

多核低冗余表示学习的稳健多视图子空间聚类方法

李 鹭^{1,2}, 王 卓¹, 于 晓 洋², 陈 德 运¹, 张 英 涛³, 孙 广 路¹

(1. 哈尔滨理工大学计算机科学与技术学院, 黑龙江 哈尔滨 150080;

2. 哈尔滨理工大学仪器科学与技术博士后流动站, 黑龙江 哈尔滨 150080;

3. 哈尔滨工业大学计算机科学与技术学院, 黑龙江 哈尔滨 150001)

摘 要: 针对高维数据冗余性、噪声干扰等问题对多视图子空间聚类性能的影响, 提出一种多核低冗余表示学习的稳健多视图子空间聚类方法。首先, 通过分析揭示数据在核空间中的冗余性和噪声影响特性, 提出采用多核学习来获得局部视图数据的稳健低冗余表示, 并利用其替代原始数据实施子空间学习。其次, 引入张量分析模型进行多视图融合, 从全局角度学习不同视图子空间表示的潜在张量低秩结构, 在捕获视图间高阶相关性的同时保持其各异性专属信息。所提方法将稳健低冗余表示学习、视图专属子空间学习以及融合潜在子空间结构学习统一到一个目标函数中, 使其在迭代中相互促进。大量实验结果表明, 所提方法在多个客观评价指标方面均优于当前主流多视图聚类方法。

关键词: 低冗余表示学习; 子空间聚类; 多视图学习; 张量分析

中图分类号: TP391

文献标识码: A

DOI: 10.11959/j.issn.1000-436x.2021217

Robust multiview subspace clustering method based on multi-kernel low-redundancy representation learning

LI Ao^{1,2}, WANG Zhuo¹, YU Xiaoyang², CHEN Deyun¹, ZHANG Yingtao³, SUN Guanglu¹

1. School of Computer Science and Technology, Harbin University of Science and Technology, Harbin 150080, China

2. Postdoctoral Station of Instrument Science and Technology, Harbin University of Science and Technology, Harbin 150080, China

3. School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China

Abstract: Considering the impact of high dimensional data redundancy and noise interference on multiview subspace clustering, a robust multiview subspace clustering method based on multi-kernel low redundancy representation learning was proposed. Firstly, by analyzing and revealing the redundancy and noise influence characteristics of data in kernel space, a multi-kernel learning method was proposed to obtain a robust low-redundancy representation of local view-specific data, which was utilized to replace the original data to implement subspace learning. Secondly, a tensor analysis model was introduced to carry out multiview fusion, so as to learn the potential low-rank tensor structure among different subspace representations from global perspective. It would capture the high-order correlation among views while maintaining their unique information. In this method, robust low-redundancy representation learning, view-specific subspace learning and fusion potential subspace structure learning were unified into the same objective function, so that they could promote each other during iterations. A large number of experimental results demonstrate that the proposed method is superior to the existing mainstream multiview clustering methods on several objective evaluation indicators.

Keywords: low-redundancy representation learning, subspace clustering, multiview learning, tensor analysis

收稿日期: 2021-09-01; **修回日期:** 2021-11-02

基金项目: 国家自然科学基金资助项目 (No.62071157); 黑龙江省自然科学基金资助项目 (No.YQ2019F011); 黑龙江省青年创新人才计划基金资助项目 (No.UNPYSCT-2018203); 黑龙江省高等学校基本科研业务费专项资金资助项目 (No.LGYC2018JQ013); 黑龙江省博士后启动基金资助项目 (No.LBH-Q19112)

Foundation Items: The National Natural Science Foundation of China (No.62071157), The Natural Science Foundation of Heilongjiang Province (No.YQ2019F011), University Nursing Program for Young Scholars with Creative Talents in Heilongjiang Province (No.UNPYSCT-2018203), The Fundamental Research Funds for the University of Heilongjiang Province (No.LGYC2018JQ013), Postdoctoral Foundation of Heilongjiang Province (No.LBH-Q19112)

1 引言

随着数据科学的发展与数据形式的多样化,现实世界的对象通常可以通过多个视角进行更完整和充分的描述。例如,图像可以通过颜色、边缘、纹理等进行描述;网页可以通过页面的文本和指向它们的超链接进行描述等,上述类型的数据称为多视图数据。多视图数据的每个视图都是关于同一对象的一种描述,不同视图通常从不同角度描述对象的某种特性,同时分析所有视图并将它们包含的信息进行融合学习的过程称为多视图学习^[1]。

子空间聚类是近年来提出的一类应用于高维数据聚类的有效方法,其通过学习高维数据本身固有的子空间结构来表征数据内在关联关系,然后利用子空间表示构造描述数据近邻关系的亲和矩阵,并对其实施图切划分算法获得聚类结果。因此,子空间表示矩阵的优劣是影响子空间聚类性能的关键因素。目前主流的子空间聚类算法中,大多假设每个样本都可以通过自身的线性组合来进行重构,并通过对自表示矩阵施加适当的约束,来有效地学习数据的子空间结构^[2]。稀疏子空间聚类(SSC, sparse subspace clustering)^[3]是子空间聚类的典型代表之一,该方法认为样本在重构时仅选择少量的近邻样本,因此考虑施加 l_1 范数从而获得具有稀疏性的自表示矩阵。与 SSC 相比,低秩表示(LRR, low-rank representation)^[4]认为有些样本远离了底层的子空间,并且利用 $l_{2,1}$ 范数将重构误差矩阵正则化。Fu 等^[5]提出了一种基于张量低秩表示和稀疏编码的子空间聚类方法,同时考虑样本的特征信息和空间结构,寻求在所有空间方向上对原始空间结构的最低秩表示。Wang 等^[6]提出了一种低秩表示约束的稳健子空间聚类模型,在寻求数据的低秩表示时将监督信息作为硬约束条件,以增强表示的鉴别能力。由于自表示的稀疏性和连通性在有效的子空间聚类中起着重要作用,Yang 等^[7]提出了一种后处理技术来优化稀疏性和连通性,利用初始亲和矩阵中包含的相关信息为每个样本找到较好的邻居,重新分配选定的邻居的系数并消除其他值,生成一个新的系数矩阵,进而改善子空间聚类性能。类似地,许多扩展的子空间聚类方法相继被提出,并得到了广泛的应用。

然而,随着数据呈现形式的发展,考虑到单独的视图并不足以全面描述数据的信息,子空间聚类

已经被扩展到多视图的情况,以获得性能的进一步提升。多视图子空间聚类是通过融合不同视图数据的自表示重构来学习一个公共的自表示矩阵,该自表示矩阵融合了不同视图的子空间结构信息,进而得到一个从全局角度进行完整数据近邻关系描述的子空间表示。Xu 等^[8]提出多视图学习方法应该充分利用共识和互补的原则以确保成功。为了实现共识原则,一个典型的策略是通过子空间学习来获得一个可以由多视图共享的潜在子空间^[9]。为了探索不同视图之间的互补信息,Gao 等^[10]提出了一种多视图子空间聚类算法,该算法对数据的每一个视图进行子空间聚类,利用一个公用的类指示矩阵来保证数据的一致性。Zhang 等^[2]提出了潜在多视图子空间聚类方法,学习基于多视图特征的潜在表示,并生成一个公共子空间表示,探索多视图间潜在的一致性信息。进一步地,Li 等^[11]提出针对子空间聚类的灵活多视图表示学习,旨在通过引入希尔伯特-施密特独立标准(HSIC, Hilbert-Schmidt independence criterion),潜在表示可以灵活地编码来自不同视图的互补信息,并探索不同视图之间的非线性关系。张苗涵等^[12]通过对隐式子空间表示施加低秩和稀疏约束来探索多视角之间的互补信息。受 SSC 和 LRR 的启发,Zhang 等^[13]提出了低秩张量约束多视图子空间聚类算法,利用不同模态展开的秩约束子空间张量,将基于 LRR 的子空间聚类扩展到多视图学习中,很好地探索来自多个源的互补信息,并大幅改进了子空间聚类的性能。该结构的多视图子空间学习策略使这些聚类方法都仅使用原始特征空间中数据的线性子空间,这并不足以捕获真实数据之间复杂的相关性。对此,Abavisani 等^[14]提出了一种基于 SSC 和 LRR 的子空间聚类算法的多模态扩展方法,并利用核学习适应数据的非线性特性。为了解决原始的高维特征向量包含噪声或冗余信息随维数的增加而呈指数级增长的问题,Kang 等^[15]提出了一种分区级别的多视图子空间聚类方法寻找来自多个分区的共识聚类,这种分区级融合方法可以更好地利用隐藏的聚类结构信息。Liu 等^[16]提出一种通过协同训练稳健数据表示进而实现多视图子空间聚类的方法,该方法将聚类与稳健数据表示的生成融合到一个模型中来获得聚类结果。

