Domain Transfer Support Vector Ranking for Person Re-Identification without Target Camera Label Information

Andy J Ma, Pong C Yuen and Jiawei Li

Department of Computer Science, Hong Kong Baptist University

香港浸會大學 HONG KONG BAPTIST UNIVERSITY

HKBU Computer Science Department 計算機科學系

Introduction

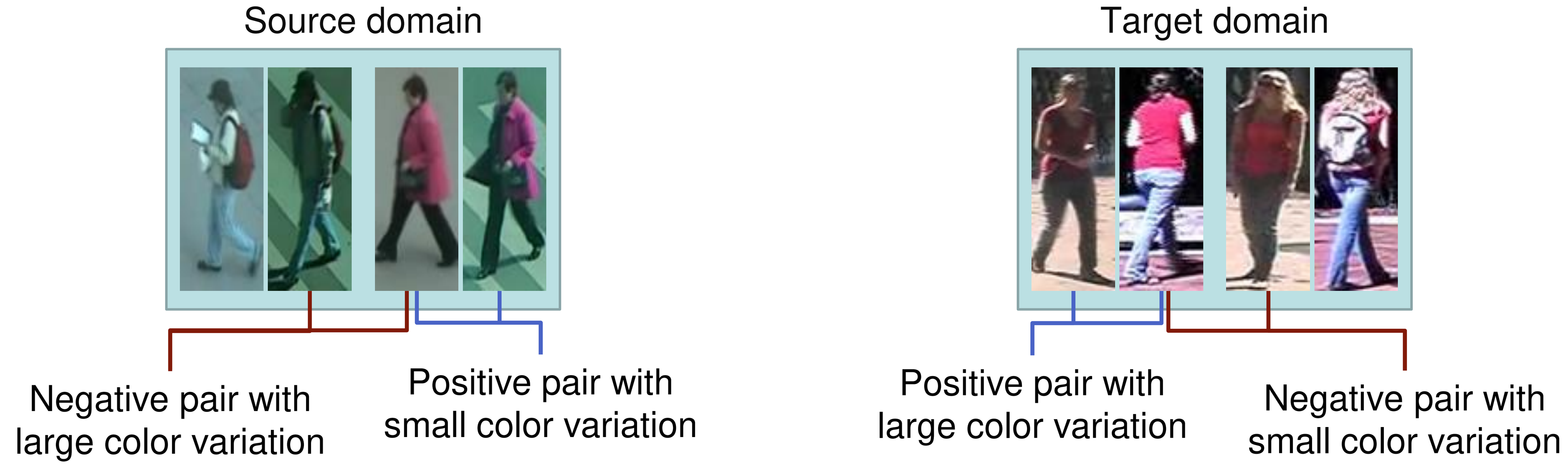
Motivation

- In large-scale camera network, label information may not be available for all cameras.
- To solve this problem, distance model learnt from source cameras could be applied to target cameras.
- However, significant inter-camera variations lead to dramatic performance deterioration.

