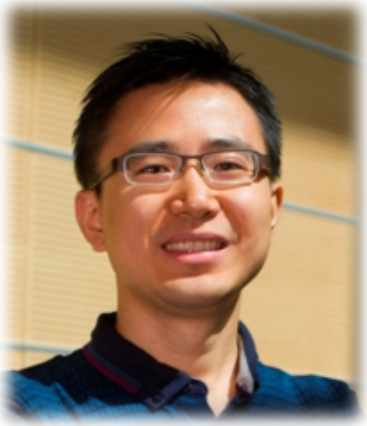
Northeastern University

Smile lab

Synergetic Media Learning Lab

Multi-view Face Representation

----FG-2017 Tutorial



Zhengming Ding **Handong Zhao** **Yun Fu**

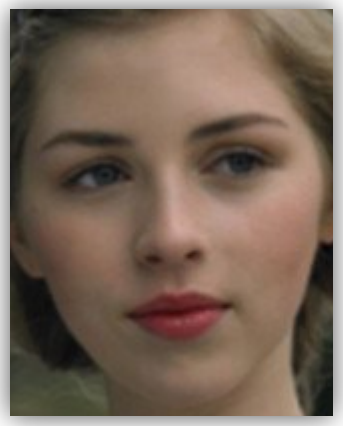
Northeastern University, Boston, USA

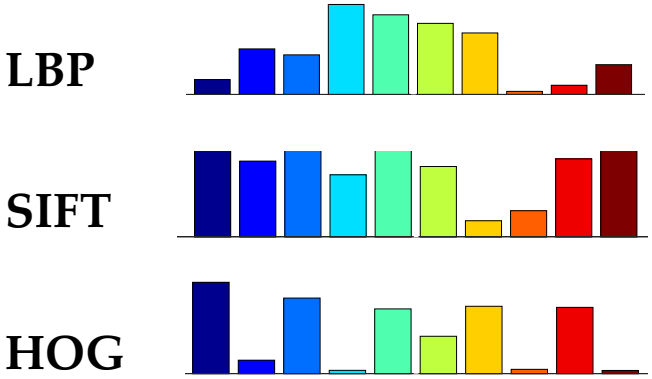
Multi-view Face

Northeastern University

Smile lab

Synergetic Media Learning Lab





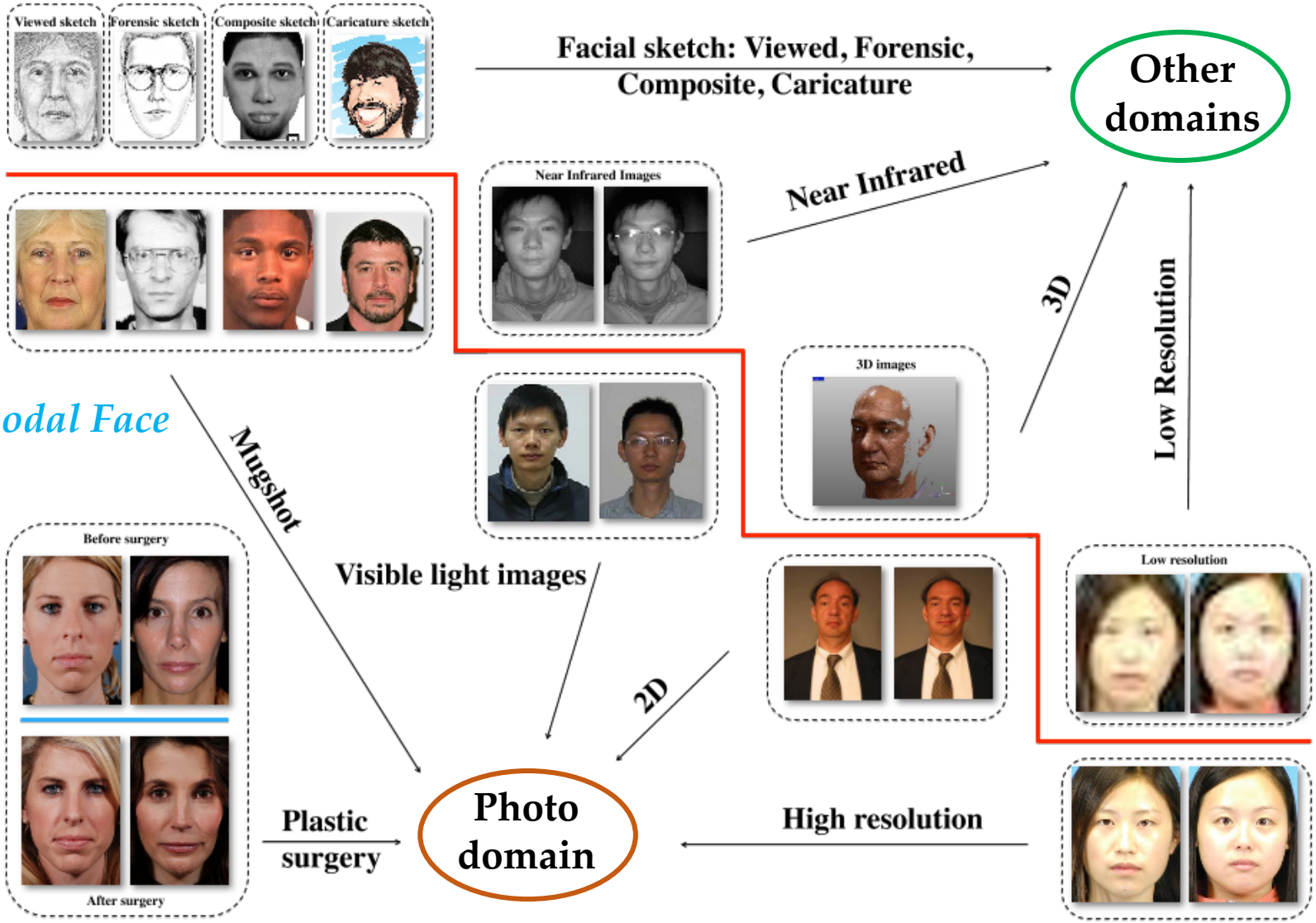
Multiple features

Multi-pose Face

Multi-view Face

Northeastern University

Smile lab — Synergetic Media Learning Lab



Outline

Northeastern University

Smile lab

Synergetic Media Learning Lab

- **Introduction & Background**
  - Multi-view Face Task
  - Multi-view Face Data
- **Unsupervised Multi-view Face Representation**
  - Methodology
  - Face Clustering, Outlier Detection
- **Supervised Multi-view Face Representation**
  - Multi-view Learning
  - Transfer Learning
- **Conclusion**

Multi-view Face Tasks

Northeastern University

Smile lab — Synergetic Media Learning Lab

- Multi-view Face Clustering
- Multi-view Face Verification

→ *Multiple feature based face*

→ *Multi-modal face* *[One Modality, One Feature]*

- Multi-view Face Identification

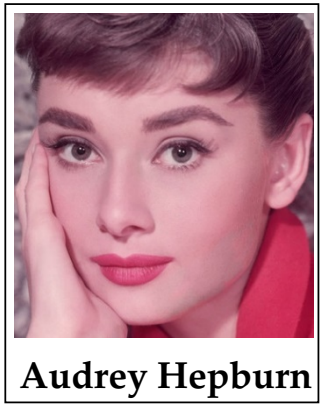
→ *Multiple feature based face*

→ *Multi-pose face*

→ *Multi-modal face* *[One Modality, One Domain]*

Face Clustering

Face Verification



Face Identification

Multi-view Face Tasks

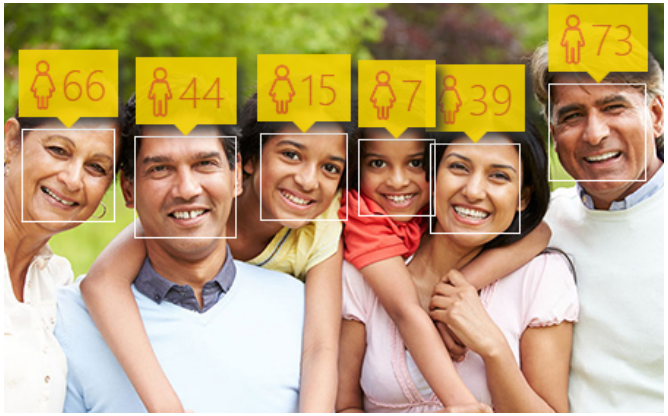
Northeastern University

Smile lab

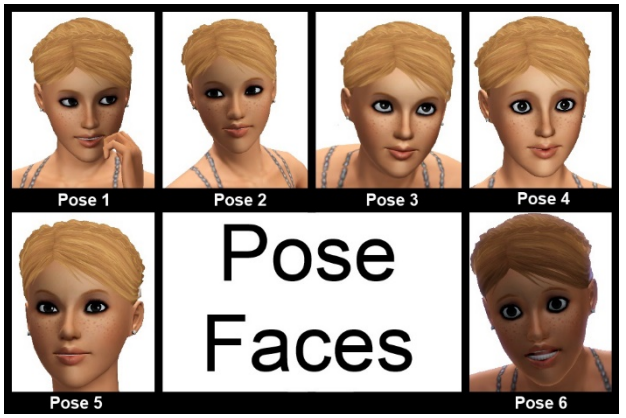
Synergetic Media Learning Lab

Multiple feature based face

- Multi-view Age Estimation

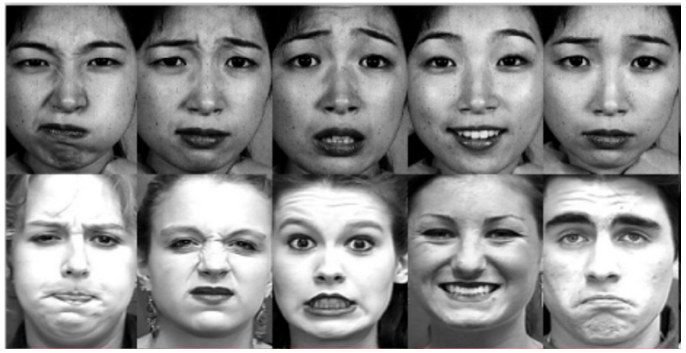


- Multi-view Pose Estimation



- Multi-view Kinship Verification

- Multi-view Facial Expression

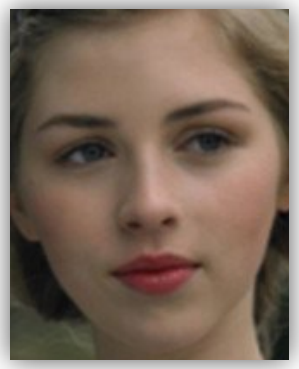


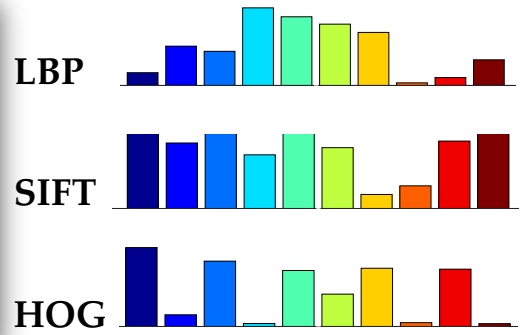
Multi-view Face Data

Northeastern University

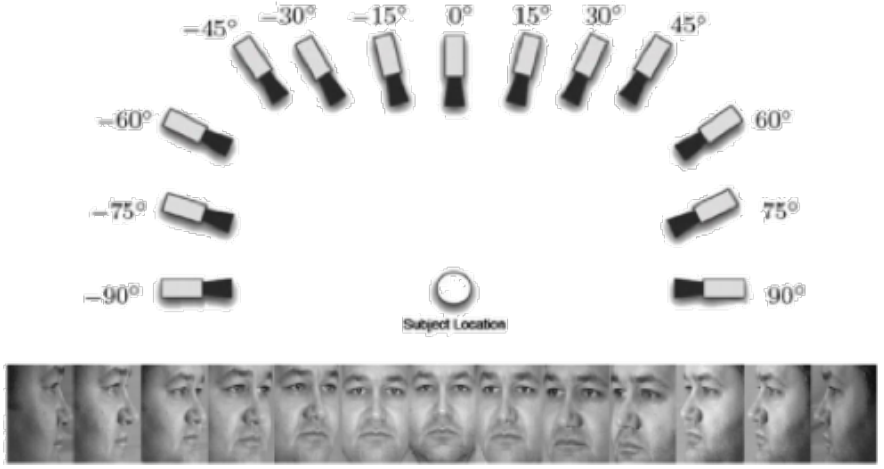
Smile lab — Synergetic Media Learning Lab

❑ **Data Correspondence [Only One Factor]→ *Heterogeneous***

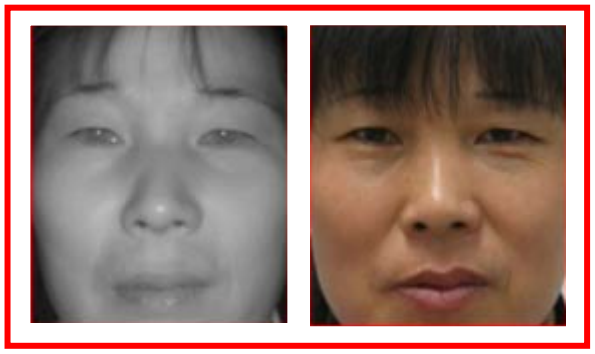




Multiple features



❑ **No Data Correspondence [Multiple Factors]**



expression

expression

occlusion

Taxonomy [*Data View*]

Northeastern University

Smile lab — Synergetic Media Learning Lab

❑ **Category 1 [Data Correspondence] (Multi-view Learning)**

- **Multiple Features,** e.g., LBP, SIFT, HOG…

  Goal: fuse various knowledge from multiple features to boost the final tasks

  ⇒ ✓ Clustering ✓ Verification ✓ Identification

- **Multi-Pose/Multi-Modal Face**

  Goal: seek a view-invariant space to mitigate the view divergence to facilitate the final task

  (***adapt*** *knowledge across different views)*

❑ **Category 2 [No Data Correspondence] (Transfer learning)**

- **Multi-Feature/Multi-Pose/Multi-Modal Face**

  Goal: transfer knowledge from well-labeled source views to unlabeled target views

Taxonomy [*Task View*]

Northeastern University

Smile lab

Synergetic Media Learning Lab

- **Unsupervised Learning [Clustering]**

  Goal: fuse various knowledge from multiple features

  Data: multiple features [*unlabeled, data correspondence*]

- **Supervised Learning [Recognition]**

  Goal: seek a view-invariant space to mitigate the view divergence to facilitate the final task (***adapt*** *knowledge across different views)*

  - **Sub-Category 1 (Multi-view Learning) [Data Correspondence]**

    Training Stage: multiple labeled view data
    Test Stage: some labeled views, some unlabeled views

  - **Sub-Category 2 (Transfer learning) [No Data Correspondence]**

    Training Stage: some source labeled views & some target unlabeled views

Outline

Northeastern University

Smile lab

Synergetic Media Learning Lab

- Introduction & Background
  - Multi-view Face Task
  - Multi-view Face Data
- **Unsupervised Multi-view Face Representation**
  - Methodology
  - Face Clustering, Outlier Detection
- Supervised Multi-view Face Representation
  - Multi-view Learning
  - Transfer Learning
- Conclusion

The number of the labeled face is limited.

Northeastern University

Smile lab

Synergetic Media Learning Lab

[Image courtesy to Lihong Wan, Hong Huo and Tao Fang]

Methodology

- Multi-view Face Clustering

Northeastern University

Smile lab

Synergetic Media Learning Lab

Canonical Correlation Analysis (CCA)

Formally, for two views $X \in R^{d \times n}$ and $Y \in R^{k \times n}$, CCA computes two projection vectors, $w_x \in R^d$ and $w_y \in R^k$, such that the following correlation coefficient is maximized:

$$\rho = \frac{w_x^T X Y^T w_y}{\sqrt{(w_x^T X X^T w_x)(w_y^T Y Y^Y w_y)}}$$

Since ρ is invariant to the scaling of w_x and w_y, CCA can be formulated equivalently as

$$\begin{aligned} \max_{w_x, w_y} \quad & w_x^T X Y^T w_y \\ \texttt{s.t.} \quad & w_x^T X X^T w_x = 1, \quad w_y^T Y Y^T w_y = 1 \end{aligned}$$

Relations between two sets of variates. Biometrika, 28(3/4):321–377, 1936

H. Hotelling.

Methodology

- Multi-view Face Clustering

Northeastern University

Smile lab

Synergetic Media Learning Lab

Canonical Correlation Analysis (CCA)

Formally, for two views $X \in R^{d \times n}$ and $Y \in R^{k \times n}$, CCA computes two projection vectors, $w_x \in R^d$ and $w_y \in R^k$, such that the following correlation coefficient is maximized:

$$\rho = \frac{w_x^T X Y^T w_y}{\sqrt{(w_x^T X X^T w_x)(w_y^T Y Y^Y w_y)}}$$

Since ρ is invariant to the scaling of w_x and w_y, CCA can be formulated equivalently as

$$\begin{aligned} \max_{w_x, w_y} \quad & w_x^T X Y^T w_y \\ \text{s.t.} \quad & w_x^T X X^T w_x = 1, \quad w_y^T Y Y^T w_y = 1 \end{aligned}$$

✓ It has a sense of consensus.

Relations between two sets of variates. Biometrika, 28(3/4):321–377, 1936

H. Hotelling.

Methodology

- Multi-view Face Clustering

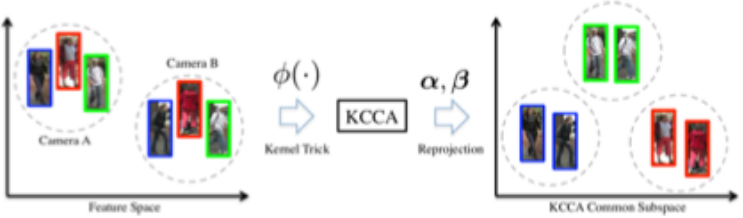
Northeastern University

Smile lab

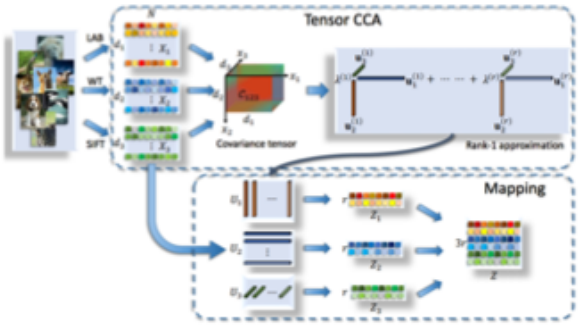
Synergetic Media Learning Lab

Extensions of CCA

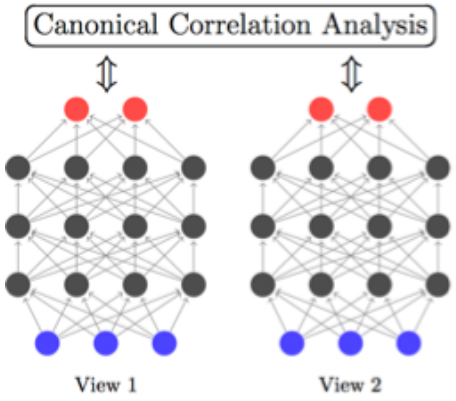
Kernel-based



Tensor-based



DeepNN-based



[**Kernel-based**] David R. Hardoon, Sándor Szedmák, John Shawe-Taylor: Canonical Correlation Analysis: An Overview with Application to Learning Methods. Neural Computation 16(12): 2639-2664 (2004)

[**Tensor-based**] Tensor Canonical Correlation Analysis for Multi-view Dimension Reduction, Yong Luo, Dacheng Tao, Kotagiri Ramamohanarao, Chao Xu, and Yonggang Wen, IEEE Transactions on Knowledge and Data Engineering (T-KDE), vol. 27, no. 11, pp. 3111-3124, 2015.

[**DeepNN-based**] Galen Andrew, Raman Arora, Jeff A. Bilmes, Karen Livescu: Deep Canonical Correlation Analysis. ICML (3) 2013: 1247-1255

Methodology

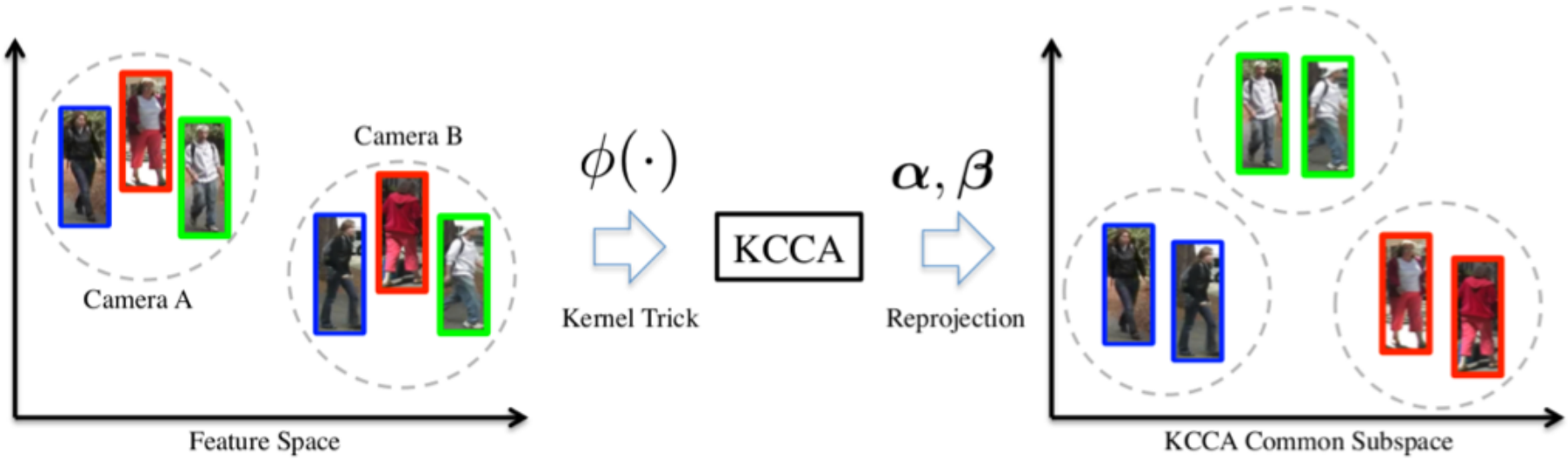
- Multi-view Face Clustering

Northeastern University

Smile lab

Synergetic Media Learning Lab

Kernel CCA



- KCCA uses the kernel trick to produce a non-linear version of CCA, by looking for functions $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ such that the random variables $\boldsymbol{\alpha}(\boldsymbol{x})$and $\boldsymbol{\beta}(\boldsymbol{y})$ have maximal correlation.

