

基于混合阶相似性的多视图聚类：一个广义的视角

陈曼笙¹⁾ 任骊安¹⁾ 王昌栋^{1),2)} 黄 栋³⁾ 赖剑煌¹⁾

¹⁾(中山大学计算机学院 广州 510006)

²⁾(广东省知识产权大数据重点实验室 广州 510006)

³⁾(华南农业大学数学与信息学院 广州 510642)

摘 要 多视图聚类已经被广泛研究,它能够采用可用的多源信息来实现更好的聚类性能.然而,大多数之前的工作仍存在两个不足:(1)它们通常关注多视图属性特征的场景,很少注意到多视图属性图数据;(2)它们主要尝试发现一致的结构或多个视图之间的关系,而忽略了多视图观测之间潜在的高阶相关性.为了解决这些问题,我们从广义角度出发,提出了一种新颖的方法,称为混合阶相似性的多视图聚类(Multiview Clustering by Hybrid-order Affinity, MCHA).它将结构图和多视图属性特征巧妙融合,同时考虑了低秩概率相似性图和混合阶的相关性.具体而言,我们通过图过滤策略构建了一组保留几何结构的视图特定的平滑表示.同时,我们将从平滑表示中学习得到的多视图概率相似性图堆叠成一个张量,并对该张量给予低秩属性的约束.这可以很好地恢复视图间更高阶的相关性.在八个基准数据集上的实验表明,我们所提出的 MCHA 方法具有最先进的有效性.

关键词 多视图聚类; 概率相似性图; 低秩张量; 高阶相关性

中图法分类号 TP18 **DOI号** 10.11897/SP.J.1016.2024.01453

Multiview Clustering by Hybrid-Order Affinity: A Generalized Perspective

CHEN Man-Sheng¹⁾ REN Li-An¹⁾ WANG Chang-Dong^{1),2)} HUANG Dong³⁾ LAI Jian-Huang¹⁾

¹⁾(School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou 510006)

²⁾(Guangdong Provincial Key Laboratory of Intellectual Property and Big Data, Guangzhou 510006)

³⁾(College of Mathematics and Informatics, South China Agricultural University, Guangzhou 510642)

Abstract Multiview clustering capable of adopting the available multisource information is extensively studied to achieve better clustering performance. However, most previous literatures still suffer from two limitations. (1) They often concentrate on the scenario of multiview attributes, paying little attention to the multiview attributed graph data. (2) They mainly attempt to discover a consensus structure or the relationships between multiple views, mostly neglecting the underlying higher-order correlations between multiview observations. To tackle these problems, we propose a novel method called Multiview Clustering by Hybrid-order Affinity (MCHA) from a generalized perspective, where the structural graph and the multiview attributes are seamlessly fused, and the low-rank probability affinity graphs with hybrid-order correlations are simultaneously considered. Specifically, a set of view-specific smooth representations preserving the geometrical structure is constructed by means of the graph filtering strategy. Meanwhile, we stack multiview probability affinity graphs learned from smooth representations

收稿日期:2023-07-30;在线发布日期:2024-04-16. 本工作由中国国家重点研发计划(2021YFF1201202),国家自然科学基金(62276277, 61976097)和广东省知识产权大数据重点实验室(2018B030322016)资助. 陈曼笙, 博士研究生, 主要研究领域为数据挖掘、多视图聚类. E-mail: chenmsh27@mail2.sysu.edu.cn. 任骊安, 本科生, 主要研究领域为数据挖掘、数据聚类. 王昌栋(通信作者), 博士, 副教授, 主要研究领域为数据聚类、网络分析、推荐算法和大数据信息安全. E-mail: wangchd3@mail.sysu.edu.cn. 黄 栋, 博士, 副教授, 主要研究领域为人工智能、大数据分析、深度学习、图神经网络. 赖剑煌, 博士, 教授, 主要研究领域为生物特征识别、数字图像处理、模式识别和机器学习.

into a tensor constrained by low-rank property, so that the higher-order correlations can be well recovered. Experiments on eight benchmark datasets indicate the state-of-the-art effectiveness of the proposed MCHA method.

Keywords multiview clustering; probability affinity graph; low-rank tensor; higher-order correlation

1 引 言

多视图数据普遍存在于许多现实世界中的应用中,它们能提供比单视图数据更加具有区分性和互补性的信息^[1-10].例如,网页数据表达可以通过上面的图片信息和文本信息来反映,其中这里的图片信息和文本信息从不同的特征角度(视图)描述了网页数据.国际新闻内容会被翻译为多种语言版本,其中不同种语言版本的内容描述了该新闻的不同视图信息.如今,我们身边有海量来自因特网和社交媒体复杂网络的数据^[11-13],它们自然地以图这种典型的非欧几里得数据结构表现出来^[14].例如在学术论文网络中,一个图视图可以描述论文之间的同

主题关系,其中的边关系表示两篇论文属于同一研究领域,而另一个图视图可以表示论文之间的合著关系,其中的边关系表示两篇论文由同一作者发表.此外,论文本身是带有属性信息的,包括代表性的词、研究领域和引文索引等文本信息,这类数据通常被称为多视图属性图数据.为了清晰起见,我们在图1中对上述数据进行了具体的图例展示.对于多视图属性数据而言,以新闻数据集为例,不同语言版本的新闻属性信息从多个视图描述了新闻数据.对于多视图属性图数据,以论文数据为例,不同论文数据之间的主题关系(是否同一个主题)或者作者关系(是否同一个作者)从不同的图视图刻画了论文数据.特别地,如何充分利用来自多个(图)视图的可用信息仍然是一个有待研究的科学问题.

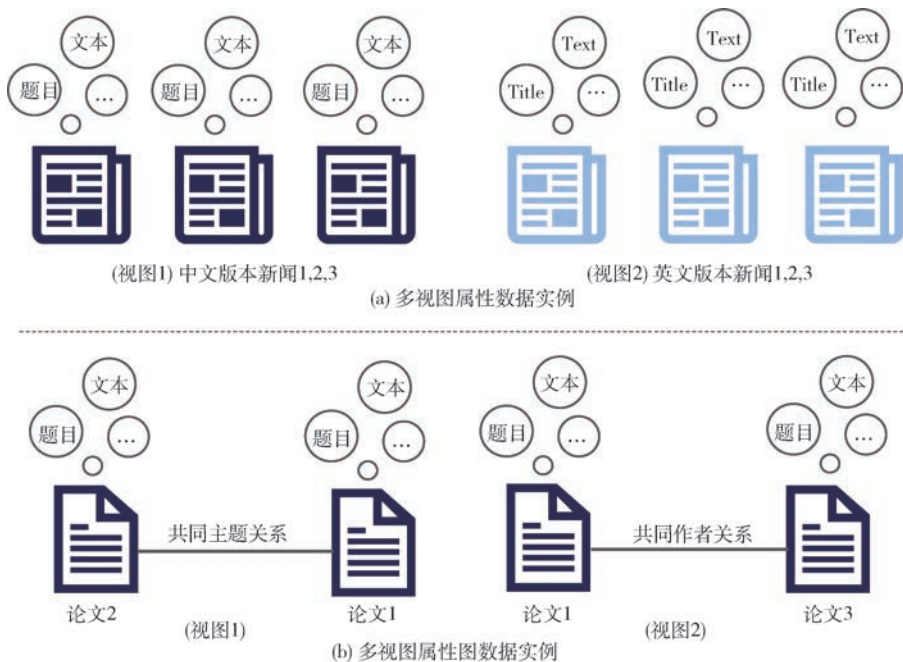


图1 多视图属性数据及多视图属性图数据的图例

多视图数据的涌现促进了多视图学习的发展.多视图学习可以结合来自不同视图的观测结果,并利用不同数据样本之间的潜在关系捕获潜在的语义线索^[15].特别地,在本工作中我们关注多视图聚

类,其目标是在没有标签信息指导的情况下,有效整合数据所有视图的不同信息,并利用它们之间的相关性去获得更鲁棒稳定的共识聚类结果.较单视图聚类而言,多视图聚类可以通过利用不同视图之

间的互补性和一致性来提高聚类的质量,更好地揭示数据的内在结构和模式.近年来,研究人员为发掘先进的多视图聚类方法做出了许多努力^[16-23].文献^[16]中的工作试图在不同的视图之间寻找类间的一致性.Zhan 等人在文献^[24]中为多视图学习设计了一种能够从多个初始图中学习全局共识图的图学习方法.Li 等人^[18]采用了深度自编码器来重构多个来自不同视图的表征,并且通过对抗训练来解耦隐空间.对于多视图核子空间聚类,文献^[25]中的工作将多个视图纳入考量,以此发现了一个共识子空间.为了捕捉不同观测结果之间的潜在语义联系,研究人员还尝试设计了一些被约束张量所约束的多视图聚类模型.文献^[26]中的工作提出了一种基于张量奇异值分解的多视图子空间聚类方法,通过张量多秩来约束由多个子空间表示堆叠而成的旋转张量.更进一步,Gao 等人通过正则化多个具有不同置信度的奇异值,提出了一种基于加权张量奇异值分解的多视图子空间聚类算法^[27].尽管以上方法已经取得了不错的效果,但绝大部分现有的多视图聚类方法仅仅考虑了多视图属性数据,而对于

多视图的属性图数据关注甚少.

有鉴于此,一些基于属性图的多视图聚类方法被设计出来^[28-30].文献^[28]中的工作尝试了基于图的自编码网络,并利用一个信息丰富的图视图和属性数据来提取节点表征,从而实现多个图视图的重构.Cheng 等人在文献^[29]中采用了多个图自编码器来学习多个表征,并且进行了一种视图间的共识学习来得到一个共同聚类.尽管这些方法效果显著,但是它们主要关注视图间的共识表示或共同关系,大多忽略了多视图观测结果之间的潜在高阶相关性.为了清晰起见,我们在图 2 中展示了传统方法的大体框架,其中对两种多视图观测结果(即多视图属性数据与多视图属性图数据)进行了展示,并对常见的局限性进行了总结.具体而言,现有的方法主要关注共识表征或共同相关性,大多忽略了不同视图观测结果之间的隐藏的高阶相关性.在实际应用中,一般存在多视图属性和由多个图数据及具有多视图属性的图结构化数据组成的多视图属性图数据,然而现有的方法很少能够在统一的模型中联合考虑这些不同类型的数据.

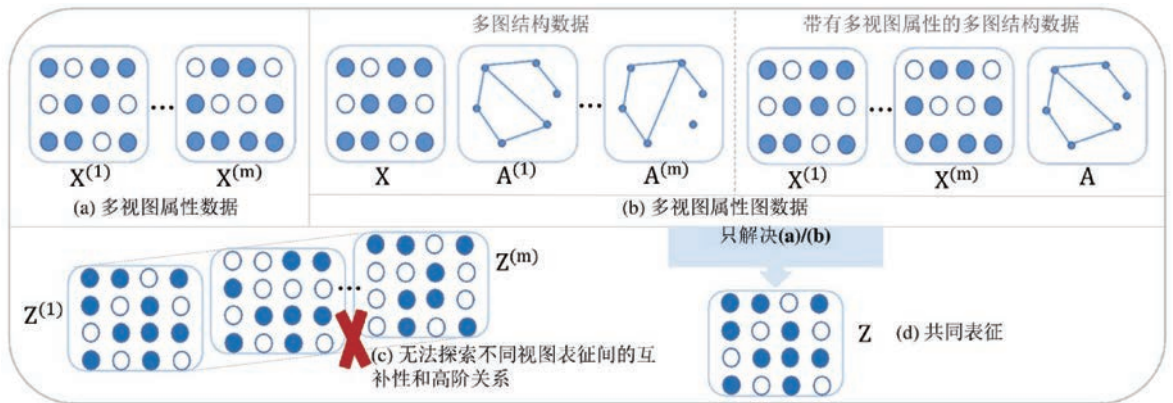


图 2 传统方法的大致框架

针对上述挑战,本篇工作提出了一种从广义视角出发的多视图聚类方法,我们将其命名为 MCHA,基于混合阶相似度的多视图聚类.我们提出的算法能够巧妙地融合结构图和多视图属性,同时考虑了具有混合阶相关性的低秩概率相似性图表示.具体来说,为了保持几何结构,我们首先通过图过滤策略构建了一组针对特定视图的平滑表示.然后,基于这些平滑的表示,我们可以利用自表达特性来学习到记录样本间二阶相关性的针对特定视图的概率相似性图.同时我们将多个视图的概率相似性图堆叠为基于张量奇异值分解的带权张量,由

此可以很好地捕捉到跨视图间的高阶相关性与来自不同奇异值的显著语义信息.因而,我们不仅能捕获到二阶相关性,而且能从多视图概率相似性图中捕获到更高阶的相关性,并用于更深入的研究中.

