Neighborhood Repulsed Metric Learning for Kinship Verification

Jiwen Lu

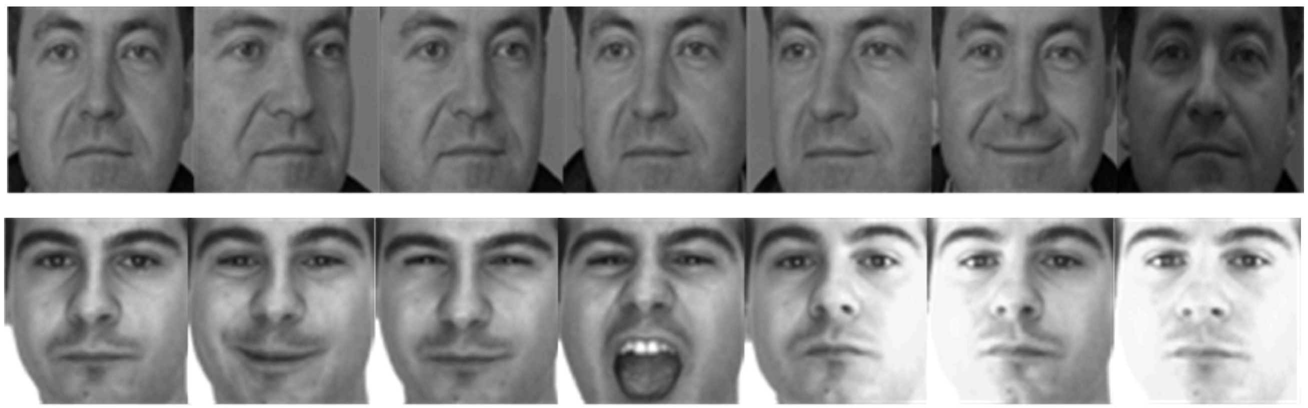
Advanced Digital Sciences Center, Singapore

Paper

Jiwen Lu, Xiuzhuang Zhou, Yap-Peng Tan, Yuanyuan Shang, Jie Zhou, Neighborhood repulsed metric learning for kinship verification, *IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI)*, vol. 36, no. 2, pp. 331-345, 2014. (Conference version appeared at CVPR2012)

Face Analysis Tasks

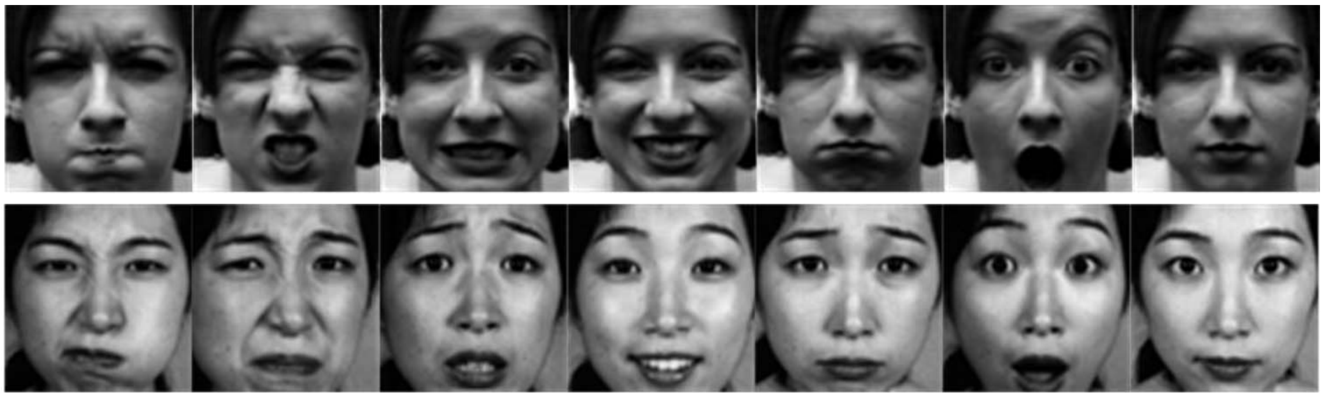
Face identification (access control/ surveillance)



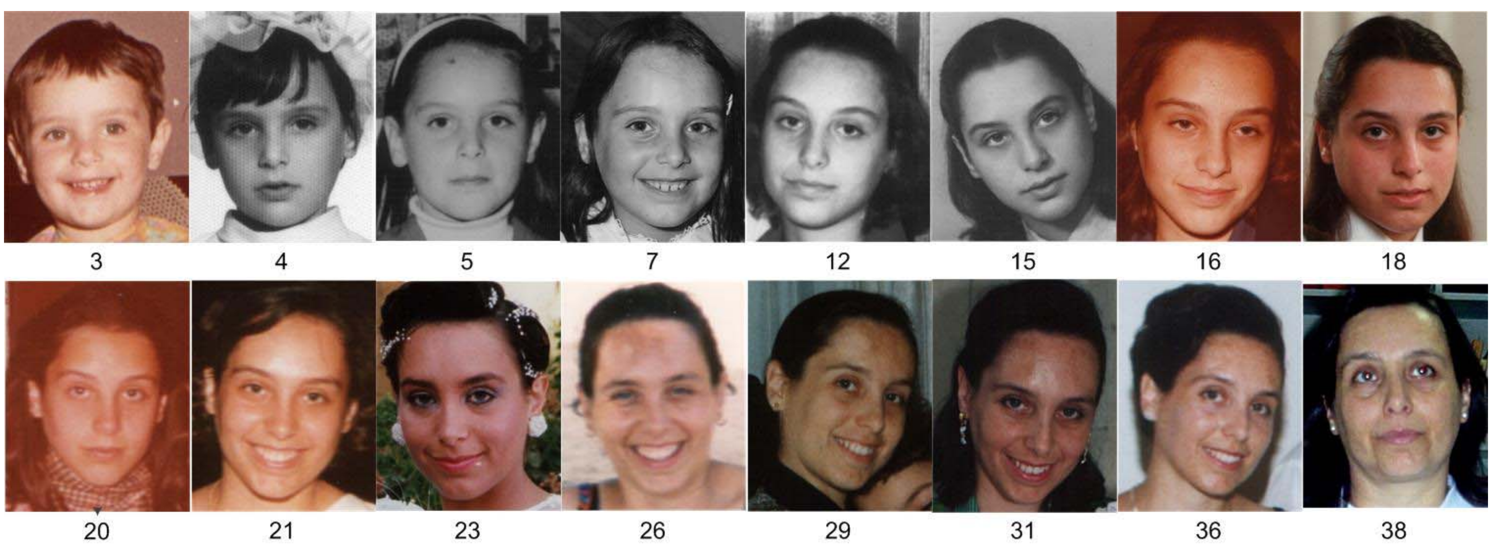
Face verification (access control/surveillance)

Face Analysis Tasks

Facial expression recognition (human-computer interaction)

