Hard Samples Rectification for Unsupervised Cross-domain Person Re-ID

Chih-Ting Liu*, Man-Yu Lee*, Tsai-Shien Chen, Shao-Yi Chien

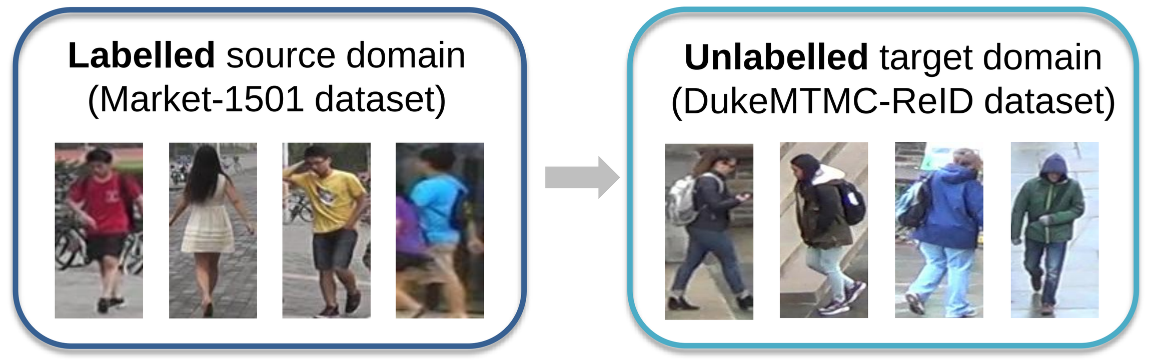
Graduate Institute of Electronics Engineering, National Taiwan University

** denotes equal contributions*

2021 IEEE International Conference on Image Processing, Anchorage, Alaska, USA

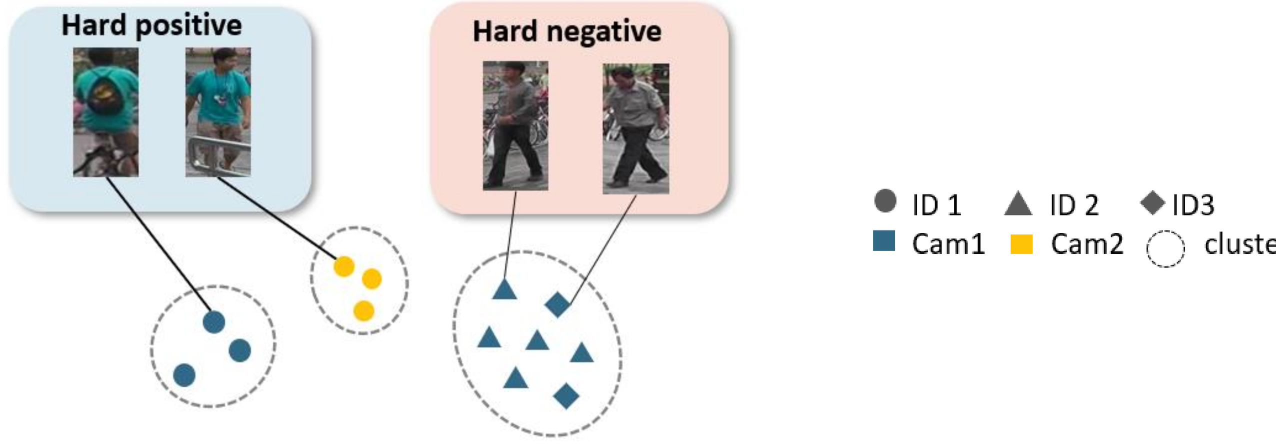
Problem Statement

- **Unsupervised** cross-domain person re-identification
  - Given labelled source domain, perform re-ID on unlabelled target domain



