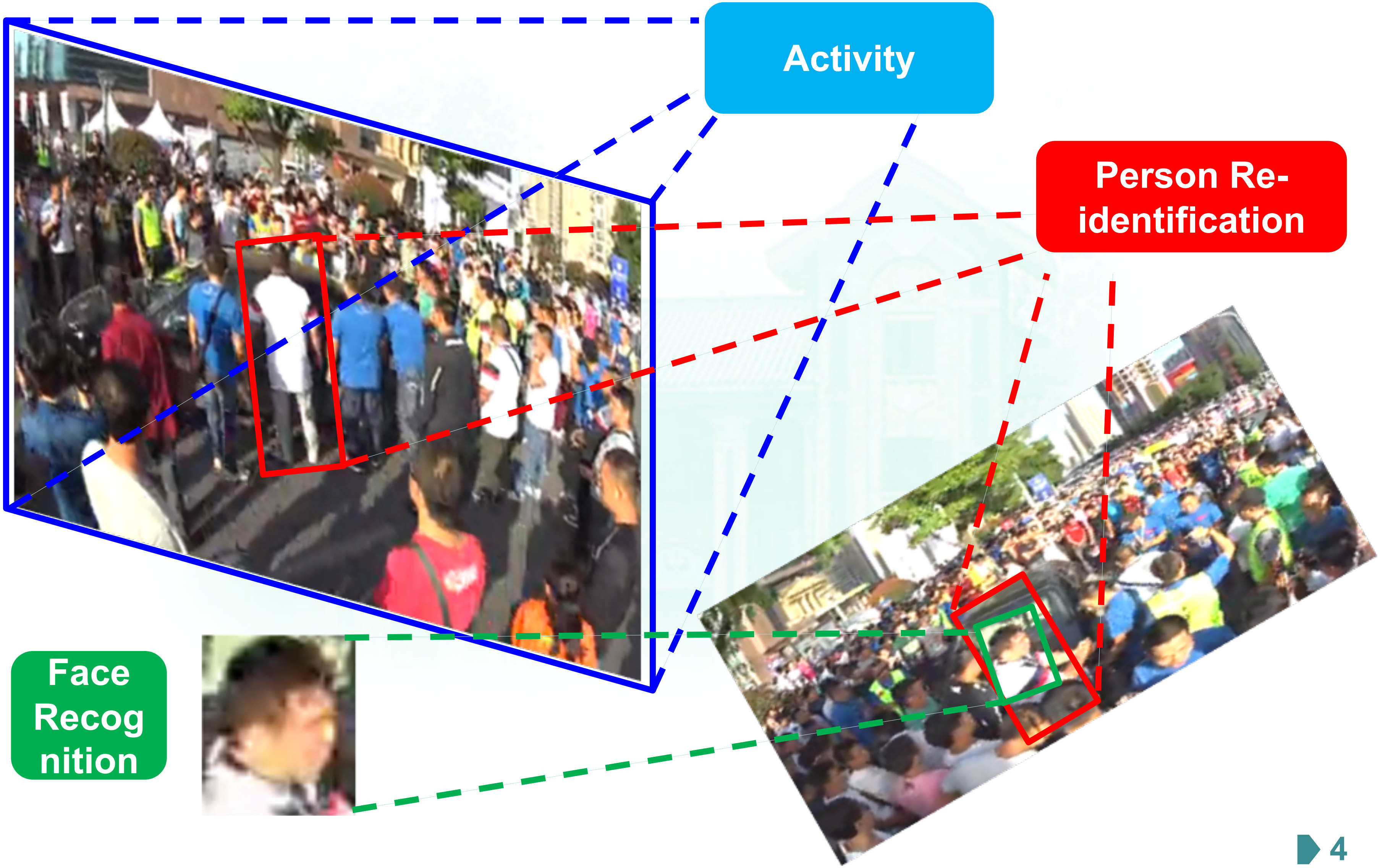
1) Detect an event

2) track persons across camera view

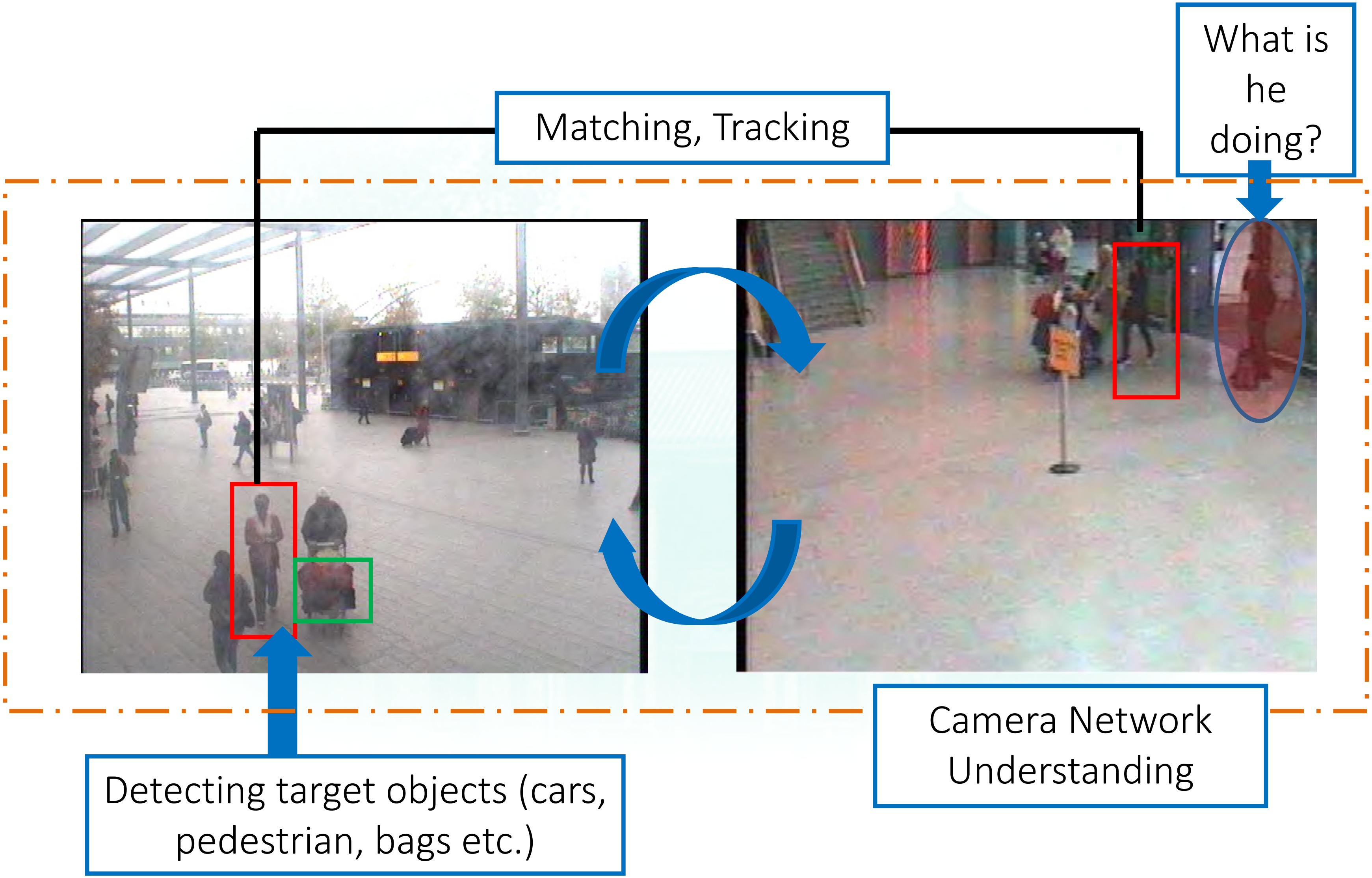
3) Identify who he/she is

Human Identification & Activity Understanding

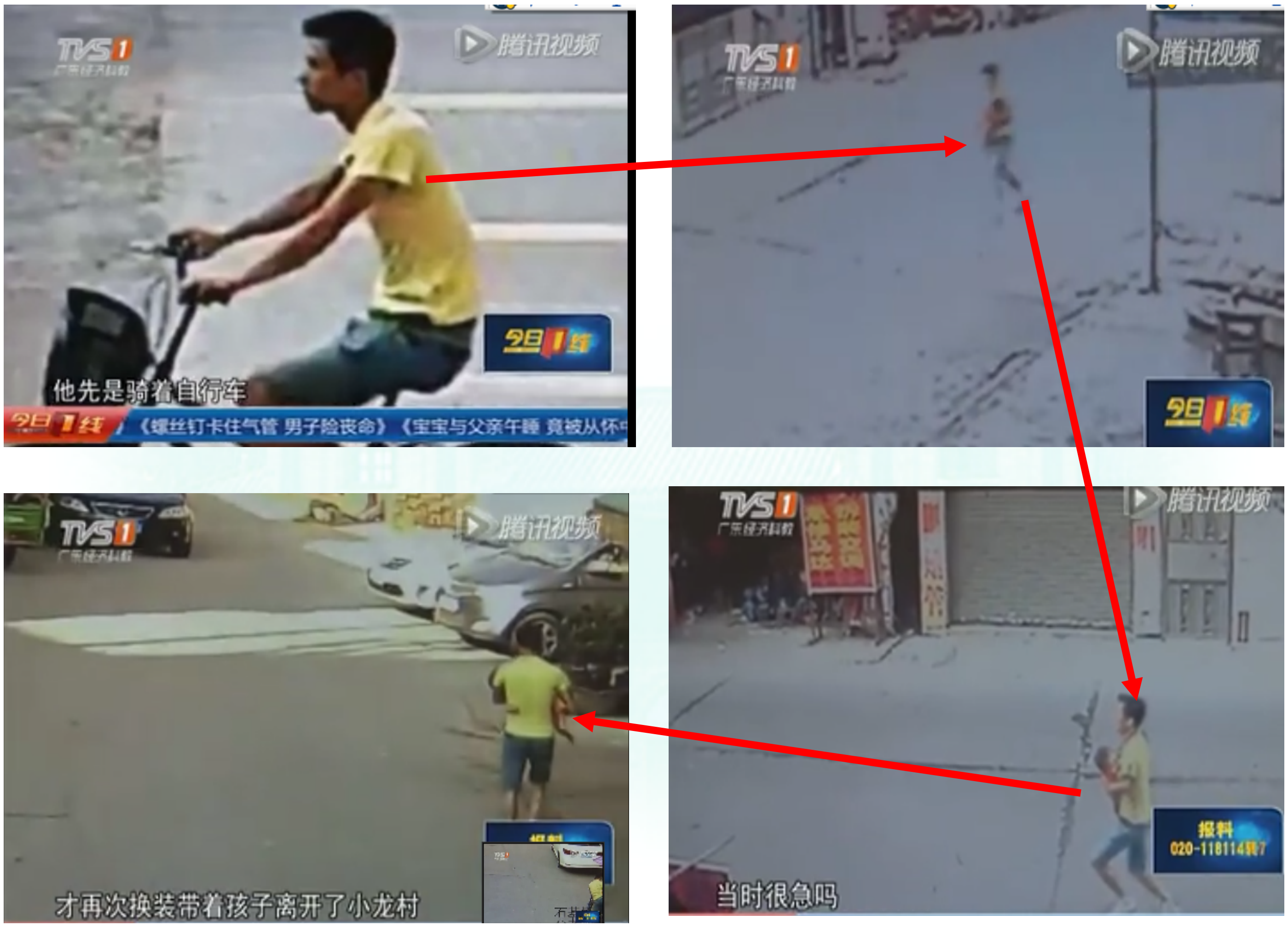
❑ Background



Person Re-identification



Person Re-identification

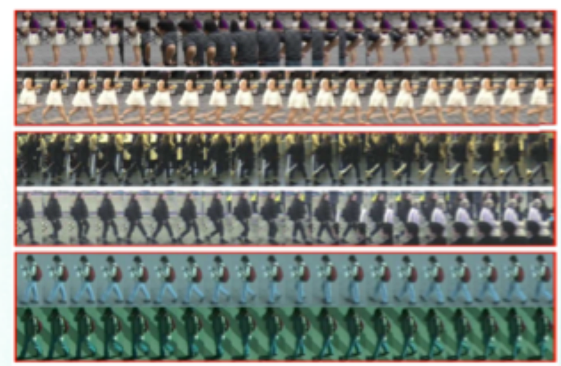


Recent Development & Question

- Pose-guided, Local, Attention-based, GAN-based, ……
[a ppt: https://share.weiyun.com/5VPtcZa]
- What should we do? I would guess we will soon have 99% matching rate this year or early next year on benchmarks

Dataset	Reference	Rank-1	mAP	
Market-1501	Harmonious Attention Network for Person Re-Identification	91.2	75.7	Single query
		93.8	82.8	Multiple queries
DukeMTMC	Learning Discriminative Features with Multiple Granularities for Person Re-Identification	88.7	78.4	

- Video-based RE-ID datasets: MARS, iLIDS-VID, PRID 2011.



- The state-of-the-art performance (Rank-1):

	MARS	iLIDS-VID	PRID 2011
CVPR 2018	82.3 (65.8)	80.2	93.2

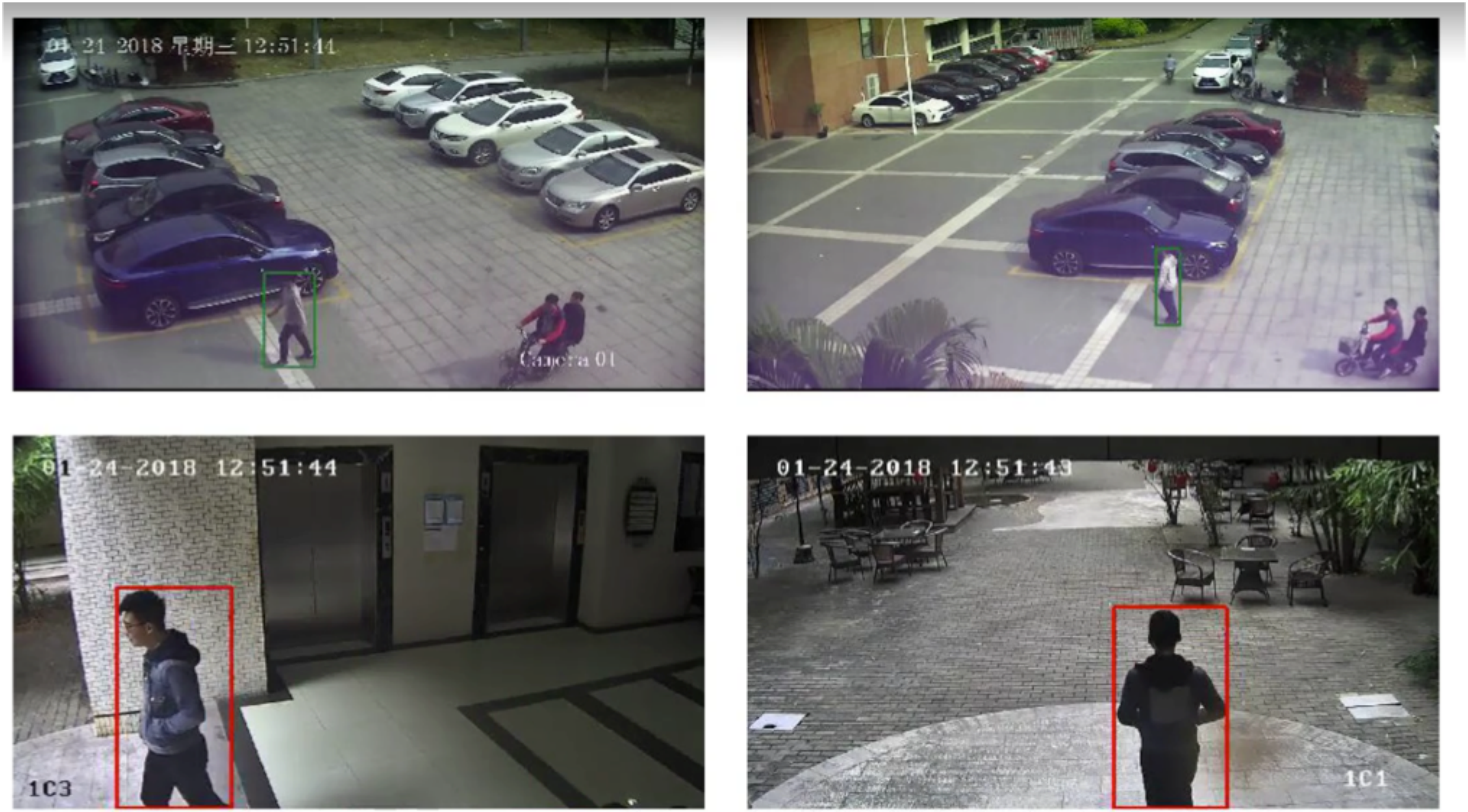
- Have we already solved it?