Canonical Correlation Analysis: An Overview with Application to Learning Methods, Neural Computation 2004
David R. Hardoon, Sándor Szedmák, John Shawe-Taylor

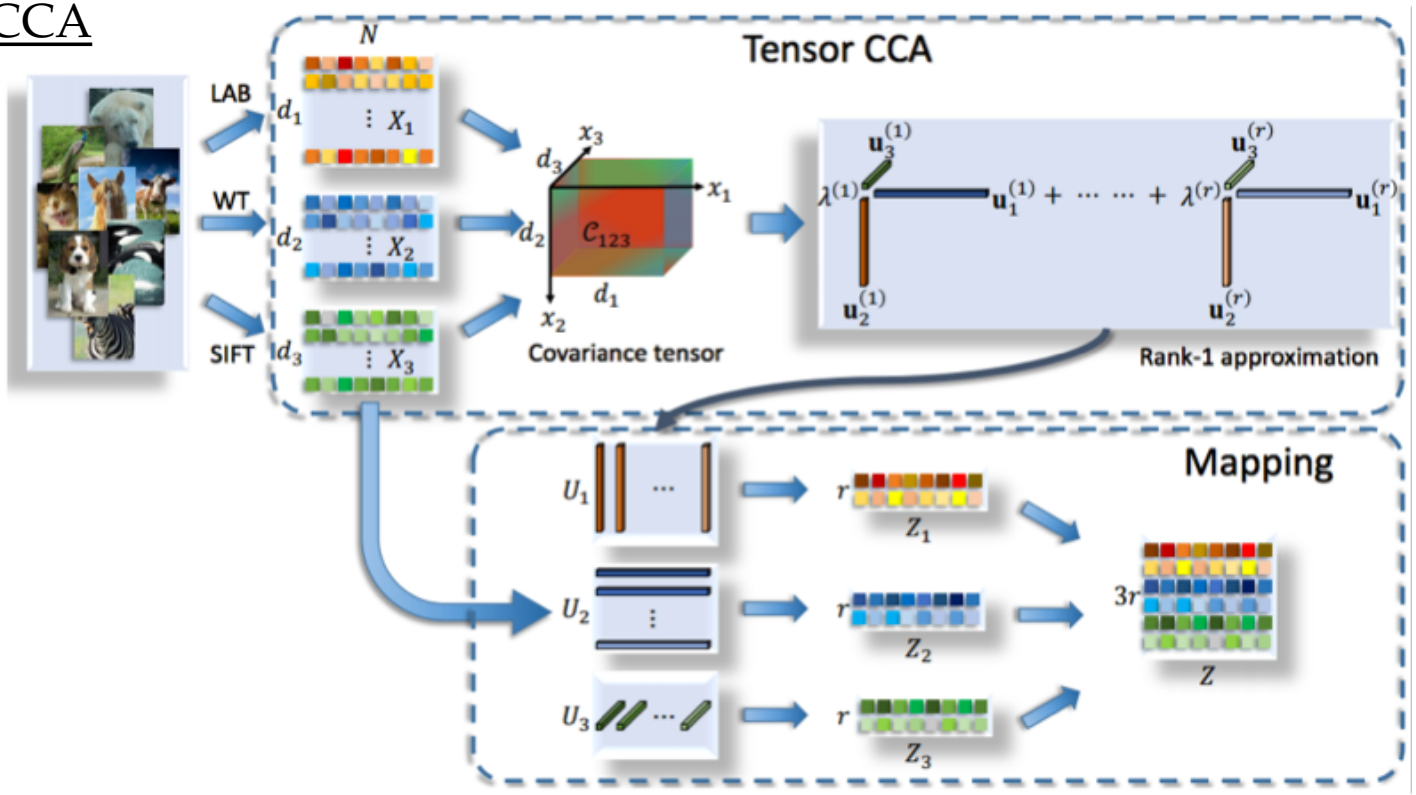
Methodology

- Multi-view Face Clustering

Northeastern University

Smile lab — Synergetic Media Learning Lab

Tensor CCA



- TCCA straightforwardly yet naturally generalizes CCA to handle the data of an arbitrary number of views by analyzing the covariance tensor of the different views.

Tensor CCA for Multi-view Dimension Reduction, TKDE'15

Yong Luo, Dacheng Tao, Kotagiri Ramamohanarao, Chao Xu, and Yonggang Wen

Methodology

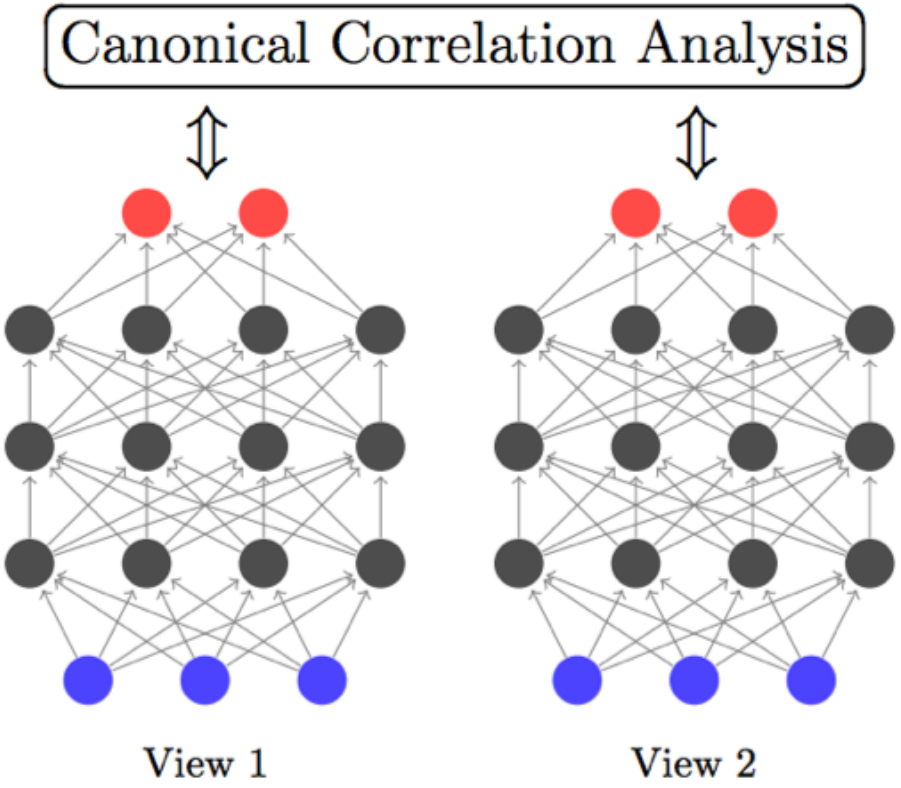
- Multi-view Face Clustering

Northeastern University

Smile lab – Synergetic Media Learning Lab

Deep CCA

- DCCA can be viewed as a nonlinear extension of the linear method CCA.
- It is an alternative to the nonparametric method kernel CCA for learning correlated nonlinear transformations.
- Unlike KCCA, DCCA does not require an inner product, and has the advantages of a parametric method: training time scales well with data size.
- The training data need not be referenced when computing the representations of unseen instances.



Deep Canonical Correlation Analysis. ICML '13
Galen Andrew, Raman Arora, Jeff A. Bilmes, Karen Livescu

Methodology

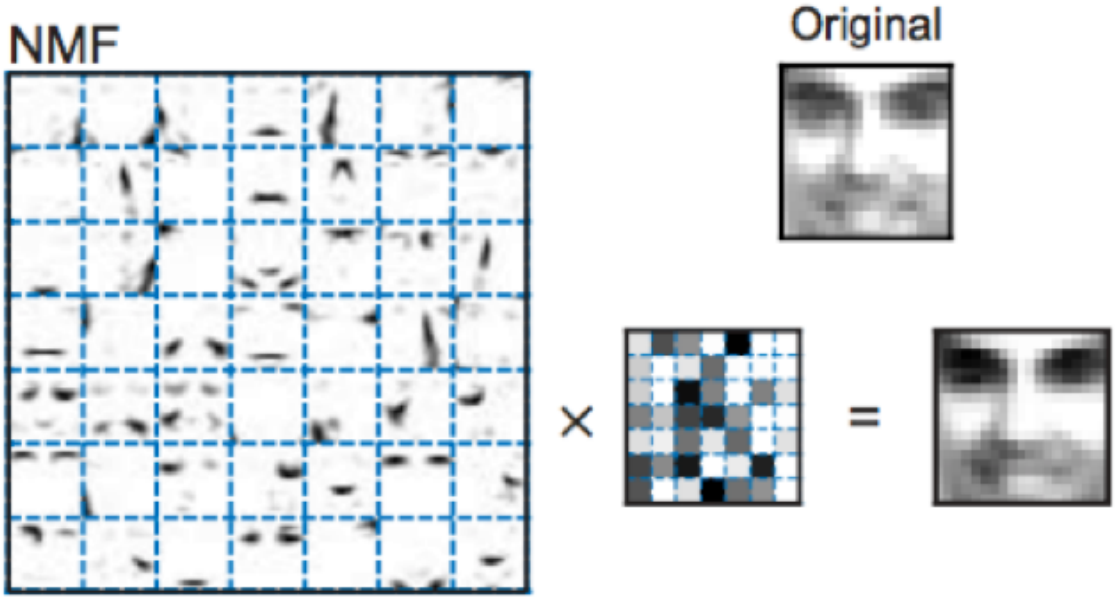
- Multi-view Face Clustering

Northeastern University

Smile lab

Synergetic Media Learning Lab

Multi-view Clustering via Joint **Nonnegative Matrix Factorization**



Objective function: $\min_{U,V} ||X - UV^T||_F^2, \ s.t. \ U \geq 0, V \geq 0$

Learning the parts of objects by non-negative matrix factorization – Nature'99

Daniel D Lee, H Sebastian Seung

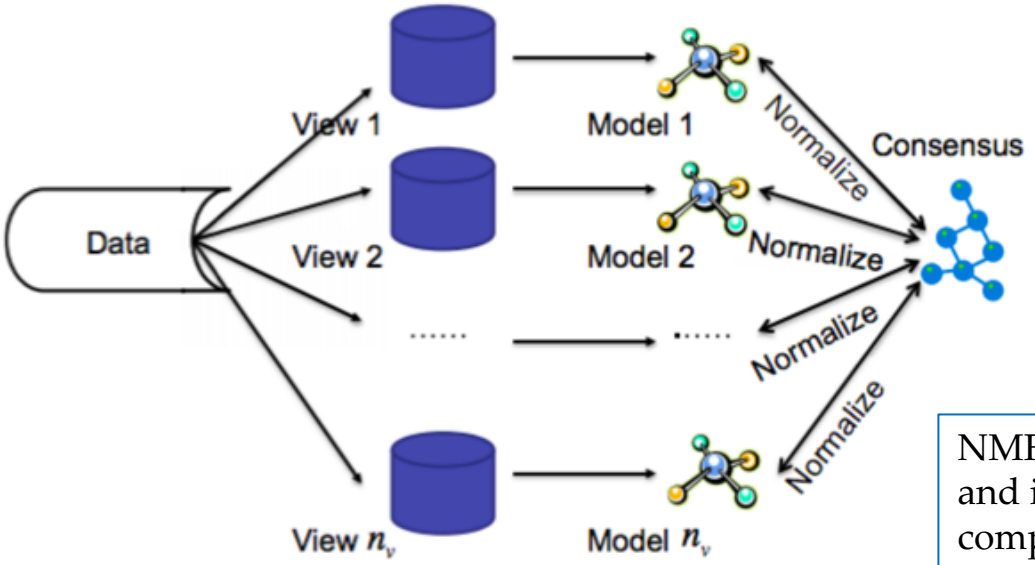
Methodology

- Multi-view Face Clustering

Northeastern University

Smile lab – Synergetic Media Learning Lab

Multi-view Clustering via Joint Nonnegative Matrix Factorization



NMF has a good interpretability, and it is reported to achieve competitive performance compared with most of the state-of-the-art unsupervised algorithms.

Objective function:

$$\sum_{v=1}^{n_v} \|X^{(v)} - U^{(v)}(V^{(v)})^T\|_F^2 + \sum_{v=1}^{n_v} \lambda_v \|V^{(v)} - V^*\|_F^2$$

$$s.t. \quad \forall 1 \le k \le K, \|U_{.,k}^{(v)}\|_1 = 1, U^{(v)}, V^{(v)}, V^* \ge 0$$

Multi-View Clustering via Joint NMF – SDM'13

Jialu Liu, Chi Wang, Jing Gao, and Jiawei Han

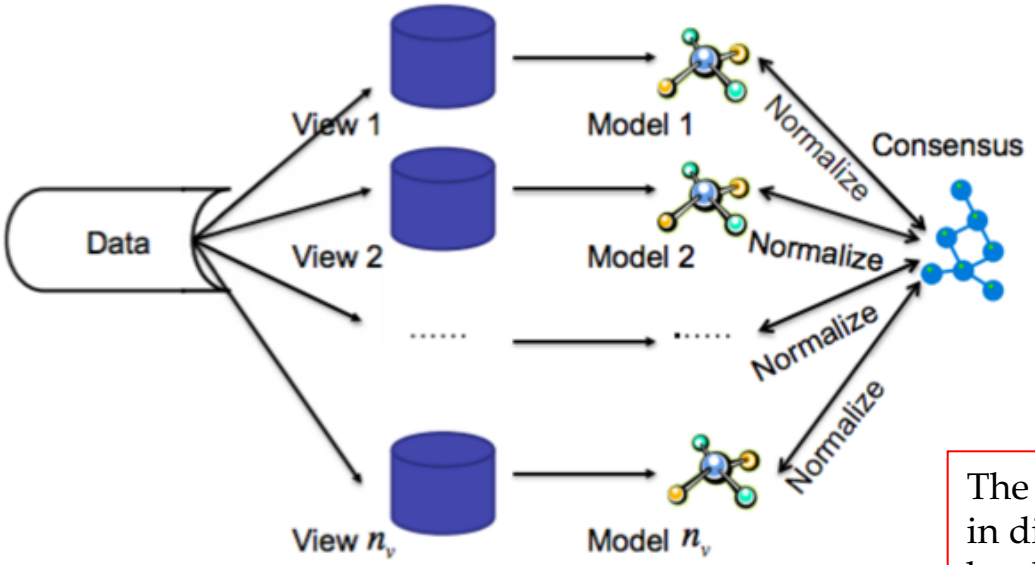
Methodology

- Multi-view Face Clustering

Northeastern University

Smile lab – Synergetic Media Learning Lab

Multi-view Clustering via Joint Nonnegative Matrix Factorization



The latent representations $V^{(v)}$ in different views are forced to be close to the consensus one V^*.

Objective function:

$$\sum_{v=1}^{n_v} \|X^{(v)} - U^{(v)}(V^{(v)})^T\|_F^2 + \sum_{v=1}^{n_v} \lambda_v \|V^{(v)} - V^*\|_F^2$$

$$s.t. \quad \forall 1 \le k \le K, \|U_{.,k}^{(v)}\|_1 = 1, U^{(v)}, V^{(v)}, V^* \ge 0$$

Multi-View Clustering via Joint NMF – SDM'13

Jialu Liu, Chi Wang, Jing Gao, and Jiawei Han

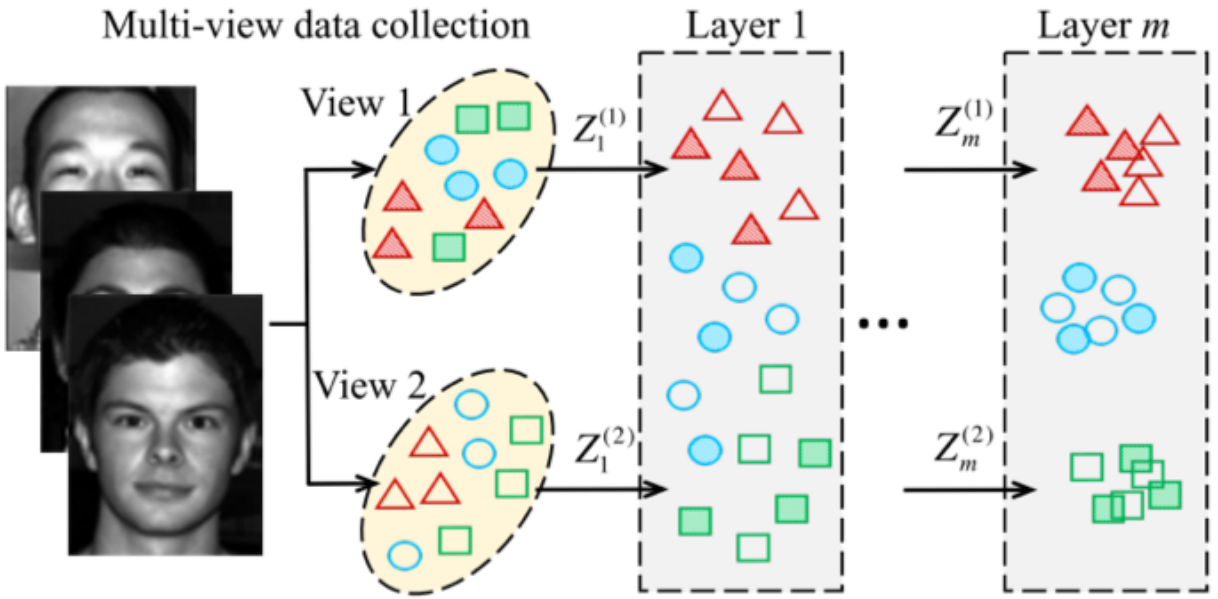
Methodology

- Multi-view Face Clustering

Northeastern University

Smile lab – Synergetic Media Learning Lab

Multi-View Clustering via Deep Matrix Factorization



Motivation:

To learn the hierarchical semantics of multi-view data in a layer-wise fashion, semi-nonnegative matrix factorization is adopted.

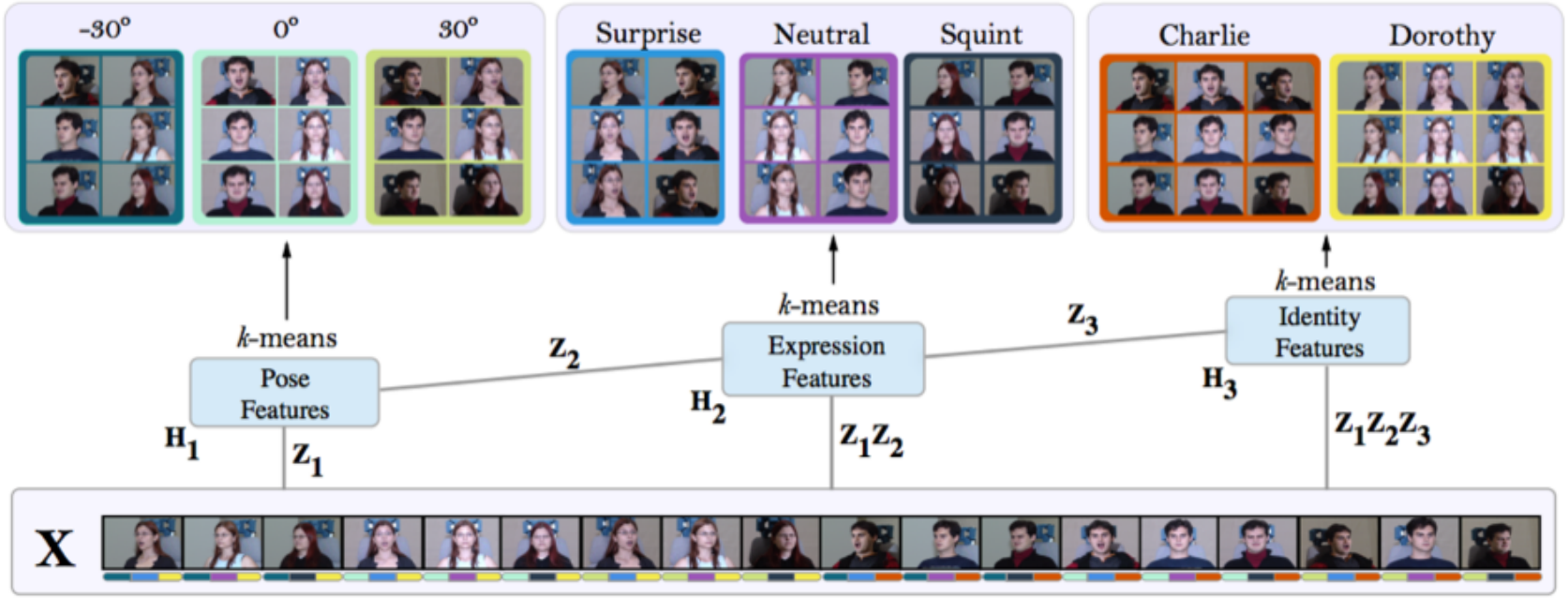
Methodology

- Multi-view Face Clustering

Northeastern University

Smile lab – Synergetic Media Learning Lab

Single-View Clustering via Deep Semi-NMF



Layer-wise formulation:

$$X^{\pm} \approx Z_1^{\pm} H_1^{+}$$
$$X^{\pm} \approx Z_1^{\pm} Z_2^{\pm} H_2^{+}$$
$$X^{\pm} \approx Z_1^{\pm} Z_2^{\pm} Z_3^{\pm} H_3^{+}$$

$$\Rightarrow \quad X^{\pm} \approx Z_1^{\pm} Z_2^{\pm} \cdots Z_m^{\pm} H_m^{+}$$

A Deep Semi-NMF Model for Learning Hidden Representations – ICML'14

George Trigeorgis, Konstantinos Bousmalis, Stefanos Zafeiriou, and Bjoern W.Schuller

Methodology

- Multi-view Face Clustering

Northeastern University

Smile lab – Synergetic Media Learning Lab

Multi-View Clustering via Deep Matrix Factorization

Objective function:

$$\min_{\substack{Z_i^{(v)}, H_i^{(v)}, \\ H_m, \alpha^{(v)}}} \sum_{v=1}^{V} (\alpha^{(v)})^{\gamma} \Big(\|X^{(v)} - Z_1^{(v)} Z_2^{(v)} \dots Z_m^{(v)} H_m\|_{\mathrm{F}}^2 + \beta \mathrm{tr}(H_m L^{(v)} H_m^{\mathrm{T}}) \Big)$$

$$\text{s.t. } H_i^{(v)} \geq 0,\ H_m \geq 0,\ \sum_{v=1}^{V} \alpha^{(v)} = 1,\ \alpha^{(v)} \geq 0$$

Decomposition on all views, where the representations on the last layer $H_m^{(v)}$ are forced to be same H_m.

Multi-View Clustering via Deep Matrix Factorization – AAAI'17

Handong Zhao, Zhengming Ding, and Yun Fu

Methodology

- Multi-view Face Clustering

Northeastern University

Smile lab – Synergetic Media Learning Lab

Multi-View Clustering via Deep Matrix Factorization

Objective function:

$$\min_{\substack{Z_i^{(v)}, H_i^{(v)}, \\ H_m, \alpha^{(v)}}} \sum_{v=1}^{V} (\alpha^{(v)})^{\gamma} \Big(\|X^{(v)} - Z_1^{(v)} Z_2^{(v)} \dots Z_m^{(v)} H_m\|_{\mathrm{F}}^2 + \beta \mathrm{tr}(H_m L^{(v)} H_m^{\mathrm{T}}) \Big)$$

Decomposition on all views, where the representations on the last layer $H_m^{(v)}$ are forced to be same H_m.

The hidden representation H are non-negative, with good interpretability.

$$\text{s.t. } H_i^{(v)} \geq 0,\ H_m \geq 0,\ \sum_{v=1}^{V} \alpha^{(v)} = 1,\ \alpha^{(v)} \geq 0$$

Multi-View Clustering via Deep Matrix Factorization – AAAI'17

Handong Zhao, Zhengming Ding, and Yun Fu

Methodology

- Multi-view Face Clustering

Northeastern University

Smile lab – Synergetic Media Learning Lab

Multi-View Clustering via Deep Matrix Factorization

Objective function:

$$\min_{\substack{Z_i^{(v)}, H_i^{(v)}, \\ H_m, \alpha^{(v)}}} \sum_{v=1}^{V} (\alpha^{(v)})^{\gamma} \Big(\|X^{(v)} - Z_1^{(v)} Z_2^{(v)} \dots Z_m^{(v)} H_m\|_{\mathrm{F}}^2 + \beta \mathrm{tr}(H_m L^{(v)} H_m^{\mathrm{T}}) \Big)$$

Decomposition on all views, where the representations on the last layer $H_m^{(v)}$ are forced to be same H_m.

$L^{(v)}$ is the graph Laplacian of the graph for view v, where each graph is constructed in k-nearest neighbor fashion.

The hidden representation H are non-negative, with good interpretability.

$$\text{s.t.} \; H_i^{(v)} \geq 0, \; H_m \geq 0, \; \sum_{v=1}^{V} \alpha^{(v)} = 1, \; \alpha^{(v)} \geq 0$$

Multi-View Clustering via Deep Matrix Factorization – AAAI'17

Handong Zhao, Zhengming Ding, and Yun Fu

Methodology

- Multi-view Face Clustering

Northeastern University

Smile lab – Synergetic Media Learning Lab

Multi-View Clustering via Deep Matrix Factorization

Multiple Features:
Intensity, LBP and Gabor

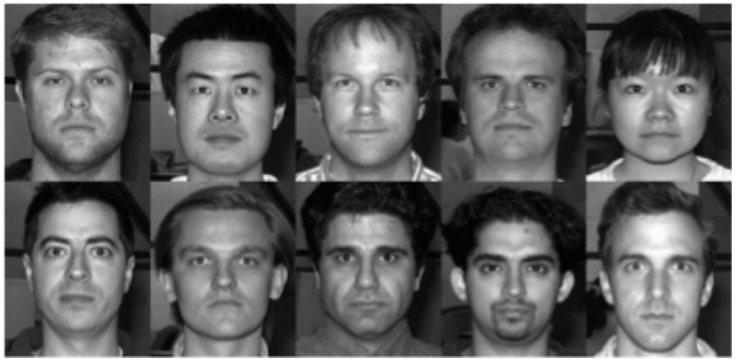


Table 1: Results (Mean ± standard deviation) on dataset Yale.

