稀疏典型相关分析：从矩阵到张量

单　位：上海大学自动化系

报告人：修贤超

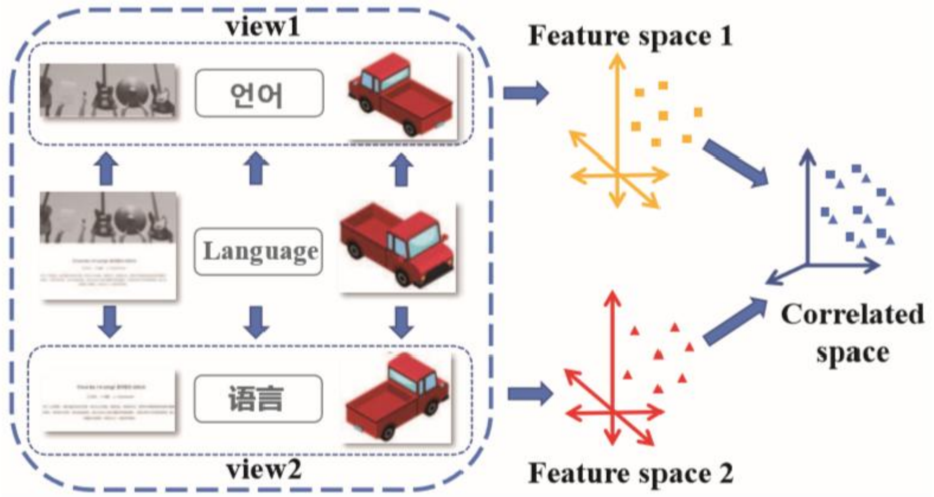
2023年5月12日-14日

汇报提纲

一	研究背景意义
二	主要研究内容
三	下一步工作安排

研究背景意义

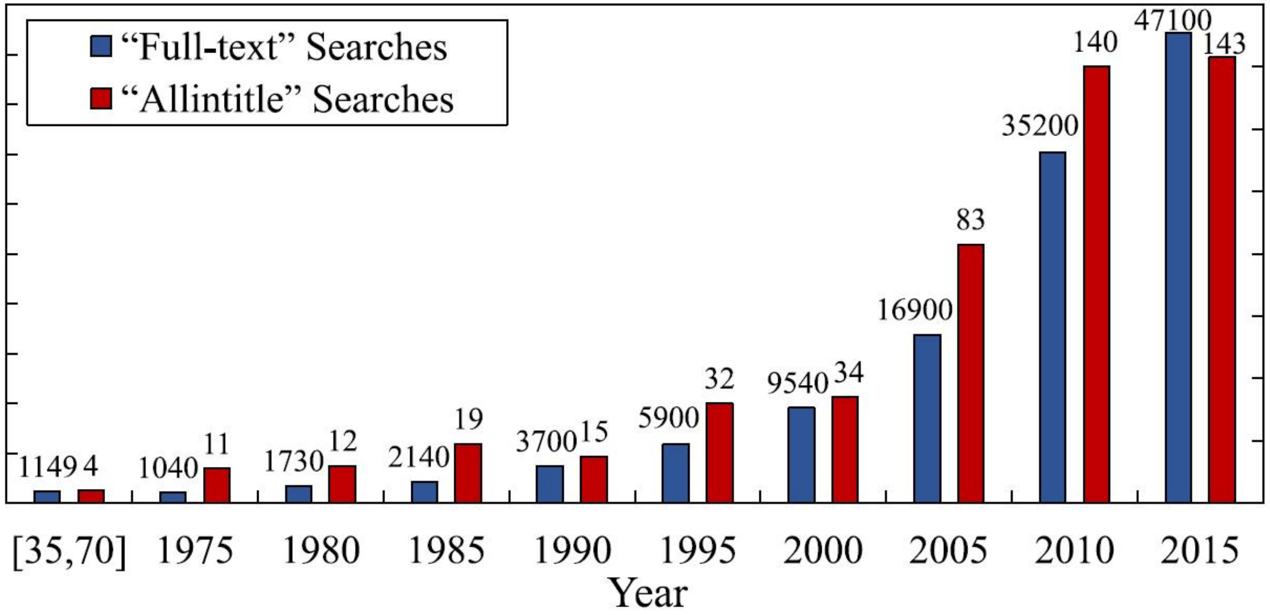
典型相关分析(Canonical Correlation Analysis, CCA)能够融合两组或多组数据来揭露其内部联系，进而提高特征选择的效率与精度。



$$\min_{A,B} \; -\mathrm{Tr}(A^T X^T Y B)$$
$$\text{s.t.} \;\; A^T X^T X A = I, \;\; B^T Y^T Y B = I$$

- Hotelling, *Relations between two sets of variates*, Biometrika, 28(3-4): 321-377, 1936.
- Hardoon-Szedmak-Shawe-Taylor, *Canonical correlation analysis: An overview with application to learning methods*, Neural Computation, 16(12): 2639-2664, 2004.

研究背景意义



- Yang-Liu-Liu-Tao, *A survey on canonical correlation analysis*, IEEE Transactions on Knowledge and Data Engineering, 33(6): 2349-2368, 2021.

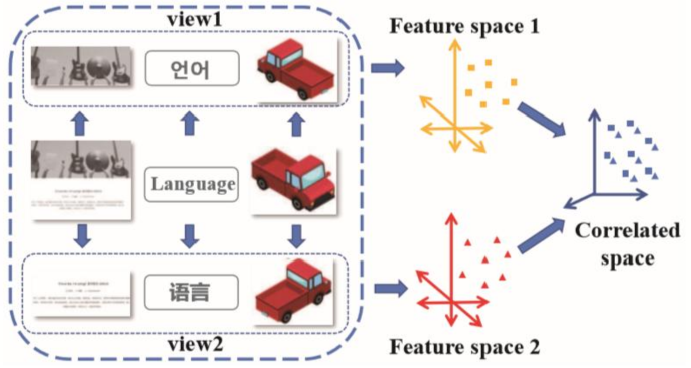
研究背景意义

统计 / 机器学习 / 优化

- 稀疏：Witten-Tibshirani-Hastie, Biostatistics, 2009
- 深度：Andrew-Arora-Bilmes-Livescu, ICML, 2013
- 协同回归：Gross-Tibshirani, Biostatistics, 2015
- 多分类: Sun-Ji-Ye, IEEE TPAMI, 2010
- 医学图像： Du-Huang-Yan-et al. , Bioinformatics, 2016
- 数据表示：Xu-Zhu-Zhang-Zhao, IEEE TC, 2019
- 最小二乘算法：Chu-Liao-Ng-Zhang, IEEE TPAMI, 2013
- 流形算法：Chen-Ma-Xue-Zou, IJOO, 2020
- 子空间算法：Mikael-Kanatsouli-Sidiropoulos, IEEE TSP, 2021

