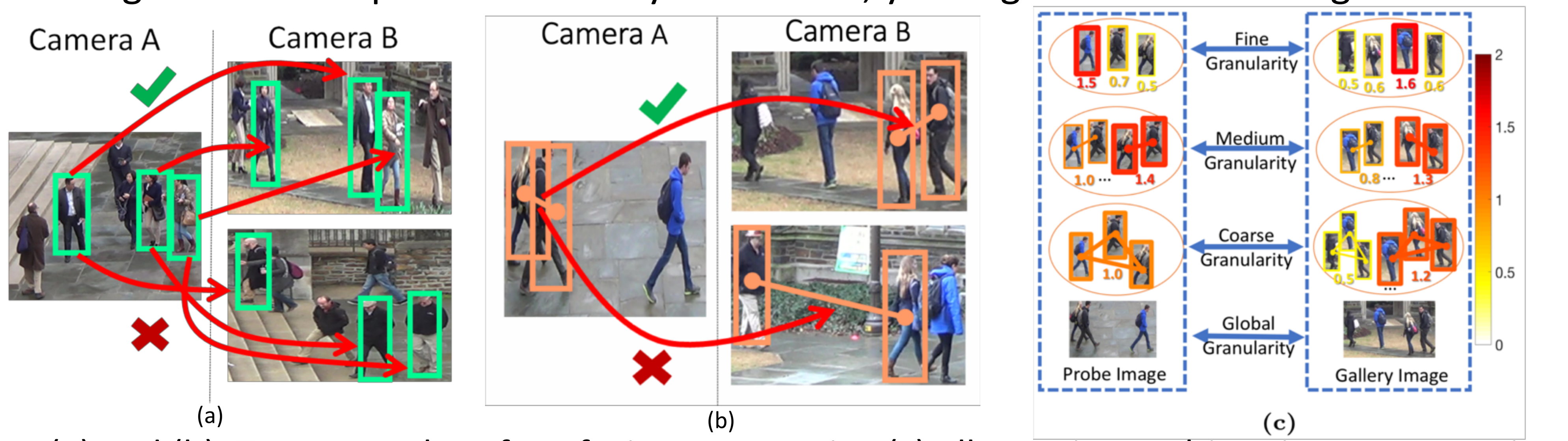
Group Re-Identification: Leveraging and Integrating Multi-grain Information

Hao Xiao[1], Weiyao Lin[1], Bin Sheng[1], Ke Lu[2], Junchi Yan[1], Jingdong Wang[3], Errui Ding[4], Yihao Zhang[5] and Hongkai Xiong[1]

[1]SEIEE, Shanghai Jiao Tong University, [2]University of Chinese Academy of Sciences, [3]Microsoft Research, [4]Baidu Inc., [5]Tencent YouTu Lab

Abstract

Group Re-ID is very challenging since it is not only interfered by view-point and human pose variations in the traditional person Re-ID, but also suffers from group layout and group member variations. To handle these issues, We propose **multi-grain representations** to characterize the appearance and spatial features of multi-grain objects and evaluate the importance weight of each object for group Re-ID. We compute the optimal group-wise matching by using a **multi-order matching process** based on the multi-grain representation and importance weights. Furthermore, we dynamically **update the importance weights** according to the current matching results and then compute a new optimal group-wise matching. The two steps are iteratively conducted, yielding the final matching results.



(a) and (b): Two examples of confusing group pairs, (c): Illustration multi-grain representation

Contributions

Importance weighting to indicate the object's discriminativity and reliability inside the probe group image for group person matching. **Multi-order matching** to integrate multi-grain representation and combines the information of both matched and unmatched objects to achieve a more reliable matching result.

Importance weight of fine-grain object: $\alpha_i = t_1(i, \mathcal{G}_{\setminus i}) + s(i, \mathcal{M}_i) + p(\mathcal{M}_i, \mathcal{M}_{\mathcal{G}_{\setminus i}})$

Static weight (stability): $t_1(i, \mathcal{G}_{\setminus i}) = \lambda_{t_1} \sum_{i' \in \mathcal{G}_{\setminus i}} \frac{\rho_i}{\rho_{i'}}$

Dynamic weight (saliency and purity): $s(i, \mathcal{M}_i) = \lambda_s \frac{d_f(\mathbf{f}_i, \mathbf{f}_{\mathcal{M}_i})}{|\mathcal{M}_i|}$ $\quad p(\mathcal{M}_i, \mathcal{M}_{\mathcal{G}_{\setminus i}}) = \sum_{i' \in \mathcal{G}_{\setminus i}} \lambda_p d_m(\mathcal{M}_i, \mathcal{M}_{i'})$

