ISSN 1004-9037,CODEN SCYCE4
Journal of Data Acquisition and Processing Vol. 41,No. 1,Jan. 2026,pp. 215－230
DOI:10. 16337/j. 1004-9037. 2026. 01. 015

http://sjcj. nuaa. edu. cn
E-mail:sjcj@nuaa. edu. cn
Tel/Fax:＋86-025-84892742

基于张量低秩学习的多视图子空间聚类方法

史德胜[1], 徐　鹤[1,2], 李　鹏[1,2]

(1. 南京邮电大学计算机学院、软件学院、网络空间安全学院,南京 210023;2. 江苏省高性能计算与智能处理工程研究中心,南京 210023)

摘　要: 多视图聚类是一种机器学习技术,通过整合多源信息可以显著提升聚类性能。然而,现有方法在处理多视图数据时未能充分利用张量低秩表示与亲和度矩阵之间的内在相关性,同时在应对数据缺失、噪声和异常值时表现不佳。为解决这些问题,提出了一种基于张量低秩学习的多视图子空间聚类方法。该方法通过对数据样本施加低秩约束,深入挖掘数据点之间的高阶关联性,精确识别数据的子空间结构。同时,引入张量奇异值分解和加权张量核范数最小化方法,对亲和度矩阵进行优化,将聚类问题转化为一个凸优化问题求解,确保了模型的鲁棒性和效率。此外,所提方法在捕捉多视图数据复杂关联性方面表现出色,能够更准确地识别数据的潜在子空间结构。实验结果表明,所提方法在4类基准数据集上的性能优于现有方法,具有高的准确性和鲁棒性。

关键词: 多视图聚类;张量低秩;子空间聚类;高阶关联性;张量分解

中图分类号: TP391　　**文献标志码:** A

引用格式: 史德胜,徐鹤,李鹏. 基于张量低秩学习的多视图子空间聚类方法[J]. 数据采集与处理,2026,41(1):215-230. SHI Desheng,XU He,LI Peng. Multi-view subspace clustering method based on tensor low-rank learning[J]. Journal of Data Acquisition and Processing,2026,41(1):215-230.

引　　言

随着现代信息技术的高速发展,市面上的大多数应用都涉及多视图数据。与单视图数据相比,多视图数据来自多个特征提取器或不同的领域,能够捕捉更丰富的特征和多样化的信息[1]。与此同时,多视图学习在人工智能和数据分析中愈发受到关注,因为其能够提供更为丰富且互补的信息,从而增强模型的区分能力和分析深度[2-3]。然而,尽管多视图数据提供了更为丰富的特征信息,但其处理过程也变得更加复杂。在实际应用中,采集到的多视图数据往往并不完整,存在视图缺失的现象[4]。例如,在网页内容处理中,有些网页可能同时包含文本、图像和视频,而另一些网页则仅有文本,从而导致视图不全。在工业监控领域,多源数据通常由一组独立的检测器进行采集,但由于检测器可能因故障或其他原因而无法正常工作,导致部分数据无法获取或丢失。在多视图聚类任务中,每个视图都提供了独特的数据特征,这些特征相互补充,共同促进了对数据的全面理解。这种多视图的协同效应对于聚类任务至关重要,因为其可以帮助算法更准确地识别数据中的模式和结构。然而,当某些视图因故障或其他原因而无法提供数据时,这种信息的不完整性会对聚类结果的准确性和稳定性产生负面影响。具体来说,缺失的视图可能导致聚类算法无法充分利用所有可用数据的协同效应,从而影响聚

基金项目: 国家重点研发计划(2019YFB2103003)。
收稿日期: 2024-09-20;**修订日期:** 2024-12-12

类结果的准确性和稳定性，即聚类算法能否准确地将相似的数据点分到同一个簇中，并且这种聚类结果在不同的数据集或不同的初始化条件下是否保持稳定。目前，国内外学者在研究多视图子空间聚类算法时，通常采用一种两步策略[5-7]。首先，对多视图数据进行张量表示学习[8-10]。在这个过程中，来自不同视图的数据被组织成一个多维张量，这种表示方式捕获了各个视图之间复杂的关系和关联特征。通过张量表示学习，研究者可以在保持每个视图独立性和多样性特征的同时，将来自不同视图的数据有效地整合到一个统一的框架中。完成张量表示学习后，得到的表示张量会被融合成一个统一的相似度亲和矩阵。这一矩阵在多视图子空间聚类中起着至关重要的作用，因为其能够反映各个数据点在所有视图下的整体相似性。通过融合不同视图的信息，亲和度矩阵能够为后续的聚类过程提供更加全面和准确的依据。最终，基于这个统一的亲和度矩阵，采用子空间聚类算法对数据进行划分，实现对多视图数据的聚类分析。然而，通过这种方式学习得到的子空间自表示特征和聚类信息是独立且顺序进行的，未考虑到张量表示学习对最终聚类计算结果的影响[11]，并且忽视了张量秩对亲和度矩阵的先验特性影响。

为了解决这一问题，需要一种能够同时学习表示张量和亲和度矩阵的方法，从而更有效地捕捉它们之间的关联性。提出了一种基于张量低秩学习的多视图子空间聚类（Tensor low rank based multi-view subspace clustering，TLR-MVSC），旨在有效减少噪声与异常值对亲和度矩阵的干扰，并通过鲁棒性学习低秩张量表示，从而优化亲和度矩阵，提升多视图聚类性能。主要贡献概要如下：

（1）提出了一种基于低秩表示（Low rank representation，LRR）的高阶张量约束方法，有效减少了噪声和异常值对多视图数据聚类的影响。通过对数据样本施加低秩约束，能够充分挖掘数据视图之间的高阶关联性，从而更准确地揭示数据的子空间结构，并增强亲和度矩阵的鲁棒性。

（2）亲和度矩阵采用基于张量奇异值分解（Tensor singular value decomposition，t-SVD）的张量核范数最小化（Tensor nuclear norm minimization，TNNM）方法进行优化。该方法不仅能更好地捕捉高维数据的全局结构，还能显著提升亲和度矩阵的质量，从而在多视图数据的聚类性能上取得更优的表现。

（3）通过实验验证，在人脸图像、新闻文本、手写数字和通用对象场景4类数据集上，采用3种标准聚类评估指标进行性能评估，并通过算法对比和评价指标分析，验证了所提方法的性能和有效性。

1 相关工作

1.1 多视图数据

近年来，为了充分挖掘缺失数据中的潜在信息，学者们对不完备多视图聚类方法进行了广泛而深入的研究[12-18]。这些研究旨在开发能够处理视图缺失问题的算法，以在不完整数据的情况下依然能够实现有效的聚类分析，从而提升多视图学习的鲁棒性和应用广泛性。

其中，Lin等[19]提出了一种通过对比学习最大化不同视图间的互信息，学习丰富一致表示，并通过双重预测最小化条件熵，有效恢复缺失视图。Wang等[20]将与单个视图相关的相似性诱导图、所有视图的统一图以及聚类指示矩阵整合到一个统一的框架中。这种方法通过融合多个视图的信息，进一步提高了聚类结果的准确性和一致性。文献[21]提出了一种新的多视图谱聚类模型，该模型同时执行图融合与谱聚类，所得融合图近似于各独立视图的原图，同时保有明确的聚类架构。Zhang等[22]通过引入各视图共享的潜在表示探索数据集内部的低秩结构。陈梅等[23]提出了一种新的基于多级自表示约束的不完备多视图聚类（Incomplete multi-view clustering based on multi-level self representation constraints，CMLC），CMLC方法利用公共潜在表示来恢复缺失值，从而有效获取缺失部分的潜在信息，以获得多视图数据的统一低秩表示。

