Person Re-Identification

A Body-based Recognition

Wei-Shi Zheng (郑伟诗)

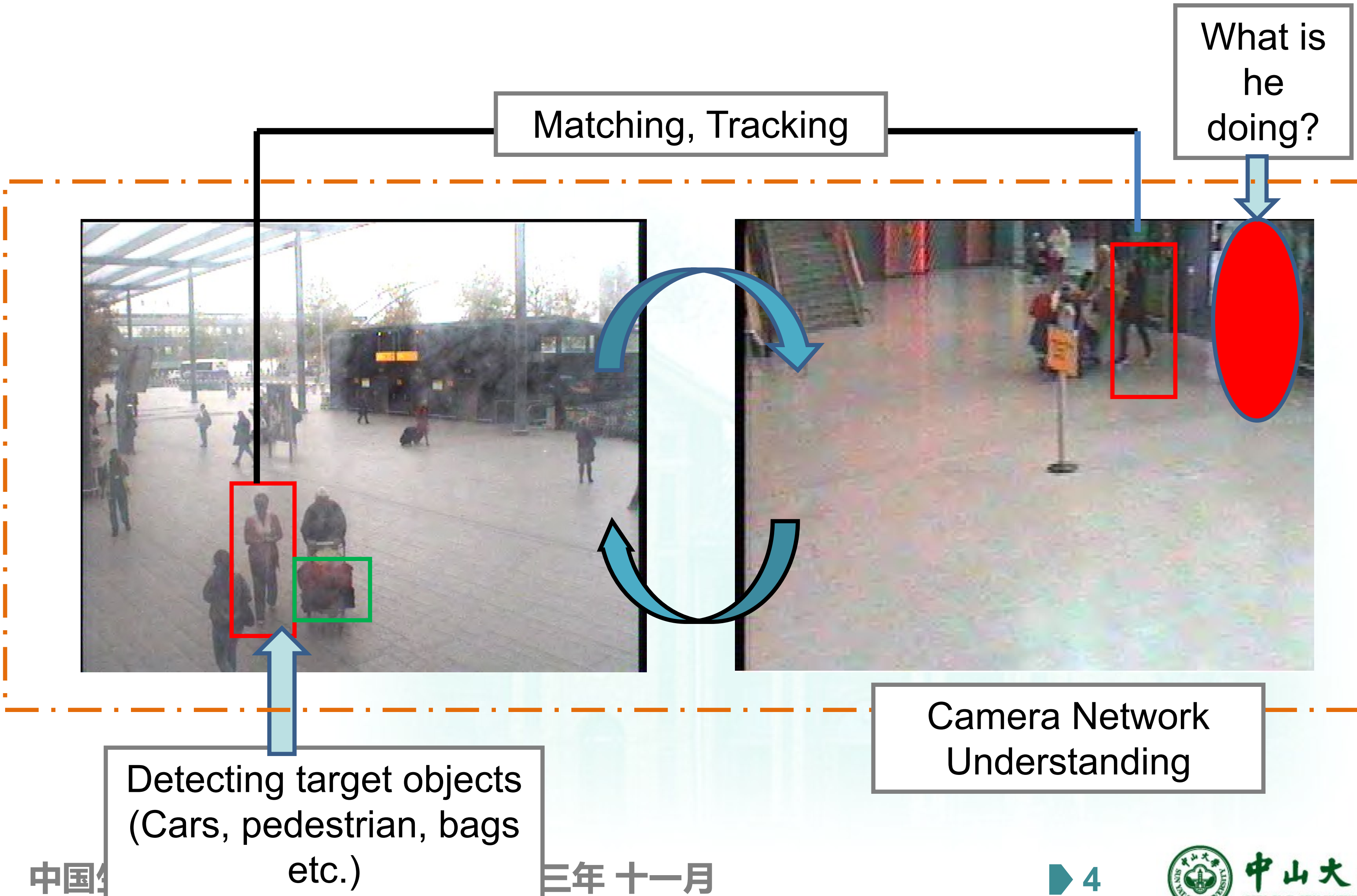
wszheng@ieee.org

Outline

- What is person re-identification
- How to describe a Person Image
- How to measure two Person Descriptors
- Current Challenges

What is Person Re-identification

Person Re-identification



Person Re-identification

The system has two basic parts：

- Capture a unique/reliable person representation

  e.g. Robust and discriminative visual descriptors need to be extracted

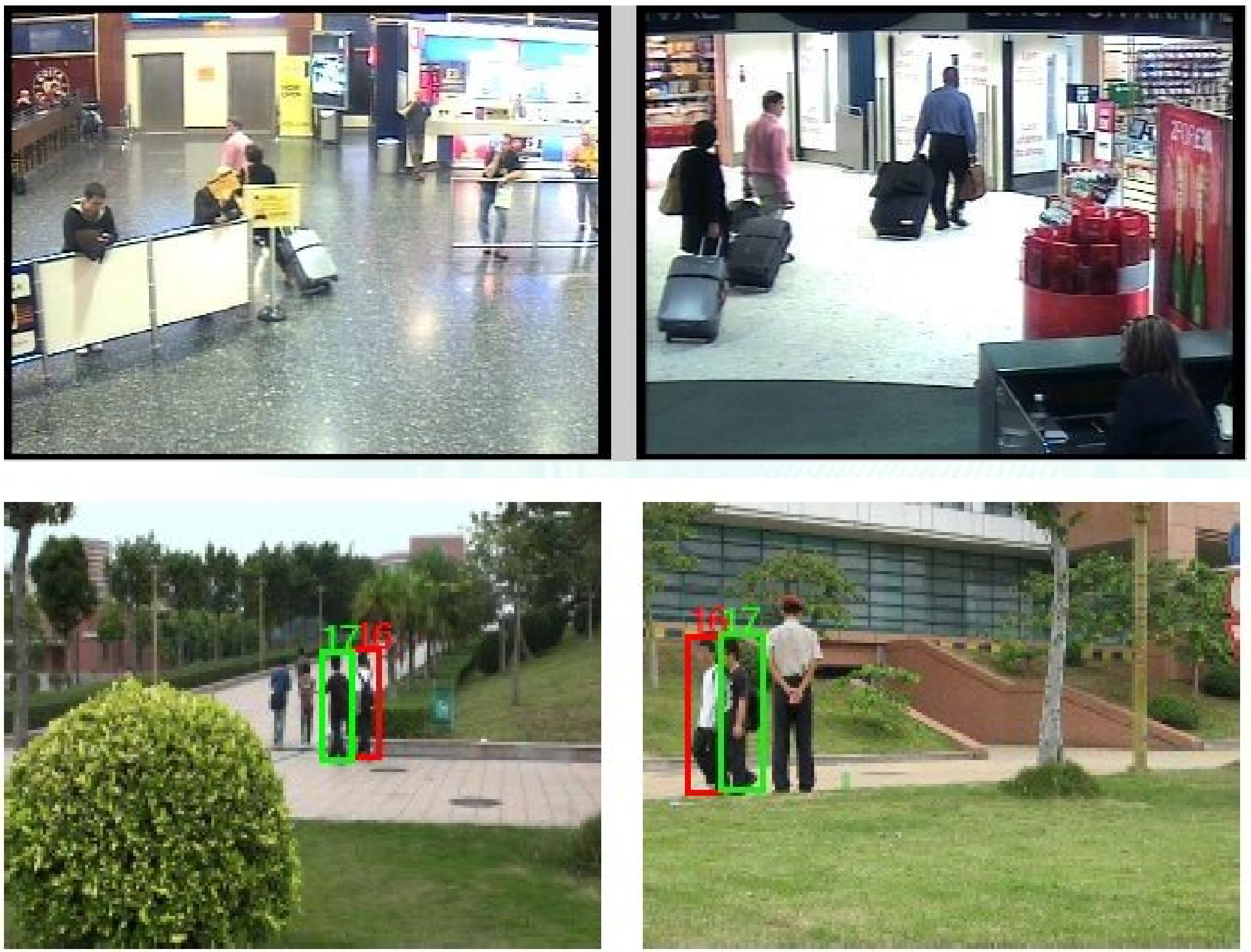
- Compare two representations

  e.g. Learning suitable distance metrics that maximise the chance of a correct correspondence.

Let us First See Some Images

Person Re-identification

Cross-camera Views (Non-overlapping)



Public Datasets

VIPeR (Benchmark Dataset, 632 / 2)

Fig. 1. Some examples from the viewpoint invariant pedestrian recognition (VIPeR) dataset [17]. Each column is one of 632 same-person example pairs.

Douglas Gray and Hai Tao, "Viewpoint Invariant Pedestrian Recognition with an Ensemble of Localized Features," ECCV 2008

Public Datasets

i-LIDS (Benchmark Dataset, 119 / 4)

Wei-Shi Zheng et al., "Person Re-identification by Probabilistic Relative Distance Comparison", CVPR 2011.

Public Datasets

ETHZ (By moving camera, 146 / 60)

W. Schwartz and L. Davis. Learning discriminative appearance-based models using partial least squares. In Brazilian Symposium on Computer Graphics and Image Processing, 2009.

Public Datasets

CAVIAR4REID (Extracted from the CAVIAR dataset)

D.S. Cheng, M. Cristani, et al., "Custom pictorial structures for re-identification," BMVC 2011.

Person Re-identification

Characteristics：

- View change
- Lighting change
- Occlusion
- (Relative) low resolution
- Short-period-of-time

Large Intra-class Variation & Large Inter-class Variation

＋ **Limited Samples**

Questions to Ask

- **How to obtain reliable features?**
  - Invariant to lighting
  - Invariant to pose variation
  - Invariant to occlusion
  - ……….
- **If features cannot be robust, what can we do?**
  - Selecting features
  - Finding a discriminant distance
  - Other cues to help?
  - ……….

Descriptors for Person Re-identification

Histogram

- Divide an Image into Stripe
- Color Bits: RGB, HSV, YCbCr
- Filter Responses (Gabor & Schmid)

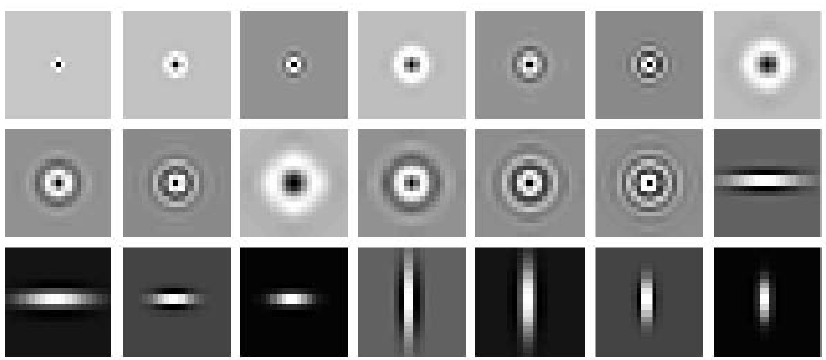
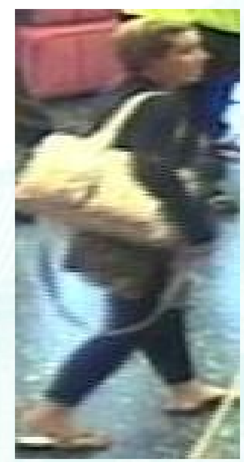
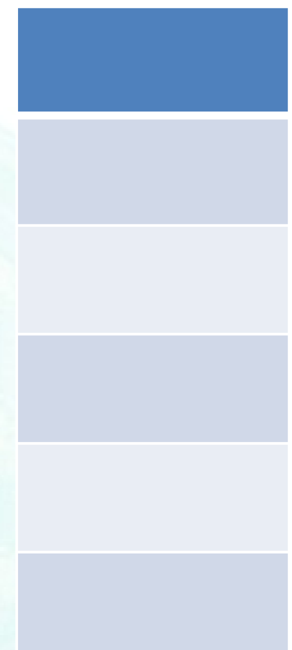


Fig. 3. The filters used in the model to describe texture. (a) Rotationally symmetric Schmid filters. (b) Horizontal and vertical Gabor filters.