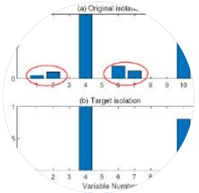
研究背景意义

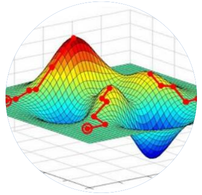
聚焦稀疏典型相关分析，**建立**优化模型与算法，**解决**实际工程问题。

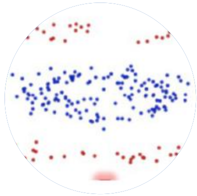


现存问题

如何提高过程变量的解释性 + 如何设计出快速的优化方法 + 如何处理过程数据的非线性







解决方案

稀疏优化 流形优化 + 随机算法 牛顿算法 + 深度学习 迁移学习

优化理论

算法设计

仿真实验

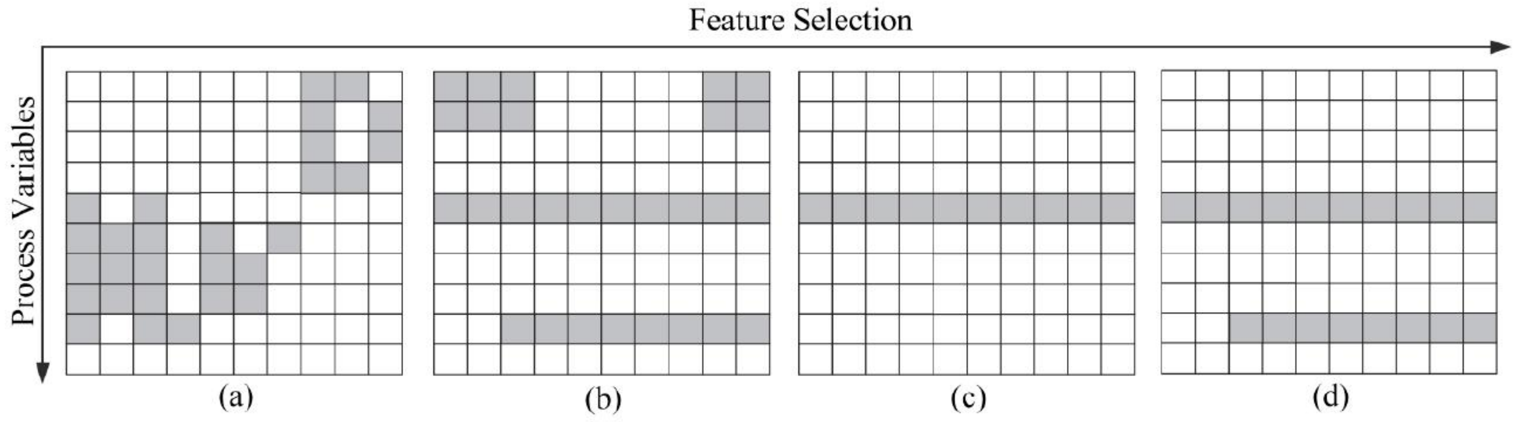
工程应用

汇报提纲

一	研究背景意义
二	主要研究内容
三	下一步工作安排

主要研究内容1

➢ 关于L1 、 L2,1 、L2,0



$$(a)\ \|\mathrm{X}\|_1 = \sum_{i=1}^{m}\sum_{j=n}^{p}|x_{ij}|$$

$$(b)\ \|\mathrm{X}\|_{2,1} = \sum_{i=1}^{m}\|\mathrm{x}_i\|_2$$

$$(c,d)\ \|\mathrm{X}\|_{2,0} = \sharp\{i : \|\mathrm{x}_i\|_2 \neq 0\}$$

- Nie-Huang-Cai-Ding, *Efficient and robust feature selection via joint L2, 1-norms minimization*, Advances in Neural Information Processing Systems, 2010.
- Liu-Ji-Ye, *Multi-task feature learning via efficient L2, 1-norm minimization*, arXiv:1205.2631.
- Cai-Nie-Huang, *Exact top-k feature selection via L2, 0-norm constraint*, Twenty-third International Joint Conference on Artificial Intelligence, 2013.
- Nie-Dong-Tian-Wang-Li, *Unsupervised feature selection with constrained L2,0-norm and optimized graph*, IEEE Transactions on Neural Networks and Learning Systems, 33(4): 1702–1713, 2022.
- Nie-Tian-Wang, *Learning feature-sparse principal subspace*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 45(4): 4858-4869, 2023.

主要研究内容1

➢ 正则 / 约束

正则

- Proximal算子
- 调节正则化参数
- 收敛性相对容易分析，如ADMM

原问题

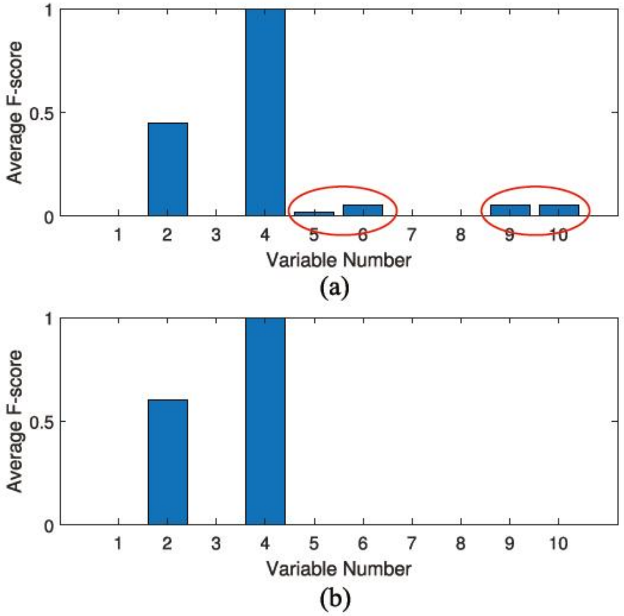
约束

- Projection算子
- 需提前告知稀疏度
- 可以与神经网络结合，如dropout

- Boyd-Parikh-Chu-Peleato-Eckstein, *Distributed optimization and statistical learning via the alternating direction method of multipliers*, Foundations and Trends in Machine learning, 3(1): 1-122, 2011.
- Parikh-Boyd, *Proximal algorithms*, Foundations and trends in Optimization, 1(3), 127-239, 2014.
- Lu, *Iterative hard thresholding methods for L0 regularized convex cone programming*, Mathematical Programming, 147(1): 125-154, 2014.
- Bian-Chen, *A smoothing proximal gradient algorithm for nonsmooth convex regression with cardinality penalty*, SIAM Journal on Numerical Analysis, 58(1): 858-883, 2020.
- Hazimeh-Mazumder, *Fast best subset selection: Coordinate descent and local combinatorial optimization algorithms*, Operations Research, 68(5): 1517-1537, 2020.

主要研究内容1

➢ 结构稀疏典型相关分析



$$\min_{\mathrm{A,B}} \; -\mathrm{Tr}(\mathrm{A}^T\mathrm{X}^T\mathrm{YB})$$
$$\text{s.t.} \;\; \mathrm{A}^T\mathrm{X}^T\mathrm{XA}=\mathrm{I},\; \mathrm{B}^T\mathrm{Y}^T\mathrm{YB}=\mathrm{I}$$

$$\Downarrow$$

$$\min_{\mathrm{A,B}} \; -\mathrm{Tr}(\mathrm{A}^T\mathrm{X}^T\mathrm{YB})$$
$$\text{s.t.} \;\; \mathrm{A}^T\mathrm{X}^T\mathrm{XA}=\mathrm{I},\; \mathrm{B}^T\mathrm{Y}^T\mathrm{YB}=\mathrm{I},$$
$$\|\mathrm{A}\|_{2,0} \le s_1,\; \|\mathrm{B}\|_{2,0} \le s_2$$

- Xiu-Pan-Yang-Liu, *Efficient and fast joint sparse constrained canonical correlation analysis for fault detection*, **IEEE Transactions on Neural Networks and Learning Systems, DOI: 10.1109/TNNLS.2022.3201881.**

主要研究内容1

➢ 交替极小化算法

$$\begin{aligned}\min_{A,B}\quad & -\mathrm{Tr}(A^TX^TYB)\\ \text{s.t.}\quad & A^TX^TXA=I,\ B^TY^TYB=I,\\ & \|A\|_{2,0}\le s_1,\ \|B\|_{2,0}\le s_2\end{aligned}$$

$$\Downarrow$$

$$\begin{aligned}\min_{A,B,C,D}\quad & -\mathrm{Tr}(C^TD)\\ \text{s.t.}\quad & C^TC=I,\ D^TD=I,\\ & \|A\|_{2,0}\le s_1,\ \|B\|_{2,0}\le s_2,\\ & XA=C,\ YB=D\end{aligned}$$

$$\Downarrow$$

$$\begin{aligned}\min_{A,B,C,D}\quad & -\mathrm{Tr}(C^TD)+\frac{\beta_1}{2}\|XA-C\|_F^2+\frac{\beta_2}{2}\|YB-D\|_F^2\\ \text{s.t.}\quad & \|A\|_{2,0}\le s_1,\ \|B\|_{2,0}\le s_2,\\ & C^TC=I,\ D^TD=I\end{aligned}$$

算法 1.1: AMA 算法

步 0. 给定数据 X, Y, 参数 $\beta_1,\beta_2>0$, $\epsilon>0$, 稀疏度 $s_1,s_2>0$.