本文的主要贡献如下:

(1)提出了一种新的通过图过滤巧妙融合结构图和多视图属性的广义多视图聚类模型.

(2)利用自表达特性和低秩张量约束来联合恢复多视图概率相似性图中的混合阶相关性,设计了一种交替最小化优化算法来求解所提出的 MCHA 方法.

(3)通过在 8 个基准数据集上进行实验,验证了所提出的 MCHA 方法具有十分先进的有效性。

本文余下部分的构成如下:我们在第 2 节中简要地介绍了一些近来有关多视图聚类的相关工作;第 3 节中给出了整篇文章使用的主要符号;第 4 节详细阐述了我们所提出的 MCHA 方法,同时介绍了优化算法和时间复杂性;第 5 节展示了在八个基准数据集中的实验结果;最后在第 6 节中我们总结了本篇文章的内容,并提出了潜在的未来研究方向。

2 相关工作

在过去的几年中,通过探索由多视图属性或多视图属性图表示的数据样本之间的关系,研究者对许多多视图聚类方法进行了广泛的研究。

针对多视图属性数据,我们粗略地将现有的多视图聚类方法分为三个主要类别,即协同训练或协同正则化风格的算法^[31]、基于图的方法^[32-35]以及基于子空间学习的方法^[25,27,36-39]。在协同训练或协同正则化风格的算法中,跨视图的共识通过交替训练的方法被最大化。文献[40]中的工作使用希尔伯特施密特独立性准则的方法研究跨不同视图间的互补信息。类似地,通过希尔伯特施密特独立性准则方法,Wang 等人利用相似性使潜在的完整样本之间的依赖度最大化^[41]。在基于图的方法中,多视图信息被用于发掘跨多个视图间的结构关系。Xia 等人在文献[32]中提出了一种鲁棒的多视图谱聚类方法(Robust Multiview Spectral Clustering, RMSC),通过低秩和稀疏分解发现了一个共享的低秩转移概率矩阵。Wang 等人共同研究了针对特定视图的相似图和共识聚类分配矩阵^[35]。Nie 等人在文献[42]中提出了一种新的多视图模型,同时考虑了聚类指标和局部结构。基于子空间学习的模型则基于所有视图的观测结果来自同一个隐藏子空间的假设,将寻找共享的隐藏子空间作为目标。Zhang 等人在文献[38]中提出了一种泛化的深度框架,用于学习视图之间共享的隐藏表示。为了探索视图之间隐藏的语义联系性,研究者们也尝试设计了一些受低秩张量约束的多视图子空间聚类模型。Wu 等人在文献[43]中利用针对特定视图的转移概率表征来构建张量。Gao 等人在文献[27]中则探索了基于张量奇异值分解的带权张量核正则最小化方法(t-SVD based Weighted Tensor Nuclear Norm Minimization, WTNNM),其中不同的矩阵奇异值以不同的置信

度进行收缩。Chen 等人在文献[44]中提出了在一个联合的框架中学习低秩表征张量和相似性矩阵的方法,既保持局部结构,也能保持多视图属性对于最终的相似性矩阵的不同重要性。Xu 等人在文献[45]中设计了针对对比多视图聚类的多层级特征学习,其中考虑了学习共识语义和重构视图独有信息间的潜在冲突。

针对多视图属性图数据,研究者也对如何从属性和图中综合集成信息做出了一些尝试^[28-30,46-48]。文献[30]中的工作提出了一种能够同时结合来自多个给定图的表征的多重属性网络。文献[46]中的工作则研究了针对特定视图的平滑节点表征和共识图。Lin 等人在文献[47]中提出在多视图属性图聚类中可以通过相邻图的累乘来提取高阶邻域信息。文献[48]中的工作则提出了适应性的图过滤策略和锚采样机制。Wang 等人在文献[49]中利用注意力机制为异质的图设计了一种图神经网络。Pan 等人设计了一种多视图对比图聚类方法,通过整合来自多个图和属性的可用信息来学习出一个共识图^[50]。Lin 等人在文献[51]中通过互信息最大化来增强不同多视图属性图之间的对偶信息,以便得到更好的聚类结果。

在本文中,我们提出了一种从广义视角出发的多视图聚类方法,基于混合阶相似度的多视图聚类(MCHA),其中巧妙地融合结构图和多视图属性,同时考虑了具有混合阶相关性的低秩概率相似性图表示。不同于现有的方法,MCHA 能够很好地处理多视图属性数据和多视图属性图数据,并考虑了数据其中的混合阶相关性。

3 符号和术语

在这部分中,我们介绍本文主要使用的符号。花体字母(例如 \mathcal{B})和大写字母(例如 \mathbf{B})分别用于表示张量和矩阵。给定一个三维的张量 $\mathcal{B} \in R^{n_1 \times n_2 \times n_3}$,符号 $\mathcal{B}(j, :, :)$, $\mathcal{B}(:, j, :)$ 和 $\mathcal{B}(:, :, j)$ 分别用于表示第 j 个纵向、横向和正向的三个切片。特别地,我们用 $\mathcal{B}^{(j)}$ 来表示 $\mathcal{B}(:, :, j)$ 。 \mathcal{B} 的快速傅里叶变换用 $\bar{\mathcal{B}} = \text{fft}(\mathcal{B}, [], 3)$ 来表示。类似地, \mathcal{B} 可以由逆傅里叶变换得到。更详细的信息可以参阅文献[49]。此外,我们在表 1 中总结了文章使用的基础符号。

一个多视图数据集可以大致定义为 $G = \{V, \mathbf{H}^{(1)}, \dots, \mathbf{H}^{(m)}, \mathbf{X}^{(1)}, \dots, \mathbf{X}^{(m)}\}$,其中 V 是含 n 个节点的节点集合, $H_{ij}^{(v)} \in \mathbf{H}^{(v)}$ 表示 v 个边集合中第 i

个和第 j 个节点的关系, $\mathbf{X}^{(v)} = [\mathbf{x}_1^{(v)}, \dots, \mathbf{x}_n^{(v)}] \in \mathbb{R}^{n \times d^v}$ 则是 d^v 维的第 v 个视图的特征空间.

邻接矩阵 $\{\mathbf{A}^{(v)}\}_{v=1}^m$ 是利用边集合来刻画的初始拓扑结构, 当 $H_{ij}^{(v)} \in \mathbf{H}^{(v)}$ 时 $a_{ij}^{(v)} = 1$, 否则 $a_{ij}^{(v)} = 0$, $\{\mathbf{D}^{(v)}\}_{v=1}^m$ 代表对应邻接图的度矩阵. 对称归一化的图拉普拉斯算子可以定义为 $\mathbf{L}^{(v)} = \mathbf{I} - (\mathbf{D}^{(v)})^{-\frac{1}{2}}(\mathbf{A}^{(v)} + \mathbf{I})(\mathbf{D}^{(v)})^{-\frac{1}{2}}$, 其中 \mathbf{I} 是单位矩阵.

表 1 基础符号和说明

符号	含义
V	包含 n 个节点的节点集合
$\mathbf{H}^{(1)}, \dots, \mathbf{H}^{(m)}$	m 个视图的边集合
$\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(m)}$	m 个视图的特征空间
$\mathbf{X}^{(v)} \in \mathbb{R}^{n \times d^v}$	第 v 个视图的特征空间
$\mathbf{A}^{(v)} \in \mathbb{R}^{n \times n}$	第 v 个视图的邻接矩阵
$\mathbf{D}^{(v)} \in \mathbb{R}^{n \times n}$	第 v 个视图的度矩阵
$\mathbf{L}^{(v)} \in \mathbb{R}^{n \times n}$	第 v 个视图的拉普拉斯矩阵
$\mathbf{M}^{(v)} \in \mathbb{R}^{n \times d^v}$	平滑化的第 v 个视图特征空间
$\mathbf{Z}^{(v)} \in \mathbb{R}^{n \times n}$	第 v 个视图的相似性图表示
$\mathbb{Z} \in \mathbb{R}^{n \times n \times m}$	多个相似性图表示构造的张量
$\mathbf{Z}_c \in \mathbb{R}^{n \times c}$	共识的相似性图表示
$\mathbf{E}^{(v)} \in \mathbb{R}^{n \times d^v}$	第 v 个视图的重构误差矩阵
$\mathbf{I} \in \mathbb{R}^{n \times n}$	单位矩阵
λ	权衡参数
$\omega \in \mathbb{R}^{1 \times m}$	加权张量核范数的权重因子
d_v	第 v 个视图的特征维度
m	视图个数
n	样本量大小
c	簇个数

4 方 法

4.1 模型表达

多视图聚类一般试图通过充分利用不同的视图中的异质特征信息来探索数据之间的连接性和相似性. 尽管现有的方法已经取得了显著的成果, 但它们大部分仅考虑了多视图的属性信息, 而没有对多视图属性图数据进行利用, 也没有发掘多视图观测结果之间的综合性和潜在的高阶相关性. 在本篇工作中, 我们从广泛的视角出发, 提出了一种新的基于混合阶相似性的多视图聚类方法 MCHA, 用于探索由多视图属性融合而成的平滑表征和图结构数据之间的潜在相关性.

图过滤. 首先我们在单视图的场景中讨论该模型. 给定一个表征 $\mathbf{X} \in \mathbb{R}^{n \times d}$, \mathbf{X} 的每一列可以看作一个 n 维的图信号. 现实世界中的信号通常在邻接

矩阵中相邻的数据样本之间是平滑的. 换句话说, 邻接的节点应当有相似的特征值. 特别地, 我们可以通过拉普拉斯-贝尔特拉米算子来衡量图信号的平滑性, 频率较低或者特征值较小的信号被证明是较为平滑的. 具体来说, 我们可以通过以下模型来获取平滑化的信号:

$$\min_{\mathbf{M}} \|\mathbf{M} - \mathbf{X}\|_F^2 + \gamma \text{Tr}(\mathbf{M}^T \mathbf{L} \mathbf{M}) \quad (1)$$

其中 $\mathbf{M} \in \mathbb{R}^{n \times d}$ 表示平滑化的信号, \mathbf{L} 表示输入数据的拉普拉斯矩阵, $\gamma > 0$ 是权衡因子. 第一项被视为保真项, 最后的图拉普拉斯正则项则被用来保证探索平滑信号的同时保持新学习的表征中的原生局部性.

通过将上面的函数对 \mathbf{M} 求导并使其为零, 我们可以得到:

$$\mathbf{M} = (\mathbf{I} + \gamma \mathbf{L})^{-1} \mathbf{X} \quad (2)$$

为了避免时间复杂度为 $O(n^3)$ 的矩阵求逆计算, 我们使用一阶泰勒展开来近似取得 \mathbf{M} . 换句话说, 我们用 $\mathbf{M} = (\mathbf{I} - \gamma \mathbf{L}) \mathbf{X}$ 来表示 \mathbf{M} . 此外我们可以用下式来重新表示第 k 阶的图过滤^[47],

$$\mathbf{M} = (\mathbf{I} - \gamma \mathbf{L})^k \mathbf{X} \quad (3)$$

其中 k 代表一个正整数. 利用第 k 阶的图过滤, 我们可以通过聚合 k 阶邻居的特征来捕获更深层的图结构信息, 使得相近的节点拥有相似的特征值. 图过滤可以将图结构信息编码进保有图几何结构的特征中. 换句话说, 图过滤是一种优雅地使用图结构来增强特征的策略.