Method	NMI	ACC	AR	F-score	Precision	Recall
BestSV	0.654 ± 0.009	0.616 ± 0.030	0.440 ± 0.011	0.475 ± 0.011	0.457 ± 0.011	0.495 ± 0.010
ConcatFea	0.641 ± 0.006	0.544 ± 0.038	0.392 ± 0.009	0.431 ± 0.008	0.415 ± 0.007	0.448 ± 0.008
ConcatPCA	0.665 ± 0.037	0.578 ± 0.038	0.396 ± 0.011	0.434 ± 0.011	0.419 ± 0.012	0.450 ± 0.009
Co-Reg	0.648 ± 0.002	0.564 ± 0.000	0.436 ± 0.002	0.466 ± 0.000	0.455 ± 0.004	0.491 ± 0.003
Co-Train	0.672 ± 0.006	0.630 ± 0.001	0.452 ± 0.010	0.487 ± 0.009	0.470 ± 0.010	0.505 ± 0.007
Min-D	0.645 ± 0.005	0.615 ± 0.043	0.433 ± 0.006	0.470 ± 0.006	0.446 ± 0.005	0.496 ± 0.006
MultiNMF	0.690 ± 0.001	0.673 ± 0.001	0.495 ± 0.001	0.527 ± 0.000	0.512 ± 0.000	0.543 ± 0.000
NaMSC	0.671 ± 0.011	0.636 ± 0.000	0.475 ± 0.004	0.508 ± 0.007	0.492 ± 0.003	0.524 ± 0.004
DiMSC	0.727 ± 0.010	0.709 ± 0.003	0.535 ± 0.001	0.564 ± 0.002	0.543 ± 0.001	0.586 ± 0.003
Ours	**0.782** ± 0.010	**0.745** ± 0.011	**0.579** ± 0.002	**0.601** ± 0.002	**0.598** ± 0.001	**0.613** ± 0.002

Multi-View Clustering via Deep Matrix Factorization – AAAI'17

Handong Zhao, Zhengming Ding, and Yun Fu

Methodology

- Multi-view Face Clustering

Northeastern University

Smile lab – Synergetic Media Learning Lab

Multi-View Clustering via Deep Matrix Factorization

Multiple Features:
Intensity, LBP and Gabor

Table 3: Results (Mean ± standard deviation) on dataset Notting-Hill.

Method	NMI	ACC	AR	F-score	Precision	Recall
BestSV	0.723 ± 0.008	0.813 ± 0.000	0.712 ± 0.020	0.775 ± 0.015	0.774 ± 0.018	0.776 ± 0.013
ConcatFea	0.628 ± 0.028	0.673 ± 0.033	0.612 ± 0.041	0.696 ± 0.032	0.699 ± 0.032	0.693 ± 0.031
ConcatPCA	0.632 ± 0.009	0.733 ± 0.008	0.598 ± 0.015	0.685 ± 0.012	0.691 ± 0.010	0.680 ± 0.014
Co-Reg	0.660 ± 0.003	0.758 ± 0.000	0.616 ± 0.004	0.699 ± 0.000	0.705 ± 0.003	0.694 ± 0.003
Co-Train	0.766 ± 0.005	0.689 ± 0.027	0.589 ± 0.035	0.677 ± 0.026	0.688 ± 0.030	0.667 ± 0.023
Min-D	0.707 ± 0.003	0.791 ± 0.000	0.689 ± 0.002	0.758 ± 0.002	0.750 ± 0.002	0.765 ± 0.003
MultiNMF	0.752 ± 0.001	0.831 ± 0.001	0.762 ± 0.000	0.815 ± 0.000	0.804 ± 0.001	0.824 ± 0.001
NaMSC	0.730 ± 0.002	0.752 ± 0.013	0.666 ± 0.004	0.738 ± 0.005	0.746 ± 0.002	0.730 ± 0.011
DiMSC	**0.799** ± 0.001	0.843 ± 0.021	0.787 ± 0.001	0.834 ± 0.001	0.822 ± 0.005	0.836 ± 0.009
Ours	0.797 ± 0.005	**0.871** ± 0.009	**0.803** ± 0.002	**0.847** ± 0.002	**0.826** ± 0.007	**0.870** ± 0.001

Multi-View Clustering via Deep Matrix Factorization – AAAI'17

Handong Zhao, Zhengming Ding, and Yun Fu

Outline

Northeastern University

Smile lab

Synergetic Media Learning Lab

- Introduction & Background
  - Multi-view Face Task
  - Multi-view Face Data
- Unsupervised Multi-view Face Representation
  - Methodology
  - **Face Clustering, Outlier Detection**
- Supervised Multi-view Face Representation
  - Multi-view Learning
  - Transfer Learning
- Conclusion

Application

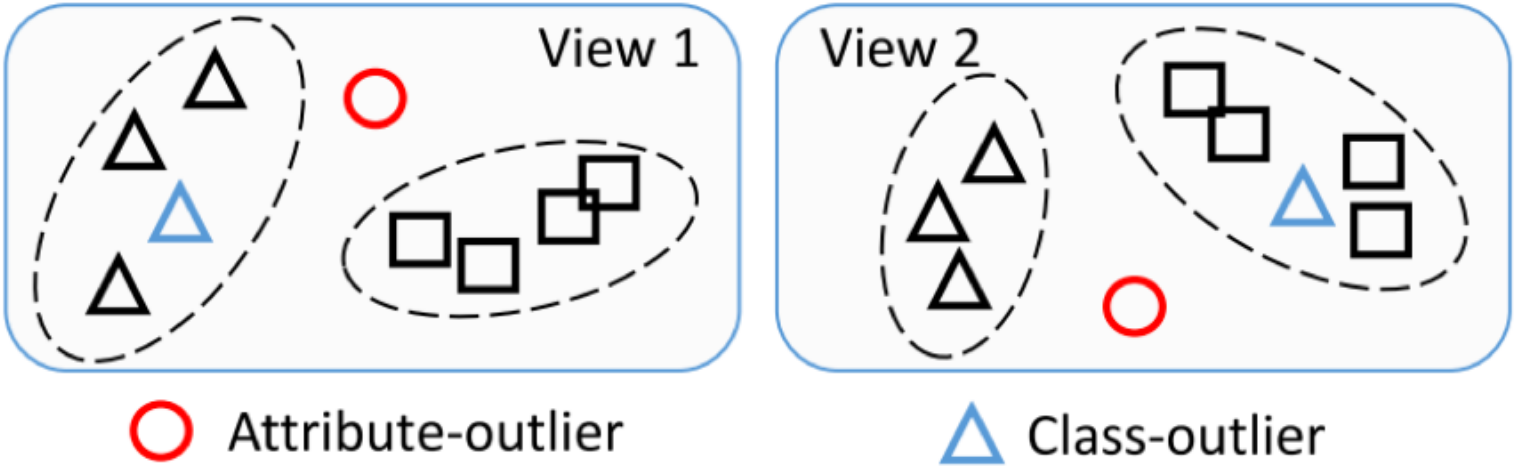
- Multi-view Face Clustering

Northeastern University

Smile lab – Synergetic Media Learning Lab

Extension 1: Outlier Detection

Outlier types in multi-view unsupervised learning scenario:



- **Attribute-outlier** is an outlier that exhibits consistent abnormal behaviors in each view, as the red circle in the figure. (Exist in traditional outlier detection.)
- **Class-outlier** is an outlier that exhibits inconsistent characteristics (e.g., cluster membership) across different views, as the blue triangle shown in the figure.

Application

- Multi-view Face Clustering

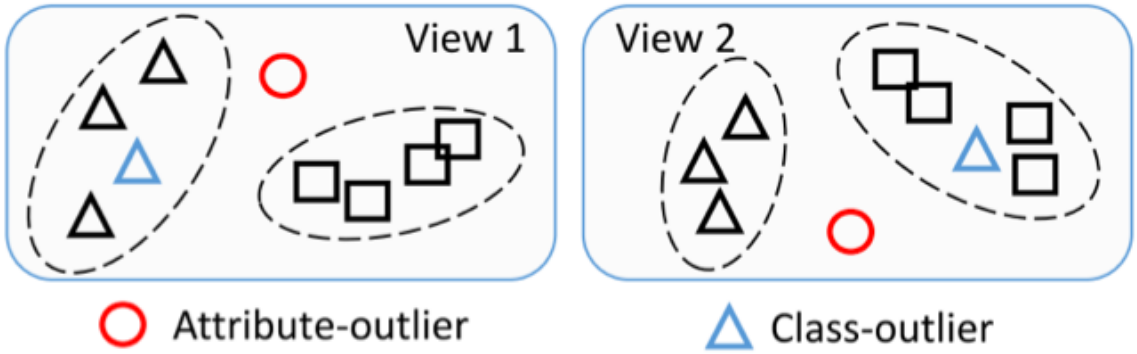
Northeastern University

Smile lab

Synergetic Media Learning Lab

Extension 1: Outlier Detection

Dual-regularized multi-view representation in outlier detection application



Objective function:

$$\min_{H^{(i)},G^{(i)},S^{(i)}} \sum_{i}^{V} \|S^{(i)}\|_{2,1} + \beta \sum_{i}^{V} \sum_{i \neq j}^{V} \|G^{(i)} - M_{ij} G^{(j)}\|_{\mathrm{F}}^{2}$$

$$\text{s.t. } X^{(i)} = H^{(i)} G^{(i)} + S^{(i)},$$

$$G_{kl} \in \{0, 1\}, \sum_{k=1}^{K} G_{kl} = 1, \forall l = 1, 2, \ldots, n$$

Dual-regularized Multi-view Outlier Detection – IJCAI'15

Handong Zhao and Yun Fu

Application

- Multi-view Face Clustering

Northeastern University

Smile lab – Synergetic Media Learning Lab

Dual-regularized multi-view representation in outlier detection application

$$\min_{H^{(i)},G^{(i)},S^{(i)}} \sum_{i}^{V} \|S^{(i)}\|_{2,1} + \beta \sum_{i}^{V} \sum_{i \neq j}^{V} \|G^{(i)} - M_{ij} G^{(j)}\|_{\mathrm{F}}^{2}$$
$$\text{s.t. } X^{(i)} = H^{(i)} G^{(i)} + S^{(i)},$$
$$G_{kl} \in \{0,1\}, \sum_{k=1}^{K} G_{kl} = 1, \forall l = 1,2,\ldots,n$$

Remark 1: *l-21* norm is defined as $\|S\|_{2,1} = \sum_{q=1}^{n} \sqrt{\sum_{p=1}^{d} |S_{pq}|^2}$
It has the power to ensure the matrix sparse in row, making it particularly suitable for sample-specific anomaly detection. This robust representation solves the outlier sensitivity problem in traditional *k*-means. Consequently, the regularization term $\|S^{(i)}\|_{2,1}$ is able to identify the attribute-outliers for *i*-th view.

Application

- Multi-view Face Clustering

Northeastern University

Smile lab – Synergetic Media Learning Lab

Dual-regularized multi-view representation in outlier detection application

$$\min_{H^{(i)},G^{(i)},S^{(i)}} \sum_{i}^{V} \|S^{(i)}\|_{2,1} + \beta \sum_{i}^{V} \sum_{i \neq j}^{V} \|G^{(i)} - M_{ij} G^{(j)}\|_{\mathrm{F}}^{2}$$

$$\text{s.t. } X^{(i)} = H^{(i)} G^{(i)} + S^{(i)},$$

$$G_{kl} \in \{0,1\}, \sum_{k=1}^{K} G_{kl} = 1, \forall l = 1,2,\ldots,n$$

Remark 1: *l-21* norm is defined as $\|S\|_{2,1} = \sum_{q=1}^{n} \sqrt{\sum_{p=1}^{d} |S_{pq}|^2}$ It has the power to ensure the matrix sparse in row, making it particularly suitable for sample-specific anomaly detection. This robust representation solves the outlier sensitivity problem in traditional *k*-means. Consequently, the regularization term $\|S^{(i)}\|_{2,1}$ is able to identify the attribute-outliers for *i*-th view.

Remark 2: In order to detect the class-outlier, i.e. sample inconsistent behavior across different views, the inconsistency with respect to each view needs to be measured. We argue that the following representation $\sum_{k=1}^{n} G_{kl}^{(i)} G_{kl}^{(j)}$ in latent space can well quantify the inconsistency of sample *l* across different views *i* and *j*.

Application

- Multi-view Face Clustering

Northeastern University

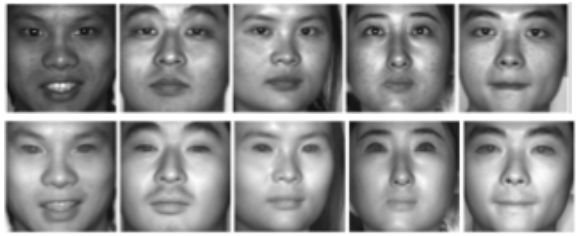
Smile lab — Synergetic Media Learning Lab

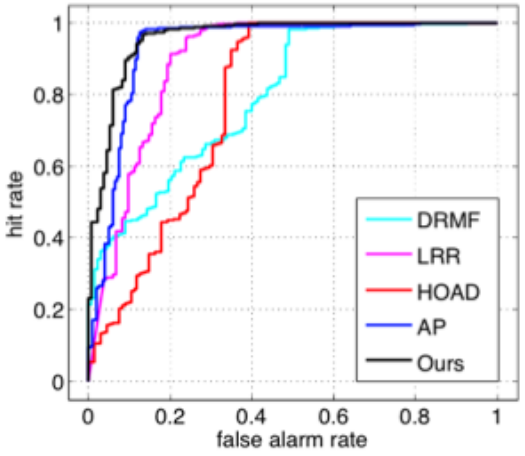
Experiment: (UCI dataset and BUAA VisNir)

Average AUC values on four UCI datasets with different settings. The setting is formatted as "DatasetName – Class-outlier Ratio (%) – Attribute-outlier Ratio (%)".

Datasets	DRMF [Xiong *et al.*, 2011]	LRR [Liu *et al.*, 2012]	HOAD [Gao *et al.*, 2013]	AP [Alvarez *et al.*, 2013]	DMOD (Ours)
iris-2-8	0.749 ± 0.044	0.779 ± 0.062	0.167 ± 0.057	0.326 ± 0.027	$\mathbf{0.868 \pm 0.036}$
iris-5-5	0.714 ± 0.038	0.762 ± 0.107	0.309 ± 0.062	0.630 ± 0.021	$\mathbf{0.865 \pm 0.047}$
iris-8-2	0.651 ± 0.037	0.740 ± 0.100	0.430 ± 0.055	0.840 ± 0.021	$\mathbf{0.882 \pm 0.043}$
breast-2-8	0.764 ± 0.013	0.586 ± 0.037	0.555 ± 0.072	0.293 ± 0.012	$\mathbf{0.816 \pm 0.038}$
breast-5-5	0.708 ± 0.034	0.493 ± 0.017	0.586 ± 0.061	0.532 ± 0.024	$\mathbf{0.809 \pm 0.020}$
breast-8-2	0.648 ± 0.024	0.508 ± 0.043	0.634 ± 0.046	0.693 ± 0.023	$\mathbf{0.778 \pm 0.019}$
ionosphere-2-8	0.705 ± 0.029	0.699 ± 0.025	0.446 ± 0.074	0.623 ± 0.033	$\mathbf{0.810 \pm 0.044}$
ionosphere-5-5	0.676 ± 0.040	0.627 ± 0.029	0.422 ± 0.051	0.761 ± 0.025	$\mathbf{0.773 \pm 0.041}$
ionosphere-8-2	0.634 ± 0.023	0.511 ± 0.014	0.448 ± 0.041	0.822 ± 0.030	$\mathbf{0.824 \pm 0.029}$
letter-2-8	0.315 ± 0.030	0.503 ± 0.011	0.536 ± 0.046	0.372 ± 0.057	$\mathbf{0.687 \pm 0.041}$
letter-5-5	0.375 ± 0.023	0.499 ± 0.012	0.663 ± 0.057	0.550 ± 0.043	$\mathbf{0.691 \pm 0.037}$
letter-8-2	0.490 ± 0.062	0.499 ± 0.016	0.569 ± 0.049	0.621 ± 0.051	$\mathbf{0.852 \pm 0.037}$

BUAA NirVis Dataset





ROC Curves of all the methods on BUAA VisNir database with both outlier levels of 5%.

We have the following observations:

- Our proposed method DMOD consistently outperforms all the other baselines in all settings.
- In most cases, single-view based methods have superior performance to multi-view based methods in outlier setting "DatasetName-2-8".
- In the experiments with setting "DatasetName-8-2", multi-view based methods perform better than single- view methods in most cases.

Dual-regularized Multi-view Outlier Detection – IJCAI'15

Handong Zhao and Yun Fu

Application

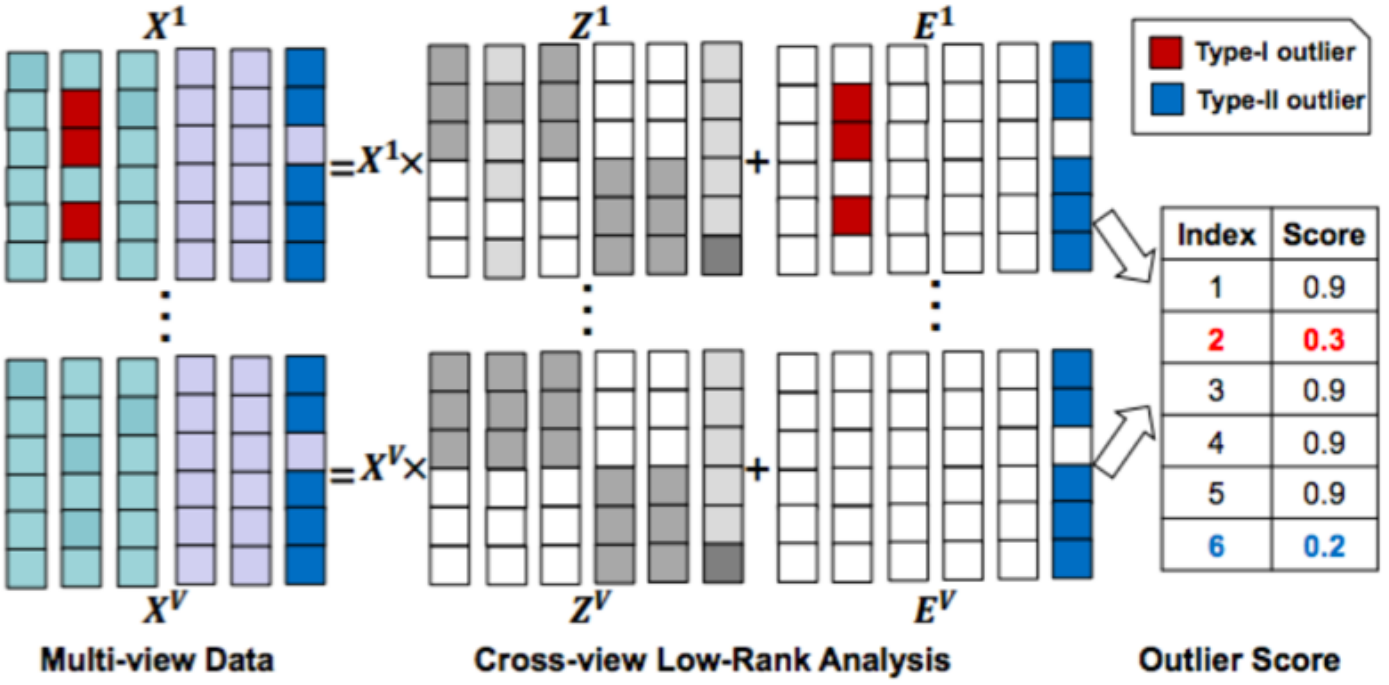
- Multi-view Face Clustering

Northeastern University

Smile lab – Synergetic Media Learning Lab

Low-rank based outlier detection.

-- Instead of modeling class-outlier by class labels (hard assignment), soft assignments with certain constraints, e.g. low-rank constraint, also work well.



Multi-View Low-Rank Analysis for Outlier Detection – SDM'15

Sheng Li, Ming Shao and Yun Fu

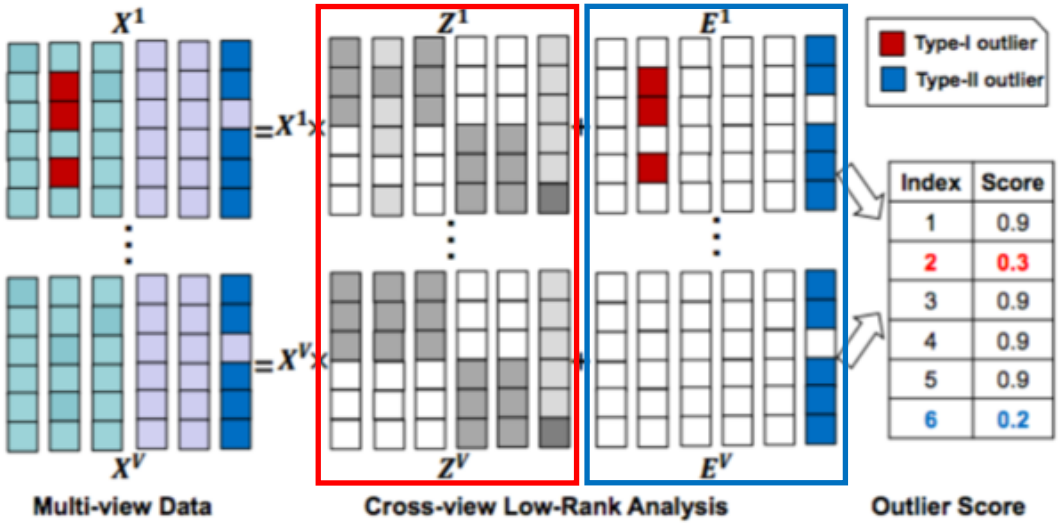
Application

- Multi-view Face Clustering

Northeastern University

Smile lab — Synergetic Media Learning Lab

Low-rank based outlier detection.



Objective function:

$$\min_{Z^{(v)},E^{(v)}} \sum_{v=1}^{2} \left(\|Z^{(v)}\|_* + \alpha \|E^{(v)}\|_{2,1} \right) + \beta \|Z^{(1)} - Z^{(2)}\|_{2,1}$$

$$\text{s.t.} \quad X^{(v)} = X^{(v)} Z^{(v)} + E^{(v)}, \; v = 1, 2,$$

Multi-View Low-Rank Analysis for Outlier Detection – SDM'15

Sheng Li, Ming Shao and Yun Fu

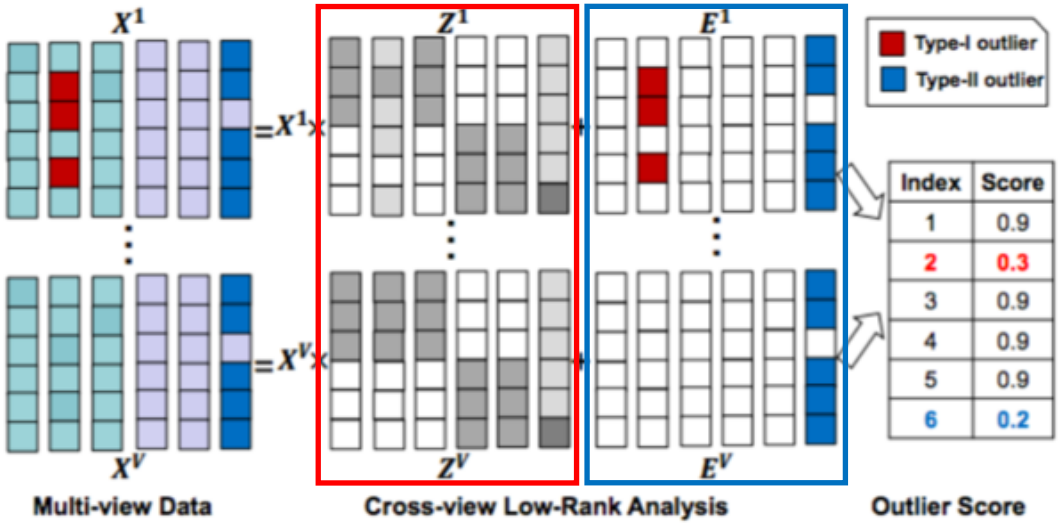
Application

- Multi-view Face Clustering

Northeastern University

Smile lab – Synergetic Media Learning Lab

Low-rank based outlier detection.



$\|\quad\|_*$ denotes trace norm, is a good surrogate of rank function.

Objective function:

$$\min_{Z^{(v)},E^{(v)}} \sum_{v=1}^{2} \left(\|Z^{(v)}\|_* + \alpha \|E^{(v)}\|_{2,1} \right) + \beta \|Z^{(1)} - Z^{(2)}\|_{2,1}$$

$$\text{s.t.} \quad X^{(v)} = X^{(v)} Z^{(v)} + E^{(v)}, \; v = 1, 2,$$

Multi-View Low-Rank Analysis for Outlier Detection – SDM'15

Sheng Li, Ming Shao and Yun Fu

Application

- Multi-view Face Clustering

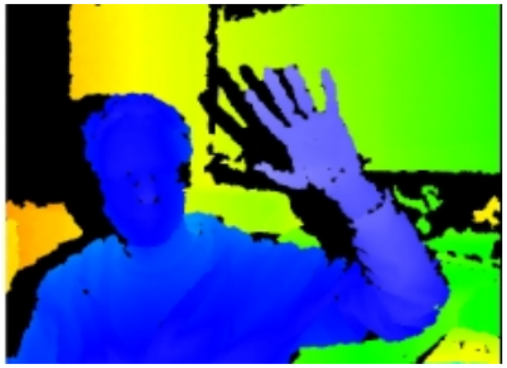
Northeastern University

Smile lab

Synergetic Media Learning Lab

Extension 2: Incomplete Scenario

When the data from one modality/more modalities are inaccessible because of sensor failure or other reasons, most traditional MVC methods would inevitably degenerate or even fail.