Douglas Gray and Hai Tao, "Viewpoint Invariant Pedestrian Recognition with an Ensemble of Localized Features," ECCV 2008

Colour Invariant Modelling

Image Acquisition Process

$$\rho_k = \int_\omega E(\lambda)S(\lambda)Q_k(\lambda)d\lambda \quad (k = 1, 2, 3).$$

$E(\lambda)$ is the spectral distribution of the illuminant

$S(\lambda)$ is the surface reflection spectral distribution

$Q_k(\lambda)$ is the sensor response function characterizing the proportion of color signal absorbed by the sensor k

$$\begin{pmatrix} R^c \\ G^c \\ B^c \end{pmatrix} = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{pmatrix} \begin{pmatrix} R^o \\ G^o \\ B^o. \end{pmatrix}$$

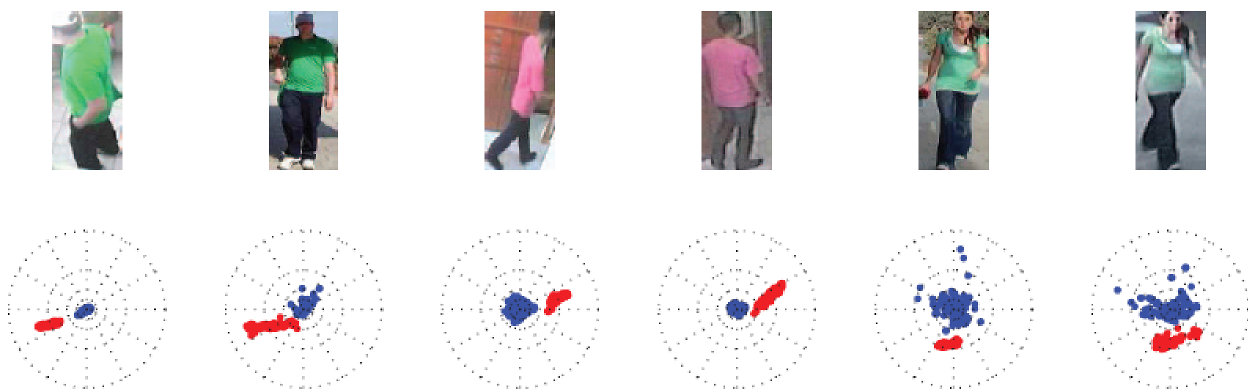
I. Kviatkovsky, A. Adam, and E. Rivlin, "Color Invariants for Person Reidentification," IEEE TPAMI, 2013

Colour Invariant Modelling

Invariant Coding by Shape Context

$$\xi_1 = \ln\frac{R}{G}, \quad \xi_2 = \ln\frac{B}{G}$$

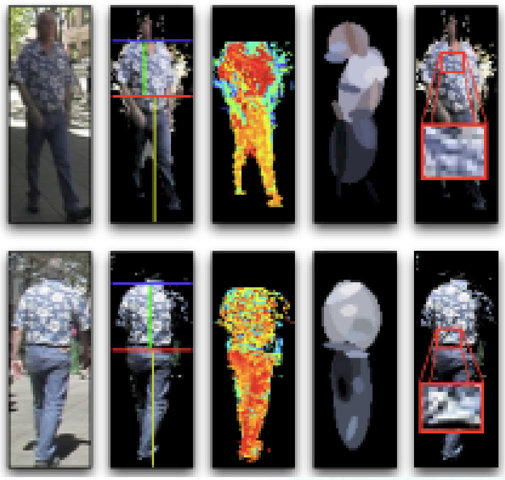
$$(\xi_1^c, \xi_2^c) = \left(\ln\frac{\alpha R^o}{\beta G^o}, \ln\frac{\gamma B^o}{\beta G^o}\right) = (\xi_1^o, \xi_2^o) + \left(\ln\frac{\alpha}{\beta}, \ln\frac{\gamma}{\beta}\right)$$

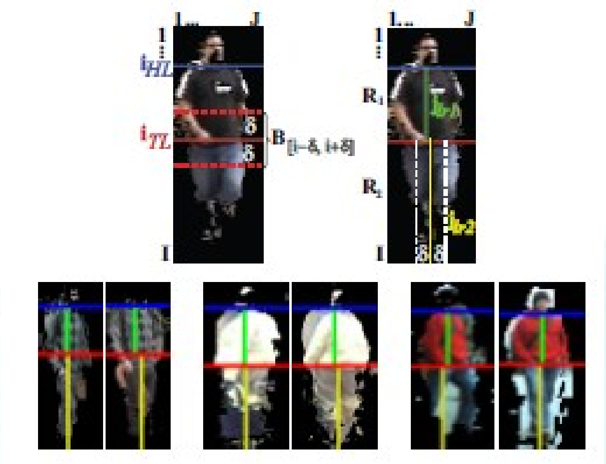


I. Kviatkovsky, A. Adam, and E. Rivlin, “Color Invariants for Person Reidentification,” IEEE TPAMI, 2013

Symmetry-Driven Accumulation of Local Features

Combining Multiple Features/Parts

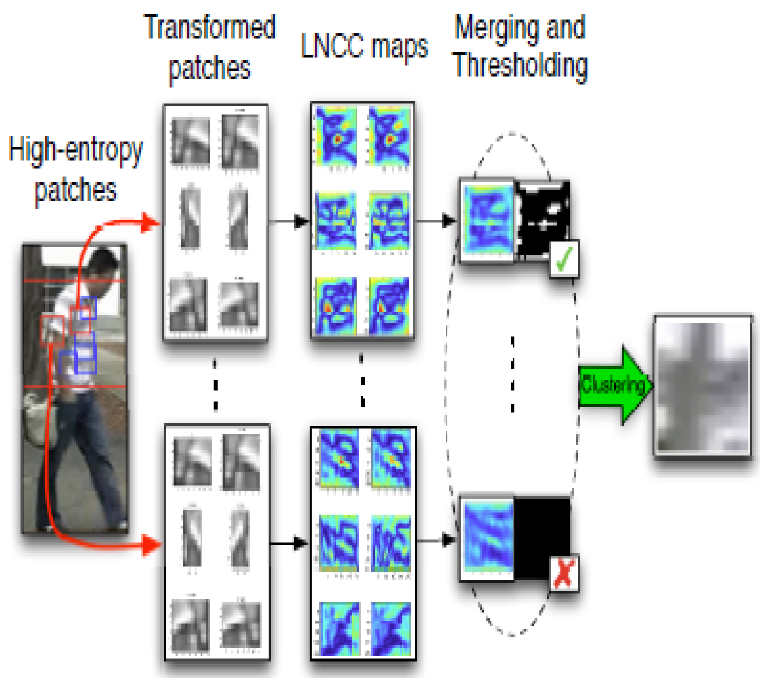




- **Find axes of asymmetry and symmetry**
  - — Two horizontal axes get three main regions isolated：head，torso and legs

M. Farenzena et al., "Person Re-Identification by Symmetry-Driven Accumulation of Local Features," CVPR 2010.

Symmetry-Driven Accumulation of Local Features



- **Different features are extracted**
  - Weight Color Histograms
  - Maximally Stable Color Regions
  - Recurrent High-Structured Patchs
- **Features matching**
  - Combining Different Metrics

M. Farenzena et al., "Person Re-Identification by Symmetry-Driven Accumulation of Local Features," CVPR 2010.

Spatiotemporal Appearance

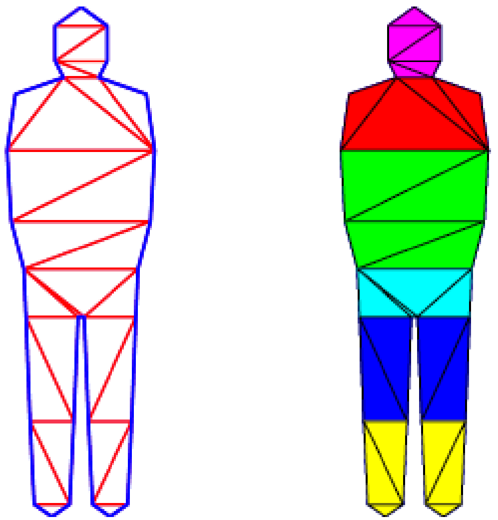
Spatiotemporal segmentation algorithm is employed to generate salient edgels

Combining normalized color and salient edgel histograms

Robust to changes in appearance of clothing

Niloofar Gheissari et al., " Person Reidentification Using Spatiotemporal Appearance," CVPR 2006.

Spatiotemporal Appearance



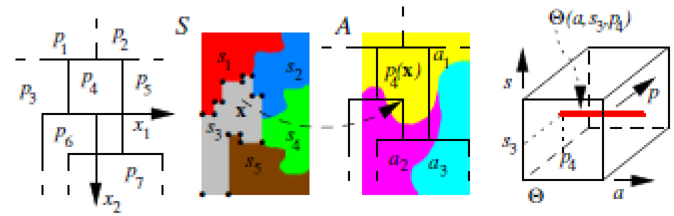
Use a decomposable triangulated graph as a novel method for model fitting to people

Niloofar Gheissari et al., " Person Reidentification Using Spatiotemporal Appearance," CVPR 2006.

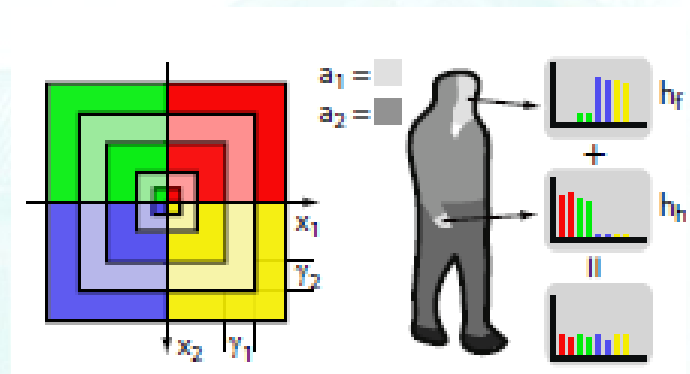
Shape and Appearance Context

Capturing the spatial relations

- The appearance context



- The shape and appearance context



Xiaogang Wang et al., "Shape and Appearance Context Modeling," ICCV 2007.

Shape and Appearance Context

- **Computing a local shape description of the of image**

HOG Log-RGB operator

$$\varphi(\mathbf{x}) \doteq (\mathrm{HOG}(\nabla \log(I_R), \mathbf{x}); \mathrm{HOG}(\nabla \log(I_G), \mathbf{x}); \mathrm{HOG}(\nabla \log(I_B), \mathbf{x}))$$

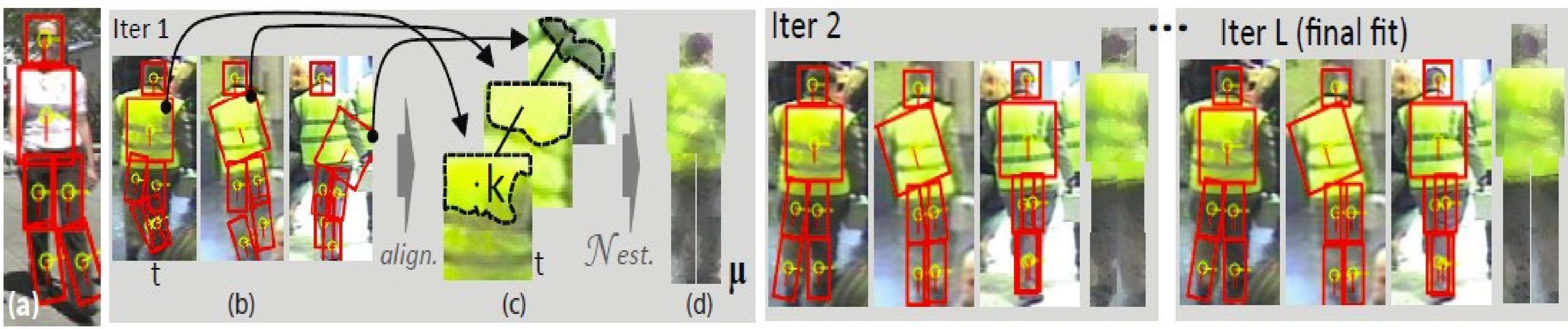
Advantages：

- Perform similar to the homomorphic filtering
- Make the descriptor robust to illumination changes

Xiaogang Wang et al., "Shape and Appearance Context Modeling," ICCV 2007.