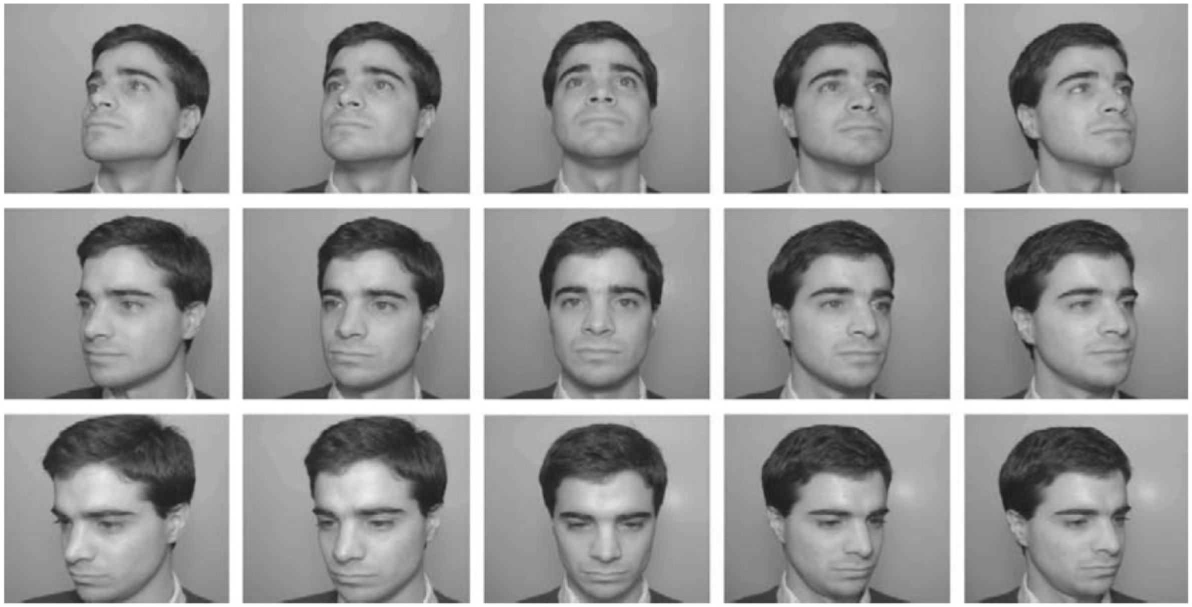


Facial age estimation (visual advertisement/social media)

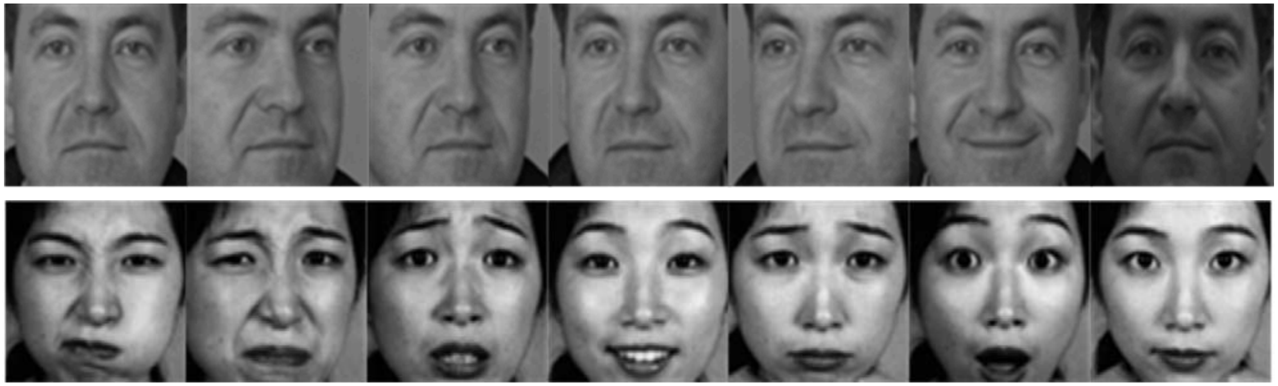


Face Analysis Tasks

Head pose estimation (human computer interaction)



Gender classification (social media analysis)



Face Analysis Tasks

Facial beauty prediction (multimedia analysis)

Face Analysis Tasks

Kinship verification (social media analysis)

Father-Son (F-S)

Father-Daughter (F-D)

Mother-Son (M-S)

Mother-Daughter (M-D)

Related Work

Local Features + SVM (Fang2010 [1])

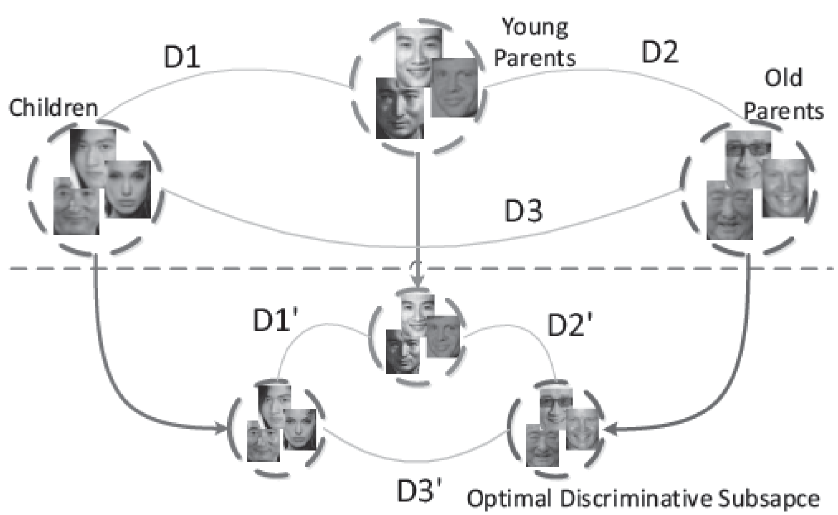
- 150 image pairs
- 50%: Caucasians
- 40%: Asians
- 7% African Americans
- 3% others;
- 40%: F-S
- 22%: F-D
- 13%: M-S
- 26%: M-D