1.2 子空间聚类

子空间聚类为解决高维多视图数据的聚类问题提供了一种有效途径[24]。这一方法通过子空间学习从数据中获取子空间表示后,构建表示矩阵并生成亲和矩阵,进而应用谱聚类[25]以获取聚类结果。子空间聚类在表示学习[26]、运动分割[27-29]和图像处理[30]等多个领域得到了广泛应用,已成为这些领域研究的重要工具。

文献[31]则提出了一种新颖的共识一步多视图子空间聚类(Consensus one-step multi-view subspace clustering,COMVSC)方法,能够优化判别分区级信息的整合,并有助于消除数据中的噪声。陈梦瑶等[32]提出了一种基于低秩表示的多视图聚类方法,能够有效提取数据的全局结构并对噪声具有较强的鲁棒性。为进一步提高聚类性能,提出了一种结合特征多样性增强和低秩表示的多视图聚类算法。

尽管这些研究在多视图数据处理和子空间聚类方面取得了显著进展,但在有效捕捉多视图数据之间的复杂关联性以及处理噪声和异常值影响方面仍存在一定不足。因此,基于张量低秩学习的多视图子空间聚类方法有望弥补这些不足,通过更精确地处理数据缺失和噪声问题,进一步提升聚类性能。

2 基于张量低秩学习的多视图子空间聚类方法

2.1 形式化表征

张量低秩约束的核心优势在于能够从多个视图中捕获数据的全局结构和高阶相关性。这种方法通过将不同视图的数据整合到一个统一的张量表示中,并施加低秩约束,从而挖掘数据点之间的深层次关系。与传统的矩阵分解不同,张量分解能够利用数据的多维结构,揭示出更加丰富的信息,这对于多视图聚类任务尤为重要。

通过对张量数据施加低秩约束,不仅能够减少噪声和异常值的影响,还能够强化视图间的协同效应。这种协同效应是通过在不同视图间,提高聚类结果的准确性和鲁棒性,从而找到共同的、一致的结构特征来实现的。具体来说,当一个视图中的数据由于某些原因(如传感器故障或数据丢失)不完整时,其他视图可以提供补偿信息,帮助恢复或估计缺失的数据,从而确保聚类算法能够充分利用所有可用信息。

此外,张量低秩约束还有助于揭示不同视图间的难以察觉的潜在关联,这些关联可能在单一视图中并不明显。通过对这些视图进行拟合计算,可以得到一个更加全面且准确的结构表示,这对于提高聚类性能至关重要。例如,在处理多源数据时,不同检测器可能捕获到的数据特征不同,但共同构成了对同一现象的完整描述。张量低秩约束通过优化这种多视图表示,使得聚类算法能够识别和利用这些跨视图的关联,从而提高聚类的准确性。

在数据科学和机器学习领域,形式化表征能够为问题建模提供明确的数学框架[32],从而更好地分析和解决实际问题。特别是在处理高维数据时,利用精确的数学描述可以帮助捕捉数据的内在结构,优化算法的性能,并提高模型的解释能力。

设 $\boldsymbol{Y}=[y_1,y_2,\cdots,y_n]\in\mathbf{R}^{d\times n}$ 为 d 维数据向量的集合,假设每个数据向量是由 k 个未知维数 $\{S_i\}_{i=1}^{k}$ 的线性子空间并集得到的,为了找到LRR $\boldsymbol{Z}\in\mathbf{R}^{n\times n}$,LRR解决了以下秩最小化问题

$$\min_{\boldsymbol{Z}}\ \mathrm{rank}(\boldsymbol{Z})\quad \text{s.t.}\quad \boldsymbol{Y}=\boldsymbol{D}\boldsymbol{Z} \tag{1}$$

式中:$\boldsymbol{D}=[d_1,d_2,\cdots,d_n]\in\mathbf{R}^{d\times n}$ 为用于数据向量线性表示的字典。

由于问题式(1)是非凸的NP难问题,LRR采用核范式作为秩函数的常用替代方法,从而获得如下凸解

$$\min_{\boldsymbol{Z}}\ \|\boldsymbol{Z}_*\|\ \text{s.t.}\ \boldsymbol{Y}=\boldsymbol{D}\boldsymbol{Z} \tag{2}$$

式中$\|\boldsymbol{Z}_*\|$表示矩阵$\boldsymbol{Z}$的核范数。上述优化问题可以利用奇异值分解有效解决,问题式(2)的解$\boldsymbol{Z}$被认为是$\boldsymbol{Y}$的LRR特征集,从整体上来看,构造的亲和度矩阵$|\boldsymbol{Z}|+|\boldsymbol{Z}|^{\mathrm{T}}$可以用于最终的谱聚类,得到$\boldsymbol{Y}$的聚类结果。

为了更好地处理高维数据的结构复杂性,可以利用张量奇异值分解这一工具。

使用张量奇异值分解的核范数求解表达式如下

$$\|\mathcal{X}\|_{\circledast}=\sum_{i=1}^{\min\{n_1,n_2\}}\sum_{j=1}^{n_3}\left|\hat{\mathcal{G}}(i,i,j)\right| \tag{3}$$

式中:$\hat{\mathcal{G}}=\mathrm{FFT}(\mathcal{G},[\,],3)$表示沿第三维方向的离散傅里叶变换,揭示了第$j$个正向切片的第$i$个奇异值。

张量奇异值分解不仅提供了一种有效的低秩近似方法,还能够捕捉数据的多维结构信息。然而,仅依赖单一的低维表示可能不足以揭示数据的全部特征。高维数据,如图像和视频,通常可以被描述为低维表示,这意味着几个参数可以表示复杂的结构。但实际上高维数据通常可以用某几个低维子空间数据结构来间接表示,具有相同低维结构的样本通常高度相似。因此,可以通过划分子空间来获得聚类结果。

子空间聚类的目标是将样本点准确地分配到数据矩阵$\boldsymbol{X}$中对应的子空间,该矩阵的样本点来源于多个子空间。子空间聚类可以通过“自表示”方式学习表征矩阵$\boldsymbol{Z}$,也称为系数矩阵$\boldsymbol{Z}$。“自表示”指的是每个数据点可以通过同一子空间内其他数据点的线性组合进行准确表示,其形式如下

$$\boldsymbol{X}=\boldsymbol{D}\boldsymbol{Z}+\boldsymbol{E} \tag{4}$$

式中:$\boldsymbol{D}$为用于对数据进行线性表示的字典,通常用$\boldsymbol{X}$来代替;$\boldsymbol{E}$表示原始数据$\boldsymbol{X}$中的噪声误差项。自表示的子空间聚类模型如下

$$\min\lambda\Omega(\boldsymbol{Z})+\Phi(\boldsymbol{E}) \tag{5}$$

式中:$\Omega(\cdot)$和$\Phi(\cdot)$表示具体的正则化策略;$\lambda>0$为一个参数,用来平衡两个正则化器。

对于包含V个不同视图的多视图数据,$\boldsymbol{X}_v\in\mathbf{R}^{d_v\times n}$表示第$v$个视图,其中$n$为样本的数量,$d_v$为$\boldsymbol{X}_v$的一个特定维度的维数。式(5)可以推广到多视图聚类(Multi-view clustering,MVC)上,定义为

$$\begin{cases}\min\sum_{v=1}^{V}\lambda\Omega(\boldsymbol{Z}_v)+\Phi(\boldsymbol{E}_v)\\ \text{s.t. }\boldsymbol{X}_v=\boldsymbol{X}_v\boldsymbol{Z}_v+\boldsymbol{E}_v\end{cases} \tag{6}$$

