Discriminative Covariance Oriented Representation Learning for Face Recognition with Image Sets

Wen Wang, Ruiping Wang , Shiguang Shan, Xilin Chen
Institute of Computing Technology (ICT), Chinese Academy of Sciences (CAS), China
wen.wang@vipl.ict.ac.cn, {wangruiping, sgshan, xlchen}@ict.ac.cn

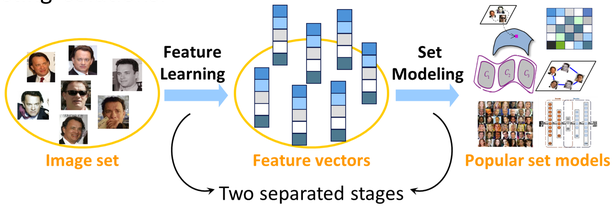
IEEE 2017 Conference on Computer Vision and Pattern Recognition
CVPR July 21-26 2017 HONOLULU

1. Problem & Motivation

◆ Problem

- Face Recognition with Image Sets

- Challenges: Large intra-class variations
- Existing solutions:



◆ Motivation:

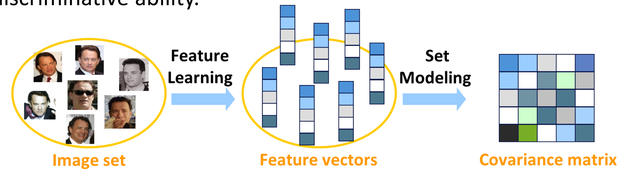
- Combine the two separated stages into a whole framework
- Image feature learning that facilitates image set classification
  - Image feature learning: Deep learning networks, e.g., CNN

$$X_i \mapsto h_i = \phi_\Theta(X_i)$$

  - Image set modeling: Set covariance matrices

$$C_i = \hat{h}_i^T \hat{h}_i, \text{ where } \hat{h}_i \text{ is the centered } h_i$$

  - Objective: Learn a feature learning network to project the images into a target feature space where set covariance matrices have maximum discriminative ability.



2. Our method (DCRL)

◆ Basic Idea

- Formulate the discrimination of set covariance by an objective function
- Solve the feature learning network by optimizing the objective function

◆ Graph Embedding Scheme

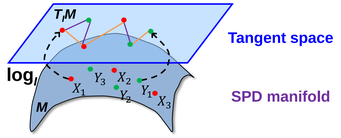
$$\Theta = \underset{\Theta}{\operatorname{argmin}} \sum_{i,j} A_{ij} LEM^2(C_i, C_j)$$

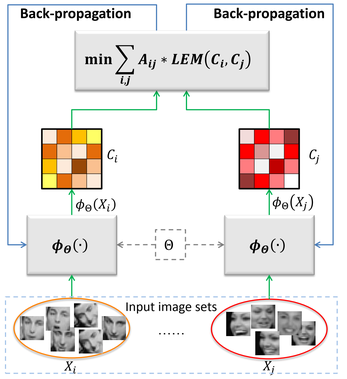
- Encode the relationship of covariance matrices with an adjacency graph

$$A_{ij} = \begin{cases} d_{ij} & if\ X_i \in N_w(X_j)\ or X_j \in N_w(X_i) \\ -d_{ij}, & if X_i \in N_b(X_j)\ or X_j \in N_b(X_i) \\ 0 & otherwise \end{cases}$$

$$d_{ij} = \exp(-LEM^2(C_i, C_j)/\sigma^2)$$

$$LEM(C_i, C_j) = \left\| \log(C_i) - \log(C_j) \right\|_F$$





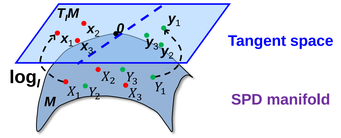
◆ Softmax Regression Scheme

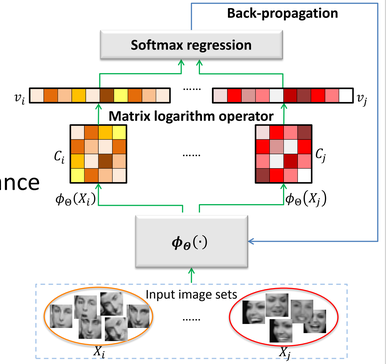
$$\Theta = \underset{\Theta}{\operatorname{argmin}} \sum_{i=1}^{n} f_{\text{softmax}}(v_i)$$

- Log-covariance vector:

$$v_i = vec(\log(C_i))$$

- Train a Softmax classifier for covariance matrices on a flat tangent space





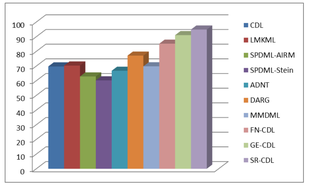
3. Experiments

◆ Three challenging databases

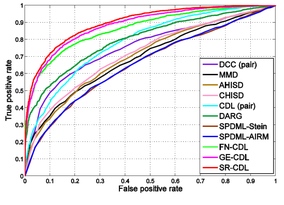
- YTC [CVPR'2008]
  - 47 subjects, 1,910 videos
- YTF [CVPR'2011]
  - 1,595 subjects, 3,425 videos
- PaSC [BTAS'2013]
  - 265 subjects, 2,802 videos

◆ Results:

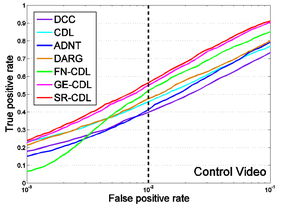
- Comparisons on YTC

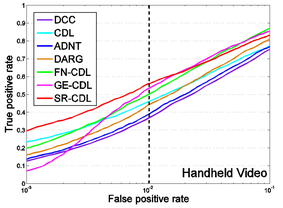


- Comparisons on YTF



- Comparisons on PaSC





◆ Conclusion

- Set covariance matrix model outperforms other set models promisingly
  -> the capability of covariance matrix to characterize image set structure
- Our DCRL shows better performance than other set covariance methods
  -> the superiority of the well learned image representations
- Our DCRL yields better results than deep methods ADNT and MMDML
  -> the explicit use of discriminative information and set covariance model