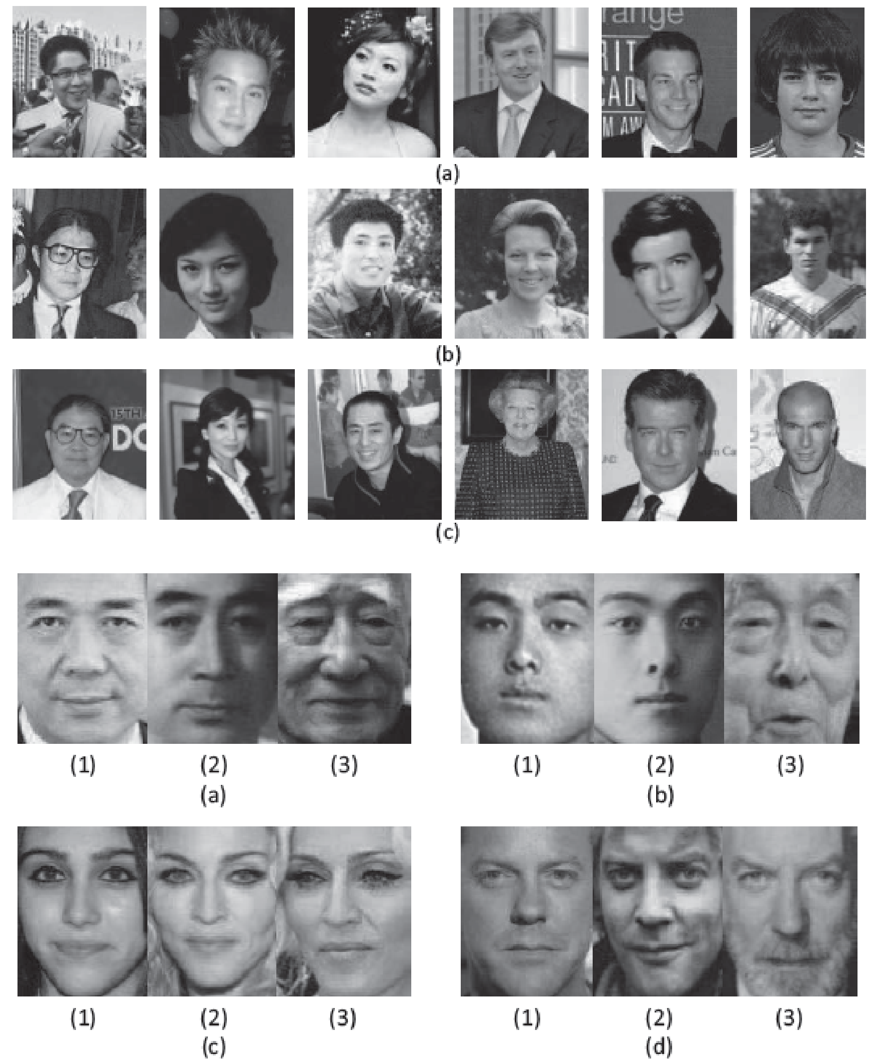
[1] Ruogu Fang, Kevin D. Tang, Noah Snavely, and Tsuhan Chen, Towards computational models of kinship verification, *ICIP*, pp. 1577-1580, 2010. (**Best Paper Award**)

Related Work

Transfer Learning (Xia2012 [2])

- 90 groups
- 3 images in each group

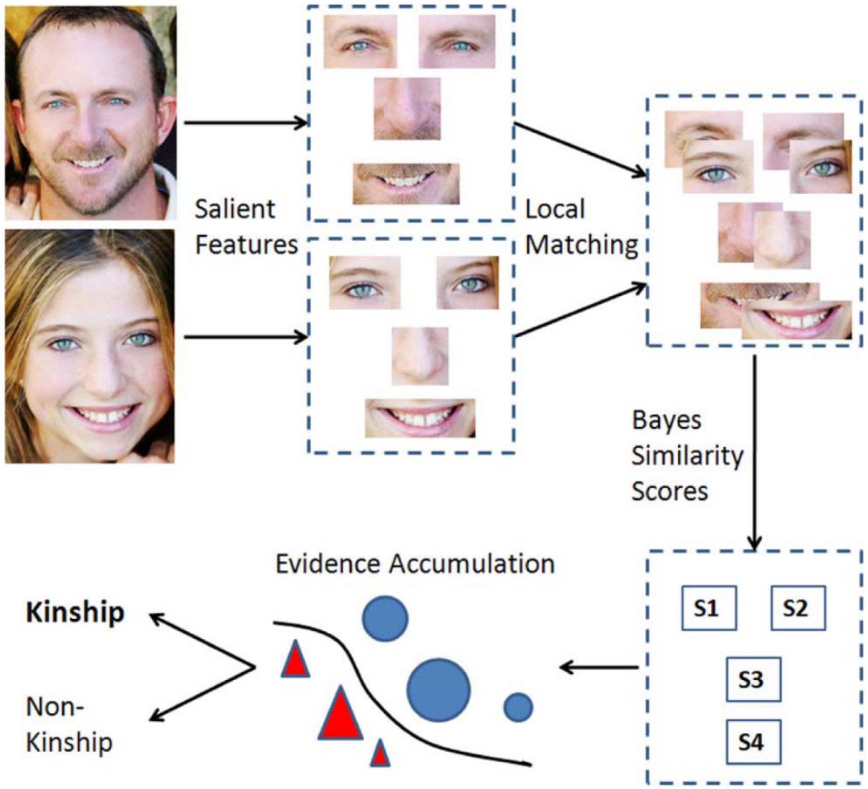


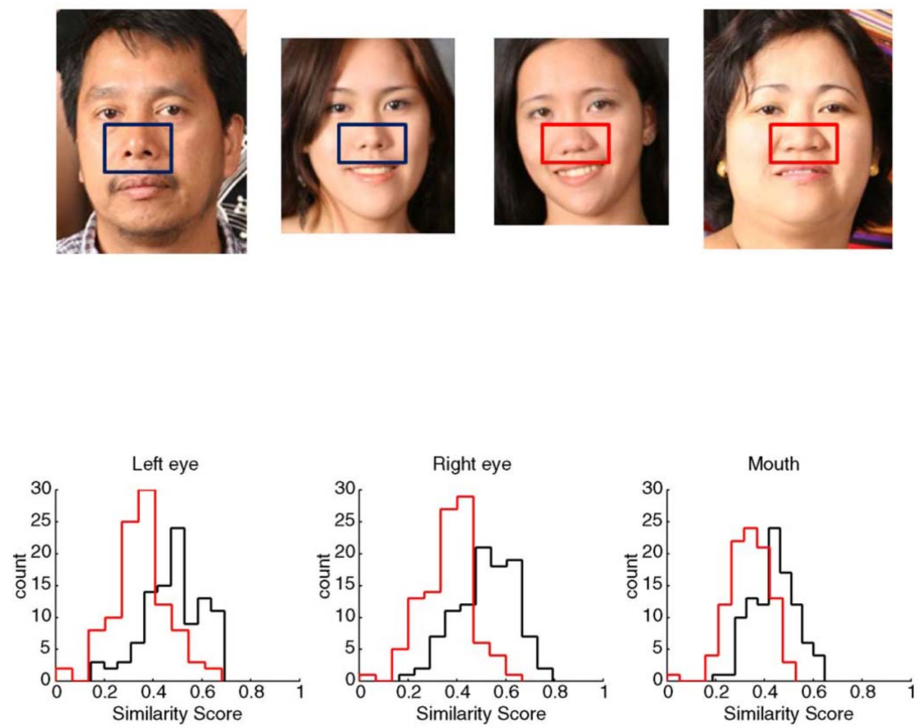


[2] Siyu Xia, Ming Shao, Jiebo Lu, and Yun Fu, Understanding kin relationships in photos, *IEEE TMM*, vol. 14, no. 8, pp. 1046-1056, 2012.