式中:$\boldsymbol{X}_v$为第v个视图的误差项,$\boldsymbol{Z}_v$为第v个视图的系数矩阵[34]。LRR是一种比较常用的策略,可以从包含错误(如噪声和异常值)的数据中恢复子空间结构。该策略在正则化项中使用秩函数$\Omega(\cdot)$,通过秩最小化方法可以恢复潜在的行空间。然而,由于秩函数的离散性,获取解变得困难。在这种情况下,可以采用基于t-SVD的核范数作为其凸松弛。此外,为了减少噪声干扰,在正则化项$\Phi(\cdot)$中可以使用$l_{2,1}$范数来增强列稀疏性。基于以上两种策略,优化问题式(6)可定义为

$$\begin{cases}\min\sum_{v=1}^{V}\lambda\|\boldsymbol{Z}_v\|_*+\|\boldsymbol{E}_v\|_{2,1}\\ \text{s.t. }\boldsymbol{X}_v=\boldsymbol{X}_v\boldsymbol{Z}_v+\boldsymbol{E}_v\end{cases} \tag{7}$$

在得到每个视图的优化结果$\boldsymbol{Z}_v$后,相似矩阵$\boldsymbol{S}\in\mathbf{R}^{n\times n}$为

$$\boldsymbol{S}=\sum_{v=1}^{V}\frac{|\boldsymbol{Z}_v|+|\boldsymbol{Z}_v|^{\mathrm{T}}}{2} \tag{8}$$

最后通过$\boldsymbol{S}$进行谱聚类,得出最终结果。

2.2 基于张量低秩学习的子空间聚类模型

常见的张量低秩约束方法在处理奇异值时,通常对每个奇异值给予平等考虑,并通过相同的参数对其进行收缩。在实际应用中,矩阵的非零奇异值之间往往存在显著差异,特别是某些较大的奇异值通常代表了矩阵中携带的关键结构信息,这种显著差异被称为先验信息,对于图像去噪、矩阵补全等任务至关重要。此外,现有的张量低秩多视图方法通常将亲和矩阵的学习与谱聚类分为两个独立的步骤,因此学习到的亲和矩阵无法很好地表征结构,限制了聚类的性能。

通过学习每个视图的自表示系数矩阵,并结合张量核范数最小化问题,利用张量低秩约束来获得更好的自表示。这种自表示充分利用了数据间的高阶相关性和视图间的互补性,是本文提出的多视图子空间聚类方法的核心创新,并在其中引入自适应约束奇异值权重来约束张量先验结构信息。在此基础上,将谱聚类纳入统一框架中,有助于使最终的融合相似矩阵能够更好地表征聚类结构,从而提升聚类的效果。图1展示了所提出模型的框架。

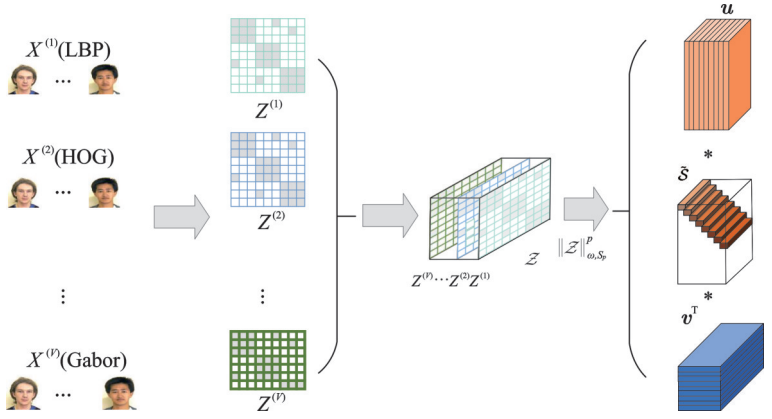

图1 基于张量低秩学习的子空间聚类模型

Fig.1 Model of subspace clustering based on tensor low-rank learning

2.3 融合自适应加权Schatten-p范数的张量低秩学习方法

张量核范数最小化[35],是一种用于张量数据处理中进行低秩张量近似和恢复的方法,一般的张量核范数最小化的模型可以表示为

$$\underset{\mathcal{X}\in\mathbf{R}^{n_1\times n_2\times n_3}}{\operatorname{argmin}}\ \frac{1}{2}\|\mathcal{X}-\mathcal{A}\|_{\mathrm{F}}^{2}+\lambda\|\mathcal{X}\|_{\circledast} \tag{9}$$

通过引入加权Schatten-p范数,来描述张量秩特性信息和充分利用张量奇异值的先验特性。对任意张量$\mathcal{X}\in\mathbf{R}^{n_1\times n_2\times n_3}$,其定义如下

$$\|\mathcal{X}\|_{\omega,S_p}=\left(\sum_{i=1}^{n_3}\|\bar{\mathcal{X}}^{(i)}\|_{\omega,S_p}^{p}\right)^{\frac{1}{p}}=\left(\sum_{i=1}^{n_3}\sum_{j=1}^{\min(n_1,n_2)}\omega_j*\sigma_j(\bar{\mathcal{X}}^{(i)})^{p}\right)^{\frac{1}{p}} \tag{10}$$

式中:ω_j为权重向量$\boldsymbol{\omega}$的第j个权重值,$\sigma_j(\bar{\mathcal{X}}^{(i)})$为$\bar{\mathcal{X}}^{(i)}$的第$j$个最大奇异值,$S_p$表示Schatten-p范数,$p$表示幂次。

现有的张量低秩约束方法通常对每组奇异值采取相同的处理方式,通过相同的参数对所有奇异值进行放大或收缩,然而在实际应用中,矩阵的非零奇异值之间往往存在显著差异,而前几个最大的奇异值通常表征了矩阵中嵌入的显著结构信息。这种被称为先验信息的显著差异在图像去噪、矩阵补全等

领域非常重要，因此通过引入自适应加权的张量Schatten-p范数，能够适当地学习每个视图的自表示系数矩阵，通过求解这一改进的约束问题，从而充分利用多视图的高阶信息和互补信息，有助于使最终的融合相似矩阵更好地表征聚类结构，从而显著提升聚类效果。目标模型表示如下

$$\begin{aligned}&\min_{\mathcal{Z},E^{(v)},F}\|\mathcal{Z}\|_{\omega,S_p}^{p}+\lambda\|E\|_{2,1}+2\alpha\mathrm{tr}\left(F^{\mathrm{T}}L_{\hat{Z}}F\right)\\&\text{s.t. }X^{(v)}=X^{(v)}Z^{(v)}+E^{(v)},v=1,2,\cdots,m\end{aligned}\tag{11}$$

式中：$Z^{(v)}$为张量$\mathcal{Z}\in\mathbf{R}^{N\times m\times N}$的侧切片，即$\mathcal{Z}(:,v,:)=Z^{(v)}$；$E^{(v)}\in\mathbf{R}^{d_v*N}$为第$v$个视图的误差矩阵；$E=[E^{(1)};E^{(2)};\cdots;E^{(m)}]$可以强制每个视图中$E^{(v)}$的列具有共同一致的大小值；$L_{\hat{Z}}=D_{\hat{Z}}-\hat{Z}$为拉普拉斯矩阵，$\hat{Z}=\frac{1}{m}\sum_{v=1}^{m}\frac{|Z^{(v)}|+|Z^{(v)}|^{\mathrm{T}}}{2}$和$D_{\hat{Z}}$为对角矩阵，其对角项为$D_{\hat{Z}}(i,i)=\sum_{j}(\hat{Z}_{ij}+\hat{Z}_{ji})$；$F\in\mathbf{R}^{c\times N}$表示聚类指标矩阵，$c$为聚类个数；$\lambda$和$\alpha$为两个平衡参数。