My Today's Focus

- Tell less about performance

- Aim to tell something of my understanding about Re-ID

Person Re-identification: Challenges



Person Re-identification: Challenges

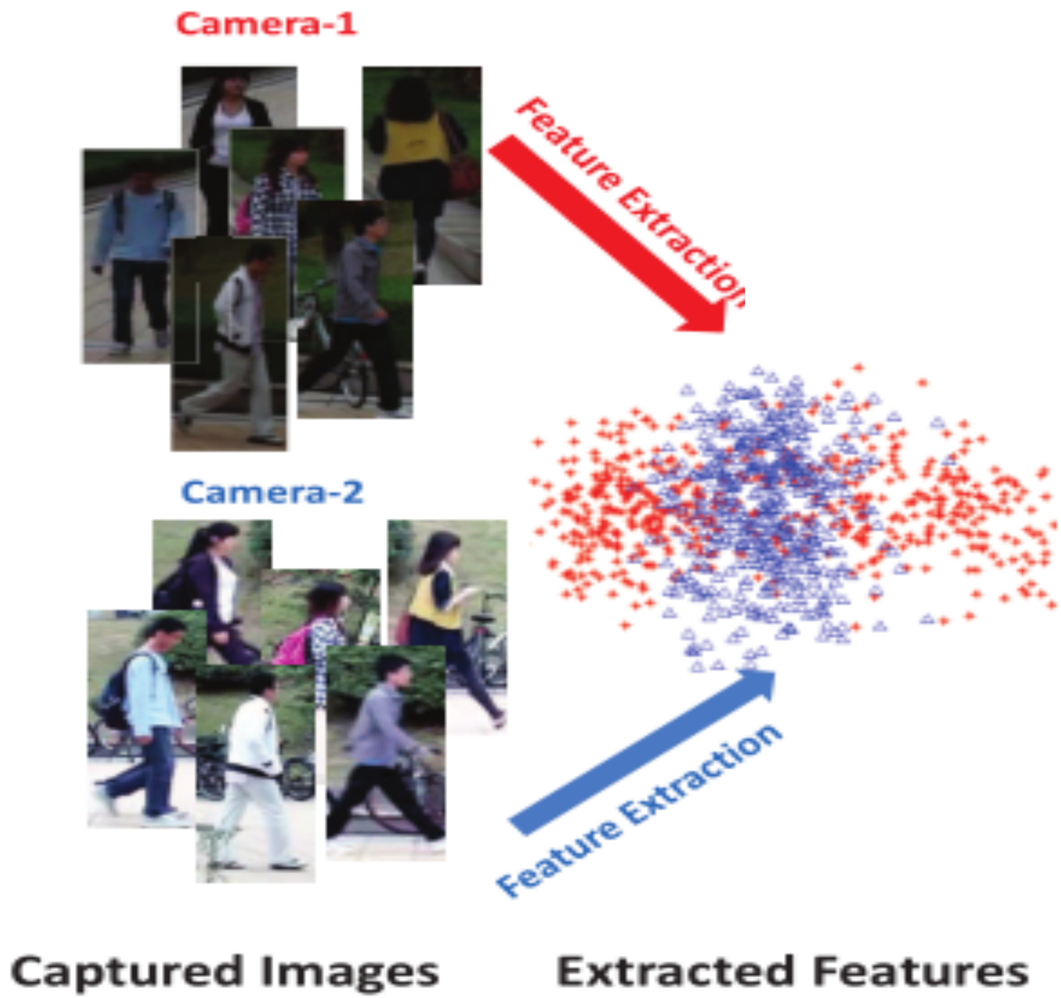
❑ Some Main Variations

View | Lighting | Occlusion | Low Resolution | Clothing Change

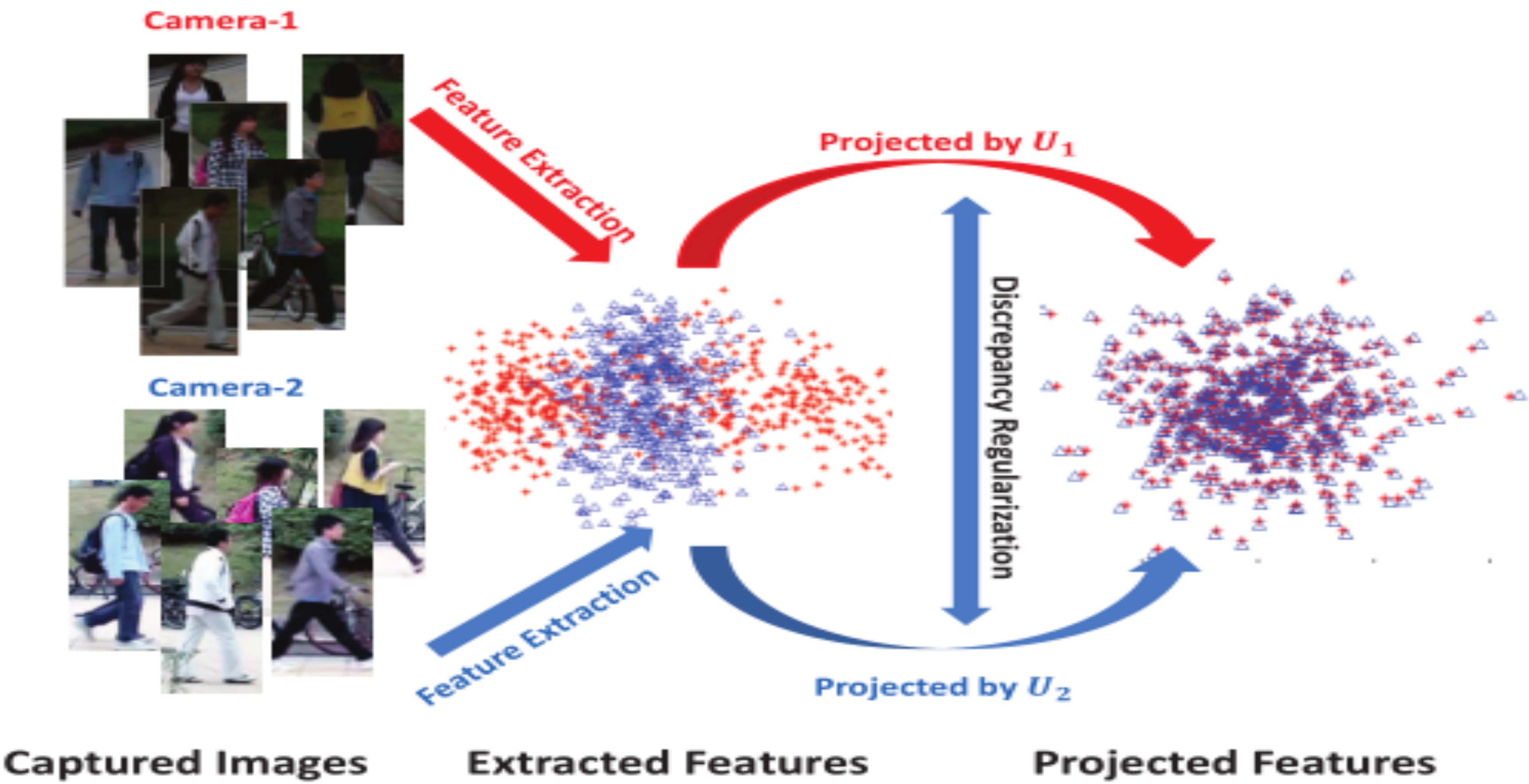
1. Connection with Cross Domain?

Person Re-ID vs. Cross-Modality

- View Bias



Asymmetric Metric for Re-ID



Learning universal feature transformation

Learning view-specific feature transformation

Asymmetric Metric for Re-ID

$$d(x_i, x_j) = \sqrt{(x_i - x_j)^T \mathbf{M}(x_i - x_j)}$$
$$= ||U^T x_i - U^T x_j||_2,$$

Learn different feature transformation for different camera views

Pseudometric

$$d(\{x_i^p, p\}, \{x_j^q, q\}) = ||U^{pT} x_i^p - U^{qT} x_j^q||_2$$
$$U^p \neq U^q$$

Non-negativity

Symmetry

$$\begin{aligned} d(\{x_i^p, p\}, \{x_j^q, q\}) &= ||U^{pT} x_i^p - U^{qT} x_j^q||_2 \\ &= ||U^{qT} x_j^q - U^{pT} x_i^p||_2 \\ &= d(\{x_j^q, q\}, \{x_i^p, p\}), \end{aligned}$$

Triangle Inequality

$$||U^{rT} x_k^r - U^{qT} x_j^q||_2 \leq ||U^{rT} x_k^r - U^{pT} x_i^p||_2 + ||U^{pT} x_i^p - U^{qT} x_j^q||_2.$$

Coincidence

$$d(\{x^p, p\}, \{x^q, q\}) = 0$$

$$\cancel{U}^{pT} x^p = \cancel{U}^{qT} x^q \quad \leftarrow \quad U^{pT} x^p = U^{qT} x^q$$

$$||U^p - U^q||_F^2 \downarrow$$

Asymmetric Metric for Re-ID

Re-ID Reformulation by Augmentation

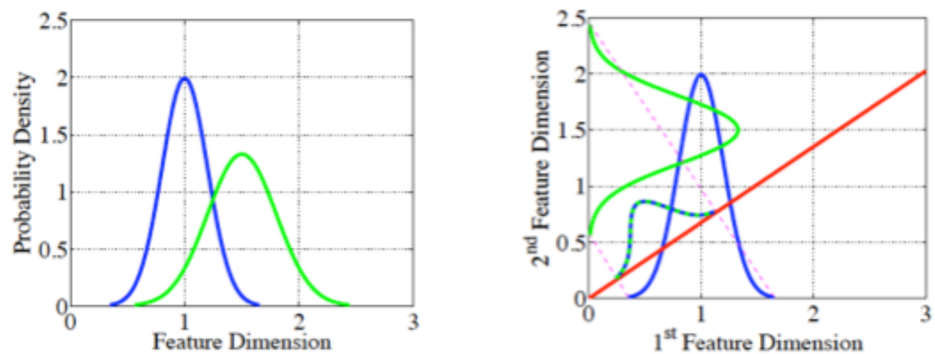
$$\tilde{X}^a_{\text{zp}} = \begin{bmatrix} I_{d\times d} \\ O_{d\times d} \end{bmatrix} X^a, \qquad \tilde{X}^b_{\text{zp}} = \begin{bmatrix} O_{d\times d} \\ I_{d\times d} \end{bmatrix} X^b$$

$$\hat{W} = \min_{W} f_{\text{obj}}(W^\top \tilde{X}_{\text{zp}})$$

$$\hat{W} = [(\hat{W}^a)^\top, (\hat{W}^b)^\top]^\top$$

View-specific transformation

$$\hat{W}^\top \tilde{X}^a_{\text{zp}} = (\hat{W}^a)^\top X^a$$
$$\hat{W}^\top \tilde{X}^b_{\text{zp}} = (\hat{W}^b)^\top X^b.$$



(a) Original 1-D feature space (b) Augmented 2-D feature space

Fig. 3. An illustration of zero padding based feature augmentation. (a) The data distribution in the original feature space from camera view a (the blue curve) and b (the green curve). (b) The augmented feature space by zero padding. The dashed blue and green curves represent the projected features with respect to the projection basis indicated by the solid red line. The two dashed lines imply feature projection operation. Note that the probability density axis is not plotted in (b) for demonstration simplicity.

Yingcong Chen, Xiatian Zhu, Wei-Shi Zheng*, and Jian-Huang Lai. Person Re-Identification by Camera Correlation Aware Feature Augmentation. IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI), 2017.

Asymmetric Metric for Re-ID

Re-ID Reformulation by Augmentation

$$\tilde{X}^a_{zp} = \begin{bmatrix} I_{d\times d} \\ O_{d\times d} \end{bmatrix} X^a, \qquad \tilde{X}^b_{zp} = \begin{bmatrix} O_{d\times d} \\ I_{d\times d} \end{bmatrix} X^b$$

$$\hat{W} = \min_{W} f_{\text{obj}}(W^\top \tilde{X}_{zp})$$

$$\hat{W} = [(\hat{W}^a)^\top, (\hat{W}^b)^\top]^\top$$

View-specific transformation

$$\hat{W}^\top \tilde{X}^a_{zp} = (\hat{W}^a)^\top X^a$$
$$\hat{W}^\top \tilde{X}^b_{zp} = (\hat{W}^b)^\top X^b.$$

Not able to measure the relationship between different view-specific transformation matrices

Do not constraint the discrepancy between feature transformation across view: Coincidence

Yingcong Chen, Xiatian Zhu, Wei-Shi Zheng*, and Jian-Huang Lai. Person Re-Identification by Camera Correlation Aware Feature Augmentation. IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI), 2017.

Asymmetric Metric for Re-ID

- Adaptive feature augmentation

$$\tilde{\boldsymbol{X}}^a_{\text{zp}} = \begin{bmatrix} \boldsymbol{I}_{d\times d} \\ \boldsymbol{O}_{d\times d} \end{bmatrix} \boldsymbol{X}^a, \qquad \tilde{\boldsymbol{X}}^b_{\text{zp}} = \begin{bmatrix} \boldsymbol{O}_{d\times d} \\ \boldsymbol{I}_{d\times d} \end{bmatrix} \boldsymbol{X}^b$$

$$\Downarrow \qquad\qquad \Downarrow$$

$$\tilde{\boldsymbol{X}}^a_{\text{craft}} = \begin{bmatrix} \boldsymbol{R} \\ \boldsymbol{M} \end{bmatrix} \boldsymbol{X}^a, \qquad \tilde{\boldsymbol{X}}^b_{\text{craft}} = \begin{bmatrix} \boldsymbol{M} \\ \boldsymbol{R} \end{bmatrix} \boldsymbol{X}^b$$

generalised

$$f_a(\tilde{\boldsymbol{X}}^a_{\text{craft}}) = \boldsymbol{W}^\top \tilde{\boldsymbol{X}}^a_{\text{craft}} = (\boldsymbol{R}^\top \boldsymbol{W}^a + \boldsymbol{M}^\top \boldsymbol{W}^b)^\top \boldsymbol{X}^a$$

$$f_b(\tilde{\boldsymbol{X}}^b_{\text{craft}}) = \boldsymbol{W}^\top \tilde{\boldsymbol{X}}^b_{\text{craft}} = (\boldsymbol{M}^\top \boldsymbol{W}^a + \boldsymbol{R}^\top \boldsymbol{W}^b)^\top \boldsymbol{X}^b$$

control the discrepancy **Between** f_a and f_b

Yingcong Chen, Xiatian Zhu, Wei-Shi Zheng*, and Jian-Huang Lai. Person Re-Identification by Camera Correlation Aware Feature Augmentation. IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI), 2017.

Asymmetric Metric for Re-ID

- **Learning:**

Camera coRrelation Aware Feature augmenTation (CRAFT)

$$\hat{\boldsymbol{W}} = \arg\min_{\boldsymbol{W}} f_{\text{obj}}(\boldsymbol{W}^\top \tilde{\boldsymbol{X}}_{\text{craft}}) + \lambda \text{tr}(\boldsymbol{W}^\top \boldsymbol{C}\boldsymbol{W})$$

Generalize any symmetric metric learning models to asymmetric ones: e.g. MFA

$$\min_{\boldsymbol{H}} \sum_{i \neq j} A^c_{ij} ||\boldsymbol{H}^\top (\ddot{\boldsymbol{x}}_i - \ddot{\boldsymbol{x}}_j)||_2^2 + \lambda \text{tr}(\boldsymbol{H}^\top \boldsymbol{H})$$

$$\text{s.t.} \sum_{i \neq j} A^p_{ij} ||\boldsymbol{H}^\top (\ddot{\boldsymbol{x}}_i - \ddot{\boldsymbol{x}}_j)||_2^2 = 1,$$

$$A^c_{ij} = \begin{cases} 1 & \text{if } i \in N^+_{k_1}(j) \text{ or } j \in N^+_{k_1}(i) \\ 0 & \text{otherwise,} \end{cases}$$

$$A^p_{ij} = \begin{cases} 1 & \text{if } (i,j) \in P_{k_2}(y_i) \text{ or } (i,j) \in P_{k_2}(y_j) \\ 0 & \text{otherwise,} \end{cases}$$

Asymmetric Metric for Re-ID

- Learning:

Camera coRrelation Aware Feature augmenTation (CRAFT)

$$\hat{\boldsymbol{W}} = \arg\min_{\boldsymbol{W}} f_{\text{obj}}(\boldsymbol{W}^\top \tilde{\boldsymbol{X}}_{\text{craft}}) + \lambda \text{tr}(\boldsymbol{W}^\top \boldsymbol{C}\boldsymbol{W})$$

$$\boldsymbol{C} = \begin{bmatrix} \boldsymbol{I} & -\beta\boldsymbol{I} \\ -\beta\boldsymbol{I} & \boldsymbol{I} \end{bmatrix}, \quad \beta = \frac{1}{1+\eta_{\text{ridge}}}$$

$$\begin{aligned}\gamma &= \|\boldsymbol{W}^a - \boldsymbol{W}^b\|^2 + \eta_{\text{ridge}}\text{tr}(\boldsymbol{W}^\top\boldsymbol{W}) \\ &= \text{tr}(\boldsymbol{W}^\top \begin{bmatrix} \boldsymbol{I} & -\boldsymbol{I} \\ -\boldsymbol{I} & \boldsymbol{I} \end{bmatrix} \boldsymbol{W}) + \eta_{\text{ridge}}\text{tr}(\boldsymbol{W}^\top\boldsymbol{W}) \\ &= (1+\eta_{\text{ridge}})\text{tr}(\boldsymbol{W}^\top\boldsymbol{C}\boldsymbol{W}),\end{aligned}$$

Camera View Discrepancy Regularization : Reduce Coincidence

Bregman discrepancy of a projection

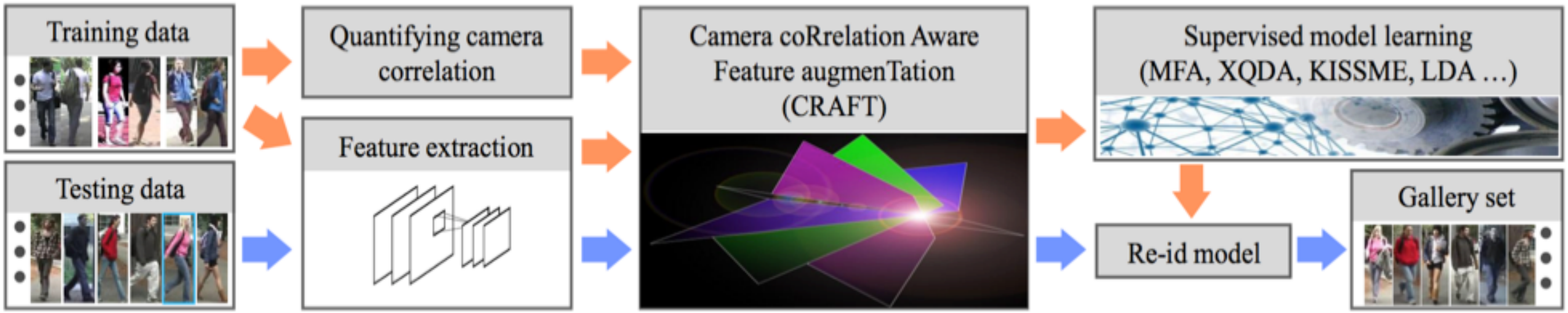
a strictly convex function $\mathcal{F} : \mathbb{R}^{d\times C} \rightarrow \mathbb{R}$

$$d_{\mathcal{F}}(\boldsymbol{U}^p, \boldsymbol{U}^q) = \mathcal{F}(\boldsymbol{U}^p) - \mathcal{F}(\boldsymbol{U}^q) - \nabla\mathcal{F}(\boldsymbol{U}^q)^T(\boldsymbol{U}^p - \boldsymbol{U}^q)$$

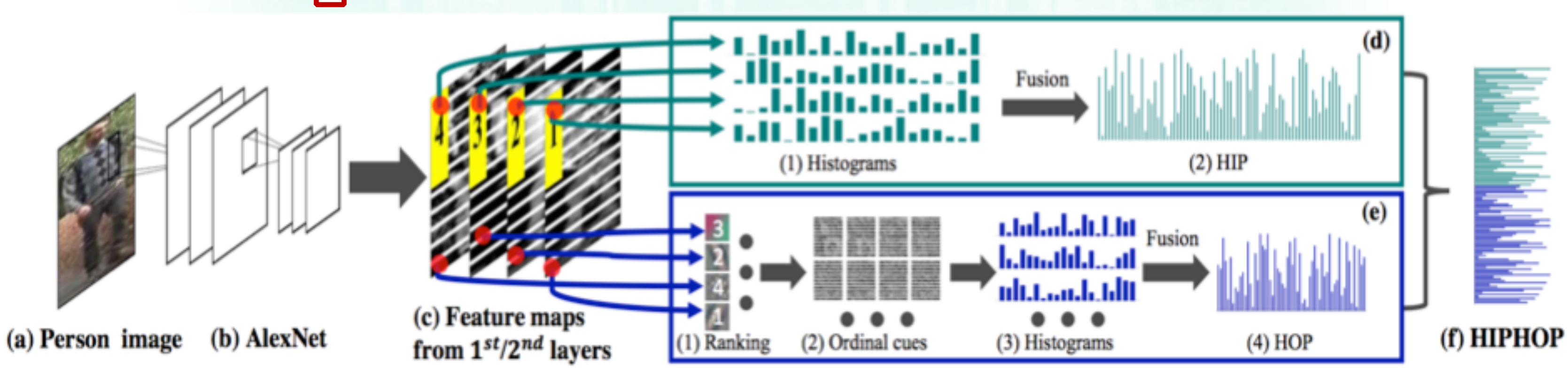
$$\mathcal{F}(\boldsymbol{x}) = \boldsymbol{x}^T\boldsymbol{x} \quad d_{\mathcal{F}}(\boldsymbol{U}^p, \boldsymbol{U}^q) = \|\boldsymbol{U}^p - \boldsymbol{U}^q\|_F^2$$

Asymmetric Metric for Re-ID

❑ A framework



to extract domain-generic and more view invariant person features



Asymmetric Metric for Re-ID

❑ Evaluation: augmentation or not augmentation?

Dataset	VIPeR [30]				CUHK01 [31]				CUHK03 [26]				Market-1501 [33]				QMUL GRID [32]			
Rank (%)	1	5	10	20	1	5	10	20	1	5	10	20	1	5	10	20	1	5	10	20
OriFeat	43.3	72.7	84.1	93.4	64.3	85.1	90.6	94.6	63.4	88.0	93.0	96.1	65.4	84.0	89.3	93.1	21.0	42.9	53.0	62.7
+ Kernelization	47.0	75.4	86.8	94.4	69.5	89.3	93.5	96.5	78.6	94.9	96.8	98.4	66.0	84.4	89.3	93.2	19.0	42.2	51.9	61.4
ZeroPad [71]	37.5	69.9	82.8	92.1	66.4	85.8	90.5	94.7	76.0	91.9	94.8	95.3	38.2	62.5	71.7	80.0	7.2	26.0	40.3	55.8
+ Kernelization	40.0	72.8	85.0	93.4	71.8	89.6	93.8	96.5	80.0	92.7	94.4	95.3	49.5	72.4	80.0	85.8	6.1	21.8	36.3	51.4
BaseFeatAug [69]	45.5	76.1	87.4	95.1	59.0	81.1	87.0	92.4	78.3	94.6	97.3	98.9	65.3	83.6	88.8	92.6	20.1	47.6	58.8	70.0
+ Kernelization	47.3	77.8	89.0	95.2	63.0	83.5	89.0	93.6	83.4	97.0	98.1	99.1	65.3	83.6	88.8	92.6	20.1	47.6	58.8	70.0
CRAFT	47.8	77.1	87.8	95.1	70.0	87.4	92.0	95.5	78.5	94.7	97.5	98.9	67.9	85.1	90.0	93.4	25.4	50.2	61.8	74.2
+ Kernelization	50.3	80.0	89.6	95.5	74.5	91.2	94.8	97.1	84.3	97.1	98.3	99.1	68.7	87.1	90.8	94.0	22.4	49.9	61.8	71.7

❑ Evaluation: augmentation vs. domain adaptation

Comparison between CRAFT and domain adaptation.