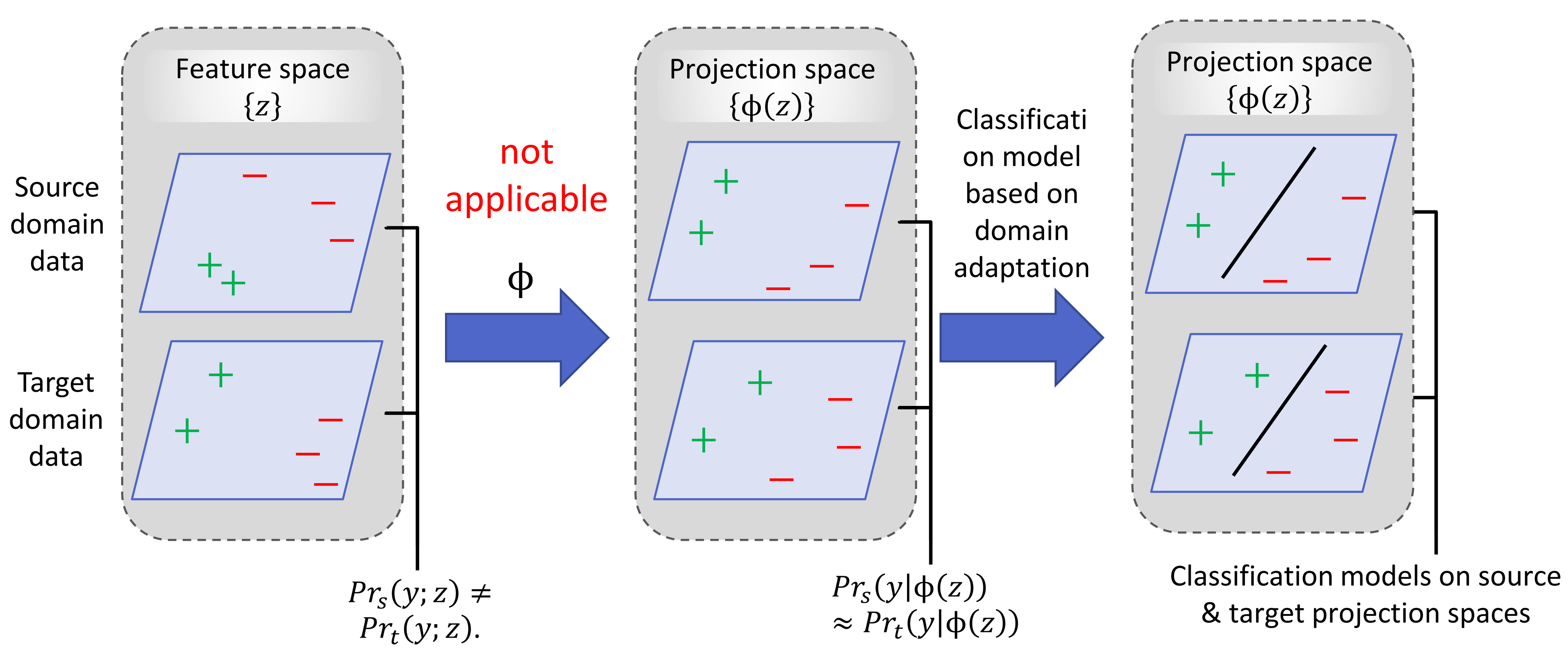
	Source domain	Target domain
View point change	45°	>90°
Lighting change	Small	Large
background	Fixed	Unfixed

Problem definition

- Due to non-trivial inter-camera variations, $Pr_s(y;z) \neq Pr_t(y;z)$.