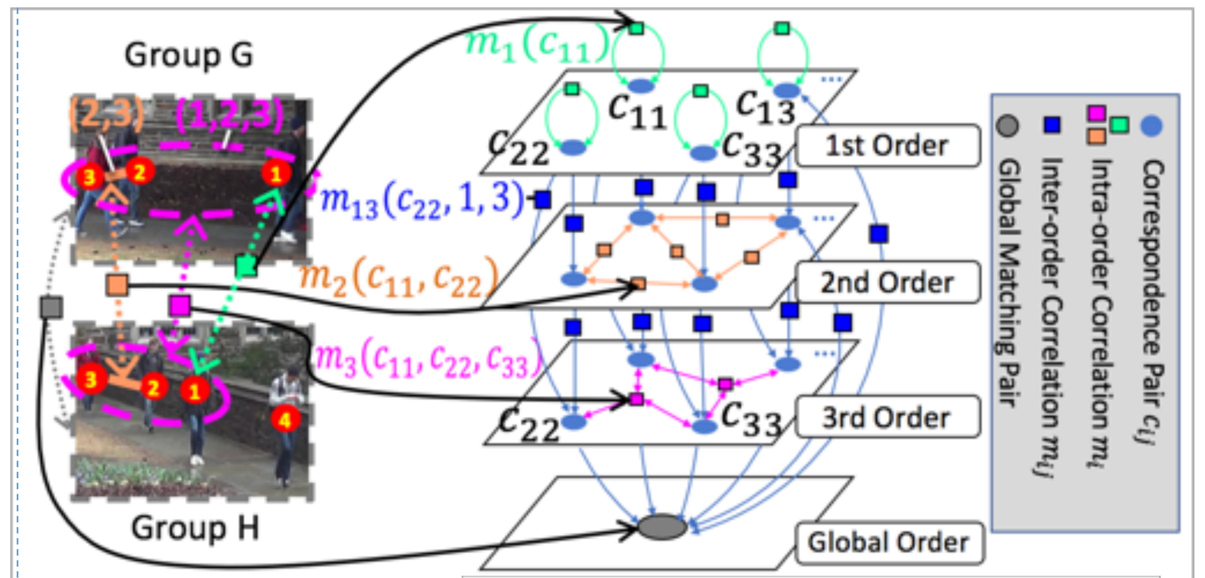
Importance weight of medium-grain and coarse-grain object: $\alpha_{i_1 i_2} = \alpha_{i_1} + \alpha_{i_2} + t_2(i_1, i_2)$

$\alpha_{i_1 i_2 i_3} = \alpha_{i_1 i_2} + \alpha_{i_2 i_3} + \alpha_{i_1 i_3} + t_3(i_1, i_2, i_3)$

Fused Matching Score:

$$S(I_p, I_g) = \sum_k \sum_{(i_1..i_k) \in \mathcal{R}_p} \frac{w_k(\mathbf{f}_{i_1..i_k}, \alpha_{i_1..i_k}, \mathbf{f}_{M(i_1..i_k)}, \alpha_{M(i_1..i_k)})}{|\mathcal{R}_p|} - \lambda_s \cdot \sum_k \left(\sum_{(i_1..i_k) \in \overline{\mathcal{R}}_p} \frac{a_{i_1..i_k}}{|\overline{\mathcal{R}}_p|} + \sum_{(j_1..j_k) \in \overline{\mathcal{R}}_g} \frac{a_{j_1..j_k}}{|\overline{\mathcal{R}}_g|} \right)$$

Our fused scheme integrates four granularities to compute the group-wise matching score. Inside each granularity, we select matched object pairs with high similarities to compute the similarity. Meanwhile, we introduce an unmatched term evaluating the importance of unmatched objects. As such, we can properly avoid misleadingly high matching scores in false group pairs and obtain a more reliable result.

