大学共同利用機関法人 情報・システム研究機構
国立情報学研究所
National Institute of Informatics

基于差异矩阵和矩阵度量的行人重识别方法
Person Re-identification via Discrepancy Matrix and Matrix Metric

王正
国立信息学研究所（日本）

大纲

研究背景

研究动机和方法

实验结果和分析

研究背景

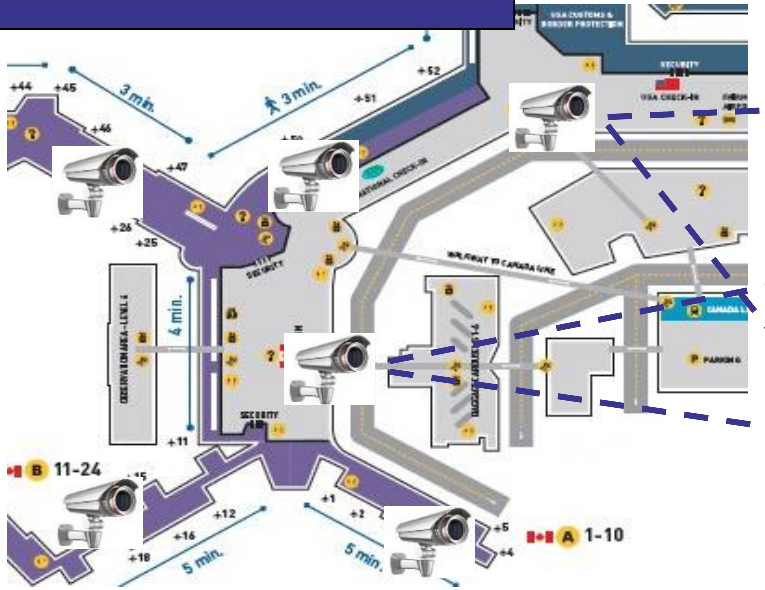
南京周克华案

查找周克华

1500 个警员，1个月时间

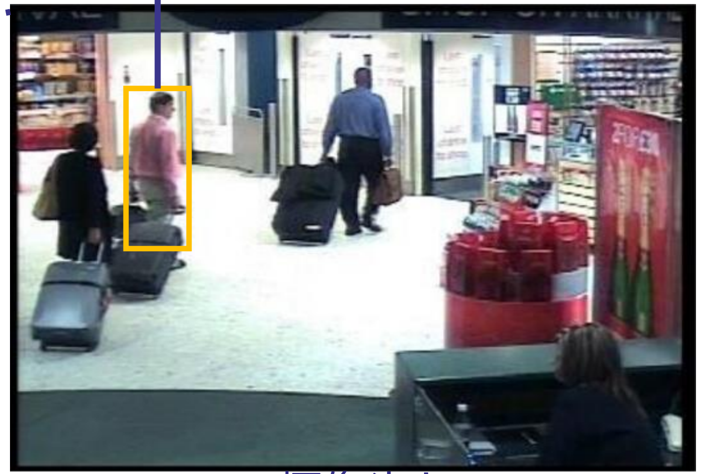
329 段视频片段

Person Re-identification



是同一人吗?