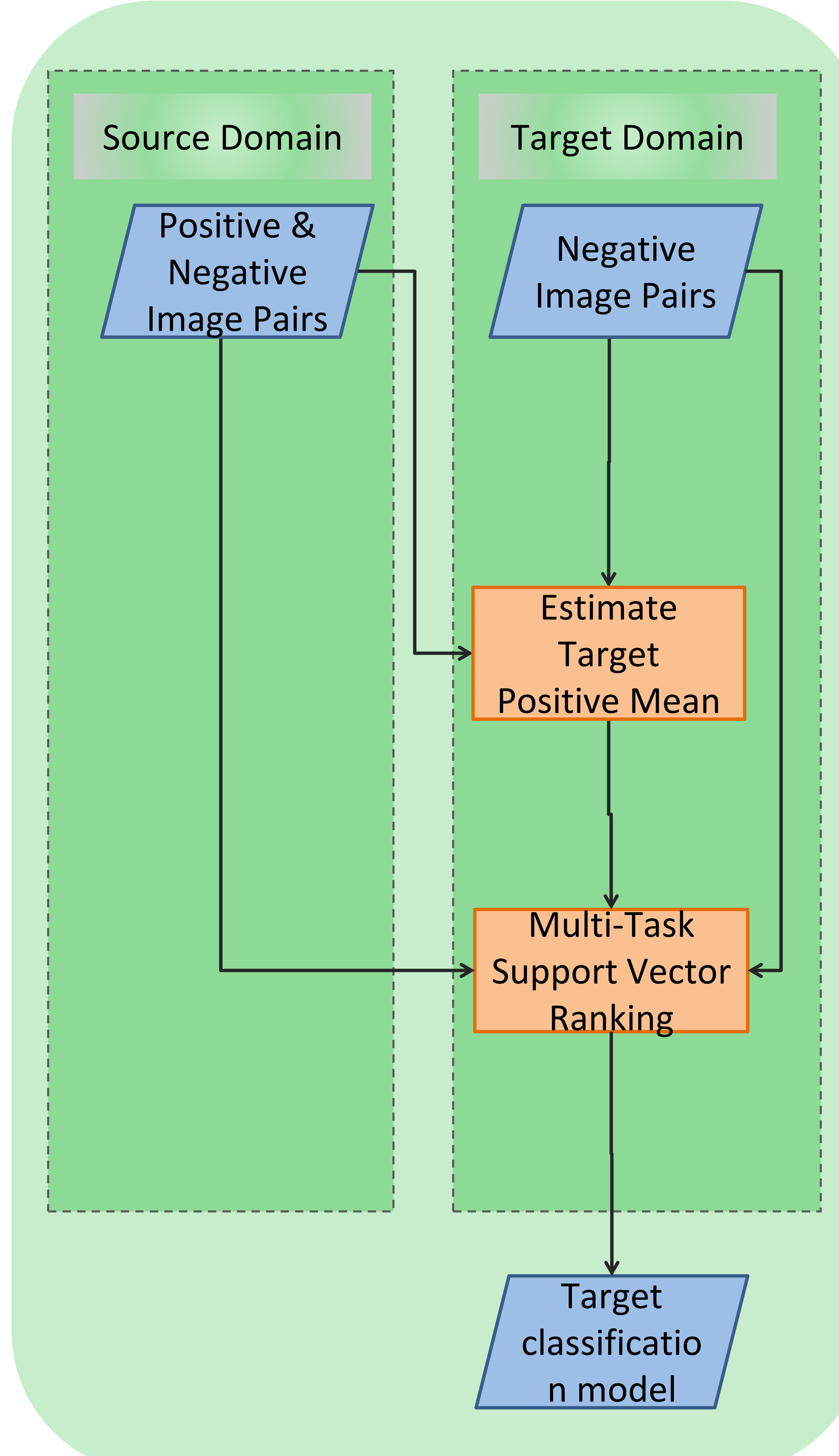


- A mapping ϕ, s.t. $Pr_s(\phi(z)) \approx Pr_t(\phi(z))$ can be learnt via unsupervised domain adaptation.
- If $Pr_s(y|\phi(z)) \approx Pr_t(y|\phi(z))$, a classification model can be learnt on projection space $\phi(z)$.



- Without label information in target domain, ϕ can not be easily estimated. So we tend to find out label information on target domain directly.
- However, above label information may not be accurate enough, multi-task learning is employed to learning a discriminative classification model on target domain.

Main Idea



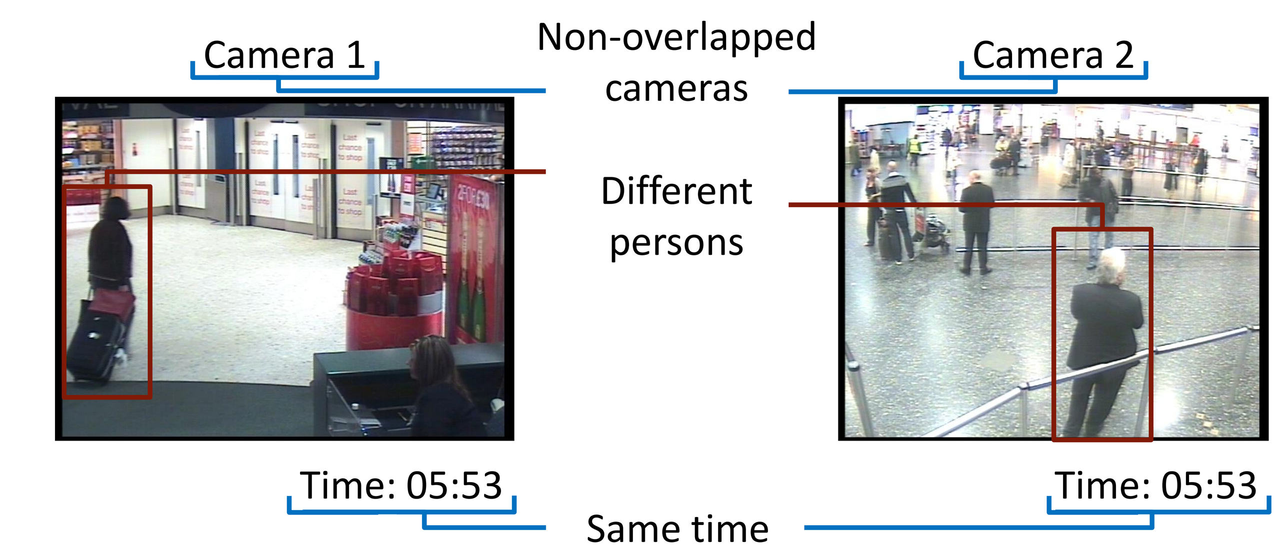
Contributions

- We propose to estimate target positive information using easily generated negative data in target domain and labeled data in source domain.
- A novel multi-task support vector ranking method is proposed to train an adaptive classification model for person re-identification.

Domain Transfer Support Vector Ranking

Generation of negative image pairs

- Same person cannot be presented at the same instant under different non-overlapping cameras.



