Consistent Re-identification in a Camera Network

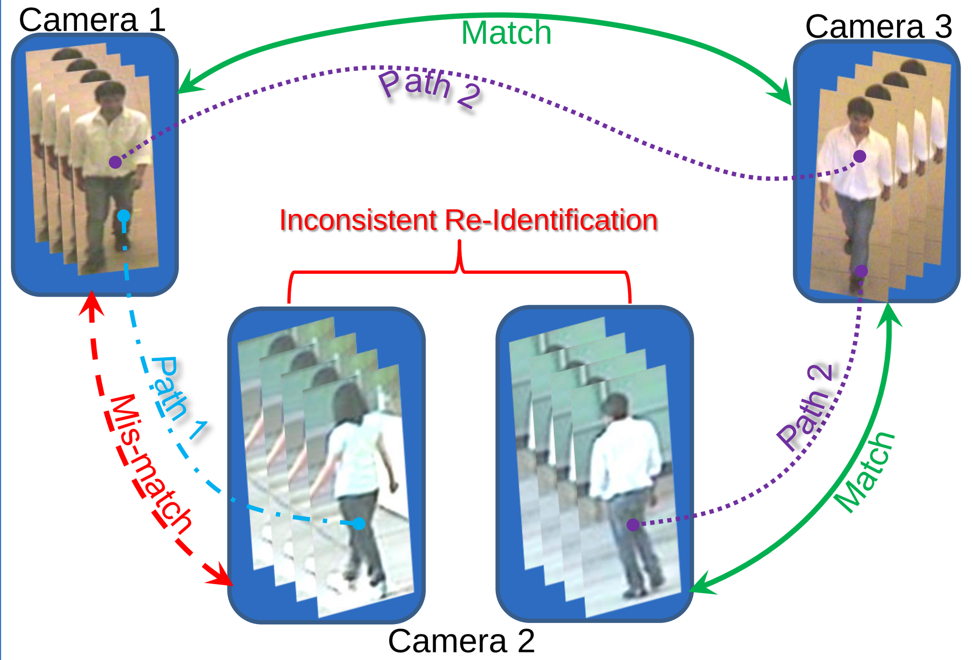
UNIVERSITY OF CALIFORNIA RIVERSIDE
Bourns College of Engineering
Video Computing Group

Abir Das*, Anirban Chakraborty* and Amit K. Roy-Chowdhury
*Joint first authors

ECCV '14
Zürich, September 6-12th, 2014

1. Problem Definition

- Existing person re-identification strategies are camera pair specific.
- High performance in camera pairwise person re-identification does not always mean consistent re-identification across multiple cameras.