尽管上述多视图子空间聚类方法取得了一定的效果,但仍存在 2 个主要的因素限制了其性能的

提升。一方面，真实环境下的数据常常受到噪声的干扰，而这些含噪数据的高维性又会使数据本身以及数据之间存有较大的冗余特性，噪声和冗余会导致自表示过程难以学习到数据本质的子空间结构；另一方面，传统多视图学习方式忽略了视图间的高阶关联性，也没有充分考虑不同视图的差异性结构，导致融合后的子空间表示矩阵性能并不理想。面向上述 2 个问题，考虑从噪声和冗余性去除、视图间高阶关联性捕获及各异性结构信息保持 2 个角度出发，本文提出一种低冗余稳健表示的多视图子空间聚类方法。为了适应数据的非线性特性，所提方法将数据在核空间进行分析，从局部角度（视图内）构建多核的低冗余稳健表示学习模型，学习适当的数据低维表达，消除噪声干扰和冗余性对子空间结构探究的影响。同时，考虑从全局角度（视图间）充分挖掘视图间的高阶关联结构，采用低秩张量约束，融合多视图信息来寻求不同视图的潜在张量子空间表示矩阵。本文稳健多视图子空间学习算法的总体结构如图 1 所示。

2 相关工作

本文主要探究多核低冗余表示学习的多视图子空间聚类方法，下面将介绍本文方法所涉及的一些基础性相关工作及局限性分析。

2.1 子空间聚类

对于给定来自 k 个类簇的含有 n 个数据的数据集 $\mathbf{X} \in \mathbb{R}^{d \times n}$ ，子空间聚类算法的目的是寻找一个利用样本自身作为字典进行数据重构的子空间表示矩阵 $\mathbf{Z} \in \mathbb{R}^{n \times n}$ ，其一般的表示形式为

$$\min_{\mathbf{Z}} \mathcal{L}(\mathbf{X}, \mathbf{XZ}) + \lambda \Omega(\mathbf{Z}) \quad (1)$$

其中， $\mathcal{L}(\cdot)$ 和 $\Omega(\cdot)$ 分别为自表示重构损失函数和正

则化约束， λ 为正则化参数。大量文献针对式(1)中的表达式设计了相应的重构损失函数和约束形式，以适应不同类型数据的子空间结构约束，不断提升子空间学习的性能。在获得子空间表示矩阵 \mathbf{Z} 后，一般利用其构造对称的亲矩阵 $\mathbf{A} = \frac{1}{2}(|\mathbf{Z}| + |\mathbf{Z}^T|)$ ，再将该亲和矩阵送入标准谱聚类算法得到最后的聚类结果。

2.2 多视图子空间聚类

对于给定含有 V 个视图的数据集合 $\{\mathbf{X}_v\}_{v=1}^V$ ，可以将式(1)简单拓展为如下的多视图子空间学习模型

$$\min_{\{\mathbf{Z}_v\}_{v=1}^V} \mathcal{L}(\{\mathbf{X}_v, \mathbf{X}_v \mathbf{Z}_v\}_{v=1}^V) + \lambda \Omega(\{\mathbf{Z}_v\}_{v=1}^V) \quad (2)$$

其中， $\mathbf{X}_v \in \mathbb{R}^{d_v \times n}$ 表示第 v 个视图的原始数据， d_v 表示第 v 个视图的特征维度， n 表示样本个数， $\mathbf{Z}_v \in \mathbb{R}^{n \times n}$ 表示第 v 个视图的子空间表示矩阵。然而，多视图学习的目的是期望融合来自不同视图的信息实现信息的共享与互补，而式(2)中的形式是对各视图数据独立地进行学习，并没有充分地考虑视图间的关联性，不是真正意义上的多视图子空间学习。为了解决这一问题并建模视图间关联性，近年来提出了大量基于一致性共享结构的子空间学习模型，这些方法大致可以分为以下 2 种主要的多视图共享方式。

1) 考虑到多视图数据是相同对象的不同表达形式，因此各视图数据应共享同一子空间结构（如文献[11,14]），基于式(2)中的模型，其共享子空间结构学习可通过式(3)实现。

$$\min_{\mathbf{Z}} \mathcal{L}(\{\mathbf{X}_v, \mathbf{X}_v \mathbf{Z}\}_{v=1}^V) + \lambda \Omega(\mathbf{Z}) \quad (3)$$

该方式认为不同的视图数据均共享相同的子

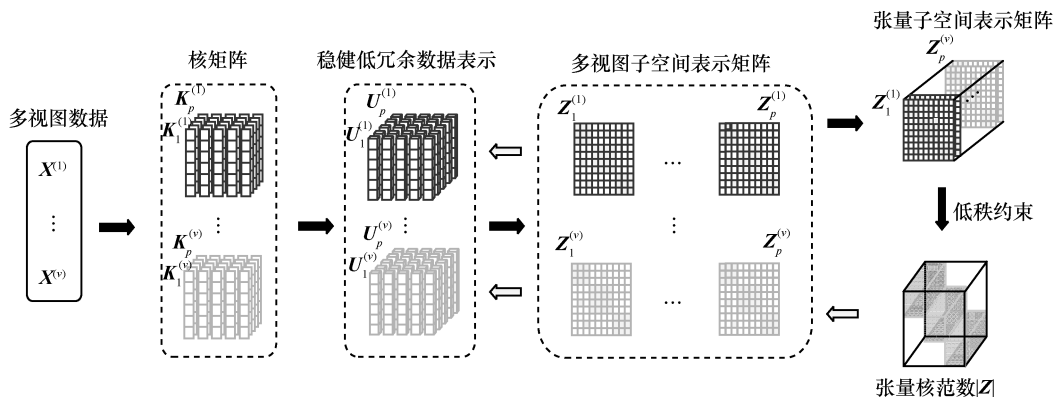


图 1 稳健多视图子空间学习算法的总体结构

空间结构 \mathbf{Z} ，通过对 \mathbf{Z} 施加对所有视图数据的重构约束来融合不同视图对子空间结构的互补性信息，并利用一个统一的正则项 $\Omega(\mathbf{Z})$ 约束子空间表示矩阵的先验特性。

2) 考虑视图间的异质或多源结构可能引起的表示矩阵差异性，构造满足一致性约束的共享子空间学习模型（如文献[15,17]），其表达式为

$$\min_{\{\mathbf{Z}_v\}_{v=1}^V, \mathbf{Z}} \mathcal{L}(\{\mathbf{X}_v, \mathbf{X}_v \mathbf{Z}_v\}_{v=1}^V) + \lambda \Omega(\{\mathbf{Z}_v\}_{v=1}^V, \mathbf{Z}) \quad (4)$$

