

鲁棒核多视图子空间聚类

程 畅, 唐科威*

辽宁师范大学数学学院, 辽宁 大连

收稿日期: 2026年8月29日; 录用日期: 2026年9月22日; 发布日期: 2026年9月30日

摘 要

多视图子空间聚类旨在利用不同视图间的互补信息和一致性信息提升聚类性能。所提出的多视图子空间聚类方法是这一方向的重要工作, 然而, 该方法基于线性子空间假设, 难以有效处理非线性多视图数据。为了解决该问题, 本文通过核诱导映射将各视图数据投影至特征空间, 使原本非线性的数据在特征空间中呈现线性子空间结构; 同时, 采用 $l_{2,1}$ 范数替代Frobenius范数约束子空间表示误差, 增强模型对噪声和离群点的鲁棒性。在多个真实数据集上的实验结果表明, 所提方法在多个指标均优于对比方法, 验证了核化策略的有效性。

关键词

多视图子空间聚类, 核方法, $l_{2,1}$ 范数

Robust Kernel Multi-View Subspace Clustering

Chang Cheng, Kewei Tang*

School of Mathematics, Liaoning Normal University, Dalian Liaoning

Received: August 29, 2026; accepted: September 22, 2026; published: September 30, 2026

Abstract

Multi-view subspace clustering aims to leverage the complementary and consensus information across different views to improve clustering performance. The proposed multi-view subspace clustering method is an important work in this direction. However, this method is based on the linear subspace assumption, making it difficult to effectively handle nonlinear multi-view data. To address this issue, this paper projects each view data into a feature space via kernel-induced mapping,

*通讯作者。

enabling the originally non-linear data to exhibit a linear subspace structure in the feature space. Meanwhile, the $l_{2,1}$ -norm is adopted to constrain the subspace representation error instead of the Frobenius norm, thereby enhancing the robustness of the model to noise and outliers. Experimental results on multiple real-world datasets demonstrate that the proposed method outperforms the compared methods across multiple metrics, validating the effectiveness of the kernelization strategy.

Keywords

Multi-View Subspace Clustering, Kernel Method, $l_{2,1}$ -Norm

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

子空间聚类是计算机视觉与机器学习领域的重要研究方向,其核心思想是假设数据分布在多个不同的低维子空间中,通过揭示高维数据中隐藏的低维子空间结构,实现数据的有效分割。现有的子空间聚类方法大致可以分为四类:基于迭代的方法、基于统计的方法、基于代数的方法以及基于谱聚类的方法[1]。其中基于谱聚类的方法是当前最主流的一类方法,该方法首先学习数据点间的亲和矩阵,再对其应用谱聚类得到聚类结果。

在基于谱聚类的方法中,稀疏子空间聚类(Sparse Subspace Clustering, SSC)[2]和低秩表示(Low-Rank Representation, LRR)[3]均基于数据的自表达特性,即每个数据点可以由同一子空间中的其他数据点线性表示,通过学习表示矩阵构建亲和矩阵,进而得到聚类结果。然而,这类方法仅使用单一数据源或单一特征来进行聚类,在面对复杂现实数据时,由于信息单一、特征表征能力有限等原因,聚类效果往往不理想。

随着传感器技术以及信息保存的飞速发展,同一对象可以从多个视角或多个特征进行描述刻画,多视图聚类因此成为机器学习领域热门的研究热点之一。多视图子空间聚类旨在充分利用各视图的互补信息与一致性信息,提升聚类性能。Gao 等人提出的多视图子空间聚类(Multi-View Subspace Clustering, MVSC)[4]是这一方向的重要工作,该方法对各视图分别学习子空间表示,并通过一个公共的聚类指示矩阵保证不同视图间聚类结构的一致性,为后续多视图子空间聚类研究奠定了重要基础。

然而, MVSC 方法基于数据点位于多个线性子空间并集上的假设,而真实数据通常呈现非线性或流形结构,线性子空间模型难以准确刻画其内在关系。核方法是解决此类非线性问题的有效工具:通过核诱导映射[5]将数据从原始输入空间映射到特征空间[6]。映射后的数据在特征空间中可能分布于多个线性子空间。基于该思想, Xiao 等人提出鲁棒核低秩表示(Robust Kernel Low-Rank Representation, RKLRR)[7],通过在映射后的特征空间中执行 LRR 并引入 $l_{2,1}$ 范数正则化增强鲁棒性。

受 RKLRR 启发,本文提出一种核化多视图子空间聚类方法,将 MVSC 推广至核空间:各视图在各自核诱导的特征空间中学习子空间表示,并通过公共聚类指示矩阵保证跨视图聚类的一致性。此外,为了放松视图间一致性约束、充分挖掘多视图互补信息的同时,保证模型对噪声的鲁棒性,本文采用 $l_{2,1}$ 范数替代 Frobenius 范数约束表示误差。具体而言,本文的主要贡献如下:

(1) 将 MVSC 方法推广至核空间, 提出鲁棒核多视图子空间聚类模型, 使其能够有效处理非线性多视图数据。

(2) 本文采用 $l_{2,1}$ 范数替代原方法中的 Frobenius 范数约束, 增强模型对噪声的鲁棒性。

(3) 通过多组基准数据集上的实验验证了所提方法的有效性。

2. 相关工作

在回顾相关工作之前, 首先介绍本文使用的主要符号, 矩阵以大写字母如 A 表示, 矩阵的列向量用下标表示(如 a_i 表示矩阵 A 的第 i 列), 行向量用上标表示(如 a^j 表示矩阵 A 的第 j 行), a_{ij} 表示矩阵 A 中第 i 行第 j 列的元素。矩阵 A 的迹记为 $\text{Tr}(A)$, 样本总数、视图总数和聚类数分别用 n, k 和 c 表示, $\mathbf{1}$ 和 $\mathbf{0}$ 分别表示全 1 和 0 的列向量。矩阵 $A \in R^{m \times n}$ 的 Frobenius 范数为 $\|A\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2}$, 矩阵 A 的 $l_{2,1}$ 范数表示为 $\|A\|_{2,1} = \sum_{j=1}^n \|a_j\|_2 = \sum_{j=1}^n \sqrt{\sum_{i=1}^m a_{ij}^2}$ 。此外, I 表示单位矩阵, 矩阵 A 的转置表示为 A^T 。