Custom Pictorial Structures

- Fitting a Pictorial Structure（PS）on images
- Use a modified HSV characterization
- Use the Maximally Stable Color Region operator（MSCR）



D.S. Cheng, M. Cristani, et al., "Custom pictorial structures for re-identification," BMVC 2011.

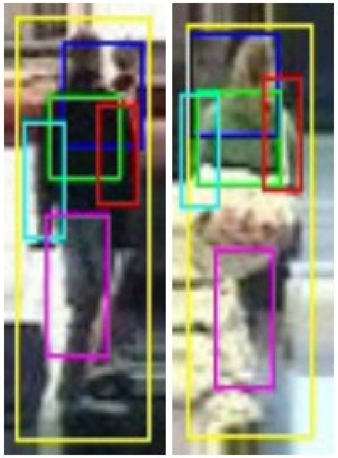
Spatial Covariance Regions

- **Human detector and human body parts detector**

  Return six regions

- **Color normalization**

  Apply histogram equalization to maximize entropy



S. Bak, E. Corvee, F. Bremond, M. Thonnat, Person re-identification using spatial covariance regions of human body parts.

Spatial Covariance Regions

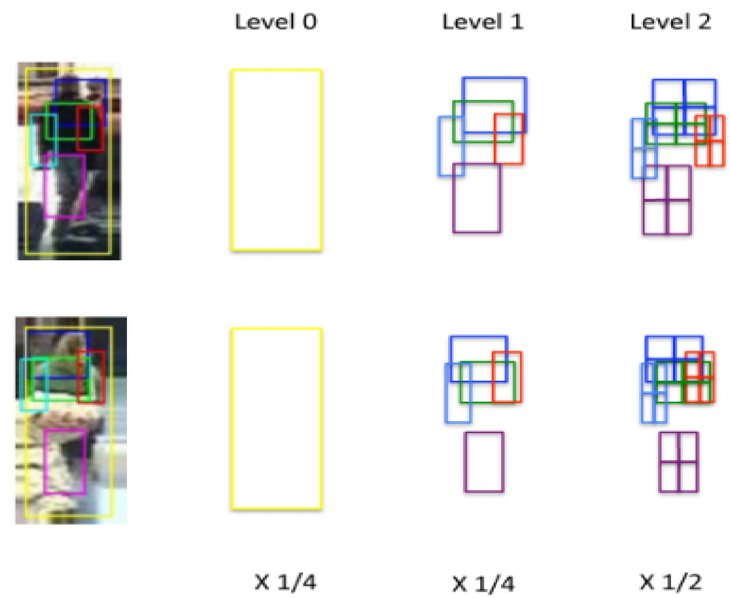
- **Covariance Regions**

Let $\{f_k\}_{k=1\ldots n}$
be the d-dimensional
feature points inside R

$$C_R = \frac{1}{n-1}\sum_{k=1}^{n}(f_k - \mu)(f_k - \mu)^T$$

μ is mean of points

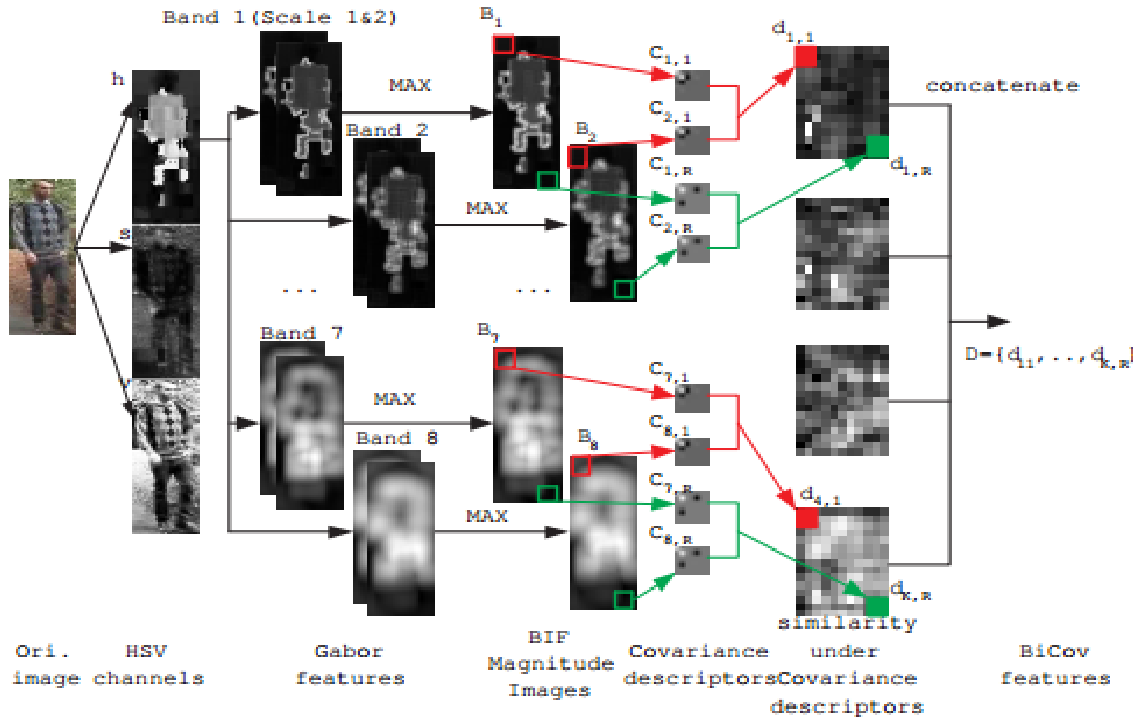
- **Spatial Pyramid Matching**



S. Bak, E. Corvee, F. Bremond, M. Thonnat, "Person re-identification using spatial covariance regions of human body parts," ICPR 2012

Covariance Descriptor based on Bio-inspired Features

BiCov is a two stages representation



B. Ma, Y. Su, F. Jurie, Bicov， “A novel image representation for person re-identification and face verification “, BMVC 2012

Covariance Descriptor based on Bio-inspired Features

- **Biologically inspired features are first extracted**
  - Use Gaber filters

$$G(\mu, v) = I(x, y) * \psi_{\mu,v}(z)$$

v is quantized into 8 orientations

$$\psi_\mu(z) = \frac{1}{8}\sum_{v=1}^{8} \psi_{\mu,v}(z) \Rightarrow G(\mu)$$

  - Capture BIF

$$B_i = \max(G(2i-1), G(2i))$$

- **Encoded by difference of covariance descriptors**

B. Ma, Y. Su, F. Jurie, Bicov，"A novel image representation for person re-identification and face verification “, BMVC 2012

Attributes

Attribute

Define the 15 binary attributes $p(a_i|\mathbf{x})$

Attribute Detection

- Train SVM to detect attributes
- Each attribute detector into a 15 dimensional vector

$$A(\mathbf{x}) = [p(a_1|\mathbf{x}), \ldots, p(a_{N_a}|\mathbf{x})]^T$$

R. Layne, T. M. Hospedales, S. Gong, "Towards person identification and re-identification with attributes", BMVC 2012

Local descriptors encoded by Fisher vectors

- **Design a 7-dimension local descriptor**

$$f(x,y,I) = (x, y, I(x,y), I_x(x,y), I_y(x,y), I_{xx}(x,y), I_{yy}(x,y))$$

- **Train the GMM model**
  - Model the data with a generative model
- **Compute image representations by using Fisher vector**
  - A powerful method for aggregating local descriptors
  - Compute the gradient of the likelihood of the data with respect to the parameters of the model

$$\nabla_\lambda \log p(M|\lambda)$$

$M = \{m_t, t = 1, \ldots, T\}$ be the set of the T local descriptors

B. Ma, Y. Su, F. Jurie, "Local descriptors encoded by fisher vectors for person re-identification", ECCV 2012

Local descriptors encoded by fisher vectors

LDFV Extensions

- **Bin-based LDFV: using spatial Information**
- **Enriched LDFV: combining LDFV with other features**
  - Weighted Color Histograms (wHSV)
  - Maximally Stable Color Regions (MSCR)
- **Supervised LDFV: using metric learning**

B. Ma, Y. Su, F. Jurie, "Local descriptors encoded by fisher vectors for person re-identification", ECCV 2012

How to Quantify the Description

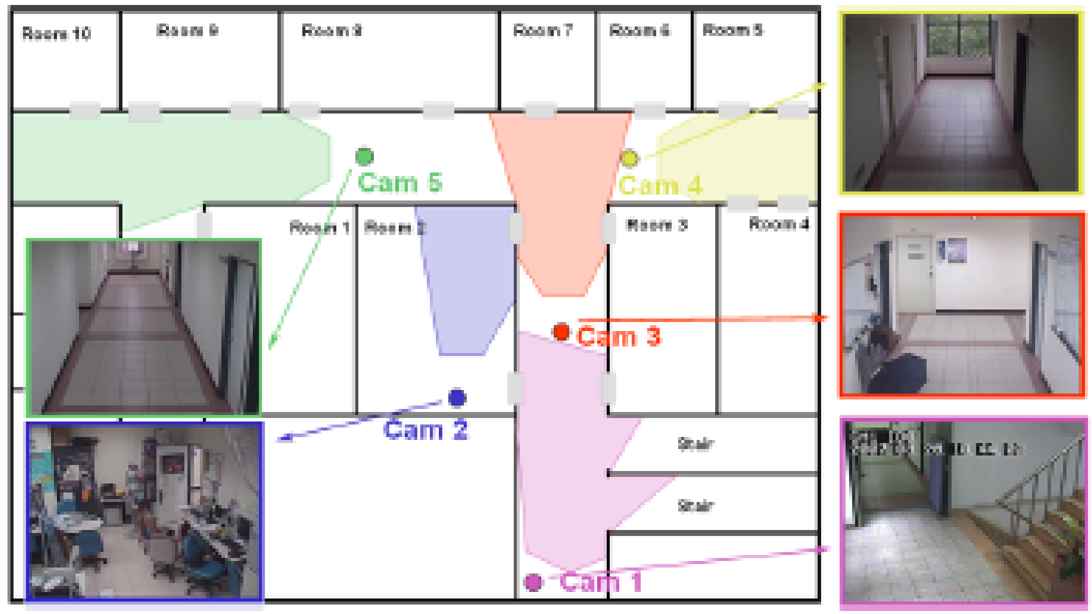
Why? Features can be distorted

Brightness Transfer Functions

$$f_{ij}(B_i) = H_j^{-1}(H_i(B_i))$$

↑

the BTF for every pair of observations Oi and Oj in the training set

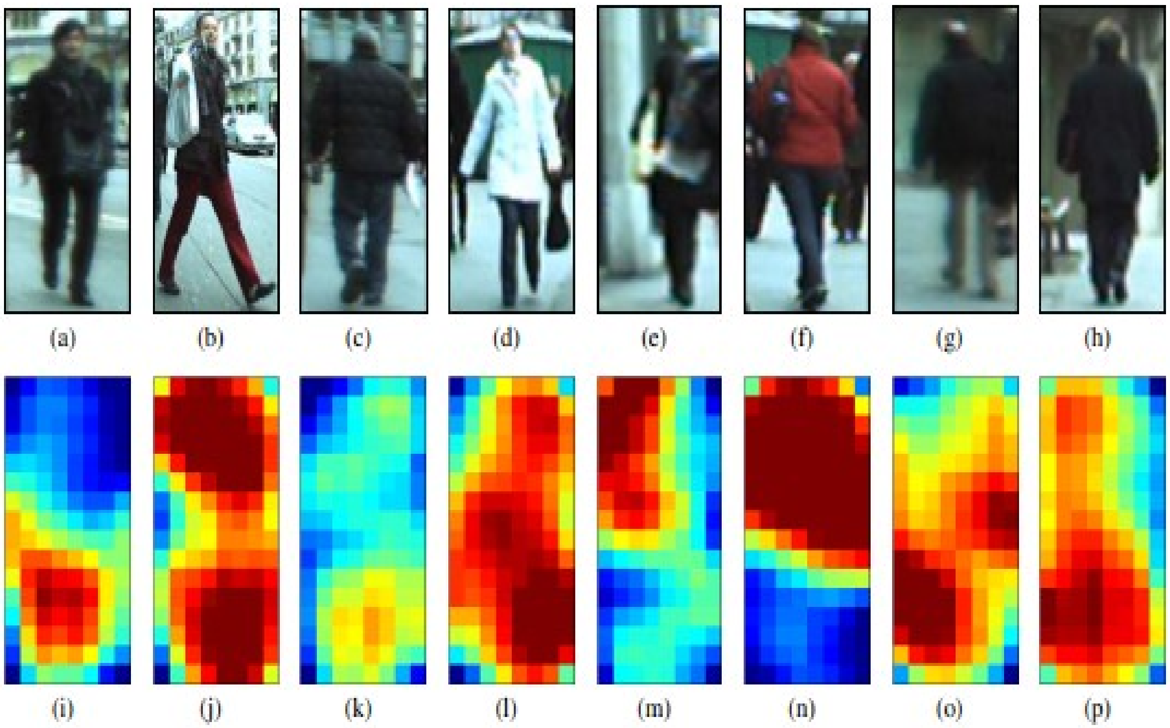


Hi **and** ***Hj*** **are normalized cumulative Histograms wrt. to each observation**

