

多样性诱导的多视图子空间聚类 (DiMSC)

作者: Xiaochun Cao

时间: 2015

汇报人: 潘岩



2. 研究背景与动机

- 多视图数据的特点

- 数据从多个视角/特征提取（如颜色、纹理、光流等）。
- 不同视图可能包含互补信息，但独立使用时存在信息冗余或不足。
- 互补性：不同视图可能包含不同的信息。例如，颜色特征可能对光照变化敏感，而纹理特征可能对光照变化不敏感。如果能够结合这些视图的互补信息，可以提升聚类性能。

- 朴素方法的局限性

- 论文中先提出了一种朴素方法（如NaMSC）独立学习视图表示，NaMSC方法的目标是将多个视图的数据表示整合到一个共同的子空间中，以便进行聚类。每个视图都使用独立的子空间表示来表示数据，但是它们并未通过任何机制来确保不同视图之间的互补性。

$$\min_{Z^{(v)}} \|X^{(v)} - X^{(v)} Z^{(v)}\|_F^2 + \alpha^{(v)} \Omega(Z^{(v)})$$

其中， $\Omega(Z^{(v)})$ 是一个平滑正则项，强制子空间表示满足分组效果。

$$\Omega(Z^{(v)}) = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n w_{ij}^{(v)} \|z_i^{(v)} - z_j^{(v)}\|_2^2 = \text{tr}(Z^{(v)} L^{(v)} Z^{(v)T})$$

- 缺乏对视图间依赖性的显式约束。

3. 相关工作

- 多视图聚类方法
 - 协同正则化谱聚类 (Co-Reg SPC)
 - 协同训练谱聚类 (Co-Train SPC)
 - 多核学习 (MKL)
- 子空间聚类方法
 - 稀疏子空间聚类 (SSC)
 - 低秩表示 (LRR)
 - 平滑表示聚类 (SMR)
- 现有方法的不足
 - 未显式约束视图间互补性 (仅假设独立性)。



4. 方法概述

- **核心思想**

- **多样性正则化**: 通过HSIC衡量视图间依赖性，强制多样性。
- **多视图联合优化**: 交替最小化策略更新各视图表示。

- **目标函数**

$$\mathcal{O} = \underbrace{\sum_v \text{重构误差}}_{\text{保证表示质量}} + \lambda_S \underbrace{\sum_v \text{平滑项}}_{\text{局部一致性}} + \lambda_V \underbrace{\sum_{v \neq w} \text{HSIC}}_{\text{多样性约束}}$$
$$O(Z^{(1)}, \dots, Z^{(V)}) = \sum_{v=1}^V \left(\|X^{(v)} - X^{(v)}Z^{(v)}\|_F^2 \right) + \sum_{v=1}^V \alpha^{(v)} \text{tr}(Z^{(v)} L^{(v)} Z^{(v)^T}) + \sum_{v \neq w} \lambda_V \text{HSIC}(Z^{(v)}, Z^{(w)})$$

其中，各项的具体含义如下：

- i. **重构误差项**:

$$\|X^{(v)} - X^{(v)}Z^{(v)}\|_F^2$$

度量了每个视图的重构误差，确保学习到的子空间表示 $Z^{(v)}$ 能够尽可能地重构原始数据 $X^{(v)}$ 。

- 2. 平滑正则项:

$$\alpha^{(v)} \text{tr}(Z^{(v)} L^{(v)} Z^{(v)^T})$$

它通过图拉普拉斯矩阵 $L^{(v)}$ 强制相似的样本在子空间表示中保持接近。此项的作用是保持数据点之间的结构一致性，确保同一类的数据点在聚类过程中被归到同一子空间。

- iii. 多样性约束项 (HSIC项):

$$\sum_{v \neq w} \lambda_V \text{HSIC}(Z^{(v)}, Z^{(w)})$$

该项是论文中提出的关键部分，利用 Hilbert–Schmidt独立性准则 (HSIC) 来强制不同视图之间的多样性。HSIC项用来度量不同视图的子空间表示之间的依赖性，目标是减少不同视图之间的冗余信息，从而提升互补性。通过最小化HSIC，确保不同视图之间的表示具有足够的多样性，避免了不同视图之间的过度相似性。

- HSIC公式：

$$\text{HSIC} = \frac{1}{(n-1)^2} \text{tr}(\mathbf{K}^{(v)} \mathbf{H} \mathbf{K}^{(w)} \mathbf{H})$$

(Gram矩阵通过 $\mathbf{Z}^{(v)T} \mathbf{Z}^{(v)}$ 计算)

- 优化流程

- 初始化各视图表示（如SMR）。
- 交替优化：固定其他视图，通过Sylvester方程求解当前视图的 $\mathbf{Z}^{(v)}$ 。
- 迭代至收敛（通常<5次迭代）。



5. 实验与结果

- 数据集
 - Yale (人脸, 光照/表情变化)
 - Extended YaleB (大范围光照变化)
 - ORL (姿态/表情变化)
 - Notting-Hill (视频人脸聚类)
- 对比方法
 - Single_best, FeatConcate, Co-Reg SPC, Co-Train SPC, NaMSC等。
- 评价指标
 - NMI, ACC, AR, F-score, Precision, Recall。
- 关键结果
 - DiMSC全面领先: 在NMI、ACC等指标上显著优于基线 (例: Extended YaleB的NMI提升4.1%)。
 - 多样性有效性: HSIC项显著减少冗余 (对比NaMSC的相似性矩阵可视化)。
 - 高效收敛: 5次迭代内收敛。



6. 结论与未来工作

- 结论
 - HSIC显式约束视图多样性，有效提升聚类性能。
 - DiMSC在多视图数据（尤其是光照/姿态变化场景）中表现优异。
- 未来方向
 - 扩展至非线性核（如高斯核）。
 - 探索更高效的多视图联合优化策略。



谢谢观看 请批评指正