2.1. 子空间聚类

子空间聚类旨在发现高维数据中潜在的低维子空间结构, 其核心假设是数据样本分布在多个低维子空间的并集上, 每个子空间对应一个聚类簇。相比于传统方法, 子空间聚类能够有效缓解维数灾难问题, 并更准确地识别不同簇所对应的判别性子空间。

对于给定的数据矩阵 $X = [x_1, x_2, \dots, x_n] \in R^{d \times n}$, 子空间聚类基于数据的自表达特性[1], 即每个样本可由同一子空间内的其他样本线性表示:

$$X = XZ + E, \quad (1)$$

其中 $Z = [z_1, z_2, \dots, z_n] \in R^{n \times n}$ 为子空间表示矩阵, $E \in R^{d \times n}$ 为噪声或误差矩阵。子空间聚类通过求解以下优化问题获得子空间结构:

$$\min_Z \|X - XZ\|_F^2 \quad \text{s.t.} \quad Z(i, i) = 0, Z^T \mathbf{1} = \mathbf{1}, \quad (2)$$

约束条件 $Z(i, i) = 0$ 为了避免每个数据样本在重构过程中被自身表示而产生平凡解, 约束条件 $Z^T \mathbf{1} = \mathbf{1}$ 表示数据点位于仿射子空间的并集内, 而非线性子空间。获得子空间表示矩阵 Z 后, 通过对称化处理:

$$W = \frac{|Z| + |Z^T|}{2}, \quad (3)$$

构建亲和矩阵, 进而应用谱聚类得到聚类结果。

2.2. 回顾 MVSC 方法

多视图子空间聚类方法旨在整合来自多个视图的互补信息, 学习统一的子空间表示或共识相似度图。例如, Gao 等人提出的 MVSC 方法是这一方向的重要工作, 将每个视图的子空间表示学习与共识嵌入对齐整合到一个联合优化框架中, 其模型可表述如下:

$$\begin{aligned} \min_{Z^{(v)}, E^{(v)}, F} \sum_v \|X^{(v)} - X^{(v)}Z^{(v)} - E^{(v)}\|_F^2 + \lambda_1 \sum_v \text{Tr}(F^T (D^{(v)} - W^{(v)}) F) + \lambda_2 \sum_v \|E^{(v)}\|_1 \\ \text{s.t.} \quad Z^{(v)T} \mathbf{1} = \mathbf{1}, Z^{(v)}(i, i) = 0, F^T F = I, \end{aligned} \quad (4)$$

其中 $X^{(v)}$ 表示第 v 个视图的特征矩阵, $Z^{(v)}$ 表示第 v 个视图的子空间表示矩阵, $E^{(v)}$ 表示第 v 个视图的误

差矩阵, $W^{(v)}$ 表示第 v 个视图的亲矩阵, $D^{(v)}$ 是每个视图的度矩阵, 为对角矩阵, 其对角线元素为 $d_{ij} = \sum_j w_{ij}$, F 表示统一的聚类指示矩阵, l_1 范数用来保证异常值矩阵的稀疏性, 参数 λ_1 和 λ_2 用于调节相应项的权重。

MVSC 的核心思想在于: 各视图的数据分布不同, 直接学习一个统一的子空间表示 Z 并不合理, 各视图的聚类划分应当一致, 因此可以通过一个公共的聚类指示矩阵 F 来约束所有视图的聚类结果保持一致, 将子空间表示学习与聚类目标联合优化。然而, MVSC 方法建立在数据位于多个线性子空间并集上的假设之上。当数据呈现非线性流形结构时, 线性子空间表示难以准确刻画数据点间的内在关系, 导致聚类性能下降。因此, 如何将 MVSC 方法扩展至非线性场景成为了一个自然的研究方向。

3. 提出方法

经过上述回顾我们可以看到, 尽管 MVSC 方法取得了一定的进展, 但仍有一些不足。MVSC 方法的核心假设是数据点分布在多个线性子空间的并集上, 然而真实世界中的数据(如图像、文本、传感器信号)往往呈现复杂的非线性流形结构。在这种情况下, 线性子空间表示难以准确刻画数据点间的内在关系, 从而导致聚类性能下降。

针对上述问题, 本文受 RKLRR 的启发引入核方法, 将 MVSC 推广至核空间, 提出鲁棒核多视图子空间聚类方法(RKMVSC)。各视图在各自核诱导的特征空间中学习子空间表示, 并通过公共聚类指示矩阵保证跨视图聚类一致性。同时, 采用 $l_{2,1}$ 范数替代 Frobenius 范数约束表示误差, 以增强模型对噪声和离群点的鲁棒性。

给定包含 k 个视图的多视图数据 $\{X^{(v)}\}_{v=1}^k$, 其中 $X^{(v)} \in R^{d_v \times n}$ 表示第 v 个视图的特征矩阵, n 为样本总数, d_v 为第 v 个视图的特征维度。对于每个视图, 我们通过核映射 Φ_v 将其映射到特征空间, 得到 $\Phi_v(X^{(v)}) = [\phi_v(x_1^{(v)}), \dots, \phi_v(x_n^{(v)})]$ 。本文的模型可表述如下:

$$\begin{aligned} \min_{Z^{(v)}, F} \sum_v \lambda_1 \left\| \Phi_v(X^{(v)}) - \Phi_v(X^{(v)}) Z^{(v)} \right\|_{2,1} + \lambda_2 \sum_v \text{Tr}(F^T (D^{(v)} - W^{(v)}) F) \\ \text{s.t. } Z^{(v)T} \mathbf{1} = \mathbf{1}, Z^{(v)}(i, i) = 0, F^T F = I. \end{aligned} \quad (5)$$