区别于式(3)中迫使表示矩阵在各视图中完全相同的做法，式(4)在每个视图学习各自的子空间表示矩阵，再利用 $\Omega(\{\mathbf{Z}_v\}_{v=1}^V, \mathbf{Z})$ 约束共享子空间与各视图子空间矩阵的关系来实现多视图信息的融合，更灵活地从不同视图中学习一个共享的子空间表示。

然而，上述基于共享式的多视图子空间学习方法在具体聚类应用中仍存在一定的局限性。一方面，上述方法中直接采用多视图原始数据进行子空间学习，而这些原始高维数据往往存在一定的冗余性，且从真实环境中获取时也不可避免地会受到噪声的干扰。数据的冗余结构和噪声会影响重构损失的自表示过程，进而不能学习到真实的子空间表示结构，导致后续聚类性能的下降；另一方面，上述的 2 种子空间共享方法更多地考虑子空间结构的一致性，却忽略了对视图数据异质性而引起的表示矩阵差异性的保持，因此在保持表示矩阵各异性的前提下更好地融合不同视图的近似结构也是本文需要解决的一个关键问题。

3 本文方法描述及其数值算法

3.1 核空间冗余性分析及低冗余表示学习

针对上述相关工作的描述，为了消除原始数据中的冗余信息和噪声，并考虑适应高维数据潜在的非线性特性，本文考虑在核空间学习数据的稳健低冗余表示，并利用其替代原始数据来学习数据的真实的本质子空间结构，实现提高子空间聚类性能的目的。给定核函数集合 $\{\phi_s(\cdot)\}_{s=1}^S$ ，第 v 个视图的第 s 个核矩阵 $\mathbf{K}_s^{(v)}$ 为

$$\mathbf{K}_s^{(v)}(i, j) = \phi_s(\mathbf{x}_i^{(v)})^\top \phi_s(\mathbf{x}_j^{(v)}) \quad (5)$$

其中， $i, j \in \{1, 2, \dots, n\}$ 表示实例样本索引， $\mathbf{x}_i^{(v)}$ 表示 \mathbf{X}_v 的第 i 列数据。由此可知，当有 V 个视图、 S 种核映射时，将会产生 $m = SV$ 个相应的核矩阵

$$\{\mathbf{K}_p\}_{p=1}^m。$$

为了在核空间对数据进行分析，本文选取多项式核函数计算 BBC-Sport 数据集中第一个视图数据的核矩阵。然后，对核矩阵进行特征分解，将得到的特征值从大到小排序并进行分布统计。图 2(a)展示了特征值分布统计曲线，横坐标表示特征值序号，纵坐标表示当前序号之前的特征值之和与特征值总和的比值。由文献[18]可知，对应于较大特征值的特征向量会携带更多的判别信息。选用特征值大小作为判别信息的近似度量，可以从图 2(a)中观察到 2 个重要的现象。一方面，尽管特征值总个数超过 500，但其前 50 个特征值已包含了该数据集中超过 60% 的主要判别信息；另一方面，BBC-Sport 数据集共有 5 个类别，若只取前 5 个特征向量，其仅包含了 35.07% 的判别信息。

根据上述现象描述，可以揭示 2 个重要结论。首先，数据样本间的关键性判别信息只包含在一小部分大特征值对应的特征向量中，而其他大多数特征向量可认为是冗余信息。类似地，本文将归一化数据加入方差为 1 的高斯白噪声，将干扰数据对应的核矩阵也进行特征分解，并利用排序后 30% 的小特征值及其特征向量重建核矩阵，如图 2(b)所示，从图 2(b)中可以看出，小特征值重建的核矩阵更偏重于表达数据中由确定数据和随机噪声交叉内积项所产生的噪声分量核。因此，可考虑通过基于核的低冗余学习方式达到去除冗余和噪声的目的。其次，在传统多核聚类方法中将学到的表示特征维度 c 设为类簇数 k 的常用做法并不理想（如 BBC-Sport 中前 $k = 5$ 个特征值对应的判别信息量仅 35.07%），可考虑利用多核来学习满足 $c > k$ 条件的携带更多判别信息的低冗余表示，并将其替代原始数据进行后续的多视图子空间聚类。

基于上述考虑，采用如下的模型进行核稳健低冗余数据表示的学习。

$$\begin{aligned} & \arg \max_U \text{Tr}(\mathbf{UKU}^\top) \\ & \text{s.t. } \mathbf{UU}^\top = \mathbf{I} \end{aligned} \quad (6)$$

其中， $\mathbf{U} \in \mathbb{R}^{c \times n}$ 表示数据的低冗余表示特征矩阵， $\mathbf{K} \in \mathbb{R}^{m \times m}$ 表示由原始数据得到的核矩阵， $\text{Tr}(\bullet)$ 表示矩阵的迹。式(6)问题的求解是通过选取核矩阵 \mathbf{K} 的前 c 个特征向量重组而来。根据上述分析，以这种方式得到的数据表示矩阵 \mathbf{U} 不仅可以在去除冗余性的同时保留其主要判别信息，还能够一定程度

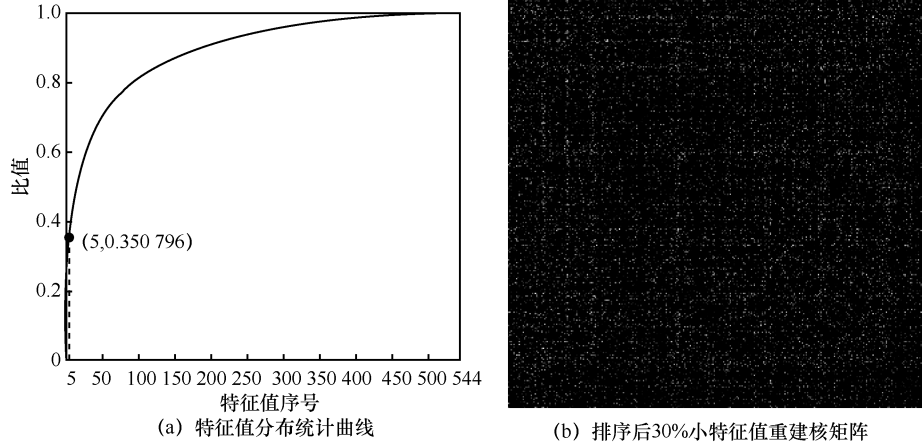


图 2 BBC-Sport 数据集的多项式核矩阵分析

上通过丢弃小特征值对应的特征向量来抑制噪声分量的干扰, 提高数据表示的稳健性。

3.2 低冗余表示的张量子空间学习

根据 3.1 节的分析, 考虑到数据的冗余性和噪声会导致自表示重构模型中难以学习到数据真实的本质子空间结构, 本文将用低冗余数据表示 U_p 代替原始数据 X_p , 提出如下基于低冗余表示的子空间学习模型。

$$\begin{aligned} \min_{Z_p} \lambda \|Z_p\|_F + \|U_p - U_p Z_p\|_F \\ \text{s.t. } U_p U_p^T = I \end{aligned} \quad (7)$$