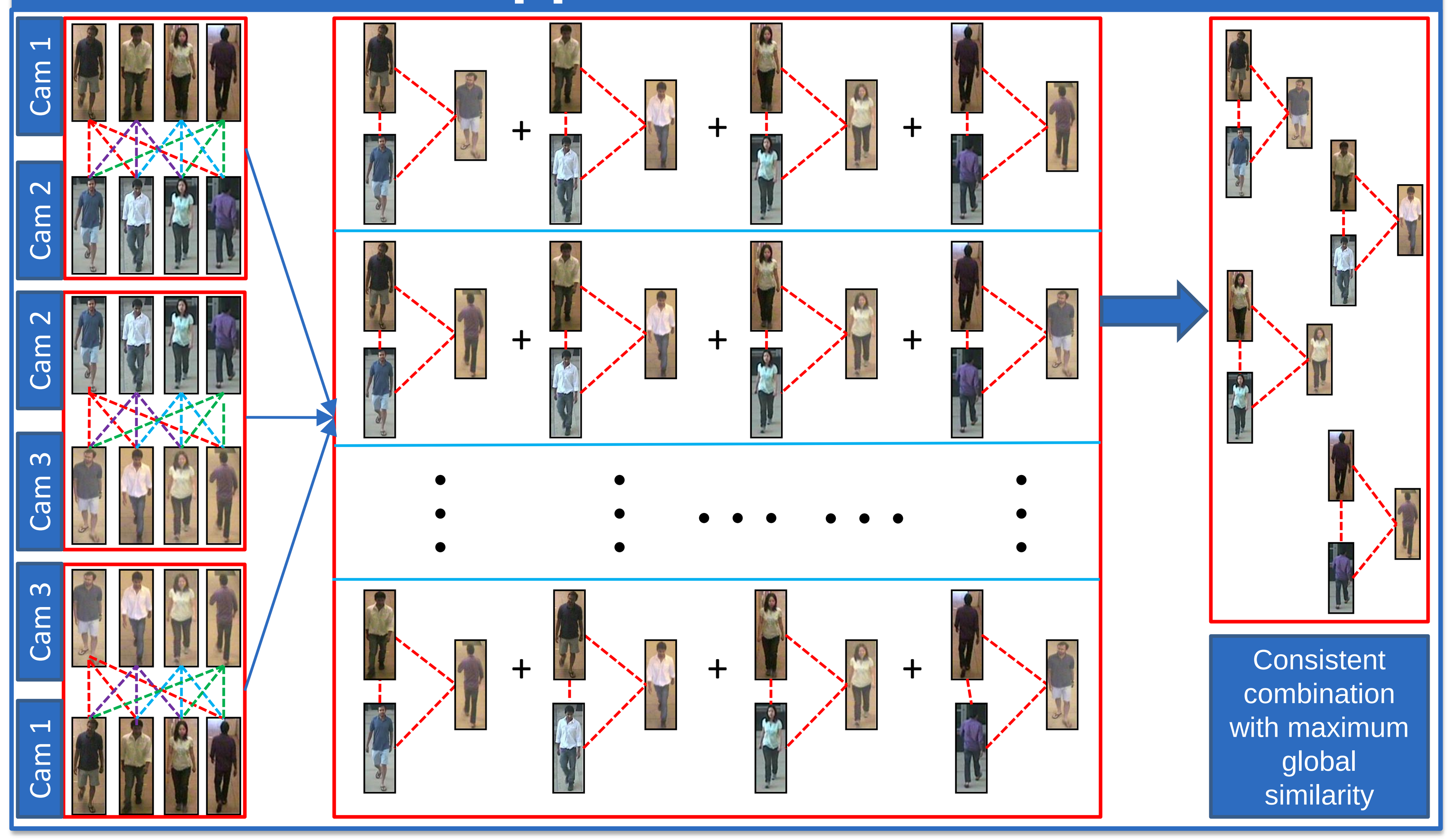
<u>We Asked:</u>

- Can the results be made consistent?
- Will re-identification performance be improved by enforcing consistency?

<u>We Addressed:</u>

- The question by maximizing global similarity across cameras with suitable consistency constraints.
- The problem is posed as a Binary Integer Program and the proposed method is termed as Network Consistent Re-identification (NCR)

2. Approach Overview



3. Approach Details

Input: Camera pairwise similarity score generated by any standard method.

Output: Optimal matching label matrix for each camera pair.

Case I: every person is present in every camera

Cost function:

$$\underset{\substack{x_{i,j}^{p,q} \\ i,j=[1,\cdots,n] \\ p,q=[1,\cdots,m]}}{\operatorname{argmax}} \left(\sum_{\substack{p,q=1 \\ p<q}}^{m} \sum_{i,j=1}^{n} c_{i,j}^{p,q} x_{i,j}^{p,q} \right)$$

m : # of cameras $\quad n$: # of persons

$c_{i,j}^{p,q}$: Similarity score between persons i and j in camera p and q respectively

$$x_{i,j}^{p,q} = \begin{cases} 1 & \text{if person } i \text{ in camera } p \text{ and person } j \text{ in camera } q \text{ are same} \\ 0 & \text{otherwise} \end{cases}$$

Association Constraint: A person from any camera can have only one match from another camera

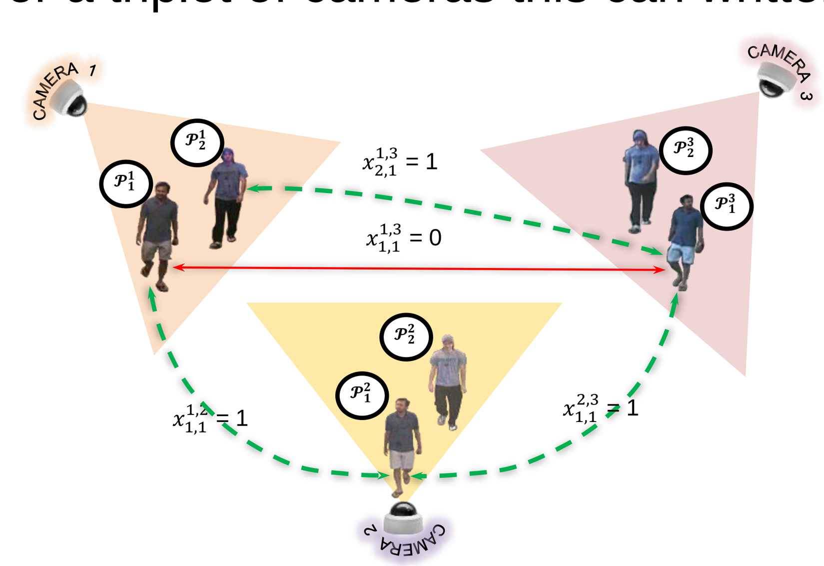
$$\sum_{j=1}^{n} x_{i,j}^{p,q} = 1 \ \forall i = 1 \text{ to } n \ \forall p,q = 1 \text{ to } m,\ p<q$$

$$\sum_{i=1}^{n} x_{i,j}^{p,q} = 1 \ \forall j = 1 \text{ to } n \ \forall p,q = 1 \text{ to } m,\ p<q$$