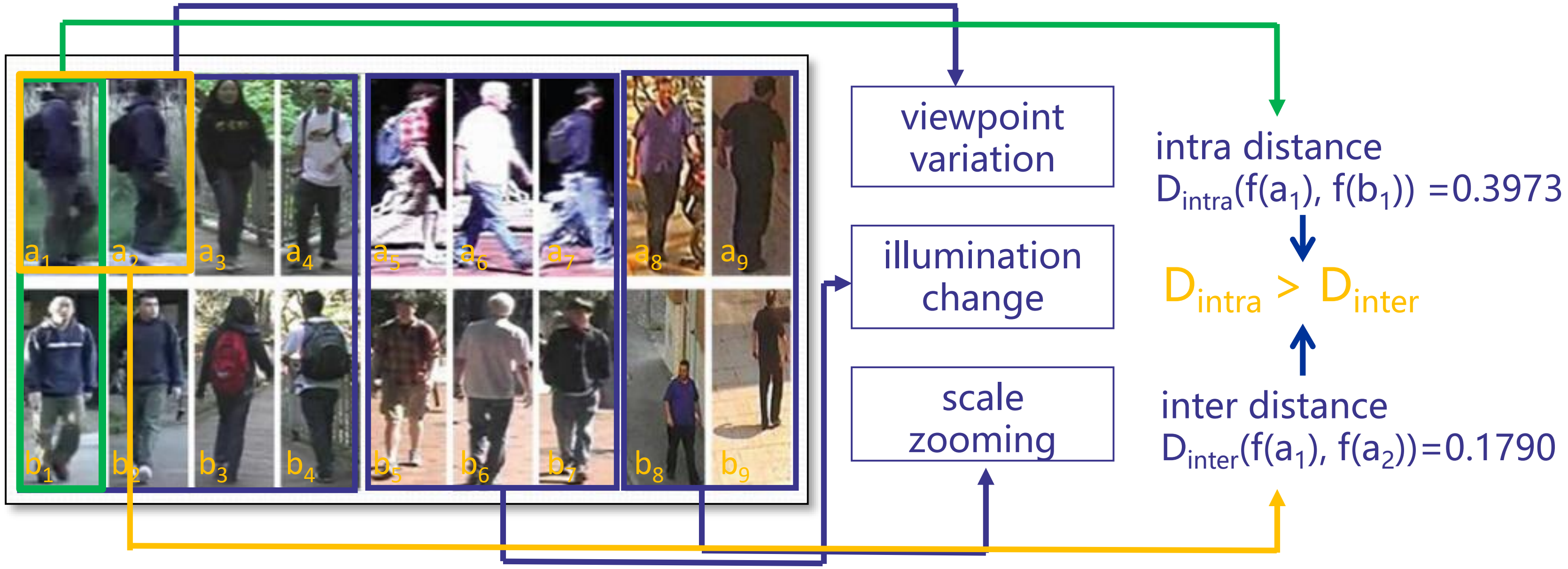
摄像头 a



摄像头 b

研究背景

挑战

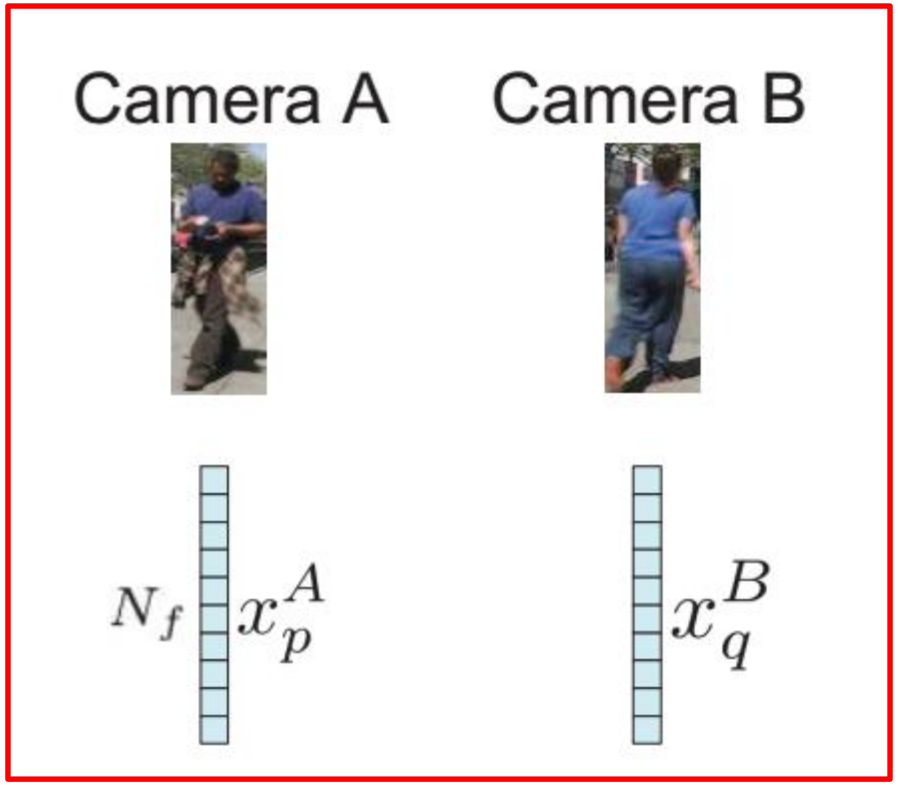


研究方向

行人表征 → **距离度量** → **排序优化**

行人表征

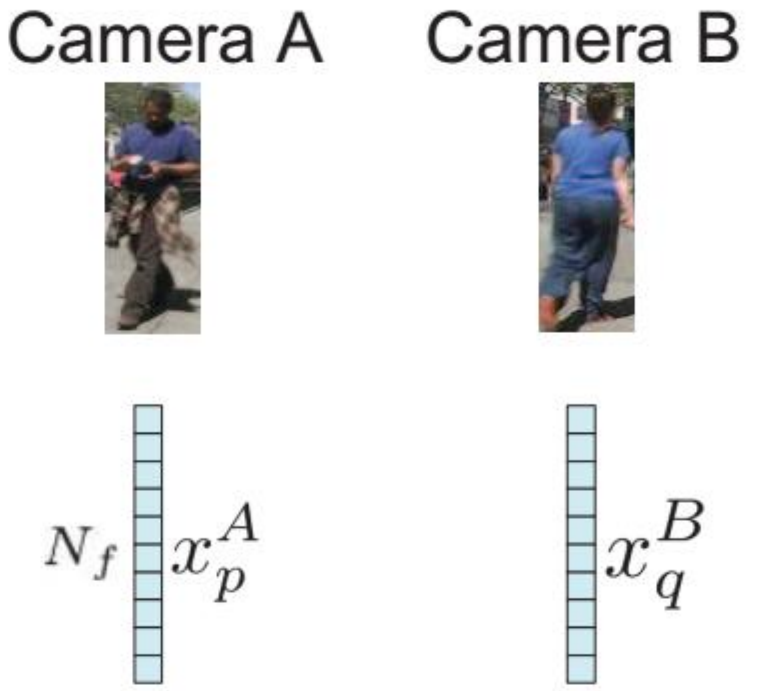
- **Zheng Wang**, Ruimin Hu, et al., Scale-adaptive Low-resolution Person Re-identification via Learning A Discriminating Surface, International Joint Conference on Artificial Intelligence (**IJCAI**), pp.2669-2675, Aug, 2016
- **Zheng Wang**, Ruimin Hu, et al., Multi-Level Fusion for Person Re-identification with Incomplete Marks, ACM international Conference on Multimedia (**ACM MM**), pp.1267-1270, Oct, 2015
- Mang Ye, Chao Liang, **Zheng Wang**, et al., Specific Person Retrieval via Incomplete Text Description, ACM International Conference on Multimedia Retrieval (**ICMR**), pp.547-550, Jun, 2015



相关工作

距离度量

- **Zheng Wang,** Ruimin Hu, Yi Yu, Junjun Jiang, Jiayi Ma, Shin'ichi Satoh, Statistical Inference of Gaussian-Laplace Distribution for Person Verification, ACM international Conference on Multimedia (**ACM MM**), 2017
- **Zheng Wang**, Ruimin Hu, et al. , Zero-Shot Person Re-identification via Cross-View Consistency, **IEEE Transactions on Multimedia**, Vol 18, No 2, pp.260-272, Feb, 2016
- **Zheng Wang**, Ruimin Hu, et al., TAICHI Distance for Person Re-identification, International Conference on Acoustics, Speech and Signal Processing (**ICASSP**), Mar, 2017
- Jin Wang, **Zheng Wang**, et al., DeepList: Learning Deep Features with Adaptive Listwise Constraint for Person Re-identification, **IEEE Transactions on Circuits and Systems for Video Technology**, Vol 27, No 3, pp.513 - 524, Mar, 2017