其中, Z_p 为第 p 个视图的子空间表示矩阵, I 为单位阵, λ 为正则化参数, $\|\cdot\|_F$ 为 Frobinus 范数。

在获得了每个局部视图数据的子空间表示矩阵 Z_p 后, 本文期望能够将多视图信息进行融合, 从全局角度建模视图间的关联性。然而, 如 2.2 节所述, 现有多视图子空间学习策略主要是假设各视图间存在一个共享子空间表示矩阵, 这种融合策略忽略了视图间的各异性和高维相关性, 对于具有二维特性的子空间表示矩阵不能实现其多维的关联性建模。为此, 考虑将各视图子空间表示矩阵 Z_p 集成来构造一个 3 阶张量矩阵 \mathcal{Z} , 并对 \mathcal{Z} 实施低秩约束来捕获不同视图之间的高阶相关性。一方面, 子空间表示系数在一定程度上反映了样本间的相似性关系, 由于是相同对象的不同视图, 因此不同的 Z_p 应包含相近的相似性信息。另一方面, 类簇的数量总是比样本数量要小得多, 所以张量 \mathcal{Z} 具有潜在的低秩特性。基于上述两点考虑, 提出如下的基于低冗余数据表示的张量子空间学习模型。

$$\begin{aligned} \min_{\{U_p\}_{p=1}^m, \{Z_p\}_{p=1}^m, \mathcal{Z}} \lambda \|\mathcal{Z}\|_{\otimes} + \sum_{p=1}^m \beta_p \|U_p - U_p Z_p\|_F^2 \\ \text{s.t. } U_p U_p^T = I, \beta^{\frac{1}{2}T} \mathbf{1} = 1, \beta = [\beta_1, \dots, \beta_m]^T \end{aligned} \quad (8)$$

其中, $\mathcal{Z} = [Z_1; Z_2; \dots; Z_m]$ ($\mathcal{Z} \in \mathbb{R}^{n \times n \times m}$) 表示由 Z_p 堆叠成的三阶张量, β 表示重构损失权重参数。考虑低秩约束的非凸性, 这里采用张量核范数 $\|\cdot\|_{\otimes}$ 约束作为其凸近似, 便于后续模型求解。

式(8)中采用张量低秩约束形式进行多视图信息融合的优势体现在 2 个层面。首先, 区别于传统共享式模型, 张量约束形式不强制要求将不同视图的信息融入一个公共的子空间矩阵中, 而是利用低秩来捕获各视图子空间矩阵的一个联合的潜在相关性结构。这个结构仍是一个由多个视图构成的具有相关性的张量体, 其在融合一致性结构的同时有效地保持了不同视图的差异化信息。其次, 在张量低秩的求解过程中, 会对张量立方体从多个维度探寻其低秩相关性信息, 相比于共享方式中采用的二维约束形式, 能更全面地挖掘视图间的互补性信息, 提高多视角融合的多维互补性。

3.3 基于稳健表示学习的多视图子空间聚类方法

联合上述的低冗余学习和张量子空间学习模型, 将多核稳健表示学习与张量子空间学习放入统一的学习框架, 提出本文稳健多视图子空间学习目标函数。

$$\begin{aligned} \min_{\{U_p\}_{p=1}^m, \{Z_p\}_{p=1}^m, \mathcal{Z}} \lambda \|\mathcal{Z}\|_{\otimes} + \sum_{p=1}^m \beta_p \|U_p - U_p Z_p\|_F^2 - \\ \sum_{p=1}^m \gamma_p \text{Tr}(U_p K_p U_p^T) \\ \text{s.t. } U_p U_p^T = I, \beta^{\frac{1}{2}T} \mathbf{1} = 1, \gamma^T \gamma = 1 \end{aligned} \quad (9)$$

其中, $\gamma = [\gamma_1, \dots, \gamma_m]$ 为权重参数。本文方法通过子空间表示学习模型将低冗余表示 U_p 、子空间表示矩阵 Z_p 以及潜在张量子空间表示矩阵 \mathcal{Z} 放入同一目标函数, 实现三者的联合学习。从式(9)目标函数中可以看出, 低冗余数据表示 U_p 将同时从低冗余表示学习和低秩张量子空间表示矩阵 \mathcal{Z} 中学习而来, 视图专属子空间学习过程在挖掘视图内相关性的同时, 还挖掘了视图间的高阶相关性。这种来自多视图融合的潜在信息也通过 Z_p 传递到 U_p 的低冗余表示学习的过程, 以获得更好的低冗余数据表示。对于潜在张量子空间结构 \mathcal{Z} 的学习, 由于本框架 Z_p 在迭代过程中是动态调整的, 随着低冗余数据表示质量的提高, 模型能够联合学习到更好的视图专属子空间表示及其融合的张量子空间表示, 进一步改善模型的稳健性。通过交替的优化低冗余表示、视图专属子空间表示和潜在张量子空间表示, 使它们在迭代过程中相互促进, 以获得目标函数的联合最优解。

在获得了张量子空间矩阵 \mathcal{Z} 后, 利用 $J = \left(\sum_{p=1}^m \left(\|\mathcal{Z}^{(p)}\| + \left(\mathcal{Z}^{(p)} \right)^T \right) \right) / m$ 计算亲和矩阵 J ($\mathcal{Z}^{(p)}$ 表示 \mathcal{Z} 沿视角方向的第 p 个切片), 并将其送入谱聚类算法得到本文方法的聚类结果。

3.4 目标函数数值求解方案

本节设计了一种有效的数值优化算法来求解式(9)中的目标函数。由于变量 \mathcal{Z} 和 Z_p 之间的耦合性, 使式(9)目标函数中张量低秩约束的求解具有一定的困难。为了解决变量间的耦合关系, 采用交替方向乘子法^[19], 引入辅助变量 \mathcal{Q} 使变量可分离, 将式(9)转化为

$$\begin{aligned} & \min_{\{U_p\}_{p=1}^m, \{Z_p\}_{p=1}^m, \mathcal{Q}} \lambda \|\mathcal{Q}\|_{\otimes} + \sum_{p=1}^m \beta_p \|U_p - U_p Z_p\|_F^2 - \\ & \sum_{p=1}^m \gamma_p \text{Tr}(U_p K_p U_p^T) \\ & \text{s.t. } \mathcal{Q} = \mathcal{Z}, \beta^{\frac{1}{2}T} \mathbf{1} = 1, \gamma^T \gamma = 1 \end{aligned} \quad (10)$$

在式(10)的基础上, 构造其相应的拉格朗日增广函数, 得到下面的可分离优化形式, 并采用交替优化技术来计算式(9)的近似解

$$\begin{aligned} & \min_{\{U_p\}_{p=1}^m, \{Z_p\}_{p=1}^m, \mathcal{Q}, \mathcal{Z}} \lambda \|\mathcal{Q}\|_{\otimes} + \|\mathcal{Q} - \mathcal{Z}\|_F^2 + \\ & \sum_{p=1}^m \beta_p \|U_p - U_p Z_p\|_F^2 - \sum_{p=1}^m \gamma_p \text{Tr}(U_p K_p U_p^T) \end{aligned}$$

$$\text{s.t. } \beta^{\frac{1}{2}T} \mathbf{1} = 1, \gamma^T \gamma = 1 \quad (11)$$

对于变量 \mathcal{Z} , 可以明显地看出各子空间表示矩阵 $\{Z_p\}_{p=1}^m$ 都是独立的, 固定其他变量, 式(11)可变为

$$\begin{aligned} & \min_{Z_p} \|Z_p - \mathcal{Q}_p\|_F^2 + \beta_p \|U_p - U_p Z_p\|_F^2 \\ & \text{s.t. } \beta^{\frac{1}{2}T} \mathbf{1} = 1 \end{aligned} \quad (12)$$

令 $Z_p - \mathcal{Q}_p = Z'_p$, $U'_p = U_p - U_p \mathcal{Q}_p$, 可以得到

$$\begin{aligned} & \min_{Z'_p} \|Z'_p\|_F^2 + \beta_p \|U'_p - U_p Z'_p\|_F^2 \\ & \text{s.t. } \beta^{\frac{1}{2}T} \mathbf{1} = 1 \end{aligned} \quad (13)$$