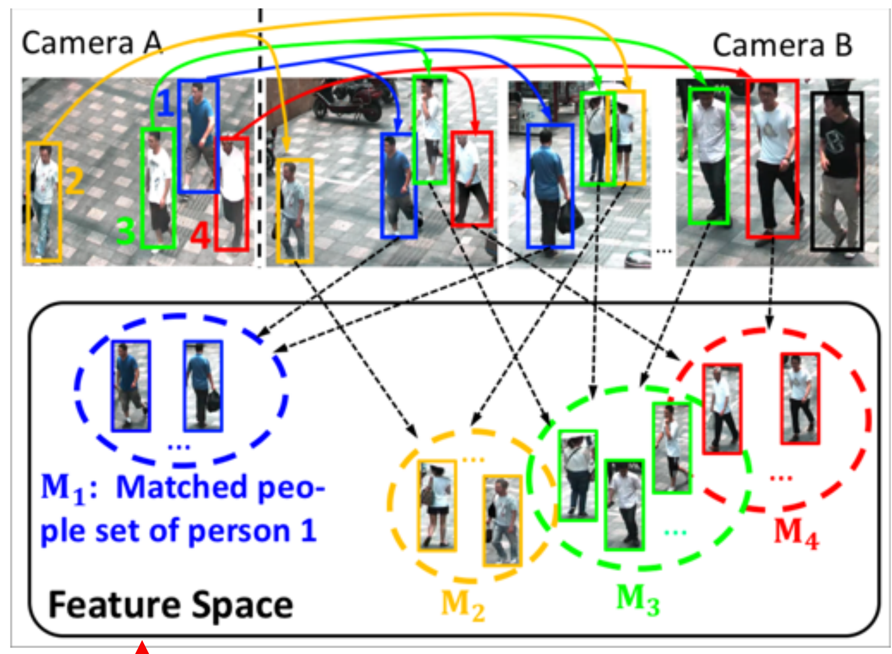


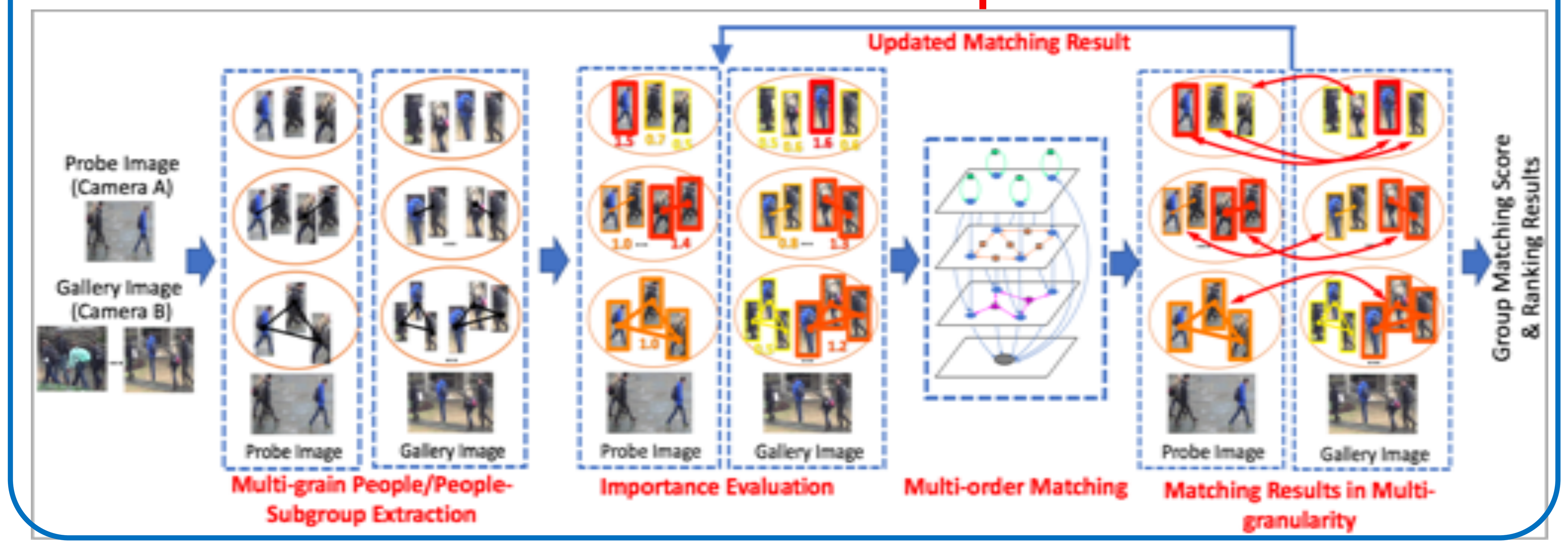
Optimization: $\mathcal{Q}(\mathcal{C}) = \mathcal{P}_1(\mathcal{C}) + \mathcal{P}_2(\mathcal{C}) + \mathcal{P}_3(\mathcal{C}) + \mathcal{P}_g(\mathcal{C}) + \sum_{k_1 \neq k_2, k_1, k_2 = 1,2,3,g} \mathcal{P}_{k_1 k_2}(\mathcal{C})$,

Where $\mathcal{P}_1(\mathcal{C})$, $\mathcal{P}_2(\mathcal{C})$, $\mathcal{P}_3(\mathcal{C})$, $\mathcal{P}_g(\mathcal{C})$ are the first-order, second-order, third-order, and global potentials, evaluating the matching quality over each subgroup of people, and $\mathcal{P}_{k_1 k_2}(\mathcal{C})$ is the inter-order potential.