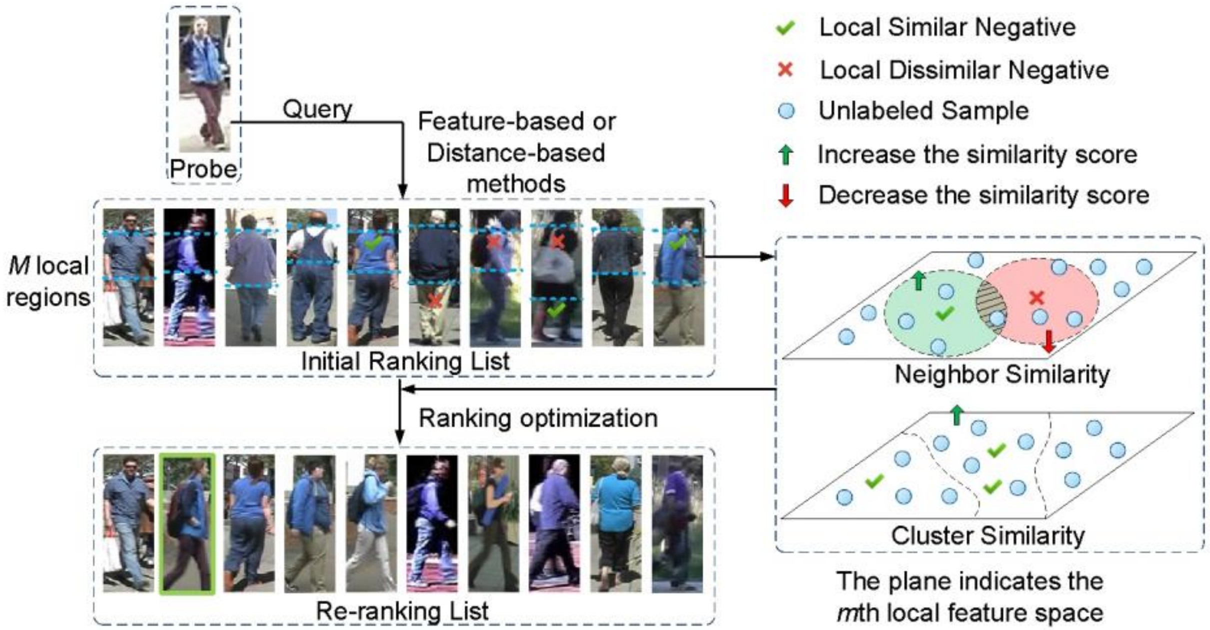
$$d(x_p^A, x_q^B) = (x_p^A - x_q^B)^\top \mathbf{M} (x_p^A - x_q^B)$$

$$\mathbf{M} = \mathbf{L}^\top \mathbf{L}$$

$$d(x_p^A, x_q^B) = \|\mathbf{L}(x_p^A - x_q^B)\|^2$$

排序优化

- **Zheng Wang**, Ruimin Hu, et al., Region-based Interactive Ranking Optimization For Person Re-identification, Pacific-Rim Conference on Multimedia (**PCM**), pp.1-10, Dec, 2014 (最佳论文奖)
- Mang Ye, Chao Liang, **Zheng Wang**, et al., Ranking Optimization for Person Re-identification via Similarity and Dissimilarity, ACM international Conference on Multimedia (**ACM MM**), pp.1239-1242, Oct, 2015
- Mang Ye, Chao Liang, Yi Yu, **Zheng Wang**, et al., Person Re-identification via Ranking Aggregation of Similarity Pulling and Dissimilarity Pushing, **IEEE Transactions on Multimedia**, Vol 18, No 12, pp.2553-2566, Dec, 2016