Dataset	VIPeR [30]				CUHK01 [31]				CUHK03 [26]				Market-1501 [33]				QMUL GRID [32]			
Rank (%)	1	5	10	20	1	5	10	20	1	5	10	20	1	5	10	20	1	5	10	20
TCA [63]	11.1	23.4	31.0	38.5	7.0	16.4	22.2	30.1	5.5	16.2	26.4	42.8	8.9	18.7	24.1	30.1	9.8	22.2	29.8	38.3
TFLDA [64]	46.4	75.8	86.7	93.9	69.6	88.7	92.8	96.2	76.7	94.4	96.5	98.0	62.5	81.3	87.0	91.6	19.5	42.5	51.6	61.8
CRAFT	**50.3**	**80.0**	**89.6**	**95.5**	**74.5**	**91.2**	**94.8**	**97.1**	**84.3**	**97.1**	**98.3**	**99.1**	**68.7**	**87.1**	**90.8**	**94.0**	**22.4**	**49.9**	**61.8**	**71.7**

❑ Evaluation: whether using Camera View Discrepancy

Evaluating the effect of our CVD regularization.

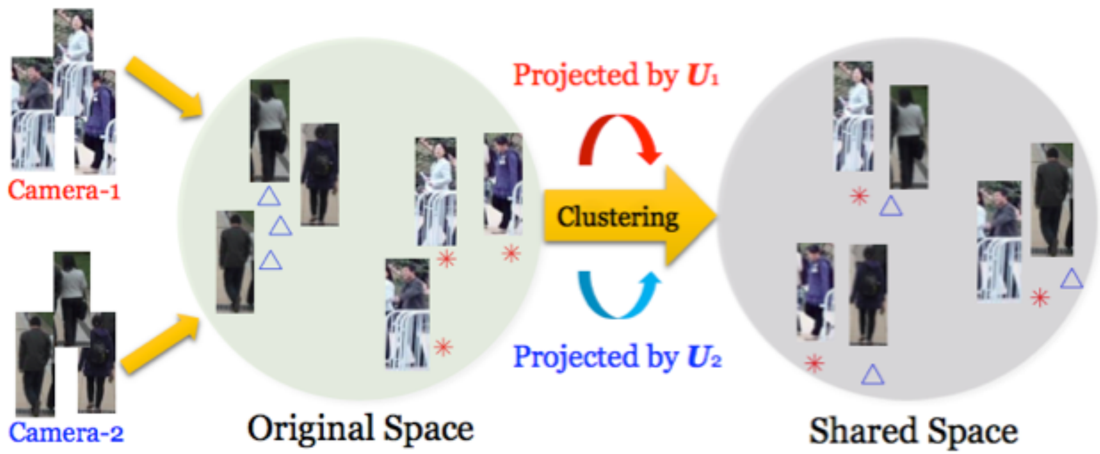
Dataset	VIPeR [30]				CUHK01 [31]				CUHK03 [26]				Market-1501 [33]				QMUL GRID [32]			
Rank (%)	1	5	10	20	1	5	10	20	1	5	10	20	1	5	10	20	1	5	10	20
CRAFT(no γ_{cvd})	46.3	77.9	88.1	95.4	73.8	90.6	94.2	96.9	83.9	97.0	98.2	**99.1**	66.6	85.9	90.7	93.7	15.8	45.0	57.7	60.0
CRAFT	**50.3**	**80.0**	**89.6**	**95.5**	**74.5**	**91.2**	**94.8**	**97.1**	**84.3**	**97.1**	**98.3**	**99.1**	**68.7**	**87.1**	**90.8**	**94.0**	**22.4**	**49.9**	**61.8**	**71.7**

Does the Asymmetric Metric Modelling Work for other setting: unsupervised, semi-supervised, …..

Asymmetric Metric for Re-ID: Unsupervised

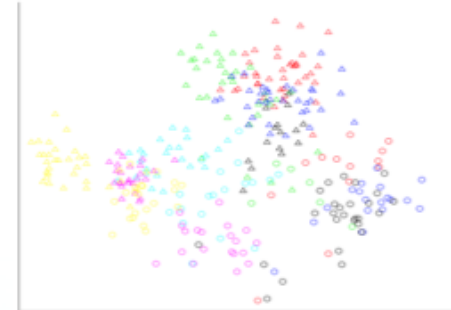
Unsupervised Learning

- Clustering-based Asymmetric MEtric Learning (CAMEL)

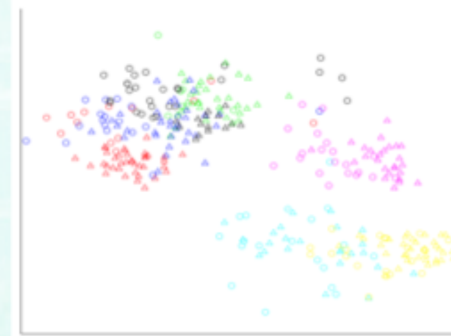


$$\min_{U^1,\cdots,U^V} \mathcal{F}_{obj} = \frac{1}{N}\sum_{k=1}^{K}\sum_{i\in C_k}\|U^{p\mathrm{T}}\mathbf{x}_i^p - \mathbf{c}_k\|^2 + \lambda\sum_{p\neq q}\|U^p - U^q\|_F^2$$

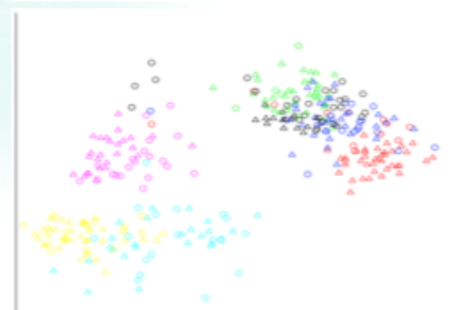
$$s.t. \quad U^{p\mathrm{T}}\Sigma^p U^p = I \quad (p = 1,\cdots,V),$$



(a) Original Feature Distribution



(b) Middle Stage of CAMEL



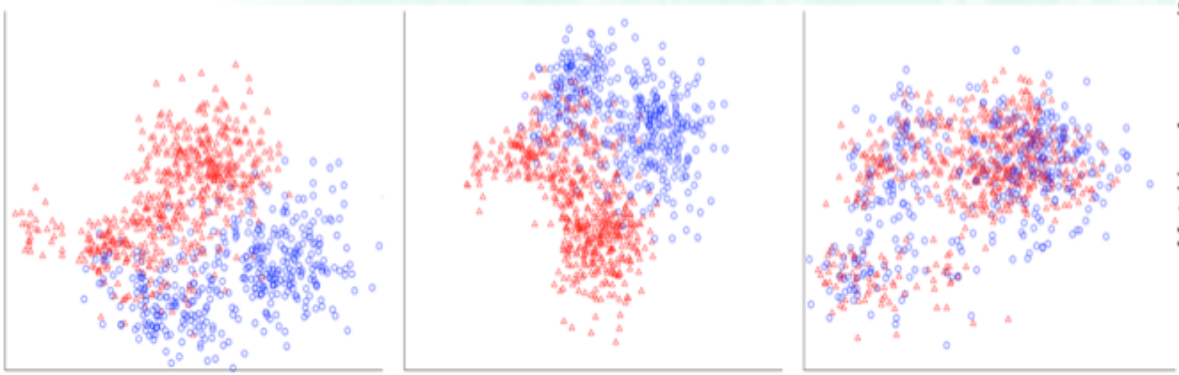
(c) Convergence Stage of CAMEL

Hongxing Yu, Ancong Wu, Wei-Shi Zheng*. Cross-view Asymmetric Metric Learning for Unsupervised Person Re-identification. In IEEE Conf. on Computer Vision (ICCV), 2017.

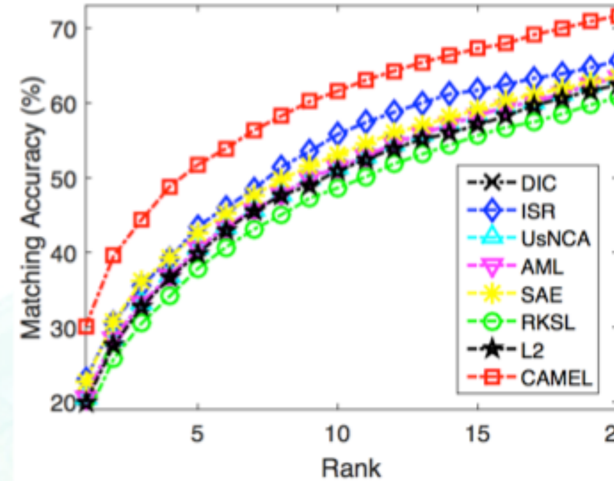
Asymmetric Metric for Re-ID: Unsupervised

Unsupervised Learning

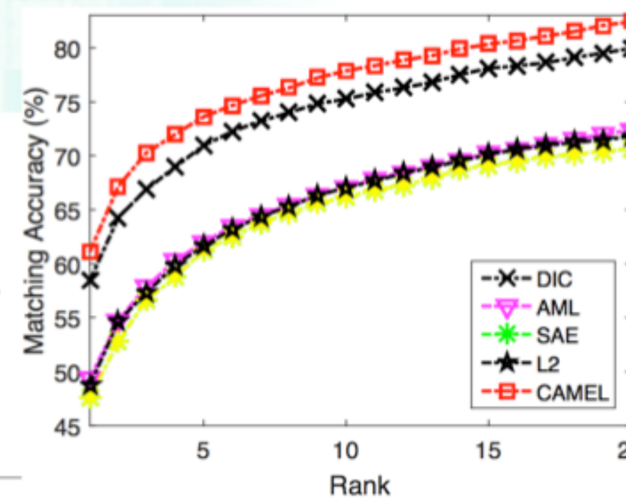
Dataset	VIPeR	CUHK01	CUHK03	SYSU	Market	ExMarket
Setting	SS	SS/MS	SS/MS	SS/MS	MS	MS
Dic [12]	32.9	49.1/52.2	23.8/29.2	20.5/26.0	57.5(28.8)	58.4(26.9)
ISR [20]	29.2	45.1/46.3	27.1/35.6	23.0/33.4	25.6(8.8)	-
RKSL [29]	27.0	38.6/40.4	20.8/26.9	17.8/22.3	41.4(17.9)	-
SAE [14]	25.1	39.2/41.9	19.3/24.4	22.7/27.6	46.8(20.3)	47.5(19.4)
L_2	23.2	40.3/43.1	20.5/26.1	19.9/24.1	47.9(19.9)	49.3(18.6)
AML [33]	23.6	37.6/39.9	18.7/23.6	20.5/24.9	47.9(19.9)	49.3(18.6)
UsNCA [24]	23.6	37.6/39.9	18.8/23.7	20.2/24.1	47.9(19.9)	-
CAMEL	**33.5**	**55.8/59.1**	**30.1/37.9**	**31.2/37.8**	**61.2(31.4)**	**61.1(29.1)**



(a) Original Feature Distribution (b) After Symmetric Metric (c) After Asymmetric Metric



(d) SYSU



(f) ExMarket

Hash Re-ID for Fast Search

- FAST Re-ID on Numbers of Cameras
  - Learning view-specific hash code for each camera

$$f_p(\boldsymbol{x}_i^p) = \boldsymbol{x}_i^p \boldsymbol{W}_p, \qquad f_g(\boldsymbol{x}_j^g) = \boldsymbol{x}_j^g \boldsymbol{W}_g$$

$$\boldsymbol{B}_p = \operatorname{sign}(\boldsymbol{X}_p \boldsymbol{W}_p) \in \{-1, 1\}^{n_p \times c},$$
$$\boldsymbol{B}_g = \operatorname{sign}(\boldsymbol{X}_g \boldsymbol{W}_g) \in \{-1, 1\}^{n_g \times c},$$

Xiatian Zhu, Botong Wu, Dongcheng Huang, Wei-Shi Zheng*(PI). Fast Open-World Person Re-Identification. IEEE Transactions on Image Processing, 2017.

Wei-Shi Zheng, Shaogang Gong, and Tao Xiang. Towards Open-World Person Re-Identification by One-Shot Group-based Verification. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), vol. 38, no. 3, pp. 591-606, 2016.

Hash Re-ID for Fast Search

Idea of the Formulation

$$\underbrace{\left(\|\boldsymbol{B}_p - \boldsymbol{X}_p\boldsymbol{W}_p\|_F^2 + \|\boldsymbol{B}_g - \boldsymbol{X}_g\boldsymbol{W}_g\|_F^2\right)}_{\text{Quantisation loss}} +$$

Cross-view Identity Verification Regularisation

$$l_{\text{ie}}(y_i, U^\top b_i) = \|U\|_F^2 + \eta_{\text{hinge}} \sum_{i=1}^{n} \varepsilon_i$$

$$\text{s.t.} \quad \forall i,j \quad u_{k_i}^\top b_i - u_j^\top b_i + y_{i,j} \geq 1 - \varepsilon_i, \quad \varepsilon_i \geq 0$$

$+$

Cross-view Identity Correlation Hashing

$$\text{cosine}(f_p(\boldsymbol{x}_i^p), f_g(\boldsymbol{x}_j^g)) = \frac{f_p(\boldsymbol{x}_i^p)(f_g(\boldsymbol{x}_j^g))^\top}{\|f_p(\boldsymbol{x}_i^p)\|_2 \|f_g(\boldsymbol{x}_j^g)\|_2}$$

$$= \frac{\boldsymbol{x}_i^p \boldsymbol{W}_p \boldsymbol{W}_g^\top \boldsymbol{x}_j^{g\top}}{\sqrt{\boldsymbol{x}_i^p \boldsymbol{W}_p \boldsymbol{W}_p^\top \boldsymbol{x}_i^{p\top}} \sqrt{\boldsymbol{x}_j^g \boldsymbol{W}_g \boldsymbol{W}_g^\top \boldsymbol{x}_j^{g\top}}}.$$

$\times$

View Context Discrepancy Regularisation

$$R_{\text{vcd}} = \|\boldsymbol{W}_p - \boldsymbol{W}_g\|_F^2.$$

Hash Re-ID for Fast Search

FAST Search

- Comparison to other related Hashing functions

TABLE III: Comparing state-of-the-art hashing methods. (Metrics: TTR (%) at varying FTRs (%), and mAP (%).)

Dataset	CUHK03 [39]						SYSU [36]						Market-1501 [40]					
Metric	Individual Verification					mAP	Individual Verification					mAP	Individual Verification					mAP
	1%	5%	10%	20%	30%	(%)	1%	5%	10%	20%	30%	(%)	1%	5%	10%	20%	30%	(%)
LSH [25]	15.03	34.87	48.12	64.66	75.20	1.91	21.21	43.73	57.42	72.17	81.17	5.48	37.00	61.56	73.43	85.12	90.92	8.28
SH [27]	11.97	27.8	39.99	54.73	65.35	1.49	17.60	34.21	45.63	60.14	70.27	4.18	38.54	58.69	69.47	80.71	86.75	9.29
SGH [65]	16.95	37.37	50.71	66.45	76.76	2.36	27.18	49.21	61.74	75.34	82.92	8.03	37.75	63.16	75.05	86.13	91.76	8.69
ITQ [66]	17.31	39.29	53.06	69.12	80.24	2.70	26.51	49.55	63.18	77.04	84.87	7.47	40.72	67.17	78.67	88.51	93.41	10.84
CCA+ITQ [66]	28.11	51.15	65.05	78.95	86.37	4.28	50.50	73.75	83.16	91.08	94.67	18.12	57.75	80.30	87.53	93.47	96.17	15.02
KSH [30]	32.29	57.54	69.78	81.73	88.96	5.49	53.23	77.28	85.88	92.62	95.62	22.29	59.03	81.83	89.01	94.26	96.41	17.34
FH [68]	20.01	40.07	52.32	67.35	77.39	1.03	29.48	50.56	62.42	75.58	83.65	8.07	28.24	48.88	60.59	73.92	81.61	5.07
SDH [31]	38.80	66.82	78.83	88.15	93.03	7.31	46.09	72.34	82.76	90.75	94.51	17.99	58.03	81.26	88.22	94.16	96.29	15.57
COSDISH [69]	13.19	29.18	40.33	56.88	68.23	1.55	38.04	61.43	72.73	83.95	89.38	11.51	39.29	62.26	73.49	83.68	89.04	8.44
CMSSH [48]	10.46	32.46	49.80	68.67	80.45	1.25	11.06	33.76	50.51	70.55	82.29	3.18	8.88	29.25	46.49	67.02	79.56	1.55
CVH [47]	2.83	10.09	17.81	31.05	42.51	0.39	5.76	19.67	31.77	49.33	62.21	1.30	3.51	13.38	22.62	37.31	50.55	0.53
CMFH [71]	11.85	31.23	46.40	64.63	75.56	1.27	25.73	54.24	68.73	82.09	89.14	6.32	24.96	52.96	67.52	81.62	89.11	4.07
SCM [51]	5.43	17.84	28.77	44.72	58.70	0.59	14.83	32.93	45.22	60.35	70.95	3.92	13.41	31.49	43.44	58.75	69.09	2.04
SePH [70]	26.98	52.69	65.88	79.29	86.24	4.18	37.15	64.01	75.75	86.09	91.56	13.56	41.88	70.39	80.72	88.89	93.09	8.80
X-ICE(hinge)	49.67	**79.60**	**89.50**	**96.09**	**98.48**	**11.66**	61.86	84.10	91.47	**96.35**	**98.26**	**29.93**	**66.52**	**88.03**	**93.66**	**97.15**	98.55	**21.47**
X-ICE(reg)	**49.96**	78.18	88.96	95.88	97.98	11.23	**63.13**	**84.86**	**91.52**	96.17	98.08	29.44	64.18	86.98	92.91	97.09	**98.59**	20.68
Metric	Set Verification					mAP	Set Verification					mAP	Set Verification					mAP
LSH [25]	4.81	13.97	21.83	35.01	45.88	1.91	7.25	18.13	27.35	41.18	52.31	5.48	17.17	33.22	43.85	57.75	67.36	8.28
SH [27]	3.91	12.00	19.93	31.72	43.41	1.49	7.22	16.36	24.39	36.62	46.96	4.18	21.52	36.77	46.41	58.33	68.00	9.29
SGH [65]	5.42	14.34	22.87	36.84	48.89	2.36	10.20	22.06	31.91	46.08	56.88	8.03	16.59	34.04	45.33	59.12	69.60	8.69
ITQ [66]	5.45	14.9	23.96	37.95	49.51	2.70	8.81	21.47	31.42	45.82	56.73	7.47	17.39	35.66	47.25	60.90	70.46	10.84
CCA+ITQ [66]	10.52	23.86	34.05	49.95	59.58	4.28	21.08	42.10	53.63	67.05	75.76	18.12	18.30	45.05	60.16	73.75	81.38	15.02
KSH [30]	11.65	27.43	38.19	52.68	63.62	5.49	22.07	43.13	55.44	69.26	77.25	22.29	21.74	47.36	61.28	74.78	82.22	17.34
FH [68]	7.66	17.92	26.82	40.15	52.08	1.03	12.82	26.19	35.86	48.86	59.28	8.07	13.32	27.17	36.78	50.13	60.84	5.07
SDH [31]	13.59	31.37	43.59	60.08	70.34	7.31	15.46	36.48	49.00	63.30	72.79	17.99	20.95	47.01	61.17	75.19	82.16	15.57
COSDISH [69]	5.07	13.37	21.71	34.27	43.74	1.55	16.86	33.10	43.36	57.20	66.38	11.51	15.80	35.00	46.81	60.49	69.66	8.44
CMSSH [48]	2.32	10.32	19.42	32.44	45.06	1.25	2.66	11.13	19.29	33.87	46.11	3.18	2.20	8.93	17.42	31.57	44.29	1.55
CVH [47]	1.31	5.62	12.06	23.29	32.45	0.39	1.75	7.65	14.44	27.18	38.12	1.30	1.57	6.83	12.83	24.39	35.74	0.53
CMFH [71]	3.49	11.91	20.06	34.46	45.35	1.27	7.25	21.42	32.99	48.97	60.98	6.32	7.95	22.60	33.55	48.81	60.49	4.07
SCM [51]	1.91	7.64	14.74	27.22	37.74	0.59	5.75	15.55	23.82	37.04	48.23	3.92	5.29	15.23	23.20	36.53	47.93	2.04
SePH [70]	9.64	23.22	33.51	49.79	60.94	4.18	12.40	30.40	42.85	57.81	67.73	13.56	11.78	32.98	46.91	63.28	73.64	8.80
X-ICE(hinge)	**16.41**	**37.50**	**50.14**	**66.56**	**77.30**	**11.66**	23.32	46.84	60.48	74.20	82.37	**29.93**	**26.81**	**52.73**	**66.47**	**79.66**	**86.16**	**21.47**
X-ICE(reg)	16.37	37.36	49.71	65.49	76.03	11.23	**25.94**	**49.94**	**62.59**	**75.91**	**83.30**	29.44	22.27	48.12	62.87	77.13	84.55	20.68

Hash Re-ID for Fast Search

❑ FAST Search

- When using more powerful features?