Positive mean estimation

$$\tilde{m}_{t+} = \tilde{m}_{t-} + m_{s+} - m_{s-} \quad (1)$$

Upper bound of proposed estimation error is

$$\|m_{t+} - \tilde{m}_{t+}\|$$
$$\leq \|m_{t-} - \tilde{m}_{t-}\| + \|(m_{t+} - m_{t-}) - (m_{s+} - m_{s-})\|$$

We need an assumption that

$$m_{t+} - m_{t-} = m_{s+} - m_{s-} \quad (2)$$

Note that estimation

$$\tilde{m}_{t+} = (N_t m_t - N_{t-}\tilde{m}_{t-})/N_t m_t \quad (3)$$

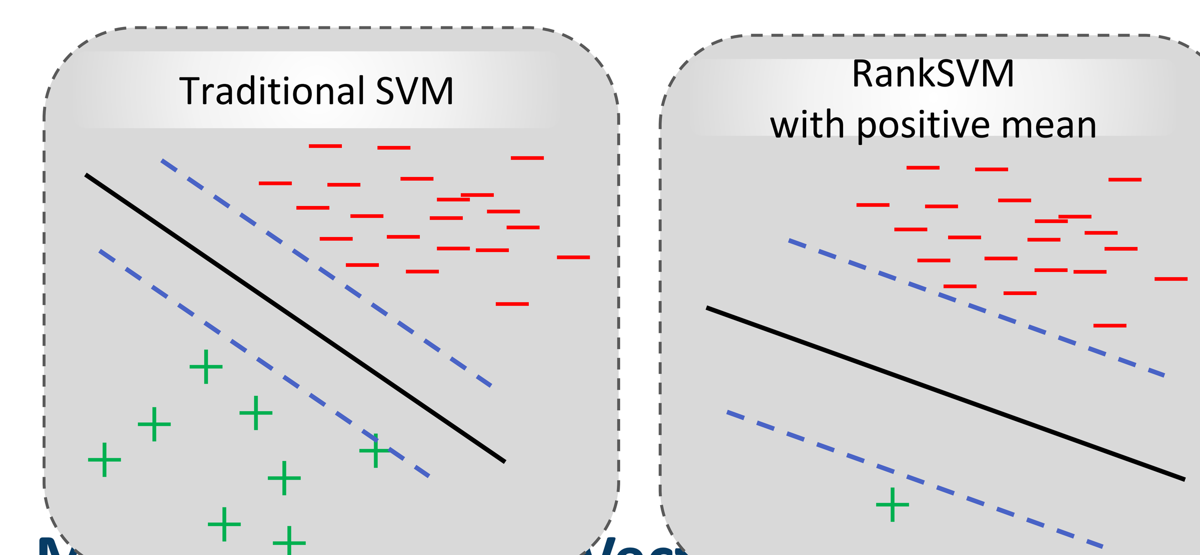
is not applicable for

a) the difficulty to estimate N_t and N_s

b) Unreliability when $N_{t-} \gg N_{t+}$

RankSVM model with positive mean

$$\min_{w_t} \frac{1}{2}\|w_t\|^2 + C\sum_{j_t,i} \xi_{j_t,i}$$
$$s.t.\ w_t^T(\tilde{m}_{t+} - z_{j_t,i}^{t-}) \geq 1 - \xi_{j_t,i},$$



Multi-Task Support Vect...

By combining the source task (traditional RankSVM) and target task (RankSVM with positive mean), we obtain a multi-task SVM problem. Parameter μ is introduced to measure the similarity between the two domains.

Experiments

Experimental setting

- Three datasets: VIPeR, PRID and i-LIDS
- VIPeR and PRID are used as target domain.
- Feature Extraction: Color histogram, Schmid and Gabor filters

Estimation errors (L1) of the positive mean

	Estimated by (3)	Estimated by (1)
i-LIDS to PRID	71.11	2.95
VIPeR to PRID	71.11	2.22
i-LIDS to VIPeR	1120.24	2.46
PRID to VIPeR	1120.24	2.18

- The estimation error estimated by (3) does not change with different source domains, since it is only based on the information in the target domain.
- Estimation with (1) has much smaller estimation errors. This convinces that the assumption given by equation (2) is reasonable for person re-identification.

Comparing with state-of-the-art re-identification methods

- The CMC curves of the learning based and non-learning based methods are shown.
- We also plot the CMC curves of the learning based methods training with the label information in the target domain as the baseline of the upper bound performance. We can see that all the learning based methods have a dramatic deterioration of performance, when the classification model is trained with the data in the source domain.
- when the source domain is i-LIDS and the target domain is PRID, the proposed DTRSVM achieves convincing performance closed to that of the upper bound using the label information in the target domain for training.