其中 $Z^{(v)} \in R^{n \times n}$ 为第 v 个视图在核空间中的子空间表示矩阵, $W^{(v)}$ 为第 v 个视图的亲矩阵, $D^{(v)}$ 为度矩阵, $F \in R^{n \times c}$ 为所有视图共享的聚类指示矩阵, λ_1, λ_2 为正则化参数。

与 MVSC 方法相比, 本文方法的改进具体体现在以下两个方面:

1) MVSC 方法依赖于数据服从线性子空间分布的假设, 而本文通过核诱导映射, 使得我们的方法能够有效处理具有非线性结构的多视图数据, 从而克服了线性模型对非线性数据建模能力不足的问题。

2) 本文采用 $l_{2,1}$ 范数来约束子空间表示误差, 使得模型对噪声和离群点具有更强的鲁棒性。

4. 模型求解

在本节中, 本文采用交替方向乘子法(Alternating Direction Method of Multipliers, ADMM) [8]对问题(5)进行求解。为便于核化求解, 本文引入辅助变量 $P^{(v)} = I - Z^{(v)}$, 则第 v 个视图的重构误差项可重写为:

$$\left\| \Phi_v(X^{(v)}) - \Phi_v(X^{(v)}) Z^{(v)} \right\|_{2,1} = \left\| \Phi_v(X^{(v)}) P^{(v)} \right\|_{2,1} = \sum_{i=1}^n \left(p_i^{(v)T} K^{(v)} p_i^{(v)} \right)^{\frac{1}{2}}, \quad (6)$$

其中 $p_i^{(v)}$ 为 $P^{(v)}$ 的第 i 列, $K^{(v)} = \Phi_v(X^{(v)})^T \Phi_v(X^{(v)})$ 为第 v 个视图的核矩阵。本文中定义了一个关于 $P \in R^{n \times n}$ 的函数 $g(P)$, 即

$$g(P) \triangleq \sum_{i=1}^n \sqrt{p_i^T K p_i}, \quad (7)$$

则问题(5)转化为如下仅依赖核矩阵的形式:

$$\begin{aligned} \min_{Z^{(v)}, P^{(v)}, F} \quad & \lambda_1 \sum_{v=1}^k g(P^{(v)}) + \lambda_2 \sum_{v=1}^k \text{Tr}(F^T L^{(v)} F) \\ \text{s.t.} \quad & Z^{(v)T} \mathbf{1} = \mathbf{1}, \quad Z^{(v)}(i, i) = 0, \quad P^{(v)} = I - Z^{(v)}, \quad F^T F = I, \end{aligned} \quad (8)$$

其中, $L^{(v)} = D^{(v)} - W^{(v)}$ 为拉普拉斯矩阵。问题(8)的增广拉格朗日函数为:

$$\zeta = \lambda_1 \sum_{v=1}^k g(P^{(v)}) + \lambda_2 \sum_{v=1}^k \text{Tr}(F^T L^{(v)} F) + \frac{\mu}{2} \sum_{v=1}^k \left\| I - Z^{(v)} - P^{(v)} + \frac{Y^{(v)}}{\mu} \right\|_F^2, \quad (9)$$

其中 $Y^{(v)}$ 为拉格朗日乘子, μ 为惩罚参数。各变量交替更新如下。

更新 $P^{(v)}$

通过固定公式(9)里的变量 $Z^{(v)}$ 、 F 和 $Y^{(v)}$, 得到关于 $P^{(v)}$ 的子问题:

$$\min_{P^{(v)}} \lambda_1 g(P^{(v)}) + \frac{\mu}{2} \left\| I - Z^{(v)} - P^{(v)} + \frac{Y^{(v)}}{\mu} \right\|_F^2, \quad (10)$$

假设 $C^{(v)} = I - Z^{(v)} + \frac{Y^{(v)}}{\mu}$, 则公式(10)中的目标函数可简化为:

$$\min_{P^{(v)}} g(P^{(v)}) + \frac{\mu}{2\lambda_1} \left\| P^{(v)} - C^{(v)} \right\|_F^2, \quad (11)$$

该问题可以分解为如下 n 个独立的子问题:

$$\min_{p_i^{(v)}} \sqrt{p_i^{(v)T} K^{(v)} p_i^{(v)}} + \frac{\mu}{2\lambda_1} \left\| p_i^{(v)T} - c_i^{(v)T} \right\|_2^2. \quad (12)$$

设 $K^{(v)}$ 的奇异值分解为 $K^{(v)} = V^{(v)} (\Sigma^{(v)})^2 V^{(v)T}$, 其中 $\Sigma^{(v)} = \text{diag}([\sigma_1^{(v)}, \dots, \sigma_r^{(v)}, 0, \dots, 0])$, r 为 $K^{(v)}$ 的秩。

令 $\tau = \frac{\mu}{2\lambda_1}$, 根据文献[6]中的定理 2, 问题(12)的闭式解为:

$$p_i^{(v)} = \begin{cases} c_i^{(v)} - V_K^{(v)} \text{diag} \left(\left\{ \frac{(\sigma_j^{(v)})^2}{\tau\alpha + (\sigma_j^{(v)})^2} \right\}_{j=1}^r \right) V_K^{(v)T} c_i^{(v)}, & \text{if } \left\| (\Sigma_u^{(v)})^{-1} V_K^{(v)T} c_i^{(v)} \right\| > \frac{1}{\tau}, \\ c_i^{(v)} - V_K^{(v)} V_K^{(v)T} c_i^{(v)}, & \text{otherwise} \end{cases} \quad (13)$$

其中 $V_K^{(v)} \in R^{n \times r}$ 由 $V^{(v)}$ 的前 r 列组成, $\Sigma_u^{(v)} = \text{diag}(\sigma_1^{(v)}, \dots, \sigma_r^{(v)})$, 标量 $\alpha > 0$ 满足:

$$\left(V_K^{(v)T} c_i^{(v)} \right)^T \text{diag} \left(\left\{ \frac{(\sigma_j^{(v)})^2}{(\tau\alpha + (\sigma_j^{(v)})^2)^2} \right\}_{j=1}^r \right) \left(V_K^{(v)T} c_i^{(v)} \right) = \frac{1}{\tau^2}, \quad (14)$$