对于多视图数据, 针对特定视图的平滑表征可以相应地被重新表示为

$$\mathbf{M}^{(v)} = (\mathbf{I} - \gamma \mathbf{L}^{(v)})^k \mathbf{X}^{(v)}, \forall v = \{1, \dots, m\}. \quad (4)$$

由于现实世界中的图不可避免地存在噪声和缺失, 平滑操作得到的表征并不会直接地用于下游聚类任务. 特别地, 本篇文章中有关平滑表征学习的图过滤过程被用于更好地处理原始多视图数据.

高阶相似性的探索. 受子空间学习的自表示特性的启发, 我们可以从平滑表征中获取可靠的、能够表征数据簇结构的相似性表征, 即

$$\begin{aligned} \min_{\mathbf{Z}^{(v)}, \mathbf{E}^{(v)}} \sum_{v=1}^m \lambda L(\mathbf{M}^{(v)T}, \mathbf{M}^{(v)T} \mathbf{Z}^{(v)}) + \Omega(\mathbf{Z}^{(v)}), \\ \text{s. t. } \mathbf{M}^{(v)T} = \mathbf{M}^{(v)T} \mathbf{Z}^{(v)} + \mathbf{E}^{(v)}, v = \{1, \dots, m\} \end{aligned} \quad (5)$$

其中 $\mathbf{Z}^{(v)}$ 是第 v 个视图的相似性图表示, $\mathbf{E}^{(v)}$ 表示相应的重构误差矩阵. $L(\cdot, \cdot)$ 是损失函数, $\Omega(\cdot)$ 表示对 $\mathbf{Z}^{(v)}$ 的正则化项, $\lambda > 0$ 则是权衡因子. 通常地, 我们可以使用弗罗贝尼乌斯范数、核范数或稀

疏 l_1 范数作为正则化器。在本篇文章中我们采用一种带权的张量核范数来捕获多个相似性图表示中的高阶相关性, 目标函数如下所示:

$$\begin{aligned} \min_{\mathbf{Z}^{(v)}, \mathbf{E}^{(v)}, \mathbf{Z}} \lambda \sum_{v=1}^m \|\mathbf{E}^{(v)}\|_{2,1} + \|\mathbf{Z}\|_{\omega, \Theta}, \\ \text{s. t. } \mathbf{M}^{(v)\top} = \mathbf{M}^{(v)\top} \mathbf{Z}^{(v)} + \mathbf{E}^{(v)}, v = \{1, \dots, m\}, \\ \mathbf{Z} = \Phi(\mathbf{Z}^{(1)}, \mathbf{Z}^{(2)}, \dots, \mathbf{Z}^{(m)}) \end{aligned} \quad (6)$$

其中 $\|\cdot\|_{\omega, \Theta}$ 表示带权的张量核范数, ω 是一个权重因子, $\Phi(\cdot)$ 表示通过多个 $\mathbf{Z}^{(v)}$ 来构造一个张量的函数, 其构造过程在图 3 中展示. $l_{2,1}$ 范数促使 $\mathbf{E}^{(v)}$ 中的每一行趋近零. 利用带权的张量核范数, 我们可以使用不同的权重因子通过软阈值函数来正则化 \mathbf{Z} 的奇异值.

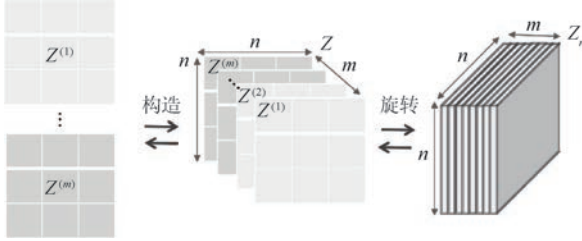


图 3 张量 \mathbf{Z} 的构造和旋转

总体目标函数. 为了使得每个单独的 $\mathbf{E}^{(v)}$ 所包含的行值具有一致的量级, 我们将不同视图对应的误差列垂直拼接起来, 由此总体的目标函数最终表示为:

$$\begin{aligned} \min_{\mathbf{Z}^{(v)}, \mathbf{E}^{(v)}, \mathbf{Z}} \lambda \|\mathbf{E}\|_{2,1} + \|\mathbf{Z}\|_{\omega, \Theta}, \\ \text{s. t. } \mathbf{M}^{(v)\top} = \mathbf{M}^{(v)\top} \mathbf{Z}^{(v)} + \mathbf{E}^{(v)}, v = \{1, \dots, m\}, \\ \mathbf{Z} = \Phi(\mathbf{Z}^{(1)}, \mathbf{Z}^{(2)}, \dots, \mathbf{Z}^{(m)}), \\ \mathbf{E} = [\mathbf{E}^{(1)}, \mathbf{E}^{(2)}, \dots, \mathbf{E}^{(m)}] \end{aligned} \quad (7)$$

特别地, 在进一步的张量计算之前, 我们需要将 \mathbf{Z} 旋转, 以此保证不同视图之间的低秩性能能够被捕获, 并且大大降低计算的复杂度, 也即是将其维度相应地从 $n \times n \times m$ 旋转为 $n \times m \times n$. 值得一提的是, 虽然旋转前后 \mathbf{Z} 的维度没有减少, 但 \mathbf{Z} 上的加权张量核范数项求解时需要进行奇异值分解, 其中旋转前后的张量所需要的时间复杂度会有很大的不同. 具体地, 旋转前的 \mathbf{Z} 奇异值分解所需复杂度是 $O(n^3 m)$, 而旋转后的 \mathbf{Z} 所需复杂度是 $O(n^2 m^2)$, 其中 $n \gg m$. 为了清楚起见, 张量旋转的过程在图 3 中展示.

一旦我们得到了相似性图表征 $\mathbf{Z}^{(v)}$, 可以通过求多个 $\mathbf{Z}^{(v)}$ 的平均值来计算出一个共识的 \mathbf{Z}_c , 即

$$\mathbf{Z}_c = \frac{1}{m} \sum_{v=1}^m \mathbf{Z}^{(v)*} \quad (8)$$

为了得到最终的聚类结果, 我们在这个共识表征 \mathbf{Z}_c 上执行谱聚类.

值得注意的是: (1) 利用公式(7)中低秩张量正则项, 我们可以很好地复原不同的表征 $\mathbf{Z}^{(v)} \in \mathbb{R}^{n \times n}$ 之间的高阶相关性. 特别地, 与矩阵(二维张量)相比, 三维张量含有更高维的性质, 换句话说, 我们可以用这样的低秩张量来复原高阶相关性. 因此, 我们所提出的模型能够很好的复原不同表征之间隐藏的二阶和三阶相关性. (2) 低阶(二阶)相似性指的是仅通过子空间学习的自表示特性所得到的相似性表征, 也即是 $\mathbf{Z}^{(v)}$; 高阶(三阶)相似性指的是通过将多个低阶相似性表征堆叠成一个高阶张量 \mathbf{Z} , 并赋予张量核范数约束以便更好地捕获到张量内部不同相似性表征之间的高阶相关性; 混合阶相似性则代表了在统一的模型中同时考虑低阶相似性和高阶相似性. 特别地, 在该模型的设计中, 由于高阶相似性依赖于低阶相似性, 因而无法单独对高阶相似性进行计算. (3) 与现有的一般仅关注多视图属性的多视图聚类方法相比, MCHA 既能够处理多视图属性数据, 也能够处理多视图属性图数据. 当输入多视图属性图数据时, 特征观测值和现有的图结构信息可以被图过滤方法巧妙地融合成平滑的表征, 然后被用于后续步骤. 此外, 当输入多视图属性(即 $\mathbf{X} = \{\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(m)}\}$)时, 我们的方法首先对于每个视图的观测结果构建相应的邻接图, 然后基于隐藏的几何结构使用图过滤方法来增强观测值, 这更有利于后续的学习任务. 此外, 我们进行了大量的实验以评估 MHCA 相对于现有最先进聚类算法的优越性.

4.2 优化策略

为了解决公式(7)中的优化问题, 我们采用了增广拉格朗日乘子的交替最小化策略, 在更新一个变量时固定其他的变量. 首先我们使用了变量分割方法使得 \mathbf{Z} 在(7)中可分离, 然后用辅助变量 $\mathbf{Q} \in \mathbb{R}^{n \times n \times m}$ 代替 \mathbf{Z} , 其增广拉格朗日函数可以重写为

$$\begin{aligned} O(\mathbf{Z}^{(1)}, \mathbf{Z}^{(2)}, \dots, \mathbf{Z}^{(m)}; \mathbf{E}^{(1)}, \mathbf{E}^{(2)}, \dots, \mathbf{E}^{(m)}; \mathbf{Q}) \\ = \lambda \|\mathbf{E}\|_{2,1} + \|\mathbf{Q}\|_{\omega, \Theta} \\ + \langle \mathbb{W}, \mathbf{Z} - \mathbf{Q} \rangle + \frac{\rho}{2} \|\mathbf{Z} - \mathbf{Q}\|_F^2 \\ + \sum_{v=1}^m \langle \mathbf{Y}^{(v)}, \mathbf{M}^{(v)\top} - \mathbf{M}^{(v)\top} \mathbf{Z}^{(v)} - \mathbf{E}^{(v)} \rangle \\ + \sum_{v=1}^m \frac{\mu}{2} \|\mathbf{M}^{(v)\top} - \mathbf{M}^{(v)\top} \mathbf{Z}^{(v)} - \mathbf{E}^{(v)}\|_F^2 \end{aligned} \quad (9)$$

其中 \mathbb{W} 和 $\mathbf{Y}^{(v)}$ 是两个拉格朗日乘子, $\rho > 0$ 和 $\mu > 0$ 是惩罚因子. 特别地, 通过最小化策略交替更新 $\mathbf{Z}^{(v)}$, $\mathbf{E}^{(v)}$ 和 \mathbb{Q} , 我们将公式(9) 中的问题分割为三个子问题.