Loop Constraint: For two persons i and j in cameras p and q respectively we take another person k in another camera r

- If the 3 persons are same then $x_{i,j}^{p,q} = x_{i,k}^{p,r} = x_{k,j}^{r,q} = 1$
- If the 3 persons are not same then at most one of $x_{i,j}^{p,q}$, $x_{i,k}^{p,r}$ and $x_{k,j}^{r,q}$ can be 1

For a triplet of cameras this can written in one line as $x_{i,j}^{p,q} \geq x_{i,k}^{p,r} + x_{k,j}^{r,q} - 1$



- Above relation expresses the consistency constraint for a triplet of cameras
- For a network with more than 3 cameras this relation needs to be held good for every triplet of the cameras

So the 'Loop constraint' exploring consistency over a camera network is given by

$$x_{i,j}^{p,q} \geq x_{i,k}^{p,r} + x_{k,j}^{r,q} - 1$$
$$\forall i,j = [1,\cdots n],\ \forall p,q,r = [1,\cdots m],\ \text{and } p<r<q$$

Case II: every person is not present in every camera

Cost function:

- Use of previous cost gives significant amount of false matches. Say, a person i in camera p does not have any match in camera q. In such a case, the previous cost function $\sum_{j=1}^{n_q} c_{i,j}^{p,q} x_{i,j}^{p,q}$ will be maximum for a target a (say) for which $c_{i,a}^{p,q}$ is maximum. This gives a false match.
- This situation can be avoided if the cost function is changed to $\sum_{j=1}^{n_q} (c_{i,j}^{p,q} - k) x_{i,j}^{p,q}$, where k has a value more than the maximum of the false match scores. This cost function will be maximum when all $x_{i,j}^{p,q}$ are 0, giving no match for person i.

So the modified cost function is

$$\underset{\substack{x_{i,j}^{p,q} \\ i=[1,\cdots,n_p] \\ j=[1,\cdots,n_q] \\ p,q=[1,\cdots,m]}}{\operatorname{argmax}} \left(\sum_{\substack{p,q=1 \\ p<q}}^{m} \sum_{i,j=1}^{n_p,n_q} \left(c_{i,j}^{p,q} - \boxed{k} \right) x_{i,j}^{p,q} \right)$$

Choosing k:

- In an ideal case where the false match and true match scores are well separated smooth functions k can be found analytically.
- In a practical scenario, k is chosen by cross validating on a set of training data

Association constraint: Now a person form a camera can have no match in another camera. So the equality constraints in the previous association constraint becomes inequality constraints.

$$\sum_{j=1}^{n_q} x_{i,j}^{p,q} \boxed{\leq} 1 \ \forall i = [1,\cdots,n_p] \ \forall p,q = [1,\cdots,m],\ p<q$$

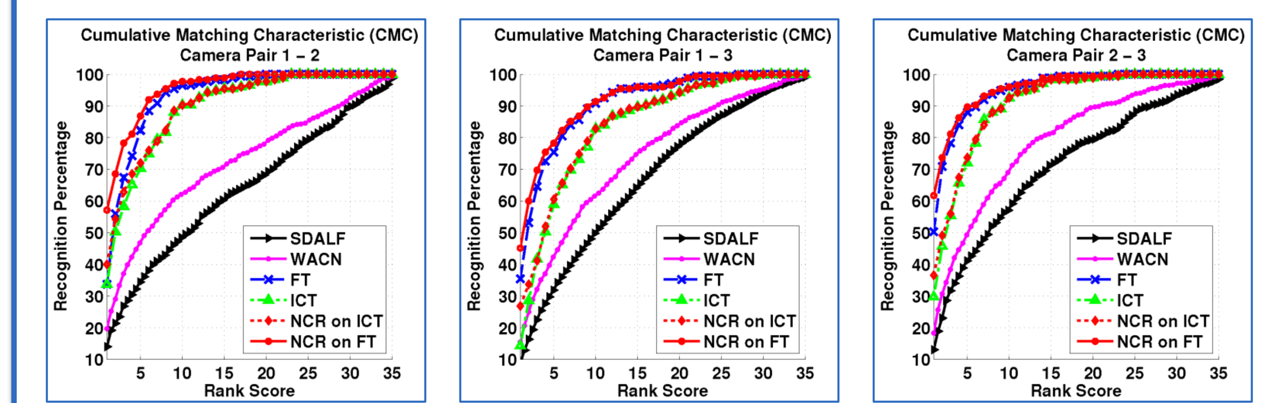
$$\sum_{i=1}^{n_p} x_{i,j}^{p,q} \boxed{\leq} 1 \ \forall j = [1,\cdots,n_q] \ \forall p,q = 1 \text{ to } m,\ p<q$$