Related Work

Salient Feature + SVM (Guo2012 [3])





[3] Guodong Guo and Xiaolong Wang, Kinship measure on salient facial features, *IEEE TIM*, vol. 61, no. 8, pp. 2322-2325, 2012.

Datasets

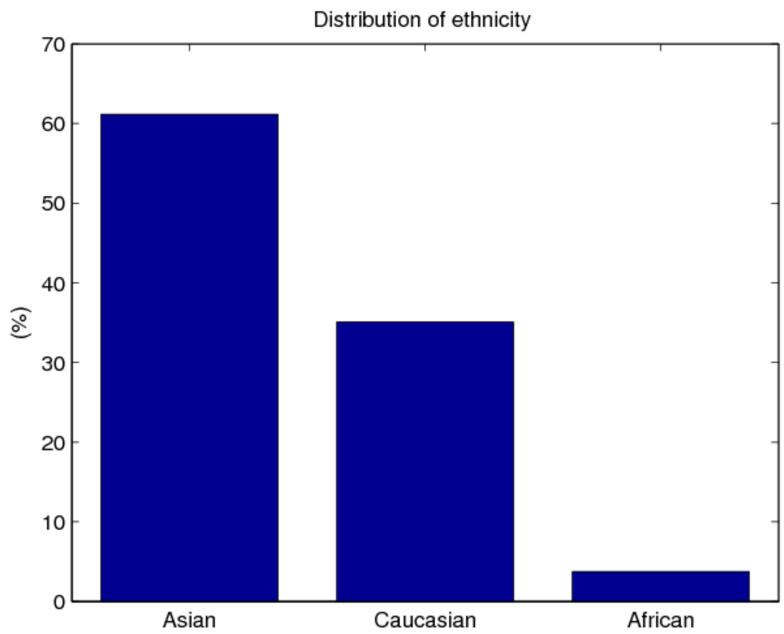
- KinFaceW-I: 500 kinship image face pairs

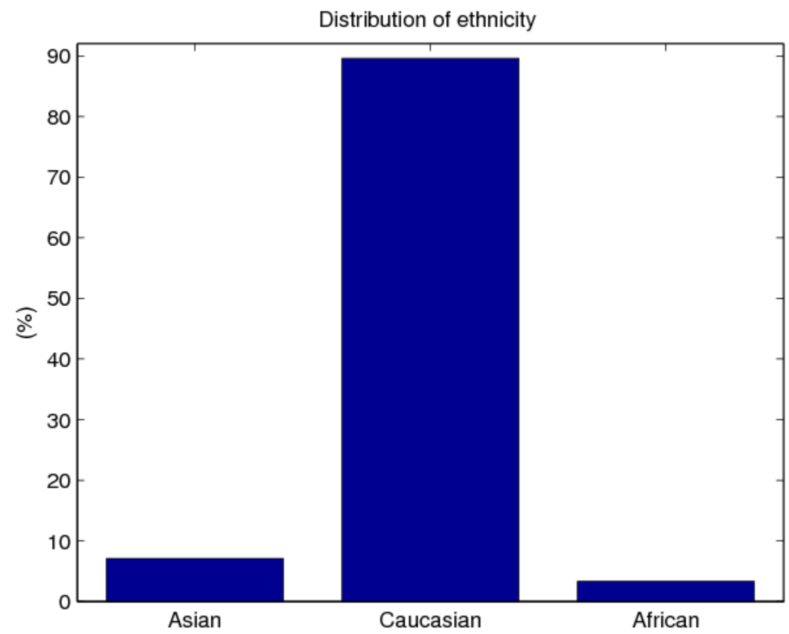
Datasets

- KinFaceW-II:1000 kinship image face pairs

Datasets

- Statistics





- Publicly available: www.kinfacew.com

Datasets

KinFaceW

Home Datasets Protocol Download Results References Contact Changes

Home

Welcome to Kinship Face in the Wild (**KinFaceW**), a database of face images collected for studying the problem of kinship verification from unconstrained face images. There are many potential applications for kinship verification such as family album organization, genealogical research, missing family members search, and social media analysis.

The aim of kinship verification is to determine whether there is a kin relation between a pair of given face images. The kinship is defined as a relationship between two persons who are biologically related with overlapping genes. Hence, there are four representative types of kin relations: Father-Son (F-S), Father-Daughter (F-D), Mother-Son (M-S) and Mother-Daughter (M-D), respectively.

News!

Sep-22-2014: The detailed information of The Kinship Verification in the Wild Evaluation can be found here, which is organized as part of FG2015.

Mahalanobis Distance

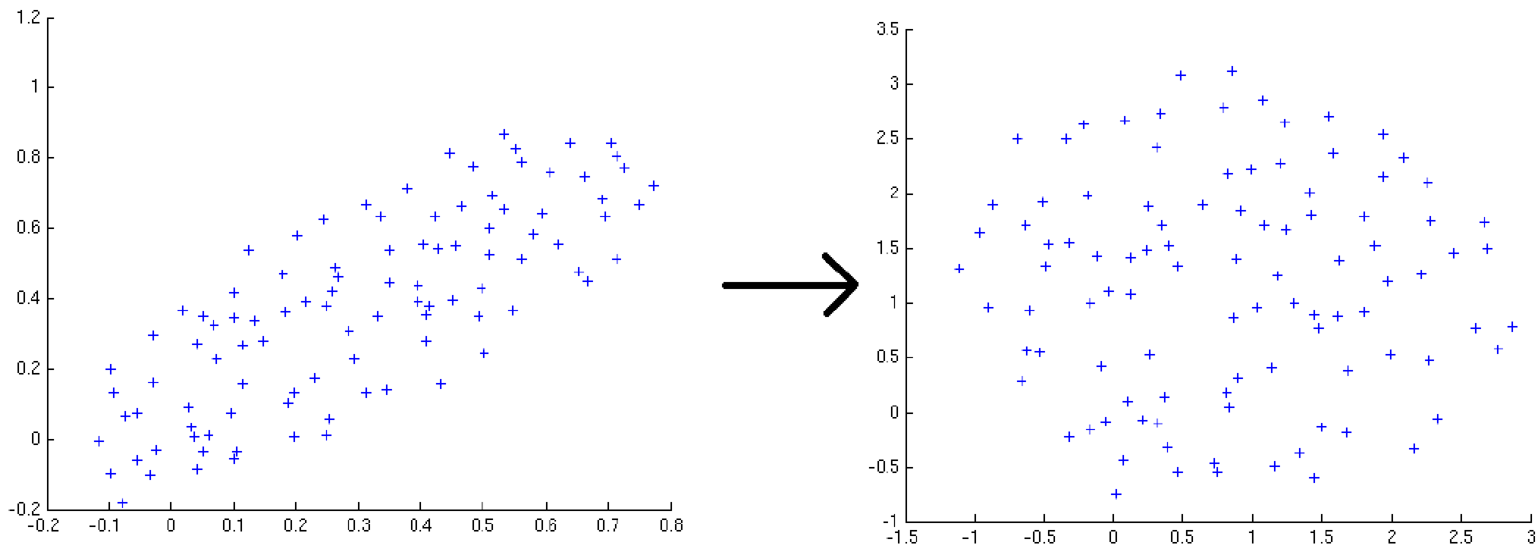
Squared Euclidean distance

$$\begin{aligned} d(\mathbf{x}_1, \mathbf{x}_2) &= \|\mathbf{x}_1 - \mathbf{x}_2\|_2^2 \\ &= (\mathbf{x}_1 - \mathbf{x}_2)^T (\mathbf{x}_1 - \mathbf{x}_2) \end{aligned}$$