$\mathbf{Z}^{(v)}$ -子问题: 固定 \mathbf{E} 和 \mathbb{Q} , 这等价于在如下问题中更新 $\mathbf{Z}^{(v)}$:

$$\begin{aligned} & \min_{\mathbf{Z}^{(v)}} \sum_{v=1}^m \langle \mathbf{Y}^{(v)}, \mathbf{M}^{(v)\top} - \mathbf{M}^{(v)\top} \mathbf{Z}^{(v)} - \mathbf{E}^{(v)} \rangle \\ & + \sum_{v=1}^m \frac{\mu}{2} \|\mathbf{M}^{(v)\top} - \mathbf{M}^{(v)\top} \mathbf{Z}^{(v)} - \mathbf{E}^{(v)}\|_F^2 \\ & + \sum_{v=1}^m (\langle \mathbf{W}^{(v)}, \mathbf{Z}^{(v)} - \mathbb{Q}^{(v)} \rangle + \frac{\rho}{2} \|\mathbf{Z}^{(v)} - \mathbb{Q}^{(v)}\|_F^2) \end{aligned} \quad (10)$$

相应地, 上面的问题可以针对每个视图分别重写为

$$\begin{aligned} & \min_{\mathbf{Z}^{(v)}} \langle \mathbf{Y}^{(v)}, \mathbf{M}^{(v)\top} - \mathbf{M}^{(v)\top} \mathbf{Z}^{(v)} - \mathbf{E}^{(v)} \rangle \\ & + \frac{\mu}{2} \|\mathbf{M}^{(v)\top} - \mathbf{M}^{(v)\top} \mathbf{Z}^{(v)} - \mathbf{E}^{(v)}\|_F^2 \\ & + \langle \mathbf{W}^{(v)}, \mathbf{Z}^{(v)} - \mathbb{Q}^{(v)} \rangle + \frac{\rho}{2} \|\mathbf{Z}^{(v)} - \mathbb{Q}^{(v)}\|_F^2 \end{aligned} \quad (11)$$

然后我们可以通过将公式(11)对 $\mathbf{Z}^{(v)}$ 求导并置导数为零来得到最优解 $\mathbf{Z}^{(v)*}$, 即

$$\begin{aligned} \mathbf{Z}^{(v)*} & = (\mathbf{I} + \frac{\mu}{\rho} \mathbf{M}^{(v)} \mathbf{M}^{(v)\top})^{-1} ((\mathbf{M}^{(v)} \mathbf{Y}^{(v)} \\ & + \mu \mathbf{M}^{(v)} \mathbf{M}^{(v)\top} - \mu \mathbf{M}^{(v)} \mathbf{E}^{(v)} - \mathbf{W}^{(v)}) / (\rho + \mathbb{Q}^{(v)}) \end{aligned} \quad (12)$$

$\mathbf{E}^{(v)}$ -子问题: 固定 $\mathbf{Z}^{(v)}$ 和 \mathbb{Q} , 这等价于在以下问题求 $\mathbf{E}^{(v)}$

$$\begin{aligned} & \min_{\mathbf{E}^{(v)}} \lambda \|\mathbf{E}\|_{2,1} \\ & + \sum_{v=1}^m \langle \mathbf{Y}^{(v)}, \mathbf{M}^{(v)\top} - \mathbf{M}^{(v)\top} \mathbf{Z}^{(v)} - \mathbf{E}^{(v)} \rangle \\ & + \sum_{v=1}^m \frac{\mu}{2} \|\mathbf{M}^{(v)\top} - \mathbf{M}^{(v)\top} \mathbf{Z}^{(v)} - \mathbf{E}^{(v)}\|_F^2 \\ & \Leftrightarrow \min_{\mathbf{E}^{(v)}} \frac{\lambda}{\mu} \|\mathbf{E}\|_{2,1} + \frac{1}{2} \|\mathbf{E} - \mathbf{F}\|_F^2 \end{aligned} \quad (13)$$

其中的 \mathbf{F} 是通过垂直拼接聚合矩阵 $\mathbf{M}^{(v)\top} - \mathbf{M}^{(v)\top} \mathbf{Z}^{(v)} + \frac{1}{\mu} \mathbf{Y}^{(v)}$ 构造的.

故 $\mathbf{E}^{(v)}$ 的最优解可以通过文献[52]计算得到

$$\mathbf{E}_{:,i}^* = \begin{cases} \frac{\|\mathbf{F}_{:,i}\|_2 - \frac{\lambda}{\mu}}{\|\mathbf{F}_{:,i}\|_2} \frac{\mu}{\lambda} \mathbf{F}_{:,i}, & \|\mathbf{F}_{:,i}\|_2 > \frac{\lambda}{\mu} \\ 0, & \text{otherwise} \end{cases} \quad (14)$$

其中上式中的 $\mathbf{F}_{:,i}$ 是 $\mathbf{F} = [\mathbf{F}^{(1)}; \mathbf{F}^{(2)}; \dots; \mathbf{F}^{(m)}]$ 的第 i 列, 且 $\mathbf{F}^{(v)} = \mathbf{M}^{(v)\top} - \mathbf{M}^{(v)\top} \mathbf{Z}^{(v)} + \frac{1}{\mu} \mathbf{Y}^{(v)}$, $v \in [1, m]$.

\mathbb{Q} -子问题: 固定 $\mathbf{Z}^{(v)}$ 和 $\mathbf{E}^{(v)}$, 我们可以通过解决以下问题来更新 \mathbb{Q} :

$$\min_{\mathbb{Q}} \|\mathbb{Q}\|_{\omega, \bullet} + \frac{\rho}{2} \|\mathbb{Q} - (\mathbf{Z} + \frac{1}{\rho} \mathbf{W})\|_F^2 \quad (15)$$

特别地, 我们需要引入以下理论来解决这个问题:

定理 1^[27]. 给定一个张量 $\mathbb{G} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, 其张量奇异值分解为 $\mathbb{U}_{\mathbb{G}} * \mathcal{S}_{\mathbb{G}} * \mathbb{V}_{\mathbb{G}}^{\top}$, $l = \min(n_1, n_2)$, 模型可以表示为

$$\operatorname{argmin}_{\mathbb{X}} \frac{1}{2} \|\mathbb{X} - \mathbb{G}\|_F^2 + \tau \|\mathbb{X}\|_{\omega, \bullet}. \quad (16)$$

模型的最优解可以具体地表示为

$$\mathbb{X}^* = \Gamma_{\tau * \omega}(\mathbb{G}) = \mathbb{U}_{\mathbb{G}} * \operatorname{ifft}(P_{\tau * \omega}(\mathbb{G})) * \mathbb{V}_{\mathbb{G}}^{\top}, \quad (17)$$

公式中的 $\mathbb{G} = \operatorname{fft}(\mathbb{G}, [], 3)$, $P_{\tau * \omega}(\mathbb{G})$ 表示一个张量, 其第 i 个正向切片为 $P_{\tau * \omega}(\overline{\mathbf{G}}^{(i)}) = \operatorname{diag}(\xi_1, \xi_2, \dots, \xi_l)$, 其中 ξ_j 是 $\operatorname{sign}(\sigma_j(\overline{\mathbf{G}}^{(i)})) \max(\sigma_j(\overline{\mathbf{G}}^{(i)}) - \tau * \omega_j, 0)$ 的值.

有鉴于此, 我们容易得到公式(15)的解, 即

$$\mathbb{Q}^* = \Gamma_{\frac{1}{\rho} * \omega}[\mathbf{Z} + \frac{1}{\rho} \mathbb{W}]. \quad (18)$$

$\mathbf{Y}^{(v)}$ 和 \mathbb{W} -子问题: 拉格朗日乘子 $\mathbf{Y}^{(v)}$ 和 \mathbb{W} 可以通过下式求解:

$$\mathbf{Y}^{(v)*} = \mathbf{Y}^{(v)} + \mu(\mathbf{M}^{(v)\top} - \mathbf{M}^{(v)\top} \mathbf{Z}^{(v)} - \mathbf{E}^{(v)}), \quad (19)$$

$$\mathbb{W}^* = \mathbb{W} + \rho * (\mathbf{Z} - \mathbb{Q}). \quad (20)$$

特别地, 本模型的伪代码概述在算法 1 中给出.

算法 1. 基于混合阶相似性的多视图聚类: 一个广义的视角.

输入: 多视图属性数据 $\{\mathbf{X}^{(v)}\}_{v=1}^m$, 或者多视图属性图数据 $\{\mathbf{X}^{(v)}, \mathbf{A}^{(v)}\}_{v=1}^m$, 图过滤阶数 $k > 0$, 图过滤参数 $\gamma > 0$, 权重系数 $\omega > 0$, 权衡参数 $\lambda > 0$, 聚类个数 c

输出: 通过在 Z_c 上执行谱聚类算法来得到聚类结果 \mathbb{C}

1. 初始化 $\mathbf{Z}^{(v)} = 0, \mathbf{E}^{(v)} = 0, \mathbf{Y}^{(v)} = 0, \mathbb{Q} = \mathbb{W} = 0, \mu = 10^{-5}, \rho = 10^{-4}, \eta = 2, \mu_{\max} = \rho_{\max} = 10^{10}, \epsilon = 10^{-7}$;
2. IF 只有 $\{\mathbf{X}^{(v)}\}_{v=1}^m$ then
3. 依据不同的视图观测值来构造 $\{\mathbf{A}^{(v)}\}_{v=1}^m$;
4. END IF
5. $\mathbf{L}^{(v)} = \mathbf{I} - (\mathbf{D}^{(v)})^{-\frac{1}{2}} (\mathbf{A}^{(v)} + \mathbf{I}) (\mathbf{D}^{(v)})^{-\frac{1}{2}}$;
6. 通过式子(4)的图过滤来计算;
7. WHILE 不收敛 DO
8. 通过解决式子(12)来更新 $\{\mathbf{Z}_{t+1}^{(v)}\}_{v=1}^m$;

9. 通过解决式子(14)来更新 \mathbf{E}_{t+1} ;
10. 通过解决式子(19)来更新 $\{\mathbf{Y}_{t+1}^{(v)}\}_{v=1}^m$;
11. 通过解决式子(18)来更新 \mathbf{Q}_{t+1} ;
12. 通过解决式子(20)来更新 \mathbb{W}_{t+1} ;
13. 分别通过 $\mu_{t+1} = \min(\eta\mu_t, \mu_{\max})$ 和 $\rho_{t+1} = \min(\eta\rho_t, \rho_{\max})$ 来更新 μ 和 ρ ;
14. 检查收敛条件: $\|\mathbf{M}^{(v)T} - \mathbf{M}^{(v)T} \mathbf{Z}_{t+1}^{(v)} - \mathbf{E}_{t+1}^{(v)}\|_{\infty} < \epsilon$, $\|\mathbf{Z}_{t+1}^{(v)} - \mathbf{Q}_{t+1}^{(v)}\|_{\infty} < \epsilon$;
15. END WHILE
16. 通过式子(8)来计算共识相似性矩阵 \mathbf{Z} .

4.3 时间复杂度及收敛性

模型的计算成本主要包括四个部分. 第一部分是通过图过滤预先计算 $\mathbf{M}^{(v)}$, 相当于将 $\mathbf{X}^{(v)}$ 左乘 $\mathbf{I} - \gamma\mathbf{L}^{(v)}$ k 次来求解问题, 其计算复杂度为 $O(n_i d k m)$, 其中 n_i 为稀疏邻接矩阵的非零元素数量, 而 $d = \sum_v d_v$. 第二部分是关于更新 $\mathbf{Z}^{(v)}$ 的, 由于逆运算的存在, 其计算复杂度通常为 $O(n^3)$. 但实际上, 如果 $d < n$, 则可以通过采用伍德伯里矩阵恒等式将 $\mathbf{Z}^{(v)}$ 的计算成本降低到 $O(n^2 d)$. 第三部分是计算误差矩阵 $\mathbf{E}^{(v)}$, 其计算成本为 $O(n^2 m)$. 最后一部分是关于更新 \mathbf{Q} 的, 其主要计算取决于维度为 $n \times m \times n$ 的旋转张量的 FFT 或者逆 FFT 计算, 以及傅里叶变换中 n 个 $n \times m$ 矩阵的奇异值分解. 因此, \mathbf{Q} 的计算成本为 $O(n^2 m \log(n))$. 因此, 所提出模型的总体计算复杂度为 $O(n_i d k m + (n^2 d + n^2 m + n^2 m \log(n)) t_a)$, 其中 t_a 表示交替优化算法的迭代次数, 主要的计算成本在于更新 \mathbf{Q} .

对于模型的收敛性分析, 式子(7)不是关于所有变量的联合凸问题, 如何得到对应的全局最优解是一个待研究的开放问题. 在本文中, 通过算法 1 中交替最小化优化策略, 我们很好地解决了式子(7)的优化求解问题. 由于每个变量的优化子问题是满足凸性且有最优解的, 因而算法 1 是能够理论收敛的. 进一步地, 在后续的小节 5.7 中, 我们通过数据实验也证实了 MCHA 的收敛性质.

5 实 验

本节对几个基准数据集进行了广泛的实验, 以展示所提出的 MCHA 方法与最先进的聚类方法相比的有效性. 为了进行全面的实验, 我们还进行了消融实验、参数敏感性和收敛性实验. MCHA 方法的源代码和真实世界数据集可以从 <https://pan.baidu.com/s/1pu5DAIRyaU6d1TCITEAcCA> 中下

载, 其中的提取码为 MCHA.

5.1 数据集

下面采用了包括多视图属性数据集和多视图属性图数据集在内的八个广泛使用的数据集.