Loop Constraint: The loop constraints remain same

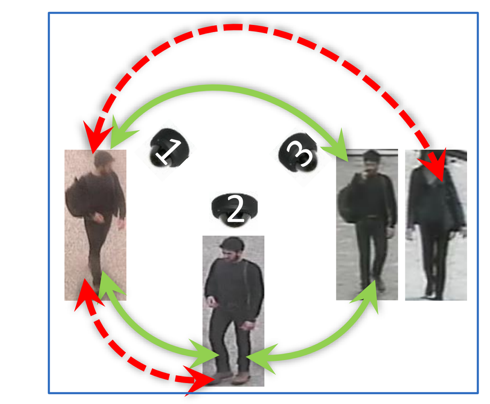
4. Experimental Results

- Most of the benchmark person re-identification datasets (*e.g.*, ETHZ, CAVIAR4REID, CUHK) are several sequences of 2 camera datasets.
- We validated the proposed approach on a publicly available WARD (3 camera dataset) and a 4 camera dataset 'Re-identification Across indoor-outdoor Dataset' (RAiD) released with this work. RAiD is available at http://www.ee.ucr.edu/~amitrc/datasets.php

Results on WARD dataset:

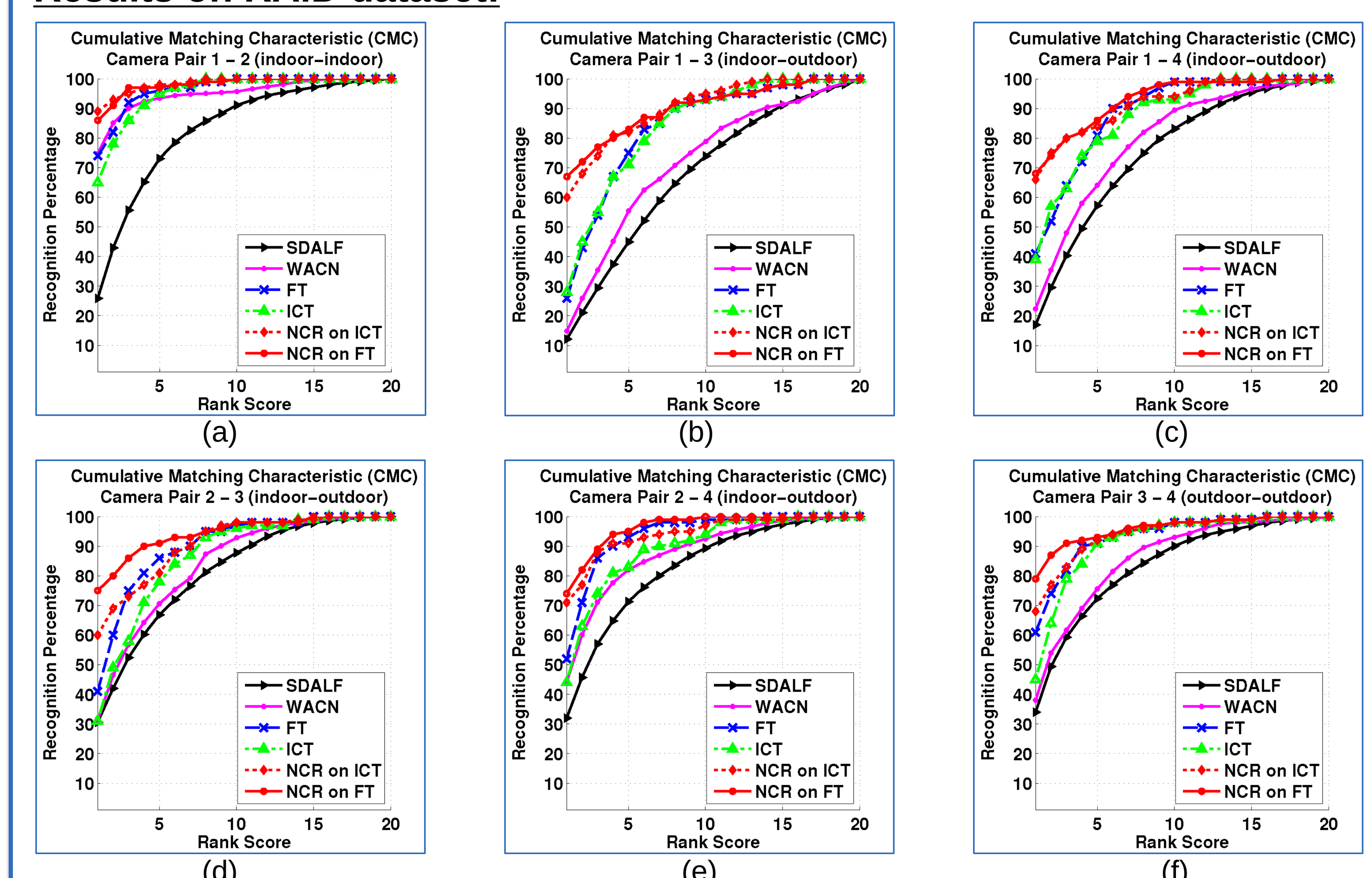


CMC curves showing comparative performance with 3 state-of-the-art and 1 baseline method. NCR applied on similarity scores generated by different methods outperforms the rest.



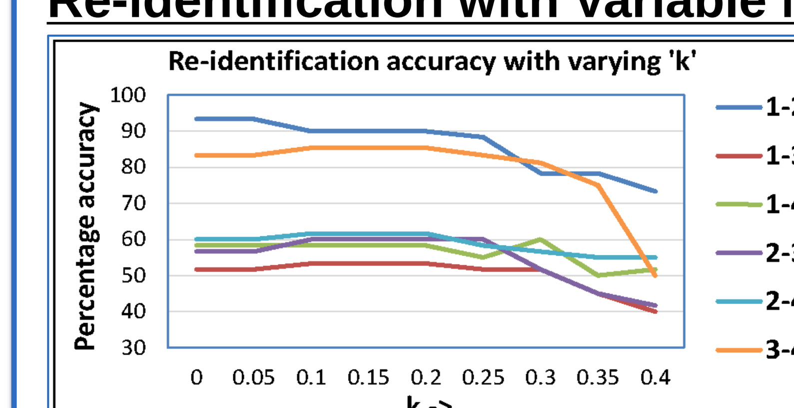
An example of correction of inconsistent re-identification on application of NCR strategy