Let $\Sigma = \sum_{i,j} (\mathbf{x}_i - \mu)(\mathbf{x}_j - \mu)^T$

The Mahalanobis distance

$$d_M(\mathbf{x}_1, \mathbf{x}_2) = (\mathbf{x}_1 - \mathbf{x}_2)^T \Sigma^{-1} (\mathbf{x}_1 - \mathbf{x}_2)$$



Metric Learning

Applying Mahalanobis distance to learn a positive semi-denite (PSD) matrix

$$d_{\mathbf{M}}(\mathbf{x}_i, \mathbf{x}_j) = \sqrt{(\mathbf{x}_i - \mathbf{x}_j)^T \mathbf{M} (\mathbf{x}_i - \mathbf{x}_j)}$$

Relationship with subspace learning

$$\begin{aligned} d_{\mathbf{M}}(\mathbf{x}_i, \mathbf{x}_j) &= \sqrt{(\mathbf{x}_i - \mathbf{x}_j)^T \mathbf{M} (\mathbf{x}_i - \mathbf{x}_j)} \\ &= \sqrt{(\mathbf{x}_i - \mathbf{x}_j)^T \mathbf{W}^T \mathbf{W} (\mathbf{x}_i - \mathbf{x}_j)} \\ &= \|\mathbf{W}\mathbf{x}_i - \mathbf{W}\mathbf{x}_j\|_2 \end{aligned}$$

where $\mathbf{M} = \mathbf{W}^T\mathbf{W}$.

Representative Metric Learning Algorithms

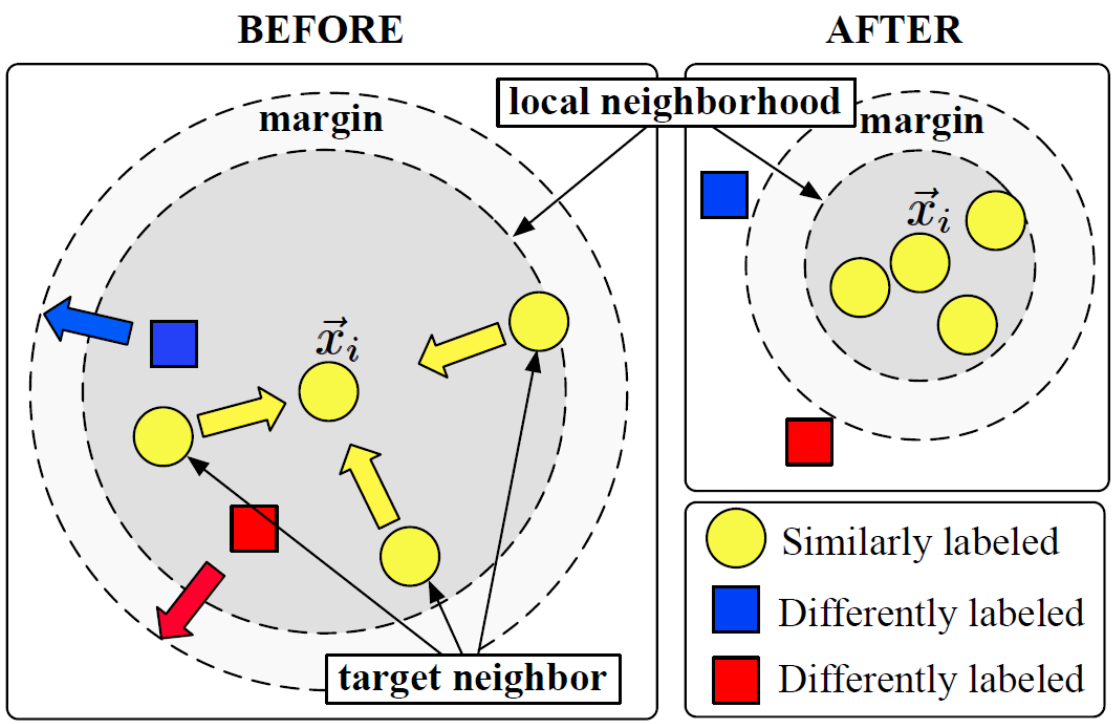
Large Margin Nearest Neighborhood (LMNN)

Minimize $\sum_{ij} \eta_{ij}(\vec{x}_i - \vec{x}_j)^\top \mathbf{M}(\vec{x}_i - \vec{x}_j) + c\sum_{ij} \eta_{ij}(1 - y_{il})\xi_{ijl}$ **subject to:**

(1) $(\vec{x}_i - \vec{x}_l)^\top \mathbf{M}(\vec{x}_i - \vec{x}_l) - (\vec{x}_i - \vec{x}_j)^\top \mathbf{M}(\vec{x}_i - \vec{x}_j) \geq 1 - \xi_{ijl}$

(2) $\xi_{ijl} \geq 0$

(3) $\mathbf{M} \succeq 0$.



[Weinberger et al, NIPS2005]

Representative Metric Learning Algorithms

Information-Theoretic Metric Learning (ITML)

$$\begin{aligned} \min_{A} \quad & \mathrm{KL}(p(\boldsymbol{x}; A_0) \| p(\boldsymbol{x}; A)) \\ \text{subject to} \quad & d_A(\boldsymbol{x}_i, \boldsymbol{x}_j) \leq u \qquad (i,j) \in S, \\ & d_A(\boldsymbol{x}_i, \boldsymbol{x}_j) \geq \ell \qquad (i,j) \in D. \end{aligned}$$

where $\mathrm{KL}(p(\boldsymbol{x}; A_o) \| p(\boldsymbol{x}; A)) = \int p(\boldsymbol{x}; A_0) \log \frac{p(\boldsymbol{x}; A_0)}{p(\boldsymbol{x}; A)} \, d\boldsymbol{x}$.