(1)MSRC^[53]. 该数据集包括 210 个图像, 它们属于 7 个类别, 其中这些类别被作为基准标签, 分别是树、建筑、飞机、奶牛、人脸、汽车和自行车. 在实验中, 我们选择了四个特征视图, 分别是 CM 特征、GIST 特征、LBP 特征和 GENT 特征.

(2)UCI^[32]. 该数据集包含 2000 个手写数字图像, 它们属于 10 个类别. 在实验中, 我们使用了三种特征类型, 分别是平均强度特征、傅里叶系数特征和形态特征.

(3)COIL-20. 该数据集由 1440 个面部图像组成, 它们属于 20 个类别^[54]. 在实验中, 我们选择了关于强度、LBP 和 Gabor 的三种特征类型.

(4)Scene-15^[55]. 该数据集由来自室内和室外环境的 4485 张图像组成, 它们属于 15 个类别, 其中每个类别包含 299 个样本. 在实验中, 我们选择了三种特征类型, 分别是 PHOW 特征、PRICoLBP 特征和 CENTRIST 特征用于表示这些图像.

(5)ACM. 该数据集来自于 ACM 的论文网络, 由节点属性和两个图组成. 节点属性与论文的关键词相关, 其中的节点数量为 3025. 两个图分别描述了共同作者和共同主题的关系, 分别表示了两篇论文属于同一作者和关注于同一领域.

(6)IMDB^[28]. 该数据集来自于 IMDB 的电影网络, 由节点属性和两个图组成. 节点属性与电影的属性相关, 其中的节点数量为 4780. 两个图分别描述了共同演员和共同导演的关系, 分别表示了两部电影由同一演员出演或由同一导演执导.

(7)Amazon photos and Amazon computers^[29]. 这两个数据集来自于亚马逊的共购网络. 每个商品的节点特征与产品评论相关, 其中的节点数量分别为 7487 和 13381, 图表征了两个商品是否是一起购买的关系. 为了获得多视图观测值, 我们通过笛卡尔积构建了其他的特征表示.

5.2 对比方法

以下最先进的聚类方法被用来与所提出的 MCHA 方法进行比较.

特别地, 对于多视图属性数据集, 采用标准谱聚类 (SPC)^[56] 来对每个单独视图进行聚类. CoTr^[16] 是一种基于共训练的多视图谱聚类方法. RMSC^[33] 通过标准的马尔可夫链来实现. LTMSC^[37]

是一种低秩张量约束的多视图子空间聚类方法. LMSC^[38]通过潜在空间发现了内在的语义信息. CSMSC^[57]恢复了一致的和视图特定的表示. tSVD^[26]采用张量奇异值分解最小化技术进行多视图子空间聚类. GMC^[36]是一种基于图的多视图聚类方法. WTNNM^[27]使用不同置信度的多个奇异值对张量奇异值分解进行正则化,并用于基于张量奇异值分解的多视图子空间聚类. AIMC^[58]通过发现潜在的完整空间来解决视图信息不足的问题. SMC^[48]是一种基于图过滤的多视图聚类方法,可用于聚类多视图属性和多视图属性图数据对于多视图属性图数据集, LINE^[59]和 GAE^[60]是图处理中两种有代表性的单视图方法. PMNE^[61]利用三种方法将多层网络投影到连续向量空间中. MNE^[62]是一种多视图网络嵌入学习方法. O2MA和 O2MAC^[28]采用图自编码器来实现属性多视图图聚类. MvAGC^[47]通过添加不同阶数的高阶邻域信息来实现多视图属性图聚类的探索. MVGC^[63]采用欧拉变换构建相应的视图描述,并用于非欧几里德图数据的聚类. DIAGC^[51]在深度学习模型中考虑不同多视图属性图之间的对偶信息来增强聚类结果的表达. MAGCN^[29]是一种针对多视图属性图结构化数据的多视图属性图卷积网络. CPLTER^[64]通过对比学习发现不同视图特征之间的共享表示. LMGEC^[65]在统一的框架中同时实现多视图属性图

表征学习和聚类划分.

5.3 实验设置

在本论文中,所提出方法的实验是在英特尔 2.4 GHz 和 128 GB 内存的机器上使用 Matlab R2019b 来完成. 对于 MCHA, 参数 γ 在 $\{0.2, 0.4, 0.6, 0.8, 1\}$ 范围内进行调整, λ 在 $\{0.001, 0.01, 0.1, 1, 5, 10\}$ 范围内进行调整, 考虑到不同视图的多个奇异值具有不同的置信度, 权重因子 ω 在 $(0, 30]$ 内进行选择. 图过滤阶数 k 被设置为 1 或 2. 对于所有基线方法, 为获得最佳聚类结果, 按照论文的建议和实验配置调整最佳参数. 我们对每种方法进行 20 次实验, 并报告其平均性能.

在实验中, 我们采用四个指标来评估聚类性能, 即准确率 (ACC)、归一化互信息 (NMI)、F1 分数 (F1) 和调整秩指数 (ARI). 对于这些评估指标, 更高的值表示更好的聚类性能. 此外, 为了更全面地展示所提出算法的算法有效性, 我们进一步利用时间 (Time) 来记录算法的运行时间成本, 其单位为秒 (s).

5.4 对比实验

在这部分中, 我们分别在表 2~表 5 中报告了多个聚类方法在八个基准数据集上获得的四个评估指标的聚类结果以及运行时间结果. 特别地, “SPC1”到“SPC4”表示我们在对应的特征视图上执行谱聚类.

表 2 不同方法在 MSRC 和 UCI 数据集上的平均性能

数据集	MSRC					UCI				
	方法	ACC	NMI	F1	ARI	Time	ACC	NMI	F1	ARI
SPC ₁ (2002)	0.4298	0.3345	0.3296	0.2205	0.07s	0.6176	0.5858	0.5063	0.4510	1.85s
SPC ₂ (2002)	0.5874	0.4808	0.4689	0.3819	0.09s	0.6847	0.5874	0.5542	0.5044	1.86s
SPC ₃ (2002)	0.5669	0.4844	0.4518	0.3621	0.08s	0.5468	0.4897	0.4271	0.3628	1.87s
SPC ₄ (2002)	0.6852	0.5242	0.5218	0.4433	0.08s	—	—	—	—	—
CoTr (2011)	0.6918	0.6156	0.5900	0.5221	5.29s	0.8401	0.7961	0.7799	0.7549	59.07s
RMSC (2014)	0.6408	0.5704	0.5431	0.4681	1.81s	0.8599	0.8225	0.8006	0.7779	97.19s
LTMSC (2015)	0.8140	0.7321	0.7013	0.6518	3.31s	0.8004	0.7689	0.7489	0.7206	206.37s
LMSC (2017)	0.6743	0.5776	0.5454	0.4703	1.93s	0.8566	0.7837	0.7628	0.7364	135.47s
CSMSC (2018)	0.8310	0.7472	0.7320	0.6880	1.62s	0.8826	0.7877	0.7841	0.7600	61.91s
tSVD ^[26] (2018)	1.0000	1.0000	1.0000	1.0000	1.94s	0.9960	0.9891	0.9920	0.9911	124.04s
GMC (2020)	0.7476	0.7144	0.6749	0.6161	1.41s	0.7355	0.8153	0.7134	0.6779	36.06s
WTNNM (2020)	1.0000	1.0000	1.0000	1.0000	6.02s	0.9965	0.9904	0.9930	0.9922	248.35s
AIMC (2022)	0.7476	0.6534	0.6140	0.5475	0.18s	0.9350	0.8797	0.8776	0.8640	0.51s
SMC (2022)	0.7333	0.6684	0.6996	0.5775	0.62s	0.8540	0.7859	0.8493	0.7203	4.97s
MCHA	1.0000	1.0000	1.0000	1.0000	1.15s	0.9980	0.9945	0.9960	0.9956	172.47s

表 3 不同方法在 COIL20 和 Scene-15 数据集上的平均性能

数据集	COIL-20					Scene-15					
	方法	ACC	NMI	F1	ARI	Time	ACC	NMI	F1	ARI	Time
SPC ₁ (2002)	0.6551	0.7564	0.5983	0.9590	0.87s	0.4256	0.4050	0.3179	0.2673	0.2673	16.92s
SPC ₂ (2002)	0.7457	0.8281	0.7121	0.9711	1.04s	0.3435	0.3515	0.2511	0.5000	0.5000	22.99s
SPC ₃ (2002)	0.6915	0.7928	0.6543	0.9652	0.86s	0.3227	0.2910	0.2145	0.1557	0.1557	22.98s
CoTr (2011)	0.7378	0.8267	0.7069	0.6912	19.58s	0.5848	0.5513	0.4594	0.4193	0.4193	260.07s
RMSC (2014)	0.7544	0.8316	0.7168	0.7020	59.85s	0.4473	0.4212	0.3417	0.2931	0.2931	1142.61s
LTMSC (2015)	0.7060	0.8099	0.6683	0.6500	415.52s	0.5657	0.5710	0.4643	0.4234	0.4234	1856.81s
LMSC (2017)	0.7307	0.8358	0.6975	0.6806	256.59s	0.5234	0.5356	0.4236	0.3768	0.3768	1383.54s
CSMSC (2018)	0.7324	0.8320	0.6947	0.6776	77.35s	0.5927	0.5684	0.4742	0.4339	0.4339	763.66s
tSVDMC (2018)	0.8253	0.9022	0.8177	0.8081	218.69s	0.9013	0.9264	0.8974	0.8897	0.8897	1197.46s
GMC (2020)	0.7910	0.9407	0.7943	0.7819	13.13s	0.3810	0.5186	0.2809	0.1905	0.1905	217.00s
WTNNM (2020)	0.8166	0.9034	0.8123	0.8024	364.07s	0.9021	0.9266	0.8981	0.8905	0.8905	2943.89s
AIMC (2022)	0.5944	0.7995	0.6061	0.5826	3.93s	0.4769	0.5633	0.3946	0.3374	0.3374	4.81s
SMC (2022)	0.7660	0.8354	0.7501	0.6851	3.66s	0.4649	0.4479	0.4473	0.2506	0.2506	22.50s
MCHA	0.8528	0.9008	0.8195	0.8101	121.49s	0.9012	0.9275	0.8978	0.8901	0.8901	2039.33s

表 4 不同方法在 ACM 和 IMDB 数据集上的平均性能

数据集	ACM					IMDB					
	方法	ACC	NMI	F1	ARI	Time	ACC	NMI	F1	ARI	Time
LINE (2015)	0.6479	0.3941	0.6594	0.3433	180.31s	0.4268	0.0031	0.2870	-0.009	0.009	486.75s
GAE (2016)	0.8216	0.4914	0.8225	0.5444	286.57s	0.4298	0.0402	0.4062	0.0473	0.0473	1886.22s
PMNE _n (2017)	0.6936	0.4648	0.6955	0.4302	130.42s	0.4958	0.0359	0.3906	0.0366	0.0366	365.24s
PMNE _r (2017)	0.6492	0.4063	0.6618	0.3453	130.42s	0.4697	0.0014	0.3183	0.0115	0.0115	365.24s
PMNE _c (2017)	0.6998	0.4775	0.7003	0.4431	130.42s	0.4719	0.0285	0.3882	0.0284	0.0284	365.24s
MNE (2018)	0.6370	0.2999	0.6479	0.2486	94.25s	0.3958	0.0017	0.3316	0.0008	0.0008	221.53s
O2MA (2020)	0.8880	0.6515	0.8894	0.6987	423.5s	0.4697	0.0524	0.4229	0.0753	0.0753	4126.37s
O2MAC (2020)	0.9042	0.6923	0.9053	0.7394	423.5s	0.4502	0.0421	0.4159	0.0564	0.0564	4126.37s
MvAGC (2021)	0.8975	0.6735	0.8986	0.7212	5.80s	0.5633	0.0371	0.3783	0.0940	0.0940	10.38s
MVGC (2022)	0.9617	0.8438	0.9608	0.8917	83.09s	0.5331	0.0623	0.3953	0.0848	0.0848	139.87s
SMC (2022)	0.8863	0.6397	0.8869	0.6929	15.94s	0.5686	0.0356	0.4048	0.0998	0.0998	16.57s
DIAGC (2022)	0.9170	0.7161	0.9177	0.7697	63.28s	0.5839	0.0658	0.4301	0.1316	0.1316	78.77s
LMGEC (2023)	0.9302	0.7513	0.9311	0.8031	3.49s	0.5893	0.0632	0.4267	0.1294	0.1294	4.96s
MCHA	0.9855	0.9289	0.9711	0.9567	644.05s	0.5203	0.0690	0.4624	0.1241	0.1241	2373.64s