K. Chen, C.-C. Lai, Y.-P. Hung, C.-S. Chen, "An Adaptive Learning Method for Target Tracking across Multiple Cameras," CVPR 2008

Partial Least Squares

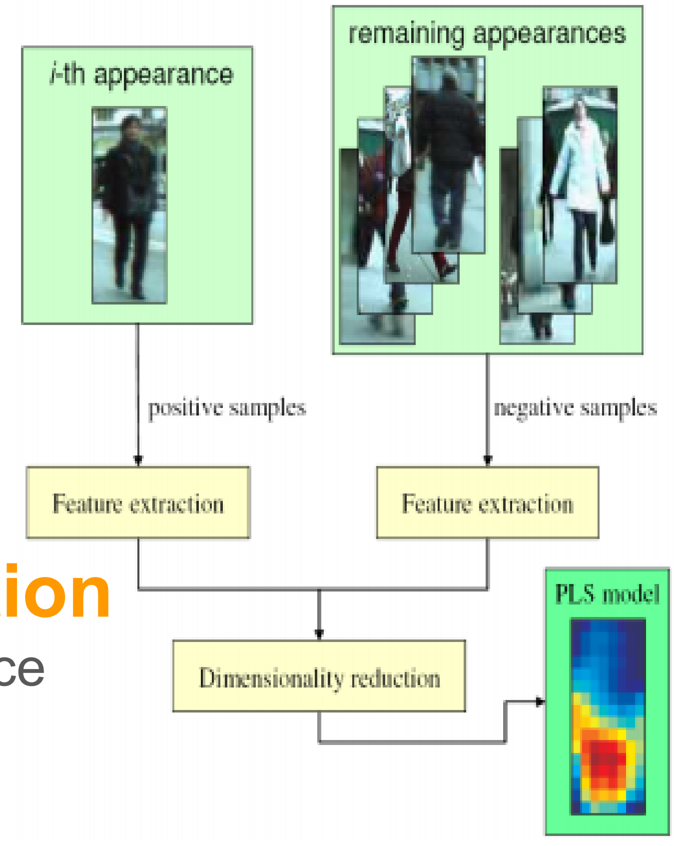
- **Learning Appearance-based Models**



W. R. Schwartz, L. S. Davis, "Learning discriminative appearance-based models using partial least squares "

Partial Least Squares

- **Feature Extraction**
  - Co-occurrence matrices
  - HOG
  - Color histograms



- **PLS for Dimension Reduction**
  - Mine Label Related Latent Subspace

W. R. Schwartz, L. S. Davis, "Learning discriminative appearance-based models using partial least squares " , Proc. Brazilian Symp. Computer Graphics and Image Processing, 2009

Feature Selection

Boosting Colour Bits and the following filter Responses (Gabor & Schmid)

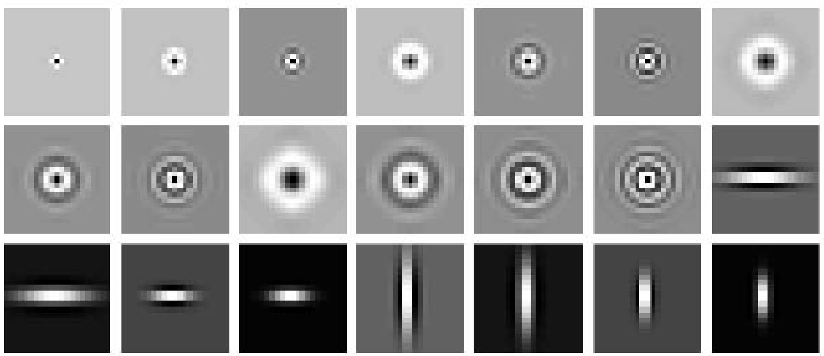
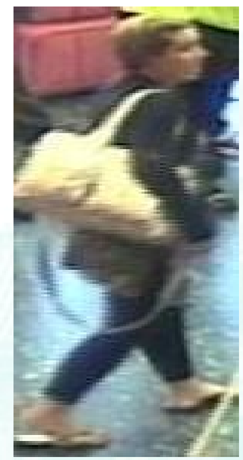
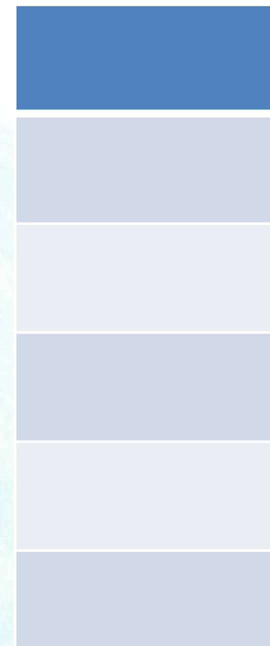


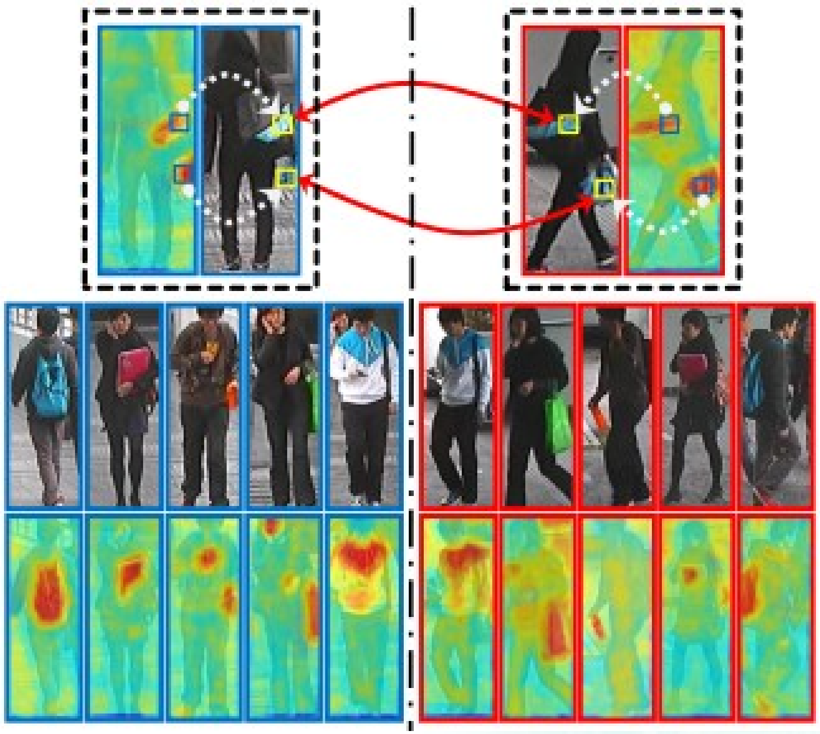
Fig. 3. The filters used in the model to describe texture. (a) Rotationally symmetric Schmid filters. (b) Horizontal and vertical Gabor filters.





Douglas Gray and Hai Tao, "Viewpoint Invariant Pedestrian Recognition with an Ensemble of Localized Features," ECCV 2008

Find Similar Patch: Salience Learning



Salience is especially designed for human matching

Human salience is incorporated in patch matching to find reliable and discriminative matched patches

Rui Zhao,Wanli Ouyang and Xiaogang Wang，"Unsupervised Salience Learning for Person Re-identification",CVPR 2013

Find Similar Patch: Salience Learning

- **Adjacency Constrained Search**
  — Do a k-nearest neighbor search for each patch
- **Unsupervised Salience Learning**
  — K-Nearest Neighbor Salience
  — One-class SVM Salience

Rui Zhao,Wanli Ouyang and Xiaogang Wang，"Unsupervised Salience Learning for Person Re-identification",CVPR 2013

Find Similar Patch: Salience Learning

- **Bi-directional Weighted Matching**
  - — Matching between a pair of images
  - — Searching for the best matched image in the gallery

Robust to viewpoint change, pose variation and articulation.

Rui Zhao,Wanli Ouyang and Xiaogang Wang，"Unsupervised Salience Learning for Person Re-identification",CVPR 2013

Measure The Differences Between Patches

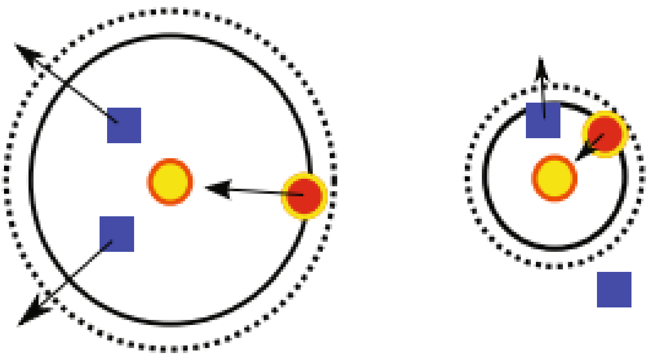
Using Metric Learning

- LMNN
- LMNN-R
- RDC(Relative Distance Comparison)
- ....

The goal is to learn a Mahalanobis metric

$$d(x, y) = (x - y)^T \mathbf{M} (x - y)$$

Pairwise Metric: LMNN (Large Margin Nearest Neighbor)



$$\mathcal{D}_{\mathbf{M}}(\boldsymbol{x}_i, \boldsymbol{x}_j) = (\boldsymbol{x}_i - \boldsymbol{x}_j)^{\top}\mathbf{M}(\boldsymbol{x}_i - \boldsymbol{x}_j), \quad \mathbf{M} = \mathbf{L}^{\top}\mathbf{L}$$

$$\mathcal{D}_{\mathbf{L}}(\boldsymbol{x}_i, \boldsymbol{x}_j) = \|\mathbf{L}(\boldsymbol{x}_i - \boldsymbol{x}_j)\|^2.$$

$$\varepsilon_{pull}(\mathbf{M}) = \sum_{i,j \rightsquigarrow i}^{N} \mathcal{D}_{\mathbf{M}}(\boldsymbol{x}_i, \boldsymbol{x}_j),$$

$$\varepsilon_{push}(\mathbf{M}) = \sum_{i,j \rightsquigarrow i} \sum_{k=1}^{N} (1 - y_{ik})\left[1 + \mathcal{D}_{\mathbf{M}}(\boldsymbol{x}_i, \boldsymbol{x}_j) - \mathcal{D}_{\mathbf{M}}(\boldsymbol{x}_i, \boldsymbol{x}_k)\right]_+$$

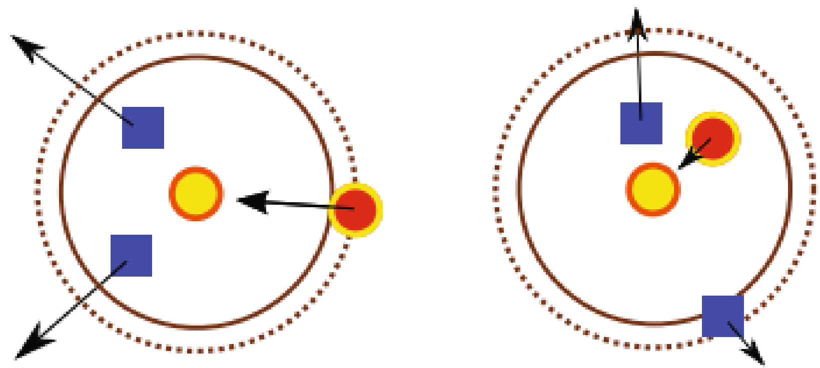
$$\varepsilon_{\mathrm{LMNN}}(\mathbf{M}) = (1-\mu) \sum_{i,j \rightsquigarrow i} \mathcal{D}_{\mathbf{M}}(\boldsymbol{x}_i, \boldsymbol{x}_j) + \mu \sum_{i,j \rightsquigarrow i} \sum_{k=1}^{N} (1 - y_{ik})\left[1 + \mathcal{D}_{\mathbf{M}}(\boldsymbol{x}_i, \boldsymbol{x}_j) - \mathcal{D}_{\mathbf{M}}(\boldsymbol{x}_i, \boldsymbol{x}_k)\right]_+ .$$