式(13)可通过如下的形式分别优化 Z'_p 的每一列 $z_p^{(i)}$ 进行求解

$$\begin{aligned} & \min_{z_p^{(i)}} \|z_p^{(i)}\|_F^2 + \beta_p \|u_p^{(i)} - U_p z_p^{(i)}\|_F^2 \\ & \text{s.t. } \beta^{\frac{1}{2}T} \mathbf{1} = 1 \end{aligned} \quad (14)$$

式(14)是二次凸规划问题, 令其导数为 0, 可得到 $z_p^{(i)}$ 的闭式解

$$z_p^{(i)} = (I + \beta_p U_p^{(i)T} U_p^{(i)})^{-1} \beta_p u_p^{(i)} U_p^{(i)T} \quad (15)$$

在获得 Z'_p 后, 根据 $Z_p - \mathcal{Q}_p = Z'_p$, 即可得到 Z_p 的解。

对于变量 \mathcal{Q} , 固定其他变量, 可得到如下的子优化问题

$$\min_{\mathcal{Q}} \lambda \|\mathcal{Q}\|_{\otimes} + \|\mathcal{Q} - \mathcal{Z}\|_F^2 \quad (16)$$

式(16)的张量核范数约束问题可根据文献[20]中基于 t-SVD 的张量核范数最小化算法进行求解, 得到变量 \mathcal{Q} 的闭式解。

对于变量 U , 由于低冗余数据表示 $\{U_p\}_{p=1}^m$ 在式(11)中是相互独立的, 固定其他变量, 可分别对每个 U_p 采用式(17)进行优化

$$\begin{aligned} & \min_{U_p} \beta_p \|U_p - U_p Z_p\|_F^2 - \gamma_p \text{Tr}(U_p K_p U_p^T) \\ & \text{s.t. } U_p U_p^T = I, \beta^{\frac{1}{2}T} \mathbf{1} = 1, \gamma^T \gamma = 1 \end{aligned} \quad (17)$$

为了便于求解, 将式(17)改写为

$$\begin{aligned} & \max_{U_p} \text{Tr}(U_p M_p U_p^T) \\ & \text{s.t. } M_p = 2\beta_p Z_p^T - \beta_p Z_p Z_p^T + \gamma_p K_p \\ & U_p U_p^T = I \end{aligned} \quad (18)$$

式(18)的迹最大化问题可通过对 M_p 的特征分解进行求解, 其中 U_p 是由矩阵 M_p 的前 c 个最大特征值所对应的特征向量组成的。

综上所述, 本文提出的目标函数的求解步骤如算法 1 所示。

算法 1 本文提出的目标函数的求解

输入 原始数据 $\{X_v\}_{v=1}^V$, 低冗余数据表示的维度 c 和正则化参数 λ

输出 张量子空间表示矩阵 \mathcal{Z}

1) 生成原始数据 $\{X_v\}_{v=1}^V$ 所对应的核矩阵 $\{K_p\}_{p=1}^m$

2) 初始化 \mathcal{Q} , $\{U_p\}_{p=1}^m$, β 和 γ

3) while not convergence do

4) 通过式(15)更新 Z_p ;

5) 构造张量 $\mathcal{Z} = [Z_1, Z_2, \dots, Z_P]$;

6) 通过式(16)更新 \mathcal{Q} ;

7) 通过式(18)更新 $\{U_p\}_{p=1}^m$;

8) 优化参数 β ;

9) 优化参数 γ ;

10) end while

3.5 参数优化

关于式(9)中的参数 β 和 γ , 实际上反映了不同视图的子空间重构损失和低冗余表示学习中的贡献度, 传统的经验设定方式不能有效获得目标函数的联合最优解。因此, 本文将这 2 个参数也看作一种特殊的变量, 与其他变量类似, 在归一化约束条件下给出如下的参数优化方案。

1) 对于变量 β , 固定式(11)中的其他变量, 关于 β 的目标函数可以简化为

$$\begin{aligned} \min_{\beta} \beta^T \mathbf{v} \\ \text{s.t. } v_p = \|U_p - U_p Z_p\|_F^2, \beta^{\frac{1}{2}T} \mathbf{1} = 1 \end{aligned} \quad (19)$$

根据柯西-施瓦兹不等式有

$$\begin{aligned} (\beta^T \mathbf{v}) \left(\sum_{p=1}^m \frac{1}{v_p} \right) &= \left(\sum_{p=1}^m (\sqrt{\beta_p} \sqrt{v_p})^2 \right) \left(\sum_{p=1}^m \left(\frac{1}{\sqrt{v_p}} \right)^2 \right) \geq \\ &\left(\sum_{p=1}^m \sqrt{\beta_p} \right)^2 = 1 \end{aligned} \quad (20)$$

当且仅当 $v_1 \sqrt{\beta_1} = v_2 \sqrt{\beta_2} = \dots = v_m \sqrt{\beta_m}$ 时, 式(20)可取得最小值。考虑式(19)中 β 的约束条件, 可

通过式(21)计算参数 β

$$\beta_p = \frac{1}{\left(v_p \sum_{q=1}^m \frac{1}{v_q} \right)^2} \quad (21)$$

2) 对于变量 γ , 固定式(11)中的其他变量, 关于 γ 的目标函数可以简化为

$$\begin{aligned} \max_{\gamma} \gamma^T \mathbf{v} \\ \text{s.t. } v_p = \text{Tr}(U_p K_p U_p^T), \gamma^T \gamma = 1 \end{aligned} \quad (22)$$

与 β 的求解相似, 有

$$(\gamma^T \mathbf{v}) \left(\sum_{p=1}^m \gamma_p v_p \right)^2 \leq \left(\sum_{p=1}^m \gamma_p^2 \right) \left(\sum_{p=1}^m v_p^2 \right) = (\gamma^T \gamma) (\mathbf{v}^T \mathbf{v}) = \mathbf{v}^T \mathbf{v} \quad (23)$$

当且仅当 $\gamma_1 / v_1 = \gamma_2 / v_2 = \dots = \gamma_m / v_m$ 存在时, 式(23)可取得最大值。考虑式(22)中 γ 的约束条件, 可通过式(24)来计算参数 γ

$$\gamma_p = \frac{v_p}{\left(\sum_{q=1}^m v_q^2 \right)^{\frac{1}{2}}} \quad (24)$$

4 实验与结果分析

4.1 数据集描述

本文在 3 个公开数据集上进行多视图聚类实验, 数据集的具体情况阐述如下。

1) BBC-Sport 数据集: 由 BBC-Sport 网站的 737 份文件组成, 对应于 5 个主题领域的体育新闻, 包括田径、板球、足球、橄榄球和网球, 共有 2 个不同的视图。

2) ORL 数据集: 由 40 个不同性别的人的 400 张面部图像组成, 其中每个人有 10 张不同拍摄角度的人脸图像。本文将对数据集中的每个样本提取灰度强度 (Gray)、局部二值模式 (LBP) 和 Gabor 这 3 种特征来表示这些人脸图像, 共有 3 个不同的视图。

3) UCI-Digits 数据集: 由对应 10 个类别的 2000 个数字图像组成, 分别提取傅里叶系数、像素平均和形态学特征 3 个不同的特征来表示这些数字图像, 共有 3 个不同的视图。