Incomplete Multi-Modal Visual Data Grouping – IJCAI'16

Handong Zhao, Hongfu Liu and Yun Fu

Application

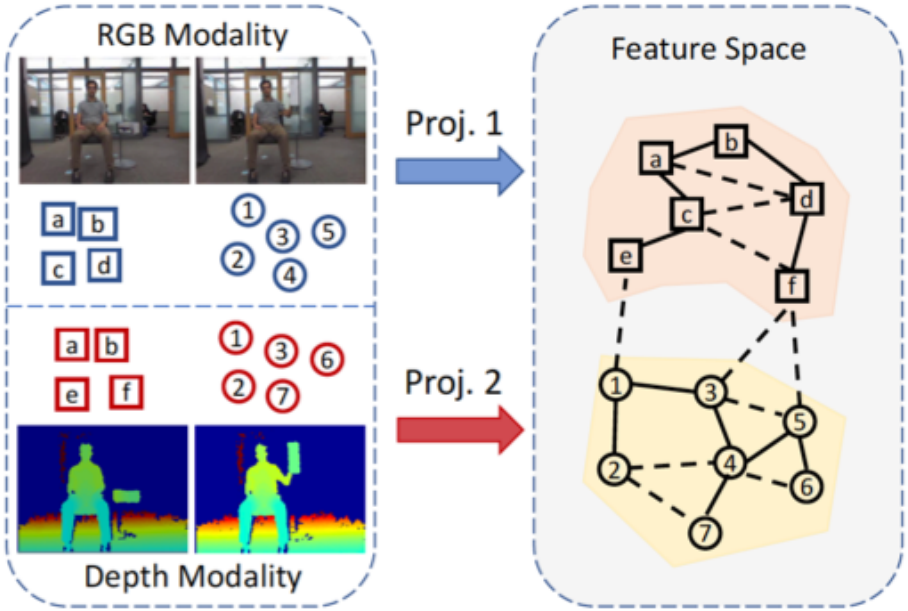
- Multi-view Face Clustering

Northeastern University

Smile lab — Synergetic Media Learning Lab

Extension 2: Incomplete Scenario
Incomplete Multi-Modal Visual Data Grouping

Motivation:



Objective function:

$$\min_{\substack{P_c,\hat{P}^{(1)},\hat{P}^{(2)}\\U^{(1)},U^{(2)},A}} \left\| \begin{bmatrix} X_c^{(1)} \\ \hat{X}^{(1)} \end{bmatrix} - \begin{bmatrix} P_c \\ \hat{P}^{(1)} \end{bmatrix} U^{(1)} \right\|_{\mathrm{F}}^2 +$$

$$\left\| \begin{bmatrix} X_c^{(2)} \\ \hat{X}^{(2)} \end{bmatrix} - \begin{bmatrix} P_c \\ \hat{P}^{(2)} \end{bmatrix} U^{(2)} \right\|_{\mathrm{F}}^2 + \mathcal{G}(P,A) + \mathcal{R}(U,A).$$

$$\text{s.t. } \forall i \; A_i^T \mathbf{1} = 1, \; A_i \succeq 0.$$

$$\mathcal{G}(P,A) = \beta \mathrm{tr}(P^{\mathrm{T}} L_A P),$$

$$\mathcal{R}(U,A) = \lambda(\|U^{(1)}\|_{\mathrm{F}}^2 + \|U^{(2)}\|_{\mathrm{F}}^2) + \gamma \|A\|_{\mathrm{F}}^2$$

Incomplete Multi-Modal Visual Data Grouping – IJCAI'16

Handong Zhao, Hongfu Liu and Yun Fu

Application

- Multi-view Face Clustering

Northeastern University

Smile lab – Synergetic Media Learning Lab

Extension 2: Incomplete Scenario
Incomplete Multi-Modal Visual Data Grouping

Objective function:

$$\min_{\substack{P_c,\hat{P}^{(1)},\hat{P}^{(2)}\\ U^{(1)},U^{(2)},A}} \left\| \begin{bmatrix} X_c^{(1)} \\ \hat{X}^{(1)} \end{bmatrix} - \begin{bmatrix} P_c \\ \hat{P}^{(1)} \end{bmatrix} U^{(1)} \right\|_{\mathrm{F}}^2 + \left\| \begin{bmatrix} X_c^{(2)} \\ \hat{X}^{(2)} \end{bmatrix} - \begin{bmatrix} P_c \\ \hat{P}^{(2)} \end{bmatrix} U^{(2)} \right\|_{\mathrm{F}}^2$$
$$+ \mathcal{G}(P, A) + \mathcal{R}(U, A)$$

$$\text{s.t. } \forall i \; A_i^T \mathbf{1} = 1, \; A_i \succeq 0.$$

$$\mathcal{G}(P, A) = \beta \mathrm{tr}(P^{\mathrm{T}} L_A P)$$

$$\mathcal{R}(U, A) = \lambda(\|U^{(1)}\|_{\mathrm{F}}^2 + \|U^{(2)}\|_{\mathrm{F}}^2) + \gamma \|A\|_{\mathrm{F}}^2$$

- Use the shared data $X_c^{(v)}$ in each view to learn the common representation P_c in low-dimensional subspace space.

Incomplete Multi-Modal Visual Data Grouping – IJCAI'16
Handong Zhao, Hongfu Liu and Yun Fu

Application

- Multi-view Face Clustering

Northeastern University

Smile lab – Synergetic Media Learning Lab

Extension 2: Incomplete Scenario
Incomplete Multi-Modal Visual Data Grouping

Objective function:

$$\min_{\substack{P_c,\hat{P}^{(1)},\hat{P}^{(2)}\\ U^{(1)},U^{(2)},A}} \left\| \begin{bmatrix} X_c^{(1)} \\ \hat{X}^{(1)} \end{bmatrix} - \begin{bmatrix} P_c \\ \hat{P}^{(1)} \end{bmatrix} U^{(1)} \right\|_{\mathrm{F}}^2 + \left\| \begin{bmatrix} X_c^{(2)} \\ \hat{X}^{(2)} \end{bmatrix} - \begin{bmatrix} P_c \\ \hat{P}^{(2)} \end{bmatrix} U^{(2)} \right\|_{\mathrm{F}}^2$$
$$+ \mathcal{G}(P, A) + \mathcal{R}(U, A)$$

$$\text{s.t. } \forall i \; A_i^T \mathbf{1} = 1, \; A_i \succeq 0.$$

$$\mathcal{G}(P, A) = \beta \mathrm{tr}(P^{\mathrm{T}} L_A P)$$

$$\mathcal{R}(U, A) = \lambda(\|U^{(1)}\|_{\mathrm{F}}^2 + \|U^{(2)}\|_{\mathrm{F}}^2) + \gamma \|A\|_{\mathrm{F}}^2$$

- Regularizers to prevent trivial solution. Here we choose a simple Frobenius norm. Its alternatives include ℓ_1 and others.

Incomplete Multi-Modal Visual Data Grouping – IJCAI'16
Handong Zhao, Hongfu Liu and Yun Fu

Application

- Multi-view Face Clustering

Northeastern University

Smile lab – Synergetic Media Learning Lab

Extension 2: Incomplete Scenario
Incomplete Multi-Modal Visual Data Grouping

Objective function:

$$\min_{\substack{P_c,\hat{P}^{(1)},\hat{P}^{(2)}\\ U^{(1)},U^{(2)},A}} \left\| \begin{bmatrix} X_c^{(1)} \\ \hat{X}^{(1)} \end{bmatrix} - \begin{bmatrix} P_c \\ \hat{P}^{(1)} \end{bmatrix} U^{(1)} \right\|_{\mathrm{F}}^2 + \left\| \begin{bmatrix} X_c^{(2)} \\ \hat{X}^{(2)} \end{bmatrix} - \begin{bmatrix} P_c \\ \hat{P}^{(2)} \end{bmatrix} U^{(2)} \right\|_{\mathrm{F}}^2$$

$$+ \mathcal{G}(P, A) + \mathcal{R}(U, A)$$

$$\text{s.t. } \forall i \; A_i^T \mathbf{1} = 1, \; A_i \succeq 0.$$

$$\mathcal{G}(P, A) = \beta \mathrm{tr}(P^{\mathrm{T}} L_A P)$$

$$\mathcal{R}(U, A) = \lambda(\|U^{(1)}\|_{\mathrm{F}}^2 + \|U^{(2)}\|_{\mathrm{F}}^2) + \gamma \|A\|_{\mathrm{F}}^2$$

- $P = [P_c; \hat{P}^{(1)}; \hat{P}^{(2)}]$
- Graph Laplacian term to preserve locality information, where L_A is the Laplacian matrix of similarity matrix A, which is learned on the latent representation P.

Application

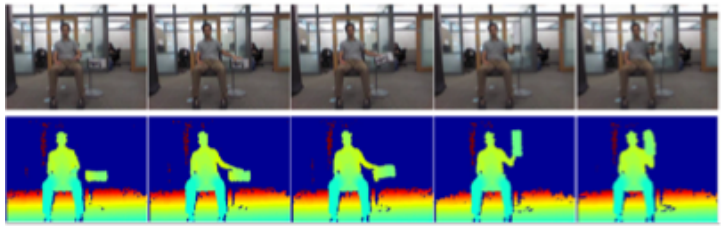
- Multi-view Face Clustering

Northeastern University

Smile lab — Synergetic Media Learning Lab

Results:

MSR Action Pairs Dataset



RGB

Depth

Table 2: NMI/Precision results on MSR Action Pairs dataset under different PER settings.

Method \ PER	0.1	0.3	0.5	0.7	0.9
BSV	0.4807 / 0.2687	0.4807 / 0.2687	0.3691 / 0.1660	0.2874 / 0.1190	0.2779 / 0.1085
Concat	0.6270 / 0.3538	0.5803 / 0.3306	0.5512 / 0.3030	0.5123 / 0.2750	0.4685 / 0.2268
MultiNMF	0.6033 / 0.4038	0.5149 / 0.2984	0.5008 / 0.2828	0.4816 / 0.2539	0.4463 / 0.2267
PVC	**0.6917** / 0.4490	0.6501 / 0.3998	0.6356 / 0.3734	0.6012 / 0.3662	0.5882 / 0.3629
Ours	0.6859 / **0.4504**	**0.6763 / 0.4431**	**0.6504 / 0.3836**	**0.6468 / 0.3774**	**0.6396 / 0.3734**

Table 3: NMI/Precision results on MSR Daily Activity dataset under different PER settings.

Method \ PER	0.1	0.3	0.5	0.7	0.9
BSV	0.2012 / 0.0826	0.1851 / 0.0765	0.1683 / 0.0680	0.1487 / 0.0641	0.1328 / 0.0626
Concat	0.2499 / 0.1137	0.2354 / 0.0997	0.2261 / 0.0843	0.2031 / 0.0755	0.1878 / 0.0758
MultiNMF	0.2077 / 0.0841	0.2057 / 0.0911	0.1924 / 0.0806	0.1823 / 0.0713	0.1655 / 0.0674
PVC	0.2605 / 0.1385	0.2487 / **0.1275**	0.2236 / 0.1086	0.2175 / 0.1049	0.2062 / 0.0902
Ours	**0.2807 / 0.1489**	**0.2554** / 0.1263	**0.2512 / 0.1241**	**0.2421 / 0.1108**	**0.2201 / 0.0907**

Incomplete Multi-Modal Visual Data Grouping – IJCAI'16

Handong Zhao, Hongfu Liu and Yun Fu

Application

- Multi-view Face Clustering

Northeastern University

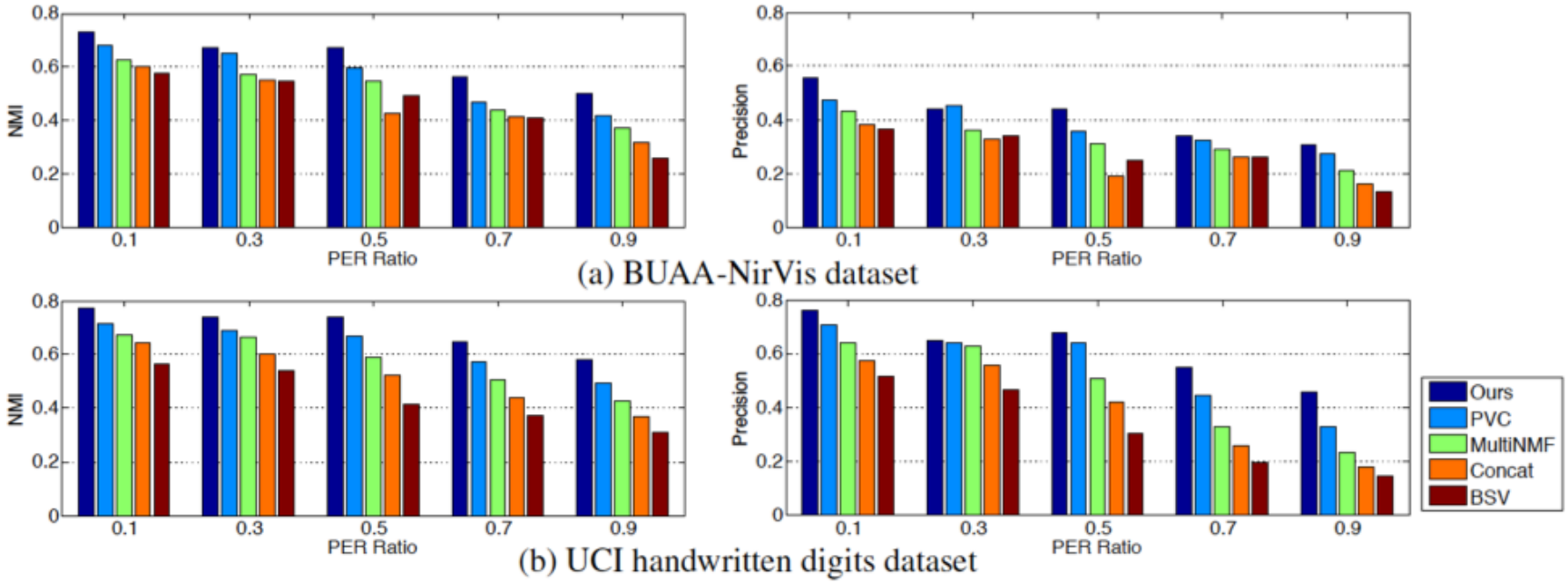
Smile lab – Synergetic Media Learning Lab

Results:

BUAA NirVis Dataset

Near Infrared

Gray-level intensity



(a) BUAA-NirVis dataset

(b) UCI handwritten digits dataset

Incomplete Multi-Modal Visual Data Grouping – IJCAI'16

Handong Zhao, Hongfu Liu and Yun Fu

Outline

Northeastern University

Smile lab

Synergetic Media Learning Lab

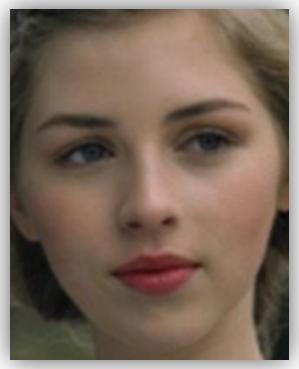
- **Introduction & Background**
  - Multi-view Face Task
  - Multi-view Face Data
- **Unsupervised Multi-view Face Representation**
  - Methodology
  - Face Clustering, Outlier Detection
- **Supervised Multi-view Face Representation**
  - *Multi-view Learning*
  - Transfer Learning
- **Conclusion**

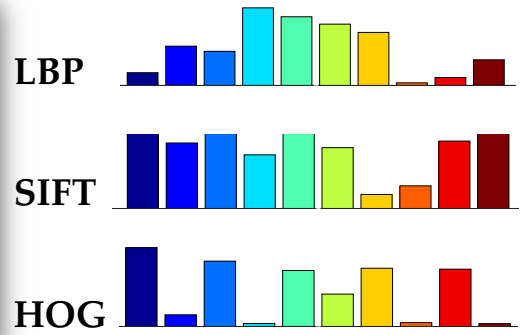
Multi-view Face Data

Northeastern University

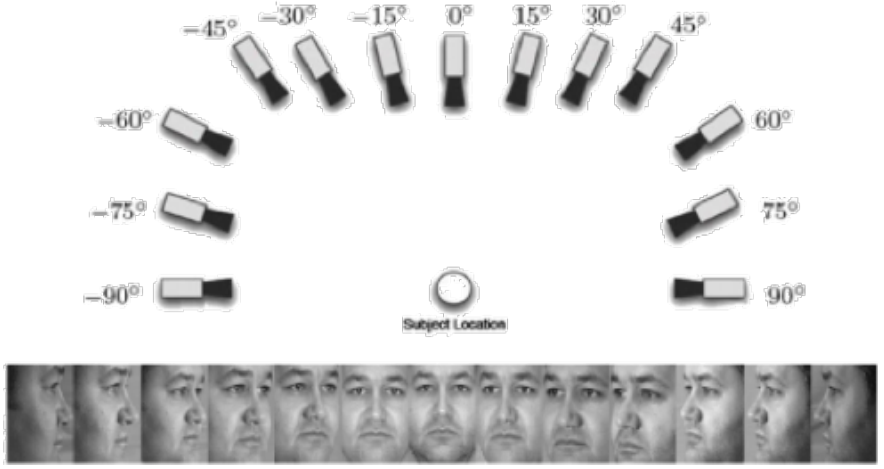
Smile lab — Synergetic Media Learning Lab

❑ **Data Correspondence [Only One Factor]→ *Heterogeneous***

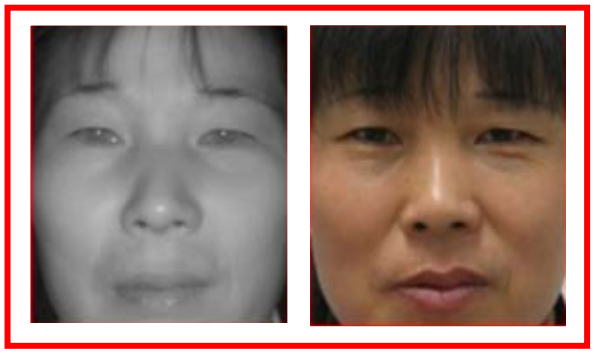




Multiple features



❑ **No Data Correspondence [Multiple Factors]**



expression

expression

occlusion

Supervised Multi-View Face

Northeastern University

Smile lab

Synergetic Media Learning Lab

Task with Labels: Face Recognition, Age Estimation, Kinship Verification…

❑ **Category 1 [Data Correspondence] (Multi-view Learning)**

- **Multiple Features,** *e.g., LBP, SIFT, HOG…*
  Goal: fuse various knowledge from multiple features
- **Multi-Pose/Multi-Modal Face**
  Goal: seek a view-invariant space to mitigate the view divergence

❑ **Category 2 [No Data Correspondence] (Transfer learning)**

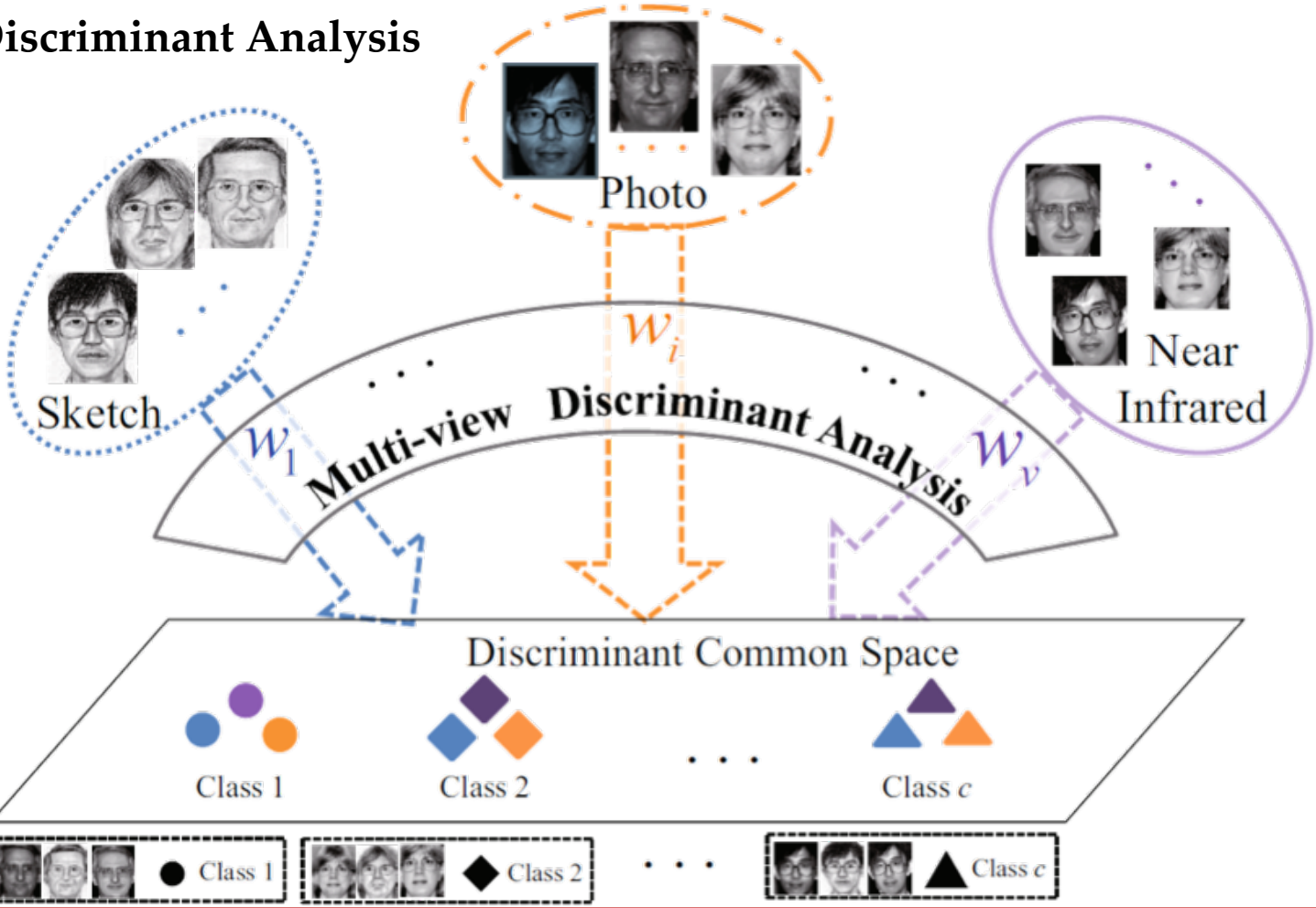
- **Multiple Feature/Multi-Pose/Multi-Modal Face**
  Goal: transfer knowledge from well-labeled sources to unlabeled targets

Supervised Multi-View Face Representation
[*Multi-view Learning*]

Northeastern University

Smile lab
Synergetic Media Learning Lab

Multi-view Discriminant Analysis



Multi-view Discriminant Analysis. –ECCV 2012 & IEEE TPAMI 2016

Meina Kan, Shiguang Shan, Haihong Zhang, Shihong Lao, and Xilin Chen

Supervised Multi-View Face Representation
[*Multi-view Learning*]

Northeastern University

Smile lab – Synergetic Media Learning Lab

Objective Function - *I*

Between-class scatter matrix

$$\mathbf{S}_B^y = \begin{bmatrix}\mathbf{w}_1^T \mathbf{w}_2^T \cdots \mathbf{w}_v^T\end{bmatrix} \begin{pmatrix}\mathbf{D}_{11} & \cdots & \mathbf{D}_{1v} \\ \vdots & \vdots & \vdots \\ \mathbf{D}_{v1} & \cdots & \mathbf{D}_{vv}\end{pmatrix} \begin{bmatrix}\mathbf{w}_1 \\ \mathbf{w}_2 \\ \vdots \\ \mathbf{w}_v\end{bmatrix} = \mathbf{W}^T\mathbf{D}\mathbf{W}$$

$$(\mathbf{w}_1^*, \mathbf{w}_2^*, \ldots, \mathbf{w}_v^*) = \arg \max_{\mathbf{w}_1,\cdots,\mathbf{w}_v} \frac{Tr(\mathbf{S}_B^y)}{Tr(\mathbf{S}_W^y)}$$