Motivation

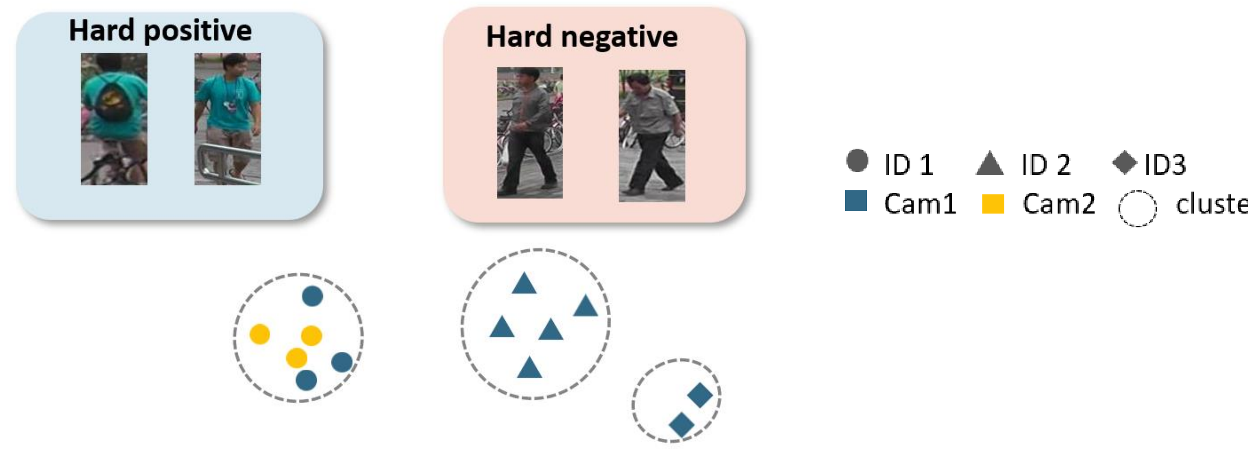
- **Clustering-based method is the mainstream.**
- **Problems in clustering-based Re-ID methods**
  - Hard positive pair → Easily be mis-clustered to different groups
  - Hard negative pair → Different people with similar appearance are in the same group



- **Goal : Rectify hard samples in clustering results**

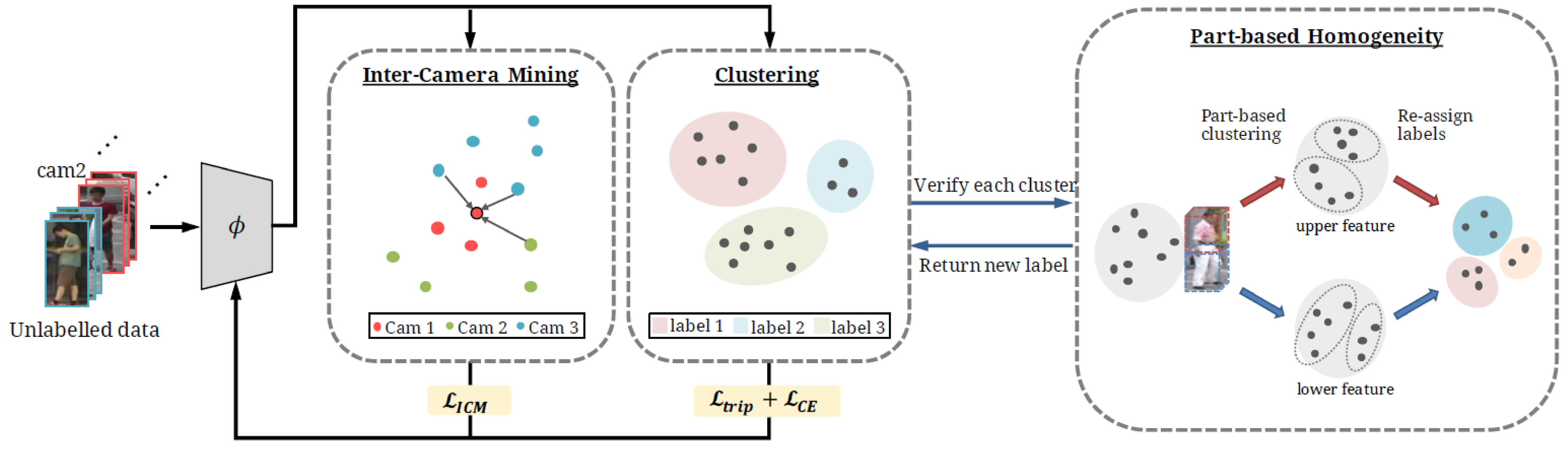
We propose **two** techniques :

1. Inter-Camera Mining **(ICM)** → rectify hard positive samples
2. Part-Based Homogeneity **(PBH)** → rectify hard negative samples



Proposed method

- **Overall architecture :** iterative clustering and CNN training



- **Inter-Camera Mining (ICM)**
  - Mine the pairs with similar features but captured under "different cameras".