Region-based Interactive Re-ranking

大纲

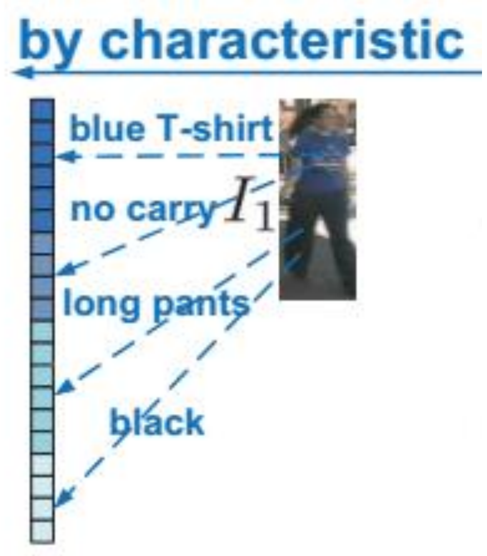
研究背景

研究动机和方法

实验结果和分析

研究动机

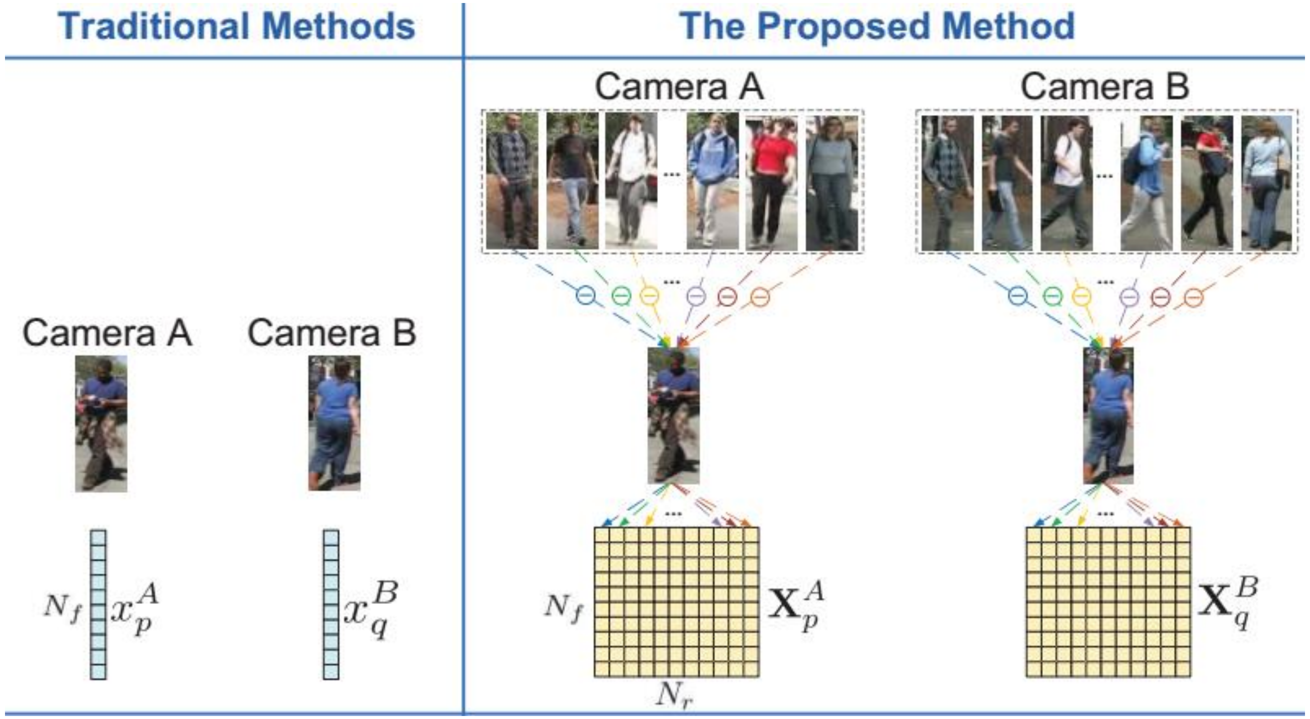
从特性表示到差异表示



猜想：
（1）抵消了摄像头变化?
（2）能突出细节?

研究动机

差异表示更好



$$f_i^A - f_i^B = v + \sigma_i$$

$$\mathbf{X}_p^A \doteq [x_p^A - f_1^A; x_p^A - f_2^A; \ldots; x_p^A - f_{N_r}^A]$$

$$\mathbf{X}_q^B \doteq [x_q^B - f_1^B; x_q^A - f_2^B; \ldots; x_q^B - f_{N_r}^B]$$

跨摄像头的成像变化　　行人个体差异

$$f_p^A - f_q^B \doteq v + \sigma$$

$$\mathbf{X}_p^A - \mathbf{X}_q^B \doteq [\sigma - \sigma_1; \sigma - \sigma_2; \ldots; \sigma - \sigma_{N_r}]$$

Method (rank@)	1	2	3	4	5	6	7	8	9	10
hand-crafted feature vector	21.3	29.7	34.6	39.4	43.9	47.5	49.9	53.3	55.6	57.6
hand-crafted discrepancy matrix	**22.9**	**33.4**	**39.5**	**43.7**	**46.5**	**49**	**51.9**	**54.7**	**56.9**	**59.7**
deep-learned feature vector	29.3	37.9	44.5	49.5	53.5	57.6	61.0	62.9	65.6	67.7
deep-learned discrepancy matrix	**31.8**	**43.1**	**49.8**	**54.2**	**57.9**	**61.1**	**63.2**	**65.3**	**68.3**	**69.7**