表 5 不同方法在 Amazon photos 和 Amazon computers 数据集上的平均性能

数据集	ACM					IMDB					
	方法	ACC	NMI	F1	ARI	Time	ACC	NMI	F1	ARI	Time
MAGCN (2020)	0.5167	0.3897	0.4736	0.2401	3783.6s	N/A	N/A	N/A	N/A	N/A	N/A
CPLTER (2021)	0.3678	0.2606	0.3067	0.0759	1029.50s	0.2417	0.1562	0.1601	0.0536	0.0536	1725.84s
MvAGC (2021)	0.6775	0.5237	0.6397	0.3968	72.22s	0.5796	0.3957	0.4117	0.3224	0.3224	215.33s
SMC (2022)	0.5473	0.5305	0.5266	0.2806	43.06s	0.4899	0.4613	0.3718	0.3384	0.3384	186.03s
LMGEC (2023)	0.7117	0.6114	0.6500	0.5123	19.42s	0.3814	0.3487	0.3627	0.2040	0.2040	26.50s
MCHA	0.7935	0.7254	0.7565	0.6421	4402.65s	0.5514	0.5245	0.4774	0.3846	0.3846	9688.34s

注: N/A 表示内存错误。

从这些表中可以看出, 我们提出的 MCHA 方法通常在所有基准数据集上获得了最佳聚类结果, 这证明了 MCHA 在处理多视图属性数据集和多视图属性图数据集方面的优越性。具体而言, 在 UCI

数据集上, MCHA 的 ACC 和 NMI 分别比第二优秀的方法 WTNNM 提高了 0.15% 和 0.41%。在 ACM 数据集上, 相对于第二优秀的方法 MVGC, MCHA 的 ACC 和 NMI 性能分别提高了 2.38% 和

8.51%。在 Amazon photos 数据集上,相对于第二优秀的方法 LMGEC, MCHA 的 ACC 和 NMI 分别提高了 8.18% 和 11.40%。值得注意的是,IMDB 数据集非常稀疏,在拓扑图结构中有许多孤立的数据点,因此在所有比较方法中,NMI 的性能通常让人不满意。

对于多视图属性场景,我们在每个单独视图上采用谱聚类(SPC)来展示它们的不同聚类能力,明确展示了权衡不同视图的重要性。通常可以观察到,通过加权低秩张量约束来探索张量中不同视图奇异值置信度的 MCHA 和 WTNNM 方法在基准数据集上的性能优于 tSVD 方法,它们在探索张量之间的高阶相关性时考虑了来自不同视图的奇异值的不同置信度。但不幸的是,WTNNM 方法无法很好地处理多视图属性图数据。

对于多视图属性图场景,依据表 4, O2MAC、MvAGC、MVGC、DIAGC 和 LMGEC 的性能优于仅考虑多个图信息的 MNE 和 PMNE 方法,进一步证实了属性和图结构信息整合的关键。尽管 O2MAC、MVGC 和 DIAGC 表现出令人印象深刻的性能,但它们只适用于多个图数据的情况,无法处理具有多视图属性的图结构化数据(包含了多个视图观测值的图数据)。从表 5 中可以看出,通过表征学习和聚类划分间的信息互学习,LMGEC 可以表现得很好。然而,通过张量捕捉多个视图之间的高阶相关性,MCHA 通常获得更好的性能。

在运行时间结果对比方面,总的而言,我们所提出的 MCHA 方法在聚类性能和时间成本上取得了相对的平衡。在多视图属性数据集中,SPC 和 AIMC 在所需时间成本上有明显的优势。对于多视

图属性图数据集,LMGEC 花费的时间较少。在未来的工作中,我们将着力于设计并优化可用于大规模多视图属性数据/多视图属性图数据的高效算法。

5.5 消融实验

在这部分中,我们进行了 MCHA 的消融实验,以研究低秩张量、图过滤和张量旋转机制的影响。具体而言,我们将每个相似性图表示 $Z^{(v)}$ 的张量核范数替换为平方弗罗贝尼乌斯范数,其他部分保持不变,并表示为 MCHA-c1。此外,我们删除了图过滤过程,直接将 $\{M^{(v)}\}_{v=1}^m$ 作为多视图观测的输入,并表示为 MCHA-c2。为了进一步验证张量旋转机制的影响,我们保持张量的原始维度(未旋转),其他部分保持不变,并表示为 MCHA-c3。我们在相应的数据集上运行这三个变体,并在表 6 中报告了有关 ACC 和 NMI 的实验结果。相比于 MCHA-c1、MCHA-c2 和 MCHA-c3, MCHA 通常获得更好的聚类性能,这证明了通过低秩张量发现高阶相关性以及通过图过滤保留底层几何结构的重要性,同时也证实了对张量进行旋转的必要性,更好地挖掘不同视图之间的低秩性。此外,我们对两个选定的数据集(即 MSRC 和 ACM 数据集)进行了可视化分析,其结果分别展示在图 4 和图 5 中。对于图 4, (a) 到 (d) 是四个特征; (e)、(f) 和 (g) 是由 MCHA-t1、MCHA-t2 和 MCHA 获得的一致性相似性图。类似地,对于图 5, (a) 是 ACM 的特征; (b)、(c) 和 (d) 是由 MCHA-t1、MCHA-t2 和 MCHA 获得的一致性相似性图。从图中可以看出, MCHA 获得的一致性相似性矩阵在这两个数据集上都揭示了一致清晰的聚类结构。相比之下, MCHA-c2 尽管在 MSRC 数据集上表现良好,但在 ACM 数据集上表现不佳。

表 6 消融实验:比较 MCHA 及其变体的 ACC/NMI 结果

变体	MSRC	UCI	COIL-20	Scene-15
MCHA	1.0000/1.0000	0.9980/0.9945	0.8528/0.9008	0.9012/0.9275
MCHA-c1	0.8175/0.7166	0.8215/0.7921	0.7076/0.7947	0.5611/0.5757
MCHA-c2	1.0000/1.0000	0.9970/0.9918	0.8243/0.8872	0.9023/0.9286
MCHA-c3	0.7810/0.6729	0.6925/0.6949	0.7403/0.7882	0.6091/0.5764
	ACM	IMDB	Amazon photos	Amazon computers
MCHA	0.9855/0.9289	0.5203/0.0690	0.7935/0.7254	0.5514/0.5245
MCHA-c1	0.9068/0.6941	0.4814/0.0652	0.7202/0.6798	0.4828/0.4970
MCHA-c2	0.9484/0.8046	0.5360/0.0343	0.5914/0.5576	0.5156/0.4305
MCHA-c3	0.8476/0.5427	0.4561/0.0547	0.5623/0.5027	0.5025/0.4693

注:其中 MCHA-c1 将关于每个相似性图表示的张量核范数替换为平方弗罗贝尼乌斯范数, MCHA-c2 去除了图过滤过程, MCHA-c3 则保持了张量的原始维度(未旋转)。

5.6 参数分析

对于提出的 MCHA 方法,我们需要调整三个

参数,即 γ 、 λ 和 ω 。在这个部分中,我们进行参数分析,以探讨这三个参数对多个数据集的影响。具体

而言, 对于 γ 和 λ 参数, 我们展示了在 MSRC、UCI、ACM 和 IMDB 数据集上, 不同参数值对 NMI 的影响, 如图 6 所示. 因此, 我们可以看出在给定的广泛参数值范围内, MCHA 也可以在多个数据集上获得稳定的聚类结果. 同样地, 对于 ω 参数, 我们在这四个数据集上报告实验的结果如图 7 所示, 其中记录了在 γ 和 λ 固定的情况下, 权重系数

取不同值时的 ACC 和 NMI 表现. 需要注意的是, 由于张量旋转的操作, ω 的维度与视图数量相关. 例如, 在给定的具有四个视图的 MSRC 数据集上, ω 可以设置为 $[4, 2, 14, 12]$, 其中每个元素表示施加在相应视图上的权重. 此外, 有了最佳的权衡参数值, MCHA 可以在不同权重系数范围内稳定地实现优越的性能.