4.2 对比实验设置及实验结果与分析

为了证明所提方法的有效性, 本文将在上述 3 个数据集上与 5 种先进的多视图子空间聚类方法进行对比。这 5 种方法分别是文献[2]中潜在多视图子空

间聚类 (LMSC, latent multiview subspace clustering)、文献[14]中多模态稀疏和低秩子空间聚类 (MSSC, multimodal sparse and low-rank subspace clustering)、文献[15]中基于划分的多视图子空间聚类 (PMSC, partition level multiview subspace clustering)、文献[11]中弹性多视图表示学习的子空间聚类 (FMR, flexible multiview representation learning for subspace clustering)以及文献[16]中协同稳健多视图子空间聚类 (CoMSC, multiview subspace clustering via co-training robust)。

本文方法采用高斯、多项式、线性 3 种核函数, 初始化参数 $\beta=\gamma=1/m$, 此后 β 和 γ 将在迭代过程中根据式(21)和式(24)分别计算。此外, 根据 3.1 节中的分析, 为了获得更判别的稳健表示, 设置低冗余数据表示的维度 c 的取值范围为 $\{k, 2k, \dots, 20k\}$ (k 为类簇数), 参数 λ 的取值范围为 $2^{\{-10, -8, \dots, 10\}}$, 并选取最优聚类性能的参数组合获得实验结果。其中, BBC-Sport 数据集取 $c=6k$ 、 $\lambda=2^{-2}$; ORL 数据集取 $c=2k$ 、 $\lambda=2^{-2}$; UCI-Digits 数据集取 $c=2k$ 、 $\lambda=2^{-2}$, 其他对比方法依据其文献描述进行参数设定。实验中采用准确度 (ACC, accuracy)、归一化互信息 (NMI, normalized mutual information) 和 F-measure 这 3 种度量方式对聚类结果进行客观评价, 各对比方法在每个数据集上运行 10 次, 并取 10 次实验结果的均值作为聚类评价指标的实验结果。表 1~表 3 中列出了所有对比方法在 3 个数据集上的聚类指标结果, 将各指标的最优结果进行加粗, 以便进行比较。

表 1 在 BBC-Sport 数据集上进行不同方法的 ACC、NMI 和 F-measure 比较

方法	ACC	NMI	F-measure
LMSC	0.915±0.00	0.839±0.00	0.900±0.00
MSSC	0.971±0.00	0.898±0.00	0.944±0.00
PMSC	0.472±0.04	0.202±0.08	0.445±0.04
FMR	0.946±0.01	0.832±0.02	0.898±0.02
CoMSC	0.963±0.00	0.881±0.00	0.923±0.00
本文方法	0.980±0.00	0.946±0.00	0.973±0.00

从表 1~表 3 可以看出, 本文方法在 3 种指标上几乎都优于其他对比方法。在 ACC 指标中, 本文方法在 BBC-Sport、ORL 和 UCI-Digits 数据集上相比次优方法依次提升 0.9%、1.4%和 4.4%; 在 NMI 指标中, 本文方法仅在 ORL 数据集上略低于 MSSC

方法, 在其他数据集上均优于对比方法, 在 BBC-Sport 和 UCI-Digits 数据集上相比次优方法分别提升 4.8%和 7.3%; 在 F-measure 指标中, 本文方法在 BBC-Sport、ORL 和 UCI-Digits 数据集上相比次优方法依次提升 2.9%、0.4%和 8.3%。上述结果表明, 在原始数据上本文方法学习到的低冗余表示能够改善局部视图的子空间学习能力, 并通过低秩张量融合多视图信息, 提高多视图子空间聚类的性能。

表 2 在 ORL 数据集上进行不同方法的 ACC、NMI 和 F-measure 比较

方法	ACC	NMI	F-measure
LMSC	0.811±0.03	0.908±0.02	0.747±0.04
MSSC	0.832±0.02	0.927±0.01	0.787±0.02
PMSC	0.551±0.04	0.732±0.03	0.394±0.06
FMR	0.809±0.02	0.898±0.01	0.741±0.02
CoMSC	0.812±0.01	0.901±0.00	0.749±0.01
本文方法	0.846±0.03	0.923±0.01	0.791±0.02

表 3 在 UCI-Digits 数据集上进行不同方法的 ACC、NMI 和 F-measure 比较

方法	ACC	NMI	F-measure
LMSC	0.763±0.05	0.718±0.02	0.682±0.04
MSSC	0.924±0.00	0.859±0.00	0.858±0.00
PMSC	0.636±0.07	0.624±0.04	0.576±0.04
FMR	0.865±0.01	0.771±0.01	0.760±0.01
CoMSC	0.930±0.00	0.870±0.00	0.866±0.00
本文方法	0.974±0.00	0.943±0.00	0.949±0.00

4.3 噪声条件下实验设置与结果

为了验证对比方法在噪声数据条件下的性能, 本文对上述数据集随机选取部分样本添加高斯白噪声进行聚类实验 (加噪样本比例是指加噪样本数占样本总数的比例), 并通过调整噪声方差和加噪样本比例来验证不同噪声强度下的模型性能。实验中, 设置噪声方差分别为 0.1 和 0.3, 加噪样本比例以 0.1 为间隔步长从 0.1 变化到 0.5, 所有视图原始数据均进行了范数归一化。

3 种指标在 2 种噪声方差下的实验结果分别如图 3 和图 4 所示。从实验结果可以看出, 随着加噪样本比例的不断上升, 本文方法在 3 种指标上几乎都优于其他对比方法。以不同加噪样本比例下的指标均值考量在 BBC-Sport、ORL 和 UCI-Digits 数据集上聚类性能的平均水平。当噪声方差为 0.1 时,