Within-class scatter matrix

$$\mathbf{S}_W^y = \begin{bmatrix}\mathbf{w}_1^T \mathbf{w}_2^T \cdots \mathbf{w}_v^T\end{bmatrix} \begin{pmatrix}\mathbf{S}_{11} & \cdots & \mathbf{S}_{1v} \\ \vdots & \vdots & \vdots \\ \mathbf{S}_{v1} & \cdots & \mathbf{S}_{vv}\end{pmatrix} \begin{bmatrix}\mathbf{w}_1 \\ \mathbf{w}_2 \\ \vdots \\ \mathbf{w}_v\end{bmatrix} = \mathbf{W}^T\mathbf{S}\mathbf{W}$$

Multi-view Discriminant Analysis – ECCV 2012 & IEEE TPAMI 2016

Meina Kan, Shiguang Shan, Haihong Zhang, Shihong Lao, and Xilin Chen

Supervised Multi-View Face Representation
[*Multi-view Learning*]

Northeastern University

Smile lab — Synergetic Media Learning Lab

Objective Function - *II*

view-consistent regularizer

Projection & Data $\mathbf{w}_i = \mathbf{X}_i \boldsymbol{\beta}_i$

$$(\mathbf{w}_1^*, \mathbf{w}_2^*, \ldots, \mathbf{w}_v^*)$$

$$= \arg\max_{\mathbf{w}_1,\ldots,\mathbf{w}_v} \frac{Tr(\mathbf{W}^T\mathbf{D}\mathbf{W})}{Tr(\mathbf{W}^T\mathbf{S}\mathbf{W}) + \lambda \sum_{i,j=1}^{v} \|\boldsymbol{\beta}_i - \boldsymbol{\beta}_j\|_2^2}$$

$$= \arg\max_{\mathbf{w}_1,\ldots,\mathbf{w}_v} \frac{Tr(\mathbf{W}^T\mathbf{D}\mathbf{W})}{Tr(\mathbf{W}^T(\mathbf{S}+\lambda\mathbf{M})\mathbf{W})}$$

$$\mathbf{X}_1 = \mathbf{R}\mathbf{X}_2$$
$$\mathbf{w}_1 = \mathbf{R}\mathbf{w}_2$$

$$\mathbf{X}_1\boldsymbol{\beta}_1 = \mathbf{R}\mathbf{X}_2\boldsymbol{\beta}_2 = \mathbf{X}_1\boldsymbol{\beta}_2$$

Results on CUFSF and HFB Datasets

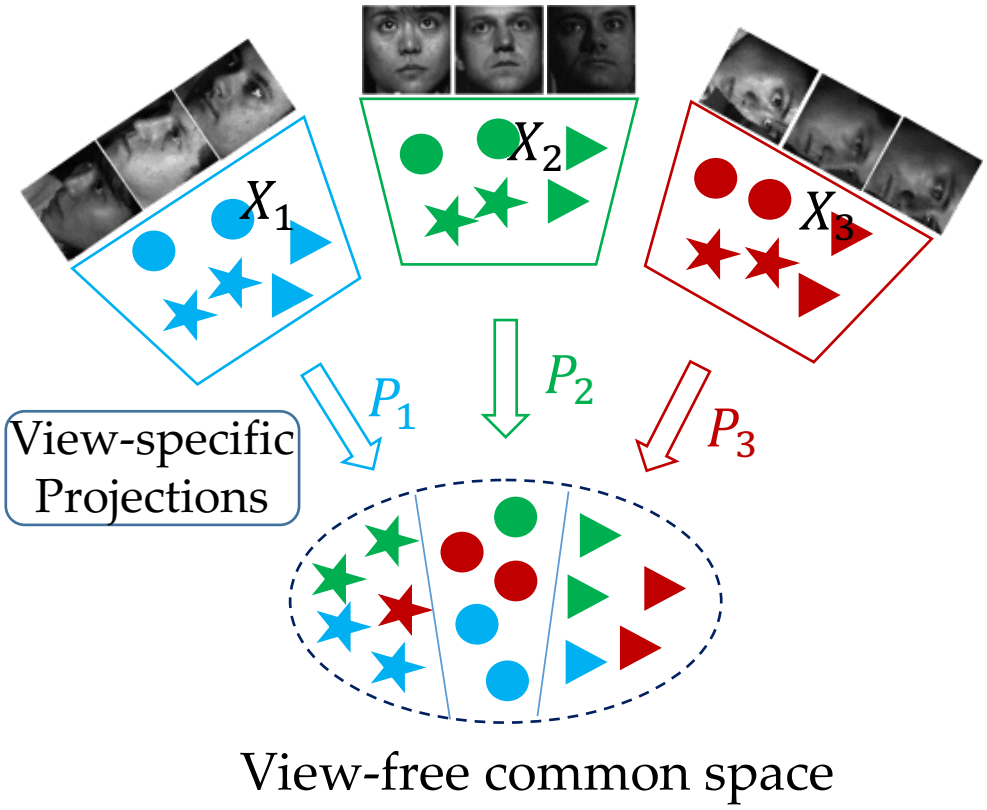
		CCA[4]**	CCA[4]+LDA	CDFE[19]	CSR[21]	PLS[6]	U-LDA[31]	GMA[28]	MvDA	MvDA-VC
CUFSF	Photo-Sketch	45.5%	45.0%	45.6%	50.2%	48.6%	46.8%	-	**53.4%**	**56.3%**
	Sketch-Photo	47.5%	50.6%	47.6%	49.0%	51.0%	53.4%	-	**55.5%**	**61.5%**
HFB	NIR-VIS	36.7%	40.0%	40.8%	26.7%	38.3%	39.1%	47.5%	**53.3%**	**59.2%**
	VIS-NIR	30.0%	40.0%	36.7%	32.5%	40.8%	40.0%	45.0%	**50.0%**	**59.2%**

Multi-view Discriminant Analysis. –ECCV 2012 & IEEE TPAMI 2016

Meina Kan, Shiguang Shan, Haihong Zhang, Shihong Lao, and Xilin Chen

Supervised Multi-View Face Representation
[*Multi-view Learning*]

Northeastern University
Smile lab
Synergetic Media Learning Lab



Previous work

Training:

Learn multiple view-specific transformations

Testing:

Project multi-view testing data via specific P_i

How about probe data with **unknown view information,**

which projection can be applied?

Low-rank Common Subspace – ICDM 2014 / IEEE TNNLS 2017

Zhengming Ding and Yun Fu

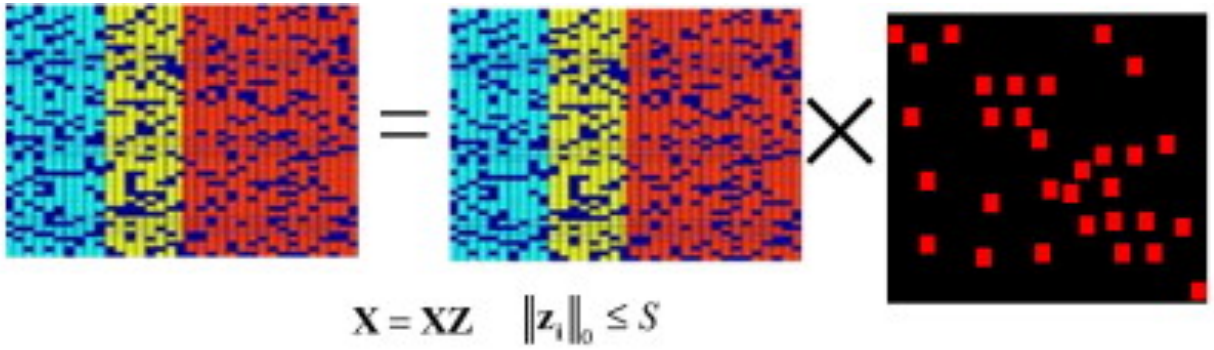
Supervised Multi-View Face Representation
[*Multi-view Learning*]

Northeastern University

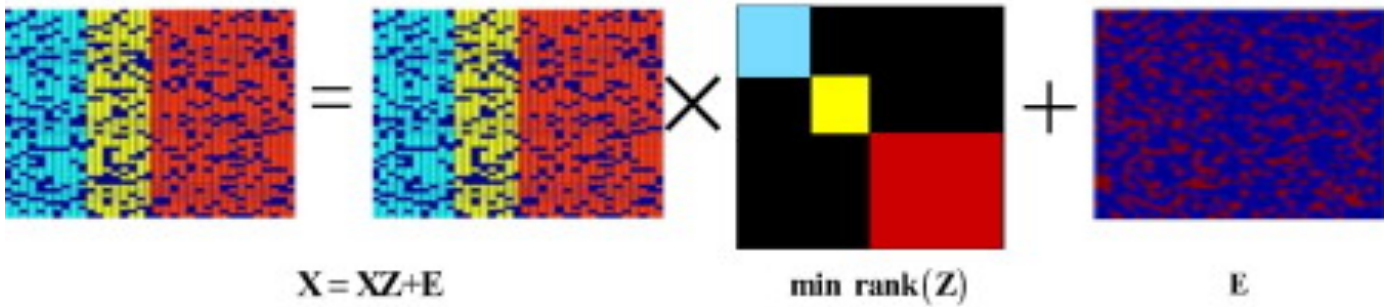
Smile lab – Synergetic Media Learning Lab

Low-rank Revisit

$$\min_{Z,E} ||Z||_* + \lambda ||E||_{2,1}, \ \text{s.t.}, \ X = XZ + E,$$



(a) Sparse representation



(b) Low-rank representation

Robust subspace segmentation by low-rank representation – ICML 2010

Guangcan Liu, Zhouchen Lin and Yong Yu

Supervised Multi-View Face Representation
[*Multi-view Learning*]

Northeastern University

Smile lab
Synergetic Media Learning Lab

Low-rank Revisit

$$\min_{Z,E} ||Z||_* + \lambda ||E||_{2,1}, \ \text{s.t.}, \ X = XZ + E,$$

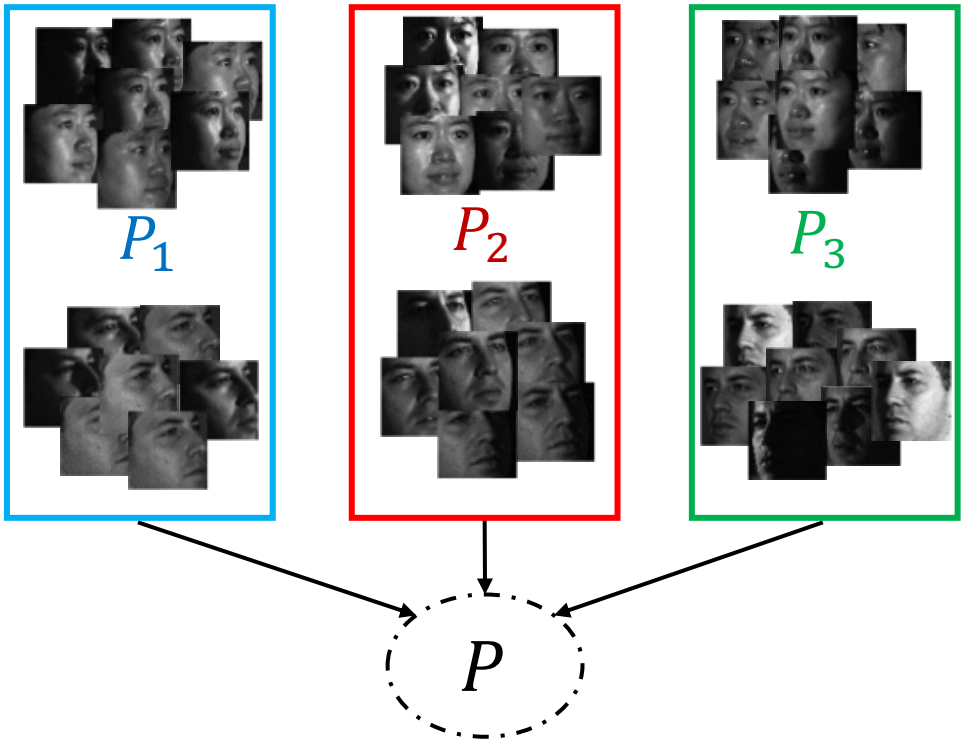
Robust subspace segmentation by low-rank representation – ICML 2010

Guangcan Liu, Zhouchen Lin and Yong Yu

Supervised Multi-View Face Representation
[*Multi-view Learning*]

Northeastern University

Smile lab – Synergetic Media Learning Lab



Low-rank Common Projection

Data have low-rank structure

$$P_i = X_i \beta_i$$

Projections also have low-rank structure

$$\min_{P,E_i,P_i} \boxed{\text{rank}(P)} + \lambda_0 \sum_{i=1}^{k} \|E_i\|_1$$
$$\text{s.t.} \quad P_i = P + E_i, \quad i = 1, \cdots, k,$$

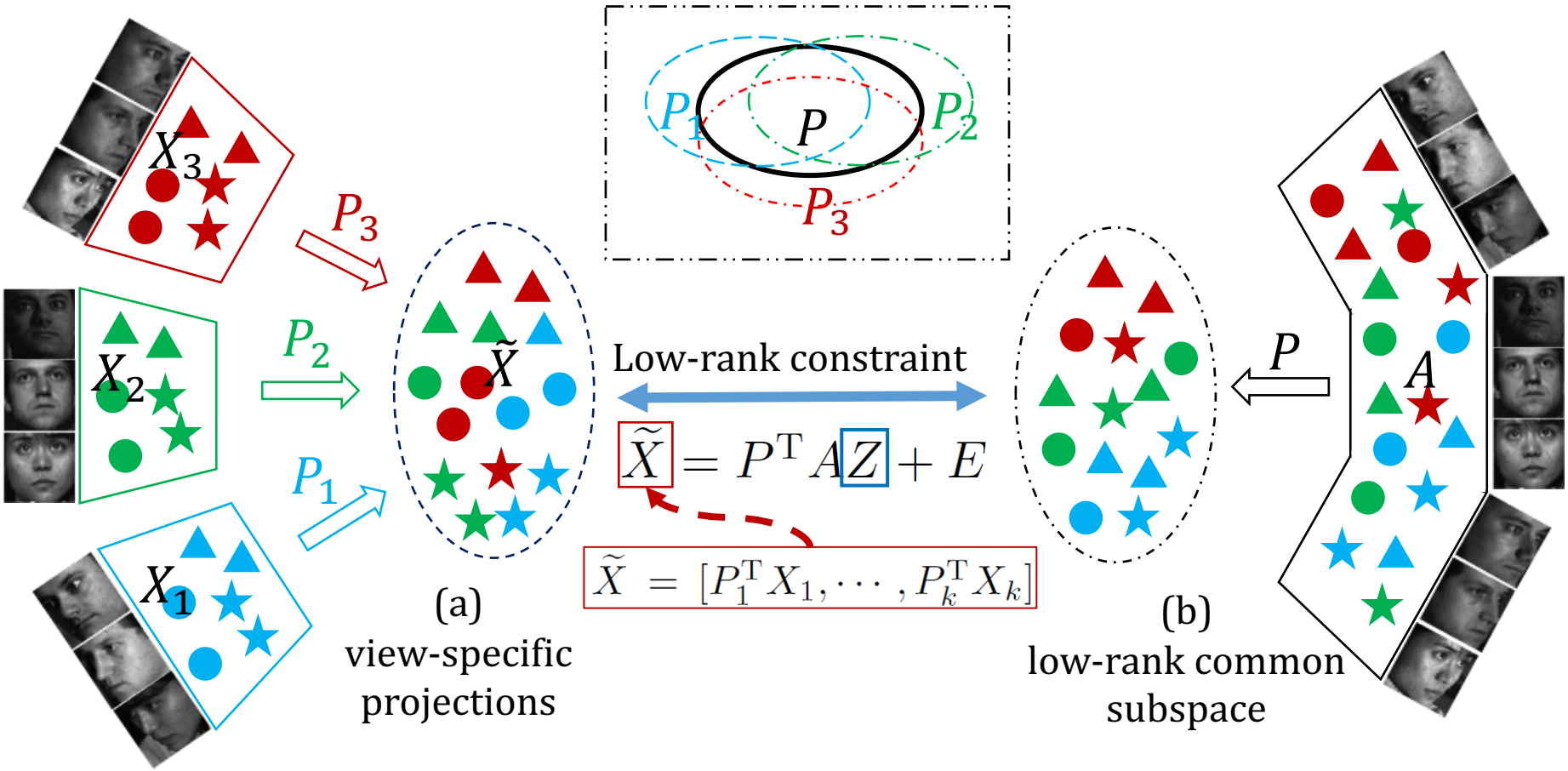
Low-rank Common Subspace – ICDM 2014 / IEEE TNNLS 2017

Zhengming Ding and Yun Fu

Supervised Multi-View Face Representation
[*Multi-view Learning*]

Northeastern University

Smile lab — Synergetic Media Learning Lab



Low-rank Common Subspace – ICDM 2014 / IEEE TNNLS 2017

Zhengming Ding and Yun Fu

Supervised Multi-View Face Representation
[*Multi-view Learning*]

Northeastern University

Smile lab — Synergetic Media Learning Lab

Model *I*

$$\min_{P,Z,E,E_i,P_i} \|Z\|_* + \|P\|_* + \lambda_0 \sum_{i=1}^{k} \|E_i\|_1 + \lambda_1 \|E\|_{2,1}$$

$$\text{s.t. } \widetilde{X} = P^{\mathrm{T}} A Z + E, \quad P^{\mathrm{T}} P = \mathrm{I},$$

$$P_i = P + E_i, \quad i = 1, \cdots, k.$$

$$\widetilde{X} = [P_1^{\mathrm{T}} X_1, \cdots, P_k^{\mathrm{T}} X_k]$$

Cross-view Alignment

Model *II*

$$\min_{\substack{\mathrm{P},\mathrm{Z_i},\mathrm{E_i},\\ \mathrm{S_i},\mathrm{P_i}}} \sum_{i=1}^{k} \left(\|\mathrm{Z_i}\|_* + \lambda_0 \|\mathrm{S_i}\|_1 + \lambda_1 \|\mathrm{E_i}\|_{2,1} \right) + \lambda_2 \Omega(\mathrm{P}, \mathrm{Z})$$

$$\text{s.t. } \mathrm{P_i^\top X_i} = \mathrm{P^\top A Z_i} + \mathrm{E_i}, \ \mathrm{P_i} = \mathrm{P} + \mathrm{S_i},$$

$$i = 1, \cdots, k, \quad \mathrm{P^\top P} = \mathrm{I}_p,$$

Differences:

- Multi-task Strategy
- Unsupervised → Supervised
- Projection Optimization

$$\begin{aligned} &\mathrm{tr}(\mathcal{S}_w) - \mathrm{tr}(\mathcal{S}_b) + \eta \|\mathrm{P^\top DZ}\|_{\mathrm{F}}^2 \\ =& \mathrm{tr}\left((\mathrm{P^\top DZ})(\mathrm{I}_{km} - \mathcal{L}_w)(\mathrm{P^\top DZ})^\top \right) \\ &- \mathrm{tr}\left((\mathrm{P^\top DZ}) \mathcal{L}_b (\mathrm{P^\top DZ})^\top \right) + \eta \|\mathrm{P^\top DZ}\|_{\mathrm{F}}^2 \\ =& \mathrm{tr}\left((\mathrm{P^\top DZ})((1+\eta)\mathrm{I}_{km} - \mathcal{L}_w - \mathcal{L}_b)(\mathrm{P^\top DZ})^\top \right) \end{aligned}$$

Low-rank Common Subspace – ICDM 2014 / IEEE TNNLS 2017

Zhengming Ding and Yun Fu

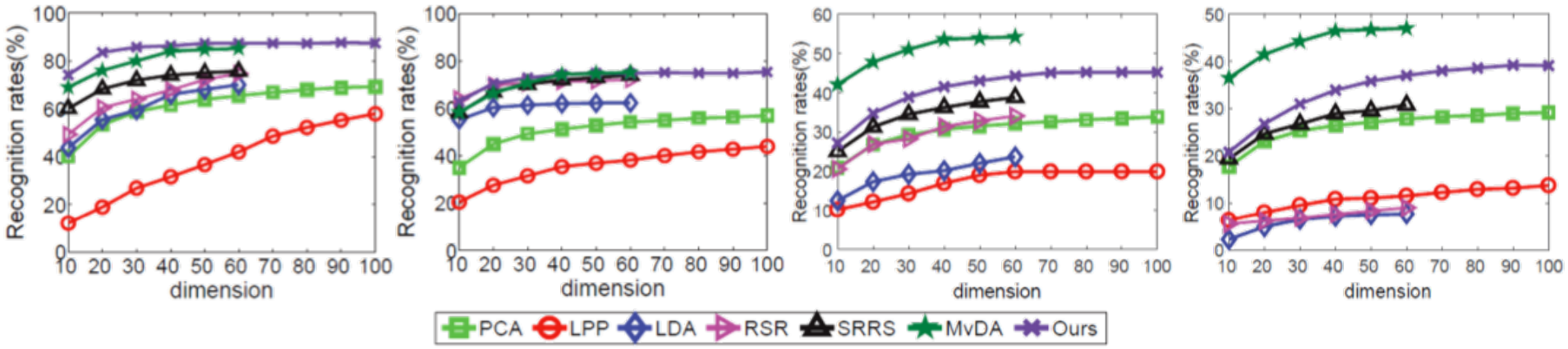
Supervised Multi-View Face Representation
[*Multi-view Learning*]

Northeastern University

Smile lab — Synergetic Media Learning Lab

Results on CMU PIE

	PCA[40]	LDA[41]	LPP[42]	RSR[43]	TFRR[44]	SRRS[25]	LRCS [12]	MvDA[13]	RMSL[4]	Ours
Case 1	69.03±0.08	70.46±0.05	57.25±0.06	77.51±0.01	77.92±0.03	78.27±0.04	87.78±0.22	85.23±0.05	**88.15±0.06**	87.10±0.07
Case 2	69.21±0.08	71.32±0.02	58.83±0.07	74.74±0.17	76.24±0.12	78.74±0.23	86.67±0.09	85.81±0.09	**87.05±0.07**	86.70±0.12
Case 3	68.52±0.12	63.51±0.75	59.25±0.56	71.10±0.04	75.29±0.07	77.45±0.02	87.38±0.39	86.12±0.12	87.40±0.17	**87.48±0.10**
Case 4	52.65±0.04	56.53±0.02	43.56±0.08	67.57±0.01	69.74±0.05	71.44±0.03	74.84±0.04	**75.36±0.18**	75.16±0.12	72.38±0.09
Case 5	34.94±0.08	24.07±0.25	19.67±0.05	29.72±0.01	33.91±0.12	38.86±0.02	44.48±0.03	**54.13±0.16**	44.93±0.11	45.88±0.09
Case 6	29.09±0.01	07.06±0.01	13.11±0.01	09.44±0.02	28.36±0.04	30.16±0.02	36.17±0.11	**47.67±0.18**	37.14±0.08	38.28±0.11



Low-rank Common Subspace – ICDM 2014 / IEEE TNNLS 2017

Zhengming Ding and Yun Fu

Supervised Multi-View Face Representation
[*Multi-view Learning*]