公式(14)可通过二分法高效求解。

更新 $Z^{(v)}$

固定 $P^{(v)}, F$ 和 $Y^{(v)}$, 关于求解 $Z^{(v)}$ 的子问题为:

$$\begin{aligned} \min_{Z^{(v)}} & \frac{\mu}{2} \left\| Z^{(v)} - \left(I - P^{(v)} + \frac{Y^{(v)}}{\mu} \right) \right\|_F^2 + \lambda_2 \text{Tr}(F^T L^{(v)} F) \\ \text{s.t. } & Z^{(v)T} \mathbf{1} = \mathbf{1}, Z^{(v)}(i, i) = 0, \end{aligned} \quad (15)$$

假设 $A^{(v)} = I - P^{(v)} + \frac{Y^{(v)}}{\mu}$, 则公式(15)中的目标函数可以简化为:

$$\begin{aligned} \min_{Z^{(v)}} & \frac{\mu}{2} \left\| Z^{(v)} - A^{(v)} \right\|_F^2 + \lambda_2 \text{Tr}(F^T L^{(v)} F) \\ \text{s.t. } & Z^{(v)T} \mathbf{1} = \mathbf{1}, Z^{(v)}(i, i) = 0, \end{aligned} \quad (16)$$

公式(16)可以直接利用文献[4]中的定理 1 求解, 我们得到该优化问题的解为:

$$\begin{aligned} z_{ij}^{(v)} &= \text{sign}(a_{ij}^{(v)}) \cdot \max \left(|a_{ij}^{(v)}| - \frac{\lambda_2 v_{ij}^{(v)}}{2\mu}, 0 \right) \quad i \neq j \\ &= \begin{cases} a_{ij}^{(v)} - \frac{\lambda_2 v_{ij}^{(v)}}{2\mu}, & \text{if } a_{ij}^{(v)} > \frac{\lambda_2 v_{ij}^{(v)}}{2\mu} \\ a_{ij}^{(v)} + \frac{\lambda_2 v_{ij}^{(v)}}{2\mu}, & \text{if } a_{ij}^{(v)} < -\frac{\lambda_2 v_{ij}^{(v)}}{2\mu} \\ 0, & \text{otherwise} \end{cases} \end{aligned} \quad (17)$$

其中 $v_{ij}^{(v)} = \left\| f^{i(v)} - f^{j(v)} \right\|_2^2$ 。

更新 F

通过固定除了 F 以外的相关的变量, 得到目标函数:

$$\min_F \sum_{v=1}^k \text{Tr}(F^T L^{(v)} F) = \min_F \text{Tr}(F^T M F) \quad \text{s.t. } F^T F = I, \quad (18)$$

其中 $M = \sum_{v=1}^k L^{(v)}$ 。该问题的最优解为 M 的 c 个最小特征值对应的特征向量。

更新拉格朗日乘子与惩罚参数

$$Y^{(v)} = Y^{(v)} + \mu(I - Z^{(v)} - P^{(v)}) \quad (19)$$

$$\mu = \min(\rho\mu, \mu_0) \quad (20)$$

其中 ρ 和 μ_0 是常数。

基于上述步骤, 本文所提 RKMVSC 算法的完整流程如算法 1 所示。

算法 1: RKMVSC 算法

输入: 多视图核矩阵 $\{K^{(v)}\}_{v=1}^k$, 参数 λ_1, λ_2 , 聚类数 c

输出: 聚类指示矩阵 F

初始化: $Z^{(v)} = 0, P^{(v)} = I, Y^{(v)} = 0, \mu = 0.5, \mu_0 = 10^6, \rho = 0.1, \varepsilon = 10^{-4}$

续表

步骤 1: 通过式(14)更新 $P^{(v)}$, 通过式(17)更新 $Z^{(v)}$, 通过式(18)更新 F

步骤 2: 通过式(19)更新拉格朗日乘子

步骤 3: 通过式(20)更新惩罚参数

步骤 4: 检查收敛条件: $\|I - Z^{(v)} - P^{(v)}\|_{\infty} \leq \varepsilon$, 不满足收敛条件, 重复步骤 1~步骤 3, 若满足收敛条件, 结束

5. 实验

在本节中, 我们在 4 个广泛使用的基准数据集上对我们的算法进行了评估, 并进行实验以证明所提出方法的有效性和优越性。

5.1. 实验设置

在实验中, 为了评估我们方法的性能, 我们将其与单视图方法以及多视图聚类方法进行了比较, 具体信息如下:

LRR-Single: 在每个视图上单独运行 LRR [3] 获取低秩子空间表示, 然后对这些表示分别进行谱聚类得到聚类结果。

SSC-Single: 在每个视图上单独运行 SSC [2] 获取子空间表示, 然后在该子空间表示上运行谱聚类方法。

MLRR-Add 和 MSSC-Add: MLRR-Add 方法首先分别在每个视图特征上运行 LRR [3] 以获得每个视图的低秩子空间表示, 然后将这些低秩子空间表示相加, 最后对其执行谱聚类。MSSC-Add 方法与 MLRR-Add 类似, 但对每个视图特征运行的是 SSC [2] 方法。

MLRR-Con 和 MSSC-Con: 与 MLRR-add 方法不同的是, MLRR-Con 是将所有视图特征进行拼接, 在拼接后的特征上运行 LRR [3], 最终执行谱聚类。MSSC-Con 方法执行的步骤相同, 但最终对拼接后的特征执行 SSC [2] 以获得子空间表示。

Centroid-based Co-regularized Multi-view Spectral Clustering (CC-MS): [9] 该方法通过引入一个共识嵌入矩阵, 并利用基于质心的协同正则化项迫使各视图的谱嵌入向该共识对齐, 使得不同视图的聚类结果彼此保持一致。