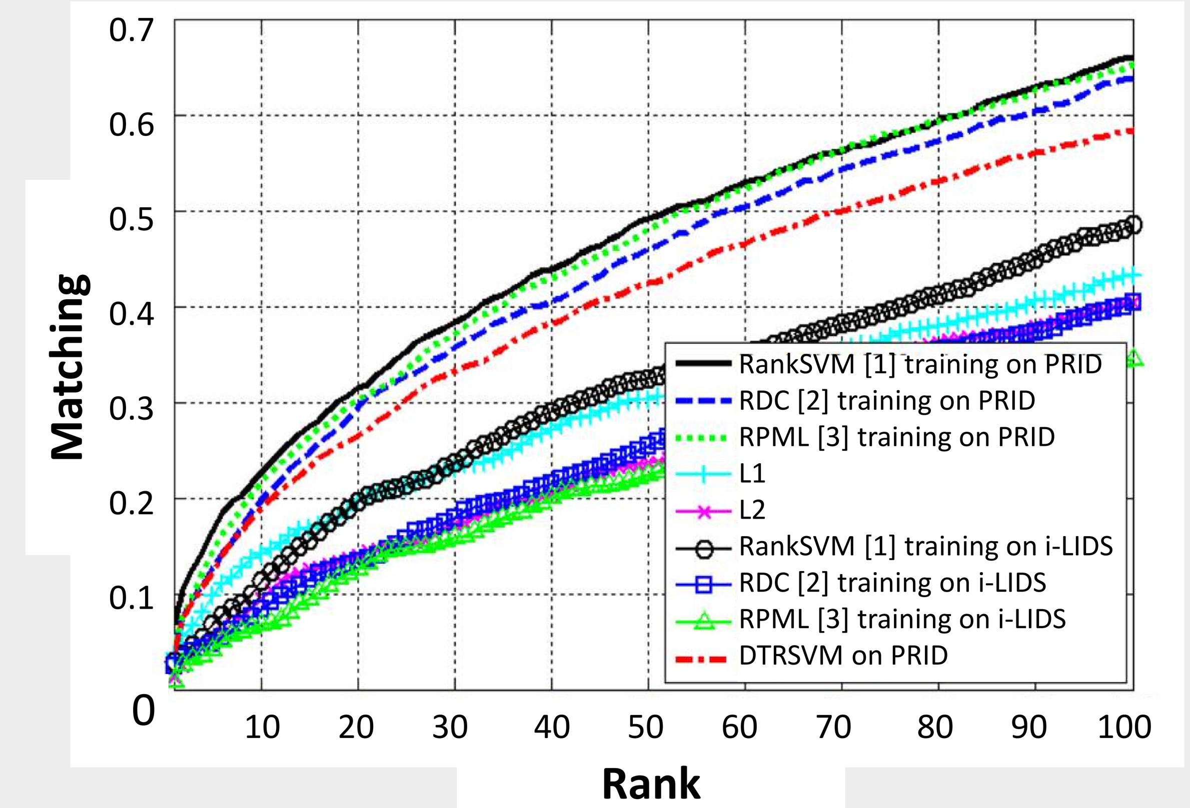
Top r ranked matching accuracy (%) on PRID dataset

Method	Source	r=1	r=10	r=20	r=30
DTRSVM	i-LIDS	3.95	**18.85**	**26.60**	**33.20**
	VIPeR	**4.60**	17.25	22.90	28.10
RankSVM [1]	i-LIDS	2.95	11.40	19.65	23.80
	VIPeR	1.05	9.70	16.20	23.35
RDC [2]	i-LIDS	2.35	8.35	13.40	18.00
	VIPeR	1.95	8.05	12.90	17.05
RPML [3]	i-LIDS	0.90	6.80	12.65	15.85
	VIPeR	1.10	11.85	17.40	21.00
L1	—	3.65	14.25	17.90	23.15
L2	—	1.35	9.55	14.00	17.25