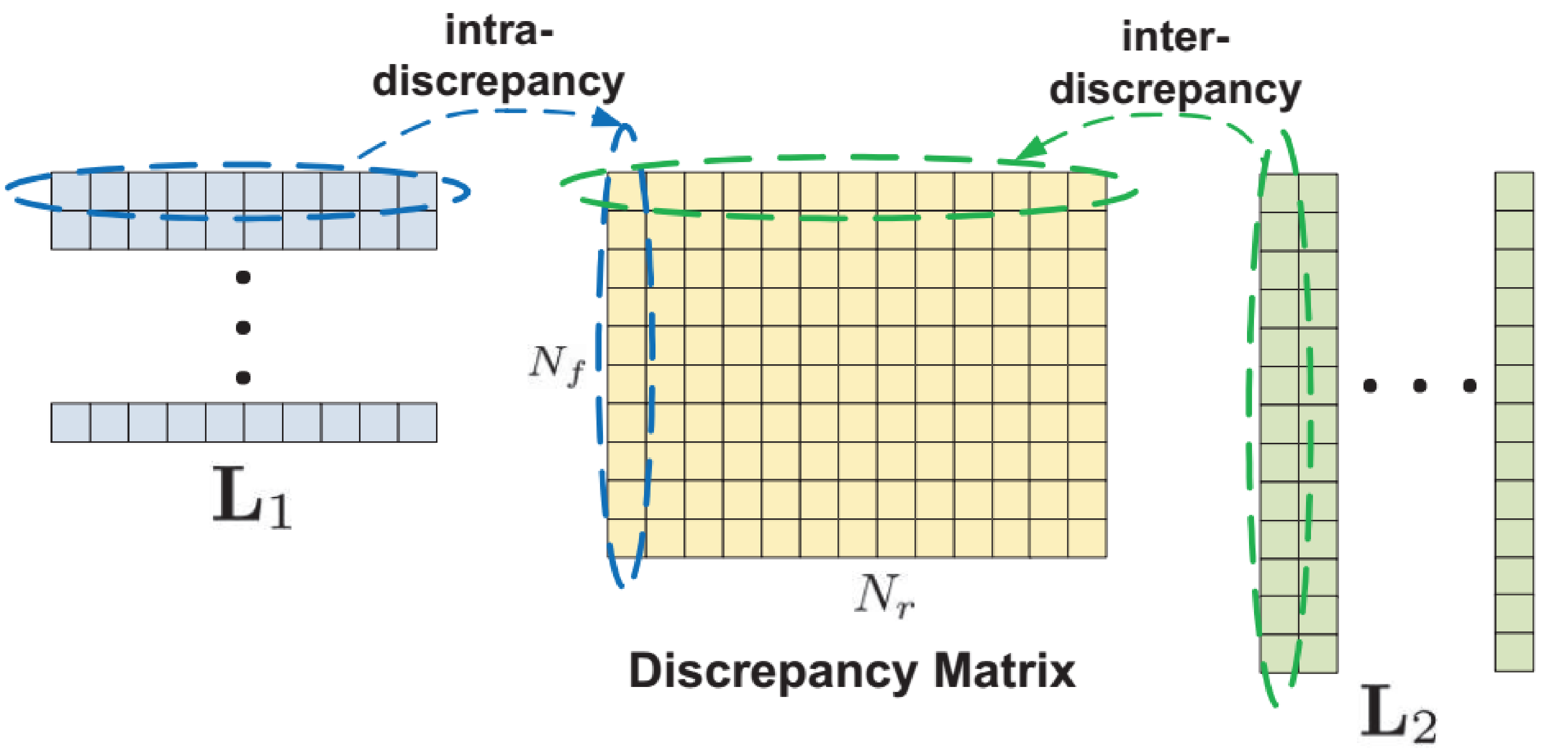
研究方法

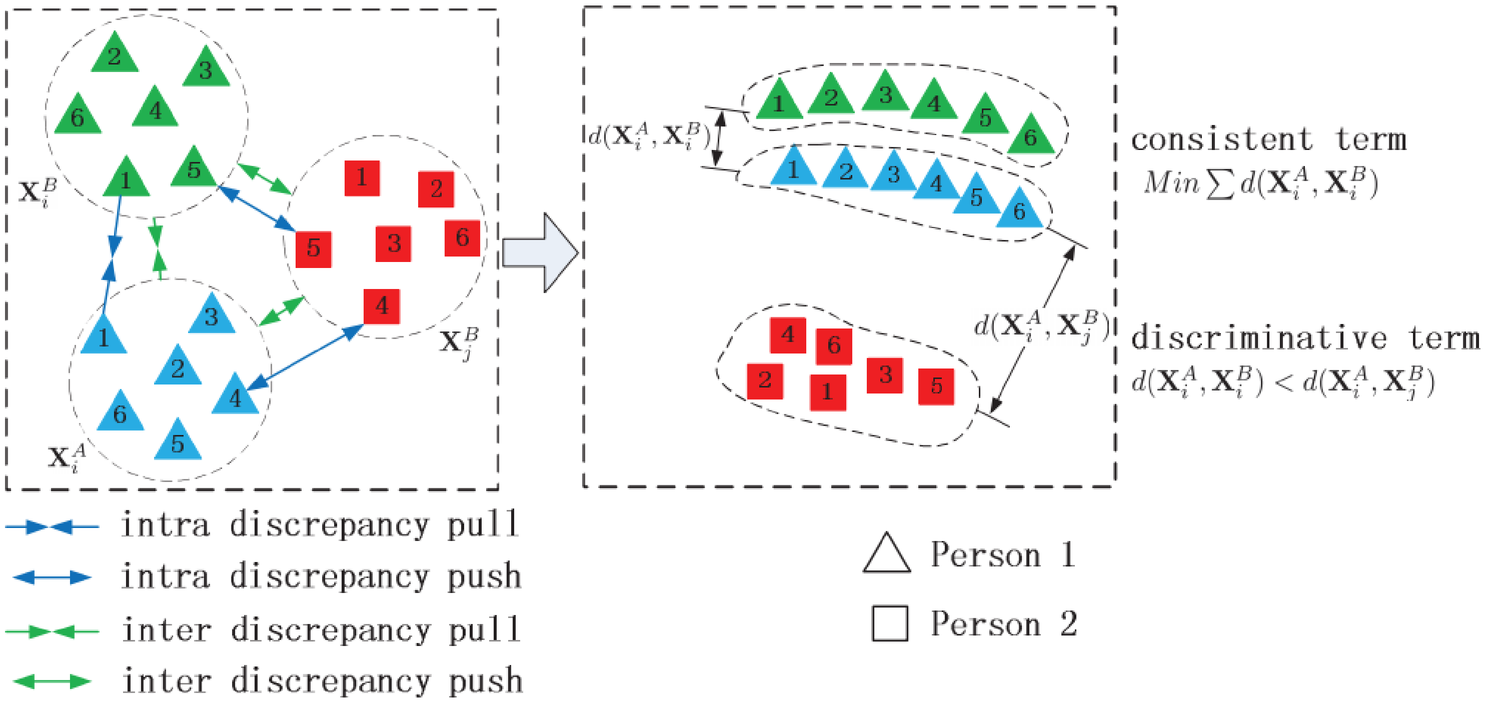
从一个投影到两个投影

$$d(x_p^A, x_q^B) = \|\mathbf{L}(x_p^A - x_q^B)\|^2 \Rightarrow d(\mathbf{X}_p^A, \mathbf{X}_q^B) = \|\mathbf{L}_1(\mathbf{X}_p^A - \mathbf{X}_q^B)\mathbf{L}_2\|_F^2$$

差异内投影矩阵

差异间投影矩阵





$$E_{con}(\mathbf{L}_1, \mathbf{L}_2) = \frac{1}{M} \sum_{i=1}^{M} d(\mathbf{X}_i^A, \mathbf{X}_i^B).$$

$$E_{dis}(\mathbf{L}_1, \mathbf{L}_2) = \frac{1}{S} \sum_{k=1}^{S} l_\beta(e(s_k))$$

$$e(s_k) = d(\mathbf{X}_i^A, \mathbf{X}_i^B) - d(\mathbf{X}_i^A, \mathbf{X}_j^B)$$

$$l_\beta(z) = \frac{1}{\beta} \log(1 + e^{\beta z})$$

$$E_{spr}(\mathbf{L}_2) = \|\mathbf{L}_2\|_{2,1}$$

$$E(\mathbf{L}_1, \mathbf{L}_2) = E_{con}(\mathbf{L}_1, \mathbf{L}_2) + E_{dis}(\mathbf{L}_1, \mathbf{L}_2) + \mu E_{spr}(\mathbf{L}_2)$$

一致项　判别项　L_2稀疏项

研究方法

迭代交替优化

$$\frac{\partial E(\mathbf{L}_1,\mathbf{L}_2)}{\partial \mathbf{L}_1}=\frac{\partial E_{con}(\mathbf{L}_1,\mathbf{L}_2)}{\partial \mathbf{L}_1}+\frac{\partial E_{dis}(\mathbf{L}_1,\mathbf{L}_2)}{\partial \mathbf{L}_1}$$

$$\frac{\partial E(\mathbf{L}_1,\mathbf{L}_2)}{\partial \mathbf{L}_2}=\frac{\partial E_{con}(\mathbf{L}_1,\mathbf{L}_2)}{\partial \mathbf{L}_2}+\frac{\partial E_{dis}(\mathbf{L}_1,\mathbf{L}_2)}{\partial \mathbf{L}_2}+\mu\frac{\partial E_{spr}(\mathbf{L}_2)}{\partial \mathbf{L}_2}$$

$$\frac{\partial E_{con}(\mathbf{L}_1,\mathbf{L}_2)}{\partial \mathbf{L}_1}=\frac{2}{M}\sum_{i=1}^{M}\mathbf{L}_1\mathbf{Z}_i\mathbf{L}_2\mathbf{L}_2^\top\mathbf{Z}_i^\top$$