The optimization function can be re-formulated as

$$\begin{aligned} \min_{A \succeq 0} \quad & D_{\ell\mathrm{d}}(A, A_0) \\ \text{s.t.} \quad & \mathrm{tr}(A(\boldsymbol{x}_i - \boldsymbol{x}_j)(\boldsymbol{x}_i - \boldsymbol{x}_j)^T) \leq u \qquad (i,j) \in S, \\ & \mathrm{tr}(A(\boldsymbol{x}_i - \boldsymbol{x}_j)(\boldsymbol{x}_i - \boldsymbol{x}_j)^T) \geq \ell \qquad (i,j) \in D, \end{aligned}$$

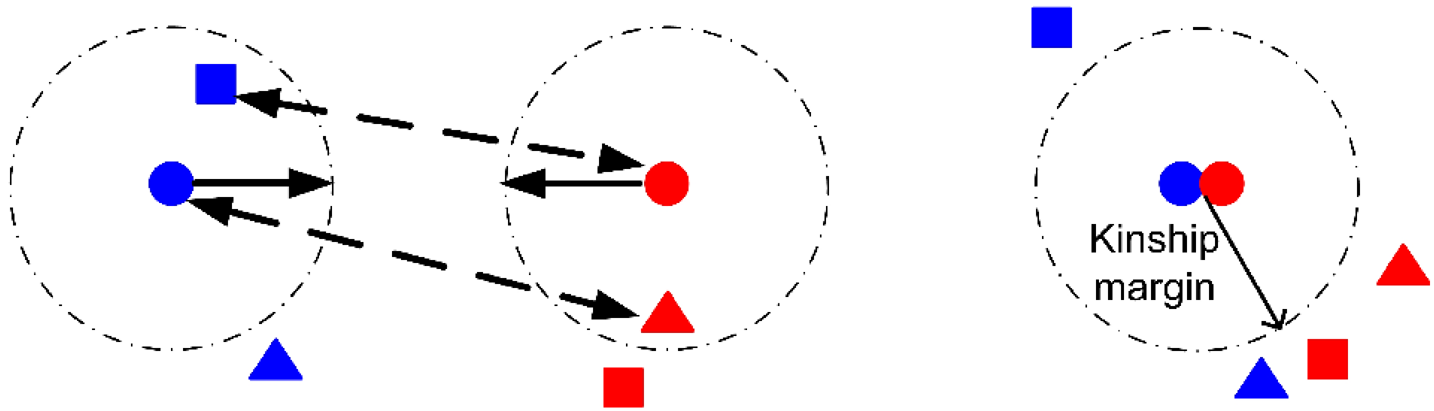
[Davis et al, ICML2007]

Neighborhood Repulsed Metric Learning

Motivations

- For the verification task, the number of negative pairs is larger than the number of positive pairs if we know the exact label information of each sample.
- The importance of different negative pairs is different. Some negative pairs are very discriminative and some are not so discriminative.
- It is desirable to identity the most informative negative pairs and ignore the less informative negative pairs to learn a discriminative metric for verification.

Neighborhood Repulsed Metric Learning



(a) (b)

$$
\begin{aligned}
&\max_A J(A) = J_1(A) + J_2(A) - J_3(A) \\
&= \frac{1}{Nk}\sum_{i=1}^{N}\sum_{t_1=1}^{k} d^2(x_i, y_{it_1}) + \frac{1}{Nk}\sum_{i=1}^{N}\sum_{t_2=1}^{k} d^2(x_{it_2}, y_i) \\
&-\frac{1}{N}\sum_{i=1}^{N} d^2(x_i, y_i) \\
&= \frac{1}{Nk}\sum_{i=1}^{N}\sum_{t_1=1}^{k}(x_i - y_{it_1})^T A (x_i - y_{it_1}) \\
&+\frac{1}{Nk}\sum_{i=1}^{N}\sum_{t_2=1}^{k}(x_{it_2} - y_i)^T A (x_{it_2} - y_i) \\
&-\frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^T A (x_i - y_i)
\end{aligned}
$$

Algorithm 1: NRML

Input: Training images: $S = \{(x_i, y_i) | i = 1, 2, \cdots, N\}$,
Parameters: neighborhood size k, iteration number T, and convergence error ε (set as 0.0001).

Output: Distance metric W.

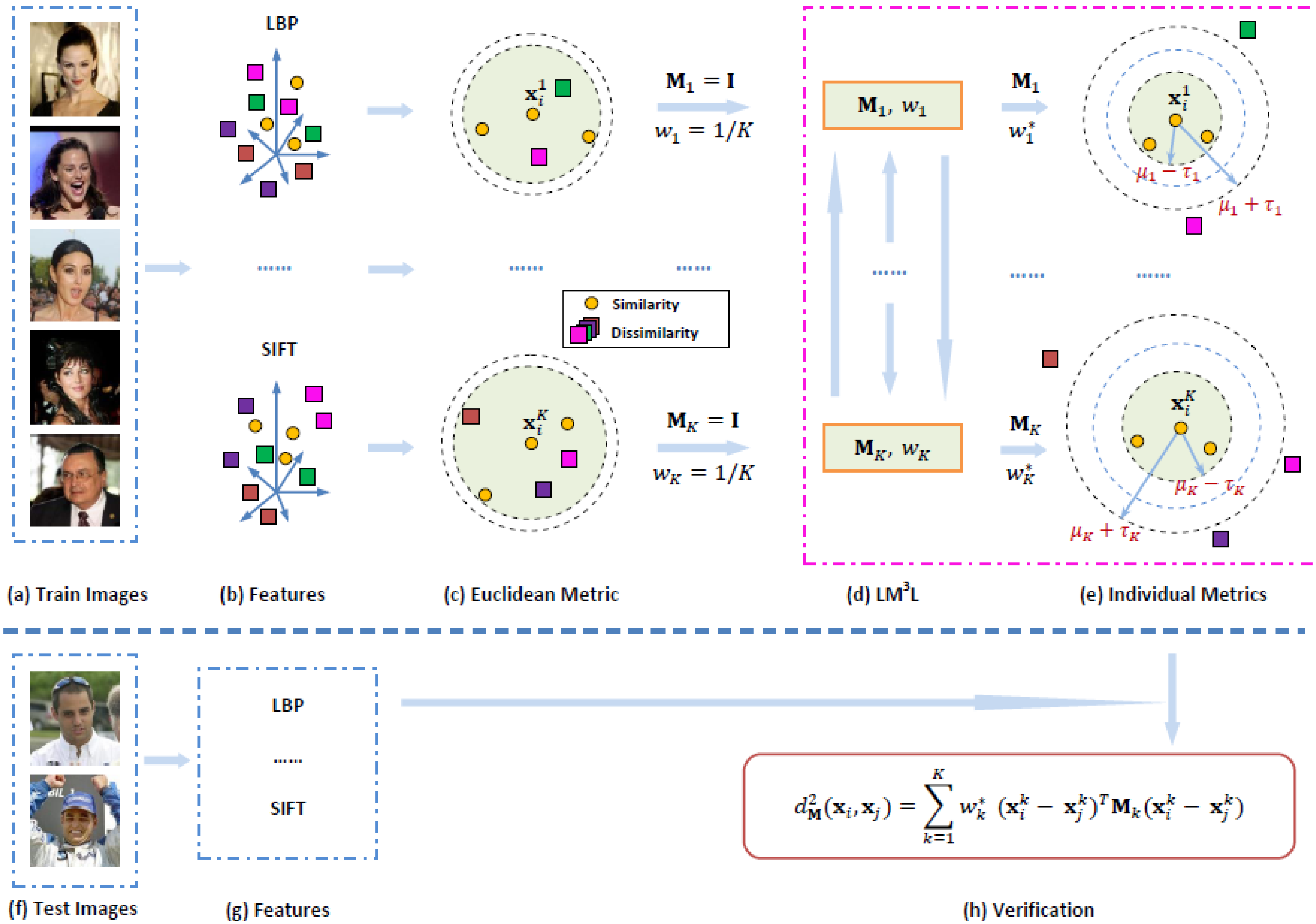
Step 1 (Initialization):
Search the k-nearest neighbors for each x_i and y_i by using the conventional Euclidean metric.