步 1. 更新 $A^{k+1}\in\arg\min_{A\in\mathcal{S}_1}\ F(A,B^k,C^k,D^k)$.

步 2. 更新 $B^{k+1}\in\arg\min_{B\in\mathcal{S}_2}\ F(A^{k+1},B,C^k,D^k)$.

步 3. 更新 $C^{k+1}\in\arg\min_{C\in\mathcal{M}_1}\ F(A^{k+1},B^{k+1},C,D^k)$.

步 4. 更新 $D^{k+1}\in\arg\min_{D\in\mathcal{M}_2}\ F(A^{k+1},B^{k+1},C^{k+1},D)$.

步 5. 如果满足停机准则, 则停止; 否则返回步 1.

主要研究内容1

子问题求解

$$\min_{\mathrm{A}\in\mathcal{S}_1} \; G(\mathrm{A}) := \frac{1}{2}\|\mathrm{XA}-\mathrm{C}\|_F^2$$

$$\Downarrow$$

算法 1.2: IIHT 算法

步 0. 给定数据 X, C, 参数 $\eta\in(0,1)$, $\sigma>0$, $\epsilon_1,\epsilon_2>0$, 稀疏度 s_1.

步 1. 更新

$$\mathrm{A}^{k+1}=\Pi_{\mathcal{S}_1}\left(\mathrm{A}^k-\alpha_k\nabla G(\mathrm{A}^k)\right),$$

这里 $\alpha_k=\alpha_0\eta^{q_k}$ 且 q_k 是满足下述条件最小的非负整数 q

$$G\left(\mathrm{A}^k(\alpha_0\eta^q)\right)\le G(\mathrm{A}^k)-\frac{\sigma}{2}\|\mathrm{A}^k(\alpha_0\eta^q)-\mathrm{A}^k\|_F^2,$$

其中 $\mathrm{A}^k(\alpha):=\Pi_{\mathcal{S}_1}(\mathrm{A}^k-\alpha\nabla G(\mathrm{A}^k))$.

$$\min_{\mathrm{C}\in\mathcal{M}_1} \; H(\mathrm{C}) := -\mathrm{Tr}(\mathrm{C}^T\mathrm{D})+\frac{\beta_1}{2}\|\mathrm{XA}-\mathrm{C}\|_F^2$$

$$\Downarrow$$

算法 1.3: MCGD 算法

步 0. 给定数据 X, A, D, 参数 $\eta\in(0,1)$, $\beta_1>0$, $\delta>0$, $\epsilon_3>0$.

步 1. 计算

$$\mathrm{V}^k\in\operatorname{argmin}_{\mathrm{V}\in\mathcal{T}_{\mathcal{M}_1}(\mathrm{C}^k)} -\mathrm{Tr}((\mathrm{C}^k+\mathrm{V})^T\mathrm{D})+\frac{\beta_1}{2}\|\mathrm{XA}-\mathrm{C}^k-\mathrm{V}\|_F^2.$$

步 2. 更新

$$\mathrm{C}^{k+1}=\mathcal{R}_{\mathrm{C}^k}\left(\gamma_k\mathrm{V}^k\right),$$

其中 $\gamma_k=\gamma_1\eta^{q_k}$, 且 q_k 是满足下述条件最小的非负整数 q

$$H\left(\mathrm{C}^{k+1}\right)\le H(\mathrm{C}^k)-\delta\gamma_1\eta^q\|\mathrm{V}^k\|_F^2.$$

- Chen-Ma-Xue-Zou, *An alternating manifold proximal gradient method for sparse principal component analysis and sparse canonical correlation analysis*, INFORMS Journal on Optimization, 2(3): 192-208, 2020.

主要研究内容1

➢ 数值试验——仿真

$$\mathbf{X} = \mathbf{u}^\top(\mathbf{v}_1 + \mathbf{e}_1), \quad \mathbf{Y} = \mathbf{u}^\top(\mathbf{v}_2 + \mathbf{e}_2)$$

where $\mathbf{u} \in \mathbb{R}^{1\times 1000i}$ is a random vector that satisfies standard normal distributions, $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^{1\times 500j}$ are given by

$$\mathbf{v}_1 = [\underbrace{1,\dots,1}_{50j}\ \underbrace{-1,\dots,-1}_{50j}\ \underbrace{0,\dots,0}_{400j}]$$

$$\mathbf{v}_2 = [\underbrace{0,\dots,0}_{400j}\ \underbrace{1,\dots,1}_{50j}\ \underbrace{-1,\dots,-1}_{50j}]$$

TABLE I

TIME (S) FOR THE SIMULATION DATASET

Problem Scale	CCA	SCCA	JSCCA	JSCCCA
(1,000;500;500)	0.08	0.09	0.09	**0.03**
(5,000;500;500)	0.35	0.32	0.36	**0.09**
(10,000;500;500)	0.65	0.61	0.62	**0.12**
(50,000;500;500)	4.20	3.28	3.15	**0.46**
(100,000;500;500)	7.22	6.44	6.34	**0.89**
(1,000;2,500;2,500)	1.16	1.25	1.19	**0.13**
(5,000;2,500;2,500)	2.95	2.84	2.73	**0.20**
(10,000;2,500;2,500)	5.47	5.35	5.22	**0.61**
(50,000;2,500;2,500)	25.82	23.66	20.45	**1.89**
(100,000;2,500;2,500)	127.04	45.39	40.10	**3.60**
(1,000;5,000;5,000)	5.25	5.97	5.74	**0.09**
(5,000;5,000;5,000)	10.78	9.85	9.72	**0.38**
(10,000;5,000;5,000)	17.25	16.73	14.95	**0.78**
(50,000;5,000;5,000)	113.06	58.12	54.07	**3.01**
(100,000;5,000;5,000)	8331.17	5249.65	4,989.36	**8.03**

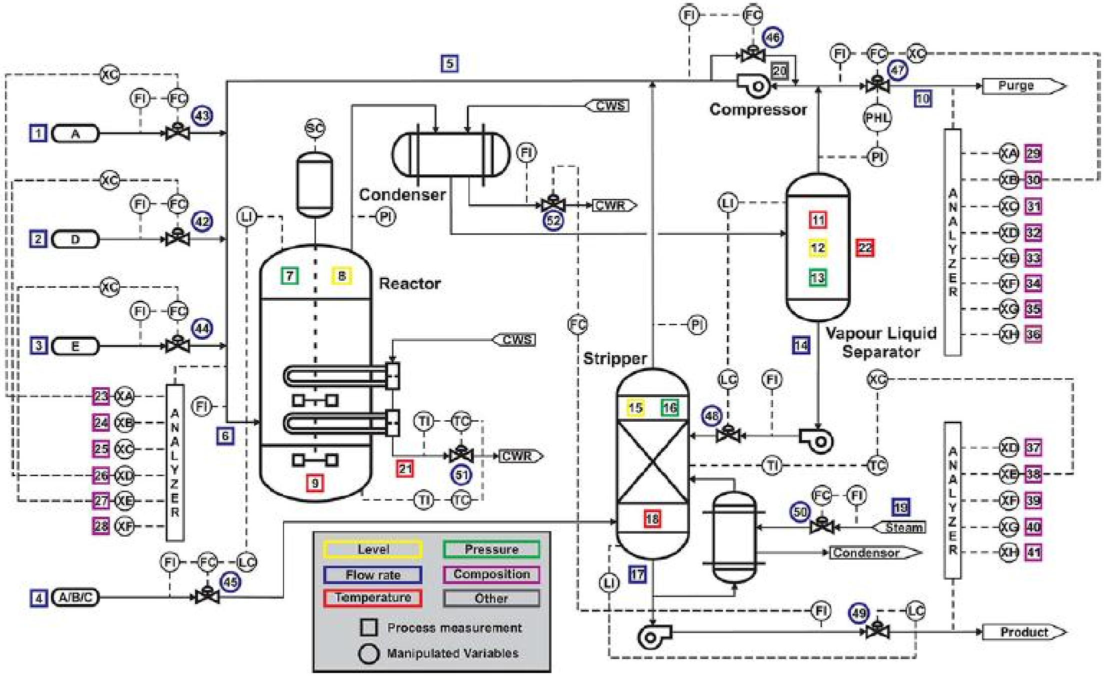
- Cai-Dan-Zhang, *L0-based sparse canonical correlation analysis with application to cross-language document retrieval*, Neurocomputing, 329: 32-45, 2019.