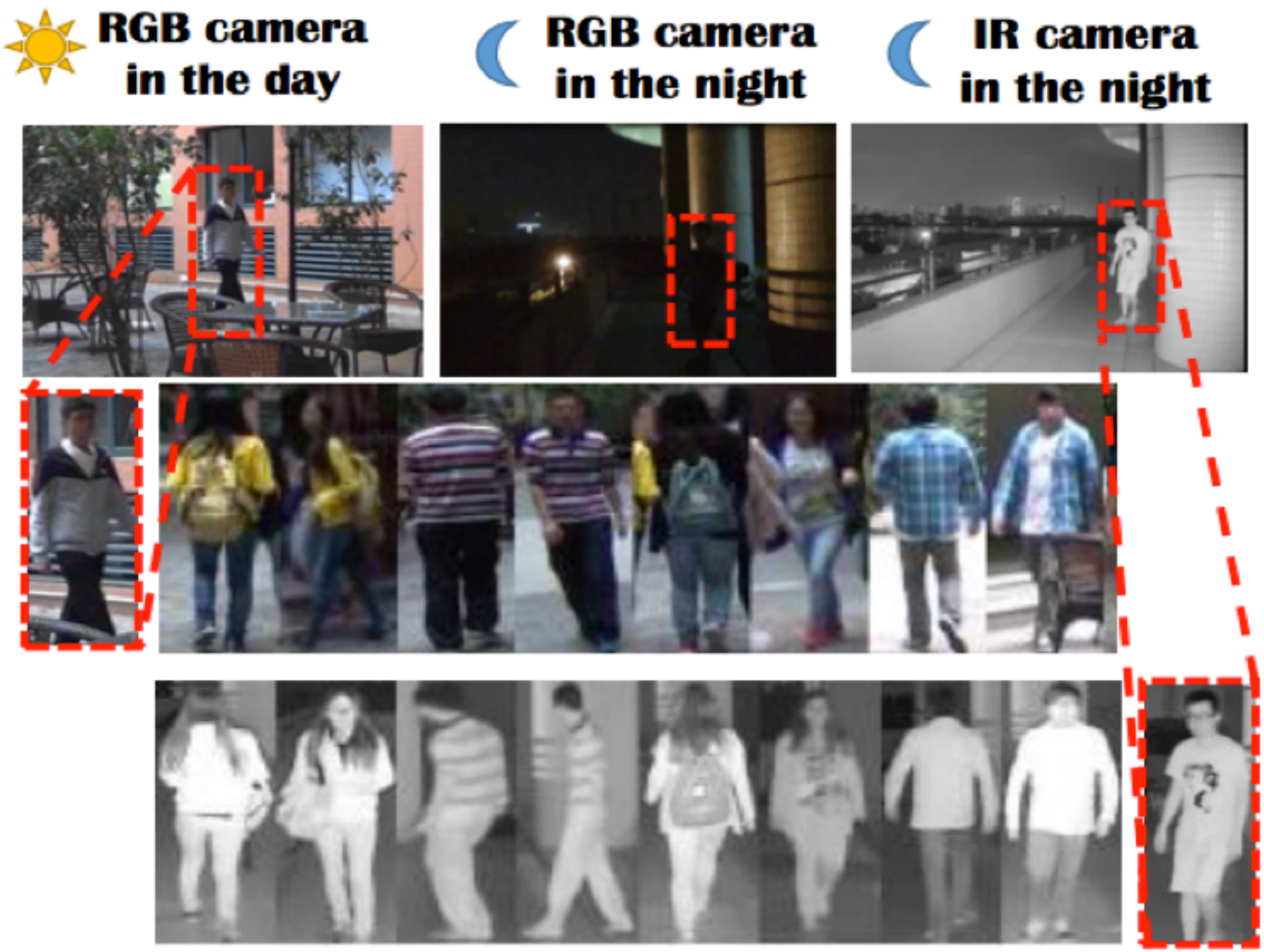
TABLE IX: Evaluating the effect of different visual features. (Metrics: TTR (%) at FTR = 1%, and mAP (%). IV: Individual Verification, SV: Set Verification.)

Method	Feature	CUHK03 [39]			SYSU [36]			Market-1501 [40]		
		IV	SV	mAP	IV	SV	mAP	IV	SV	mAP
DCNN [74]	Deep	47.87	14.38	13.62	58.77	19.64	29.69	78.37	31.58	33.65
KSH [30]	LOMO	32.29	11.65	5.49	53.23	22.07	22.29	59.03	21.74	17.34
	Deep	51.08	**19.00**	14.77	60.38	21.55	31.54	79.50	34.39	34.96
SDH [31]	LOMO	38.80	13.59	7.31	46.09	15.46	17.99	58.03	20.95	15.57
	Deep	36.88	15.93	9.11	52.93	16.60	22.23	71.00	35.61	26.73
X-ICE(hinge)	LOMO	49.67	16.41	11.66	61.86	23.32	29.93	66.52	26.81	21.47
	Deep	51.79	18.29	**15.33**	63.08	23.35	32.95	**80.90**	**41.52**	**37.34**
X-ICE(reg)	LOMO	49.96	16.37	11.23	63.13	25.94	29.44	64.18	22.27	20.68
	Deep	**52.62**	17.63	15.21	**64.05**	**24.57**	**33.07**	80.37	41.01	37.23

2. How to match heterogeneous person images across camera views?

Person Re-ID vs. Cross-Modality

- Matching between Heterogeneous Images



RGB-Infrared Re-ID

Cross-Modality Learning: RGB-IR Re-ID

- Deep zero-padding

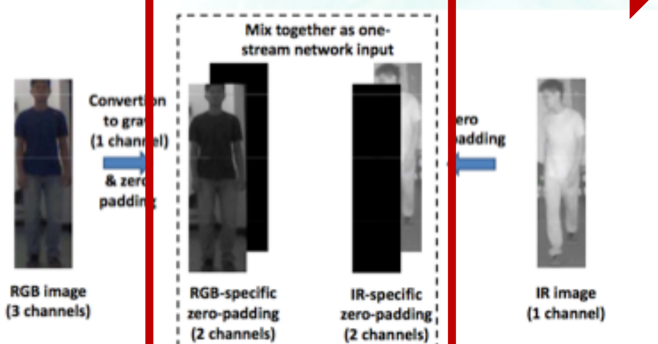
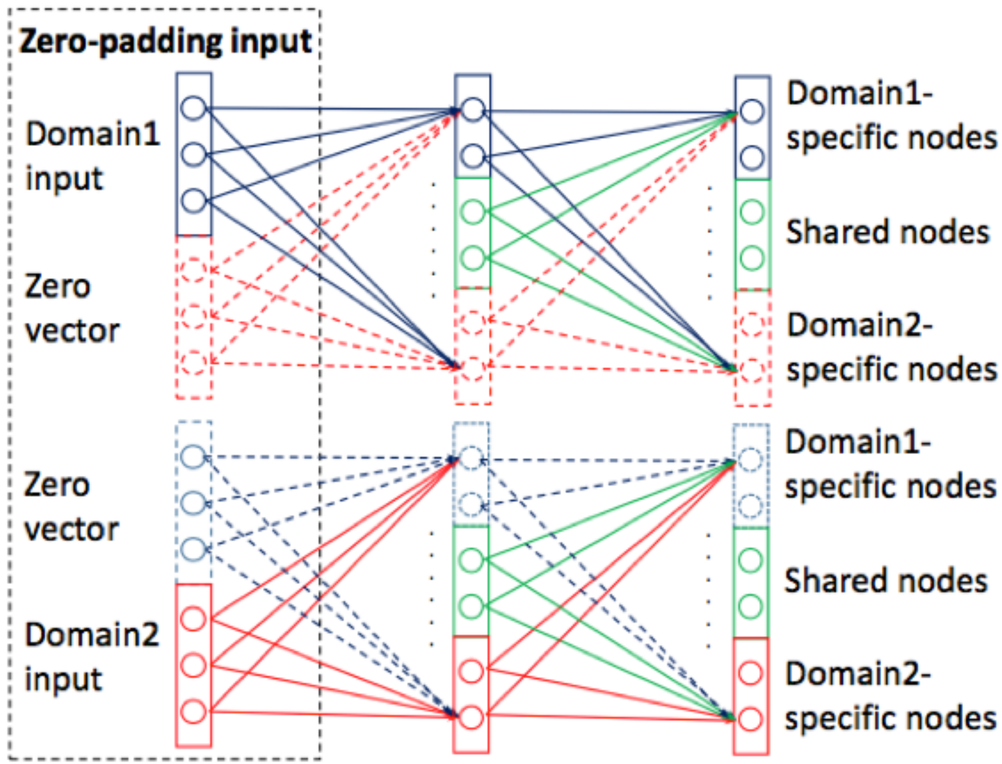


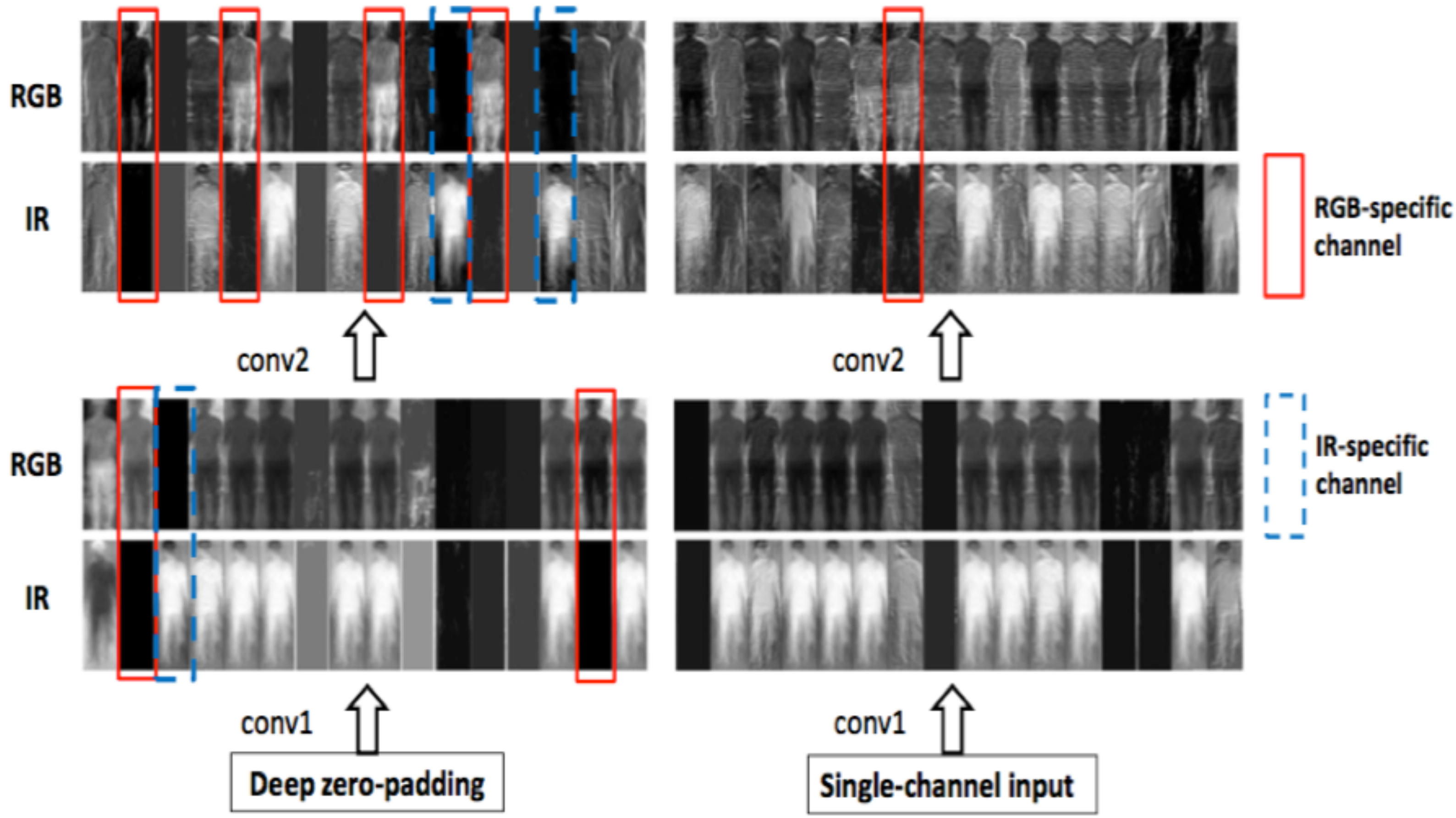
Figure 6. Deep zero-padding for RGB and infrared (IR) images.



Ancong Wu, Wei-Shi Zheng*(PI), Hongxing Yu, Shaogang Gong, Jianhuang Lai. RGB-Infrared Cross-Modality Person Re-Identification. In IEEE Conf. on Computer Vision (ICCV), 2017.

RGB-Infrared Re-ID

- Cross-Modality Learning: RGB-IR Re-ID



RGB-Infrared Re-ID

❑ Cross-Modality Learning: RGB-IR Re-ID

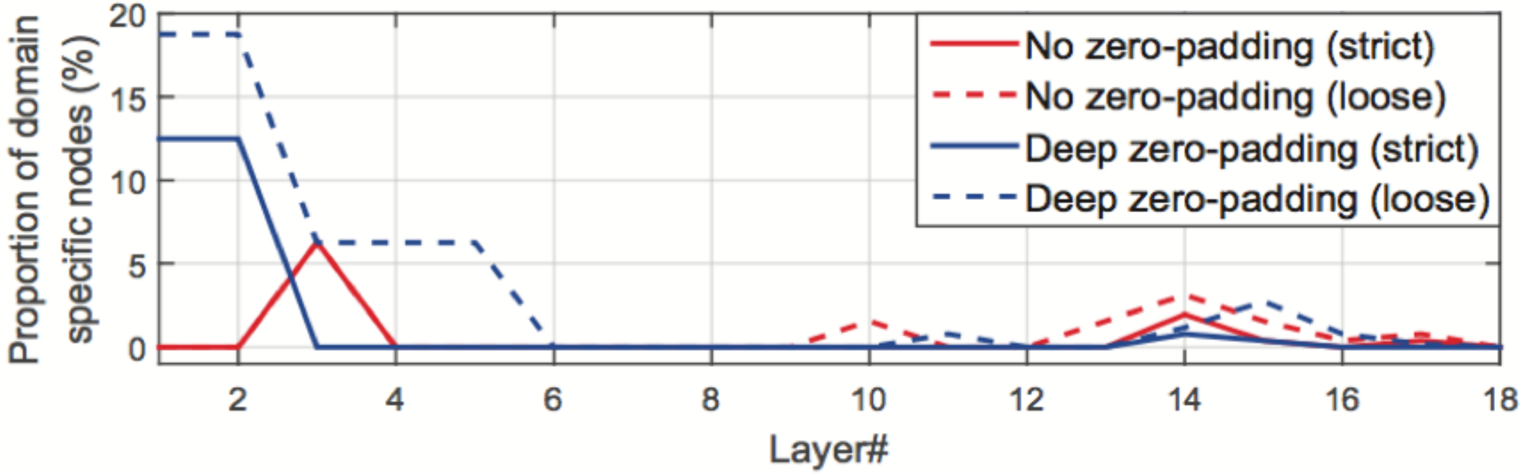


Figure 8. Relation between proportion of domain-specific nodes and layer depth. The x-axis denotes layer depth from bottom to top of the network, and the y-axis denotes the proportion of domain-specific nodes. The strict threshold is $T = 0.01\ std(x_i^{(l)})$ and the loose threshold is $T = 0.05\ std(x_i^{(l)})$ ($std(x_i^{(l)})$ is the standard deviation of the output of the i-th node in layer l). Generally, the proportion of domain-specific nodes using deep zero-padding is higher than that without zero-padding.