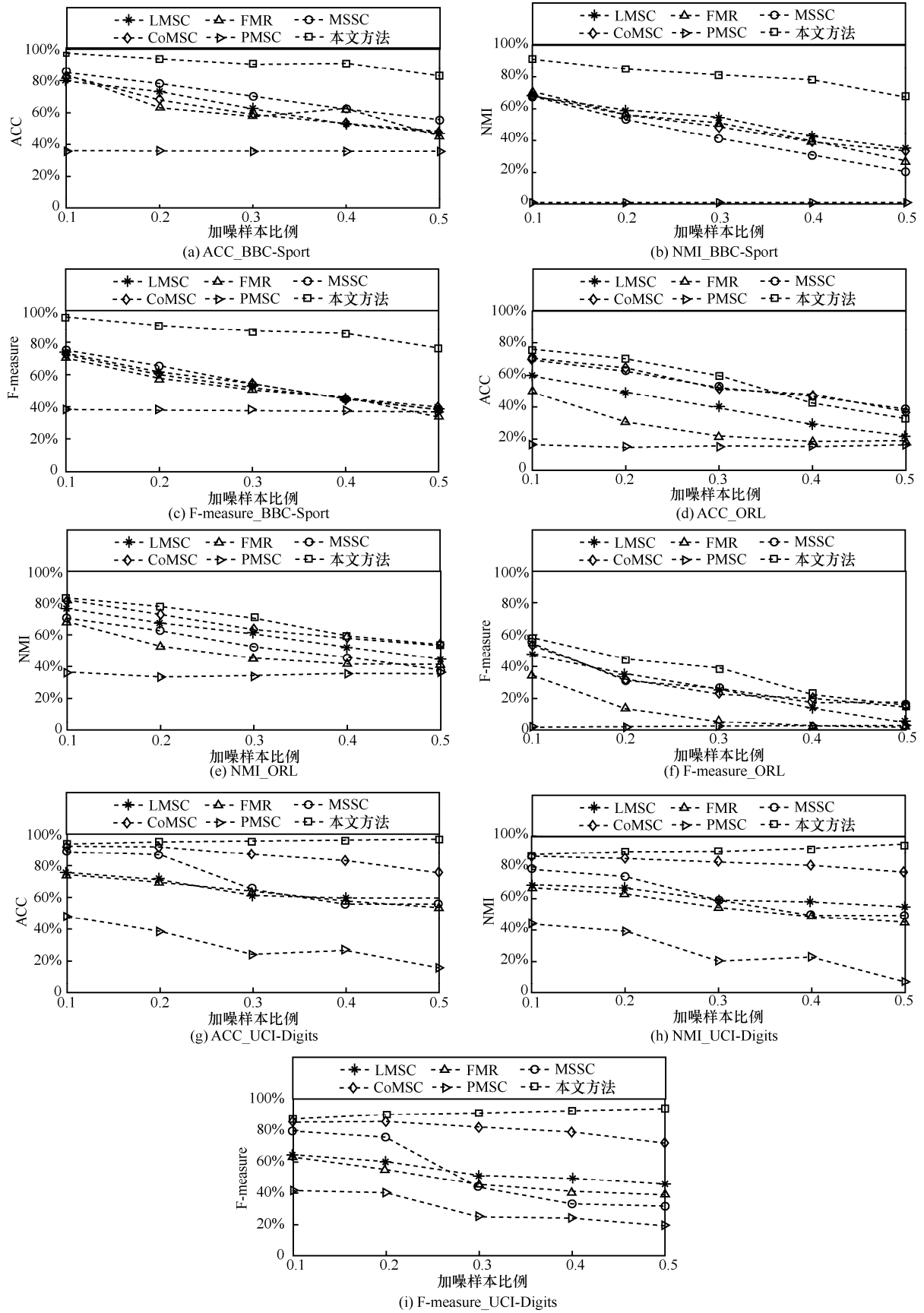


图 3 不同方法在噪声方差为 0.1 时的聚类结果

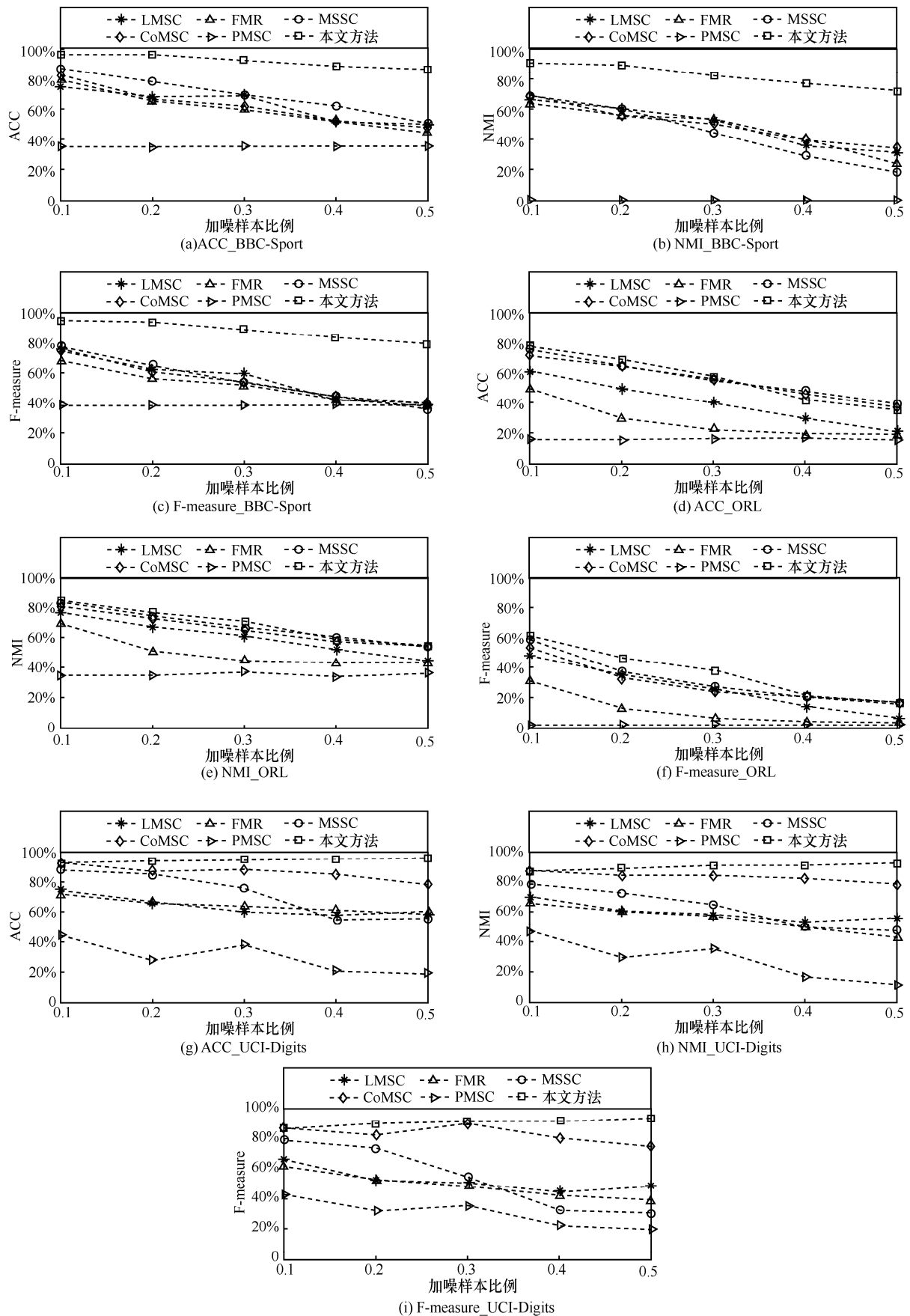


图 4 不同方法在噪声方差为 0.3 时的聚类结果

在 ACC 指标中, 本文方法在 3 个数据集上依次高于次优方法 20%、2.53% 和 9.55%; 在 NMI 指标中, 本文方法在 3 个数据集上依次高于次优方法 17.26%、2.27% 和 7.84%; 在 F-measure 指标中, 本文方法在 3 个数据集上依次高于次优方法 31.7%、6.96% 和 10.22%。当噪声方差为 0.3 时, 本文方法在 BBC-Sport 数据集上各评价指标的实验结果均有较明显的优势; 对于其他 2 个数据集的实验结果, 在 ACC 指标中, 本文方法在 ORL 和 UCI-Digits 数据集上分别高于次优方法 1% 和 8.5%; 在 NMI 指标中, 本文方法在 ORL 和 UCI-Digits 数据集上分别高于次优方法 1.72% 和 7.1%; 在 F-measure 指标中, 本文方法在 ORL 和 UCI-Digits 数据集上分别高于次优方法 5.14% 和 8.76%。综合上述验证实验结果, 说明了本文方法在噪声干扰下具有有效性和稳健性。

4.4 复杂度与收敛性分析

本文方法由 5 个子问题组成, 在每次迭代时, 变量 \mathcal{Z} 的求解需计算式(15)中逆矩阵, 其复杂度为 $\mathcal{O}(n^3)$ 。对于变量 \mathcal{Q} , 首先要对 $n \times m \times n$ 的三维张量进行快速傅里叶变换和逆变换, 复杂度为 $\mathcal{O}(mn^2 \log(n))$; 其次, 还要对变换后张量的各切片做奇异值分解, 复杂度为 $\mathcal{O}(m^2 n^2)$ 。因此, 变量 \mathcal{Q} 的复杂度为 $\mathcal{O}(m^2 n^2 + mn^2 \log(n))$ 。在 U 子问题中, 需要对每个视图进行特征分解, 复杂度为 $\mathcal{O}(mn^3)$ 。优化参数 β 和 γ 的时间复杂度为 $\mathcal{O}(cn^2)$ 。由于视图数 m 远小于样本数 n ($m \ll n$), 因此本文方法进行 t 次迭代的时间复杂度为 $\mathcal{O}(tmn^3)$ 。在实验对比方法中依据其文献描述, LMSC、MSSC、PMSC 以及 FMR 的时间复杂度均为 $\mathcal{O}(tn^3)$, CoMSC 的时间复杂度为 $\mathcal{O}(tmn^3)$, 可以看出本文方法与 CoMSC 方法的时间复杂度略高于其他对比方法。