在式(11)所表示的目标模型中，Schatten-p范数的权重向量需要手动设置。当视图间存在巨大差异和复杂数据分布时，为不同奇异值选取适当权重极具挑战性，限制了算法灵活性。考虑到张量数据中的奇异值差异明显，较大奇异值通常反映张量核心结构，因此设计了自动加权机制以增强算法适应性与灵活性。具体来说，对于张量$\bar{\mathcal{X}}$的第i个奇异值$\sigma_i(\bar{\mathcal{X}})$，可以定义相应的权重元素为

$$\omega_i=\frac{L}{\sigma_i(\bar{\mathcal{X}})+\tau}+\omega_i'\tag{12}$$

2.4 张量低秩学习模型优化

在传统的张量低秩近似问题中，通常假设张量的所有奇异值在优化过程中具有相同的重要性，并通过一个固定的参数对其进行统一的放缩。然而，研究中观察到不同视图数据的贡献程度存在显著差异，特别是当某些视图因数据缺失或噪声干扰而变得不可靠时。为了解决这一问题，加入的自适应加权张量Schatten-p范数方法，以动态调整不同奇异值的权重，从而更好地捕捉多视图数据的互补信息和高阶相关性。

具体来说，通过引入一个自适应权重向量来调整张量核范数，其中每个权重因子是基于对应奇异值的重要性自动确定的。这种方法不仅考虑了奇异值的大小，还考虑了在不同视图数据中的作用和可靠性。通过这种方式，能够自适应地突出重要视图的贡献，并抑制不可靠视图的影响，从而提高结果的准确性。

为了实现这一自适应加权机制，设计了一种基于数据驱动的权重更新策略。在每次迭代中，根据当前估计的亲和度矩阵来更新权重向量，使得较大的奇异值(对应于关键结构信息)获得较大的权重，而较小的奇异值(可能对应于噪声或异常值)获得较小的权重。这种动态调整策略使得优化算法能够更加灵活地应对数据的复杂性和变化性。

引入辅助变量$\mathcal{Q},\mathcal{J}$后构建增广拉格朗日函数，模型式(11)重构为无约束优化问题

$$\begin{aligned}&\mathcal{L}(Z^{(1)},Z^{(2)},\cdots,Z^{(m)},\mathcal{J},E^{(1)},E^{(2)},\cdots,E^{(m)},F)=\|\mathcal{J}\|_{\omega,S_p}^{p}+\lambda\|E\|_{2,1}+2\alpha\mathrm{tr}\left(F^{\mathrm{T}}L_{\hat{Z}}F\right)+\\&\sum_{v=1}^{m}\left(\begin{array}{l}\langle Y^{(v)},X^{(v)}-X^{(v)}Z^{(v)}-E^{(v)}\rangle\\+\frac{\mu}{2}\|X^{(v)}-X^{(v)}Z^{(v)}-E^{(v)}\|_{\mathrm{F}}^{2}\end{array}\right)+\langle\mathcal{Q},\mathcal{Z}-\mathcal{J}\rangle+\frac{\rho}{2}\|\mathcal{Z}-\mathcal{J}\|_{\mathrm{F}}^{2}\end{aligned}\tag{13}$$

式中：矩阵$Y^{(v)}$和张量$\mathcal{Q}$代表两个拉格朗日乘子，μ和ρ为惩罚参数。

因此，模型(13)可以分解为求解下列4个子问题。

子问题1:求解$\boldsymbol{Z}^{(v)}$,变量$\boldsymbol{E}$、$\boldsymbol{\mathcal{J}}$、$\boldsymbol{F}$是固定的。

$$\begin{aligned}&\arg\ \min\ 2\alpha\mathrm{tr}\left(\boldsymbol{F}^{\mathrm{T}}\boldsymbol{L}_{\hat{Z}}\boldsymbol{F}\right)+\langle\boldsymbol{\mathcal{Q}},\boldsymbol{\mathcal{Z}}-\boldsymbol{\mathcal{J}}\rangle+\\&\sum_{v=1}^{m}\left(\langle\boldsymbol{Y}^{(v)},\boldsymbol{X}^{(v)}-\boldsymbol{X}^{(v)}\boldsymbol{Z}^{(v)}-\boldsymbol{E}^{(v)}\rangle+\frac{\mu}{2}\|\boldsymbol{X}^{(v)}-\boldsymbol{X}^{(v)}\boldsymbol{Z}^{(v)}-\boldsymbol{E}^{(v)}\|_{\mathrm{F}}^{2}\right)+\frac{\rho}{2}\|\boldsymbol{\mathcal{Z}}-\boldsymbol{\mathcal{J}}\|_{\mathrm{F}}^{2}=\\&\arg\min\ 2\alpha\mathrm{tr}\left(\boldsymbol{F}^{\mathrm{T}}\boldsymbol{L}_{\hat{Z}}\boldsymbol{F}\right)+\langle\boldsymbol{\mathcal{Q}},\boldsymbol{\mathcal{Z}}-\boldsymbol{\mathcal{J}}\rangle+\\&\sum_{v=1}^{m}\left(\frac{\mu}{2}\left\|\boldsymbol{X}^{(v)}-\boldsymbol{X}^{(v)}\boldsymbol{Z}^{(v)}-\boldsymbol{E}^{(v)}+\frac{\boldsymbol{Y}^{(v)}}{\mu}\right\|_{\mathrm{F}}^{2}+\frac{\rho}{2}\left\|\boldsymbol{Z}^{(v)}-\boldsymbol{J}^{(v)}+\frac{\boldsymbol{W}^{(v)}}{\rho}\right\|_{\mathrm{F}}^{2}\right)+\frac{\rho}{2}\|\boldsymbol{\mathcal{Z}}-\boldsymbol{\mathcal{J}}\|_{\mathrm{F}}^{2}\end{aligned}\tag{14}$$

$$2\mathrm{tr}(\boldsymbol{F}^{\mathrm{T}}\boldsymbol{L}_{\hat{z}}\boldsymbol{F})=\mathrm{tr}\left(\boldsymbol{P}^{\mathrm{T}}\hat{\boldsymbol{Z}}\right)=\mathrm{tr}\left(\boldsymbol{P}^{\mathrm{T}}\left(\frac{1}{m}\sum_{v=1}^{m}\frac{|\boldsymbol{Z}^{(v)}|+|\boldsymbol{Z}^{(v)}|^{\mathrm{T}}}{2}\right)\right)=\frac{1}{2m}\sum_{v=1}^{m}\mathrm{tr}\left(\boldsymbol{P}^{\mathrm{T}}|\boldsymbol{Z}^{(v)}|+\boldsymbol{P}^{\mathrm{T}}|\boldsymbol{Z}^{(v)}|^{\mathrm{T}}\right)\tag{15}$$

式中:$\boldsymbol{P}=[P_1,\cdots,P_j,\cdots,P_N]$,$P_j=\left[\|F_1-F_j\|_2^2;\cdots;\|F_N-F_j\|_2^2\right]$,$F_j$表示$\boldsymbol{F}$的第$j$行。然后,可以得到

此时,模型式(14)可以写成

$$\mathrm{argmin}\ \frac{\alpha}{2m}\mathrm{tr}\left(\boldsymbol{P}^{\mathrm{T}}|\boldsymbol{Z}^{(v)}|+\boldsymbol{P}^{\mathrm{T}}|\boldsymbol{Z}^{(v)}|^{\mathrm{T}}\right)+\frac{\mu}{2}\left\|\boldsymbol{X}^{(v)}-\boldsymbol{X}^{(v)}\boldsymbol{Z}^{(v)}-\boldsymbol{E}^{(v)}+\frac{\boldsymbol{Y}^{(v)}}{\mu}\right\|_{\mathrm{F}}^{2}+\frac{\rho}{2}\left\|\boldsymbol{Z}^{(v)}-\boldsymbol{J}^{(v)}+\frac{\boldsymbol{W}^{(v)}}{\rho}\right\|_{\mathrm{F}}^{2}\tag{16}$$