K. Q.Weinberger and L. K. Saul, "Distance Metric Learning for Large Margin Nearest Neighbor Classification," JMLR 2009.

LMNN-R

Use a metric learning framework to obtain a robust metric for large margin nearest neighbor classification with rejection

$$R = \frac{1}{NK} \sum_{m,l \rightsquigarrow m} \mathcal{D}_{\mathbf{M}}(\boldsymbol{x}_m, \boldsymbol{x}_l)$$



$$\varepsilon_{\mathbf{LMNN\text{-}R}}(\mathbf{M}) = (1-\mu)\varepsilon_{pull}(\mathbf{M}) + \mu\varepsilon^{*}_{push}(\mathbf{M})$$

$$\varepsilon^{*}_{push}(\mathbf{M}) = \sum_{i=1}^{N}\sum_{k=1}^{N}(1-y_{ik})\left[1 + \frac{1}{NK}\left(\sum_{m,l \rightsquigarrow m} \mathcal{D}_{\mathbf{M}}(\boldsymbol{x}_m, \boldsymbol{x}_l)\right) - \mathcal{D}_{\mathbf{M}}(\boldsymbol{x}_i, \boldsymbol{x}_k)\right]_{+}$$

M. Dikmen, E. Akbas, T. S. Huang, N. Ahuja, Pedestrian recognition with a learned metric, in: Asian Conference in Computer Vision (ACCV), 2010.

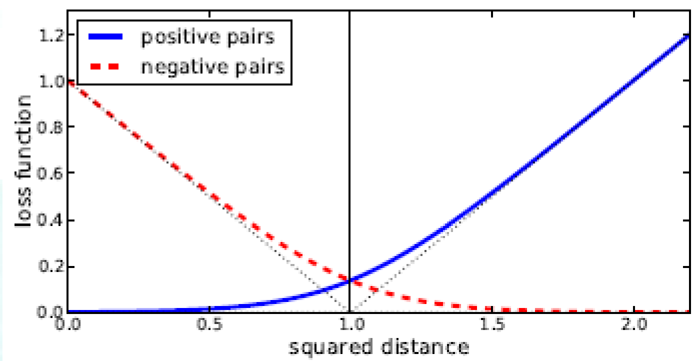
Pairwise Constrained Component Analysis

A Learned Bayesian Metric

$$\min_{L} E(L) = \sum_{n=1}^{c} \ell_\beta \left(y_n (D_L^2(\mathrm{x}_{i_n}, \mathrm{x}_{j_n}) - 1) \right)$$

$$D_L^2(\mathrm{x}, \mathrm{y}) = \|L(\mathrm{x} - \mathrm{y})\|_2^2$$

$$\ell_\beta(x) = \frac{1}{\beta} \log(1 + e^{\beta x})$$



Loss function for pos. (solid) and neg. (dashed) pairs.

A. Mignon and F. Jurie, "PCCA: A New Approach for Distance Learning from Sparse Pairwise Constraints," CVPR 2012

Pairwise Constrained Component Analysis

$$\mathcal{L}(\mathbf{L}) = \frac{1}{|\mathcal{S}|} \sum_{\underset{\uparrow}{(i,j)\in\mathcal{S}}} ||\mathbf{L}(x_i - x_j)||^2 - \frac{1}{|\mathcal{D}|} \sum_{\underset{\uparrow}{(i,j)\in\mathcal{D}}} ||\mathbf{L}(x_i - x_j)||^2 .$$

Similar pair

dissimilar pair

$$\begin{aligned} &\min \ \mathcal{L}(\mathrm{M}) \\ &s.t.\ \mathrm{M} \succeq 0\,,\ \ \mathbf{L}\boldsymbol{\Sigma}_S\mathbf{L}^\top = \mathbf{I}\,,\ \ \mathbf{L}\boldsymbol{\Sigma}_D\mathbf{L}^\top = \mathbf{I} \end{aligned}$$

M. Hirzer, P. M. Roth, M. KOstinger, and H. Bischof, "Relaxed Pairwise Learned Metric for Person Re-identification," ECCV 2012

Attribute Metric Learning

- **Model and Fusion**
  - Build on SDALF
  - integrate attribute-based distance

$$
\begin{aligned}
d(I_p, I_q) &= (1-\beta_{ATTR}) \cdot d_{SDALF}(SDALF(I_p), SDALF(I_q)) \\
&+ \beta_{ATTR} \cdot d_{ATTR}(ATTR(I_p), ATTR(I_q)).
\end{aligned}
$$

- **Attribute Metric Learning**
  - Define the distance between attribute profiles $A(\mathbf{x})$

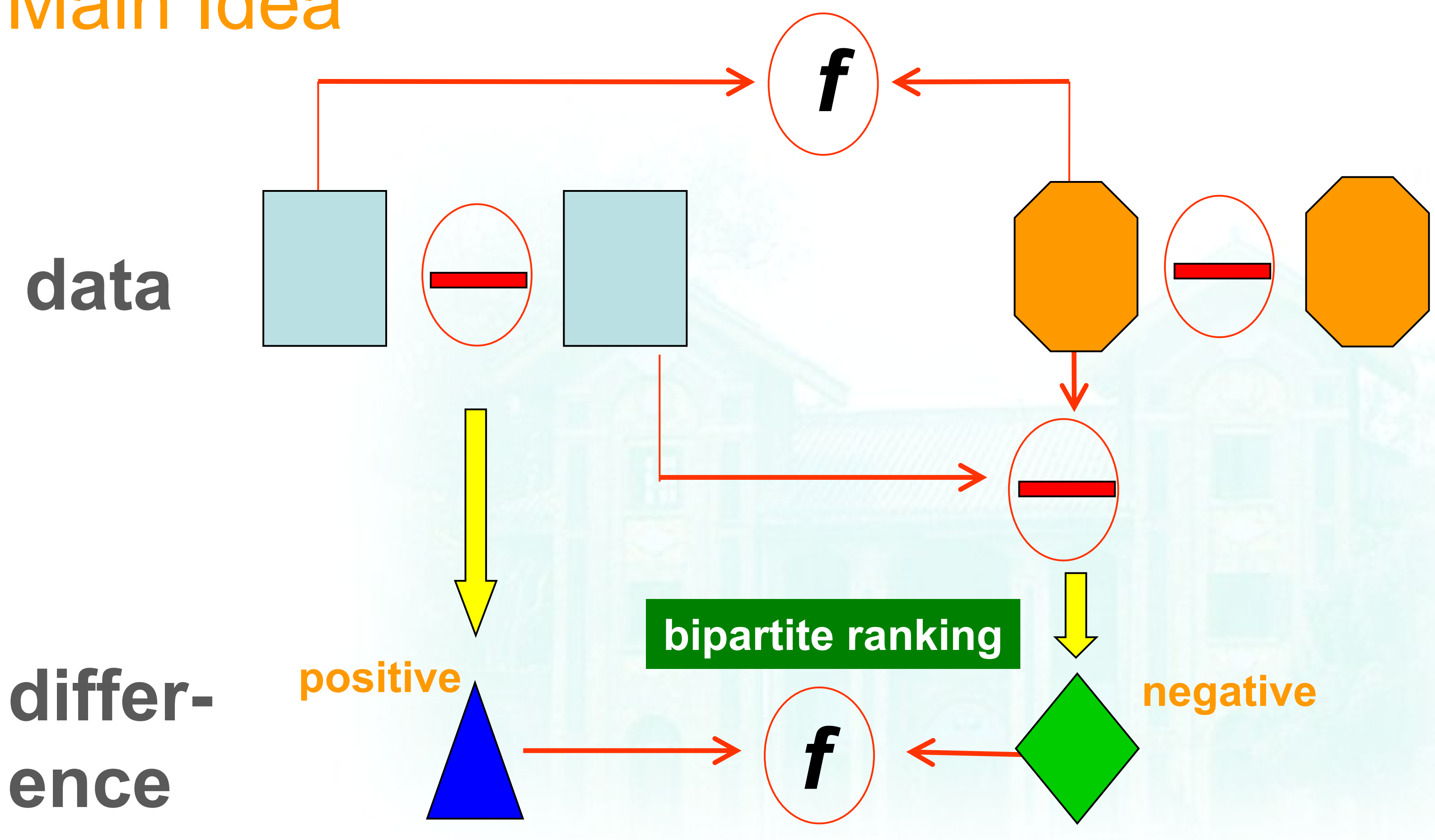
$$
d_{ATTR}(I_p, I_q; \Lambda) = (A(\mathbf{x}_p) - A(\mathbf{x}_q))^T \Lambda (A(\mathbf{x}_p) - A(\mathbf{x}_q))
$$

$$
\begin{aligned}
&\min_{\Lambda} \mathcal{KLD}\left(p(\mathbf{x}; \Lambda_0) || p(\mathbf{x}; \Lambda)\right) \; s.t. \\
&d_A(\mathbf{x}_i, \mathbf{x}_j) \le u \quad if \quad (i,j) \in S, \\
&d_A(\mathbf{x}_i, \mathbf{x}_j) \ge l \quad if \quad (i,j) \in D,
\end{aligned}
$$

R. Layne, T. M. Hospedales, S. Gong, "Towards person identification and re-identification with attributes", ECCV Workshop, 2012

Triple based Learning: Bipartite Ranking

Main Idea



Triple based Learning: Bipartite Ranking

Preliminary work: RankSVM

$$\min_{\mathbf{w}} \frac{1}{2}\|\mathbf{w}\|^2 + \beta \sum_{i=1}^{|\mathbb{O}|} \max\left(0, 1 - \mathbf{w}^T\left(\mathbf{x}_i^p - \mathbf{x}_i^n\right)\right)^2$$

$\mathbf{x}_i^p$ Positive Data Difference

$\mathbf{x}_i^n$ Related Negative Data Difference

(1) Maximising the margin between difference souces of data difference
(2) Quantifying first-order feature vectors
(3) Sensitive to parameter

Triple based Learning: Bipartite Ranking

Relative Distance Comparison

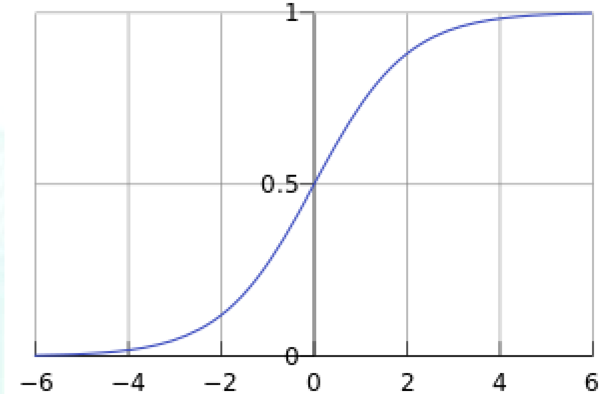
difference vector

$$f(\mathbf{x}) = \mathbf{x}^T \mathbf{M} \mathbf{x}, \ \mathbf{M} \succeq 0$$

$$f(\mathbf{x}_i^p) < f(\mathbf{x}_i^n)$$

positive difference vector

negative Difference vector



$$C_f(\mathbf{x}_i^p, \mathbf{x}_i^n) = \left(1 + \exp\left\{f(\mathbf{x}_i^p) - f(\mathbf{x}_i^n)\right\}\right)^{-1}$$

soft margin measure

Triple based Learning: Bipartite Ranking

RDC: Modelling & Characteristic

$$\min_f r(f, \mathbb{O}), \;\; r(f, \mathbb{O}) = -\log\left(\prod\nolimits_{\mathbb{O}_i} C_f(\mathbf{x}_i^p, \mathbf{x}_i^n)\right)$$

$$\Downarrow$$

$$\min_{\mathbf{W}} r(\mathbf{W}, \mathbb{O}), \;\; s.t. \;\; \mathbf{w}_i^T \mathbf{w}_j = 0, \;\; \forall i \neq j$$

$$r(\mathbf{W}, \mathbb{O}) = \sum\nolimits_{\mathbb{O}_i} \log(1 + \exp\{||\mathbf{W}^T \mathbf{x}_i^p||^2 - ||\mathbf{W}^T \mathbf{x}_i^n||^2\})$$