RGB-Infrared Re-ID

Cross-Modality Learning: RGB-IR Re-ID

- SYSU RGB-IR Re-ID Dataset

Feature	Metric	All-search								Indoor-search							
		Single-shot				Multi-shot				Single-shot				Multi-shot			
		r1	r10	r20	mAP	r1	r10	r20	mAP	r1	r10	r20	mAP	r1	r10	r20	mAP
One-stream network (deep zero-padding)	Euclidean	**14.80**	**54.12**	**71.33**	**15.95**	**19.13**	**61.40**	**78.41**	**10.89**	**20.58**	**68.38**	**85.79**	**26.92**	**24.43**	**75.86**	**91.32**	**18.64**
One-stream network	Euclidean	12.04	49.68	66.74	13.67	16.26	58.14	75.05	8.59	16.94	63.55	82.10	22.95	22.62	71.74	87.82	15.04
Asymmetric FC layer network	Euclidean	9.30	43.26	60.38	10.82	13.06	52.11	69.52	6.68	14.59	57.94	78.68	20.33	20.09	69.37	85.80	13.04
Lin's	GSM	5.29	33.71	52.95	8.00	6.19	37.15	55.66	4.38	9.46	48.98	72.06	15.57	11.36	51.34	73.41	9.03
HIPHOP	CRAFT	1.80	14.56	26.29	3.40	1.92	16.00	28.31	1.77	2.86	23.40	41.94	7.16	3.01	25.53	44.97	3.43
HOG	Euclidean	2.76	18.25	31.91	4.24	3.82	22.77	37.63	2.16	3.22	24.68	44.52	7.25	4.75	29.06	49.38	3.51
	KISSME	2.12	16.21	29.13	3.53	2.79	18.23	31.25	1.96	3.11	25.47	46.47	7.43	4.10	29.32	50.59	3.61
	LFDA	2.33	18.58	33.38	4.35	3.82	20.48	35.84	2.20	2.44	24.13	45.50	6.87	3.42	25.27	45.11	3.19
	CCA	2.74	18.91	32.51	4.28	3.25	21.82	36.51	2.04	4.38	29.96	50.43	8.70	4.62	34.22	56.28	3.87
	CDFE	2.09	16.68	30.51	3.75	2.47	19.11	34.11	1.86	2.80	23.39	44.46	6.91	3.28	27.31	48.61	3.24
	GMA	1.07	10.42	20.91	2.52	1.03	10.29	20.73	1.39	1.84	17.97	36.14	5.64	1.80	18.10	35.79	2.63
	SCM	1.86	15.16	28.27	3.57	2.40	17.45	31.22	1.66	3.30	25.82	46.23	7.52	3.90	28.84	51.64	3.22
	CRAFT	2.59	17.93	31.50	4.24	3.58	22.90	38.59	2.06	3.03	24.07	42.89	7.07	4.16	27.75	47.16	3.17
LOMO	Euclidean	1.75	14.14	26.63	3.48	1.96	15.06	27.30	1.85	2.24	22.53	41.53	6.64	2.24	22.79	41.80	3.31
	KISSME	2.23	18.95	32.67	4.05	2.65	20.36	34.78	2.45	3.83	31.09	52.86	8.94	4.46	34.35	58.43	4.93
	LFDA	2.98	21.11	35.36	4.81	3.86	24.01	40.54	2.61	4.81	32.16	52.50	9.56	6.27	36.29	58.11	5.15
	CCA	2.42	18.22	32.45	4.19	2.63	19.68	34.82	2.15	4.11	30.60	52.54	8.83	4.86	34.40	57.30	4.47
	CDFE	3.64	23.18	37.28	4.53	4.70	28.23	43.05	2.28	5.75	34.35	54.90	10.19	7.36	40.38	60.33	5.64
	GMA	1.04	10.45	20.81	2.54	0.99	10.50	21.06	1.47	1.79	17.90	36.01	5.63	1.71	18.11	36.17	2.88
	SCM	1.54	14.12	26.27	3.34	1.66	15.17	28.41	1.57	2.86	24.34	44.53	7.06	2.89	25.81	48.33	3.02
	CRAFT	2.34	18.70	32.93	4.22	3.03	21.70	37.05	2.13	3.89	27.55	48.16	8.37	2.45	20.20	38.15	2.69

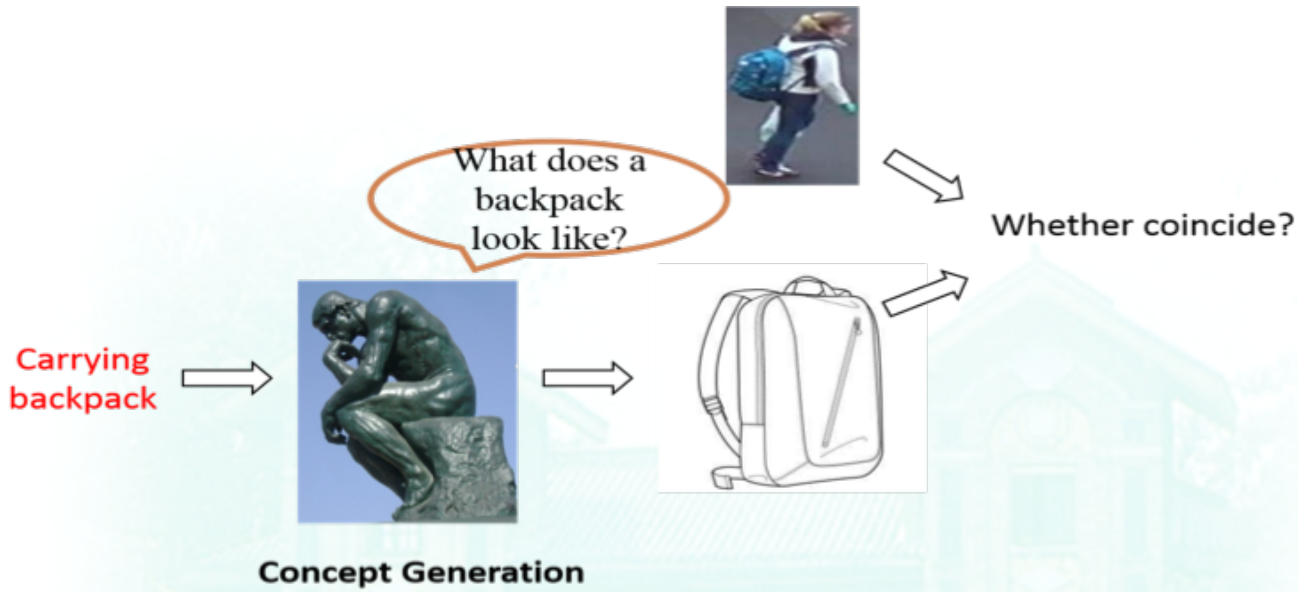
When the input is not image?

Attribute-Image Person Re-ID

- Match person images with specific attribute descriptions in surveillance environment.

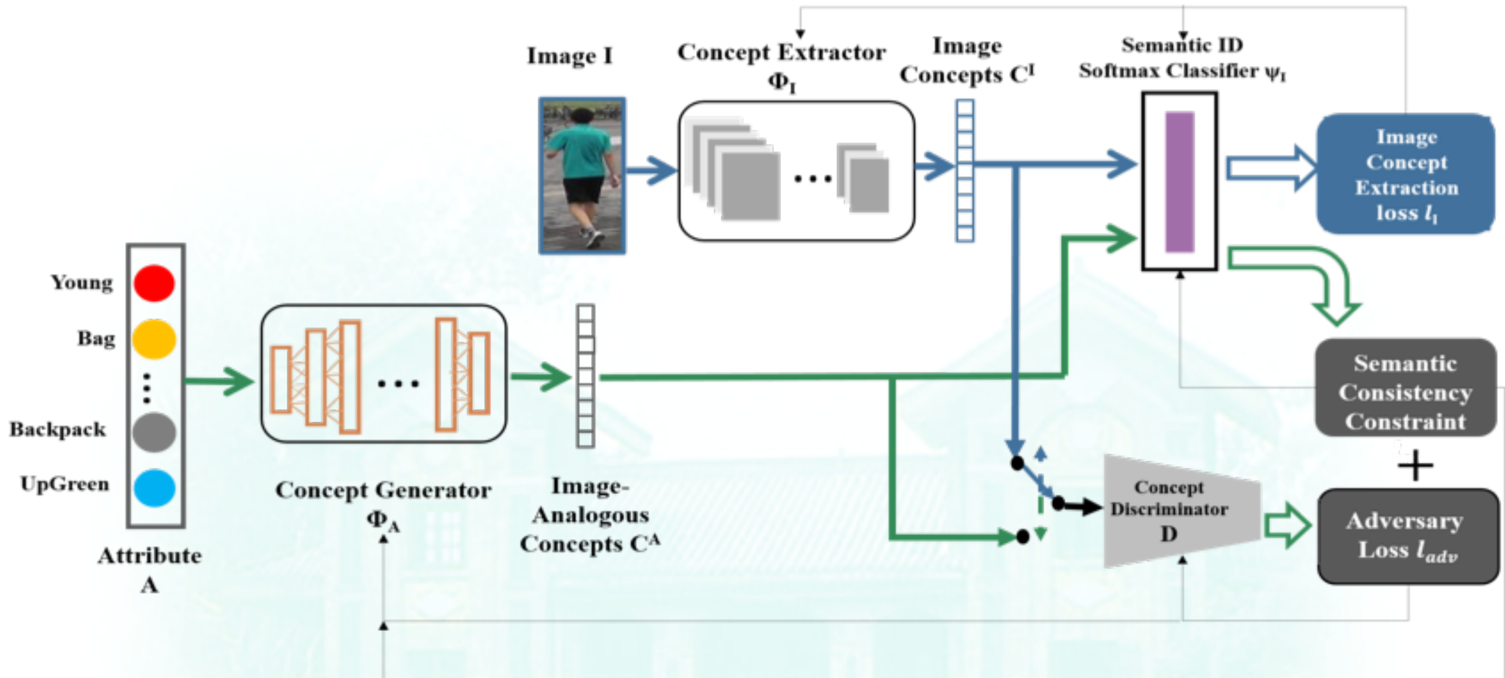
Zhou Yin, Wei-Shi Zheng*(PI), et al. Adversarial Attribute-Image Person Re-identification, IJCAI 2018

Attribute-Image Person Re-ID



- Intuitively, when we hold some attribute description in mind, e.g., “carrying backpack”, we generate an obscure and vague imagination on how a backpack may look like, which we refer to as a concept.
- We model this generation process and match the generated concepts with image perceptions.

Attribute-Image Person Re-ID



- **Image Concept Extraction loss** l_I: Our model learns a semantically discriminative structure of low-level person images.
- **Semantic Consistency Constraint + Adversary loss** l_{adv}: Our model generates the corresponding aligned image-analogous concept for high-level attribute.

Attribute-Image Person Re-ID

Method	Market				Duke				PETA			
	rank1	rank5	rank10	mAP	rank1	rank5	rank10	mAP	rank1	rank5	rank10	mAP
DeepCCAE [Wang *et al.*, 2015]	8.12	23.97	34.55	9.72	33.28	59.35	67.64	14.95	14.24	22.09	29.94	14.45
DeepCCA [Andrew *et al.*, 2013]	29.94	**50.70**	**58.14**	17.47	36.71	58.79	65.11	13.53	14.44	20.77	26.31	11.49
2WayNet [Eisenschtat and Wolf, 2017]	11.29	24.38	31.47	7.76	25.24	39.88	45.92	10.19	23.73	38.53	41.93	15.38
DeepMAR [Li *et al.*, 2015]	13.15	24.87	32.90	8.86	36.60	57.70	67.00	14.34	17.80	25.59	31.06	12.67
CMCE [Li *et al.*, 2017]	35.04	**50.99**	56.47	**22.80**	39.75	56.39	62.79	15.40	31.72	39.18	48.35	26.23
ours w/o adv	33.83	48.17	53.48	17.82	39.30	55.88	62.50	15.17	36.34	48.48	53.03	25.35
ours w/o sc	2.08	4.80	4.80	1.00	5.26	9.37	10.87	1.56	3.43	4.15	4.15	5.80
ours w/o adv+MMD	34.15	47.96	57.20	18.90	41.77	**62.32**	**68.61**	14.23	**39.31**	48.28	**54.88**	**31.54**
ours w/o adv+DeepCoral	**36.56**	47.61	55.92	20.08	**46.09**	**61.02**	68.15	**17.10**	35.62	**48.65**	53.75	27.58
ours	**40.26**	49.21	**58.61**	**20.67**	**46.60**	59.64	**69.07**	**15.67**	**39.00**	**53.62**	**62.20**	**27.86**

Our model:

- Outperforms traditional cross modality retrieval methods (DeepCCAE, DeepCCA, 2WayNet, CMCE).
- Outperforms classical pedestrian attribute recognition model (DeepMAR).
- Outperforms other variants of our model, which also generate homogenous distributions under semantic consistency regularization for the two modalities (MMD, DeepCoral).

Attribute-Image Person Re-ID

Method	Market				Duke				PETA			
	rank1	rank5	rank10	mAP	rank1	rank5	rank10	mAP	rank1	rank5	rank10	mAP
DeepCCAE [Wang *et al.*, 2015]	8.12	23.97	34.55	9.72	33.28	59.35	67.64	14.95	14.24	22.09	29.94	14.45
DeepCCA [Andrew *et al.*, 2013]	29.94	50.70	58.14	17.47	36.71	58.79	65.11	13.53	14.44	20.77	26.31	11.49
2WayNet [Eisenschtat and Wolf, 2017]	11.29	24.38	31.47	7.76	25.24	39.88	45.92	10.19	23.73	38.53	41.93	15.38
DeepMAR [Li *et al.*, 2015]	13.15	24.87	32.90	8.86	36.60	57.70	67.00	14.34	17.80	25.59	31.06	12.67
CMCE [Li *et al.*, 2017]	35.04	50.99	56.47	22.80	39.75	56.39	62.79	15.40	31.72	39.18	48.35	26.23
ours w/o adv	33.83	48.17	53.48	17.82	39.30	55.88	62.50	15.17	36.34	48.48	53.03	25.35
ours w/o sc	2.08	4.80	4.80	1.00	5.26	9.37	10.87	1.56	3.43	4.15	4.15	5.80
ours w/o adv+MMD	34.15	47.96	57.20	18.90	41.77	62.32	68.61	14.23	39.31	48.28	54.88	31.54
ours w/o adv+DeepCoral	36.56	47.61	55.92	20.08	46.09	61.02	68.15	17.10	35.62	48.65	53.75	27.58
ours	40.26	49.21	58.61	20.67	46.60	59.64	69.07	15.67	39.00	53.62	62.20	27.86

Part of query attributes

Teenager	Backpack	Bag	DownBlue	UpWhite	UpPink
☑	☒	☒	☑	☑	☒

Attribute-Image Person Re-ID

Effects of different generation strategies:

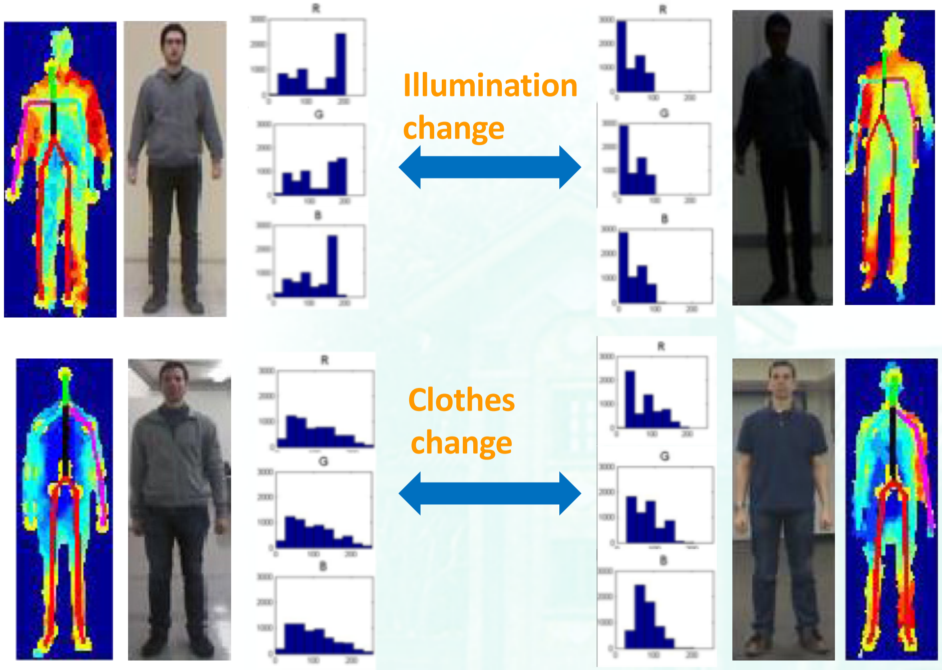
Method	Market	Duke	PETA
A2Img (proposed)	**40.3**	**46.6**	39.0
Img2A (reverse of the proposed)	36.0	43.7	**43.6**
Real Images	8.13	20.01	19.85

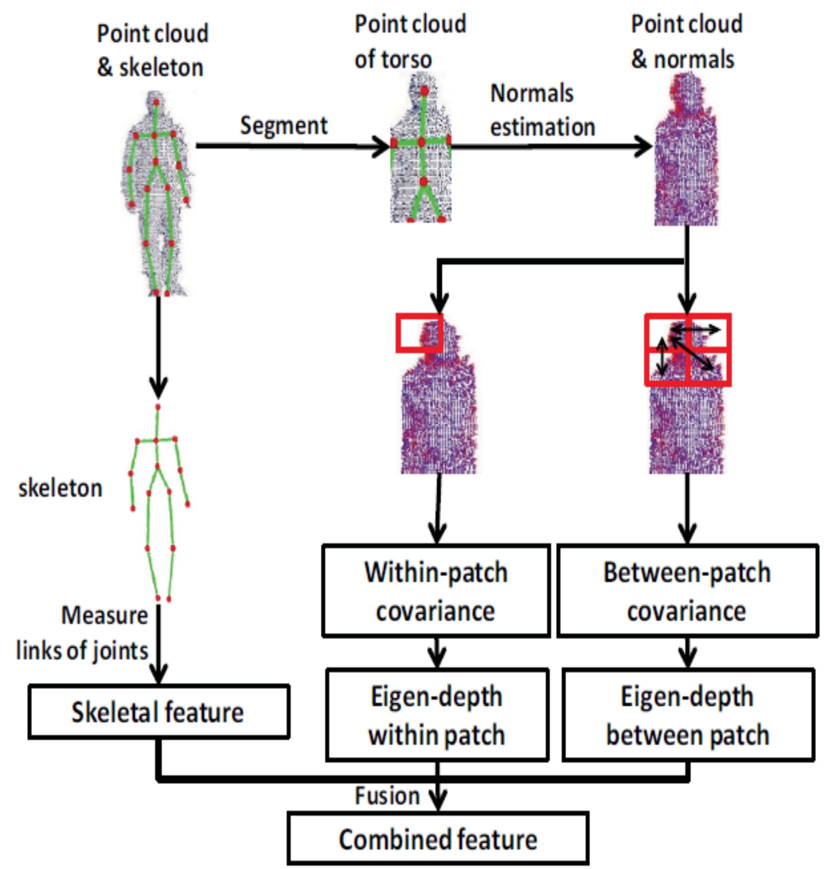
- Generation from attributes to image is better than generation from image to attributes. (A2Img vs. Img2A):
  - Estimating the manifold of images from the training data is more reliable than estimating that of attributes
- Generation in feature space is better than generation in real image space. (A2Img vs. Real Images):
  - Generating real pedestrian image is difficult. Generating noisy low-level images and then eliminating these noise to extract discriminative concepts is not necessary

When dressing differently?

Depth Re-ID

- Something to see





In these cases, appearance cues are not reliable.

Depth Re-ID

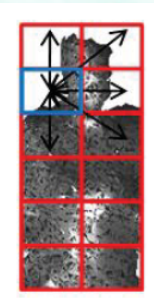
❑ Depth descriptors

- **Within-patch Covariance**

$$C_1 = \frac{1}{m-1}\sum_{i=1}^{m}(f_{1i}-\mu_1)(f_{1i}-\mu_1)^T,$$

where μ_1 is the mean of the feature vectors of R_1.

- **Between-patch Covariance**



$$C_{12} = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}(f_{1i}-f_{2j})(f_{1i}-f_{2j})^T.$$

- **Eigen-depth feature**

$$\mathbf{x}_p = [\ln\lambda_{p,1}\ \ln\lambda_{p,2}\ \ldots\ \ln\lambda_{p,6}]^T$$

Eigen-depth feature is rotation invariant.

Depth Re-ID

❑ Metric

Theorem 1: Computing the Euclidean distance between $\mathbf{x}_1$ and $\mathbf{x}_2$ is equivalent to computing the geodesic distance between covariance matrix $\mathbf{C}_1$ and the rotation normalized covariance matrix $\mathbf{C}_2^N$ on the Riemannian manifold.

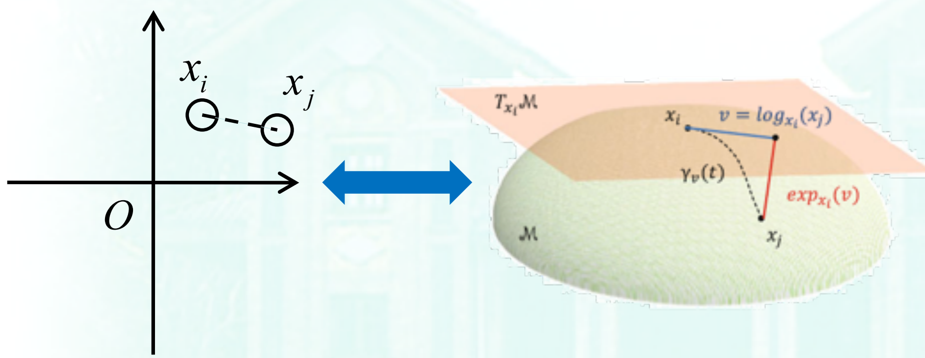
$$\mathbf{C}_1 = \mathbf{U}_1 \mathrm{diag}(\lambda_{1,1}, \lambda_{1,2}, ..., \lambda_{1,6}) \mathbf{U}_1^T$$

$$\mathbf{C}_2 = \mathbf{U}_2 \mathrm{diag}(\lambda_{2,1}, \lambda_{2,2}, ..., \lambda_{2,6}) \mathbf{U}_2^T$$

$$\mathbf{C}_2^N = \mathbf{U}_1 \mathrm{diag}(\lambda_{2,1}, \lambda_{2,2}, ..., \lambda_{2,6}) \mathbf{U}_1^T$$

$$\mathbf{x}_1 = [\ln\lambda_{1,1} \ \ln\lambda_{1,2} \ ... \ \ln\lambda_{1,6}]^T$$

$$\mathbf{x}_2 = [\ln\lambda_{2,1} \ \ln\lambda_{2,2} \ ... \ \ln\lambda_{2,6}]^T$$



$$\|\mathbf{x}_2 - \mathbf{x}_1\|_2 = \sqrt{\sum_{i=1}^{6} (\ln\lambda_{2,i} - \ln\lambda_{1,i})^2} = \sqrt{\sum_{i=1}^{6} \ln^2 \frac{\lambda_{2,i}}{\lambda_{1,i}}}$$

$$\mathbf{C}_1^{-1}\mathbf{C}_2^N = (\mathbf{U}_1 \mathrm{diag}(\lambda_{1,i}) \mathbf{U}_1^T)^{-1} (\mathbf{U}_1 \mathrm{diag}(\lambda_{2,i}) \mathbf{U}_1^T)$$
$$= \mathbf{U}_1 \mathrm{diag}(\frac{\lambda_{2,i}}{\lambda_{1,i}}) \mathbf{U}_1^T.$$

$$dist(C_1, C_2^N) = \sqrt{\sum_{k=1}^{6} ln^2 \lambda_k(C_1, C_2^N)}$$

$$\lambda_k \mathbf{C}_1 \mathbf{z}_k - \mathbf{C}_2^N \mathbf{z}_k = 0.$$

Extracting Eigen-depth feature converts covariance matrices on Riemannian manifold to feature vectors in Euclidean space.