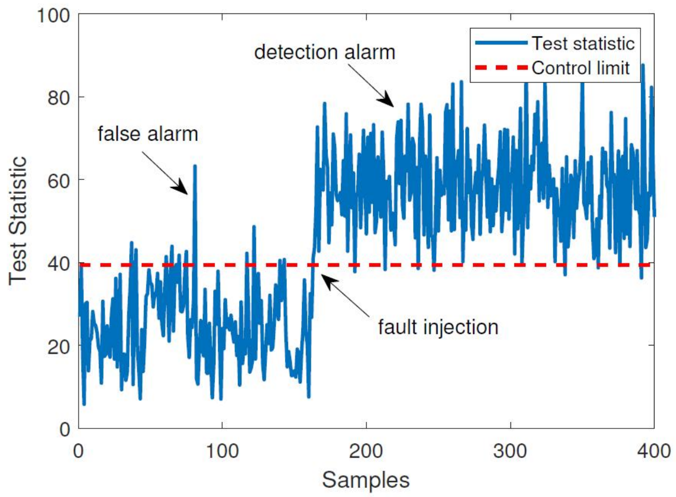
主要研究内容1

- **故障诊断**是过程控制领域的研究热点之一，包括故障检测、故障隔离、故障修复、寿命预测等。

主要研究内容1

➢ 数值试验——TE过程





$$\mathrm{FDR} = \mathrm{prob}(T^2 > J_{th,T^2} \mid f \neq 0)$$

$$\mathrm{FAR} = \mathrm{prob}(T^2 > J_{th,T^2} \mid f = 0)$$

主要研究内容1

➢ 数值试验——TE过程

Fault No.	CCA		SCCA		JSCCA		JSCCCA	
	FDR	FAR	FDR	FAR	FDR	FAR	FDR	FAR
IDV(01)	99.25%	**0.00%**	99.38%	**0.00%**	99.50%	**0.00%**	**99.62%**	**0.00%**
IDV(02)	98.62%	0.63%	99.25%	**0.00%**	99.25%	**0.00%**	**99.37%**	**0.00%**
IDV(03)	32.80%	2.50%	36.64%	1.88%	38.20%	**0.00%**	**43.98%**	**0.00%**
IDV(04)	**100%**	1.88%	**100%**	0.63%	**100%**	**0.00%**	**100%**	**0.00%**
IDV(05)	27.63%	1.88%	31.00%	0.63%	34.50%	**0.00%**	**38.36%**	**0.00%**
IDV(06)	99.75%	0.63%	99.90%	**0.00%**	**100%**	**0.00%**	**100%**	**0.00%**
IDV(07)	**100%**	1.88%	**100%**	0.63%	**100%**	0.63%	**100%**	**0.00%**
IDV(08)	93.25%	1.88%	95.00%	0.63%	95.24%	**0.00%**	**97.88%**	**0.00%**
IDV(09)	31.20%	3.13%	35.25%	2.50%	38.50%	0.63%	**39.87%**	**0.00%**
IDV(10)	27.50%	1.25%	34.82%	**0.63%**	36.24%	**0.63%**	**40.12%**	**0.63%**
IDV(11)	66.43%	0.63%	68.71%	**0.00%**	74.00%	**0.00%**	**77.89%**	**0.00%**
IDV(12)	90.87%	1.88%	93.87%	1.25%	94.50%	**0.63%**	**95.75%**	**0.63%**
IDV(13)	91.36%	0.63%	92.12%	0.63%	93.87%	**0.00%**	**95.61%**	**0.00%**
IDV(14)	86.00%	1.88%	87.25%	**0.63%**	88.12%	**0.63%**	**90.37%**	**0.63%**
IDV(15)	35.10%	2.50%	38.57%	1.25%	41.23%	0.63%	**43.57%**	**0.00%**
IDV(16)	15.78%	7.50%	18.23%	4.38%	22.75%	3.13%	**25.24%**	**1.25%**
IDV(17)	33.75%	3.13%	35.12%	3.13%	37.00%	3.13%	**39.75%**	**2.50%**
IDV(18)	87.88%	1.88%	90.18%	0.63%	92.54%	0.63%	**95.86%**	**0.00%**
IDV(19)	22.24%	**1.25%**	25.87%	**1.25%**	28.06%	**1.25%**	**29.78%**	**1.25%**
IDV(20)	47.50%	0.63%	52.75%	**0.00%**	54.37%	**0.00%**	**55.63%**	**0.00%**
IDV(21)	89.62%	1.25%	90.36%	1.25%	93.75%	**0.63%**	**96.85%**	**0.63%**

主要研究内容1

➢ 数值试验——TE过程

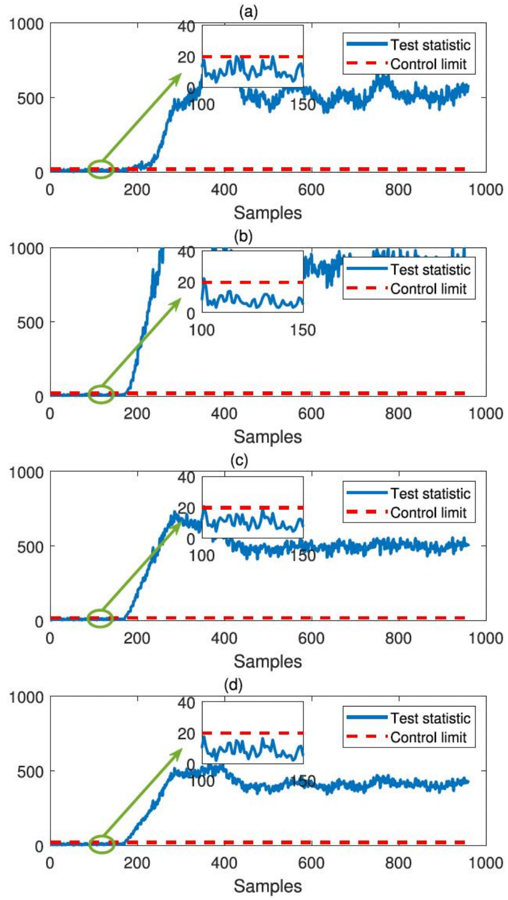


Fig. 3. Detection performance for IDV(02) in the TEP. (a) CCA. (b) SCCA. (c) JSCCA. (d) JSCCCA.

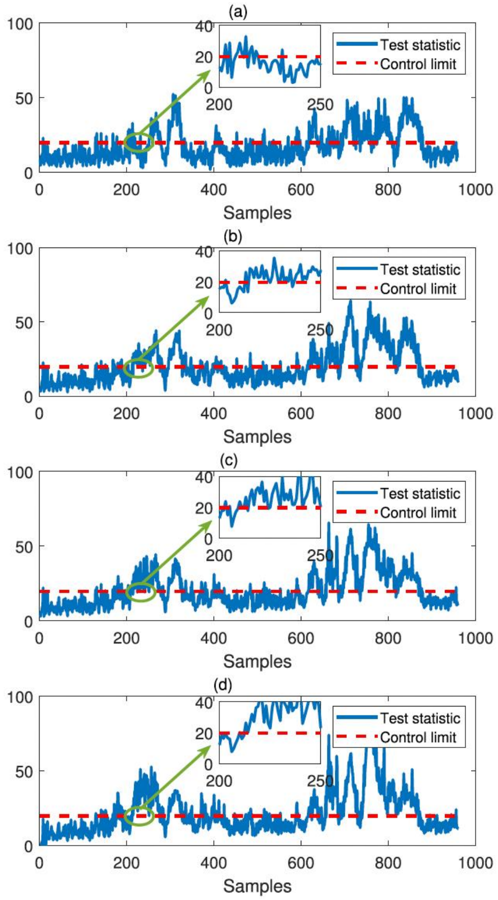


Fig. 4. Detection performance for IDV(10) in the TEP. (a) CCA. (b) SCCA. (c) JSCCA. (d) JSCCCA.