Pairwise Co-regularized Multi-view Spectral Clustering (PC-MS): [9] 该方法与 CC-MS 类似, 区别在于 PC-MS 是对每两个视图施加一个协同正则化项, 使得不同视图之间彼此保持一致。

MVSC: [4] 该方法通过一个统一的聚类指示矩阵来引导各视图的子空间表示学习, 从而保证不同视图之间的聚类一致性。

在实验中, 所有对比方法的参数均通过网格搜索从 $\{10^{-8}, 10^{-7}, \dots, 10^3, 10^4\}$ 范围内选取, 而本文提出的方法由于参数过小导致算法中二分法迭代的初始区间失去有效解, 因此我们方法的参数通过网格搜索从 $\{10^{-5}, 10^{-4}, \dots, 10^3, 10^4\}$ 范围内选取。并且由于所有方法均需执行 K 均值聚类步骤, 为了确保公平, 我们固定特定参数, 通过运行十次取平均的策略报告了每种方法获得的最佳聚类结果。表 1 汇总了用于评估的 4 个真实世界数据库, 其中 $\#d_i$ 表示第 i 个视图的特征维度。

5.1.1. 数据集介绍

在实验中, 为评估我们提出方法的有效性, 我们选取了四个真实世界的多视图数据集, 具体描述如下:

MSRC-v1 数据集: [10] 数据库是由微软剑桥研究院提供的图像数据库, 包含涵盖 9 个物体类别的 240 幅图像。在实验中, 我们采用了一个包含 7 个类别和 210 幅图像的多视图子集, 该子集包含 6 种特征视

图, 其特征维度分别为 1302、48、512、100、256 和 210。

Caltech101-7 数据集: [11]该数据集是广泛使用的 Caltech101 图像数据集中的子集。在实验中, 我们采用了一个包含 7 个代表性类别, 包含 441 个样本的多视图子集, 该子集包含 6 中特征视图, 其特征维度分别为 1302、48、512、100、256 和 441。

ORL 数据集: [12]该数据集是由 40 个人的 400 张人脸图像组成的。在实验中, 我们将所有图像均调整为 32×32 像素尺寸, 并提取灰度、LBP 和 Gabor 的特征, 形成具有三个视图的多视图 ORL 数据集, 其特征维度分别为 1024、36 和 1024。

3Sources 数据集: [13]该数据集是一个多视图文本数据集, 包含来自三大主流在线新闻来源(BBC、Reuters and Guardian)的文本数据。实验中, 我们从这三个来源中联合选取 169 篇文章构建该多视角数据集, 每个媒体来源作为一个独立的视图, 以词频特征表示, 维度分别为 3560、3631 和 3068。

Table 1. Dataset information
表 1. 数据集信息

数据集	$d1$	$d2$	$d3$	$d4$	$d5$	$d6$	样本数	类别数
MSRC-v1	1032	48	512	100	256	210	210	7
Caltech101-7	1302	48	512	100	256	441	441	6
ORL	1024	36	1024	-	-	-	400	40
3Sources	3560	3631	3068	-	-	-	169	6

5.1.2. 评价指标

在实验中, 我们采用六项指标评估性能: 包括**准确率(ACC)**、**归一化互信息(NMI)**、**调整兰德指数(ARI)**、**F-score**、**Precision** 和 **Recall**。具体而言, ACC、NMI、F-score、Precision 和 Recall 的取值范围为 $[0, 1]$, 而 ARI 的取值范围为 $[-1, 1]$, 上述指标数值越高代表性能越好。

5.2. 实验结果

表 2~5 主要展示了在 4 个多视图数据集上的运行结果, 结果表示形式为各个方法最优结果的均值 $\pm$ 标准差。其中, 最优结果以粗体表示, 次优结果以下划线表示。

Table 2. Experimental results on the MSRC-v1 dataset
表 2. 在 MSRC-v1 数据库上的实验结果

数据集	ACC	NMI	ARI	F-score	Precision	Recall
LRR-Single1	0.654 $\pm$ 0.044	0.530 $\pm$ 0.026	0.442 $\pm$ 0.038	0.521 $\pm$ 0.033	0.511 $\pm$ 0.034	0.532 $\pm$ 0.032
LRR-Single2	0.665 $\pm$ 0.033	0.548 $\pm$ 0.011	0.452 $\pm$ 0.017	0.529 $\pm$ 0.014	0.519 $\pm$ 0.019	0.541 $\pm$ 0.009
LRR-Single3	0.668 $\pm$ 0.025	0.555 $\pm$ 0.011	0.466 $\pm$ 0.020	0.543 $\pm$ 0.017	0.525 $\pm$ 0.018	0.563 $\pm$ 0.017
LRR-Single4	0.660 $\pm$ 0.021	0.564 $\pm$ 0.011	0.452 $\pm$ 0.014	0.529 $\pm$ 0.012	0.522 $\pm$ 0.012	0.536 $\pm$ 0.012
LRR-Single5	0.224 $\pm$ 0.014	0.046 $\pm$ 0.006	0.001 $\pm$ 0.003	0.156 $\pm$ 0.008	0.140 $\pm$ 0.003	0.178 $\pm$ 0.019
LRR-Single6	0.444 $\pm$ 0.024	0.284 $\pm$ 0.012	0.189 $\pm$ 0.008	0.305 $\pm$ 0.007	0.295 $\pm$ 0.008	0.316 $\pm$ 0.009
SSC-Single1	0.757 $\pm$ 0.027	0.642 $\pm$ 0.019	0.576 $\pm$ 0.025	0.637 $\pm$ 0.021	0.614 $\pm$ 0.025	0.662 $\pm$ 0.019