(1) Mainly concerning the relative distance comparison
(2) Quantifying second-order feature vectors
(3) no improtance weight
(4) low-rank

Wei-Shi Zheng et al.,
"Re-identification by Probabilistic Relative Distance Comparison"
IEEE Trans. on PAMI, 2013

Triple based Learning: Bipartite Ranking

Entry-wise Absolute Difference Vector

$$\mathbf{x} = d(\mathbf{z}, \mathbf{z}') = \left|\mathbf{z} - \mathbf{z}'\right|, \ \mathbf{x}(k) = \left|\mathbf{z}(k) - \mathbf{z}'(k)\right|$$

$$f(|\mathbf{x}_{ij}|) = |\mathbf{z}_i - \mathbf{z}_j|^T \mathbf{M} |\mathbf{z}_i - \mathbf{z}_j| = ||\mathbf{W}^T |\mathbf{x}_{ij}| \ ||^2$$

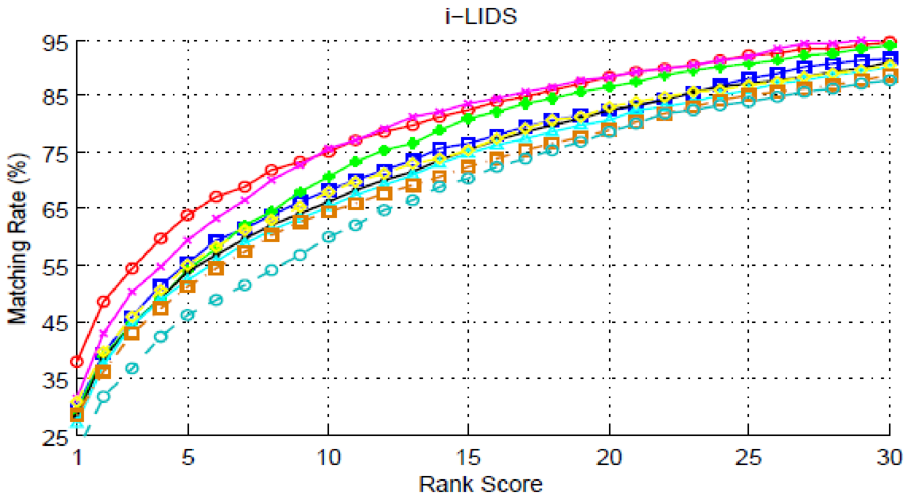
$$\left|\left||\mathbf{x}_{ij}| - |\mathbf{x}_{ij'}|\right|\right| \leq \left|\left|\mathbf{x}_{ij} - \mathbf{x}_{ij'}\right|\right|$$

$$upper(\left|\left|\mathbf{W}^T(|\mathbf{x}_{ij}| - |\mathbf{x}_{ij'}|)\right|\right|) \leq upper(\left|\left|\mathbf{W}^T(\mathbf{x}_{ij} - \mathbf{x}_{ij'})\right|\right|)$$

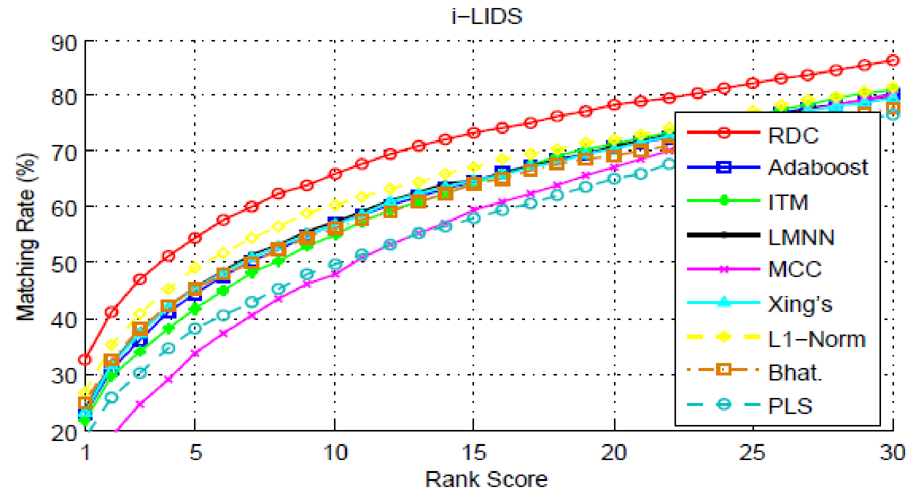
Relative Distance Learning can be more robust in the absolute distance space

Triple based Learning: Bipartite Ranking

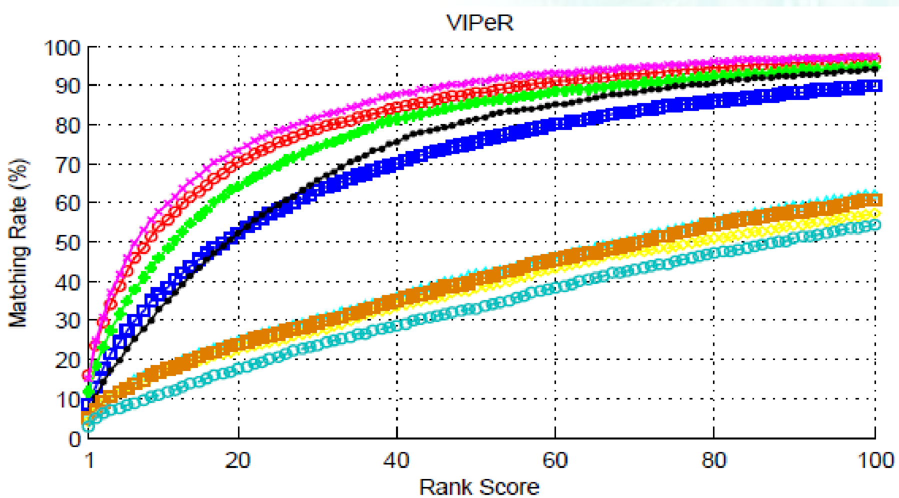
Re-identification (i-LIDS&VIPeR)



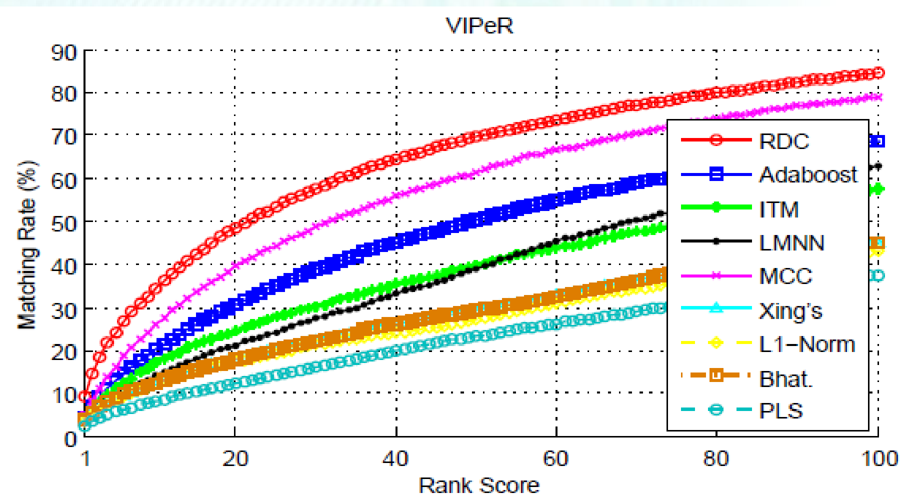
(a) $p = 50$



(b) $p = 80$



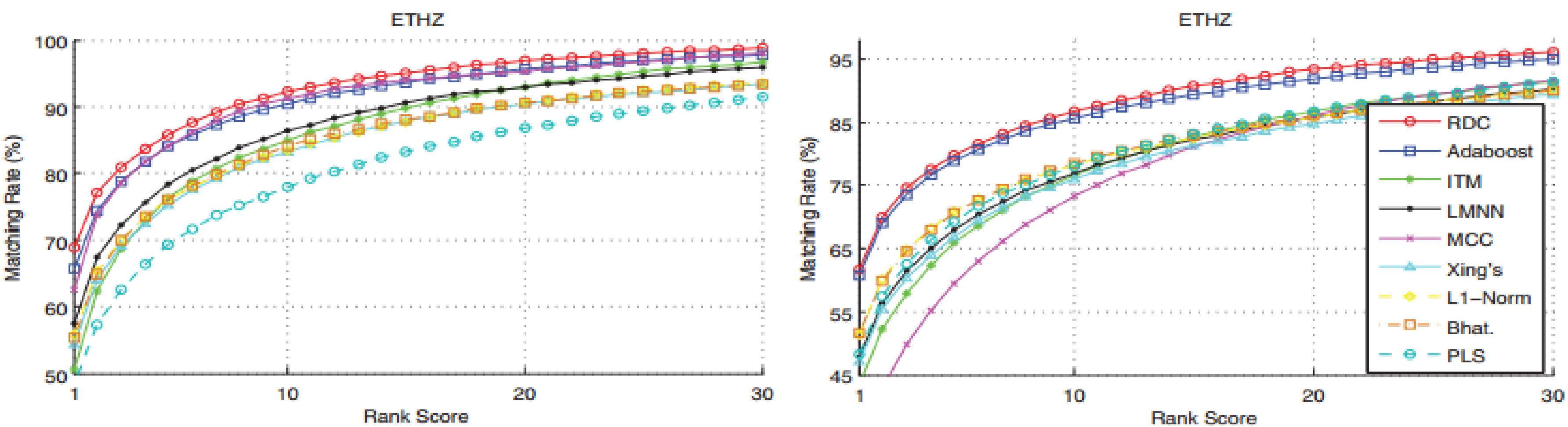
(a) $p = 316$



(b) $p = 532$

Triple based Learning: Bipartite Ranking

Re-identification (ETHZ)



(a) $p = 70$ (b) $p = 120$

Ensemble Metric Learning

Ensemble RDC: Motivation

- RDC: Large space complexity

$$O(q\cdot((\frac{1}{L}-\frac{1}{L^2})\cdot N^3+(\frac{1}{L}-1)\cdot N^2)) \Rightarrow O(q\cdot((\frac{b^2}{L}-\frac{b}{L^2})\cdot N^3+(\frac{b}{L}-b^2)\cdot N^2))$$

- RDC: Trapped in locally optimal solution

Ensemble RDC: Modelling

- Randomly dividing the set into small groups
- Learning a set of weak RDC models
- Boosting them

Wei-Shi Zheng et al.,
"Re-identification by Probabilistic Relative Distance Comparison"
IEEE Trans. on PAMI, 2013

Ensemble Metric Learning

Using Ensemble Learning to boost weak RDC metric

$$f_S(\mathbf{x}) = \sum_{i=1}^{H} \beta_i \cdot f_{w,i}(\mathbf{x})$$

Algorithm 2: Algorithm of Ensemble RDC

Data: Pairwise relevant difference vector set $\mathbb{O}$, a set of weak RDC models $\{f_{w,i}\}_{i=1}^{H}$, Initial distribution D

begin

  $D_1 \longleftarrow D$;

  for $t = 1, \cdots, T$ **do**

    Select the best weak RDC mdoel f_{w,k_t} by Eq. (22);

    Compute the weight α_t by Eq. (24);

    Update the distribution D_{t+1} by Eq. (23).

  end

end

Output: $f_s(\mathbf{x}) = \sum_{t=1}^{T} \alpha_t \cdot f_{w,k_t}(\mathbf{x}) = \sum_{i=1}^{H} \beta_i \cdot f_{w,i}(\mathbf{x})$