主要研究内容2

➢ 深度典型相关分析

$$
\begin{aligned}
\min_{A,B} \quad & -\operatorname{Tr}(A^T X^T Y B) \\
\text{s.t.} \quad & A^T X^T X A = I,\ B^T Y^T Y B = I, \\
& \|A\|_{2,0} \le s_1,\ \|B\|_{2,0} \le s_2
\end{aligned}
$$

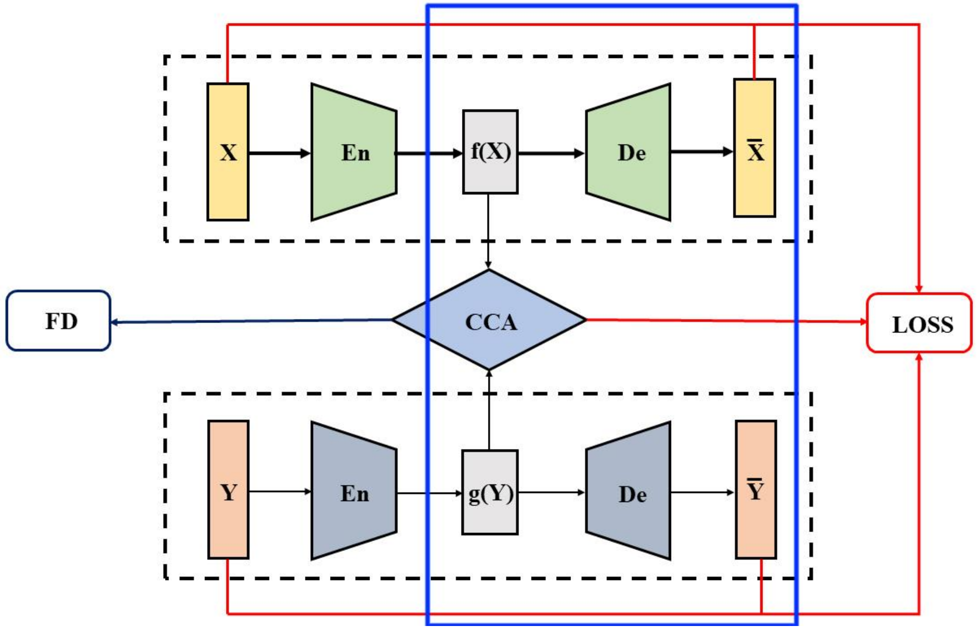
$$\Downarrow$$

$$
\begin{aligned}
\min_{A,B} \ & \text{loss1} = -\operatorname{Tr}(A^T f(X) g(Y)^T B) \\
\text{s.t.} \ & A^T f(X) f(X)^T A = I,\ B^T g(Y) g(Y)^T B = I, \\
& \|A\|_{2,0} \le s_1, \|B\|_{2,0} \le s_2
\end{aligned}
$$

- Xiu-Miao-Yang-Liu, *Deep canonical correlation analysis using sparsity constrained optimization for nonlinear process monitoring*, IEEE Transactions on Industrial Informatics, 18(10): 6690-6699, 2022.

主要研究内容2

➢ **解码模块: 重新映射回原始输入空间, 并进行误差拟合**



$$\text{loss2} = \sum_{i=1}^{N} \|\mathrm{x}_i - \bar{\mathrm{x}}_i\|^2 = \sum_{i=1}^{N} \|\mathrm{x}_i - p(f(\mathrm{x}_i))\|^2,$$

$$\text{loss3} = \sum_{i=1}^{N} \|\mathrm{y}_i - \bar{\mathrm{y}}_i\|^2 = \sum_{i=1}^{N} \|\mathrm{y}_i - q(g(\mathrm{y}_i))\|^2.$$

□ Wang-Arora-Livescu-Bilmes, *On deep multi-view representation learning*, In International Conference on Machine Learning (ICML), 1083-1092, 2015.

主要研究内容2

➢ 数值试验——TE过程

Fault No.	CCA		CCA-SCO		KCCA		KCCA-SCO		DCCA		DCCA-SCO	
	FDR	FAR	FDR	FAR	FDR	FAR	FDR	FAR	FDR	FAR	FDR	FAR
IDV(1)	99.75%	0.63%	99.75%	0.63%	99.88%	0.63%	99.88%	0.00%	99.88%	0.00%	99.88%	0.00%
IDV(2)	96.50%	0.63%	97.25%	0.63%	98.38%	0.00%	98.38%	0.00%	98.47%	0.00%	99.50%	0.00%
IDV(4)	100%	1.88%	100%	1.25%	100%	1.25%	100%	0.63%	100%	0.63%	100%	0.00%
IDV(5)	100%	3.75%	100%	3.25%	100%	2.50%	100%	2.50%	100%	2.50%	100%	1.88%
IDV(6)	100%	4.38%	100%	4.38%	100%	3.75%	100%	3.25%	100%	3.25%	100%	3.25%
IDV(7)	100%	3.75%	100%	3.25%	100%	2.50%	100%	1.75%	100%	1.25%	100%	0.63%
IDV(8)	96.50%	1.88%	97.50%	1.88%	97.85%	0.63%	98.25%	0.63%	98.88%	0.63%	99.38%	0.00%
IDV(10)	86.88%	1.25%	87.75%	0.63%	89.58%	0.63%	89.75%	0.63%	90.38%	0.00%	93.88%	0.00%
IDV(11)	76.50%	0.63%	77.50%	0.63%	78.50%	0.63%	79.63%	0.63%	80.13%	0.63%	84.50%	0.00%
IDV(12)	99.00%	1.25%	99.00%	0.63%	99.25%	0.00%	99.37%	0.00%	99.50%	0.63%	99.75%	0.00%
IDV(13)	95.75%	0.63%	96.13%	0.63%	96.50%	0.63%	96.50%	0.63%	96.75%	0.00%	96.88%	0.00%
IDV(14)	100%	1.88%	100%	1.25%	100%	0.63%	100%	0.63%	100%	0.63%	100%	0.63%
IDV(16)	93.00%	7.50%	94.38%	5.63%	95.63%	1.25%	96.63%	1.25%	96.63%	1.25%	98.75%	0.63%
IDV(17)	94.13%	3.13%	94.13%	2.50%	94.25%	2.50%	95.13%	1.75%	95.75%	1.25%	96.38%	1.25%
IDV(18)	90.88%	1.88%	91.25%	1.25%	92.50%	0.00%	92.75%	0.00%	93.50%	0.63%	95.63%	0.00%
IDV(19)	92.00%	1.25%	92.63%	1.25%	94.25%	1.25%	94.25%	0.63%	94.93%	0.63%	95.50%	0.63%
IDV(20)	86.88%	0.63%	87.13%	0.63%	87.75%	0.63%	87.75%	0.00%	88.88%	1.25%	89.38%	0.00%

主要研究内容2

➢ 数值试验——TE过程

TABLE II
DETECTION TIME IN SECONDS

CCA	CCA-SCO	KCCA	KCCA-SCO	DCCA	DCCA-SCO
0.044	0.072	15.676	22.556	2.806	3.097

TABLE III
MONITORING PERFORMANCE UNDER DIFFERENT LAYERS

Hidden layers	1	2	3
FDR	89.75%	93.88%	93.88%
FAR	0.63%	0.00%	0.00%

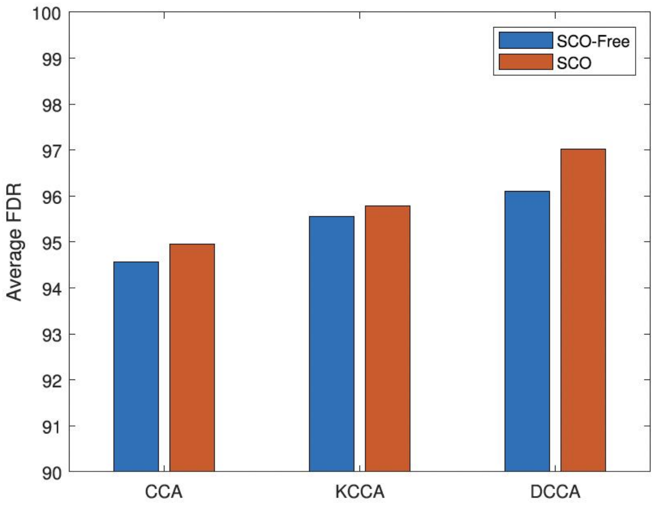


Fig. 10. Illustrations of the improvement.