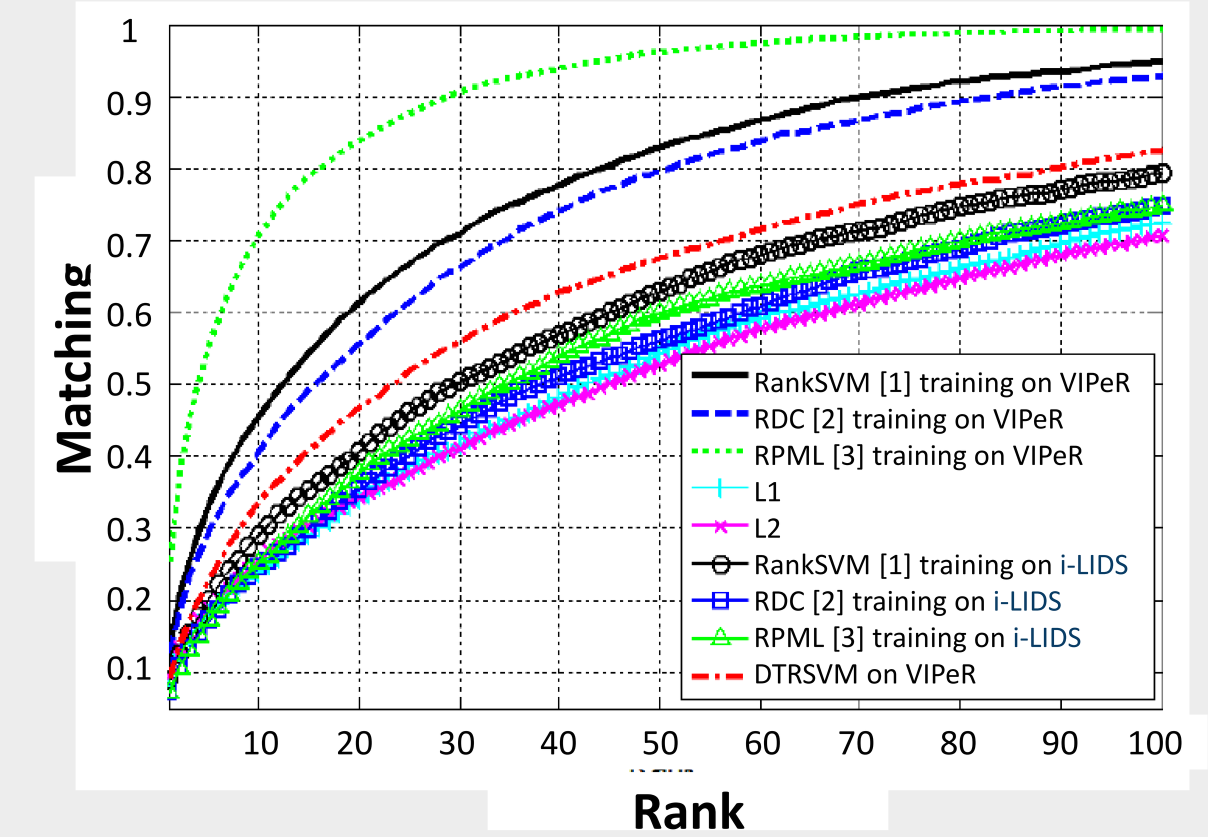
Performance influence of parameter μ:

- the regularization parameter μ measures the degree of relevance of the source and target domains.
- The best performance is achieved by different values of μ under different transfer learning scenarios. This implies that the degree of relevance differs with different source or target domain.

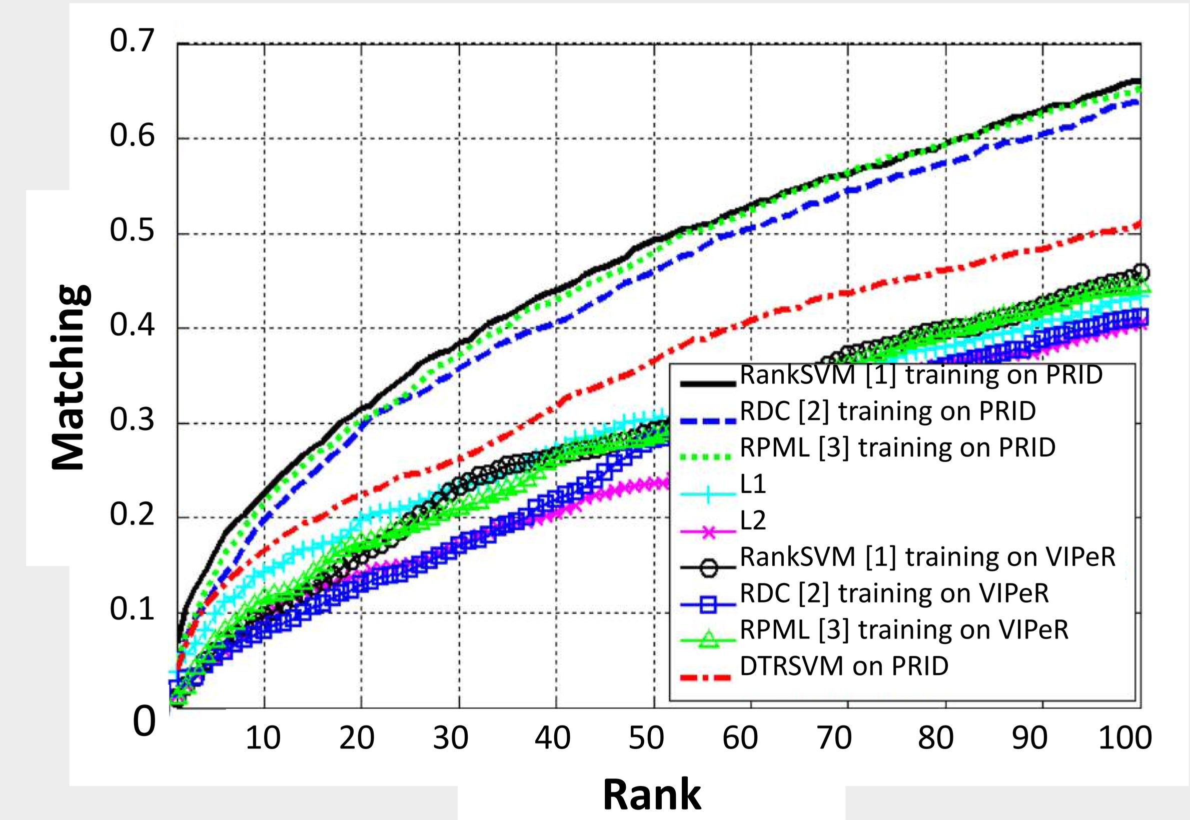
- CMC curves of source: i-LIDS target: PRID