Depth Re-ID

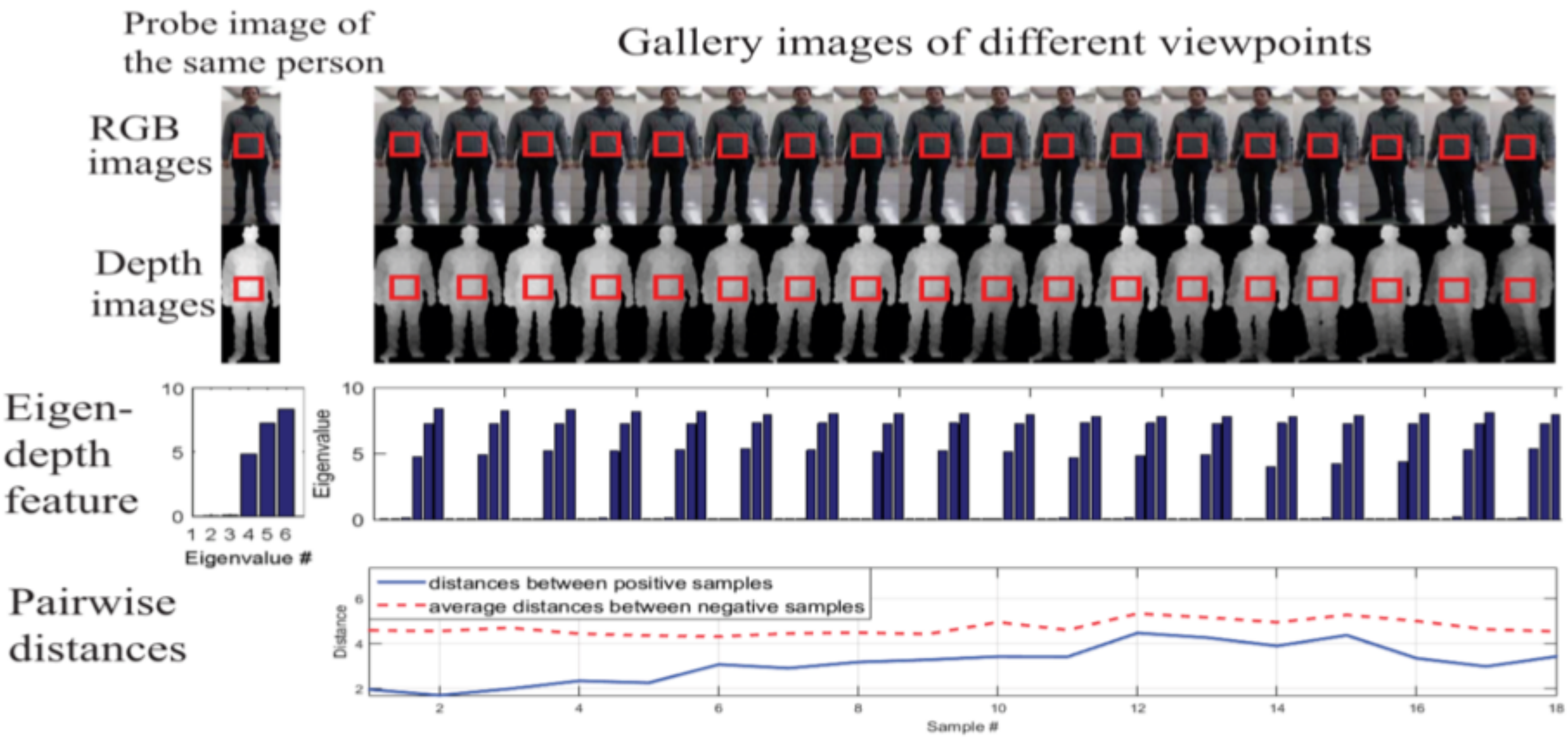
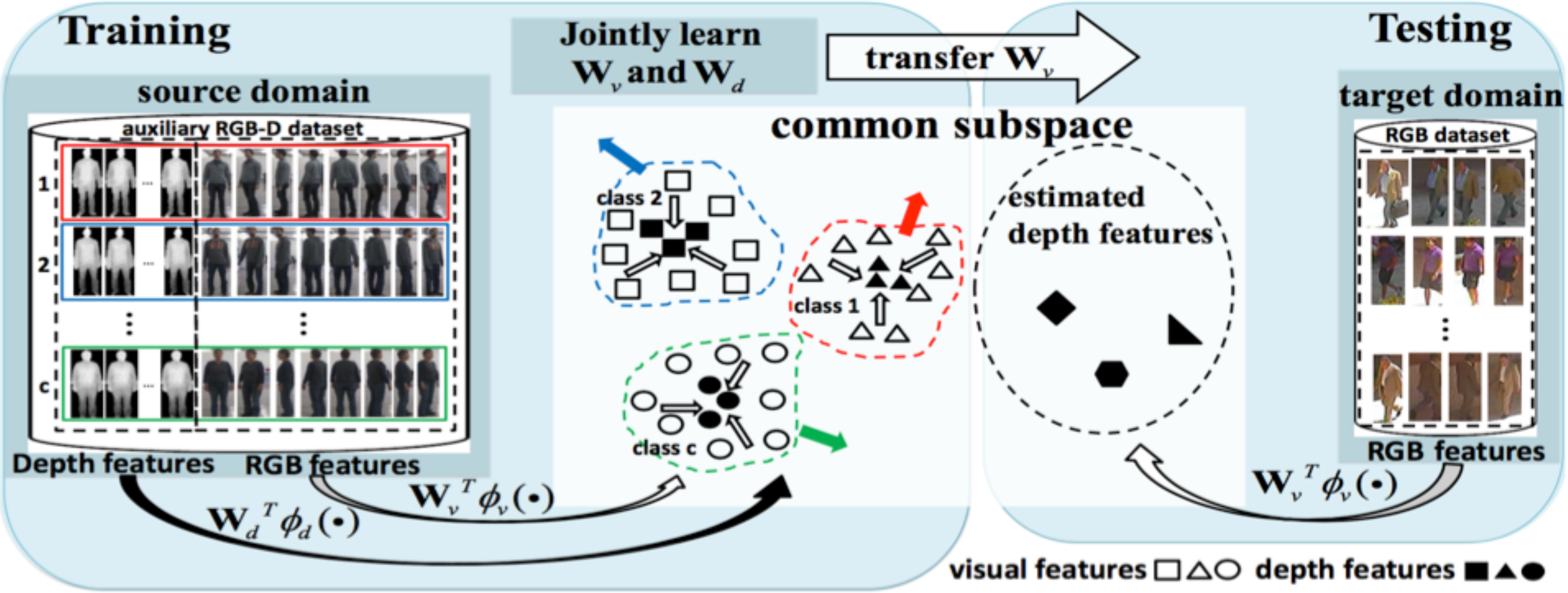


Fig. 4. Visualization of the logarithms of eigenvalues of a fixed voxel and comparison of distance of positive pair (i.e., samples from the same class) and average distance of negative pair (i.e., samples from different classes). The first row shows RGB and depth images of the same person. The second row shows within-voxel Eigen-depth feature of the fixed voxel indicated by red bounding boxes. The third row shows the comparison between the distance of positive pair and the distance of negative pair.

Depth Re-ID

Transferring Depth



学习RGB信息、深度(Depth)信息、骨架之间的关联关系

Dataset	Probe	ED	ED+SKL	3D RAM [4]	PCM [3]	PCM+SKL [3]	SKL [3]
PAVIS	Walking2	54.4	**57.0**	41.3	-	-	28.6
IAS-Lab	TestingA	44.0	**49.9**	48.3	28.6	25.6	22.5
RGBD-ID	TestingB	55.5	**66.6**	63.7	43.7	63.3	55.5

Ancong Wu, **Wei-Shi Zheng*(PI),** and Jian-Huang Lai. Robust Depth-based Person Re-identification. IEEE Transactions on Image Processing, 2017

Depth Re-ID

Fig. 10. Examples of images in IAS-Lab RGBD-ID. All samples were captured from multiple views. Compared to "Training", samples in "TestingA" changed clothes and some samples in "TestingB" were captured in dark environment.

Fig. 6. Examples of images in "Walking1" and "Walking2" in PAVIS. Most persons in "Walking2" dressed different clothes from "Walking1".

TABLE III

PAVIS DATASET: RANK-1 AND RANK-5 ACCURACIES (%), INCLUDING RESULTS OF OUR PROPOSED METHODS AND COMPARISONS WITH RGB-BASED APPEARANCE FEATURES AND DEPTH-BASED FEATURES

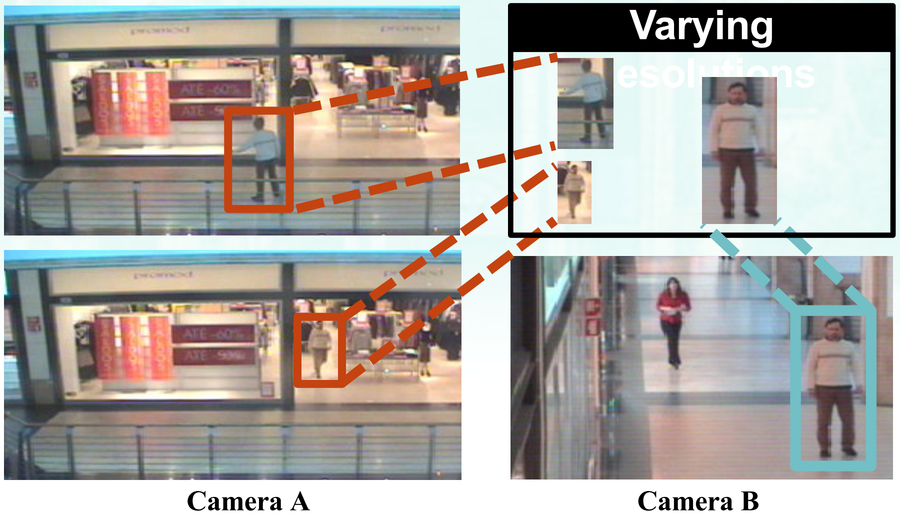
Setting	Single-shot		Multi-shot	
Rank	1	5	1	5
RGB-based appearance features				
LOMO [12]	12.05	35.03	19.74	44.36
ELF18 [50]	52.15	77.85	52.62	78.26
Color Hist [11]	47.90	74.97	48.92	74.82
HOG [13]	45.03	73.49	45.33	73.95
LBP [80]	42.92	71.33	45.64	72.36
Depth-based features				
RIFT2M [5]	7.13	22.77	8.77	27.69
Fehr's [6]	24.26	51.64	30.56	58.67
Skeleton [2]	33.13	67.85	37.33	71.13
Proposed				
DVCov (depth voxel covariance)	61.49	81.23	66.00	82.92
DVCov+SKL	**67.64**	**87.33**	**71.74**	**88.46**
ED (Eigen-depth feature)	44.67	72.10	51.59	76.15
ED+SKL	55.95	84.77	61.23	87.64

TABLE IV

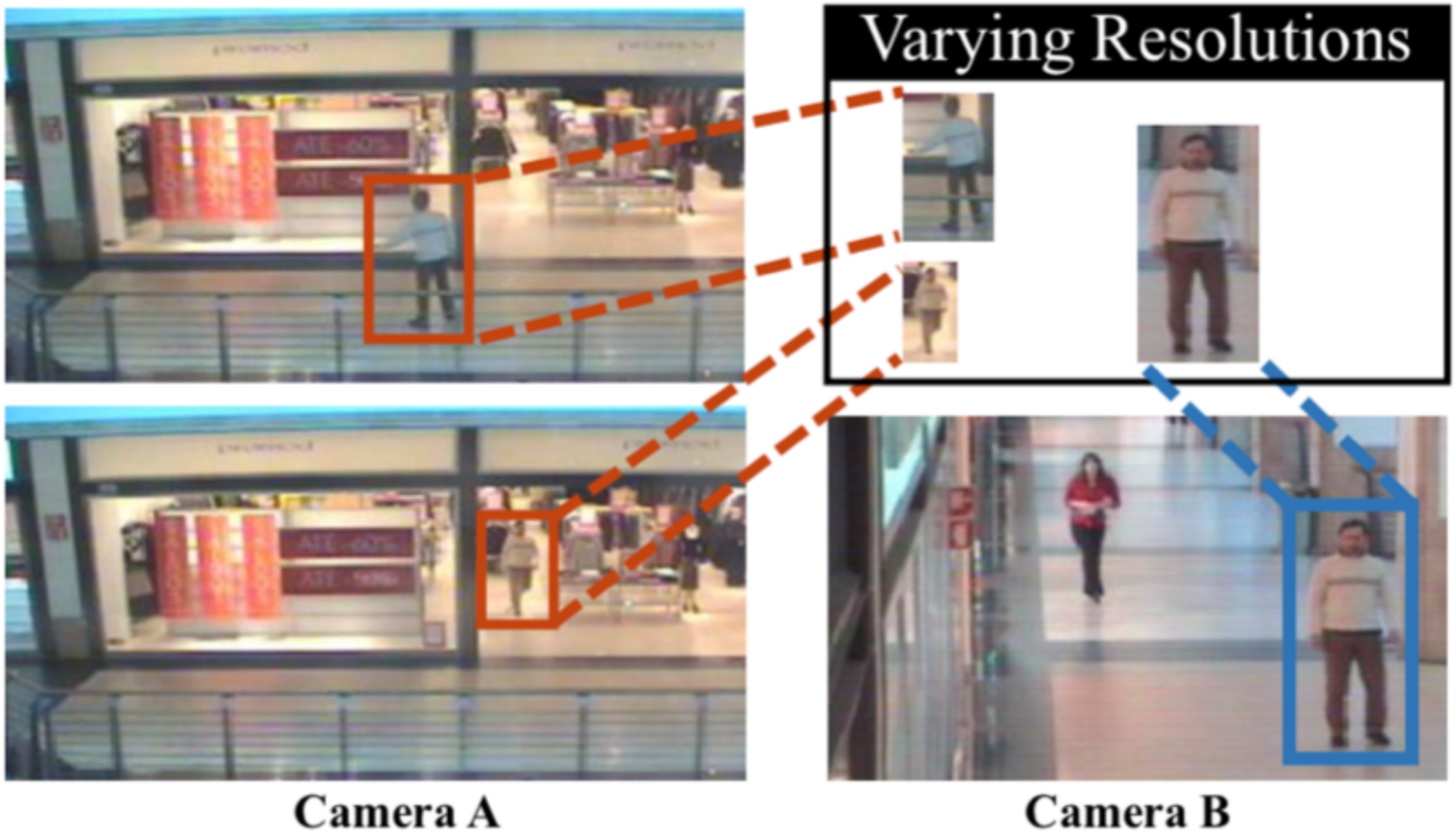
BIWI RGBD-ID DATASET "STILL" AND "WALKING": RANK-1 AND RANK-5 ACCURACIES (%), INCLUDING RESULTS OF OUR PROPOSED METHODS AND COMPARISONS WITH RGB-BASED APPEARANCE FEATURES AND DEPTH-BASED FEATURES

Probe	Still				Walking			
Setting	Single-shot		Multi-shot		Single-shot		Multi-shot	
Rank	1	5	1	5	1	5	1	5
RGB-based appearance features								
LOMO [12]	9.07	28.21	18.17	35.47	8.74	23.33	10.31	25.39
ELF18 [50]	2.79	18.18	4.11	19.13	1.32	16.03	1.50	16.77
Color Hist [11]	7.02	25.47	10.61	31.92	5.43	19.56	5.86	21.70
HOG [13]	8.42	25.69	12.35	30.39	6.38	21.00	6.94	23.29
LBP [80]	7.37	26.04	10.87	33.57	4.87	20.04	5.34	23.31
Depth-based features								
RIFT2M [5]	4.04	19.52	4.34	20.78	3.25	17.46	3.75	18.31
Fehr's [6]	12.08	38.17	14.06	43.78	9.33	32.39	12.09	39.60
Skeleton [2]	21.34	53.32	26.55	62.73	14.52	42.36	16.94	47.18
Proposed								
DVCov	16.32	45.93	23.07	58.89	12.58	39.22	17.24	45.93
DVCov+SKL	23.49	57.06	34.37	72.77	16.59	46.67	21.40	54.12
ED	28.98	61.85	36.22	**73.11**	20.90	51.98	28.71	63.85
ED+SKL	**30.52**	**67.86**	**39.38**	72.13	**24.47**	**60.63**	**29.96**	**65.18**