续表

SSC-Single2	0.389 ± 0.014	0.268 ± 0.025	0.147 ± 0.021	0.270 ± 0.017	0.260 ± 0.017	0.281 ± 0.018
SSC-Single3	0.653 ± 0.047	0.565 ± 0.038	0.468 ± 0.053	0.544 ± 0.045	0.526 ± 0.043	0.564 ± 0.048
SSC-Single4	0.570 ± 0.033	0.490 ± 0.024	0.369 ± 0.031	0.459 ± 0.027	0.448 ± 0.026	0.470 ± 0.028
SSC-Single5	0.663 ± 0.017	0.533 ± 0.019	0.444 ± 0.030	0.524 ± 0.024	0.502 ± 0.031	0.549 ± 0.018
SSC-Single6	0.404 ± 0.020	0.233 ± 0.019	0.146 ± 0.020	0.266 ± 0.013	0.263 ± 0.013	0.269 ± 0.013
MLRR-Add	0.907 ± 0.004	0.833 ± 0.006	0.806 ± 0.007	0.833 ± 0.006	0.828 ± 0.006	0.838 ± 0.007
MLRR-Con	0.245 ± 0.013	0.057 ± 0.007	0.007 ± 0.006	0.162 ± 0.006	0.144 ± 0.005	0.185 ± 0.011
MSSC-Add	0.811 ± 0.031	0.726 ± 0.020	0.658 ± 0.030	0.706 ± 0.026	0.698 ± 0.027	0.713 ± 0.026
MSSC-Con	0.605 ± 0.039	0.512 ± 0.017	0.392 ± 0.022	0.481 ± 0.019	0.457 ± 0.020	0.509 ± 0.020
CC-MSC	0.887 ± 0.012	0.795 ± 0.019	0.758 ± 0.023	0.792 ± 0.019	0.785 ± 0.021	0.799 ± 0.018
PC-MSC	0.891 ± 0.007	0.813 ± 0.006	0.762 ± 0.012	0.795 ± 0.011	0.787 ± 0.012	0.803 ± 0.009
MVSC	0.513 ± 0.070	0.392 ± 0.073	0.259 ± 0.086	0.383 ± 0.061	0.328 ± 0.081	0.472 ± 0.024
RKMVSC	0.929 ± 0.000	0.868 ± 0.000	0.835 ± 0.000	0.858 ± 0.000	0.849 ± 0.000	0.869 ± 0.000

Table 3. Experimental results on the caltech101-7 dataset

表 3. 在 Caltech101-7 数据库上的实验结果

数据集	ACC	NMI	ARI	F-score	Precision	Recall
LRR-Single1	0.637 ± 0.025	0.495 ± 0.016	0.469 ± 0.031	0.560 ± 0.028	0.539 ± 0.025	0.584 ± 0.050
LRR-Single2	0.473 ± 0.031	0.324 ± 0.024	0.209 ± 0.023	0.344 ± 0.021	0.331 ± 0.017	0.358 ± 0.028
LRR-Single3	0.700 ± 0.035	0.522 ± 0.017	0.497 ± 0.031	0.585 ± 0.025	0.550 ± 0.028	0.624 ± 0.026
LRR-Single4	0.674 ± 0.020	0.518 ± 0.007	0.506 ± 0.023	0.586 ± 0.022	0.588 ± 0.009	0.586 ± 0.036
LRR-Single5	0.275 ± 0.010	0.088 ± 0.006	0.047 ± 0.007	0.200 ± 0.007	0.205 ± 0.006	0.195 ± 0.012
LRR-Single6	0.695 ± 0.016	0.609 ± 0.008	0.577 ± 0.007	0.644 ± 0.006	0.665 ± 0.011	0.625 ± 0.005
SSC-Single1	0.648 ± 0.038	0.555 ± 0.018	0.499 ± 0.028	0.587 ± 0.021	0.555 ± 0.033	0.622 ± 0.015
SSC-Single2	0.458 ± 0.010	0.298 ± 0.013	0.184 ± 0.006	0.315 ± 0.006	0.322 ± 0.008	0.309 ± 0.013
SSC-Single3	0.659 ± 0.031	0.544 ± 0.019	0.475 ± 0.033	0.565 ± 0.026	0.543 ± 0.036	0.590 ± 0.017
SSC-Single4	0.668 ± 0.026	0.561 ± 0.026	0.504 ± 0.032	0.591 ± 0.026	0.553 ± 0.027	0.635 ± 0.031
SSC-Single5	0.556 ± 0.023	0.316 ± 0.004	0.295 ± 0.011	0.427 ± 0.014	0.375 ± 0.008	0.498 ± 0.040
SSC-Single6	0.687 ± 0.041	0.634 ± 0.013	0.599 ± 0.044	0.666 ± 0.037	0.659 ± 0.039	0.674 ± 0.047
MLRR-Add	0.818 ± 0.037	0.777 ± 0.014	0.766 ± 0.036	0.804 ± 0.031	0.815 ± 0.026	0.794 ± 0.036
MLRR-Con	0.285 ± 0.026	0.071 ± 0.028	0.044 ± 0.024	0.299 ± 0.010	0.184 ± 0.012	0.806 ± 0.094
MSSC-Add	0.784 ± 0.028	0.683 ± 0.010	0.670 ± 0.011	0.724 ± 0.009	0.722 ± 0.012	0.727 ± 0.014

续表

MSSC-Con	0.556 ± 0.032	0.324 ± 0.006	0.301 ± 0.016	0.431 ± 0.019	0.381 ± 0.008	0.500 ± 0.051
CC-MSC	0.744 ± 0.054	0.775 ± 0.023	0.725 ± 0.062	0.770 ± 0.052	0.779 ± 0.054	0.761 ± 0.052
PC-MSC	0.764 ± 0.044	0.775 ± 0.023	0.738 ± 0.065	0.780 ± 0.055	0.788 ± 0.048	0.773 ± 0.061
MVSC	0.521 ± 0.047	0.326 ± 0.036	0.178 ± 0.027	0.370 ± 0.016	0.261 ± 0.017	0.643 ± 0.050
RKMVSC	0.816 ± 0.001	0.724 ± 0.001	0.674 ± 0.001	0.729 ± 0.001	0.713 ± 0.001	0.746 ± 0.001