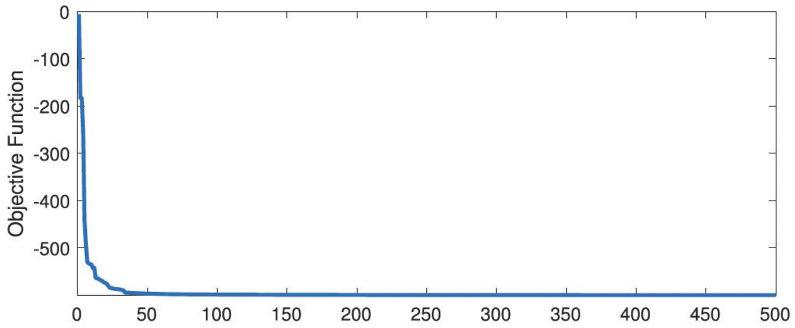


Fig. 9. Convergence of objective values.

主要研究内容2

➢ 数值试验——柴油发电机数据

TABLE IV
SELECTED VARIABLES IN THE DG PROCESS

Var.	Description
1	oil inlet temperature
2	cooling water outlet temperature
3	exhaust gas into supercharger A temperature
4	exhaust gas into supercharger B temperature
5	high temperature cooling water inlet pressure
6	starting air pressure
7	oil inlet pressure
8	running speed
9	running load
10	operating hours
11	bearing temperature
12	U-phase winding temperature
13	V-phase winding temperature
14	W-phase winding temperature

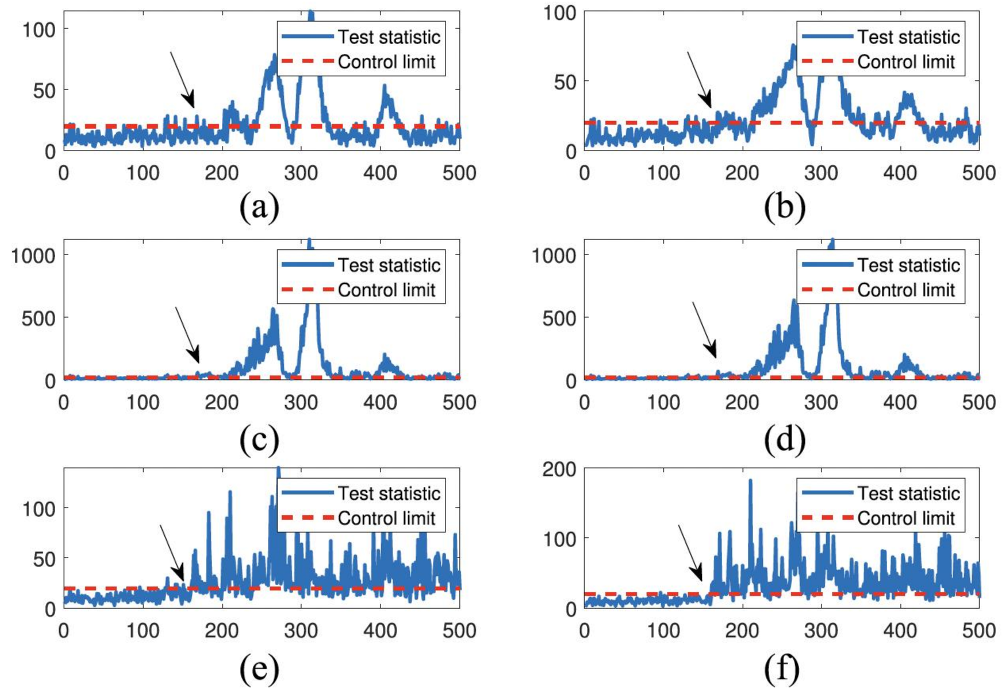
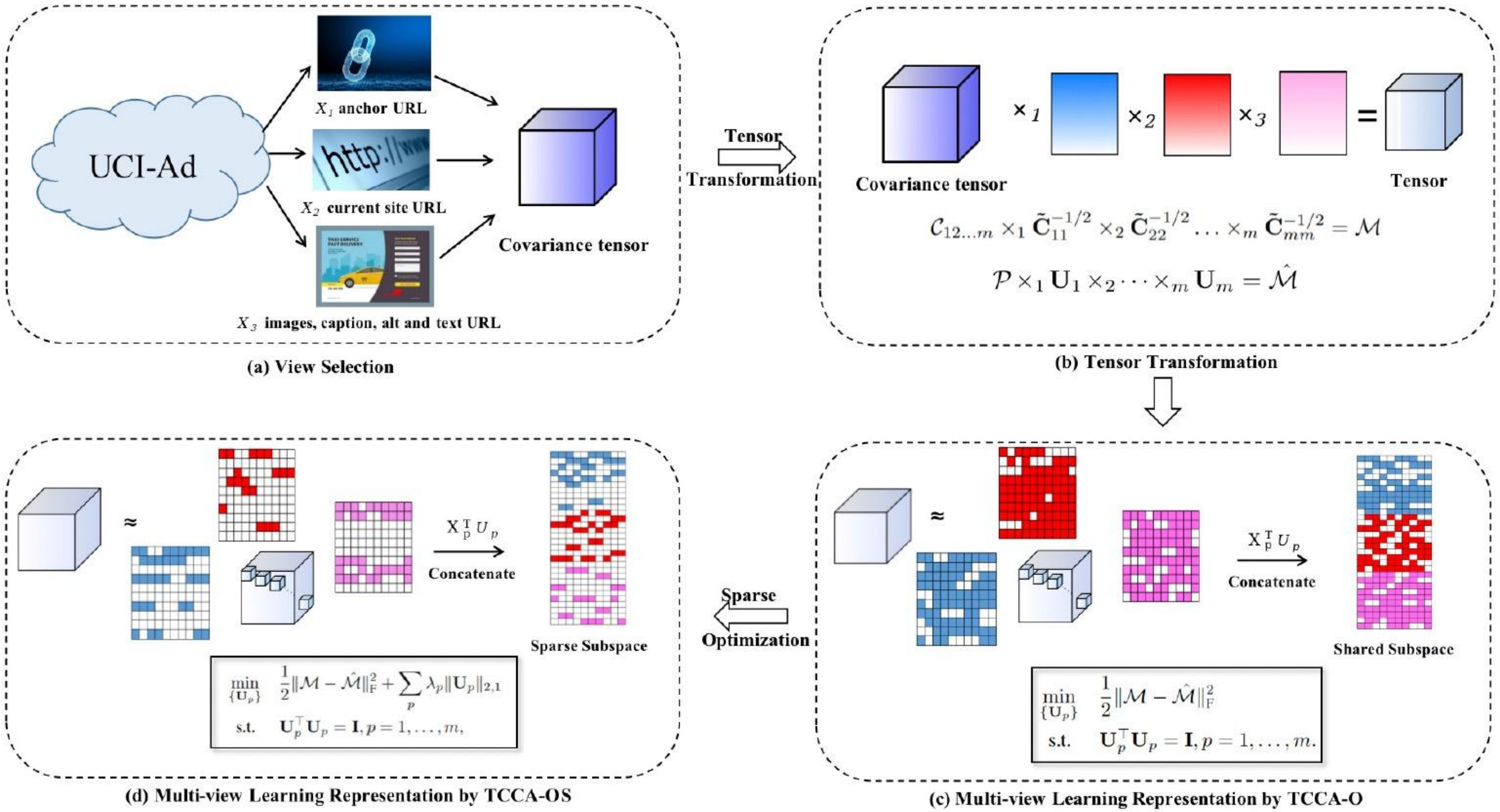


Fig. 11. Monitoring results for the DG process. (a) CCA. (b) CCA-SCO. (c) KCCA. (d) KCCA-SCO. (e) DCCA. (f) DCCA-SCO.