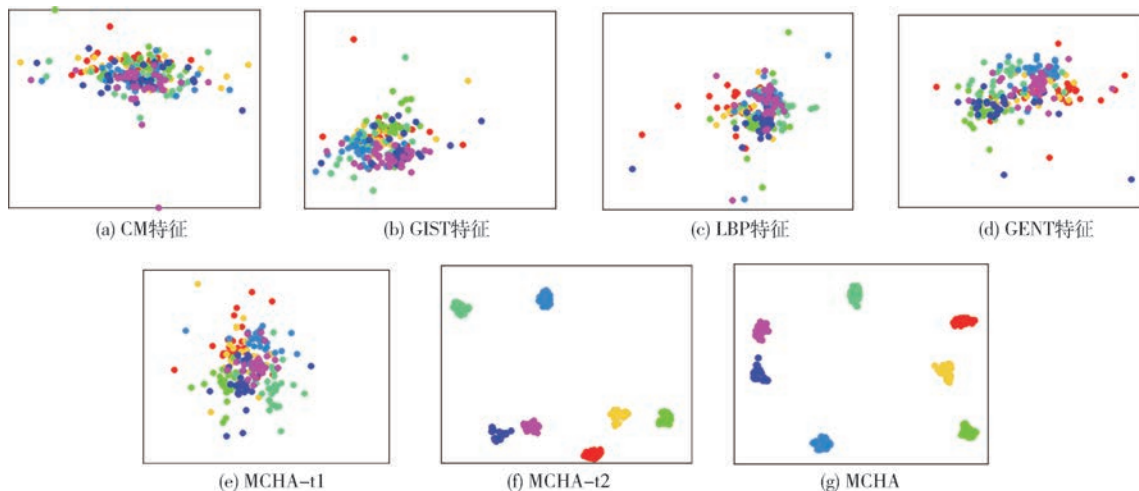


图 4 在 MSRC 上的 T-SNE 可视化

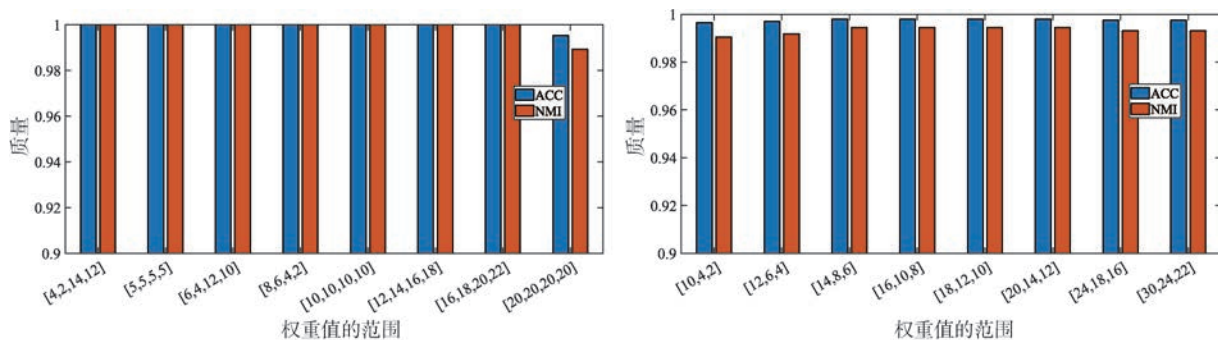


图 5 在 ACM 上的 T-SNE 可视化

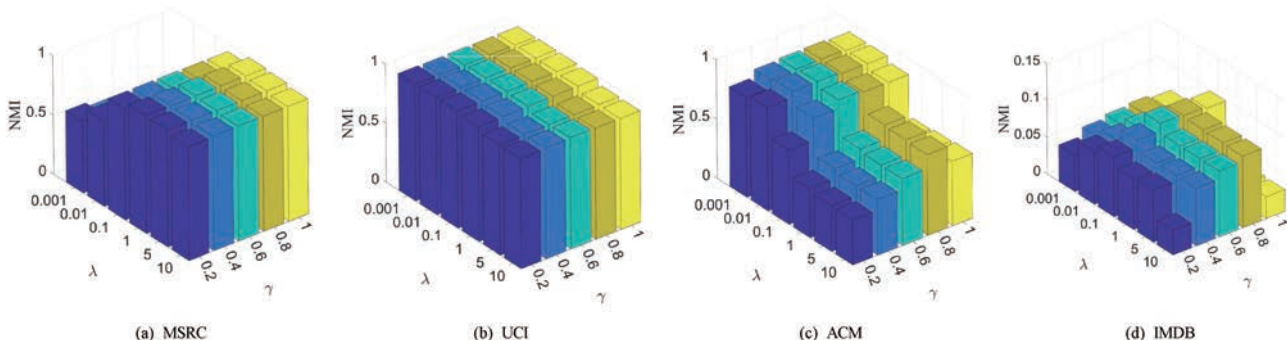


图 6 MCHA 在四个基准数据集上关于 γ 和 λ 的参数分析

5.7 收敛性分析

为了验证 MCHA 的收敛性质, 我们在 MSRC、

UCI、ACM 和 IMDB 数据集上进行了收敛性实验, 如图 8 所示. 在该图中, 记录了重构误差

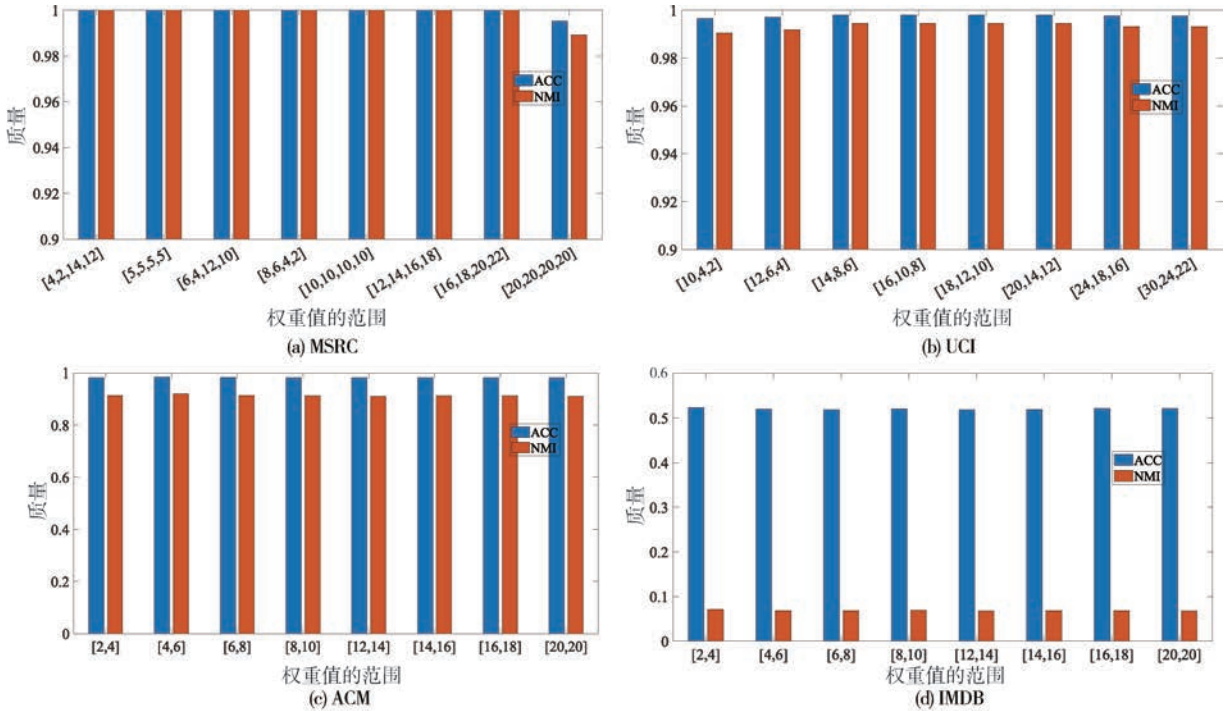


图7 MCHA在四个基准数据集上关于加权系数 $\omega \in \mathbb{R}^{1 \times m}$ 的参数分析

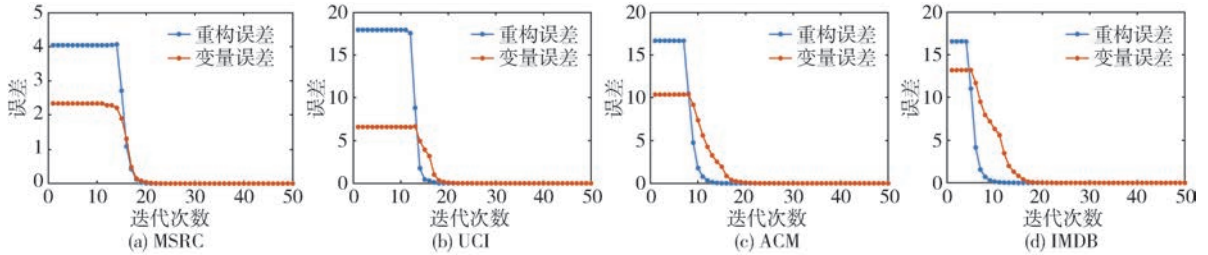


图8 MCHA在四个基准数据集上的收敛分析

$\sum_{v=1}^m \|\mathbf{M}_{(t+1)}^{(v)\top} - \mathbf{M}_{(t+1)}^{(v)\top} \mathbf{S}_{(t+1)}^{(v)} - \mathbf{E}_{(t+1)}^{(v)}\|_{\infty}$ 和变量误差 $\sum_{v=1}^m \|\mathbf{S}(t+1)^{(v)} - \mathbf{Q}(t+1)^{(v)}\|_{\infty}$ 在相应迭代步骤时的误差结果。显然,我们可以观察到MCHA具有良好的收敛性能,并能够实现卓越的聚类性能。

6 总结与展望

在本文中,我们从广义的角度提出了一种新型模型,称为混合阶相似性的多视图聚类(MCHA)。在这个框架中,通过多个图和多视图属性的巧妙结合,我们构建了一组视图特定的光滑表示,以此保留原始的几何结构。基于这些光滑表示,可以学习到多个概率相似性图,并用于记录跨视图的二阶相关性。同时,我们将多视图概率相似性图堆叠到由加权张量核范数约束的张量中,这可以很好地捕获来自不同视图奇异值的显著信息和底层的高阶相关性。在8个真实世界数据集上的广泛实验结果验证

了MCHA的优越性能。由于MCHA具有关于样本量的二次时间复杂度,在未来的工作中,我们将尝试更多方法来减轻在处理大规模或海量多视图属性图数据时的负担。

参 考 文 献

- [1] Xu Y M, Wang C D, Lai J H. Weighted multi-view clustering with feature selection. *Pattern Recognition*, 2016, 53: 25-35
- [2] Kang Z, Wen L J, Chen W Y, Xu Z L. Low-rank kernel learning for graph-based clustering. *Knowledge-Based Systems*, 2019, 163: 510-517
- [3] Huang L, Wang C D, Chao H Y. MVStream: Multiview data stream clustering. *IEEE Transactions on Neural Networks and Learning Systems*, 2019, 31(9): 3482-3496
- [4] Sun S L, Mao L, Dong Z, Wu L D. *Multiview Machine Learning*. Singapore, 2019
- [5] Chen M S, Huang L, Wang C D, Huang D. Multi-view clustering in latent embedding space//*Proceedings of the 34th*

- AAAI Conference on Artificial Intelligence. New York, USA, 2020, 3513-3520
- [6] Sun Y, Ren Z, Hu P, Peng D Z, Wang X. Hierarchical consensus hashing for cross-modal retrieval. *IEEE Transactions on Multimedia*, 2024, 26:824-836
- [7] Zhong G, Pun C M. Latent lowrank graph learning for multi-modal clustering//37th IEEE International Conference on Data Engineering. Chania, Greece, 2021, 492-503
- [8] Zou X, Tang C, Zheng X, Sun K, Zhang W, Ding D Q. Inclusivity induced adaptive graph learning for multi-view clustering. *Knowledge-Based Systems*, 2023, 267: 110424
- [9] Xing L, Zhao H Q, Lin Z P, Chen B D. Mixture correntropy based robust multi-view k -means clustering. *Knowledge-Based Systems*, 2023, 262: 110231
- [10] Li L K, Kang Z, Long B. Riemannian manifold based multi-view spectral clustering algorithm. *Computer Engineering*, 2023, 49(1): 113-120, 129(in Chinese)
(李林珂, 康昭, 龙波. 基于黎曼流形的多视角谱聚类算法. *计算机工程*, 2023, 49(1): 113-120, 129)
- [11] Wu Y, Zhong Z N, Xiong W, Chen L, Jing N. An efficient method for attributed graph clustering. *Chinese Journal of Computers*, 2013, 36(8): 1704-1713(in Chinese)
(吴烨, 钟志农, 熊伟, 陈萃, 景宁. 一种高效的属性图聚类方法. *计算机学报*, 2013, 36(8): 1704-1713)
- [12] Chen J Y, He H H. Research on density-based clustering algorithm for mixed data with determine cluster centers automatically. *Acta Automatica Sinica*, 2015, 41(10): 1798-1813 (in Chinese)
(陈晋言, 何辉豪. 基于密度的聚类中心自动确定的混合属性数据聚类算法研究. *自动化学报*, 2015, 41(10): 1798-1813)
- [13] Luo H, Han J Q. Graph clustering based on flexibly balanced constraint. *Acta Automatica Sinica*, 2023, 49(4): 778-789 (in Chinese)
(罗辉, 韩纪庆. 基于灵活平衡约束的图聚类方法. *自动化学报*, 2023, 49(4): 778-789)
- [14] Liang Y J, Hu T T, Zhao P X. Efficient structural clustering in large uncertain graphs//36th IEEE International Conference on Data Engineering. Dallas, USA, 2020, 1966-1969
- [15] Xu C, Tao D C, Xu C. A survey on multi-view learning. *CoRR*, 2013, abs/1304.5634
- [16] Kumar A, Daum'e III H. A cotraining approach for multi-view spectral clustering//Proceedings of the 28th International Conference on Machine Learning. Bellevue, USA, 2011, 393-400
- [17] Cao X C, Zhang C Q, Fu H Z, Liu S, and Zhang H. Diversity-induced multi-view subspace clustering//IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA, 2015, 586-594
- [18] Li Z Y, Wang Q Q, Tao Z Q, Gao Q X, Yang Z H. Deep adversarial multi-view clustering network//Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence. Macao, China, 2019, 2952-2958
- [19] Ren Z W, Sun Q S. Simultaneous global and local graph structure preserving for multiple kernel clustering. *IEEE Transactions on Neural Networks and Learning Systems*, 2021, 32(5): 1839-1851
- [20] Liu B Y, Huang L, Wang C D, Fan S H, Yu Philip S. Adaptively weighted multiview proximity learning for clustering. *IEEE Transactions on Cybernetics*, 2021, 51(3): 1571-1585
- [21] Chen M S, Wang C D, Lai J H. Low-rank tensor based proximity learning for multi-view clustering. *IEEE Transactions on Knowledge and Data Engineering*, 2023, 35(5): 5076-5090
- [22] Liu Z, Li Y, Yao L N, Wang X Z, Nie F P. Agglomerative neural networks for multiview clustering. *IEEE Transactions on Neural Networks and Learning Systems*, 2022, 33(7): 2842-2852
- [23] Khan A, Maji P. Multimanifold optimization for multiview subspace clustering. *IEEE Transactions on Neural Networks and Learning Systems*, 2022, 33(8): 3895-3907
- [24] Zhan K, Zhang C Q, Guan J P, Wang J S. Graph learning for multiview clustering. *IEEE Transactions on Cybernetics*, 2018, 48(10): 2887-2895
- [25] Zhang G Y, Zhou Y R, He X Y, Wang C D, Huang D. On-estep kernel multi-view subspace clustering. *Knowledge-Based Systems*, 2020, 189: 105126
- [26] Xie Y, Tao D C, Zhang W S, Liu Y, Zhang L, and Qu Y Y. On unifying multi-view self-representations for clustering by tensor multi-rank minimization. *International Journal of Computer Vision*, 2018, 126(11): 1157-1179
- [27] Gao Q X, Xia W, Wan Z X, Xie D Y, Zhang P. Tensor-svd based graph learning for multi-view subspace clustering//The Thirty-Fourth AAAI Conference on Artificial Intelligence. New York, USA, 2020, 3930-3937
- [28] Fan S H, Wang X, Shi C, Lu E M, Lin K, Wang B. One2multi graph autoencoder for multi-view graph clustering //Proceedings of the Web Conference 2020. Taipei, China, 2020, 3070-3076
- [29] Cheng J F, Wang Q Q, Tao Z Q, Xie D Y, Gao Q X. Multi-view attribute graph convolution networks for clustering//Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, 2020, 2973-2979
- [30] Park C Y, Kim D, Han J W, Yu H. Unsupervised attributed multiplex network embedding//Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence. New York, USA, 2020, 5371-5378
- [31] Chen M S, Lin J Q, Li X L, Liu B Y, Wang C D, Huang D, Lai J H. Representation learning in multi-view clustering: A literature review. *Data Science and Engineering*, 2022, 7(3): 225-241
- [32] Xia R K, Pan Y, Du L, Yin J. Robust multiview spectral clustering via low-rank and sparse decomposition//Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence. Quebec City, Canada, 2014, 2149-2155
- [33] Shu L, Latecki J L. Integration of single-view graphs with diffusion of tensor product graphs for multi-view spectral clustering//Proceedings of The 7th Asian Conference on Ma-