Table 4. Experimental results on the ORL dataset
表 4. 在 ORL 数据库上的实验结果

数据集	ACC	NMI	ARI	F-score	Precision	Recall
LRR-Single1	0.674 ± 0.033	0.816 ± 0.015	0.557 ± 0.035	0.568 ± 0.034	0.533 ± 0.038	0.607 ± 0.029
LRR-Single2	0.315 ± 0.020	0.537 ± 0.012	0.120 ± 0.014	0.141 ± 0.014	0.135 ± 0.014	0.148 ± 0.014
LRR-Single3	0.708 ± 0.013	0.852 ± 0.008	0.612 ± 0.019	0.621 ± 0.019	0.566 ± 0.019	0.688 ± 0.023
SSC-Single1	0.707 ± 0.033	0.846 ± 0.012	0.605 ± 0.029	0.614 ± 0.028	0.563 ± 0.034	0.676 ± 0.023
SSC-Single2	0.377 ± 0.013	0.608 ± 0.007	0.195 ± 0.008	0.214 ± 0.008	0.202 ± 0.008	0.227 ± 0.008
SSC-Single3	0.717 ± 0.023	0.858 ± 0.011	0.624 ± 0.028	0.634 ± 0.027	0.574 ± 0.033	0.708 ± 0.021
MLRR-Add	0.717 ± 0.031	0.850 ± 0.019	0.617 ± 0.043	0.626 ± 0.042	0.583 ± 0.045	0.677 ± 0.038
MLRR-Con	0.709 ± 0.028	0.842 ± 0.012	0.605 ± 0.027	0.614 ± 0.026	0.569 ± 0.024	0.668 ± 0.030
MSSC-Add	0.690 ± 0.035	0.835 ± 0.014	0.587 ± 0.032	0.597 ± 0.031	0.558 ± 0.036	0.642 ± 0.027
MSSC-Con	0.715 ± 0.021	0.849 ± 0.012	0.606 ± 0.029	0.616 ± 0.028	0.552 ± 0.032	0.697 ± 0.025
CC-MSC	0.654 ± 0.024	0.804 ± 0.013	0.532 ± 0.028	0.544 ± 0.027	0.507 ± 0.031	0.586 ± 0.024
PC-MSC	0.662 ± 0.021	0.809 ± 0.008	0.539 ± 0.021	0.550 ± 0.021	0.508 ± 0.024	0.601 ± 0.018
MVSC	0.431 ± 0.033	0.550 ± 0.038	0.110 ± 0.038	0.145 ± 0.035	0.085 ± 0.025	0.559 ± 0.028
RKMVSC	0.780 ± 0.014	0.880 ± 0.004	0.696 ± 0.012	0.703 ± 0.012	0.661 ± 0.017	0.752 ± 0.008

Table 5. Experimental results on the 3Sources dataset
表 5. 在 3Sources 数据库上的实验结果

数据集	ACC	NMI	ARI	F-score	Precision	Recall
LRR-Single1	0.625 ± 0.070	0.566 ± 0.046	0.485 ± 0.081	0.592 ± 0.066	0.661 ± 0.067	0.536 ± 0.066
LRR-Single2	0.770 ± 0.049	0.619 ± 0.030	0.616 ± 0.060	0.699 ± 0.051	0.744 ± 0.035	0.661 ± 0.068
LRR-Single3	0.704 ± 0.043	0.640 ± 0.026	0.562 ± 0.034	0.655 ± 0.028	0.720 ± 0.037	0.602 ± 0.040
SSC-Single1	0.709 ± 0.015	0.569 ± 0.021	0.512 ± 0.017	0.628 ± 0.014	0.618 ± 0.033	0.642 ± 0.048
SSC-Single2	0.693 ± 0.062	0.524 ± 0.034	0.510 ± 0.059	0.614 ± 0.047	0.668 ± 0.048	0.570 ± 0.050
SSC-Single3	0.745 ± 0.017	0.627 ± 0.014	0.599 ± 0.014	0.687 ± 0.011	0.723 ± 0.015	0.655 ± 0.016

续表

MLRR-Add	0.782 ± 0.036	0.676 ± 0.025	0.653 ± 0.046	0.734 ± 0.036	0.731 ± 0.035	0.738 ± 0.038
MLRR-Con	0.746 ± 0.050	0.640 ± 0.033	0.625 ± 0.062	0.711 ± 0.054	0.715 ± 0.020	0.712 ± 0.096
MSSC-Add	0.766 ± 0.030	0.667 ± 0.030	0.593 ± 0.039	0.687 ± 0.034	0.688 ± 0.027	0.689 ± 0.060
MSSC-Con	0.758 ± 0.022	0.640 ± 0.027	0.598 ± 0.042	0.684 ± 0.035	0.747 ± 0.026	0.630 ± 0.043
CC-MSc	0.504 ± 0.022	0.442 ± 0.016	0.276 ± 0.038	0.448 ± 0.024	0.441 ± 0.039	0.458 ± 0.025
PC-MSc	0.485 ± 0.030	0.411 ± 0.028	0.260 ± 0.021	0.422 ± 0.017	0.449 ± 0.021	0.398 ± 0.023
MVSC	0.614 ± 0.085	0.365 ± 0.100	0.314 ± 0.110	0.527 ± 0.056	0.404 ± 0.070	0.785 ± 0.068
RKMVSC	0.792 ± 0.002	0.690 ± 0.007	0.665 ± 0.013	0.738 ± 0.009	0.786 ± 0.016	0.697 ± 0.003

从整体来看, 本文方法在绝大多数数据集和评价指标上均取得了最优或次优的聚类性能。与单视图方法(LRR-single、SSC-Single)及其简单融合策略(MLRR-Add、MSSC-Add、MLRR-Con、MSSC-Con)相比, 本文方法在 MSRC-v1 和 ORL 数据集上取得了全面提升。以 MSRC-v1 数据集为例, 本文方法在各个指标上分别超出最优单视图方法 MLRR-Add 约 2.2%、3.5%、2.9%、2.5%、2.1%和 3.1%。在 ORL 数据集上, 本文方法相比最优单视图方法 SSC-Single3, 六项指标分别提升 6.3%、2.2%、7.2%、6.9%、7.8%和 4.4%。上述结果验证了多视图融合策略的有效性, 表明不同视图间的互补信息能够有效提高聚类性能。