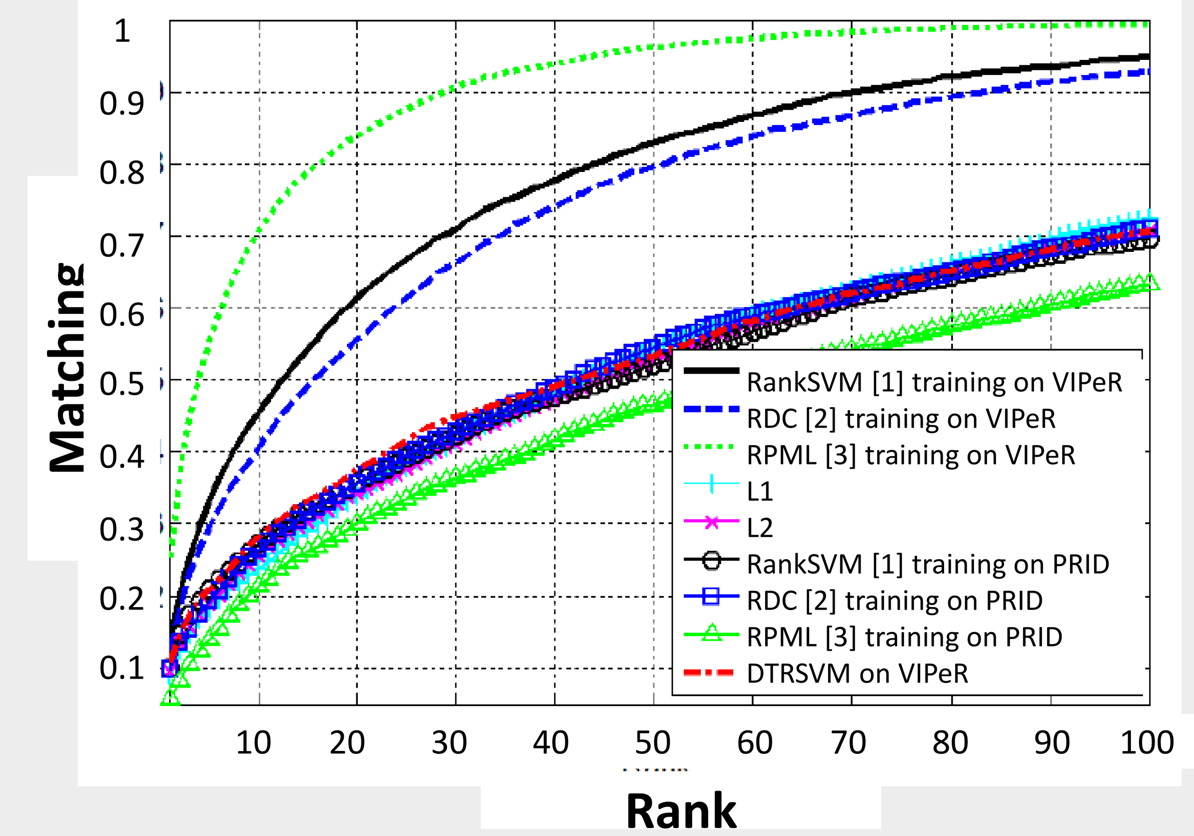
- CMC curves of source: i-LIDS target: VIPeR