$$\frac{\partial E_{con}(\mathbf{L}_1,\mathbf{L}_2)}{\partial \mathbf{L}_2}=\frac{2}{M}\sum_{i=1}^{M}\mathbf{Z}_i^\top\mathbf{L}_1^\top\mathbf{L}_1\mathbf{Z}_i\mathbf{L}_2$$

$$\frac{\partial E_{dis}(\mathbf{L}_1,\mathbf{L}_2)}{\partial \mathbf{L}_1}=\frac{2}{S}\sum_{k=1}^{S}g(e(s_k))(\mathbf{L}_1\mathbf{U}_k\mathbf{L}_2\mathbf{L}_2^\top\mathbf{U}_k^\top-\mathbf{L}_1\mathbf{V}_k\mathbf{L}_2\mathbf{L}_2^\top\mathbf{V}_k^\top)$$

$$\frac{\partial E_{dis}(\mathbf{L}_1,\mathbf{L}_2)}{\partial \mathbf{L}_2}=\frac{2}{S}\sum_{k=1}^{S}g(e(s_k))(\mathbf{U}_k^\top\mathbf{L}_1^\top\mathbf{L}_1\mathbf{U}_k\mathbf{L}_2-\mathbf{V}_k^\top\mathbf{L}_1^\top\mathbf{L}_1\mathbf{V}_k\mathbf{L}_2)$$

$$\frac{\partial E_{spr}(\mathbf{L}_2)}{\partial \mathbf{L}_2}=2\mathbf{D}\mathbf{L}_2$$

Algorithm 1 Learning the matrix metric $\mathbf{L}_1$ and $\mathbf{L}_2$

Input: The training data: Positive samples with pair form $\{(\mathbf{X}_i^A,\mathbf{X}_i^B)\}$, and Negative Samples with triple form $\{(\mathbf{X}_i^A,\mathbf{X}_i^B,\mathbf{X}_j^B)_k\}$.
Output: The optimal matrix $\mathbf{L}_1^*$ and $\mathbf{L}_2^*$.

```
Initialize L_1 and L_2;
for n = 1 to MaxIter do
    Fix L_2^n;
    Compute ∇E(L_1) = ∂E(L_1,L_2)/∂L_1 as Eq. 8, Eq. 10, and Eq. 12;
    Choose a proper step λ_1 as [61];
    Compute L_1^{n+1} = L_1^n − λ_1 ∇E(L_1);
    Fix L_1^{n+1};
    Compute ∇E(L_2) = ∂E(L_1,L_2)/∂L_2 as Eq. 9, Eq. 11, and Eq. 13;
    Choose a proper step λ_2 as [61];
    Compute L_2^{n+1} = L_2^n − λ_2 ∇E(L_2);
    if converge then
        break;
    end if
end for
```

F. Nie, H. Huang, X. Cai, and C. H. Ding, “Efficient and robust feature selection via joint l2, 1-norms minimization,” in Adv. Neural Inform. Process. Syst., 2010, pp. 1813–1821.