模型式(16)的解为

$$\begin{aligned}\boldsymbol{Z}_t^{(v)}=&(\mu\boldsymbol{X}^{(v)T}\boldsymbol{X}^{(v)}+\rho\boldsymbol{I})^{-1}\cdot(\mu\boldsymbol{X}^{(v)\mathrm{T}}\boldsymbol{X}^{(v)}+\boldsymbol{X}^{(v)\mathrm{T}}\boldsymbol{Y}^{(v)}+\rho\boldsymbol{J}^{(v)})-\mu\boldsymbol{X}^{(v)\mathrm{T}}\boldsymbol{E}^{(v)}-\boldsymbol{W}^{(v)}-\\&\frac{\alpha}{2M}\left(\boldsymbol{P}\odot\mathrm{sign}(\boldsymbol{Z}_{t-1}^{(v)})+\boldsymbol{P}^{\mathrm{T}}\odot\mathrm{sign}(\boldsymbol{Z}_{t-1}^{(v)})^{\mathrm{T}}\right)\end{aligned}\tag{17}$$

子问题2:求解$\boldsymbol{E}^{(v)}$,变量$\boldsymbol{Z}^{(v)}$,$\boldsymbol{\mathcal{J}}$、$\boldsymbol{F}$是固定的。

$$\begin{aligned}&\arg\min_{E}\lambda\|\boldsymbol{E}\|_{2,1}+\sum_{v=1}^{m}\langle\boldsymbol{Y}_v,\boldsymbol{X}^{(v)}-\boldsymbol{X}^{(v)}\boldsymbol{Z}^{(v)}-\boldsymbol{E}^{(v)}\rangle+\sum_{v=1}^{m}\frac{\mu}{2}\|\boldsymbol{X}^{(v)}-\boldsymbol{X}^{(v)}\boldsymbol{Z}^{(v)}-\boldsymbol{E}^{(v)}\|_{\mathrm{F}}^{2}=\\&\arg\min_{E}\frac{\lambda}{\mu}\|\boldsymbol{E}\|_{2,1}+\frac{1}{2}\|\boldsymbol{E}-\boldsymbol{D}\|_{\mathrm{F}}^{2}\end{aligned}\tag{18}$$

模型最优解由文献[23]得到

$$E_{:,i}^{*}=\begin{cases}\dfrac{\|\boldsymbol{D}_{:,i}\|_2-\dfrac{\lambda}{\mu}}{\|\boldsymbol{D}_{:,i}\|_2}\boldsymbol{D}_{:,i} & \|\boldsymbol{D}_{:,i}\|_2>\dfrac{\lambda}{\mu}\\0 & \text{其他}\end{cases}\tag{19}$$

式中:$\boldsymbol{D}_{:,i}$表示$\boldsymbol{D}=[D^1;D^2;\cdots;D^m]$的第$i$列;$\boldsymbol{D}^j=\boldsymbol{X}^j-\boldsymbol{X}^j\boldsymbol{Z}^j+\dfrac{1}{\mu}\boldsymbol{Y}_j$,$j=1,2,\cdots,m$。

子问题3:求解$\boldsymbol{\mathcal{J}}$,变量$\boldsymbol{Z}^{(v)}$、$\boldsymbol{E}^{(v)}$、$\boldsymbol{F}$是固定的。

$$\boldsymbol{\mathcal{J}}^{*}=\arg\min_{\mathcal{J}}\|\boldsymbol{\mathcal{J}}\|_{\omega,S_p}^{p}+\langle\boldsymbol{\mathcal{Q}},\boldsymbol{\mathcal{Z}}-\boldsymbol{\mathcal{J}}\rangle+\frac{\rho}{2}\|\boldsymbol{\mathcal{Z}}-\boldsymbol{\mathcal{J}}\|_{\mathrm{F}}^{2}=\arg\ \min_{\mathcal{J}}\|\boldsymbol{\mathcal{J}}\|_{\omega,S_p}^{p}+\frac{\rho}{2}\left\|\boldsymbol{\mathcal{Z}}-\boldsymbol{\mathcal{J}}+\frac{\boldsymbol{\mathcal{Q}}}{\rho}\right\|_{\mathrm{F}}^{2}\tag{20}$$

最优解为

$$\boldsymbol{\mathcal{J}}^{*}=\Gamma_{\frac{1}{\rho}n_3\omega}\left(\boldsymbol{\mathcal{Z}}+\frac{1}{\rho}\boldsymbol{\mathcal{Q}}\right)\tag{21}$$

子问题4:求解F,变量$Z^{(v)}$、$E^{(v)}$、$\mathcal{J}$是固定的。

在这种情况下,式(15)变成

$$\begin{aligned} F &= \arg\min_{F} \operatorname{tr}\left(F^{\mathrm{T}} L_{\hat{Z}} F\right) \\ &\text{s.t. } F^{\mathrm{T}} F = I, F \in \mathrm{R}^{N \times c} \end{aligned} \tag{22}$$

通过模型推导分析,得出其伪代码如表1所示。

表1 TLR-MVSC伪代码

Table 1 TLR-MVSC pseudo-codes

算法 TLR-MVSC

输入:多视图聚类矩阵:$\boldsymbol{X}^{(1)},\boldsymbol{X}^{(2)},\cdots,\boldsymbol{X}^{(m)}$,参数$\alpha$、$\lambda$和类别数$K$;

输出:聚类结果C

初始化:$\boldsymbol{Z}^{(v)}=0,\boldsymbol{E}^{(v)},\boldsymbol{Y}^{(v)}=0,i=1,2,\cdots,m,\mathcal{J}=0,\mathcal{Q}=0,\mu=10^{-5},\rho=10^{-4},\mu_{\max}=\rho_{\max}=10^{10}$

while 不收敛 do:

1. 通过式(17)更新$\boldsymbol{Z}^{(v)},v=1,2,\cdots,m$;
2. 通过式(20)更新$\boldsymbol{E}$;
3. 通过$\boldsymbol{Y}^{(v)}=\boldsymbol{Y}^{(v)}+\mu(\boldsymbol{X}^{(v)}-\boldsymbol{X}^{(v)}\boldsymbol{Z}^{(v)}-\boldsymbol{E}^{(v)})$更新$\boldsymbol{Y}^{(v)},v=1,2,\cdots,m$;
4. 得到$\mathcal{Z}=\Theta(\boldsymbol{Z}^{(1)},\boldsymbol{Z}^{(2)},\cdots,\boldsymbol{Z}^{(m)})$;
5. 通过式(21)更新$\mathcal{J}$;
6. 通过$\mathcal{Q}=\mathcal{Q}+\rho(\mathcal{Z}-\mathcal{J})$更新$\mathcal{Q}$;
7. 通过式(22)更新$\boldsymbol{F}$;
8. 更新参数μ和ρ:$\min(\eta\mu,\mu_{\max}),\min(\eta\rho,\rho_{\max})$;
9. 得到$\boldsymbol{J}^{(1)},\boldsymbol{J}^{(1)},\cdots,\boldsymbol{J}^{(m)}=\Theta^{-1}(\mathcal{J})$;
10. 检查收敛条件

$\|\boldsymbol{X}^{(v)}-\boldsymbol{X}^{(v)}\boldsymbol{Z}^{(v)}-\boldsymbol{E}^{(v)}\|_{\infty}<\varepsilon$ & $\|\boldsymbol{Z}^{(v)}-\boldsymbol{J}^{(v)}\|_{\infty}<\varepsilon$

end while

11. 通过$\boldsymbol{S}=\frac{1}{m}\sum_{v=1}^{m}(|\boldsymbol{Z}^{(v)}|+|\boldsymbol{Z}^{(v)}|^{\mathrm{T}})$得到亲和度矩阵$\boldsymbol{S}$;
12. 在$\boldsymbol{S}$上进行谱聚类得到最终结果C

3 实验与结果分析

3.1 数据集

为了评估算法效能,在如表2所示类别数据集上进行实验,这些类别包括人脸图像、新闻文本、手写数字和通用对象场景,分别使用了广泛接纳的数据集进行验证。

(1) Yale:在本研究中,采用了由15位参与者提供的总计165张灰度图像作为数据集。每位参与者贡献了11张图像,这些图像通过引入不同的照明条件、面部表情及遮挡因素,展现了丰富的多样性。在实验过程中,该数据集被有效利用以评