Northeastern University

Smile lab — Synergetic Media Learning Lab

Deep Multi-View Learning

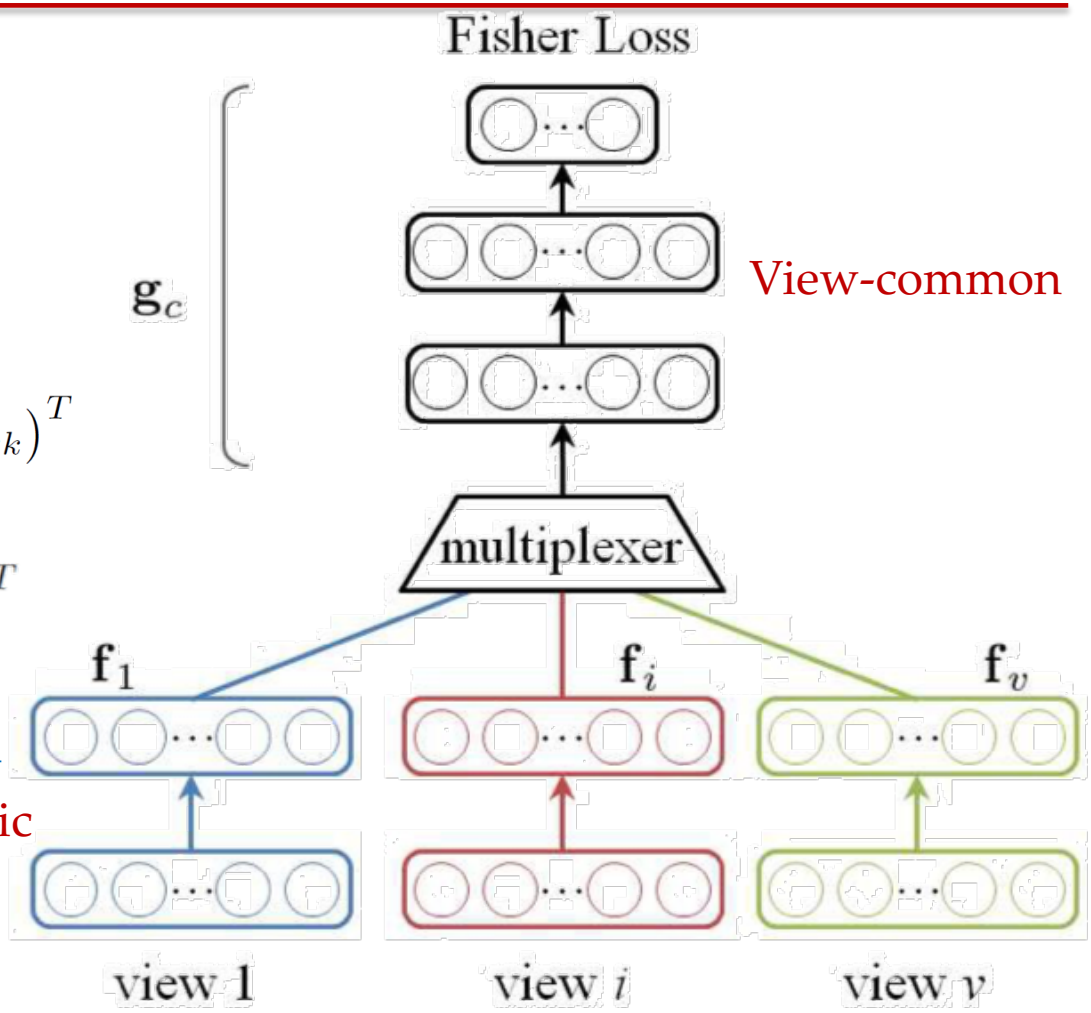
$$\min_{\mathbf{g}_c,\mathbf{f}_1,\cdots,\mathbf{f}_v} Tr\left(\frac{\mathbf{S}_W^y}{\mathbf{S}_B^y}\right)$$

$$\mathbf{S}_W^y = \sum_{k=1}^{c}\sum_{i=1}^{v}\sum_{j=1}^{n_{ki}} \left(\mathbf{y}_{jk}^i - \boldsymbol{\mu}_k\right)\left(\mathbf{y}_{jk}^i - \boldsymbol{\mu}_k\right)^T$$

$$\mathbf{S}_B^y = \sum_{k=1}^{c} n_k \left(\boldsymbol{\mu}_k - \boldsymbol{\mu}\right)\left(\boldsymbol{\mu}_k - \boldsymbol{\mu}\right)^T$$

Fisher-like Loss

$$\mathbf{y}_j^i = \mathbf{g}_c\left(\mathbf{f}_i\left(\mathbf{x}_j^i\right)\right)$$



Multi-view Deep Network for Cross-view Classification– CVPR 2016

Meina Kan, Shiguang Shan and Xilin Chen

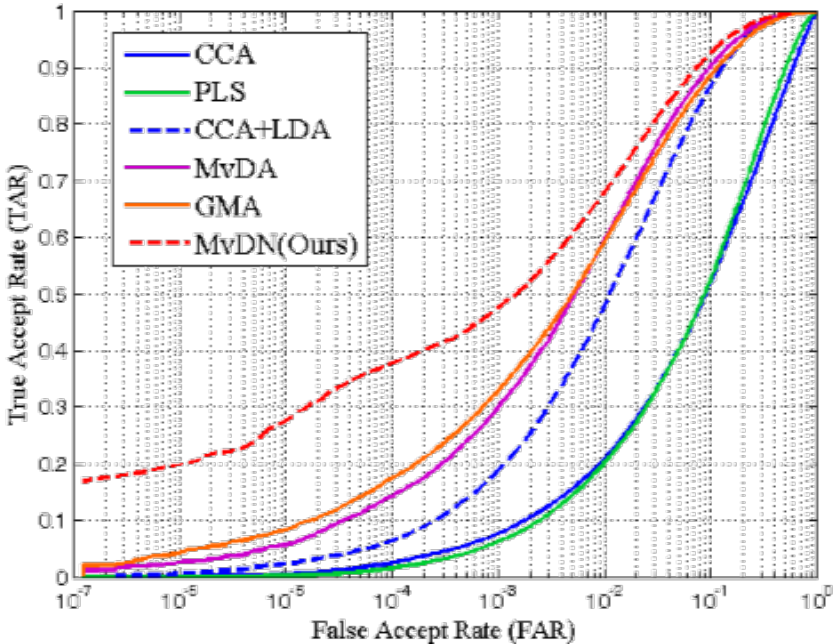
Supervised Multi-View Face Representation
[*Multi-view Learning*]

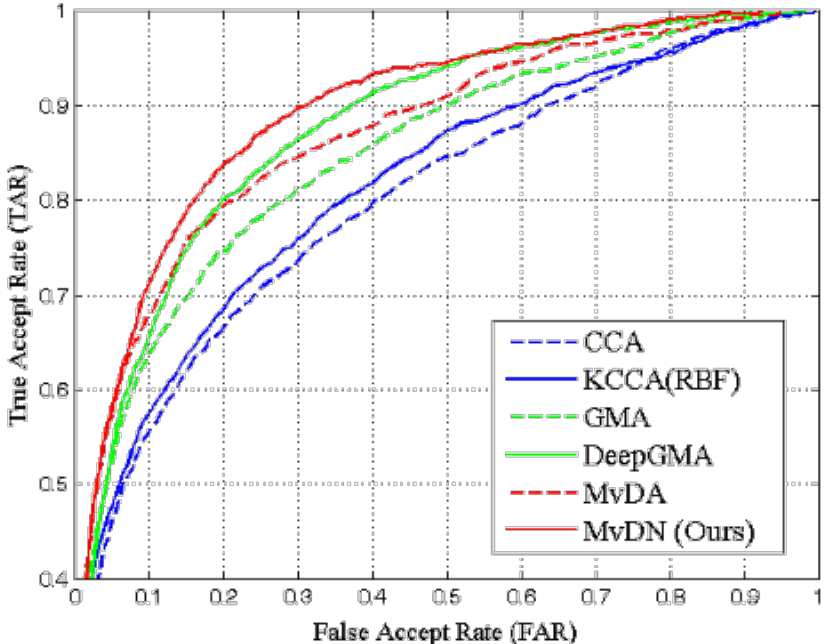
Northeastern University

Smile lab — Synergetic Media Learning Lab

Table 1. Evaluation of face recogntion across view angle on the MultiPIE dataset.

Methods	-90°	-75°	-60°	-45°	-30°	-15°	15°	30°	45°	60°	75°	90°	Average
PLS	0.319	0.775	0.892	0.934	0.883	0.981	0.981	0.934	0.906	0.873	0.723	0.268	0.789
MCCA	0.409	0.742	0.822	0.723	0.685	0.920	0.906	0.798	0.747	0.779	0.714	0.376	0.718
PLS+LDA	0.380	0.798	0.869	0.944	0.920	0.995	0.986	0.967	0.883	0.850	0.709	0.319	0.802
MCCA+LDA	0.488	0.662	0.817	0.887	1.00	1.00	1.00	0.995	0.831	0.803	0.676	0.568	0.811
MvDA	0.568	0.723	0.845	0.920	0.967	1.00	1.00	0.991	0.897	0.864	0.714	0.559	0.837
GMA	0.526	0.732	0.845	0.901	1.00	1.00	1.00	1.00	0.906	0.859	0.718	0.573	0.838
MvDN (Ours)	**0.704**	**0.822**	**0.883**	**0.911**	**0.991**	**1.00**	**1.00**	**0.991**	**0.930**	**0.911**	**0.798**	**0.709**	**0.887**





Outline

Northeastern University

Smile lab

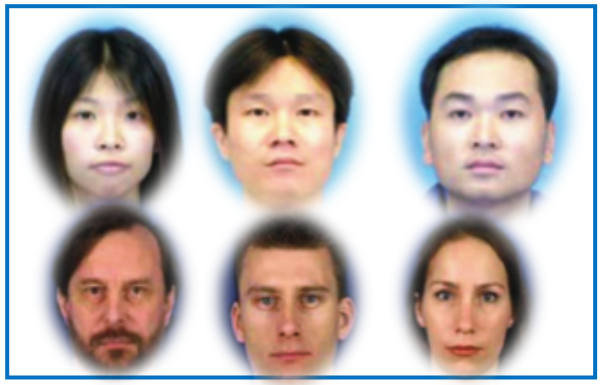
Synergetic Media Learning Lab

- **Introduction & Background**
  - Multi-view Face Task
  - Multi-view Face Data
- **Unsupervised Multi-view Face Representation**
  - Methodology
  - Face Clustering, Outlier Detection
- **Supervised Multi-view Face Representation**
  - Multi-view Learning
  - *Transfer Learning*
- **Conclusion**

Supervised Multi-View Face Representation
[*Transfer Learning*]

Northeastern University
Smile lab
Synergetic Media Learning Lab

[No Data Correspondence]



Source

Target

- Source Views (labeled)

$$D_{s,M} = \{(x_{i,M}, y_{i,M}), i = 1,2\cdots, N \sim P_s(X,Y)\}$$

mismatch

- Target Views ([un]labeled)

$$D_T = \{(x_i, ?), i = 1,2\cdots, N \sim P_T(X,Y)\}$$

Performance degrades significantly!

- **Objective**

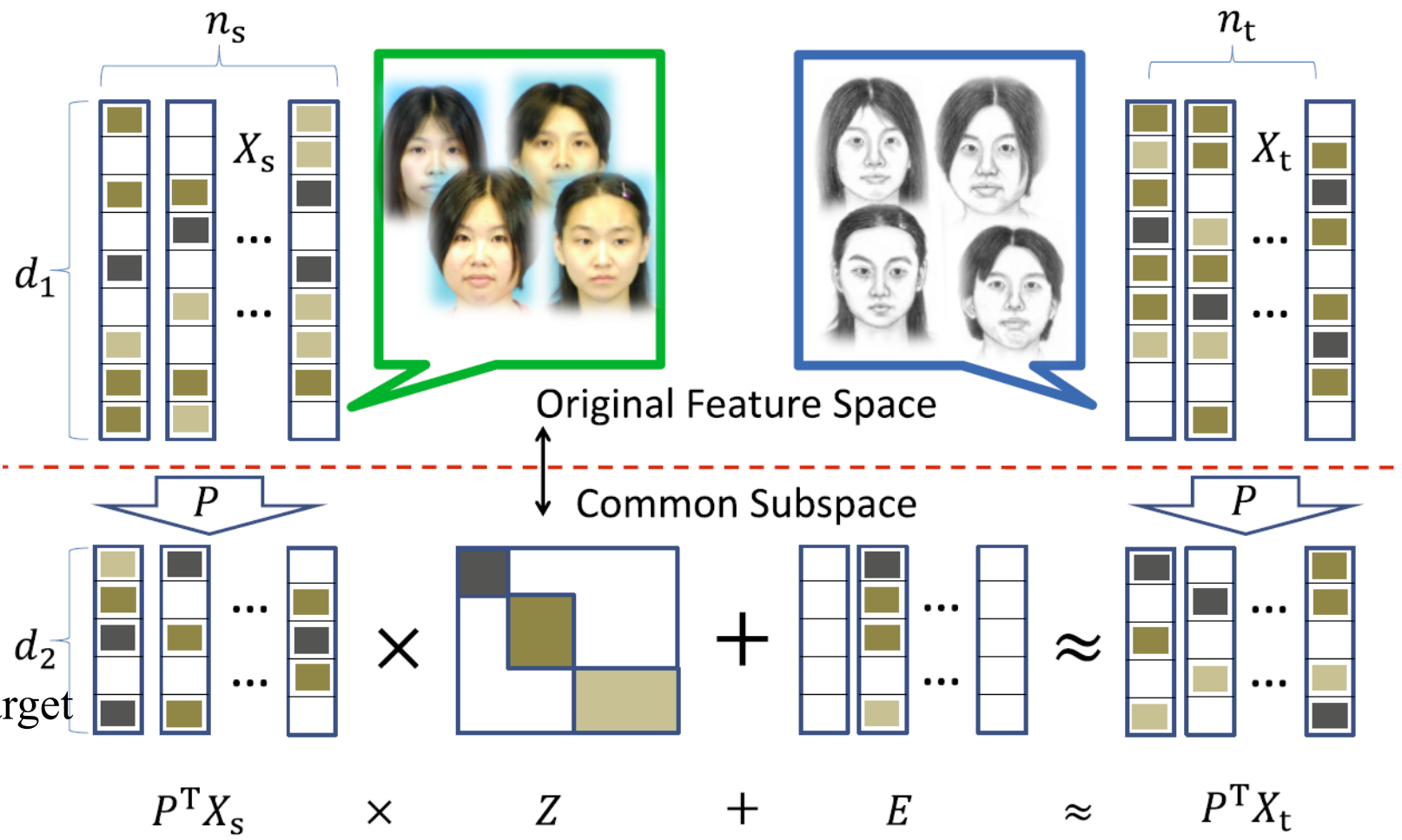
  Train classification model to work well on the target

Learn view-invariant features

Supervised Multi-View Face Representation
[*Transfer Learning*]

Northeastern University

Smile lab – Synergetic Media Learning Lab



Low-Rank Transfer Subspace Learning – ICDM 2012, IJCV 2014

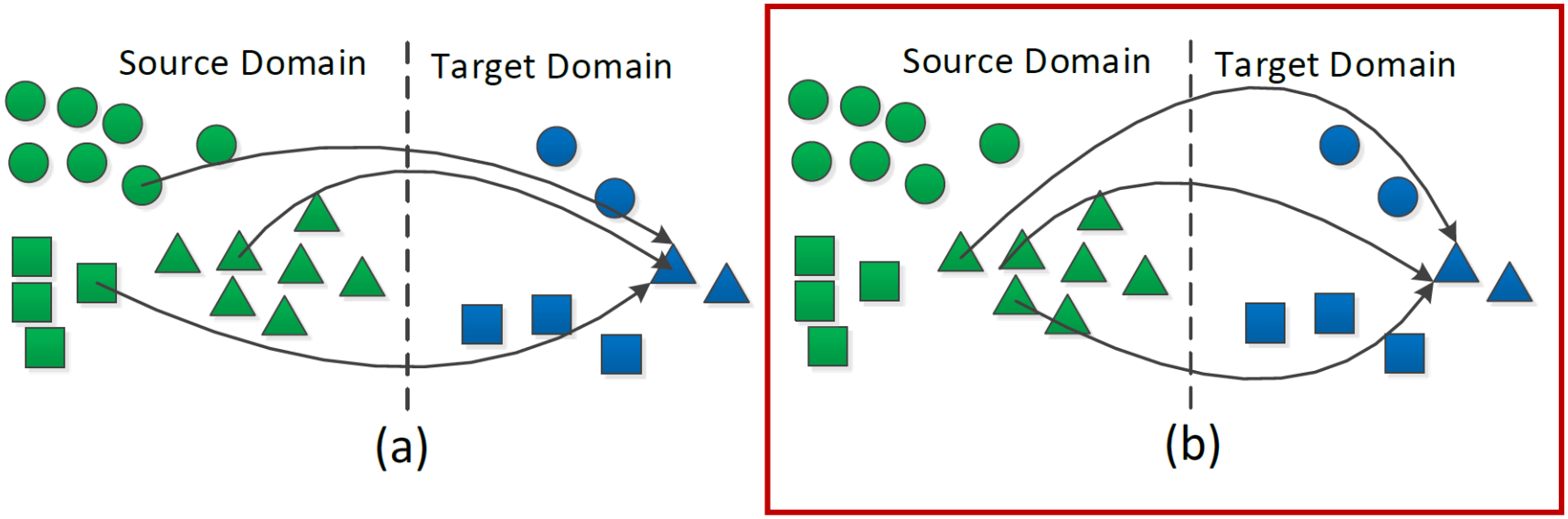
Ming Shao and Yun Fu

Supervised Multi-View Face Representation
[*Transfer Learning*]

Northeastern University
Smile lab
Synergetic Media Learning Lab

Goal: Align the source and target data in a learned subspace by minimizing the reconstruction error

- Incorporate *low-rank* constraint to enforce locality aware reconstruction
- Low-rank and sparse constraints are used to compensate for the noisy data
- LTSL generalizes traditional subspace learning techniques to transfer learning scenario



Low-Rank Transfer Subspace Learning – ICDM 2012, IJCV 2014

Ming Shao and Yun Fu

Supervised Multi-View Face Representation
[*Transfer Learning*]

Northeastern University

Smile lab – Synergetic Media Learning Lab

Objective Function

$\mathrm{Tr}(P^T U_1 P)$

$$\min_{Z,E,P} F(P, X_s) + \lambda_1 \|Z\|_* + \lambda_2 \|E\|_{2,1}$$

$$\text{s.t.,} \; P^T X_t = P^T X_s Z + E, \; \mathbf{1}_{n_s}^T Z = \mathbf{1}_{n_t}^T, \; P^T U_2 P = \mathbf{I}_{d_2}$$

U_1 and U_2 are selected based on subspace learning methods

Method	U_1	U_2
PCA	$-\Sigma$	$\mathbf{I}_{d_1}$
LDA	S_w	S_b
LPP[a]	$X_s \mathcal{L} X_s^T$	$X_s D X_s^T$
NPE	$X_s(\mathbf{I}_{n_s} - W)^T(\mathbf{I}_{n_s} - W)X_s^T$	$X_s X_s^T$
MFA[b]	$X_s(D - W)X_s^T$	$X_s(D_p - W_p)X_s^T$
DLA[c]	$X_s L X_s^T$	$\mathbf{I}_{d_1}$

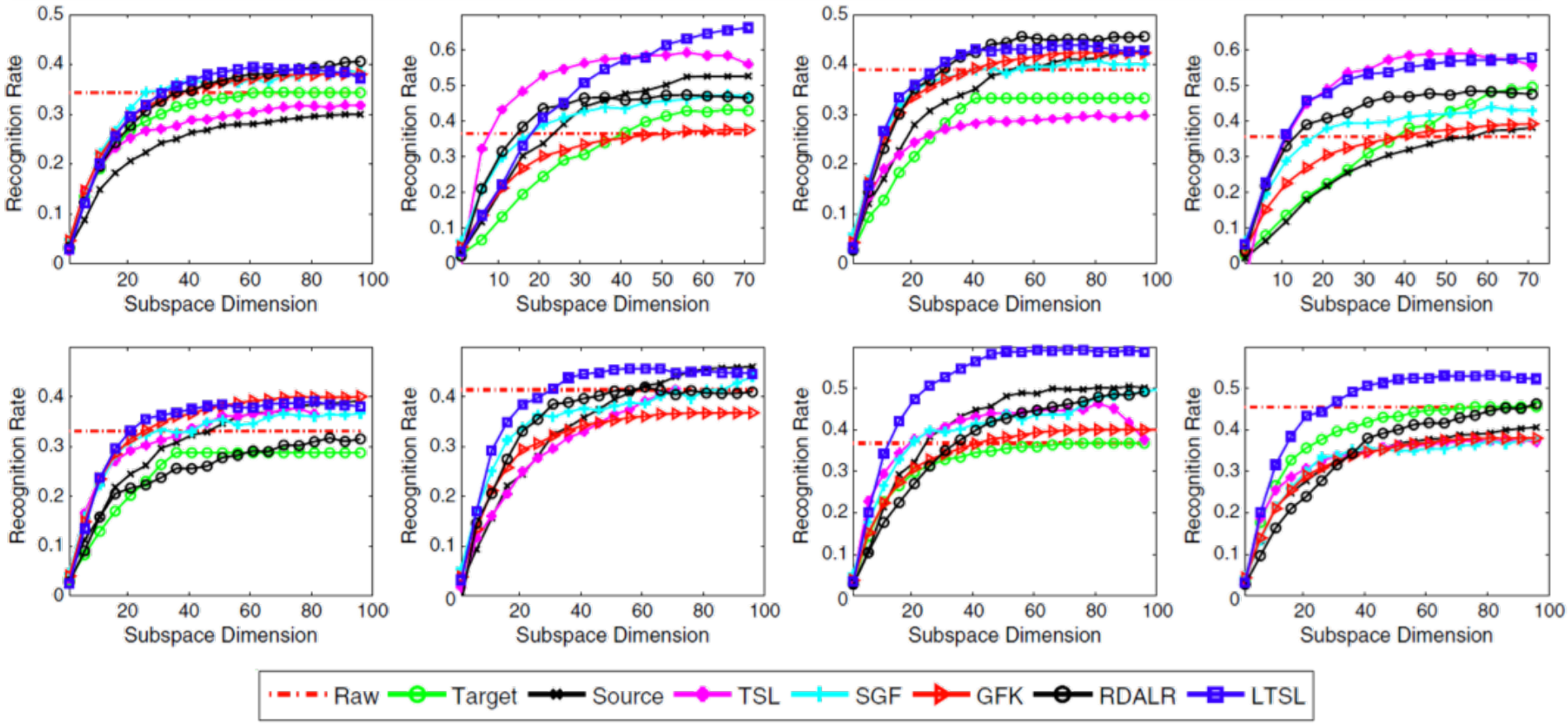
Low-Rank Transfer Subspace Learning – ICDM 2012, IJCV 2014

Ming Shao and Yun Fu

Supervised Multi-View Face Representation
[*Transfer Learning*]

Northeastern University

Smile lab
Synergetic Media Learning Lab



BUAA–VISNIR: Subspace learning methods are PCA, LDA, ULPP, SLPP, UNPE, SNPE, MFA, DLA

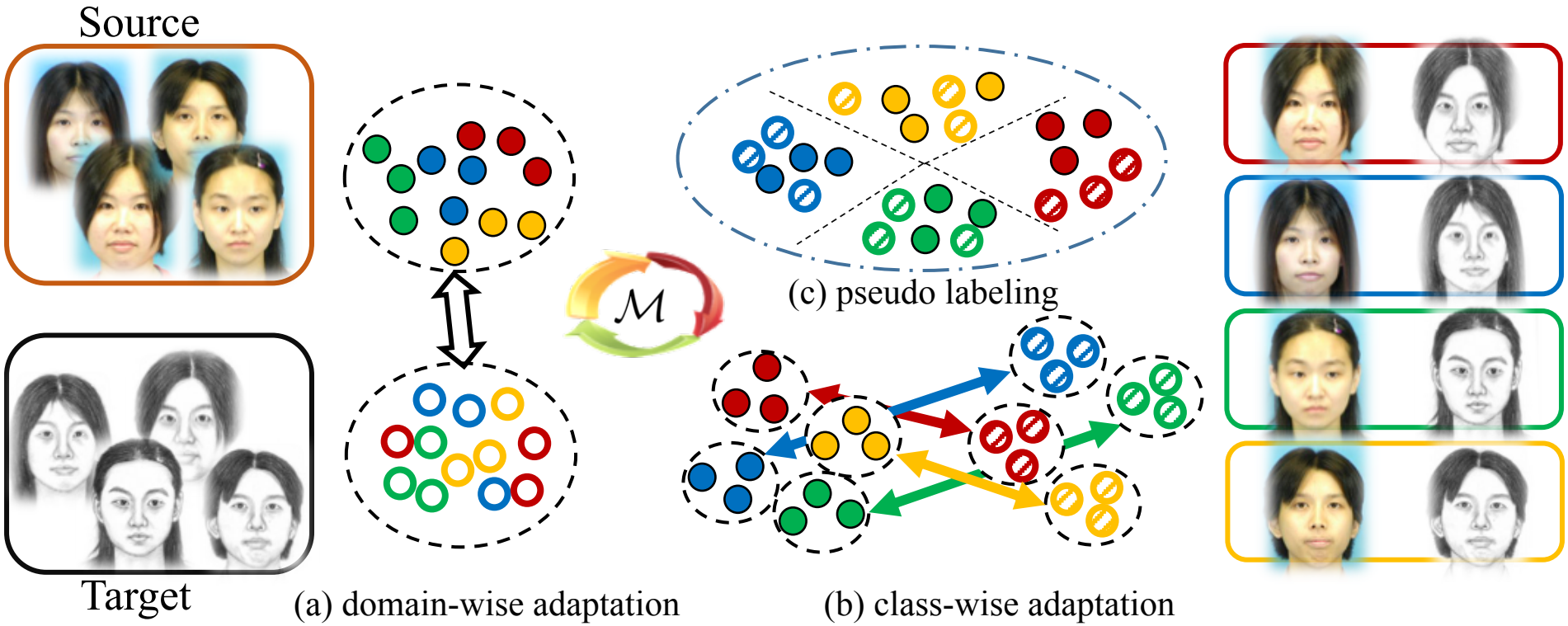
Low-Rank Transfer Subspace Learning – ICDM 2012, IJCV 2014