大纲

- 研究背景
- 研究动机和方法
- **实验结果和分析**

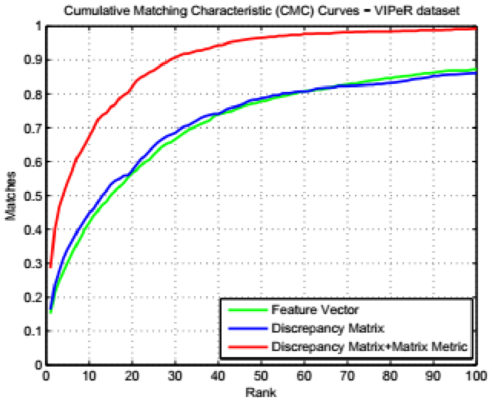
实验和分析

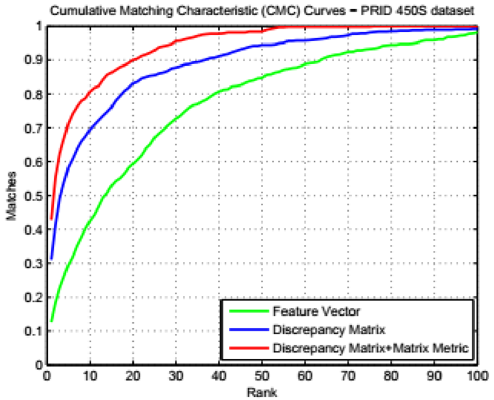
(a) VIPeR dataset

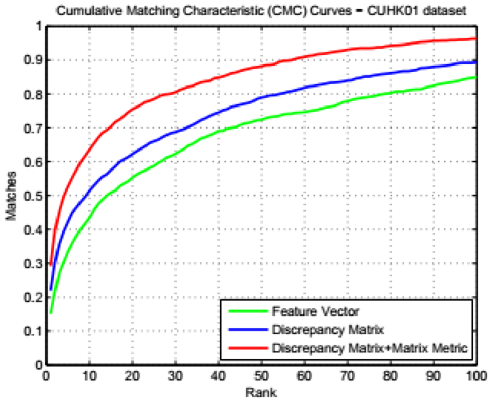
(b) PRID 450S dataset

(c) CUHK01 dataset

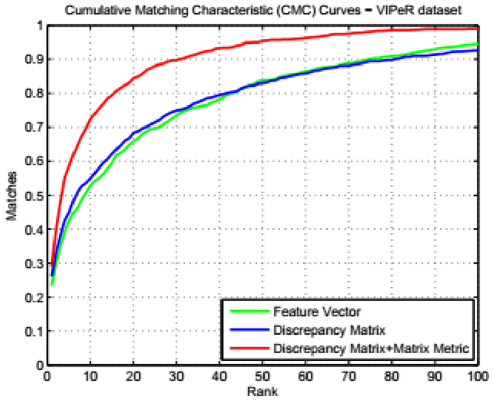
GoG

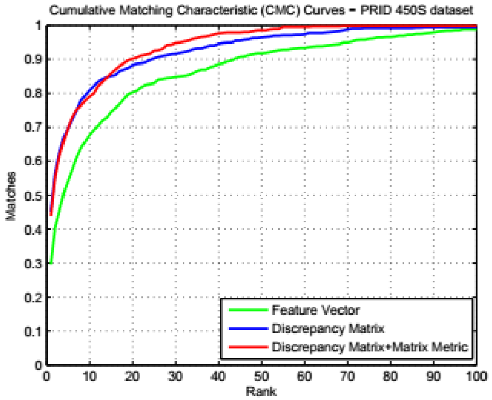


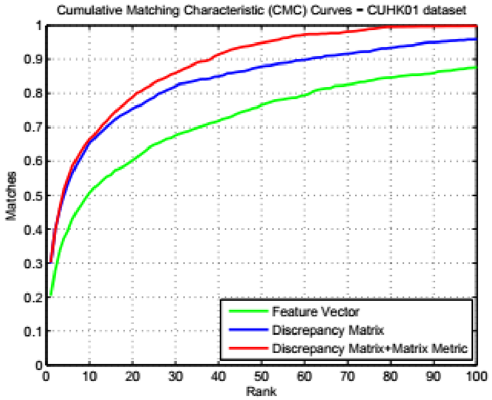




FTCNN



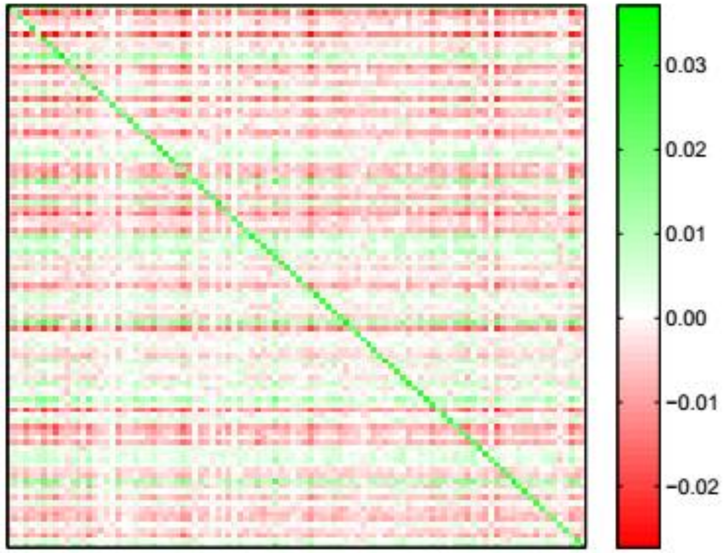




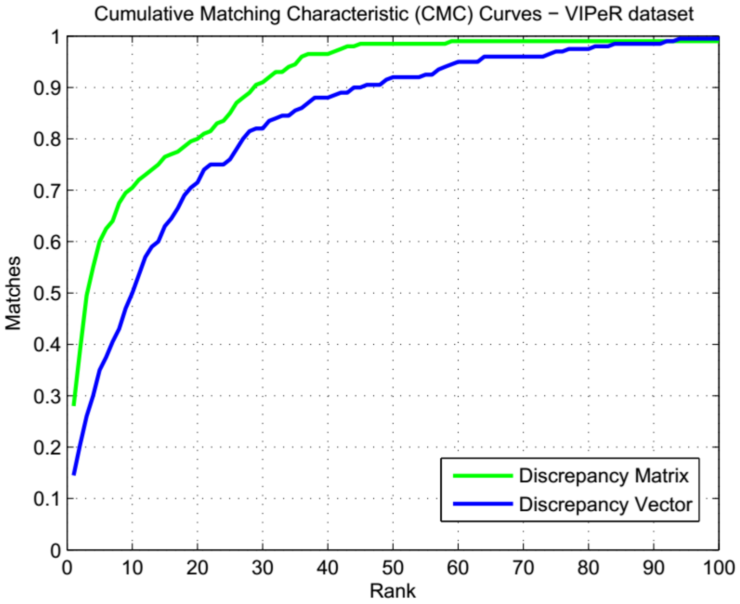
（1）差异矩阵
（2）矩阵度量投影
都有用

实验和分析

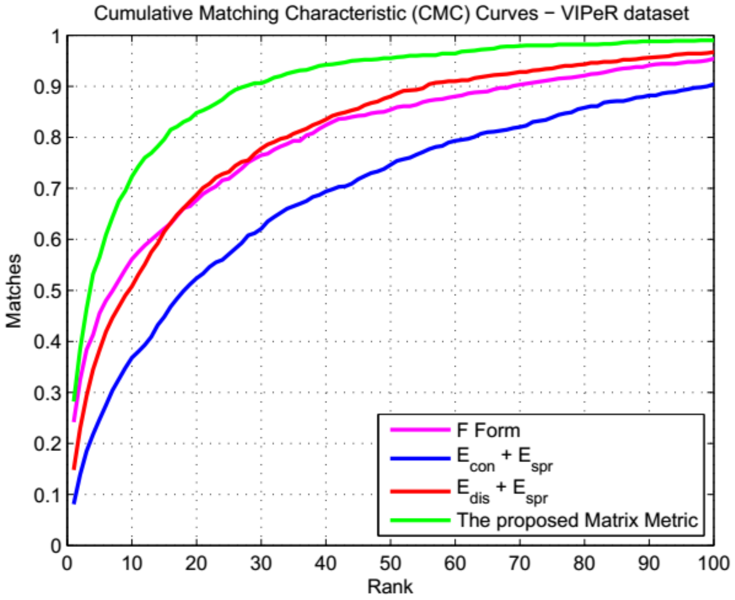
差异间投影的稀疏性



向量 V.S. 矩阵



判别项和一致项



与前沿方法的比较

ViPER

Method (rank@)	1	5	10	20
ELF [22]	12.0	-	43.0	60.0
BiCov [39]	20.6	43.2	56.1	68.0
SDALF [23]	19.9	38.4	49.4	66.0
eSDC [42]	26.3	46.4	58.6	72.8
MidFilter [64]	29.1	52.5	65.9	79.9
SCNCD [24]	37.8	68.5	81.2	90.4
RD [37]	33.3	65.1	78.3	88.5
PRDC [19]	15.7	38.4	53.9	70.1
KISSME [33]	19.6	48.0	62.2	77.0
PCCA [49]	19.3	48.9	64.9	80.3
LADF [34]	30.0	64.0	80.0	92.0
LOMO+XQDA [25]	40.0	68.5	80.5	91.0
DeepMetric [65]	28.2	59.3	73.4	86.4
DeepRanking [66]	38.4	69.2	81.3	90.4
DeepFeature+RDC [28]	40.5	60.8	70.4	84.4
DeepList [30]	40.5	69.1	80.1	91.2
LOMO+NFST [51]	42.2	71.4	82.9	92.0
(1) GoG [26]+XQDA	37.3	67.4	77.2	89.6
(2) FTCNN [56]+XQDA	31.2	59.8	74.0	83.5
(3) FTCNN+**DM**^{3}	37.3(↑6.1)	67.4(↑7.6)	80.3(↑6.3)	89.5(↑6.0)
Combine (1) and (2)	38.3	67.2	77.0	89.3