Step 2 (Local optimization):
For $r = 1, 2, \cdots, T$, repeat
- **2.1**. Compute H_1, H_2 and H_3, respectively.
- **2.2**. Solve the eigenvalue problem in Eq. (9).
- **2.3**. Obtain $W^r = [w_1, w_2, \cdots, w_l]$.
- **2.4**. Update the k-nearest neighbors of x_i and y_i by W^r.
- **2.5**. If $r > 2$ and $|W^r - W^{r-1}| < \varepsilon$, go to Step 3.

Step 3 (Output distance metric):
Output distance metric $W = W^r$.

[4] Jiwen Lu, Xiuzhuang Zhou, Yap-Peng Tan, Yuanyuan Shang, Jie Zhou, Neighborhood repulsed metric learning for kinship verification, *IEEEE PAMI*, vol. 36, no. 2, pp. 331-345, 2014.

Multi-view Neighborhood Repulsed Metric Learning



Multi-view Neighborhood Repulsed Metric Learning

$$\max_{W,\beta} \sum_{p=1}^{K} \beta_p tr[W^T(H_1^p + H_2^p - H_3^p)W]$$

$$\text{subject to} \quad W^TW = I, \sum_{p=1}^{K} \beta_p = 1, \beta_p \geq 0.$$

$$\max_{W,\beta} \sum_{p=1}^{K} \beta_p^q tr[W^T(H_1^p + H_2^p - H_3^p)W]$$

$$\text{subject to} \quad W^TW = I, \sum_{p=1}^{K} \beta_p = 1, \beta_p \geq 0.$$

Algorithm 2: MNRML

Input: Training images: $S^p = \{(x_i^p, y_i^p)|i = 1, 2, \cdots, N\}$ be the pth view set of N pairs of kinship images, Parameters: neighborhood size k, iteration number T, tuning parameter q, and convergence error ε (set as 0.0001).

Output: Distance metric W.

Step 1 (Initialization):

1.1. Set $\beta = [1/K, 1/K, \cdots, 1/K]$;

1.2. Obtain W^0 by solving Eq. (18).

Step 2 (Local optimization):

For $r = 1, 2, \cdots, T$, repeat

2.1. Compute β by using Eq. (16).

2.2. Obtain W^r by solving Eq. (18).

2.3. If $r > 2$ and $|W^r - W^{r-1}| < \varepsilon$, go to Step 3.

Step 3 (Output distance metric):

Output distance metric $W = W^r$.

Experimental Results

- Comparisons with existing metric learning methods

Method	Feature	F-S	F-D	M-S	M-D	Mean
CSML	LBP	63.7	61.2	55.4	62.4	60.7
	LE	61.1	58.1	60.9	70.0	62.5
	SIFT	66.5	60.0	60.0	56.4	59.8
	TPLBP	57.3	61.5	63.2	57.0	59.7
NCA	LBP	61.7	62.2	56.4	62.4	60.7
	LE	62.1	57.1	61.9	69.0	62.3
	SIFT	67.5	61.0	61.0	57.4	60.8
	TPLBP	56.3	60.5	62.2	56.0	58.7
LMNN	LBP	62.7	63.2	57.4	63.4	61.7
	LE	63.1	58.1	62.9	70.0	63.3
	SIFT	69.5	63.0	63.0	59.4	62.8
	TPLBP	57.3	61.5	63.2	57.0	59.7
NRML	LBP	64.7	65.2	59.4	65.4	63.7
	LE	64.1	59.1	63.9	71.0	64.3
	SIFT	70.5	64.0	64.0	60.4	63.8
	TPLBP	59.3	63.5	65.2	60.0	62.9
MNRML	All	**72.5**	**66.5**	**66.2**	**72.0**	**69.9**

On KinFaceW-I dataset.

Method	Feature	F-S	F-D	M-S	M-D	Mean
CSML	LBP	66.0	65.5	64.8	65.0	65.3
	LE	71.8	68.1	73.8	74.0	71.9
	SIFT	62.0	58.9	56.8	57.4	58.8
	TPLBP	66.4	62.6	62.8	64.9	64.2
NCA	LBP	67.0	66.5	65.8	66.0	66.3
	LE	73.8	70.1	74.8	75.0	73.5
	SIFT	63.0	59.9	58.8	59.4	60.4
	TPLBP	67.4	63.6	63.8	66.9	66.5
LMNN	LBP	68.0	68.5	68.8	67.0	68.2
	LE	74.8	71.1	75.8	76.0	74.5
	SIFT	65.0	57.9	58.8	59.4	60.4
	TPLBP	68.4	65.6	65.8	67.9	68.1
NRML	LBP	69.0	69.5	69.8	69.0	69.5
	LE	76.8	73.1	76.8	77.0	75.7
	SIFT	68.0	60.9	60.8	61.4	62.8
	TPLBP	70.4	67.6	67.8	69.9	70.1
MNRML	All	**76.9**	**74.3**	**77.4**	**77.6**	**76.5**

On KinFaceW-II dataset.

Experimental Results

- Comparisons with multi-view learning methods

Method	KinFaceW-I	KinFaceW-II
MSE	68.4	74.3
MKL	67.5	73.4
MNRML	69.9	76.5

- Comparisons with human observers

Method	F-S	F-D	M-S	M-D	Mean
HumanA	61.00	58.00	66.00	70.00	63.75
HumanB	67.00	65.00	75.00	77.00	71.00

Correct verification accuracy on the KinFaceW-I dataset.

Method	F-S	F-D	M-S	M-D	Mean
HumanA	61.00	61.00	69.00	73.00	66.75
HumanB	70.00	68.00	78.00	80.00	74.00

Correct verification accuracy on the KinFaceW-II dataset.