Results on RAiD dataset:

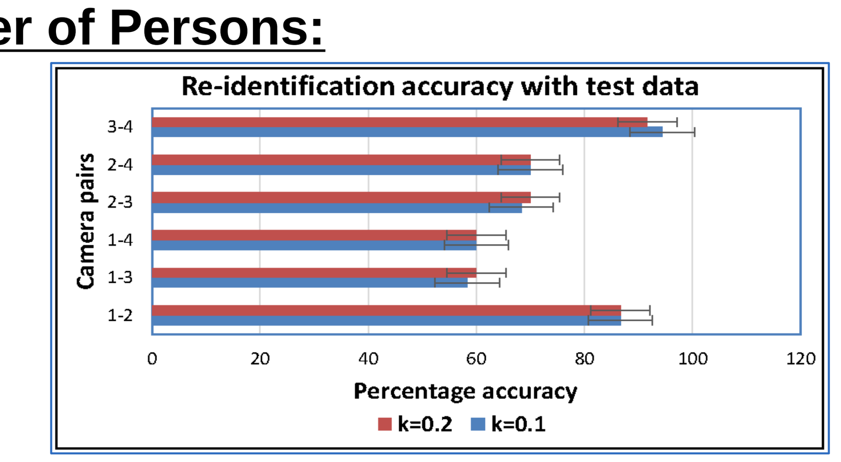


CMC curves showing comparative performance with the same methods. NCR applied on similarity scores generated by different methods outperforms the rest. For camera pairs with large illumination variation (1-3, 1-4, 2-3 and 2-4) the relative performance improvement is significantly large.

Re-identification with Variable Number of Persons:



Variation of training accuracy with k. Camera pairs (1-2 and 3-4) with all the same persons show a decreasing trend while the rest of the camera pairs show an increasing and then decreasing trend.



Re-identification accuracy on test data for two different values of k.

$$\text{Accuracy} \triangleq \frac{(\#\text{ true match} + \#\text{ true no-match})}{\#\text{ of unique people in the testset}}$$

Acknowledgement:

This work was partially supported by NSF grants IIS-1316934 and CNS-1330110.