表2 实验数据集

Table2 Experimental data sets

数据集	样本数量	类别	视图	种类
Yale	165	15	3	面部图像
ORL	400	40	3	面部图像
3Sources	169	6	3	新闻故事
BBCSport	685	5	4	新闻故事
UCI-Digits	2 000	10	3	手写教学
Scene-15	4 485	15	3	通用对象

估并比较不同计算机视觉算法的表现,利用文献[24]中的方法选择3种类型的特征作为不同的视图。

(2) ORL:该数据库由英国剑桥实验室开发,包含40位个体在不同时间点、照明条件、面部表情状态以及是否佩戴眼镜等变量下的共计400张面部图像。此数据库与Yale数据库均隶属于面部图像领域,并且通过提取强度、局部二值模式(LBP)及高斯小波包(Gabor)特征,实现了对于多视图数据的多元表示。

(3) 3Sources:利用了一个多元视角文本数据集,其包含了416篇独立的新闻报道,这些信息源自于BBC、路透社及《卫报》三家在线新闻媒体平台。所述数据集被细分为六大主题标签,具体包括商业、娱乐、健康、政治、体育及技术领域。这一数据集的特性在于其由3种不同类型的特征组成,形成了一个多层次、多维度的数据结构。

(4) BBCSport:归类于新闻故事领域,其源自于2004—2005年间BBCSport网站上所发布的5个主题领域的体育新闻文章,此资源广泛应用于机器学习研究作为评估基准。该数据集由685个包含5种不同特征的文档组成,共同构成了一个多视图数据集。

(5) UCI-Digits:此数据集由源自荷兰公共事业地图库的手写数字(范围为0至9)构成,共计10类阿拉伯数字,每类别包含200张图像。根据文献[35]的指引,从该数据集中提取了3个特征作为多视图数据的基础:傅里叶系数、像素均值以及形态特征。

(6) Scene-15[36]:该数据库共有15个场景类别,总计4 485张图像。所有的图像都来源于广泛的室内和室外环境,如工业、卧室、厨房、办公室、商店等。在实验中,通过文献[37]中的方法提取了3种类型的图像特征,包含1 240维的CENTRIST特征、1 800维的PHOW特征和1 180维的PRI-CoLBP 。

3.2 评价指标与对比算法

本次实验采用3项标准聚类评估指标,包括准确度(ACC)、归一化互信息(NMI)及纯度(Purity),以评测聚类效能。

ACC的定义为

$$\mathrm{ACC}=\frac{\sum_{i=1}^{N}\delta(b_i,\mathrm{map}(a_i))}{N} \tag{23}$$

NMI定义为

$$\mathrm{NMI}(C,C')=\frac{\mathrm{MI}(C,C')}{\max(H(C),H(C'))} \tag{24}$$

聚类评估指标Purity为

$$\mathrm{Purity}=\frac{n_1}{n_1+n_2} \tag{25}$$

式中:n_1表示一对样本在实际分类下同属一类,在预测分类下亦同属一类的样本数量;n_2则指在实际分类下非同类,在预测分类下却误判为同类的样本对数。

在多视图聚类算法中,张量化方法通过将多视图数据表示为张量并施加低秩约束,有效提取了视图之间的高阶相关性。例如,t-SVD-MSC[38]、AMGL[39]、ETLMSC[40]和WTNNM[41]通过不同的张量核范数约束来捕捉视图信息,减少噪声影响,而MLAN则自适应地调整图的权重以优化聚类效果。

通过矩阵旋转构建多视图子空间表示的张量,整合张量的低秩约束与多秩约束,以统一多视角自表示,优化聚类性能。

ETLMSC(Essential tensor learning for multi-view spectral clustering)[40],构建基于马尔可夫链的多视图转移概率矩阵为张量,通过张量奇异值分解(t-SVD)驱动的张量核范数约束,高效捕获多视图间的

高阶关联信息。

WTNNM, tensor-SVD基于图学习的多视图子空间聚类[41],引入了加权张量核范数与t-SVD结合,有效约束多视图子空间表示形成的张量结构,精确评估各视图信息价值,显著降低噪声干扰,优化聚类效果。

MLAN(Multi-view clustering and semi-supervised classification with adaptive neighbors)[42],不仅集成聚类与局部流形结构学习功能,还能自适应调整每张图的权重,最终生成的优化图结构能精确对应所需的类别数量。

3.3 计算复杂度与收敛性分析

计算复杂度主要在求解变量与迭代过程,具体为:$O(mN^2d_v)$,$O(mN^2\log(mN)+m^2N^2)$, $O(mN^2)$和$O(N^3)$,这里m代表视图数量,N为每个视图的样本数,d_v为第v个视图的维度。综合考虑迭代周期与特定条件后,迭代次数T以及条件$m\ll N$,本算法的总复杂度为$O(T(N^3+2mN^2d_v+mN^2\log(mN)))$。

计算复杂度的分析还不足以完全评估一个算法的实际性能,为了全面理解和验证所提方法的有效性,还需要对其收敛性进行深入探讨。当变量个数为2时,拉格朗日乘子法求解的收敛性得到良好证明[43];对于变量个数大于或等于3的情形,算法收敛性未能证实。鉴于此,本文通过实验验证模型收敛性,具体地,在6个数据集上评估其收敛表现。图2列出变量误差$\|\mathcal{Z}^{(v)}-\mathcal{J}^{(v)}\|_\infty$迭代次数与论文主题相关。由图可知,所有迭代在达到20次后稳定,证实方法有效收敛。实验结果显示,所提算法展现出优越的收敛性能。

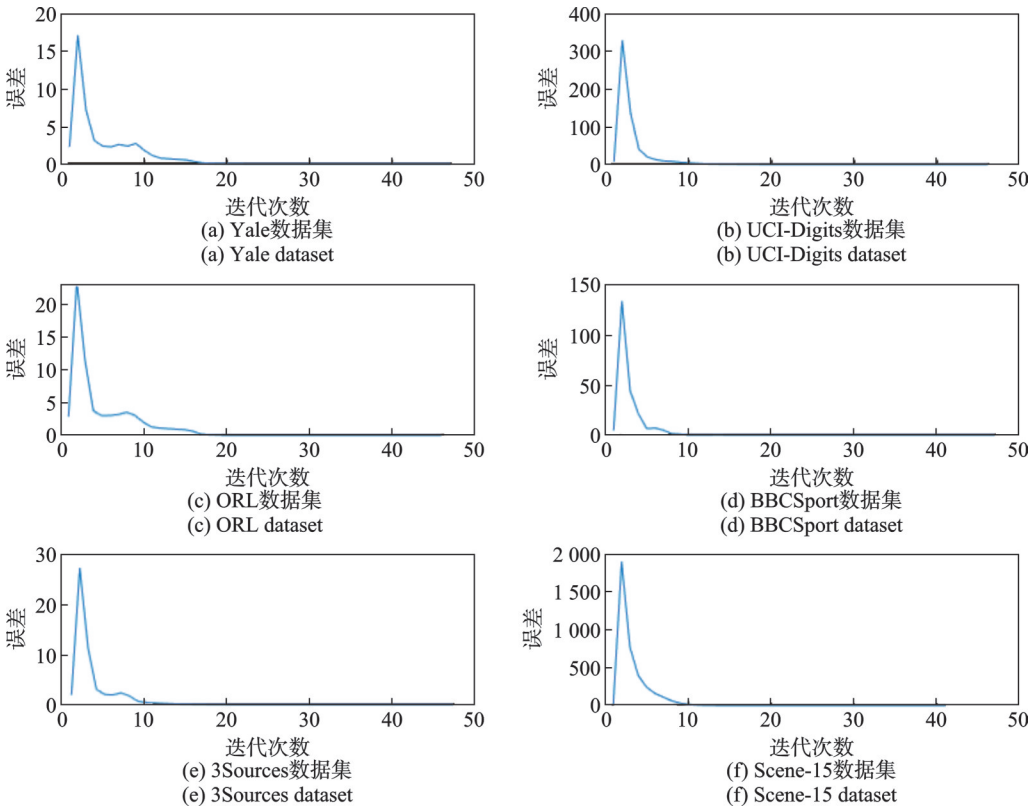

图2 在不同数据集上的收敛性实验

Fig.2 Convergence experiments on different datasets

3.4 实验参数说明

在实验参数设定过程中,为了确保所提出算法在不同数据集上的泛化能力以及能够取得最优的聚类效果,对关键参数α和λ进行了广泛的调整和优化。具体来说,参数α和λ的选择对于算法的性能至关重要,因此在实验中对这两个参数进行了系统的搜索,通过在预设的参数范围内不断调整并测试不同组合,以此来找出在各个数据集上表现最优的参数设置。