除此之外, 对比 CC-MSc, PC-MSc 以及 MVSC 等传统多视图子空间聚类方法, 本文方法有明显优势。MVSC 方法采用线性假设, 难以捕捉数据的非线性结构, 而由于各类数据普遍存在的非线性分布的问题, 导致其聚类性能较差。而 CC-MSc 与 PC-MSc 虽然通过视图协同正则实现分区对齐, 但本质上仍受限于线性建模。以 3Sources 数据集为例, 我们所提的方法的 ACC 较 CC-MSc、PC-MSc 以及 MVSC 分别提升 28.8%、30.7%、17.8%。实验结果表明核化策略有效弥补了 MVSC 等线性模型的缺陷, 能够更好地挖掘多视图非线性子空间结构。

在表 3 中, 虽然 RKMVSC 在 Caltech101-7 数据集上仅在 ACC 上取得次优结果, 但仍处于前列。在 Caltech101-7 数据集上 RKMVSC 的 ACC (0.816)略低于 MLRR-Add (0.818), 这一微小差距可能源于本文对所有视图统一采用高斯核, 以保证比较的公平性与实现的简洁性, 该策略未能针对各视图的数据分布自适应地选取核函数与参数, 在特定数据分布下的适应性仍有提升空间。尽管如此, RKMVSC 的标准差 (0.001)远小于 MLRR-Add (0.037), 表明本文方法在聚类稳定性上具有显著优势, 结果波动更小, 聚类结果在不同初始化下更为稳定。

综合四个数据集的实验结果, 本文提出的核化多视图子空间聚类方法在保持高稳定性的同时, 能够有效利用多视图的互补信息, 并在非线性数据上取得优于线性模型的聚类表现。

6. 结论

本文针对 MVSC 方法在线性假设下处理非线性多视图数据的局限性, 提出了一种鲁棒核多视图子空间聚类方法。本文方法 RKMVSC 通过核诱导映射将各视图数据投影至特征空间进行子空间学习, 并采用 $l_{2,1}$ 范数约束表示误差以增强模型鲁棒性。实验结果表明, 所提方法在多个基准数据集上取得了较好的聚类效果, 验证了核化策略与 $l_{2,1}$ 范数约束在多视图子空间聚类中的有效性。

基金项目

国家自然科学基金面上项目(12571563)。

参考文献

- [1] Vidal, R. (2011) Subspace Clustering. *IEEE Signal Processing Magazine*, **28**, 52-68. <https://doi.org/10.1109/msp.2010.939739>
- [2] Elhamifar, E. and Vidal, R. (2013) Sparse Subspace Clustering: Algorithm, Theory, and Applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **35**, 2765-2781. <https://doi.org/10.1109/tpami.2013.57>
- [3] Liu, G., Lin, Z., Yan, S., Sun, J., Yu, Y. and Ma, Y. (2013) Robust Recovery of Subspace Structures by Low-Rank Representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **35**, 171-184. <https://doi.org/10.1109/tpami.2012.88>
- [4] Gao, H., Nie, F., Li, X. and Huang, H. (2015) Multi-View Subspace Clustering. 2015 *IEEE International Conference on Computer Vision (ICCV)*, Santiago, 7-13 December 2015, 4238-4246. <https://doi.org/10.1109/iccv.2015.482>
- [5] Xu, X.X., Tsang, I.W. and Dong Xu, (2013) Soft Margin Multiple Kernel Learning. *IEEE Transactions on Neural Networks and Learning Systems*, **24**, 749-761. <https://doi.org/10.1109/tnnls.2012.2237183>
- [6] Shawe-Taylor, J. and Cristianini, N. (2004) Kernel Methods for Pattern Analysis. Cambridge University Press. <https://doi.org/10.1017/cbo9780511809682>
- [7] Xiao, S., Tan, M., Xu, D. and Dong, Z.Y. (2016) Robust Kernel Low-Rank Representation. *IEEE Transactions on Neural Networks and Learning Systems*, **27**, 2268-2281. <https://doi.org/10.1109/tnnls.2015.2472284>
- [8] Boyd, S., Parikh, N., Chu, E., Peleato, B. and Eckstein, J. (2011) Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. *Foundations and Trends® in Machine Learning*, **3**, 1-122. <https://doi.org/10.1561/22000000016>
- [9] Kumar, A., Rai, P. and Daumé, H. (2011) Co-Regularized Multi-View Spectral Clustering. *Proceedings of the 25th International Conference on Neural Information Processing Systems*, Granada, 12-15 December 2011, 1413-1421.
- [10] Winn, J. and Jovic, N. (2005) LOCUS: Learning Object Classes with Unsupervised Segmentation. *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, Beijing, 17-21 October 2005, 756-763. <https://doi.org/10.1109/iccv.2005.148>
- [11] Li, F.F., Fergus, R. and Perona, P. (2004) Learning Generative Visual Models from Few Training Examples: An Incremental Bayesian Approach Tested on 101 Object Categories. 2004 *Conference on Computer Vision and Pattern Recognition Workshop*, Washington, 27 June-2 July 2004, 178. <https://doi.org/10.1109/cvpr.2004.383>
- [12] Samaria, F.S. and Harter, A.C. (1994) Parameterisation of a Stochastic Model for Human Face Identification. *Proceedings of 1994 IEEE Workshop on Applications of Computer Vision*, Sarasota, 5-7 December 1994, 138-142. <https://doi.org/10.1109/acv.1994.341300>
- [13] Wen, J., Xu, Y. and Liu, H. (2020) Incomplete Multiview Spectral Clustering with Adaptive Graph Learning. *IEEE Transactions on Cybernetics*, **50**, 1418-1429. <https://doi.org/10.1109/tcyb.2018.2884715>