KISS Metric Learning

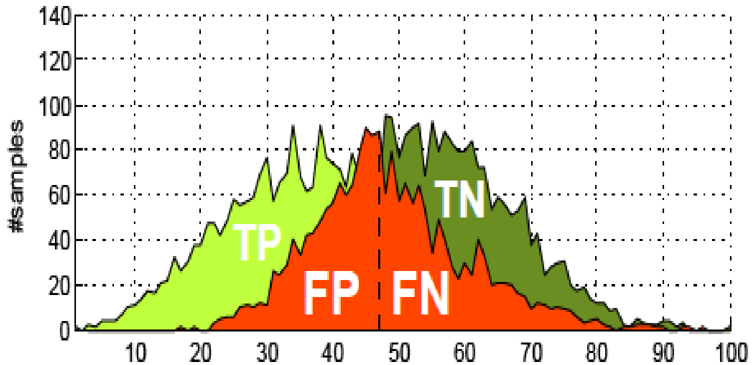
A Bayesian Metric

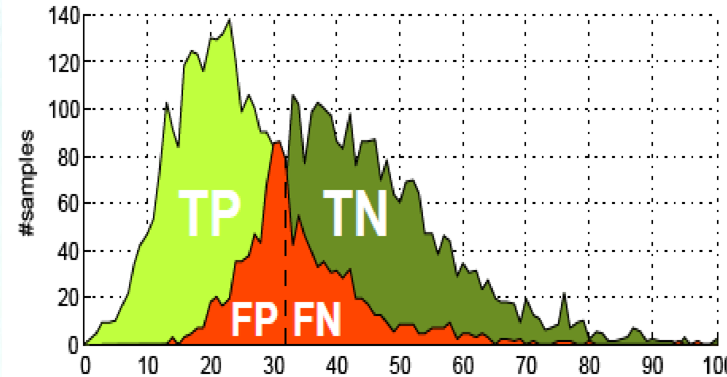
$$\delta(\mathbf{x}_{ij}) = \log\left(\frac{p(\mathbf{x}_{ij}|H_0)}{p(\mathbf{x}_{ij}|H_1)}\right) = \log\left(\frac{f(\mathbf{x}_{ij}|\theta_0)}{f(\mathbf{x}_{ij}|\theta_1)}\right) \Rightarrow \delta(\mathbf{x}_{ij}) = \log\left(\frac{\frac{1}{\sqrt{2\pi|\Sigma_{y_{ij}=0}|}}\exp(-1/2\ \mathbf{x}_{ij}^T\ \Sigma_{y_{ij}=0}^{-1}\ \mathbf{x}_{ij})}{\frac{1}{\sqrt{2\pi|\Sigma_{y_{ij}=1}|}}\exp(-1/2\ \mathbf{x}_{ij}^T\ \Sigma_{y_{ij}=1}^{-1}\ \mathbf{x}_{ij})}\right)$$

$$\Sigma_{y_{ij}=1} = \sum_{y_{ij}=1}(\mathbf{x}_i - \mathbf{x}_j)(\mathbf{x}_i - \mathbf{x}_j)^T$$

$$\Sigma_{y_{ij}=0} = \sum_{y_{ij}=0}(\mathbf{x}_i - \mathbf{x}_j)(\mathbf{x}_i - \mathbf{x}_j)^T$$

$$\delta(\mathbf{x}_{ij}) = \mathbf{x}_{ij}^T(\Sigma_{y_{ij}=1}^{-1} - \Sigma_{y_{ij}=0}^{-1})\mathbf{x}_{ij}.$$



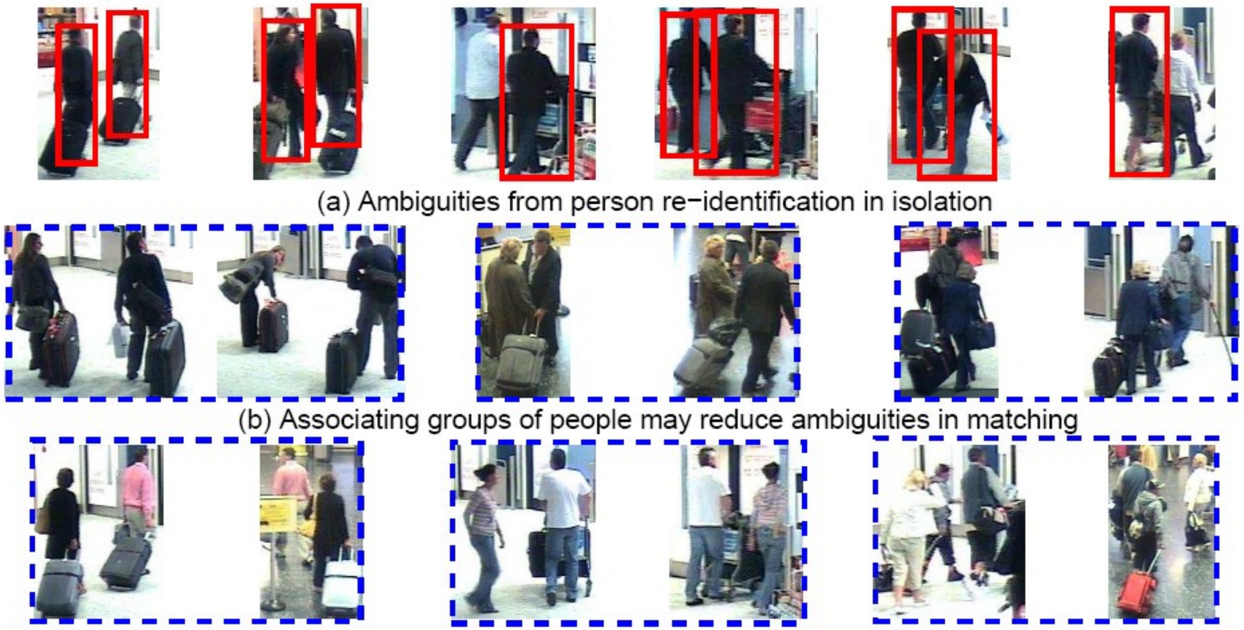


M. Kostinger, M. Hirzer, P. Wohlhart, P. M. Roth, H. Bischof, "Large Scale Metric Learning from Equivalence Constraints," CVPR 2012.

Context-aware Person Re-identificaiton

Associating Groups of People

- Associating Group of People vs. Individuals



(a) Ambiguities from person re−identification in isolation

(b) Associating groups of people may reduce ambiguities in matching

(c) Difficult examples of associating groups of people

- **Group Information Plays as Context to help Individual Identification**

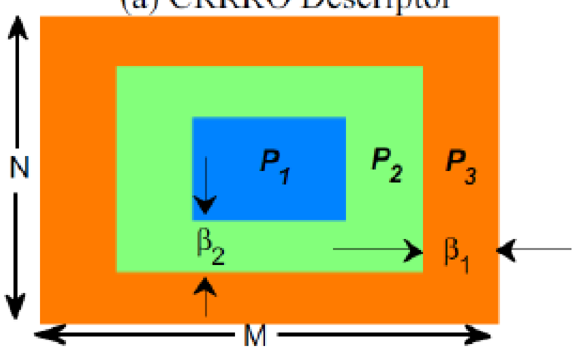
Wei-Shi Zheng et al., "Associating Groups of People," BMVC 2009.

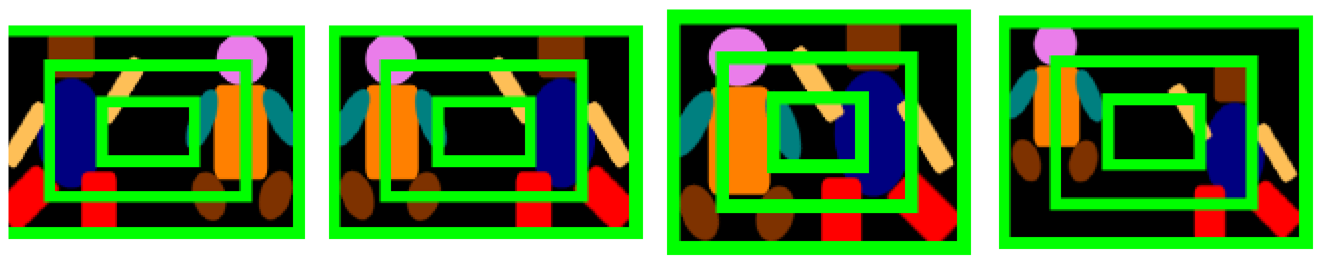
Associating Groups of People

Modelling: Group Descriptor

- **A rectangle ring descriptor**: rotation invariant

(a) CRRRO Descriptor





intra ratio-occurrence map $\mathbf{H}_i$

$$\mathbf{H}_i(a,b) = \frac{\mathbf{h}_i(a)}{\mathbf{h}_i(a) + \mathbf{h}_i(b) + \varepsilon}$$

$$\mathbf{T}_r^i = \{\mathbf{H}_i, \mathbf{S}_i, \mathbf{G}_i\}$$

inter ratio-occurrence maps $\mathbf{S}_i$ and $\mathbf{G}_i$

$$\mathbf{G}_i(a,b) = \frac{\mathbf{g}_i(a)}{\mathbf{g}_i(a) + \mathbf{h}_i(b) + \varepsilon}, \quad \mathbf{S}_i(a,b) = \frac{\mathbf{s}_i(a)}{\mathbf{s}_i(a) + \mathbf{h}_i(b) + \varepsilon}, \quad \mathbf{g}_i = \sum_{j=1}^{i-1} \mathbf{h}_j, \quad \mathbf{s}_i = \sum_{j=i+1}^{\ell} \mathbf{h}_j$$

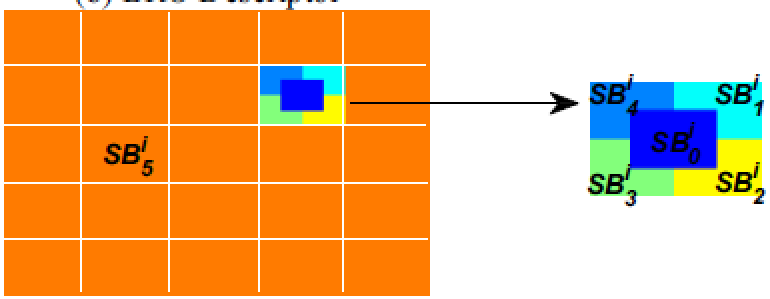
Associating Groups of People

Modelling: Group Descriptor

- **A block based occurrence descriptor**:

  for large non-center-rotational changes in people's positions

(b) BRO Descriptor



$$\mathbf{T}^i_b = \{\mathbf{H}^i_j\}_{j=0}^{4\gamma+1} \bigcup \{\mathbf{O}^i_j\}_{j=1}^{2}$$

intra ratio-occurrence map $\mathbf{H}^i_j$ between visual words in each block region SB^i_j

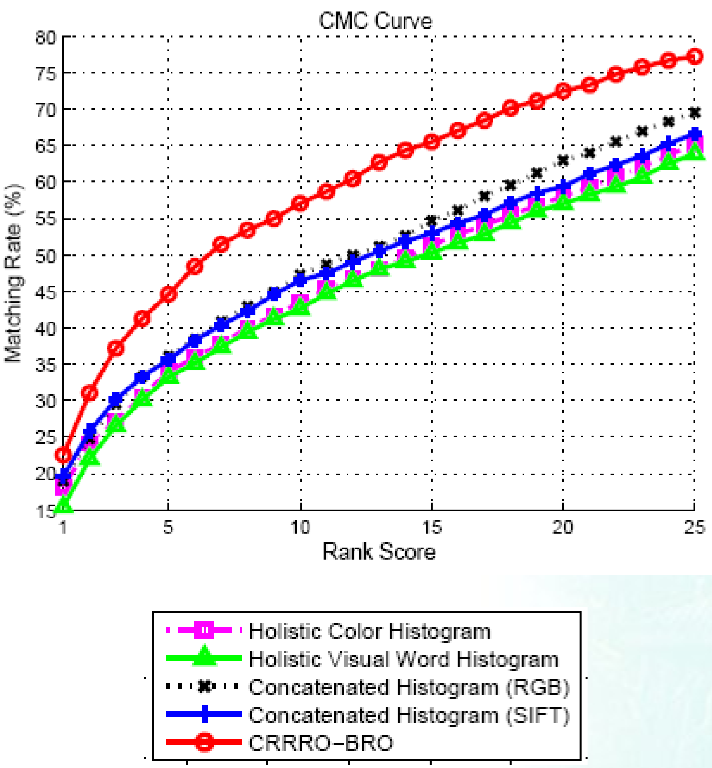
$$\mathbf{H}_i(a,b) = \frac{\mathbf{h}_i(a)}{\mathbf{h}_i(a)+\mathbf{h}_i(b)+\varepsilon}$$

inter ratio-occurrence maps $\mathbf{O}^i_j$ between block B_i and its complementary region $SB^i_{4\gamma+1}$

$$\mathbf{O}^i_1(a,b) = \frac{\mathbf{t}_i(a)}{\mathbf{t}_i(a)+\mathbf{z}_i(b)+\varepsilon} \qquad \mathbf{O}^i_2(a,b) = \frac{\mathbf{z}_i(a)}{\mathbf{z}_i(a)+\mathbf{t}_i(b)+\varepsilon}$$