Combine (1) and (3)	**42.7**	**74.3**	**85.1**	**93.1**

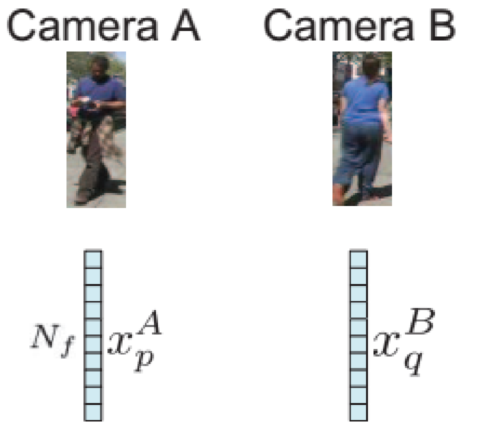
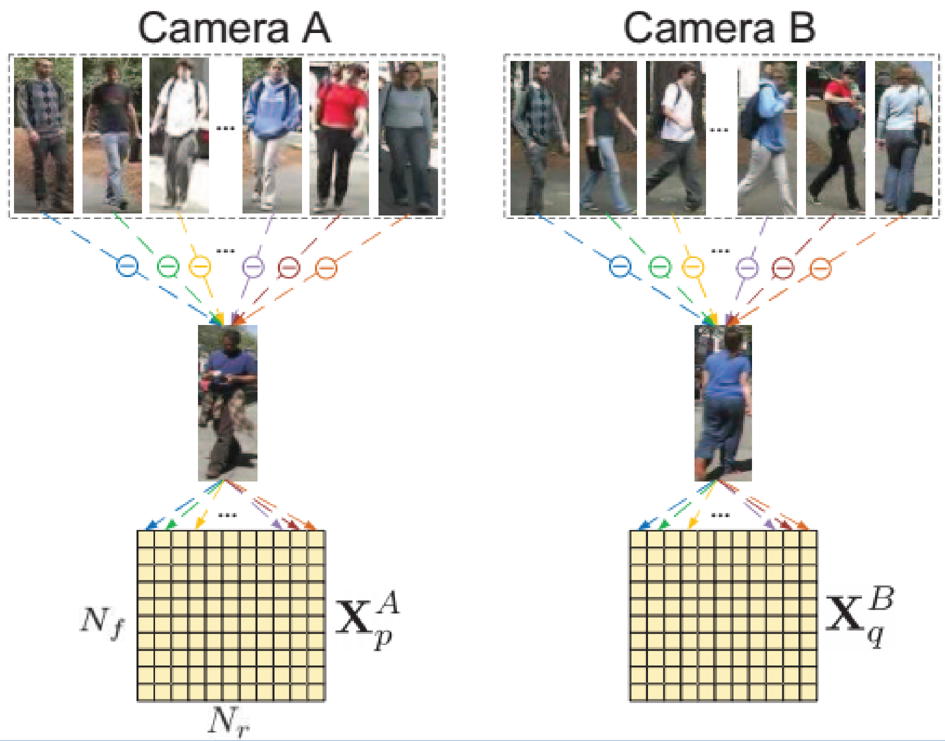
PRID 450S

Method (rank@)	1	5	10	20
SCNCD [24]	41.6	68.9	79.4	87.8
KISSME [33]	33.0	59.8	71.0	79.0
CBRA [67]	26.4	57.1	71.0	83.2
CSL [68]	44.4	71.6	82.2	89.8
Mirror [69]	55.4	79.3	87.8	93.9
DRML [70]	56.4	-	82.2	90.2
(1) GoG [26]+XQDA	51.6	76.8	88.8	94.2
(2) FTCNN [56]+XQDA	50.2	74.2	84.8	93.7
(3) FTCNN+**DM**^{3}	56.7(↑6.5)	83.1(↑8.9)	88.4(↑3.6)	94.7(↑1.0)
Combine (1) and (2)	51.8	76.9	87.0	94.2
Combine (1) and (3)	**61.0**	**85.8**	**92.0**	**96.7**

CUHK01

Method (rank@)	1	5	10	20
SDALF [23]	9.9	22.6	30.3	41.0
TML [63]	20.0	43.5	56.0	69.3
SalMatch [71]	28.4	45.8	55.7	67.9
MidFilter [64]	34.3	55.1	65.0	74.9
RD [37]	31.1	-	68.5	79.1
ImprovedDeep [72]	47.5	71.0	80.0	-
(1) GoG [26]+XQDA	44.5	71.1	78.1	89.0
(2) FTCNN [56]+XQDA	41.1	63.5	73.6	85.8
(3) FTCNN+**DM**^{3}	43.7(↑2.6)	70.1(↑6.6)	77.4(↑3.8)	88.7(↑2.9)
Combine (1) and (2)	42.1	70.1	78.3	89.6
Combine (1) and (3)	**49.7**	**77.3**	**86.1**	**91.4**

总结

	Traditional Methods	The Proposed Method
Feature description		
Distance metric	$d(x_p^A, x_q^B) = \lVert \mathbf{L}(x_p^A - x_q^B) \rVert^2$	$d(\mathbf{X}_p^A, \mathbf{X}_q^B) = \lVert \mathbf{L}_1(\mathbf{X}_p^A - \mathbf{X}_q^B)\mathbf{L}_2 \rVert_F^2$

从特性向量到差异矩阵

从向量度量到矩阵度量

谢谢！