主要研究内容3

➢ 张量典型相关分析



- Sun-Xiu-Luo-Liu, *Learning high-order multi-view representation by new tensor canonical correlation analysis*, IEEE Transactions on Circuits and Systems for Video Technology, DOI: 10.1109/TCSVT.2023.3263853.

主要研究内容3

➢ 模型与算法

$$
\begin{aligned}
\min_{\mathcal{P},\{\mathbf{U}_p\}} \quad & \frac{1}{2}\|\mathcal{M}-\hat{\mathcal{M}}\|_{\mathrm{F}}^2 \\
\text{s.t.} \quad & \mathbf{U}_p^\top \mathbf{U}_p = \mathbf{I}, p = 1,\ldots,m
\end{aligned}
$$

$\Downarrow$

Algorithm 1 Optimization Algorithm for TCCA-O

Input: Multi-view data $\mathbf{X}$, parameters d_p, r. Calculate covariance matrix $\mathbf{C}_{pp}$, covariance tensor $\mathcal{C}_{12\cdots m}$, and tensor $\mathcal{M}$.

For $p = 1, \cdots, m$

1: Compute the matrix $\mathcal{M}_{(p)} = \mathbf{C}_{pp}\mathcal{C}_{12\cdots m(p)}(\mathbf{C}_{mm-1} \otimes \cdots \otimes \mathbf{C}_{pp+1} \otimes \mathbf{C}_{pp-1} \otimes \cdots \otimes \mathbf{C}_{11})^\top \in \mathbb{R}^{d_p \times (\prod_{i\neq p} d_i)}$;

2: Compute the singular value decomposition (SVD) of $\mathcal{M}_{(p)}$ and obtain the r leading left singular matrix $\mathbf{U}_p$;

3: Compute $\mathcal{P} = \mathcal{M} \times_1 \mathbf{U}_1^\top \times_2 \cdots \times_m \mathbf{U}_m^\top$.

End for

$$
\begin{aligned}
\min_{\mathcal{P},\{\mathbf{U}_p\}} \quad & \frac{1}{2}\|\mathcal{M}-\hat{\mathcal{M}}\|_{\mathrm{F}}^2 + \sum_p \lambda_p \|\mathbf{U}_p\|_{2,1} \\
\text{s.t.} \quad & \mathbf{U}_p^\top \mathbf{U}_p = \mathbf{I}, p = 1,\ldots,m
\end{aligned}
$$

$\Downarrow$

Algorithm 2 Optimization Algorithm for TCCA-OS

Input: Multi-view data $\mathbf{X}$, parameters $d_p, r, \lambda_p, t, \epsilon_1, \epsilon_2$. Calculate covariance matrix $\mathbf{C}_{pp}$, covariance tensor $\mathcal{C}_{12\cdots m}$, and tensor $\mathcal{M}$.

Initialize: Factor matrices $\{\mathbf{U}_p^0\}$ from the Tucker decomposition of $\mathcal{M}$ and $\{\mathbf{B}_p^0\}$.

While not converged **do**

1: Update $\mathcal{P}$

2: Update $\{\mathbf{V}_p\}$

3: Update $\{\mathbf{U}_p\}$

4: Update multipliers $\{\mathbf{B}_p\}$

5: Check the convergence conditions

End while

主要研究内容3

➢ 数值试验——数据集

Datasets	Views	Dim	Instance	Class
Caltech101-7	Gabor Wavelet moments CENTRIST HOG	48 40 254 1984	1474	7
NUS-WIDE	color auto-correlogram wavelet texture bag of visual words	144 128 500	11647	10
UCI-Ad	image, caption, alt text current site anchor URL	588 495 472	3279	3
BBC	View 1 View 2 View 3	4569 4633 4665	685	5

- ◻ Cai-Liu-Xiao-Lin, *Semi-supervised multi-view clustering based on orthonormality-constrained nonnegative matrix factorization*, Information Sciences, 536: 171-184, 2020.

主要研究内容3

➢ 数值试验——分类

TABLE II
THE CLASSIFICATION ACCURACY (%) OF ALL COMPARED METHODS UNDER BEST DIMENSIONS.

Methods	Caltech101-7	NUS-WIDE	UCI-Ad	BBC
KNN	80.93 (± 1.42)	29.20 (± 0.32)	92.42 (± 0.35)	80.88 (± 1.14)
CCA	83.33 (± 1.55)	29.72 (± 0.50)	88.17 (± 0.9)	83.12 (± 1.13)
SCCA	83.17 (± 1.85)	30.09 (± 0.36)	92.91 (± 0.42)	83.32 (± 2.63)
TCCA	87.83 (± 2.72)	30.17 (± 1.03)	94.78 (± 1.59)	74.63 (± 2.29)
TCCA-O	93.37 (± 0.95)	33.67 (± 0.84)	95.35 (± 0.64)	84.98 (± 1.75)
TCCA-OS	**93.69** (± **1.24**)	**33.73** (± **0.81**)	**96.07** (± **0.46**)	**87.56** (± **1.94**)

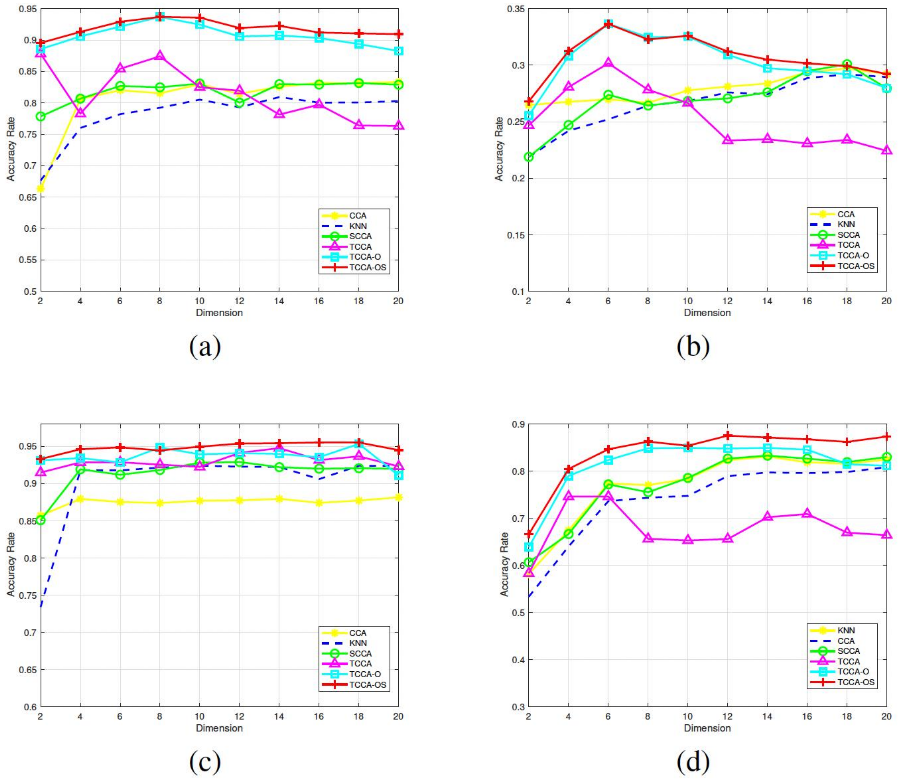


Fig. 3. The classification accuracy under different dimensions on (a) the Caltech101-7 dataset, (b) the NUS-WIDE dataset, (c) the UCI-Ad dataset, (d) the BBC dataset.