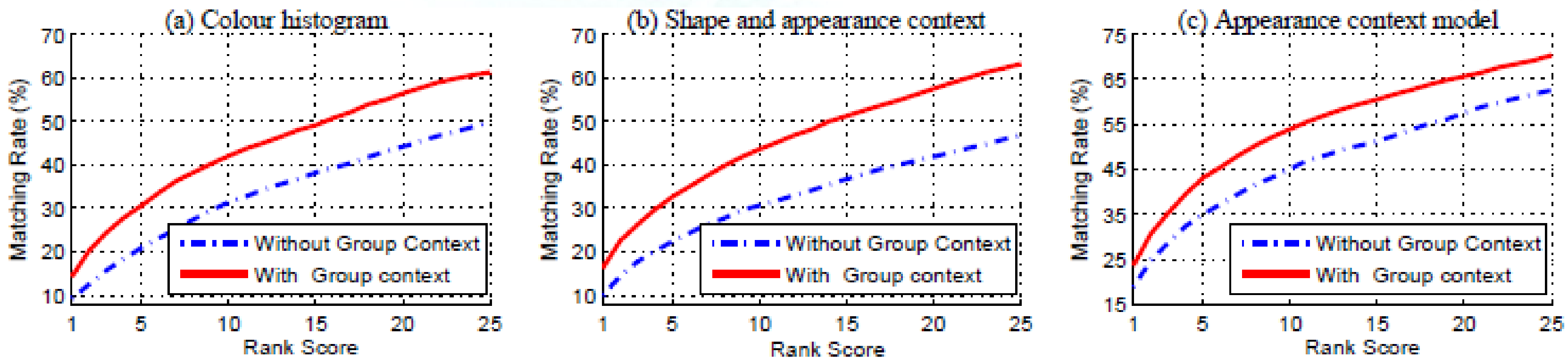
where $\mathbf{z}_i$ and $\mathbf{t}_i$ are the histograms of visual words of block B_i and image region $SB^i_{4\gamma+1}$

Associating Groups of People



Associating Groups of People

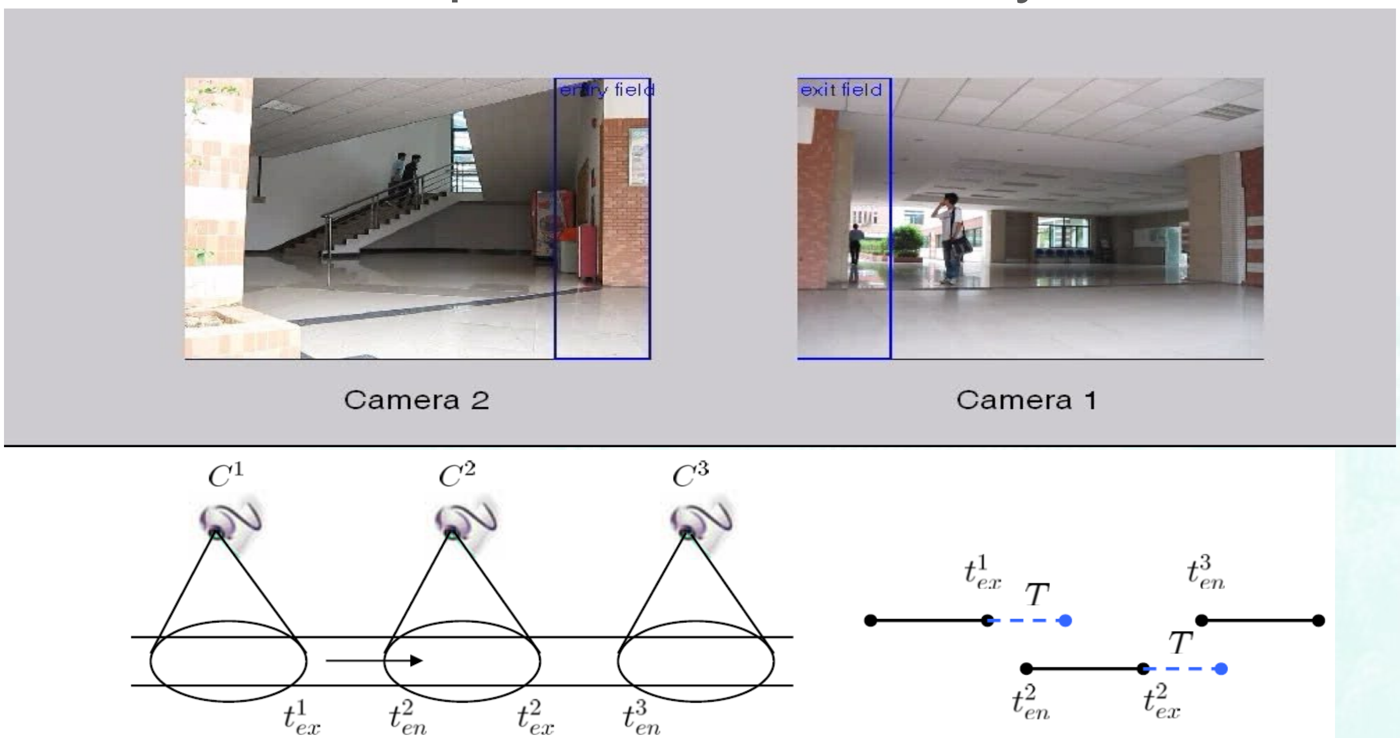
Person Re-identification using Group Context



Context-Aware Person Re-identification

A Spatial Temporal Mode

– We further incoporate the time delay information



G. Lian, J. H. Lai, Wei-Shi Zheng, "Spatial-temporal Consistent Labeling of Tracked Pedestrians across Non-overlapping Camera Views," Pattern Recognition 2011.

Transfer Learning in Person Re-identification

Adaptive Metric Learning

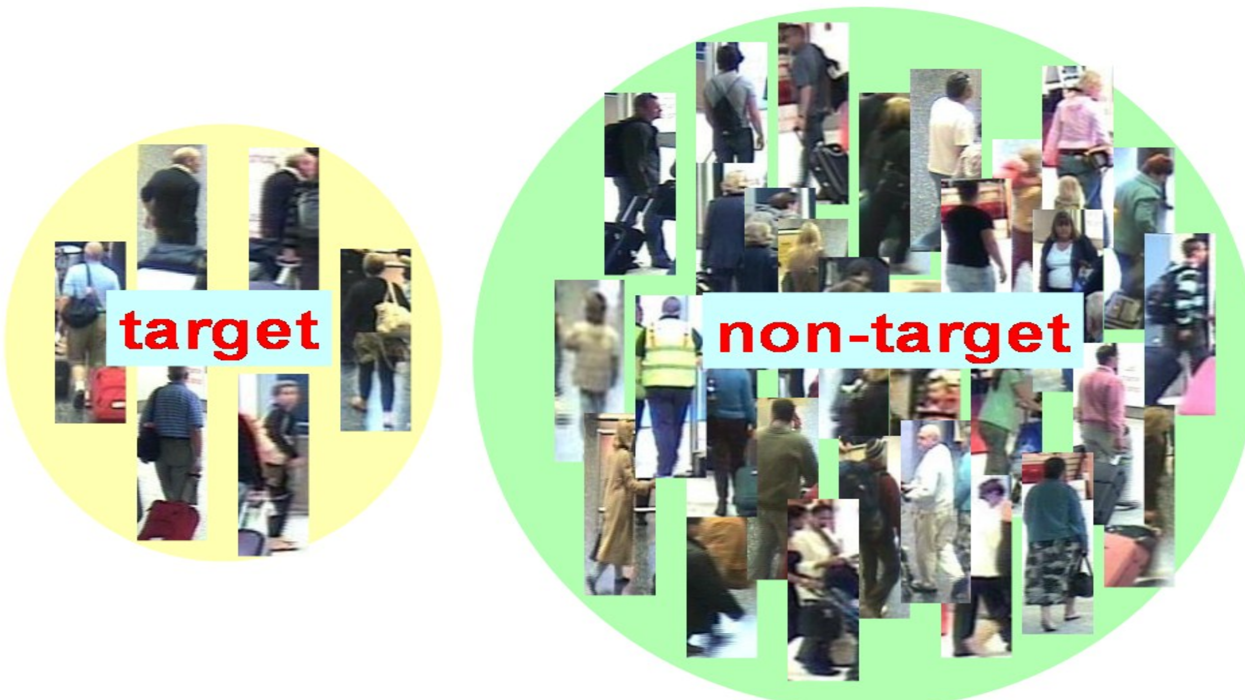
- **Assumption:**
  - Similar guys share similar neighbours
- **For a target**
  - Select similar training Samples and their neighbours
  - Obtain the corresponding pairs in the training set
  - Weight the pairs
  - Learn a metric

W. Li, R. Zhao and X. Wang, "Human Reidentification with Transferred Metric Learning," ACCV 2012

Person Verification

Target Based

Utilising Large amount of non-target unlabelled data



Wei-Shi Zheng, et al., "Transfer Re-identification: From Person to Set-based Verification", CVPR 2012.

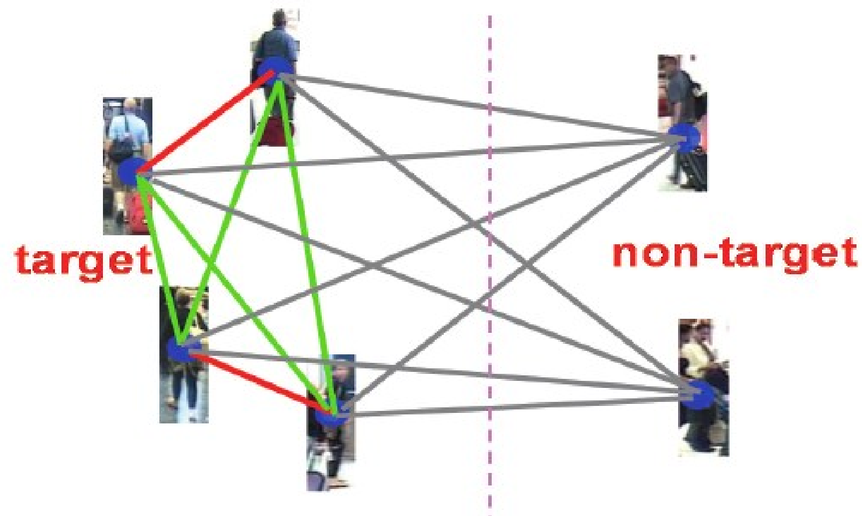
Person Verification

Modelling & Transferring

- Matching all people

  ⬇

  Matching people on the watch list against non-target people



- (1) Performing verification a small target set of people
  (2) Considering the effect of non-target ones

- Transfer RankSVM & Transfer RDC

Summary &
Some Challenges

Summary

- Person Re-identification is very important

- Spatial, Occlusion & Lighting Invariant Features are commanded

- Selection and Metric Learning are used to reduce false matching

Open Issues

- Cloths: if people change their cloths?
- Low resolution
- Occlusion
- Person Images are from Different Sources

………

■ Contact:

Wei-Shi Zheng
wszheng@ieee.org

■ Specially Thank:

Xiang Li

Q & A. Welcome!

Thank you!