Ming Shao and Yun Fu

Supervised Multi-View Face Representation
[*Transfer Learning*]

Northeastern University
Smile lab
Synergetic Media Learning Lab

Robust Transfer Metric Learning



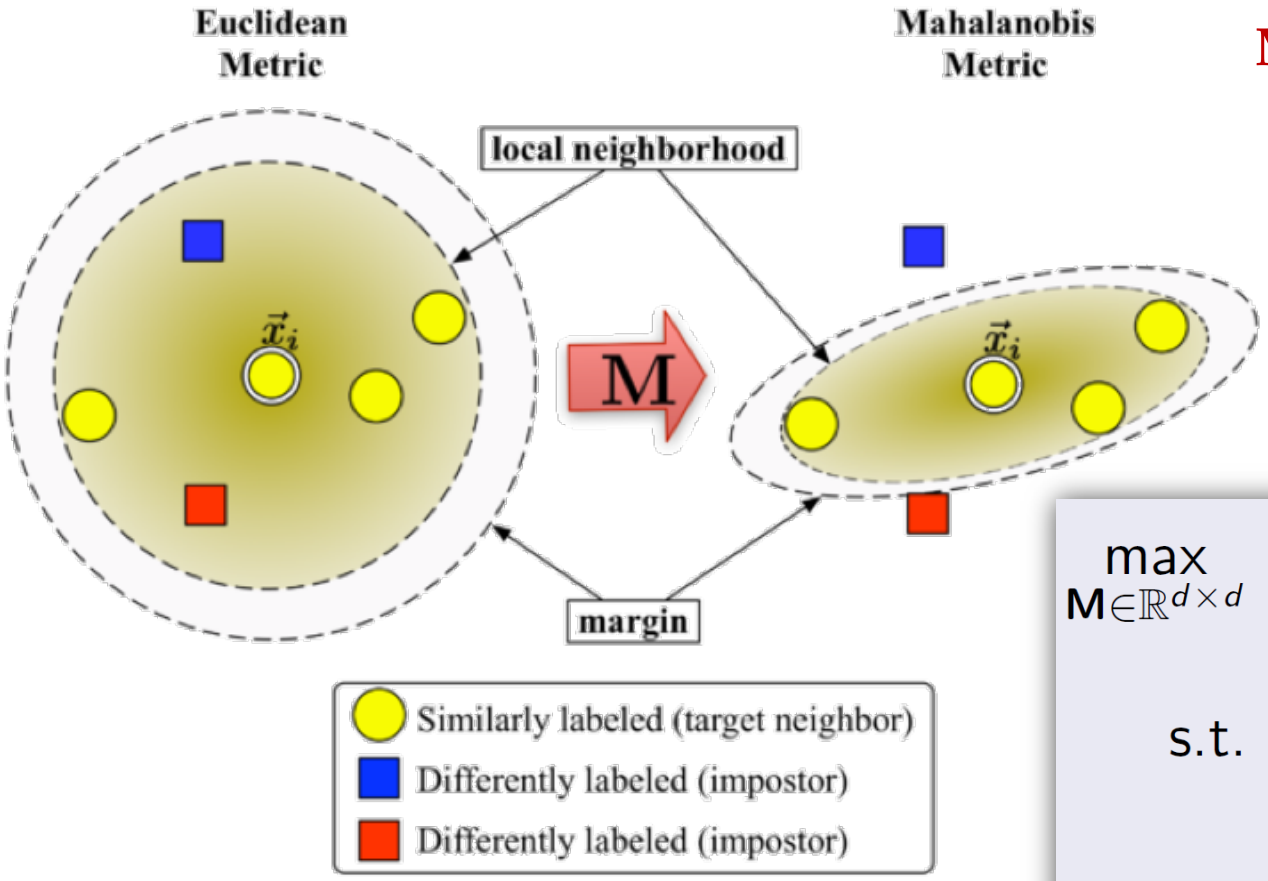
Robust Transfer Metric Learning – IEEE TIP 2016

Zhengming Ding and Yun Fu

Supervised Multi-View Face Representation
[*Transfer Learning*]

Northeastern University

Smile lab — Synergetic Media Learning Lab



Metric Learning Revisit

$$\mathcal{M} = PP^{\top}$$

$$\max_{\mathbf{M}\in\mathbb{R}^{d\times d}} \sum_{(\mathbf{x}_i,\mathbf{x}_j)\in\mathcal{D}} d_{\mathbf{M}}(\mathbf{x}_i,\mathbf{x}_j)$$

$$\text{s.t.} \sum_{(\mathbf{x}_i,\mathbf{x}_j)\in\mathcal{S}} d^2_{\mathbf{M}}(\mathbf{x}_i,\mathbf{x}_j) \leq 1,$$

$$\mathbf{M} \succeq 0.$$

Distance Metric learning: A Comprehensive Survey.
Yang Liu, and Rong Jin.

Supervised Multi-View Face Representation
[*Transfer Learning*]

Northeastern University

Smile lab – Synergetic Media Learning Lab

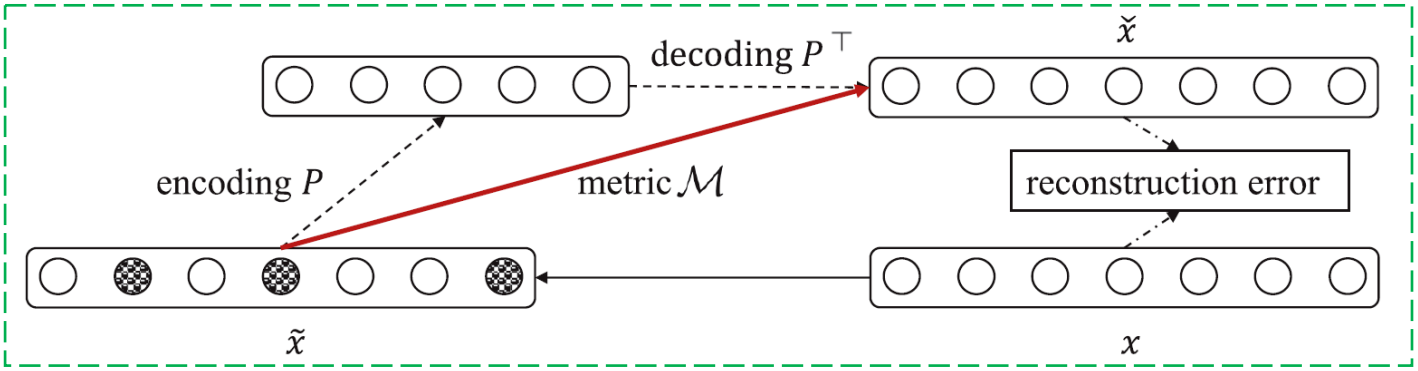
Knowledge Adaptation Marginal Denoising Low-rank Constraint

$$\min_{\mathcal{M}\in\mathbb{S}_+^d} \sum_{i=0}^{c} \boxed{\mathrm{tr}(\Phi^i\mathcal{M})} + \alpha\boxed{\|\bar{X}-\mathcal{M}\tilde{X}\|_{\mathrm{F}}^2} + \lambda\boxed{\mathrm{rank}(\mathcal{M})}$$

$$\sum_{i=0}^{c}(\mu_s^i-\mu_t^i)^\top\mathcal{M}(\mu_s^i-\mu_t^i)$$

$$\mathcal{M} = PP^\top$$

$$\sum_{i=r+1}^{d}\left(\sigma_i(\mathcal{M})\right)^2$$



Robust Transfer Metric Learning – IEEE TIP 2016

Zhengming Ding and Yun Fu

Supervised Multi-View Face Representation
[*Transfer Learning*]

Northeastern University

Smile lab
Synergetic Media Learning Lab

Results on CMU-PIE [*Five Poses*]

Config\Methods	PCA [39]	ITML [40]	GFK [23]	JDA [41]	TCA [42]	CDML [19]	mSDA [11]	Ours
C05→C07	25.43	27.75	26.15	58.81	40.76	53.22	28.35	60.12
C05→C09	26.87	27.33	27.27	54.23	41.79	53.12	26.91	55.21
C05→C27	30.98	31.60	31.15	84.50	59.63	80.12	30.39	85.19
C05→C29	17.21	19.00	17.59	49.75	29.35	48.23	21.76	52.98
C07→C05	25.13	26.05	25.24	57.62	41.81	52.39	28.27	58.13
C07→C09	47.43	48.71	47.37	62.93	51.47	54.23	44.19	63.92
C07→C27	53.98	55.54	54.25	75.82	64.73	68.36	55.39	76.16
C07→C29	27.12	29.53	27.08	39.89	33.70	37.34	28.08	40.38
C09→C05	21.67	22.99	28.69	50.96	34.69	43.54	24.83	53.12
C09→C07	42.87	44.20	43.16	57.95	47.70	54.87	42.59	58.67
C09→C27	45.97	47.34	46.41	68.45	56.23	62.76	50.25	69.81
C09→C29	26.43	28.25	26.78	39.95	33.15	38.21	27.83	42.13
C27→C05	34.12	35.83	34.24	80.58	55.64	75.12	32.89	81.12
C27→C07	62.09	64.46	62.92	82.63	67.83	80.53	63.10	83.92
C27→C09	72.67	74.39	73.35	87.25	75.86	83.72	74.70	89.51
C27→C29	37.43	39.46	37.38	54.66	40.26	52.78	34.81	56.26
C29→C05	21.08	21.40	20.35	23.17	26.98	27.34	25.85	29.11
C29→C07	24.61	25.72	24.62	31.74	29.90	30.82	26.33	33.28
C29→C09	28.19	29.72	28.49	38.17	29.90	36.34	28.63	39.85
C29→C27	31.05	31.93	31.33	45.99	33.64	40.61	32.98	47.13
Average	35.13	36.56	35.69	57.25	44.75	53.69	36.41	58.80

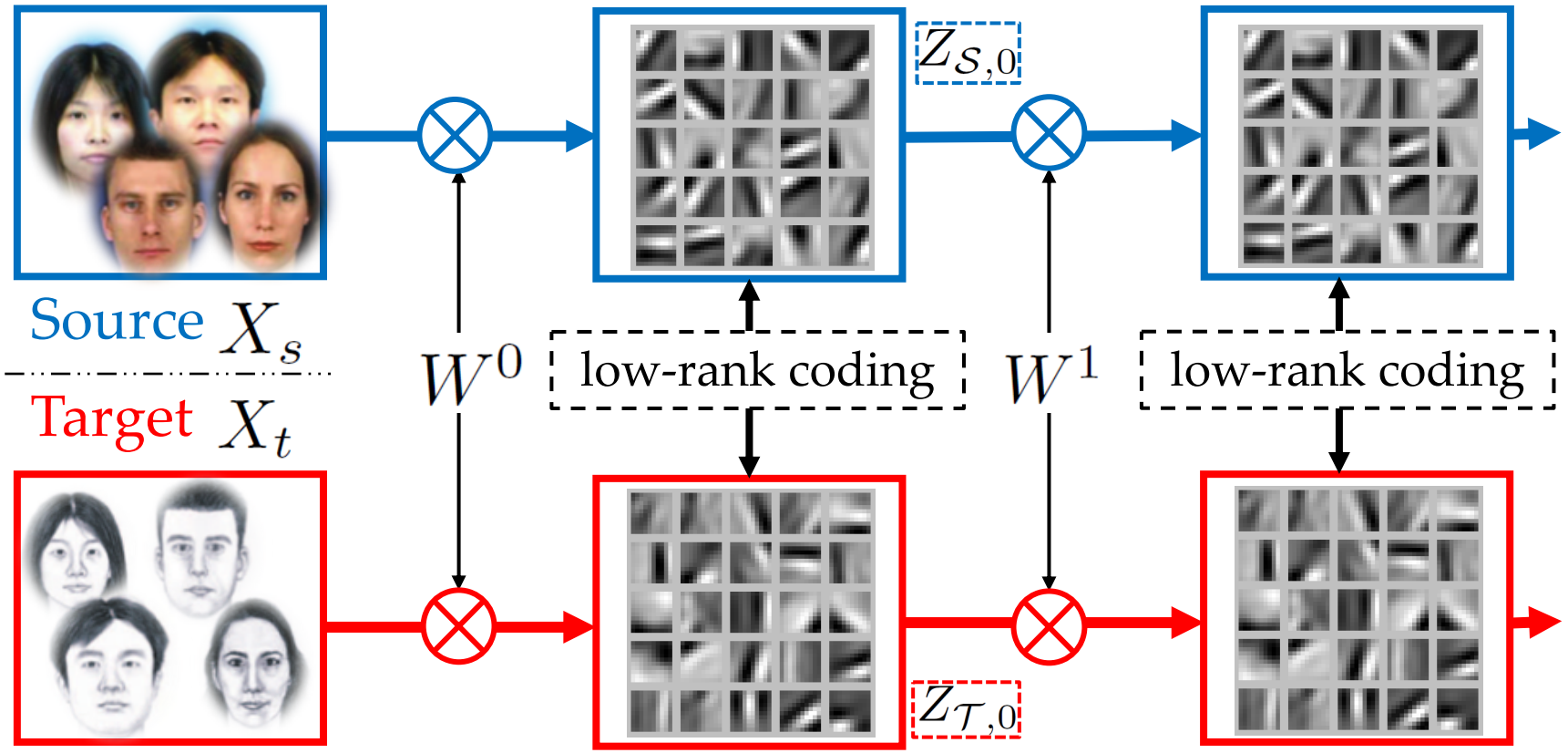
Robust Transfer Metric Learning – IEEE TIP 2016

Zhengming Ding and Yun Fu

Supervised Multi-View Face Representation
[*Transfer Learning*]

Northeastern University

Smile lab
Synergetic Media Learning Lab



Deep Low-Rank Coding for Transfer Learning – IJCAI 2015

Zhengming Ding, Ming Shao and Yun Fu

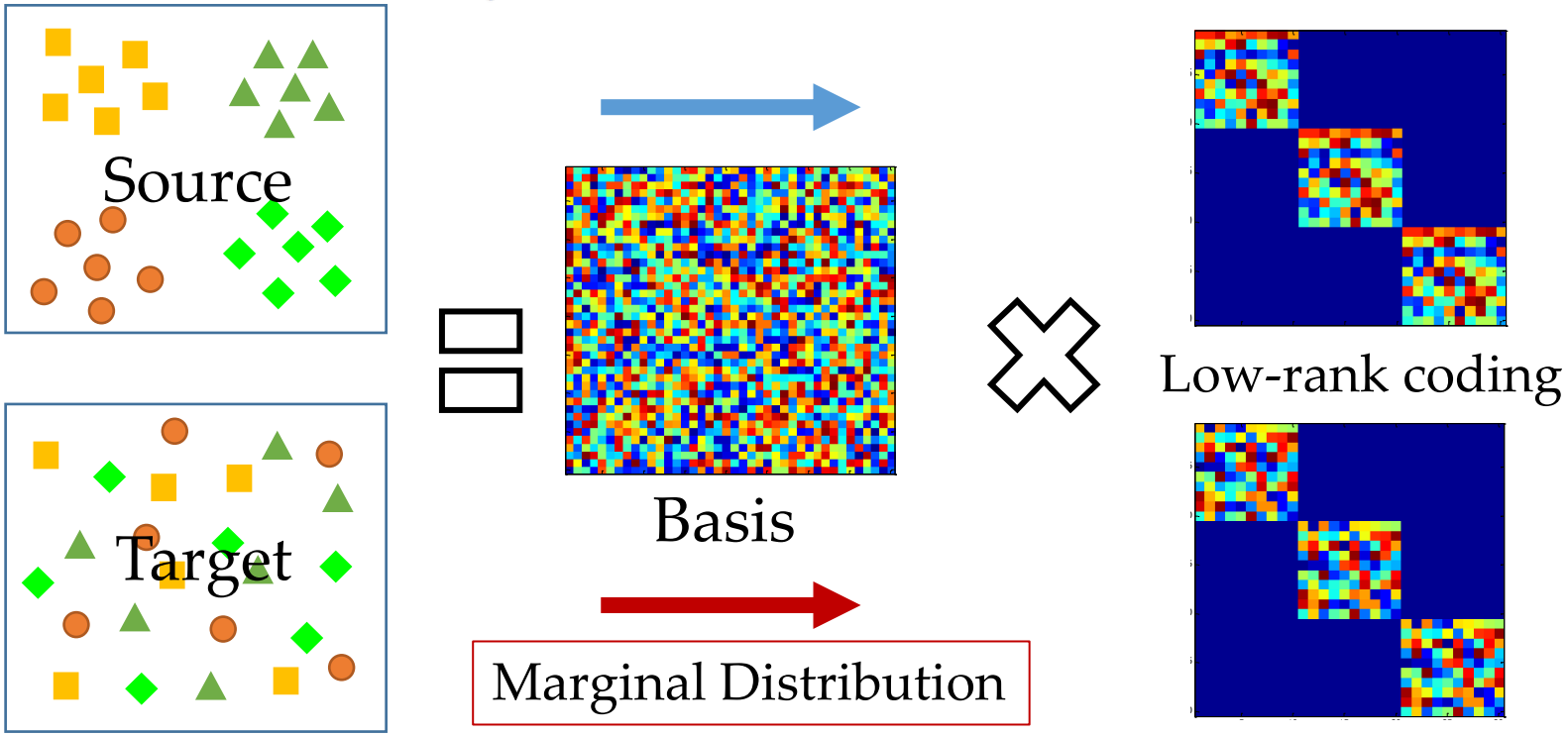
Supervised Multi-View Face Representation
[*Transfer Learning*]

Northeastern University
Smile lab
Synergetic Media Learning Lab

Single-Layer Objective Function

Low-rank Coding

$$\min_{Z,W} \operatorname{rank}(Z) + \lambda\Omega(W), \quad \text{s.t.} \quad WX = WX_{\mathcal{S}}Z$$



Deep Low-Rank Coding for Transfer Learning – IJCAI 2015

Zhengming Ding, Ming Shao and Yun Fu

Supervised Multi-View Face Representation
[*Transfer Learning*]

Northeastern University

Smile lab – Synergetic Media Learning Lab

Single-Layer Objective Function

Structure Constraint

$$\min_{W,Z} \|Z\|_* + \lambda \mathrm{tr}\left[(\bar{X} - W\tilde{X})^{\mathrm{T}}(\bar{X} - W\tilde{X})\right] + \alpha \|Z_l - H\|_{\mathrm{F}}^2$$

$$\text{s.t. } WX = WX_{\mathcal{S}}Z,$$

Marginalized denoising regularizer [1]

$\bar{X}$, $\tilde{X}$ are the *m*-repeated clean and corrupted data

Iterative Structure Matrix

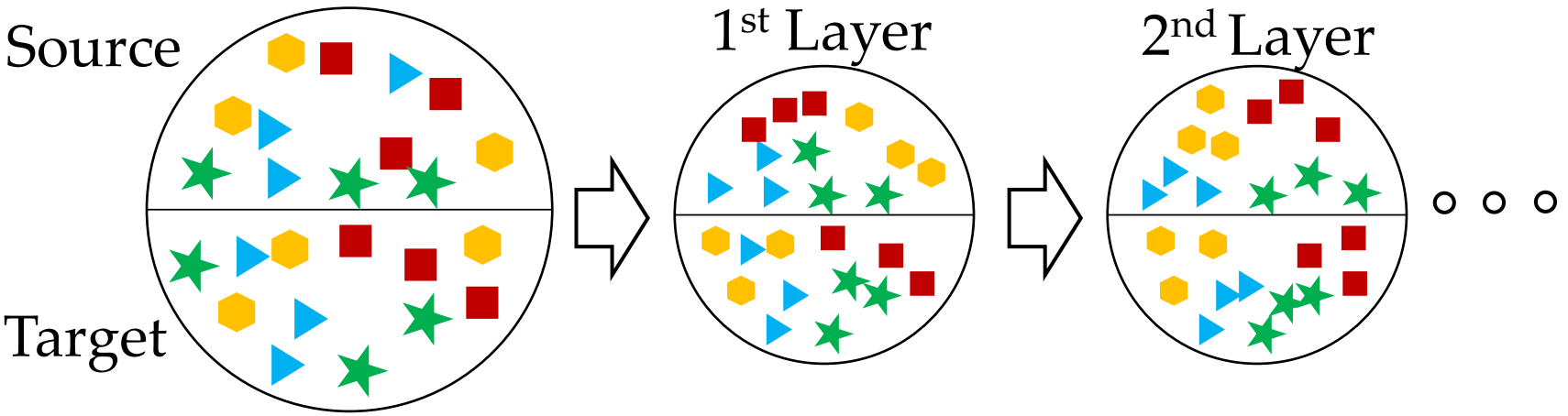
- Make the learned transformation *W* more robust to noisy data
- Discriminative and robust low-rank coding for two domains

[1] Chen et al. *Marginalized Denoising Autoencoders for Domain Adaptation*. ICML, 2012

Supervised Multi-View Face Representation
[*Transfer Learning*]

Northeastern University

Smile lab – Synergetic Media Learning Lab



Algorithm 2 Algorithm of Deep Low-Rank Coding (DLRC)

Input: $X_{\mathcal{S}}, X_{\mathcal{T}}, L$ is the number of layers,

for $k = 1$ **to** L **do**

1. Use **Algorithm 1** to learn coding $Z_{\mathcal{S},k}$ and $Z_{\mathcal{T},k}$;
2. Set $X_{\mathcal{S},k+1} = Z_{\mathcal{S},k}$ and $X_{\mathcal{T},k+1} = Z_{\mathcal{T},k}$;
3. Update H_k via Eq. (10);

end for

output: Low-rank codings $\{Z_{\mathcal{S},k}, Z_{\mathcal{T},k}\}, (k = 1, \cdots, L)$.

Deep Low-Rank Coding for Transfer Learning – IJCAI 2015

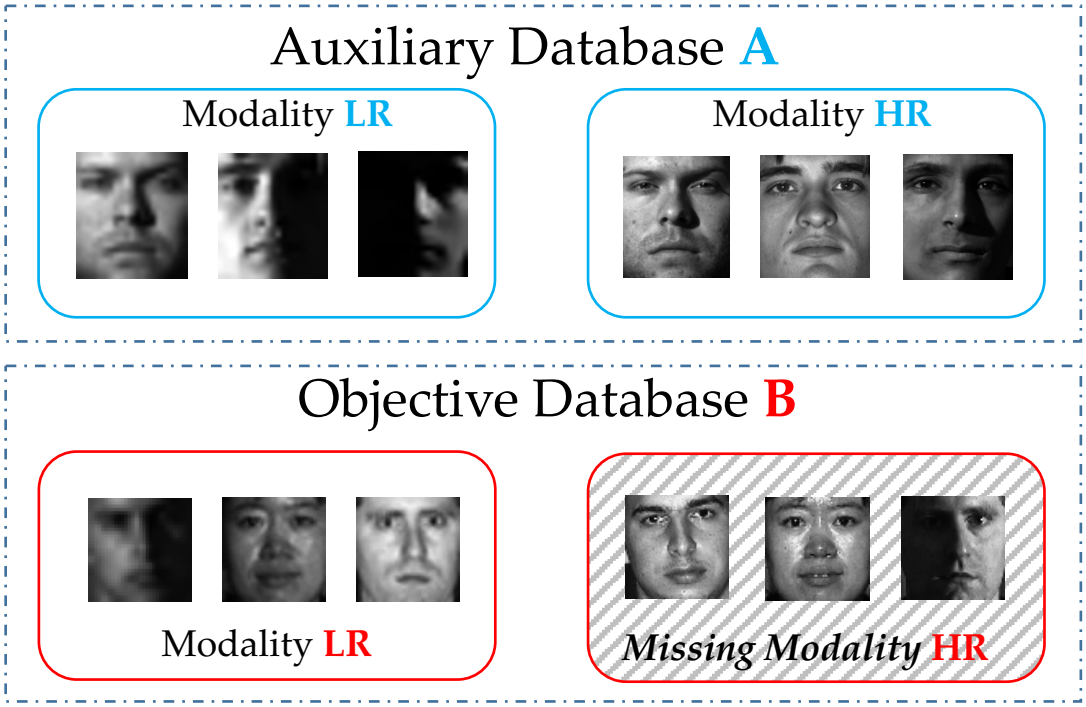
Zhengming Ding, Ming Shao and Yun Fu

Supervised Multi-View Face Representation
[*Transfer Learning*]

Northeastern University

Problem: The target modality is missing in the training stage.