3. Low-resolution Person Re-identification



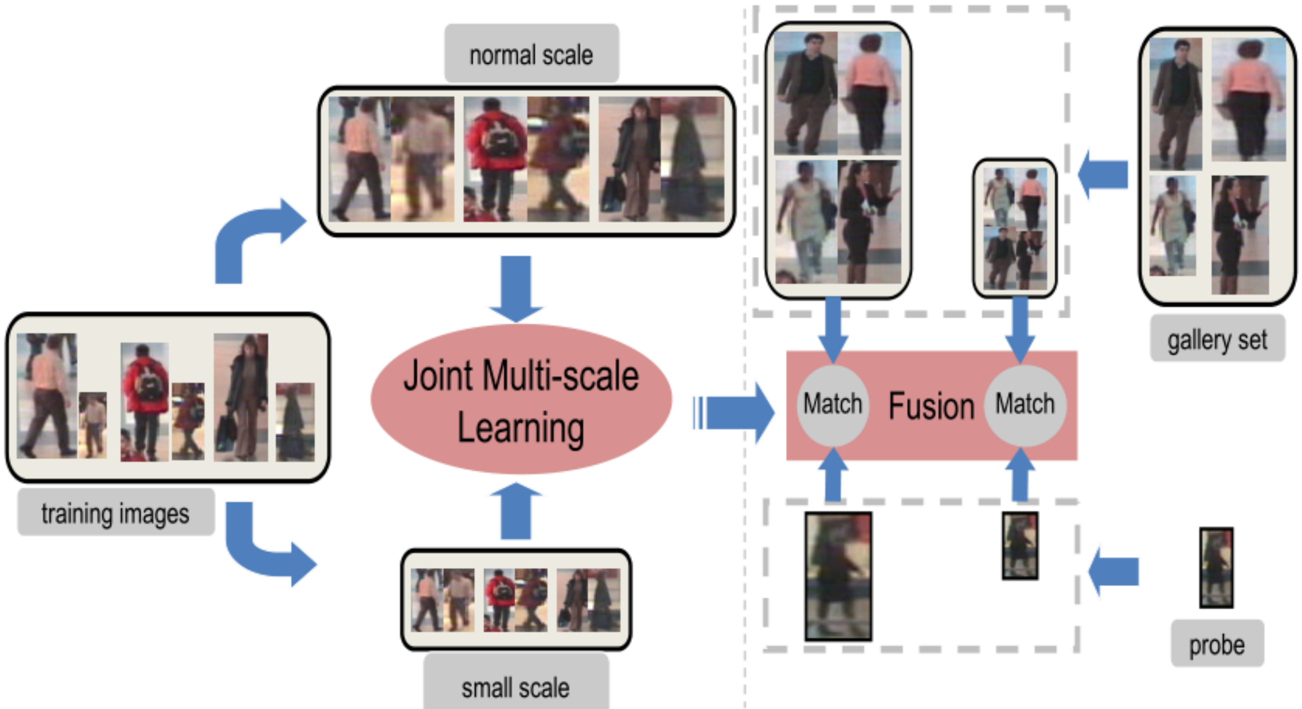
Low-resolution Re-ID



Low-resolution Re-ID

❑ Low-resolution Re-ID

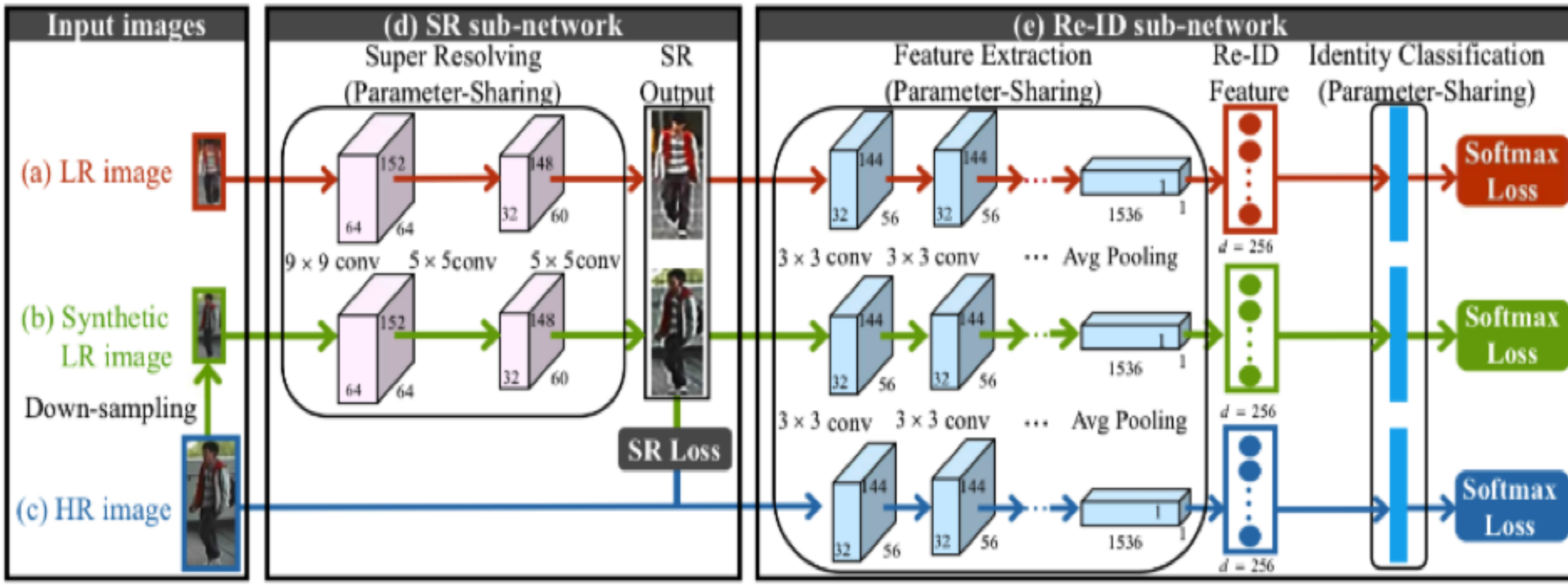
- JUDEA : joint multi-scale discriminant component analysis



Xiang Li, Wei-Shi Zheng*(PI), Xiaojuan Wang, Tao Xiang, Shaogang Gong. Multi-scale Learning for Low-resolution Person Re-identification. IEEE Conf. on Computer Vision (ICCV), 2015.

Low-resolution Re-ID

- Super- resolution and Identity joiNt learninG (SING)



Jiening Jiao, Wei-Shi Zheng*(PI), Ancong Wu, Xiatian Zhu, and Shaogang Gong. Deep Low-resolution Person Re-identification. AAAI 2018

Low-resolution Re-ID

Qualitative examples of super-resolved person images.

Low-resolution Re-ID

Results

Table 2: Comparing combinations of image super-resolution and person re-id schemes (%).

Super-Resolution Method	Re-ID Method	CAVIAR				MLR-CUHK03				MLR-SYSU				MLR-VIPeR			
		r=1	r=5	r=10	r=20	r=1	r=5	r=10	r=20	r=1	r=5	r=10	r=20	r=1	r=5	r=10	r=20
Bilinear	XQDA	22.7	61.4	81.9	98.8	45.5	78.0	87.8	93.7	39.3	67.4	77.2	85.6	37.6	65.7	78.5	89.7
Bicubic	XQDA	24.3	63.4	83.1	**99.2**	45.1	78.1	87.7	93.3	40.0	66.9	77.2	85.5	37.8	66.0	78.9	89.3
SRCNN	XQDA	24.8	64.3	84.0	**99.1**	44.7	77.8	87.5	93.1	40.3	67.3	77.4	85.6	36.5	65.1	78.9	89.8
Bilinear	NFST	23.3	60.5	82.2	99.0	48.0	47.9	46.2	49.0	41.6	69.0	79.5	87.7	**39.7**	**68.4**	**81.0**	**90.4**
Bicubic	NFST	24.5	61.1	82.1	99.2	47.9	74.8	83.6	92.8	42.4	69.3	**80.5**	**88.0**	**39.2**	**67.9**	**80.3**	**90.7**
SRCNN	NFST	25.0	61.2	82.9	99.2	49.0	74.8	85.0	92.1	**43.2**	**69.6**	80.3	88.0	38.6	67.1	79.5	90.1
Bilinear	DGD	25.3	61.0	82.6	98.4	58.5	86.0	92.2	96.0	39.6	66.4	74.8	82.5	23.1	45.9	56.6	67.7
Bicubic	DGD	27.4	63.4	83.0	98.5	62.5	88.7	93.7	96.5	41.5	67.4	76.9	84.7	25.0	51.3	59.2	69.3
SRCNN	DGD	**28.4**	**66.3**	**85.9**	98.5	**63.8**	**89.3**	**93.9**	**96.8**	42.6	68.2	77.1	85.5	25.3	48.4	57.3	66.5
SING		**33.5**	**72.7**	**89.0**	98.6	**67.7**	**90.7**	**94.7**	**97.4**	**50.7**	**75.4**	**83.1**	**88.1**	33.5	57.0	66.5	76.6

Low-resolution Re-ID

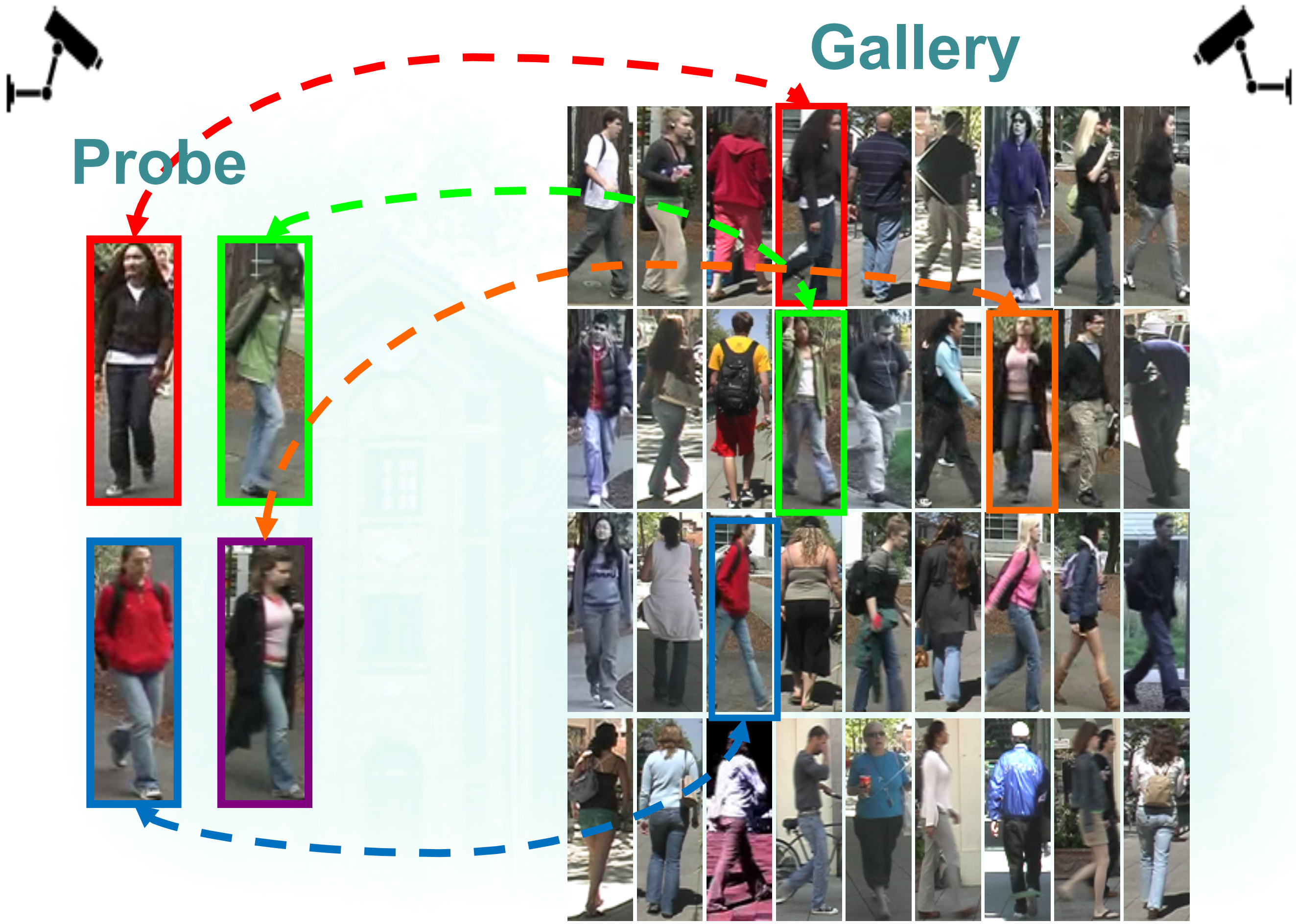
❑ Results

Table 1: Comparing state-of-the-art LR re-id methods (%). The $1^{st}/2^{nd}$ best results are indicated in red/blue.

CAVIAR	r=1	r=5	r=10	r=20
JUDEA	22.0	60.1	80.8	98.1
SLD^2L	18.4	44.8	61.2	83.6
SDF	14.3	37.5	62.5	95.2
SING	33.5	72.7	89.0	98.6
MLR-CUHK03	**r=1**	**r=5**	**r=10**	**r=20**
JUDEA	26.2	58.0	73.4	87.0
SLD^2L	-	-	-	-
SDF	22.2	48.0	64.0	80.0
SING	67.7	90.7	94.7	97.4
MLR-SYSU	**r=1**	**r=5**	**r=10**	**r=20**
JUDEA	18.3	41.9	54.5	68.0
SLD^2L	20.3	34.8	43.4	55.4
SDF	13.3	26.7	42.9	66.7
SING	50.7	75.4	83.1	88.1
MLR-VIPeR	**r=1**	**r=5**	**r=10**	**r=20**
JUDEA	26.0	55.1	69.2	82.3
SLD^2L	20.3	44.0	62.0	78.2
SDF	9.52	38.1	52.4	68.0
SING	33.5	57.0	66.5	76.6

More ……

Cross-set Re-ID

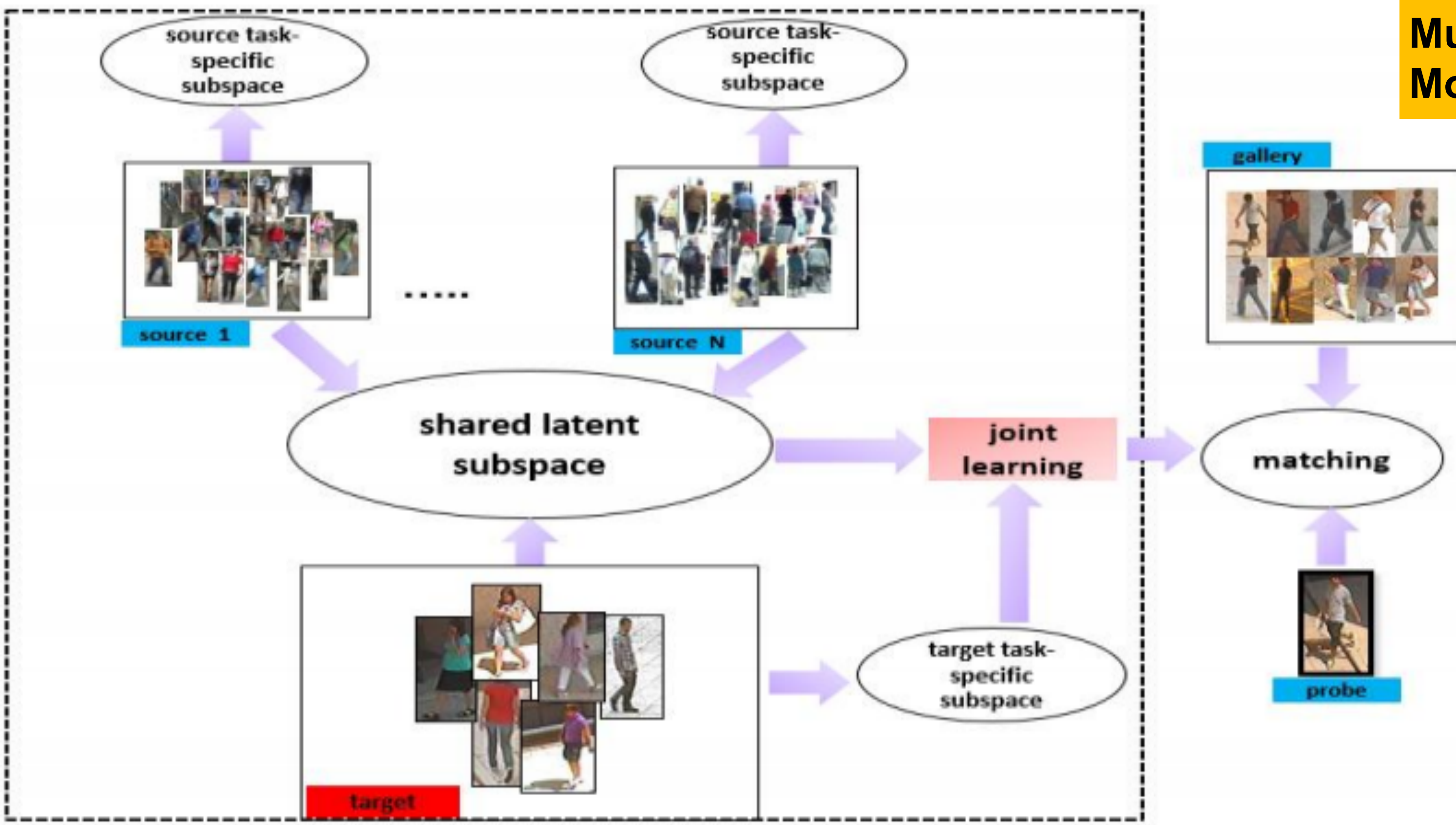


Labelling images across camera views is costly

Cross-scenario Re-ID

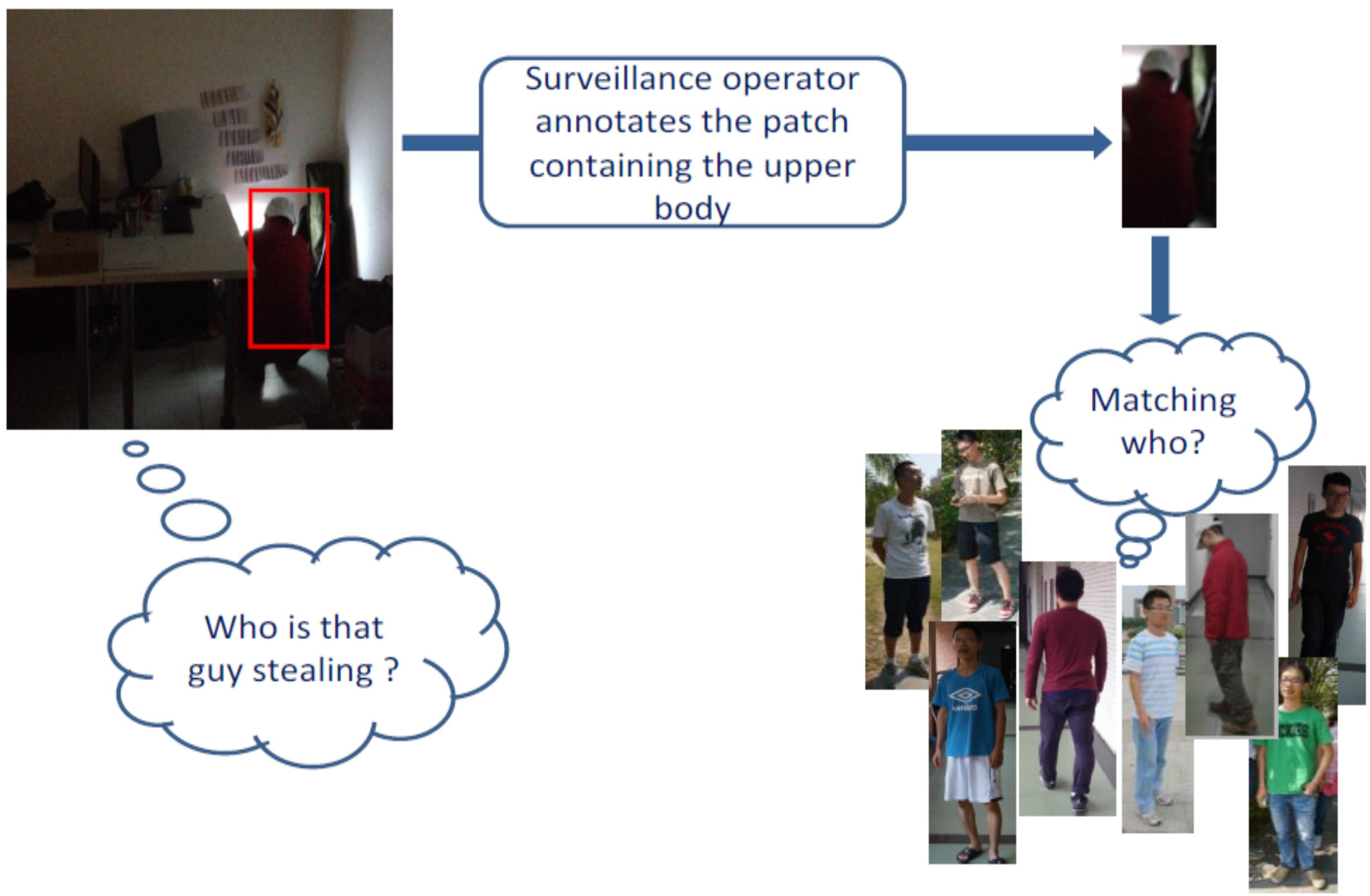
- Transferring between sets

An Asymmetric Multi-task Modelling

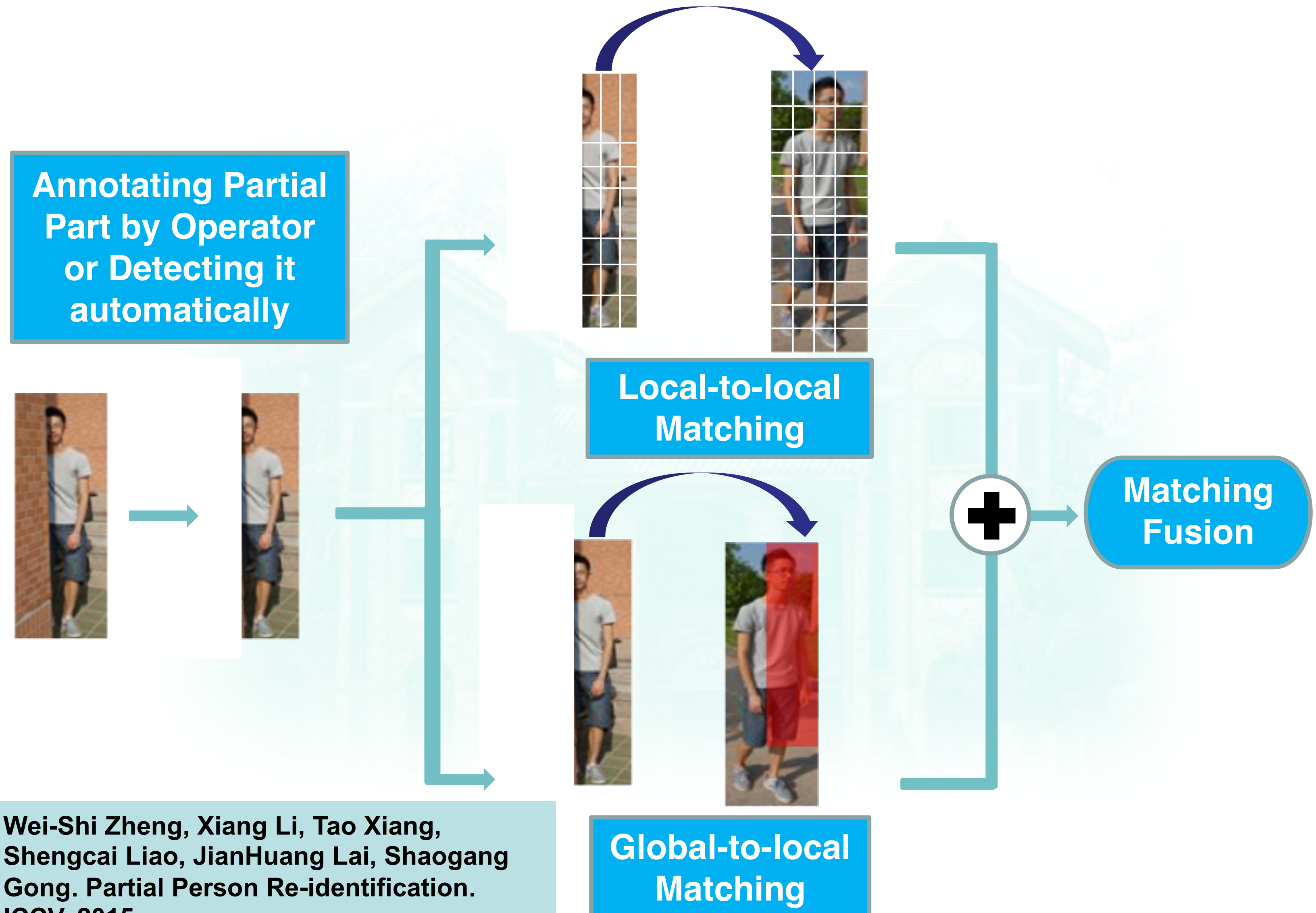


Xiaojuan Wang, Wei-Shi Zheng*(PI), Xiang Li, and Jianguo Zhang. Cross-scenario Transfer Person Re-identification. IEEE Transactions on Circuits and Systems for Video Technology, vol. 26, no. 8, pp. 1447-1460, 2016.