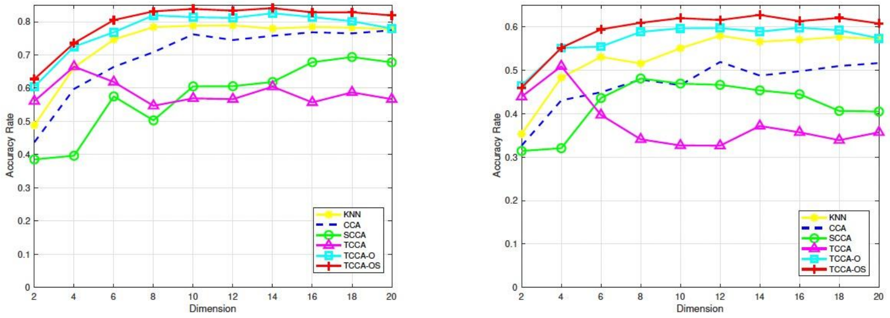
- Luo-Tao-Ramamohanarao-Xu-Wen, *Tensor canonical correlation analysis for multi-view dimension reduction*, IEEE Transactions on Knowledge and Data Engineering, 27(11): 3111-3124, 2015.

主要研究内容3

➢ 数值试验——鲁棒性分析

TABLE III
THE CLASSIFICATION ACCURACY (%) UNDER BEST DIMENSIONS ON THE NUS-WIDE DATASET WITH DIFFERENT NOISE LEVELS.

Methods	30 %	60 %	90 %
KNN	25.51 (± 0.21)	23.28 (± 0.53)	21.63 (± 1.91)
CCA	24.95 (± 0.57)	21.52 (± 0.85)	21.95 (± 0.97)
SCCA	25.27 (± 0.13)	24.31 (± 0.83)	22.08 (± 0.61)
TCCA	27.57 (± 0.73)	24.06 (± 0.15)	22.36 (± 1.28)
TCCA-O	30.11 (± 3.09)	25.62 (± 0.70)	23.41 (± 1.62)
TCCA-OS	**30.77** (± **0.50**)	**26.13** (± **0.70**)	**23.86** (± **1.74**)



(a) (b)

Fig. 6. The classification accuracy under different dimensions on the BBC dataset with (a) Gaussian noise, (b) non-Gaussian noise.

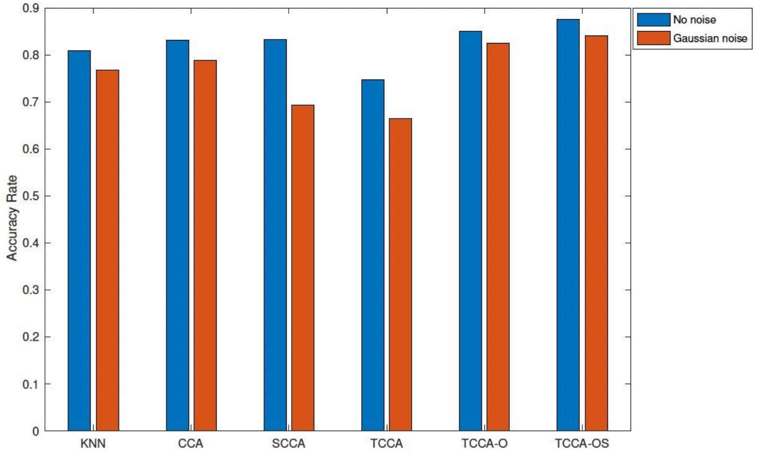


Fig. 7. The performance of all compared methods without noise and with Gaussian noise.

汇报提纲

一	研究背景意义
二	主要研究内容
三	下一步工作安排

下一步工作安排

➢ 非凸张量典型相关分析

预处理
Autoencoder

→

模型
Structured Sparsity

↓

算法
Manifold Optimization

←

应用
Multi-view Learning

$$\min_{\mathcal{P},\{\mathbf{U}_p\}} \quad \frac{1}{2}\|\mathcal{M}-\hat{\mathcal{M}}\|_{\mathrm{F}}^2+\sum_p \lambda_p\|\mathbf{U}_p\|_{2,1}$$
$$\text{s.t.} \quad \mathbf{U}_p^\top\mathbf{U}_p=\mathbf{I}, p=1,\ldots,m$$

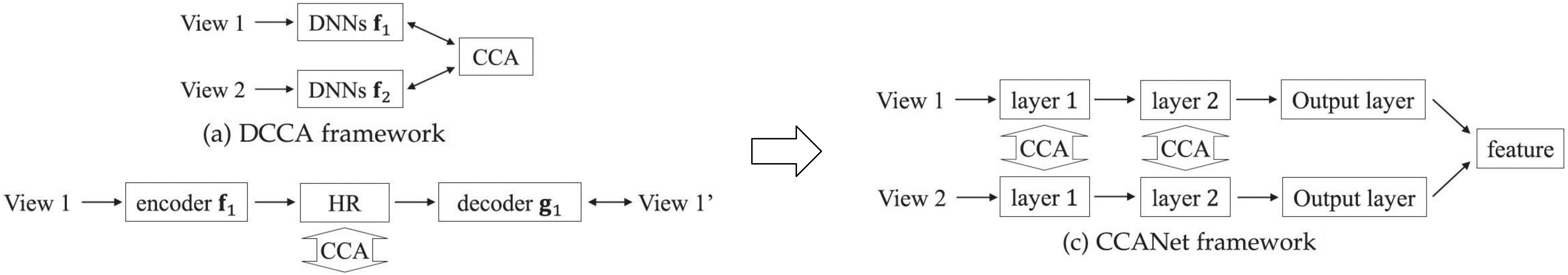
$$\Downarrow$$

$$\min_{\mathcal{P},\{\mathbf{U}_p\}} \quad \frac{1}{2}\|\mathcal{M}-\hat{\mathcal{M}}\|_{\mathrm{F}}^2+\sum_p \lambda_p\phi(\|\mathbf{U}_p^i\|_2)$$
$$\text{s.t.} \quad \mathbf{U}_p^\top\mathbf{U}_p=\mathbf{I}, p=1,\ldots,m$$

- Ng. *Sparse autoencoder,* CS294A Lecture Notes, 72: 1-19, 2011.
- Yang-Pong-Chen. *Alternating direction method of multipliers for a class of nonconvex and nonsmooth problems with applications to background/foreground extraction*, SIAM Journal on Imaging Sciences, 10(1): 74-110, 2017.
- Chen-Ma-So-Zhang. *Proximal gradient method for nonsmooth optimization over the Stiefel manifold*, SIAM Journal on Optimization, 30(1): 210-239, 2020.
- Li-Nie-Bian-Wu-Li. *Sparse PCA via L2,p-norm regularization for unsupervised feature selection*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 45(4): 5322-5328, 2023.
- Zhang-Zheng-Wang-Tang-Zhou-Lin. *Structured sparsity optimization with non-convex surrogates L2,0-norm: A unified algorithmic framework*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023.

下一步工作安排

➢ 深度张量典型相关分析



(a) DCCA framework

(b) DCCAE framework

(c) CCANet framework

- Yang-Liu-Liu, *Tensor canonical correlation analysis networks for multi-view remote sensing scene recognition*, IEEE Transactions on Knowledge and Data Engineering, 34(6): 2948–2961, 2022.
- Shu-Qu-Zhu, *D-GCCA: Decomposition-based generalized canonical correlation analysis for multi-view high-dimensional data*, Journal of Machine Learning Research, 23(169), 1-64, 2022.
- Wong-Wang-Chan-Zeng, *Deep tensor CCA for multi-view learning*, IEEE Transactions on Big Data, 8(6): 1664-1677, 2022.
- Li-Qiu-Zhang, *Robust low transformed multi-rank tensor completion with deep prior regularization for multi-dimensional image recovery*, IEEE Transactions on Big Data, 2023.

谢谢各位专家，

请您批评指正！

Email: xcxiu@shu.edu.cn Homepage：https://xianchaoxiu.github.io/