Embedding shared low-rank and feature correlation for multi-view data analysis

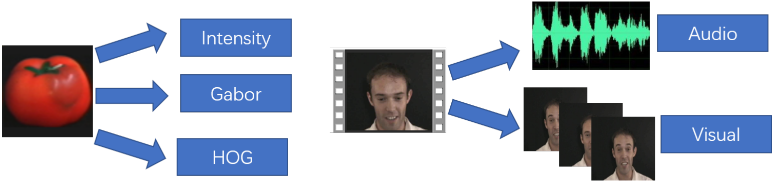
Zhan Wang[1], Lizhi Wang[1], Lei Zhang[1], Hua Huang[2]

[1]School of Computer Science and Technology, Beijing Institute of Technology, Beijing, China
[2] School of Artificial Intelligence, Beijing Normal University, Beijing, China

ICPR 2020

Outline

Motivation



- The image or video can be represented as multiple views.
- Exploring structure and common component form multi-view data is important.

Motivation

- Multi-view learning is a solution for multi-view data analysis and multi-view subspace learning is a kind of representative method.
- Conventional multi-view subspace learning methods such as CCA, PLS, have been proposed.
  - The correlations among multiple heterogeneous features can not directly characterize the structure information and the consensus of multi-view data
  - These methods cannot deal with noisy data well.
- Our proposed method: Integrating self-expressiveness and correlation-based subsapce learning.

Related works

- Low-rank representation:

$$\min_{\mathrm{H},\mathrm{O}} \|\mathrm{H}\|_* + \eta\|\mathrm{O}\|_{2,1}, s.t.\ \mathrm{A} = \mathrm{AH} + \mathrm{O}, \tag{1}$$

where $\mathrm{A} \in \mathbb{R}^{d\times n}$: data matrix, $\mathrm{H} \in \mathbb{R}^{n\times n}$: representative matrix, $\mathrm{O} \in \mathbb{R}^{d\times n}$: the error component.

- Correlation-based multi-view subspace learning

$$\max_{\mathrm{P}_1,\cdots,\mathrm{P}_m} \sum_i \sum_{j,i\neq j} tr(\mathrm{P}_i^T \mathrm{C}_{ij} \mathrm{P}_j), s.t.\ \mathrm{P}_i^T \mathrm{C}_{ii} \mathrm{P}_i = \mathrm{I}, i = 1, \cdots, m, \tag{2}$$

where $\mathrm{C}_{ij} = \mathrm{X}_i \mathrm{X}_j^T (i \neq j)$ and $\mathrm{C}_{ii} = \mathrm{X}_i \mathrm{X}_i^T$.

Proposed method

$$\begin{aligned}
\min_{\substack{\mathrm{U},\mathrm{V}\\ \mathrm{Z},\mathrm{P}_i,\mathrm{E}_i}} \; & \|\mathrm{U}\|_* + \|\mathrm{V}\|_* + \lambda \sum_i \|\mathrm{E}_i\|_{2,1} \\
& - \alpha \sum_i \sum_{j, i \neq j} tr(\mathrm{P}_i^T \mathrm{C}_{ij} \mathrm{P}_j) + \beta \sum_i \|\mathrm{P}_i\|_{2,p}^p \\
s.t. \; & \mathrm{P}_i^T \mathrm{X}_i = \mathrm{P}_i^T \mathrm{X}_i \mathrm{Z} + \mathrm{E}_i, \mathrm{Z} = \mathrm{UV}, \mathrm{P}_i^T \mathrm{P}_i = \mathrm{I} \\
& i = 1, 2, \cdots, m.
\end{aligned} \tag{3}$$

- Multiple views have common representation matrix.
- Double nuclear norm is used to construct more accurate low-rank representation matrix.
- Correlation analysis reveal the correlation of multi-view data.

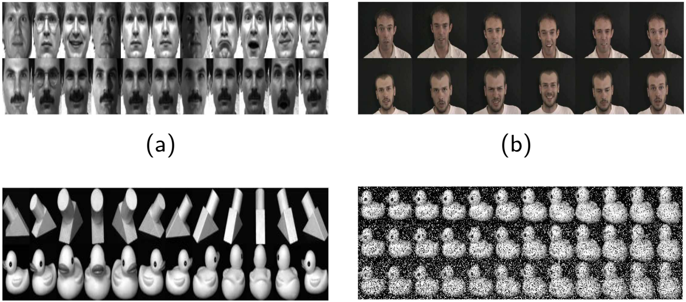
Proposed method

$$\begin{aligned}\min_{\substack{U,V\\Z,P_i,E_i}} & \|U\|_* + \|V\|_* + \lambda \sum_i \|E_i\|_{2,1} \\ & - \alpha \sum_i \sum_{j,i\neq j} tr(P_i^T C_{ij} P_j) + \beta \sum_i \|P_i\|_{2,p}^p \\ s.t.\ & P_i^T X_i = P_i^T X_i Z + E_i, Z = UV, P_i^T P_i = I \\ & i = 1, 2, \cdots, m.\end{aligned} \tag{4}$$

- $\ell_{2,p}$ norm on projection matrix is used to select a subset with representative features from the multi-view data
- A unify framework, which explores the structure and correlation, is develop for seeking the robust projection.
- ADMM algorithm is adopted to solve the optimization problem.