Partial Re-ID



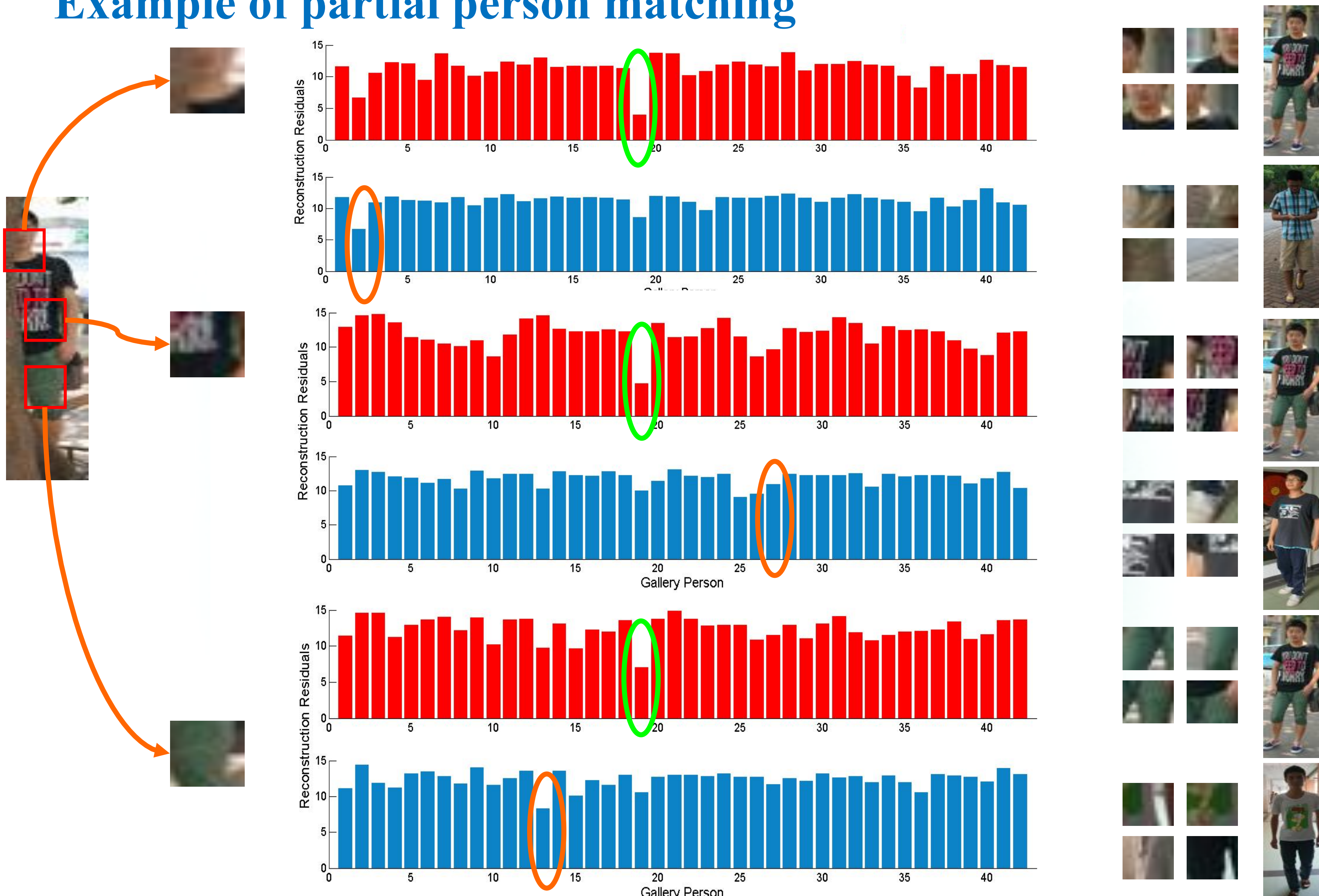
Partial Re-ID



Wei-Shi Zheng, Xiang Li, Tao Xiang, Shengcai Liao, JianHuang Lai, Shaogang Gong. Partial Person Re-identification. ICCV, 2015.

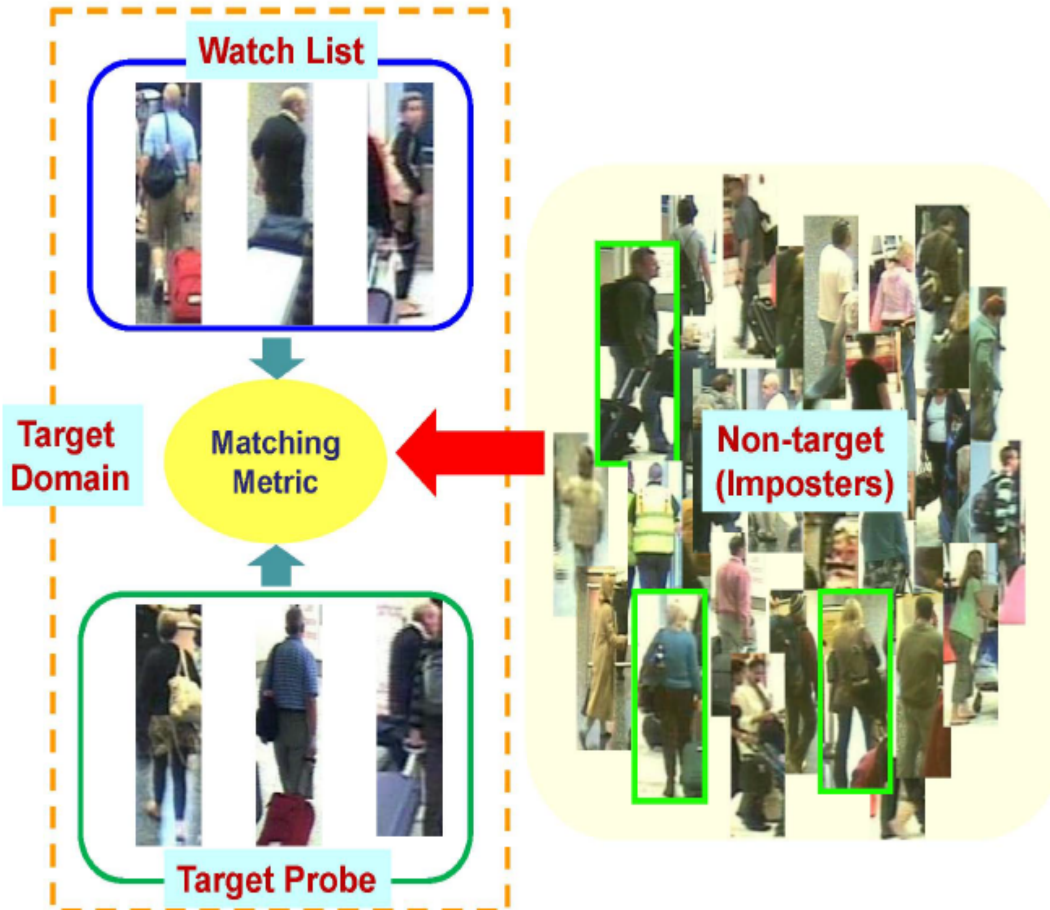
Partial Re-ID

Example of partial person matching



One-Shot Open-World Group-based Re-id

❑ Motivation



Open-world person re-identification setting

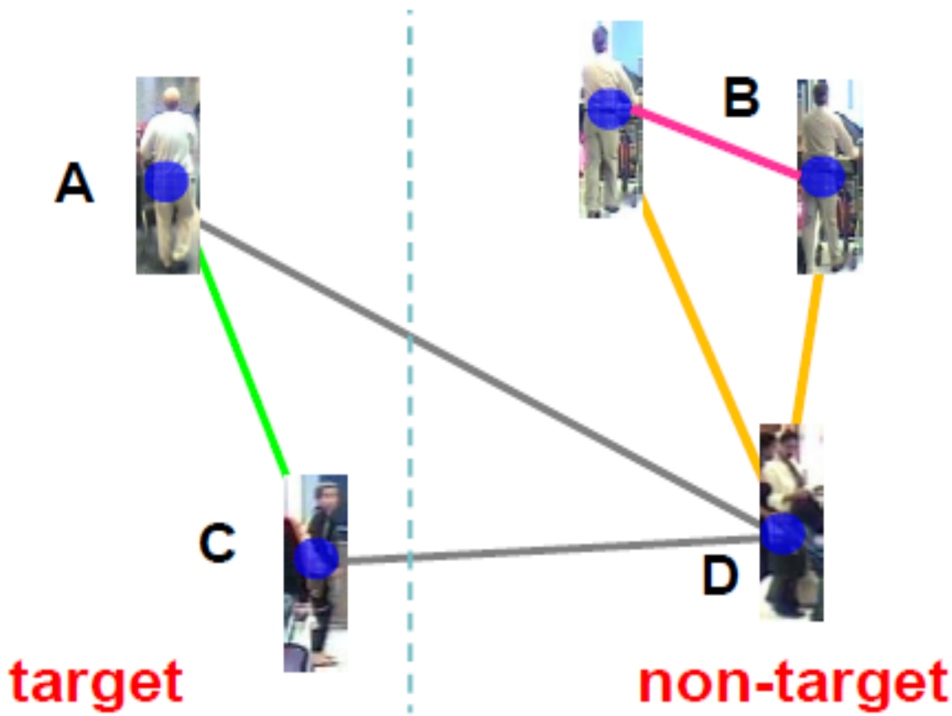
1) A large amount of non-target imposters captured along with the target people on the watch list.

2) Their images will also appear in the probe set and some of them will look visually similar to the target people

Wei-Shi Zheng, Shaogang Gong, and Tao Xiang. Towards Open-World Person Re-Identification by One-Shot Group-based Verification. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), vol. 38, no. 3, pp. 591-606, 2016.

One-Shot Open-World Group-based Re-id

- Knowledge to transfer



Enrich intra-class variation

Approximate target intra-inter class pair (magenta line and green line)

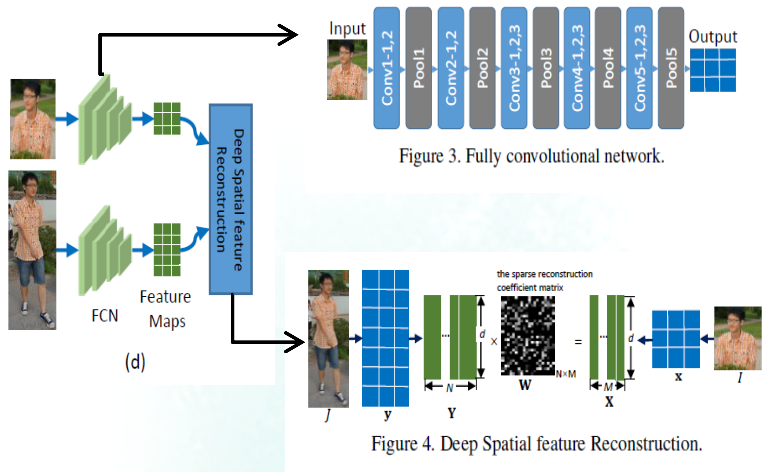
Enrich inter-class variation

Target specific non-target intra-inter class pair
(magenta line and yellow line)

Enrich group separation

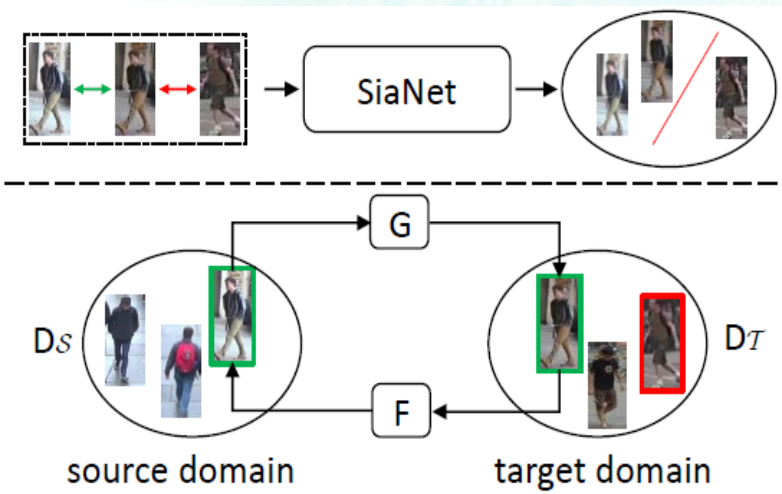
Group separation intra-inter class pair
(green line and grey line)

More



Lingxiao He, et al., CVPR 2018.

S.Li et al., CVPR 2017



Weijian Deng, et al,. CVPR, 2018.

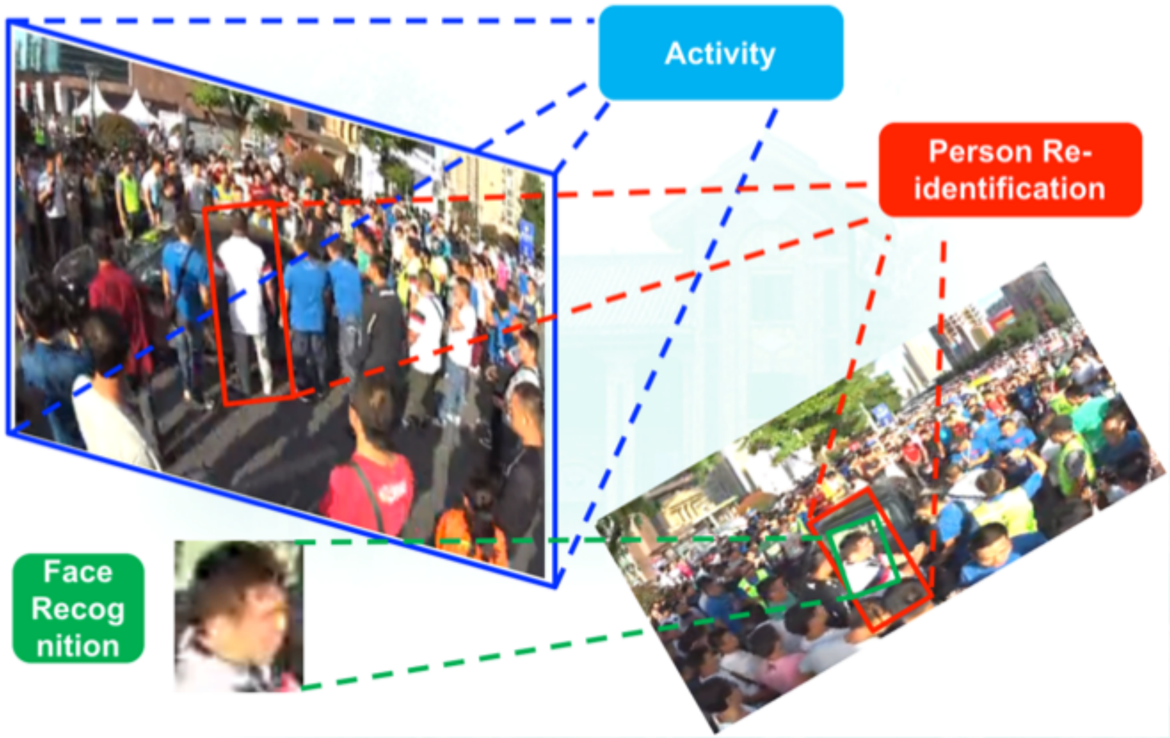
Longhui Wei, et al,. CVPR, 2018.

Summary

Connection with Machine Learning for Cross Modalities/Cross Domains

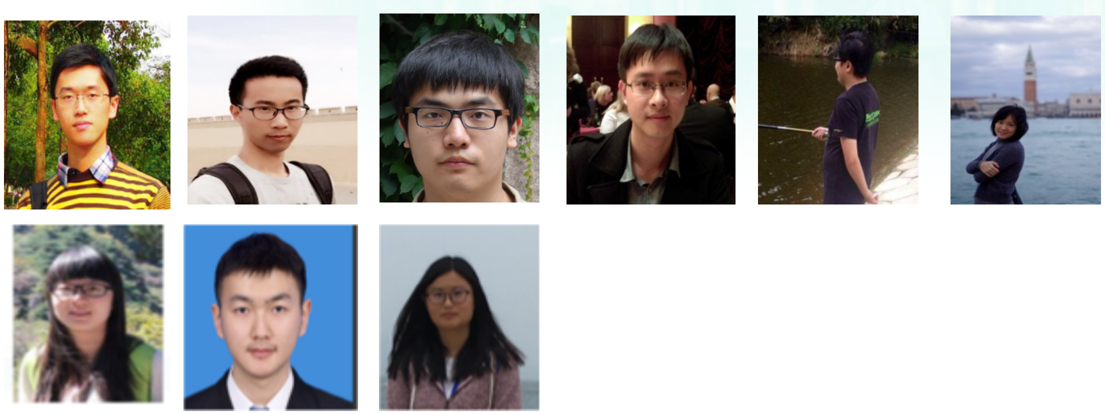
- Asymmetric Metric Learning
- Hash Re-ID
- Attribute-Image Re-ID
- Partial Re-ID
- Unsupervised Re-ID
- RGB-Infrared Re-ID
- Depth Re-ID
- Cross-scenario Re-ID
- Low-resolution Re-ID
- Open-world Re-ID

Person Re-identification



Intelligence Science & System Lab
iSEE
智能科学与系统实验室

Thanks to my students



http://isee.sysu.edu.cn/~zhwshi

EMAIL ME：wszheng@ieee.org