Discriminative Multi-Metric Learning

$$\min_{M_1,\ldots,M_K,\alpha} J = \sum_{k=1}^{K} \alpha_k f_k(M_k) + \lambda g_k(W_1,\ldots,W_K)$$

$$\text{subject to} \quad \sum_{k=1}^{K} \alpha_k = 1, \alpha_k \geq 0.$$

where

$$f_k(M_k) = -\log\Big(\prod_{O_1^k} P\big(g(x_i^k, y_i^k) < g(x_i^k, y_j^k)\big)\Big) - \log\Big(\prod_{O_2^k} P\big(g(x_i^k, y_i^k) < g(x_l^k, y_i^k)\big)\Big)$$

$$g_k(W_1,\ldots,W_K) = \sum_{\substack{k_1,k_2=1\\k_1\neq k_2}}^{K} \sum_{i=1}^{N} \|W_{k_1}^T x_i^{k_1} - W_{k_2}^T x_i^{k_2}\|_F^2$$

$$\min_{W_1,\ldots,W_K,\alpha} J = \sum_{k=1}^{K} \alpha_k f_k(W_k) + \lambda \sum_{\substack{k_1,k_2=1\\k_1\neq k_2}}^{K} \sum_{i=1}^{N} \|W_{k_1}^T x_i^{k_1} - W_{k_2}^T x_i^{k_2}\|_F^2$$

where

$$f_k(W_k) = \prod_{O_1^k} \log\big(1 + \exp(\|W_k^T x_{ik}^p\|^2 - \|W_k^T x_{ik}^n\|^2)\big)$$

[5] Haibin Yan, Jiwen Lu, Weihong Deng, Xiuzhuang Zhou, Discriminative multimetric learning for kinship verification, *IEEE TIFS,* vol. 9, no. 7, pp. 1169-1178, 2014.

Discriminative Multi-Metric Learning

$$\min_{W_k} J(W_k) = \alpha_k f_k(W_k) + \lambda \sum_{l=1, l\neq k}^{K} G(W_k)$$

where

$$G(W_k) = \sum_{i=1}^{N} \|W_k^T x_i^k - W_l^T x_i^l\|_2^2$$

$$\frac{\partial f_k(W_k)}{\partial W_k} = \prod_{O_1^k} \frac{2 + \exp(\|W_k^T x_{ik}^p\|^2 - \|W_k^T x_{ik}^n\|^2)}{1 + \exp(\|W_k^T x_{ik}^p\|^2 - \|W_k^T x_{ik}^n\|^2)} \times (x_{ik}^p x_{ik}^{pT} - x_{ik}^n x_{ik}^{nT}) W_k$$

$$\frac{\partial G(W_k)}{\partial W_k} = 2\lambda(K-1) W_k \sum_{i=1}^{N} (x_i^k)^T x_i^k - 2\lambda W_k \sum_{\substack{l=1 \\ l\neq k}}^{K} \sum_{i=1}^{N} (x_i^l)^T x_i^l$$

$$W_k^{t+1} = W_k^t - \eta(\alpha_k \frac{f_k(W_k)}{W_k} + \lambda \sum_{l=1, l\neq k}^{K} \frac{\partial G(W_k)}{\partial W_k})$$

Iteration is terminated: $J(W_k^t) - J(W_k^{t+1}) < \varepsilon$ or $\|W_k^{t+1} - W_k^t\| < \varepsilon$

Discriminative Multi-Metric Learning

$$\min_{\alpha} \quad J(\alpha) = \sum_{k=1}^{K} \alpha_k^r f_k(W_k)$$

$$\text{subject to} \quad \sum_{k=1}^{K} \alpha_k = 1, \quad \alpha_k > 0.$$

The Lagrange function can be constructed as:

$$L(\alpha, \zeta) = \sum_{k=1}^{K} \alpha_k^r f_k(W_k) - \zeta(\sum_{k=1}^{K} \alpha_k - 1)$$

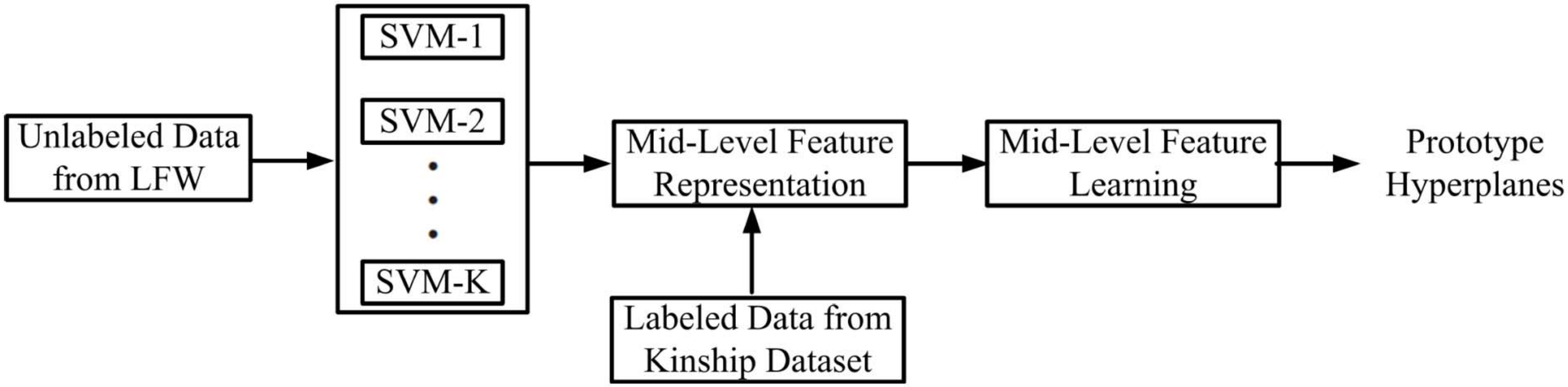
Let $\frac{\partial L(\alpha,\zeta)}{\partial \alpha_k} = 0$ and $\frac{\partial L(\alpha,\zeta)}{\partial \zeta} = 0$, we have

$$r\alpha_k^{r-1} f_k(W_k) - \zeta = 0$$

$$\sum_{k=1}^{K} \alpha_k - 1 = 0$$

$$\alpha_k = \frac{(1/f_k(W_k))^{1/(r-1)}}{\sum_{k=1}^{K} (1/f_k(W_k))^{1/(r-1)}}$$

Prototype-Based Discriminative Feature Learning



$$\begin{aligned} \max H(B) &= H_1(B) + H_2(B) - H_3(B) \\ &= \frac{1}{Mk}\sum_{i=1}^{M}\sum_{t_1=1}^{k}\|f(x_i) - f(y_{it_1})\|_2^2 \\ &+ \frac{1}{Mk}\sum_{i=1}^{M}\sum_{t_2=1}^{k}\|f(x_{it_2}) - f(y_i)\|_2^2 \\ &- \frac{1}{M}\sum_{i=1}^{M}\|f(x_i) - f(y_i)\|_2^2 \\ \text{subject to} \quad & \|\beta_k\|_1 \leq \gamma, k = 1, 2, \cdots, K. \end{aligned}$$

[5] Haibin Yan, Jiwen Lu, Xiuzhuang Zhou, Prototype-based discriminative feature learning for kinship verification, *IEEE TCYB,*2014, accepted.

Summary and Future Work

- Metric learning is effective for kinship verification.
- Feature learning for kinship verification?
- Deep learning for kinship verification?