为了分析本文数值算法的收敛性, 以 BBC-Sport 数据集为例, 以迭代次数和目标函数值分别为横、纵坐标绘制收敛性曲线, 如图 5 所示。从图 5 中可以看出, 本文方法在 BBC-Sport 数据集上迭代 10 次左右即可达到收敛, 说明了本文数值算法具有高效稳定的收敛性能。

5 结束语

本文提出了一种核稳健低冗余表示学习的多视图子空间聚类方法。该方法利用多核学习捕获数据的非线性相关性信息, 并利用特征分解方法去除

数据中的冗余信息和噪声干扰获得了具有低冗余性的稳健数据表示。为了在探索多视图间的互补性信息的同时能够探索多视图间的差异性信息, 将每个视图的子空间表示矩阵重组为张量形式, 利用张量低秩约束从全局角度挖掘了视图间的高阶相关性且保持了局部视图的各异性子空间结构。将本文提出的多视图聚类方法在 3 个公开数据集上与主流的先进多视图子空间聚类方法进行对比实验, 并对本文模型的参数和收敛性进行分析, 实验结果表明了本文方法的优越性和稳健性。

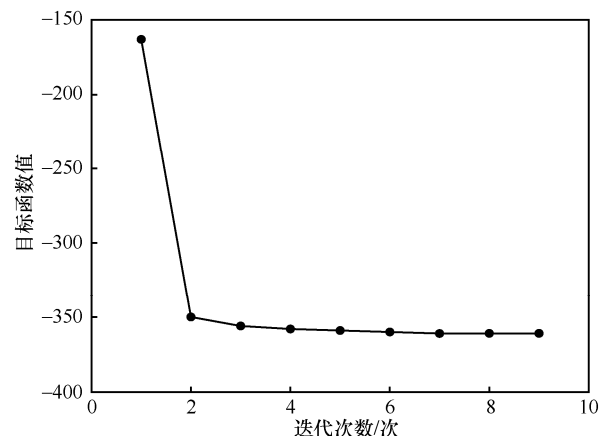


图 5 本文方法在 BBC-Sport 数据集上的收敛性曲线

参考文献:

- [1] ZHAO J, XIE X J, XU X, et al. Multi-view learning overview: recent progress and new challenges[J]. Information Fusion, 2017, 38: 43-54.
- [2] ZHANG C Q, HU Q H, FU H Z, et al. Latent multi-view subspace clustering[C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2017: 4333-4341.
- [3] ELHAMIFAR E, VIDAL R. Sparse subspace clustering: algorithm, theory, and applications[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(11): 2765-2781.
- [4] LIU G C, LIN Z C, YAN S C, et al. Robust recovery of subspace structures by low-rank representation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(1): 171-184.
- [5] FU Y F, GAO J B, TIEN D, et al. Tensor LRR and sparse coding-based subspace clustering[J]. IEEE Transactions on Neural Networks and Learning Systems, 2016, 27(10): 2120-2133.
- [6] WANG J, WANG X, TIAN F, et al. Constrained low-rank representation for robust subspace clustering[J]. IEEE Transactions on Cybernetics, 2017, 47(12): 4534-4546.
- [7] YANG J F, LIANG J, WANG K, et al. Subspace clustering via good neighbors[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(6): 1537-1544.
- [8] XU C, TAO D, XU C. A survey on multi-view learning[J]. arXiv Preprint, arXiv:1304.5634, 2013.
- [9] CHAUDHURI K, KAKADE S M, LIVESCU K, et al. Multi-view

- clustering via canonical correlation analysis[C]//Proceedings of Proceedings of the 26th Annual International Conference on Machine Learning. New York: ACM Press, 2009: 129-136.
- [10] GAO H C, NIE F P, LI X L, et al. Multi-view subspace clustering[C]//Proceedings of 2015 IEEE International Conference on Computer Vision (ICCV). Piscataway: IEEE Press, 2015: 4238-4246.
- [11] LI R H, ZHANG C Q, HU Q H, et al. Flexible multi-view representation learning for subspace clustering[C]//Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence. [S.n.:s.l.], 2019: 2916-2922.
- [12] 张苗涵, 曹容玮, 李晨, 等. 隐式低秩稀疏表示的多视角子空间聚类[J]. 模式识别与人工智能, 2020, 33(4): 344-352.
- ZHANG Z H, CAO R W, LI C, et al. Latent low-rank sparse multi-view subspace clustering[J]. Pattern Recognition and Artificial Intelligence, 2020, 33(4): 344-352.
- [13] ZHANG C Q, FU H Z, LIU S, et al. Low-rank tensor constrained multiview subspace clustering[C]//Proceedings of 2015 IEEE International Conference on Computer Vision (ICCV). Piscataway: IEEE Press, 2015: 1582-1590.
- [14] ABAVISANI M, PATEL V M. Multimodal sparse and low-rank subspace clustering[J]. Information Fusion, 2018, 39: 168-177.
- [15] KANG Z, ZHAO X J, PENG C, et al. Partition level multiview subspace clustering[J]. Neural Networks, 2020, 122: 279-288.
- [16] LIU J Y, LIU X W, YANG Y X, et al. Multiview subspace clustering via co-training robust data representation[J]. IEEE Transactions on Neural Networks and Learning Systems, 2021, PP(99): 1-13.
- [17] CAO X C, ZHANG C Q, FU H Z, et al. Diversity-induced Multi-view subspace clustering[C]//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2015: 586-594.
- [18] SCHÖLKOPF B, SMOLA A, MÜLLER K R. Nonlinear component analysis as a kernel eigenvalue problem[J]. Neural Computation, 1998, 10(5): 1299-1319.
- [19] BOYD S. Distributed optimization and statistical learning via the alternating direction method of multipliers[J]. Foundations and

Trends® in Machine Learning, 2010, 3(1): 1-122.

- [20] KILMER M E, BRAMAN K, HAO N, et al. Third-order tensors as operators on matrices: a theoretical and computational framework with applications in imaging[J]. SIAM Journal on Matrix Analysis and Applications, 2013, 34(1): 148-172.

[作者简介]



李骞(1986-), 男, 黑龙江哈尔滨人, 博士, 哈尔滨理工大学副教授、博士生导师, 主要研究方向为模式识别、机器学习等。

王卓(1997-), 女, 黑龙江牡丹江人, 哈尔滨理工大学硕士生, 主要研究方向为数据挖掘、机器学习。

于晓洋(1962-), 男, 黑龙江哈尔滨人, 博士, 哈尔滨理工大学教授、博士生导师, 主要研究方向为三维视觉检测、图像处理等。

陈德运(1962-), 男, 黑龙江哈尔滨人, 博士, 哈尔滨理工大学教授、博士生导师, 主要研究方向为探测与成像技术、图像处理等。

张英涛(1975-), 女, 黑龙江哈尔滨人, 博士, 哈尔滨工业大学副教授, 主要研究方向为人工智能与信息处理等。

孙广路(1979-), 男, 黑龙江哈尔滨人, 博士, 哈尔滨理工大学教授、博士生导师, 主要研究方向为机器学习、网络安全等。