α在$[10^{-8}, 10^{-7}, 10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 1]$范围内调整,$\lambda$在$[0.1, 0.2, 0.4, 0.6, 0.8, 1]$范围内调整。具体来说,所提方法在Yale数据集上当$\alpha=10^{-7}$,$\lambda=0.1$时获得最好的性能;在UCI-Digits数据集上,当$\alpha=10^{-8}$,$\lambda=0.1$时获得最好的性能;在ORL数据集上,当$\alpha=10^{-7}$,$\lambda=0.5$时获得最好的性能;在BBCSport数据集上,当$\alpha=10^{-8}$,$\lambda=1$时获得最好的性能;在3Sources数据集上,当$\alpha=10^{-5}$,$\lambda=0.4$时获得最好的性能;在Scene-15数据集上,当$\alpha=10^{-8}$,$\lambda=0.1$时获得最好的性能。每个实验重复10次,表3~8列出了所有方法的各项评价指标。

表3 Yale数据集实验结果

Table 3 Experimental results on Yale dataset

方法	评价指标					
	ACC	NMI	Purity	*F*-score	Recall	AR
t-SVD-MSC	0.874±0.013	0.918±0.010	0.883±0.012	0.834±0.020	0.865±0.018	0.832±0.002
ETLMSC	0.659±0.042	0.697±0.038	0.659±0.043	0.533±0.044	0.550±0.048	0.501±0.021
WTNNM	0.832±0.000	0.882±0.000	0.911±0.000	0.861±0.000	0.831±0.000	0.840±0.000
MLAN	0.594±0.000	0.493±0.000	0.665±0.001	0.312±0.003	0.428±0.016	0.253±0.004
本文方法	0.981±0.000	0.977±0.000	0.981±0.000	0.963±0.000	0.964±0.000	0.960±0.000

表4 UCI-Digits数据集实验结果

Table 4 Experimental results on UCI-Digits dataset

方法	评价指标					
	ACC	NMI	Purity	*F*-score	Recall	AR
t-SVD-MSC	0.965±0.000	0.919±0.000	0.965±0.000	0.935±0.000	0.922±0.000	0.918±0.000
ETLMSC	0.942±0.000	0.902±0.000	0.942±0.000	0.905±0.000	0.887±0.000	0.879±0.000
WTNNM	0.984±0.000	0.962±0.000	0.982±0.000	0.972±0.000	0.970±0.000	0.970±0.000
MLAN	0.721±0.006	0.710±0.006	0.771±0.006	0.720±0.047	0.767±0.007	0.706±0.006
本文方法	0.995±0.000	0.986±0.006	0.995±0.000	0.990±0.000	0.990±0.000	0.988±0.000

表5 ORL数据集实验结果

Table 5 Experimental results on ORL dataset

方法	评价指标					
	ACC	NMI	Purity	*F*-score	Recall	AR
t-SVD-MSC	0.962±0.008	0.990±0.003	0.973±0.006	0.960±0.009	0.979±0.006	0.959±0.001
ETLMSC	0.958±0.024	0.931±0.005	0.970±0.016	0.960±0.020	0.984±0.010	0.959±0.002
WTNNM	0.982±0.00	0.991±0.000	0.991±0.000	0.990±0.000	0.990±0.000	0.990±0.000
MLAN	0.565±0.002	0.798±0.000	0.652±0.001	0.277±0.021	0.657±0.001	0.250±0.024
本文方法	0.986±0.000	0.991±0.000	0.967±0.000	0.990±0.000	0.990±0.000	0.988±0.000

表6 BBCSport数据集实验结果

Table 6 Experimental results on BBCSport dataset

方法	评价指标					
	ACC	NMI	Purity	*F*-score	Recall	AR
t-SVD-MSC	0.556±0.000	0.512±0.000	0.649±0.000	0.500±0.000	0.510±0.000	0.396±0.000
ETLMSC	0.642±0.000	0.610±0.000	0.739±0.000	0.421±0.001	0.410±0.003	0.390±0.000
WTNNM	0.573±0.000	0.531±0.000	0.655±0.000	0.512±0.000	0.520±0.000	0.414±0.000
MLAN	0.243±0.000	0.395±0.000	0.420±0.001	0.110±0.000	0.480±0.003	0.009±0.001
本文方法	0.867±0.000	0.865±0.000	0.904±0.000	0.841±0.000	0.845±0.000	0.869±0.000

表7 3Sources数据集实验结果

Table 7 Experimental results on 3Sources dataset

方法	评价指标					
	ACC	NMI	Purity	*F*-score	Recall	AR
t-SVD-MSC	0.558±0.004	0.438±0.009	0.624±0.012	0.498±0.000	0.509±0.000	0.395±0.000
ETLMSC	0.529±0.008	0.610±0.002	0.565±0.000	0.739±0.000	0.617±0.000	0.539±0.000
WTNNM	0.841±0.000	0.900±0.001	0.627±0.000	0.810±0.000	0.833±0.000	0.810±0.000
MLAN	0.172±0.001	0.154±0.001	0.192±0.001	0.475±0.000	0.592±0.000	0.347±0.000
本文方法	0.884±0.000	0.938±0.000	0.865±0.002	0.837±0.000	0.834±0.000	0.825±0.000

表8 Scene-15数据集实验结果

Table 8 Experimental results on Scene-15 dataset

方法	评价指标					
	ACC	NMI	Purity	*F*-score	Recall	AR
t-SVD-MSC	0.892±0.000	0.919±0.000	0.922±0.000	0.883±0.000	0.892±0.000	0.874±0.000
ETLMSC	0.871±0.000	0.891±0.000	0.906±0.000	0.853±0.000	0.892±0.000	0.842±0.000
WTNNM	0.901±0.001	0.931±0.000	0.931±0.000	0.901±0.000	0.833±0.000	0.891±0.000
MLAN	0.342±0.003	0.493±0.003	0.354±0.002	0.262±0.003	0.740±0.043	0.167±0.045
本文方法	0.971±0.000	0.991±0.000	0.975±0.000	0.987±0.000	0.987±0.000	0.986±0.000

3.5 实验结果与分析

实验结果表明，基于张量的聚类方法（如t-SVD-MSC、ETLMSC、WTNNM和TLR-MVSC）在多视图数据处理中表现出的聚类性能显著优于非张量聚类方法（如MLAN）。这种优势主要归因于非张量方法在处理多视图数据时，无法最大化利用嵌入于不同视图间的视图相似度矩阵所蕴含的互补信息与空间结构信息，从而导致聚类效果的限制。例如在Scene-15数据集上，基于张量的聚类方法相较非张量方法的聚类效果至少提升了25%，这表明，对于多视图数据中视图间存在较大差异的情况，张量方法能够更好地捕捉视图间的关联性和互补性，从而显著提升聚类性能。