Algorithm 1: Inter-Camera Mining

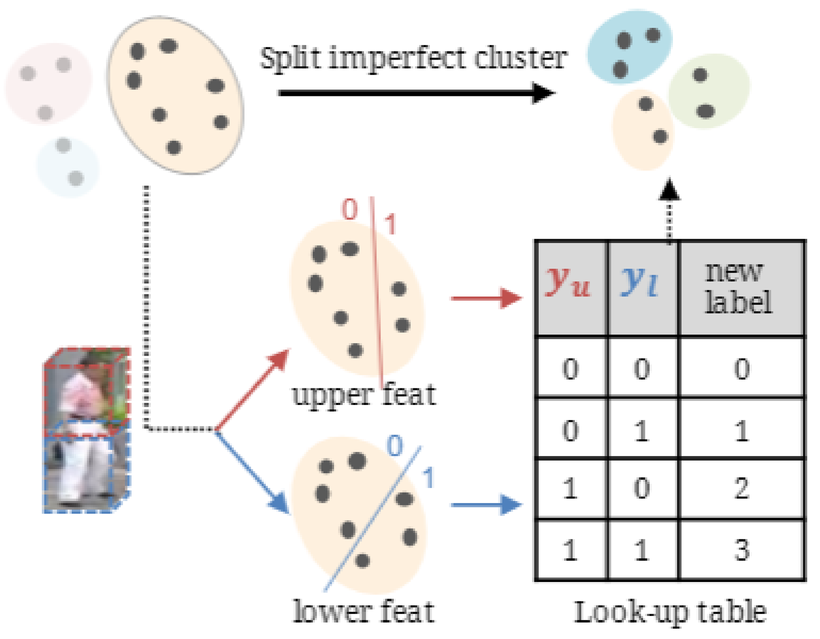
Input: Image feature vectors $\{\phi(I_i^t)\}_{i=1}^{N_t}$ and its camera ID $\{c_i\}_{i=1}^{N_t}$ on target domain

Output: Possible hard positive pairs

Calculate similarity matrix $\mathbf{S} \in \mathbb{R}^{N_t \times N_t}$.
for i=1 ; $i \leq N_t$; i=i+1 **do**
    Sort $\mathbf{S}[i]$ in descending order.
    $Rank(I_i^t)$ = top-K images $\{I_j^t\}_{j=1}^{K}$ in $\mathbf{S}[i]$ with $c_j \neq c_i$
end for
Choose image pairs (I_i^t, I_j^t) conform to $I_j^t \in Rank(I_i^t)$ and $I_i^t \in Rank(I_j^t)$. → Best buddies pairs [1]
return all chosen pairs.

- **Part-Based Homogeneity (PBH)**

1. Use Silhouette score [2] to define imperfect cluster
2. Split features into fine-grained parts
3. Separately cluster features of each part into 2 groups.
4. According to the labels, the imperfect cluster can be split into at most 4 sub-groups.



Experiment Results

(a) Ablation Studies

Experimental setting	$\mathcal{L}_{CE}$ (labelled)	$\mathcal{L}_{trip}$ (unlabelled)	$\mathcal{L}_{ICM}$	PBH	Duke → Marktet R1	Duke → Marktet mAP	Market → Duke R1	Market → Duke mAP
Direct Transfer					50.1	20.9	36.2	18.3
Baseline	✓	✓			72.9	46.3	60.2	42.2
Baseline w/ PBH	✓	✓		✓	74.5	47.1	63.5	44.6
Baseline w/ $\mathcal{L}_{ICM}$	✓	✓	✓		83.8	63.3	73.5	54.4
HSR (Ours)	✓	✓	✓	✓	**85.3**	**65.2**	**76.1**	**58.0**

- **Direct Transfer** : testing with only pretrained model
- **Baseline** : iterative clustering & training with triplet loss and cross-entropy loss

(b) Comparison to state-of-the-arts

Methods	Duke → Market R1	Duke → Market mAP	Market → Duke R1	Market → Duke mAP
PUL [7]	45.5	20.5	30.0	16.4
CAMEL [6]	54.5	26.3	-	-
SPGAN [3]	58.1	26.9	46.9	26.4
HHL [4]	62.2	31.4	46.9	27.2
MAR [19]	67.7	40.0	67.1	48.0
PAST [8]	78.4	54.6	72.4	54.3
SSG [9]	80.0	58.3	73.0	53.4
pMR-SADA [20]	83.0	59.8	74.5	55.8
GDS-H [10]	81.1	61.2	73.1	55.1
HSR (Ours)	**85.3**	**65.2**	**76.1**	**58.1**

Clustering-based methods

Reference

[1] Tali Dekel, et al. Best-buddies similar-ity for robust template matching, In CVPR, 2015

[2] Peter J Rousseeuw. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of computational and applied mathematics, 1987

[3] **Market**: Liang Zheng, Liyue Shen, Lu Tian, Shengjin Wang, Jingdong Wang, and Qi Tian. Scalable person re-identification: A benchmark. In ICCV, 2015.

[4] **Duke** : E. Ristani, F. Solera, R. Zou, R. Cucchiara, and C. Tomasi. Performance measures and a data set for multi-target, multicamera tracking. In European Conference on Computer Vision workshop on Benchmarking Multi-Target Tracking, 2016