Idea: With the help of existed data (auxiliary database **A** and modality LR of **B**), the missing modality HR of database **B** can be recovered to facilitate the recognition performance.



Missing Modality Transfer Learning

Missing Modality Transfer Learning – AAAI 2014 & TIP 2015

Zhengming Ding, Ming Shao and Yun Fu

Supervised Multi-View Face Representation
[*Transfer Learning*]

Northeastern University

Smile lab – Synergetic Media Learning Lab

Latent Low-rank Recovery

Observed data Hidden data

$$\min_{Z} \|Z\|_*, \quad \text{s.t.} \quad X_O = [X_O, X_H]Z$$

$$\begin{aligned} X_O &= [X_O, X_H]Z^*_{O,H} = X_O Z^*_{O|H} + X_H Z^*_{H|O} \\ &= X_O Z^*_{O|H} + X_H V_H V_O^T \\ &= X_O Z^*_{O|H} + U\Sigma V_H^T V_H V_O^T \\ &= X_O Z^*_{O|H} + U\Sigma V_H^T V_H \Sigma^{-1} U^T X_O. \end{aligned}$$

$$\begin{aligned} \min_{Z_{O|H}, L_{H|O}} \quad & \text{rank}\left(Z_{O|H}\right) + \text{rank}\left(L_{H|O}\right), \\ \text{s.t.} \quad & X_O = X_O Z_{O|H} + L_{H|O} X_O. \end{aligned}$$

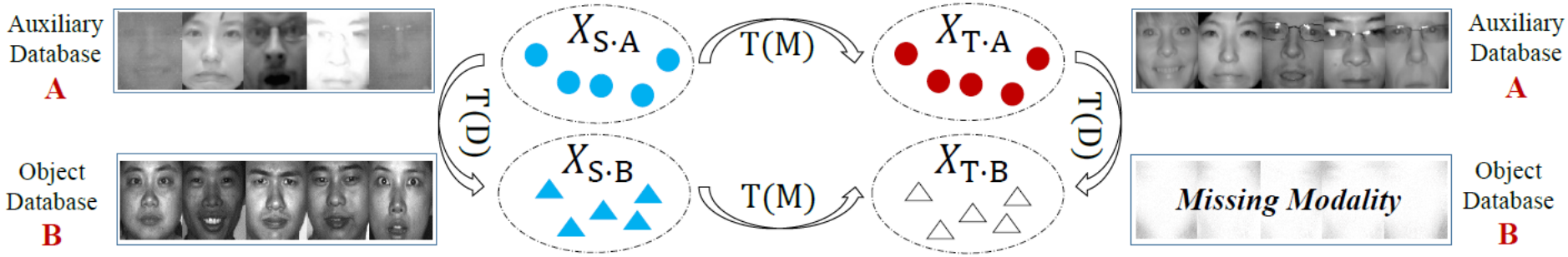
Latent Low-Rank Representation – ICCV 2011

Guangcan Liu and Shuicheng Yan

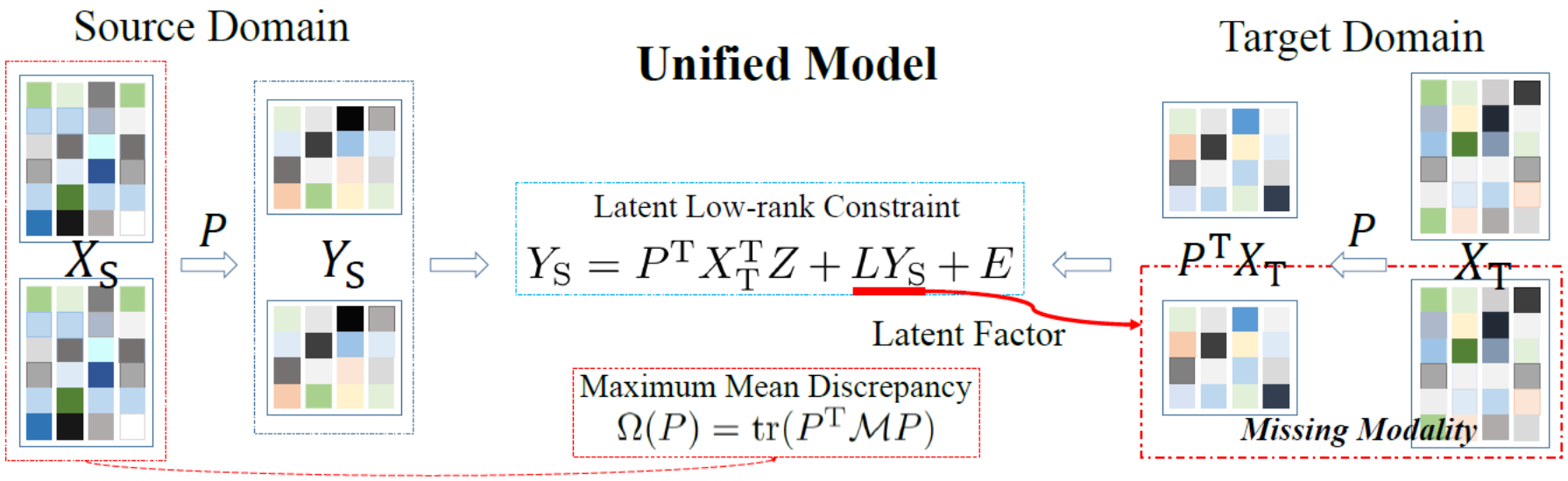
Supervised Multi-View Face Representation
[*Transfer Learning*]

Northeastern University

Smile lab — Synergetic Media Learning Lab



Illustration



Missing Modality Transfer Learning – AAAI 2014 & TIP 2015

Zhengming Ding, Ming Shao and Yun Fu

Supervised Multi-View Face Representation
[*Transfer Learning*]

Northeastern University

Smile lab
Synergetic Media Learning Lab

Model *I*

$$\min_{P,Z,L,E} \|Z\|_* + \|L\|_* + \lambda\|E\|_1 + \alpha \mathrm{tr}(P^{\mathrm{T}} U_1 P),$$
$$\text{s.t.} \quad P^{\mathrm{T}} X_{\mathrm{S}} = P^{\mathrm{T}} X_{\mathrm{T}} Z + L P^{\mathrm{T}} X_{\mathrm{S}} + E, P^{\mathrm{T}} U_2 P = \mathrm{I}_{\mathrm{p}}$$

Different Subspace

Latent Factor

Model *II*

$$\min_{P,Z,L,E} \|Z\|_* + \|L\|_* + \lambda\|E\|_1 + \alpha\|P\|_{2,1} + \beta\Omega(P)$$
$$\text{s.t.} \quad Y_{\mathrm{S}} = P^{\mathrm{T}} X_{\mathrm{T}} Z + L Y_{\mathrm{S}} + E, P^{\mathrm{T}} P = \mathrm{I}_{\mathrm{p}},$$

$$\|\frac{1}{n_a}\sum_{i=1}^{n_a} P^{\mathrm{T}} x_i - \frac{1}{n_b}\sum_{j=n_a+1}^{n} P^{\mathrm{T}} x_j\|_{\mathrm{F}}^2 = \|P^{\mathrm{T}}\mu_{\mathrm{A}} - P^{\mathrm{T}}\mu_{\mathrm{B}}\|_{\mathrm{F}}^2 = \mathrm{tr}(P^{\mathrm{T}}\mathcal{M}P)$$

Differences:

- Pre-learned Low-dimensional Features
- Sparse Projection

Missing Modality Transfer Learning – AAAI 2014 & TIP 2015

Zhengming Ding, Ming Shao and Yun Fu

Supervised Multi-View Face Representation
[*Transfer Learning*]

Northeastern University
Smile lab
Synergetic Media Learning Lab

BUAA & Oulu Face Database

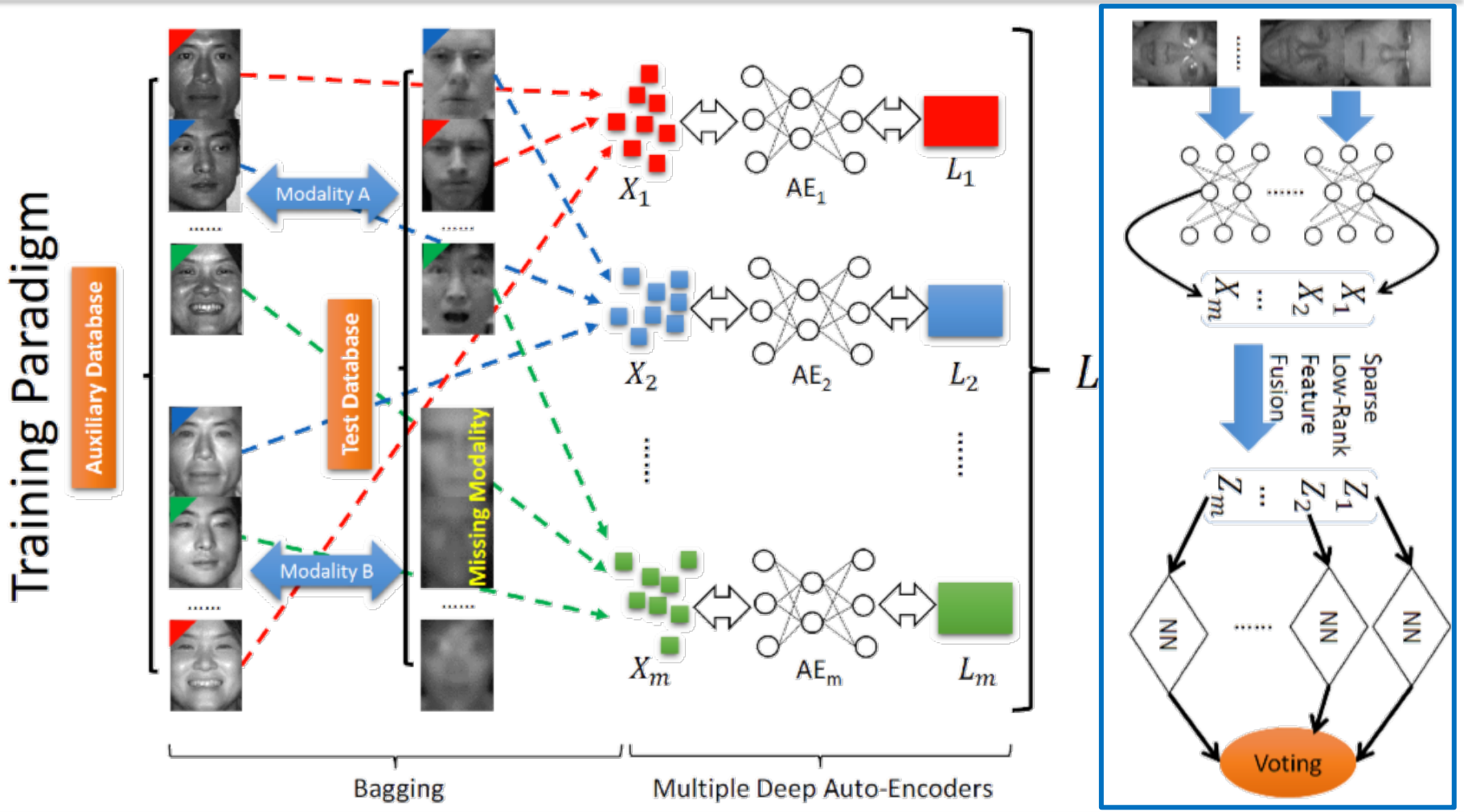
Methods		TSL [42]	RDALR [5]	GFK [43]	LTSL [4]	DASA [2]	L^2TSL [13]	Ours
Case 1	PCA	35.82±0.76	40.21±0.67	38.34±0.83	47.21±0.54	59.43±0.62	52.32±0.67	59.02±0.46
	LDA	31.31±0.32	38.52±0.52	12.7±0.12	42.38±0.43	11.59±0.10	48.72±0.42	62.48±0.30
	ULPP	29.28±0.45	42.84±0.37	40.21±0.25	50.81±0.85	41.31±0.83	59.68±0.48	58.32±0.65
	SLPP	36.86±0.38	47.27±0.42	39.56±0.36	53.57±0.52	18.17±0.15	63.71±0.62	66.68±0.35
Case 2	PCA	37.06±0.34	33.76±0.39	42.39±0.49	38.39±0.46	38.37±0.49	49.79±0.52	51.21±0.56
	LDA	28.39±0.12	34.57±0.23	15.84±0.18	41.38±0.38	18.76±0.07	43.23±0.36	52.98±0.25
	ULPP	38.29±0.31	39.88±0.42	39.29±0.23	41.28±0.35	37.48±0.51	49.34±0.35	55.71±0.48
	SLPP	46.88±0.51	50.28±0.28	48.39±0.39	56.79±0.53	34.76±0.22	60.73±0.58	61.54±0.63
Case 3	PCA	39.26±0.23	41.37±0.25	39.59±0.38	41.89±0.33	42.25±0.10	48.34±0.43	50.21±0.26
	LDA	42.25±0.51	36.58±0.24	26.87±0.38	50.76±0.63	23.83±0.29	56.82±0.42	64.77±0.54
	ULPP	47.37±0.43	42.39±0.62	28.38±0.35	48.24±0.32	52.58±0.11	50.83±0.42	58.89±0.43
	SLPP	45.75±0.38	48.28±0.41	45.38±0.47	54.78±0.52	41.08±0.06	55.71±0.32	64.58±0.55
Case 4	PCA	31.59±0.54	39.77±0.62	39.29±0.71	43.35±0.58	48.00±0.20	46.32±0.48	51.25±0.43
	LDA	40.34±0.42	42.38±0.33	38.36±0.51	48.28±0.35	29.03± 0.71	67.54±0.34	74.47±0.76
	ULPP	39.26±0.51	47.57±0.35	42.89±0.72	52.38±0.53	56.50±0.09	58.23±0.32	61.91±0.20
	SLPP	36.25±0.24	49.39±0.29	29.38±0.35	58.89±0.25	38.50±0.11	68.54±0.32	73.31±0.39

Missing Modality Transfer Learning – AAAI 2014 & TIP 2015

Zhengming Ding, Ming Shao and Yun Fu

Supervised Multi-View Face Representation
[*Transfer Learning*]

Northeastern University
Smile lab
Synergetic Media Learning Lab



Missing Modality Face Recognition – FG 2015

Ming Shao, Zhengming Ding, and Yun Fu

Supervised Multi-View Face Representation
[*Transfer Learning*]

Northeastern University
Smile lab
Synergetic Media Learning Lab

Objective Function

Training Stage

$$\min_{\substack{W_1,W_2,\\ b_1,b_2}} \mathcal{L}(X) + \lambda_1(\|W_1\|_{\mathrm{F}}^2 + \|W_2\|_{\mathrm{F}}^2) + \lambda_2 \sum_{i=1}^{d} \mathrm{KL}(\rho||\hat{\rho}_i)$$

$$\mathcal{L}(X) = \min_{W_1,b_1,W_2,b_2} \frac{1}{2n} \sum_{i=1}^{n} \|x_i - h(x_i)\|_2^2$$

Test Stage

$$\min_{Z,E} \|Z\|_* + \lambda_1\|E\|_{2,1} + \lambda_2\|\Psi(Z)\|_{2,1}$$
$$\text{s.t.} \quad X = XZ + E.$$

$$\begin{bmatrix} (Z_{11})_1 & (Z_{12})_1 & \cdots & (Z_{nn})_1 \\ (Z_{11})_2 & (Z_{12})_2 & \cdots & (Z_{nn})_2 \\ \vdots & \vdots & \ddots & \vdots \\ (Z_{11})_m & (Z_{12})_m & \cdots & (Z_{nn})_m \end{bmatrix}$$

Missing Modality Face Recognition – FG 2015

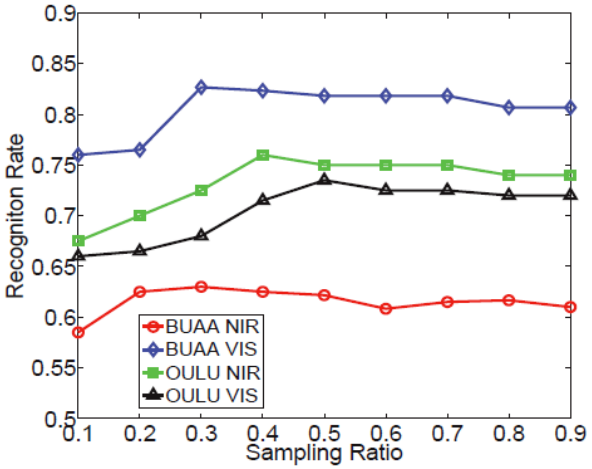
Ming Shao, Zhengming Ding, and Yun Fu

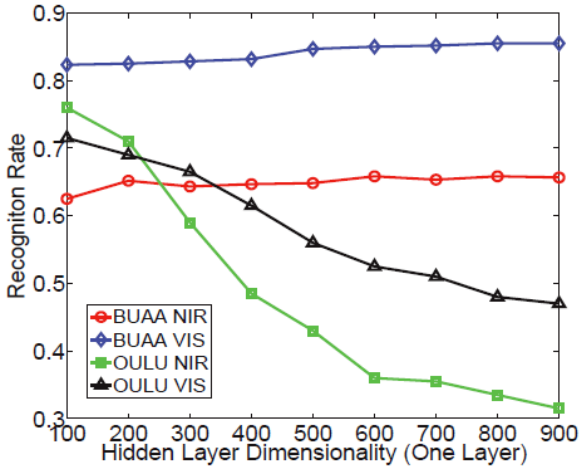
Supervised Multi-View Face Representation
[*Transfer Learning*]

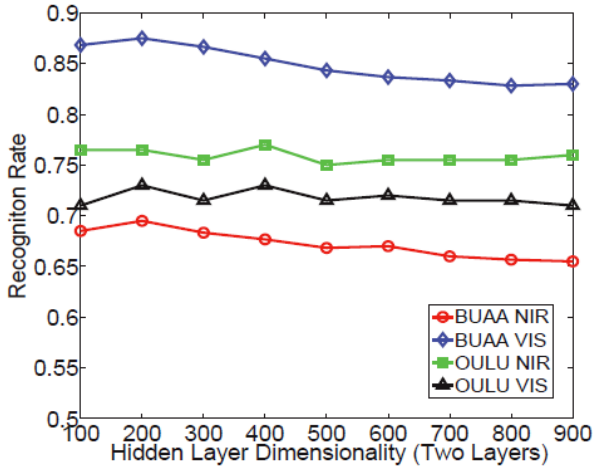
Northeastern University

Smile lab
Synergetic Media Learning Lab

BUAA & Oulu Face Database







Methods	Case 1				Case 2				Case 3				Case 4			
	PCA	LDA	ULPP	SLPP	PCA	LDA	ULPP	SLPP	PCA	LDA	ULPP	SLPP	PCA	LDA	ULPP	SLPP
TSL [24]	35.8	31.3	29.2	36.8	37.0	28.3	38.2	46.8	39.2	42.2	47.3	45.7	31.5	40.3	39.2	36.2
RDALR [13]	40.2	38.5	42.8	47.2	33.7	34.5	39.8	50.2	41.3	36.5	42.3	48.2	39.7	42.3	47.5	49.3
GFK [11]	38.3	12.7	40.2	39.5	42.3	15.8	39.2	48.3	39.5	26.8	28.3	45.3	39.2	38.3	42.8	29.3
LTSL [22]	47.2	42.3	50.8	53.5	38.3	41.3	41.2	56.7	41.8	50.7	48.2	54.7	43.3	48.2	52.3	58.8
L^2TSL [9]	52.3	48.7	59.7	**63.7**	49.8	43.2	49.3	60.7	48.3	**56.8**	50.8	55.7	46.3	67.5	58.2	**68.5**
Ours-I	46.22				**66.12**				50.92				50.72			
Ours-II	**74.03**				**89.83**				**79.06**				**73.47**			

Missing Modality Face Recognition – FG 2015

Ming Shao, Zhengming Ding, and Yun Fu

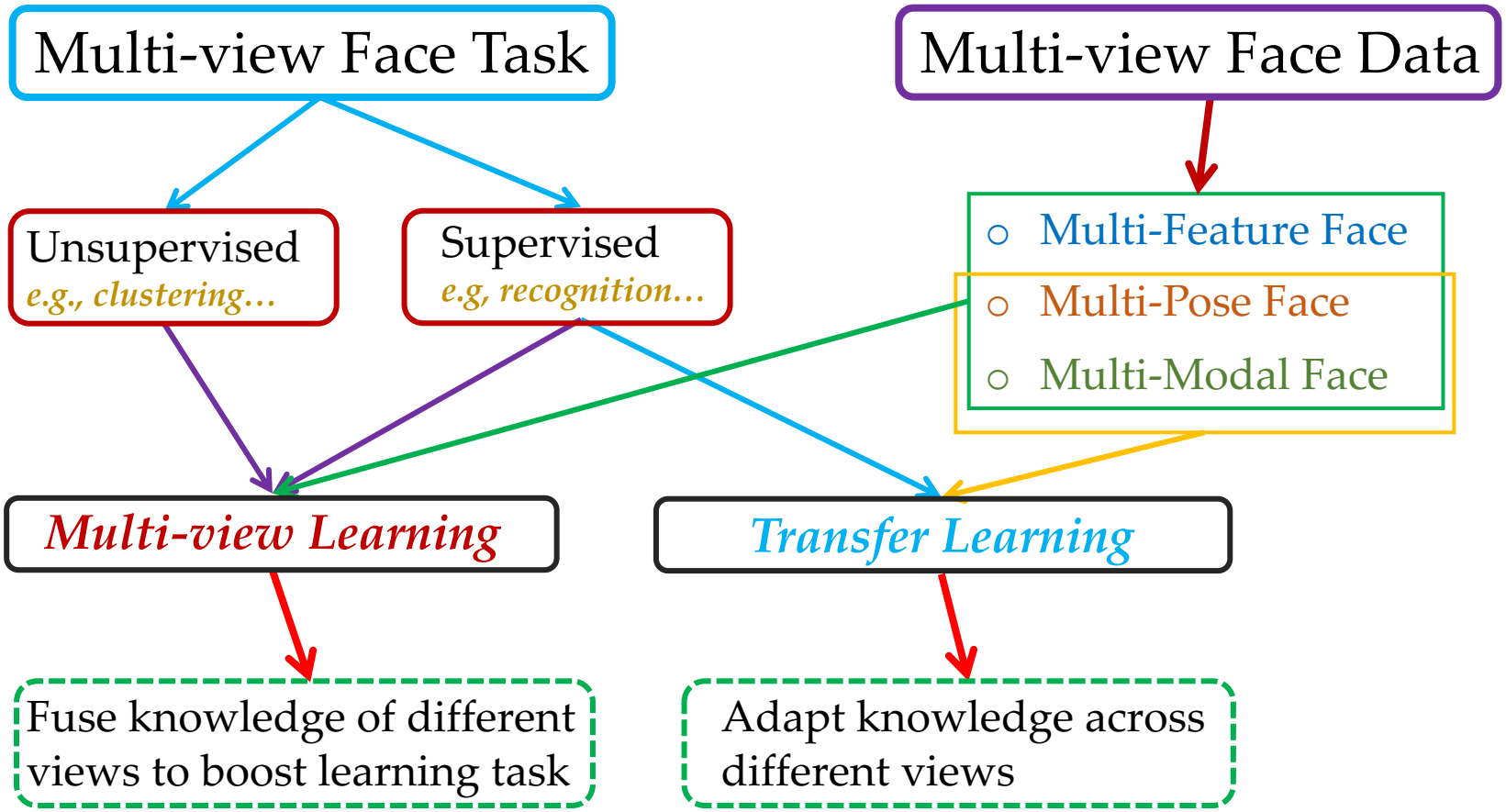
Outline

Northeastern University

Smile lab
Synergetic Media Learning Lab

- **Introduction & Background**
  - Multi-view Face Task
  - Multi-view Face Data
- **Unsupervised Multi-view Face Representation**
  - Methodology
  - Face Clustering, Outlier Detection
- **Supervised Multi-view Face Representation**
  - Multi-view Learning
  - Transfer Learning
- **Conclusion**

Conclusion [Flowchart]

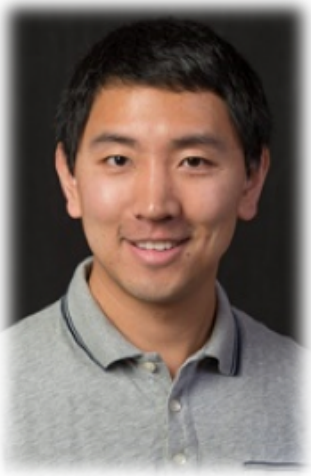


Collaborators

Northeastern University

Smile lab

Synergetic Media Learning Lab



Ming Shao

Sheng Li

Hongfu Liu

Thank you!

Q& A