Experiments

- Experimental datasets: Yale face, eNTERFACE'05 audio-visual emotion, and COIL-20 object object datasets



(a) (b)

(c) (d)

Figure: Some sample images used in the experiments. (a) Yale face dataset, (b) eNTERFACE'05 emotion dataset and (c) COIL-20 object dataset.

Experiments

- Experimental setting:
  Yale:intensity feature (4096), LBP feature (3304) and Gabor feature (6750);
  eNTERFACE'05: acoustic feature IS10 (1582) and dynamic textures feature HOG-TOP (1728);
  COIL-20: intensity feature (4096), HOG feature (1764) and Gabor feature (1280).
- Comparison methods:
  Single-view algorithms, including PCA and LRE;
  Multi-view subspace learning algorithms, including MCCA, SCCA, PLS and MDcR.

Experiments

Recognition performances (%) on Yale dataset

Method	$l = 40$	$l = 50$	$l = 60$
PCA_{FC}	69.33±3.46	72.86±3.34	75.00±4.20
LRE_{FC}	69.92±2.31	74.10±3.93	77.89±3.37
MCCA	70.75±3.67	73.33±3.10	77.89±3.37
SCCA	67.50±3.83	70.38±2.11	72.78±3.79
PLS	70.58±4.30	72.86±3.15	75.00±2.58
MDcR	73.00±2.81	76.67±3.68	81.78±3.60
ESLRFC	**73.75±3.50**	**77.71±3.94**	**82.89±3.40**

Recognition performances (%) on eNTERFACE'05 dataset

Method	$l = 60$	$l = 80$	$l = 100$
PCA_{FC}	58.70±1.48	65.80±2.40	71.87±1.07
LRE_{FC}	58.75±1.43	65.90±2.37	71.90±1.10
MCCA	53.21±1.93	61.35±2.30	67.69±2.74
SCCA	56.33±1.97	63.16±2.25	69.03±2.03
PLS	53.90±1.34	60.84±2.65	64.95±2.90
MDcR	**63.30±1.58**	70.57±2.18	76.38±2.69
ESLRFC	63.08±1.71	**71.65±2.60**	**77.95±2.19**

Experiments

Recognition performances (%) on COIL-20 dataset

Method	$l = 5$	$l = 10$	$l = 15$
PCA_{FC}	84.78±1.74	90.59±0.96	95.32±0.59
LRE_{FC}	84.96±1.41	91.31±1.06	95.38±0.79
MCCA	83.62±1.57	91.06±1.14	94.79±0.86
SCCA	83.75±1.74	90.67±0.98	95.21±0.57
PLS	85.01±1.44	91.23±1.01	95.26±0.79
MDcR	85.23±1.61	91.42±1.22	95.51±0.66
ESLRFC	**85.78±1.93**	**92.19±0.97**	**95.88±0.73**

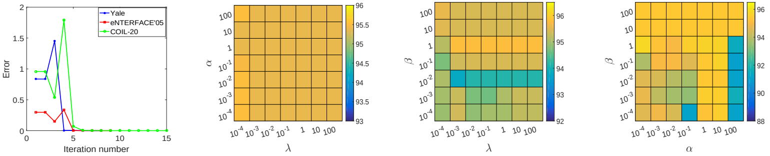
Recognition performances (%) on noisy COIL-20 dataset.

Method	20% noise	30% noise	40% noise
PCA_{FC}	92.98±0.68	90.04±0.56	84.54±1.46
LRE_{FC}	93.71±0.63	92.61±0.64	90.72±0.91
MCCA	92.90±0.70	90.08±0.85	83.80±1.17
SCCA	89.97±0.83	84.19±0.80	78.97±1.59
PLS	92.49±0.75	88.61±0.92	84.08±0.69
MDcR	94.10±0.56	92.62±0.52	89.85±0.78
ESLRFC	**94.48±0.90**	**93.14±0.81**	**90.86±1.04**

Experiments

Ablation study (%) of our method on three datasets

Method	Yale	eNTERFACE'05	COIL-20
ESLRFC$_\lambda$	82.11±2.84	76.13±2.26	94.87±0.66
ESLRFC$_\alpha$	81.67±3.15	68.85±0.84	95.34±0.65
ESLRFC$_\beta$	81.56±2.36	74.36±1.94	94.59±0.72
ESLRFC	**82.89±3.40**	**77.95±2.19**	**95.88±0.73**



(a) (b) (c) (d)

Figure: The sensitive analysis of ESLRFC. (a) convergence, (b),(c) and (d) parameter influences.

Conclusion

- Propose a novel unsupervised method, named embedding shared low-rank and feature correlation (ESLRFC), for multi-view data analysis.
  - A robust shared low-rank model is first proposed to explore the global structure and consistency of multi-view data in the embedding subspace, and to make our method robust to noise.
  - A kind of correlation analysis is developed to explore the correlations among multi-view features and remove the redundancy of each view.
- Seeking robust projection matrices for obtaining common subspace.
- Good results on the classification tasks

Thank you!