- chine Learning. Hong Kong, China, 2015, 45: 362-377
- [34] Nie F P, Li J, Li X L. Self-weighted multiview clustering with multiple graphs // Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence. Melbourne, Australia, 2017, 2564-2570
- [35] Wang H, Yang Y, Liu B. GMC: graph-based multi-view clustering. *IEEE Transactions on Knowledge and Data Engineering*, 2020, 32(6): 1116-1129
- [36] Zhang C Q, Fu H Z, Liu S, Liu G C, Cao X C. Low-rank tensor constrained multiview subspace clustering // Proceedings of the 2015 IEEE International Conference on Computer Vision. Santiago, Chile, 2015, 1582-1590
- [37] Zhang C Q, Hu Q H, Fu H Z, Zhu P F, Cao X C. Latent multi-view subspace clustering // Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA, 2017, 4279-4287
- [38] Zhang C Q, Fu H Z, Hu Q H, Cao X C, Xie Y, Tao D C, Xu D. Generalized latent multi-view subspace clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(1): 86-99
- [39] Xu H L, Zhang X D, Xia W, Gao Q X, Gao X B. Lowrank tensor constrained co-regularized multi-view spectral clustering. *Neural Networks*, 2020, 132: 245-252
- [40] Kumar A, Rai P, Daum' e III H. Co-regularized multiview spectral clustering // Proceedings of the Advances in Neural Information Processing Systems 24: 25th Annual Conference on Neural Information Processing Systems 2011. Granada, Spain, 2011, 1413-1421
- [41] Wang X B, Lei Z, Guo X J, Zhang C Q, Shi H L, Li S Z. Multi-view subspace clustering with intactness-aware similarity. *Pattern Recognition*, 2019, 88: 50-63
- [42] Nie F P, Cai G H, Li J, Li X L. Auto-weighted multiview learning for image clustering and semi-supervised classification. *IEEE Transactions on Image Processing*, 2018, 27(3): 1501-1511
- [43] Wu J L, Lin Z C, Zha H B. Essential tensor learning for multi-view spectral clustering. *IEEE Transactions on Image Processing*, 2019, 28(12): 5910-5922
- [44] Chen Y Y, Xiao X L, Zhou Y C. Multi-view subspace clustering via simultaneously learning the representation tensor and affinity matrix. *Pattern Recognition*, 2020, 106: 107441
- [45] Xu J, Tang H Y, Ren Y Z, Peng L, Zhu X F, He L F. Multi-level Feature Learning for Contrastive Multi-view Clustering // IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA, 2022, 16030 - 16039
- [46] Lin Z P, Kang Z, Zhang L Z, Tian L. Multi-view attributed graph clustering. *IEEE Transactions on Knowledge and Data Engineering*, 2023, 35(2): 1872-1880
- [47] Lin Z P, Kang Z. Graph filter-based multi-view attributed graph clustering // Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence. Montreal, Canada, 2021, 2723-2729
- [48] Liu L, Chen P, Luo G C, Kang Z, Luo Y G, Han S C. Scalable multi-view clustering with graph filtering. *Neural Computing and Applications*, 2022, 34(19): 16213-16221
- [49] Wang X, Ji H Y, Shi C, Wang B, Ye Y F, Cui P, Yu P S. Heterogeneous graph attention network // The World Wide Web Conference (WWW). San Francisco, USA, 2019, 2022-2032
- [50] Pan E L, Kang Z. Multi-view contrastive graph clustering // Advances in Neural Information Processing Systems 34: Annual Conference on Neural Information Processing Systems, virtual, 2021, 2148-2159
- [51] Lin J Q, Chen M S, Zhu X R, Wang C D, Zhang H Z. Dual Information Enhanced Multi-view Attributed Graph Clustering. *arXiv preprint arXiv:2211.14987*, 2022
- [52] Liu G C, Lin Z C, Yan S C, Sun J, Yu Y, Ma Y. Robust recovery of subspace structures by low-rank representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 35(1): 171-184
- [53] Winn J, Jojic N. Locus: Learning object classes with unsupervised segmentation // Proceedings of the 10th IEEE International Conference on Computer Vision. Beijing, China, 2005, 756-763
- [54] Nene S A, Nayar S K, Murase H, et al. Columbia object image library (coil-100). 1996
- [55] Oliva A, Torralba A. Modeling the shape of the scene: A holistic representation of the spatial envelope. *International Journal of Computer Vision*, 2001, 42(3): 145-175
- [56] Ng A Y, Jordan M I, Weiss Y. On spectral clustering: Analysis and an algorithm // Proceedings of the Advances in Neural Information Processing Systems. Vancouver, Canada, 2001, 849-856
- [57] Luo S R, Zhang C Q, Zhang W, Cao X C. Consistent and specific multiview subspace clustering // Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence. New Orleans, USA, 2018, 3730-3737
- [58] Chen M S, Liu T, Wang C D, Huang D, Lai J H. Adaptively-weighted integral space for fast multiview clustering // The 30th ACM International Conference on Multimedia. Lisboa, Portugal, 2022, 3774-3782
- [59] Tang J, Qu M, Wang M Z, Zhang M, Yan J, Mei Q Z. LINE: Large scale information network embedding // Proceedings of the 24th International Conference on World Wide Web. Florence, Italy, 2015, 1067-1077
- [60] Kipf T N., Welling M. Variational graph auto-encoders. *arXiv preprint arXiv:1611.07308*, 2016
- [61] Zhang H M, Qiu L W, Yi L L, Song Y Q. Scalable multiplex network embedding // Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence. Stockholm, Sweden, 2018, 3082-3088
- [62] Liu W Y, Chen P Y, Yeung S L, Suzumura T, Chen L L. Principled multilayer network embedding // 2017 IEEE International Conference on Data Mining Workshops. New Orleans, USA, 2017, 134-141
- [63] Xia W, Wang S, Yang M, Gao Q X, Han J G, Gao X B.

Multi-view graph embedding clustering network: Joint self-supervision and block diagonal representation. *Neural Networks*, 2022, 145: 1-9

- [64] Lin Y J, Gou Y B, Liu Z T, Li B Y, Lv J C, Peng X. COMPLETER: incomplete multi-view clustering via contrastive prediction// *IEEE Conference on Computer Vision and Pat-*

tern Recognition, 2021, 11174-11183

- [65] Fettal C, Labiod L, Nadif M. Simultaneous Linear Multi-view Attributed Graph Representation Learning and Clustering// *Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining*. Singapore, 2023, 303-311



CHEN Man-Sheng, Ph. D. candidate. Her research interests include data mining and multi-view clustering.

REN Li-An, undergraduate student. His research interests include data mining and data clustering.

WANG Chang-Dong, Ph. D., associate professor. His research interests include machine learning and data mining.

HUANG Dong, Ph. D., associate professor. His research interests include machine learning and data mining.

LAI Jian-Huang, Ph. D., professor. His research interests include machine learning and data mining.

Background

Multi-view data are ubiquitous in many real-world applications, providing more discriminative and complementary information than the single-view data. For instance, we can represent a webpage by multiview information about image and text. Nowadays, massive complex network data from the Internet and social media are all around us, in which they are naturally represented as graphs, the typical non-Euclidean data structures. For instance, in an academic network, one graph view could describe the co-subject relationship, indicating that two papers are in the same field, while another graph view represents the co-author relationship, meaning that two papers are proposed by the same author; papers themselves are described by multiple features, such as representative words, research fields and citation index, which is named the attributed graph data. Specifically, how to take full advantage of the available information from multiple (graph) views remains a scientific problem.

The emergence of multiview data promotes the development of multiview clustering. In the past few years, a great number of multiview clustering methods have been studied extensively by exploring the relationships between data samples represented by multiview attributes or multiview attribute graph. Aiming at multiview attribute data, we roughly divide the existing multiview clustering methods into three main categories, namely co-training or co-regularized style algorithms, graphbased methods and subspace learning based models. However, most existing multiview clustering approaches merely take multiview attributes into consideration, paying little attention to multiview attributed graph data. Therefore, some attempts are made to comprehensively integrate the information from

the attribute and graph. Fan et al. attempted to deploy one informative graph view and attribute data to extract the node embeddings. Pan et al. developed a multiview contrastive graph clustering method. Despite impressive performance, they mainly focus on the consensus representation or common relationships across views, mostly neglecting the underlying higher-order correlations between multiview observations.

According to the aforementioned challenges, a novel method called Multiview Clustering by Hybrid-order Affinity (MCHA) from a generalized perspective is developed in this paper, where the structural graph and the Multiview attributes are seamlessly fused, and the low-rank probability affinity graph representations with hybrid-order correlations are simultaneously considered. To be specific, a set of view-specific smooth representations are first constructed to preserve the geometrical structures by means of a graph filtering strategy. Based on the smooth representations, the view-specific probability affinity graph recording the second-order correlations across samples can be then learned via the self-expression property. Meanwhile, we stack the multiview probability affinity graphs into a tensor singular value decomposition based weighted tensor, where the cross-view higher-order correlations as well as the salient semantic information from different singular values can be well captured. Consequently, not only the second-order correlations but also higher-order ones from Multiview probability affinity graphs are captured for deeper study.

This work was supported by National Key Research and Development Program of China (2021YFF1201202), NSFC (62276277, 61976097) and Guangdong Provincial Key Laboratory of Intellectual Property and Big Data (2018B030322016).