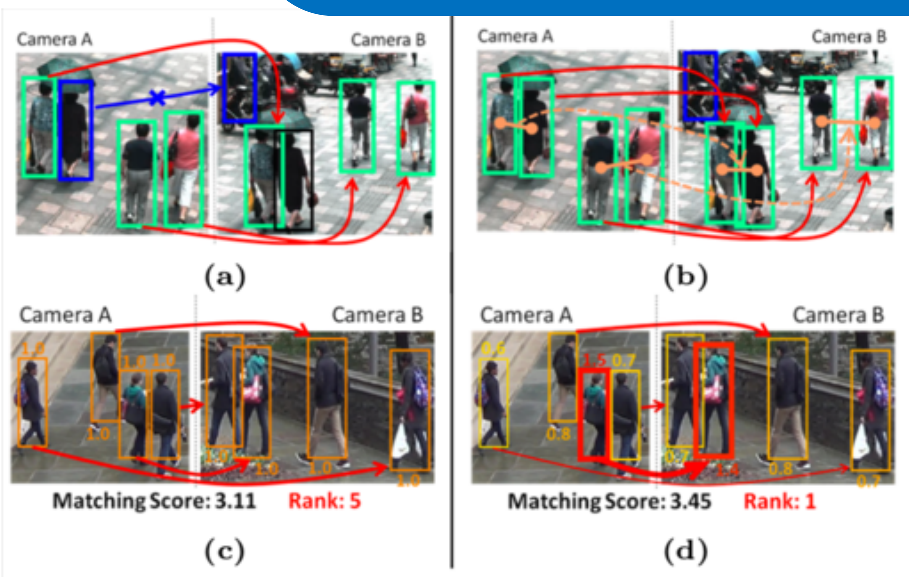
Framework & Algorithm

Given the probe group image captured from one camera, our goal is to find the matched group images from a set of gallery group images captured from another camera. We represent each group image by a set of multi-grain objects, and extract the features for the multi-grain objects. The matching process is an iterative process. We compute the static and dynamic importance weights of multi-grain objects for the probe and gallery images according to the intermediate matching results. Then, we use a multi-order matching algorithm to compute intermediate matching results, which are used to update the dynamic importance weights. We perform the two stages iteratively, and obtain the final matching results.





Comparisons



Matching results by
(a) only using information of individuals;
(b) using multi-grain information;
(c) setting equal importance weights for all individuals/subgroups;
(d) using our importance evaluation process to obtain importance weights.

Table 1: Group Re-ID results on Road Group dataset

Rank	1	5	10	15	20	30
Global	15.8	31.6	43.0	48.6	54.8	61.7
Fine	62.0	82.2	89.6	95.1	96.5	97.3
Fine+Medium	66.7	87.2	93.3	96.0	96.8	97.3
Fine+Medium+Coarse	71.1	89.4	94.1	97.0	97.3	97.5
Equal-weight	55.8	78.0	88.1	92.1	93.6	97.8
Proposed-no spatial	69.6	88.6	94.0	96.2	96.5	97.4
Proposed-auto	72.3	90.6	94.1	97.1	97.5	98.0
Proposed-GT	**76.0**	**91.8**	**95.3**	**97.2**	**98.0**	**98.0**

Table 2: Matching and Re-ID results on Road Group dataset (*MA*: matching accuracy between individuals; *R*-1*(Gr)*: Rank-1 CMC for group Re-ID; *R*-1*(In)*: Rank-1 CMC for individual Re-ID)

Method	Single	No inter	No dis	Hyp-E	Proposed-Auto
MA(%)	83.1	87.0	88.2	86.4	**88.2**
R-1(Gr)	62.0	70.1	65.8	55.1	**72.3**
R-1(In)	60.1	68.9	63.4	53.3	**71.4**

The results show that a) assigning importance weights to different people/people subgroups is significant in guaranteeing group Re-ID performances; b) Our proposed importance evaluation process is effective in finding proper importance weights, such that reliable and discriminative people/people subgroups are highlighted to obtain satisfactory results; (c) our multi-order matching process can make better use of the multi-grain information in groups during matching.

CMC curves for different methods. Dataset from left to right: i-LIDS MCTS, DukeMTMC Group, Road Group