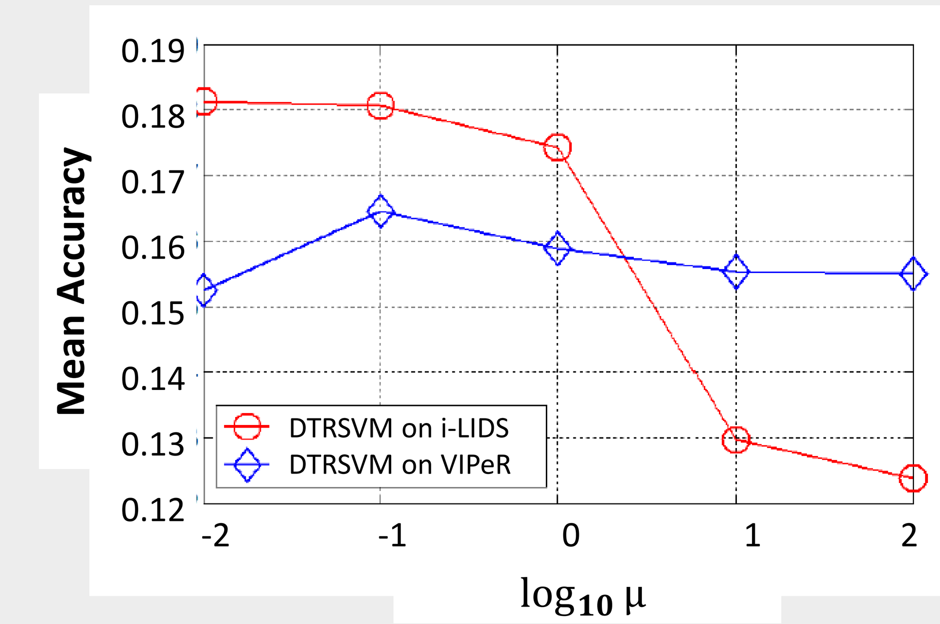
- CMC curves of source: VIPeR target: PRID



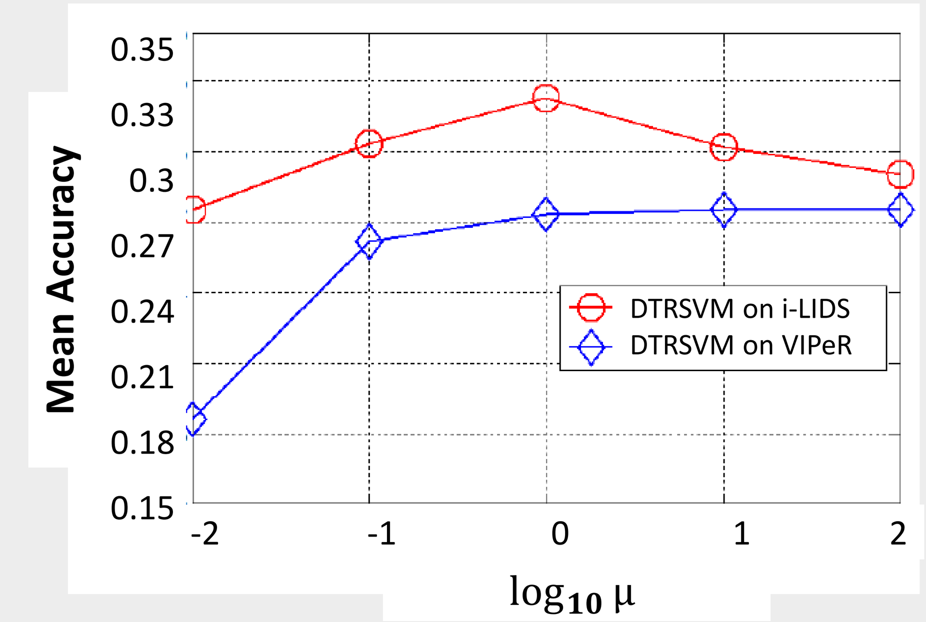
- CMC curves of source: PRID target: VIPeR



- Results on PRID with different μ



- Results on VIPeR with different μ



Conclusions

The performance deteriorates dramatically when using the learnt model trained on source domain
to target domain.
While proposed DTRSVM outperforms existing methods without using the label information in the target domain for training.
Different source domains have an effect on the performance of the proposed DTRSVM.

References

1. B. Prosser, W.-S. Zheng, S. Gong, and T. Xiang. Person reidentification by support vector ranking. In *BMVC*, 2010.
2. W.-S. Zheng, S. Gong, and T. Xiang. Reidentification by relative distance comparison. *TPAMI*, 35(3):653–668, 2013.
3. M. Hirzer, P. M. Roth, M. K¨ostinger, and H. Bischof. Relaxed pairwise learned metric for person re-identification. In *ECCV*, 2012.