在多种基于张量的聚类方法中，本文方法在Scene-15数据集上，与t-SVD-MSC比较，各项指标均分别实现提升9%、8%、6%、12%、11%和13%。这一优势主要源于本文方法充分考虑了不同视图信息的重要性差异。同时，本文提出的张量核范数最小化策略，结合了奇异值差异性的分析和动态权重调

整机制,不仅提升了模型的鲁棒性,还在多视图数据复杂性处理方面实现了更高的准确性和一致性。

实验还探讨了核范数最小化参数p对聚类效果的调节作用,发现其对不同数据集的性能优化具有显著影响。图3为本文方法在不同数据集上的应用,显示了其多样化的聚类效果。从图表分析可见,在值变动下,该方法的聚类效能展现出一定的波动性。在保持其他参数不变的前提下,实验结果显示,Yale数据集在$p=0.5$时达到最佳聚类效果,而ORL数据集在$p=0.9$时表现最优秀。这种现象表明,参数p对奇异值间差异性的调节能够有效强化视图间的互补信息利用,显著降低了光照等外部因素对聚类效果的干扰。通过优化该参数,本文方法成果实现了聚类性能的全面提升,进一步验证了其在多视图数据处理中适应性和鲁棒性上的显著优势。

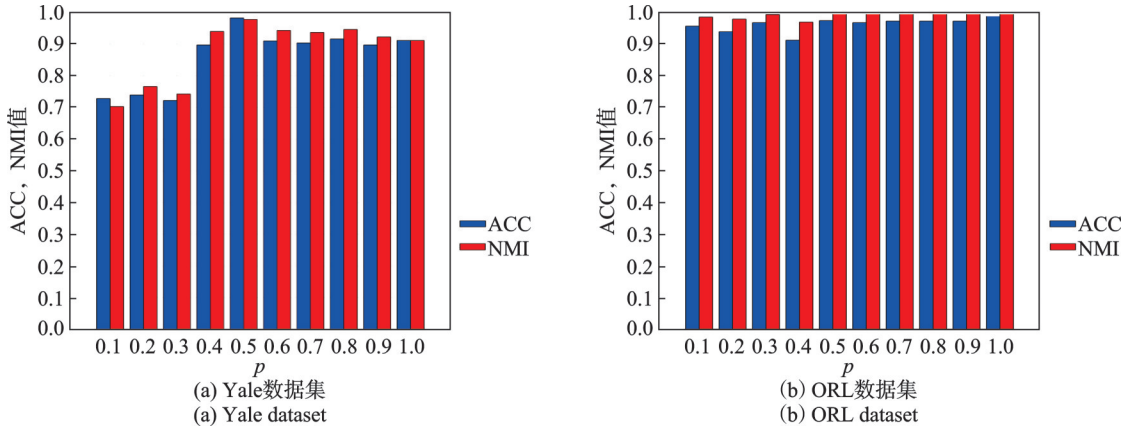

图3 参数p对聚类效果的影响

Fig.3 Effect of parameter p on clustering effect

4 结束语

提出了一种基于张量低秩学习的多视图子空间聚类方法,以解决现有多视图聚类方法中忽视表示张量与亲和度矩阵高度相关性以及在处理数据缺失、噪声和异常值时表现不足的问题。通过施加低秩约束,该方法有效地挖掘了数据点之间的高阶关联性,从而增强了对数据子空间结构的识别精度,并提高了亲和度矩阵的鲁棒性。同时,采用基于张量奇异值分解的核范数约束技术,对亲和度矩阵的求解过程进行优化。实验结果表明,该方法比现有方法在四类基准数据集中的表现更好,并验证了其有效性和处理多视图数据的先进性。在未来工作中,将尝试更多方法来挖掘数据点间的高阶关联性来进一步提升性能。

参考文献:

[1] 陈曼笙,蔡晓莎,林家祺,等.张量学习诱导的多视图谱聚类[J].计算机学报,2024,47(1):52-68.
CHEN Mansheng, CAI Xiaosha, LIN Jiaqi, et al. Tensor learning-induced multi-view spectral clustering[J]. Chinese Journal of Computers, 2024, 47(1): 52-68.

[2] LIU W, TSANG I W. Making decision trees feasible in ultrahigh feature and label dimensions[J]. Journal of Machine Learning Research, 2017, 18(1): 2814-2849.

[3] LIU W, TSANG I W, MULLER K. An easy-to-hard learning paradigm for multiple classes and multiple labels[J]. Journal of Machine Learning Research, 2017, 18: 1-38.

[4] ZHANG C, LI H, LV W, et al. Enhanced tensor low-rank and sparse representation recovery for incomplete multi-view clustering[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2023, 37(9): 11174-11182.

[5] WANG Q, CHENG J, GAO Q, et al. Deep multi-view subspace clustering with unified and discriminative learning[J]. IEEE Transactions on Multimedia, 2020, 23: 3483-3493.

[6] GAO Q, XIA W, WAN Z, et al. Tensor-SVD based graph learning for multi-view subspace clustering[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(4): 3930-3937.

[7] KANG Z, ZHOU W, ZHAO Z, et al. Large-scale multi-view subspace clustering in linear time[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(4): 4412-4419.

[8] YANG M, LUO Q, LI W, et al. Multiview clustering of images with tensor rank minimization via nonconvex approach[J]. SIAM Journal on Imaging Sciences, 2020, 13(4): 2361-2392.

[9] CHEN Y, XIAO X, ZHOU Y. Multi-view subspace clustering via simultaneously learning the representation tensor and affinity matrix[J]. Pattern Recognition, 2020, 106: 107441.

[10] WU J, XIE X, NIE L, et al. Unified graph and low-rank tensor learning for multi-view clustering[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(4): 6388-6395.

[11] SI X, YIN Q, ZHAO X, et al. Consistent and diverse multi-view subspace clustering with structure constraint[J]. Pattern Recognition, 2022, 121: 108196.

[12] SHAO W, HE L, YU P S. Multiple incomplete views clustering via weighted nonnegative matrix factorization with l2, 1 regularization[C]//Proceedings of Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Porto: Springer, 2015:318-334.

[13] GAO H, PENG Y, JIAN S. Incomplete multi-view clustering[C]//Proceedings of Intelligent Information Processing VIII: 9th IFIP TC 12 International Conference. Australia: Springer, 2016: 245-255.

[14] WANG H, ZONG L, LIU B, et al. Spectral perturbation meets incomplete multi-view data[C]//Proceedings of the 28th International Joint Conference on Artificial Intelligence. [S.l.]: [s.n.], 2019: 3677-3683.

[15] HU M, CHEN S. Doubly aligned incomplete multi-view clustering[J]. arXiv preprint arXiv:1903.02785, 2019: 2262-2268.

[16] LIU J, TENG S, FEI L, et al. A novel consensus learning approach to incomplete multi-view clustering[J]. Pattern Recognition, 2021: 115(2): 107890.

[17] LI X L, CHEN M S, WANG C D, et al. Refining graph structure for incomplete multi-view clustering[J]. IEEE Transactions on Neural Networks and Learning Systems, 2022, 12: 2300-2313.

[18] 崔金荣,黄诚.改进的自步深度不完备多视图聚类[J].数据采集与处理,2022,37(5):1036-1048.
CUI Jinrong, HUANG Cheng. Improved self-paced deep incomplete multi-view clustering [J]. Journal of Data Acquisition & Processing, 2022, 37(5): 1036-1048.

[19] LIN Y, GOU Y, LIU Z, et al. Completer: Incomplete multi-view clustering via contrastive prediction[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE. 2021: 11174-11183.

[20] WANG L, YANG Y, LIU B. GMC: Graph-based multi-view clustering[J]. IEEE Transactions on Knowledge and Data Engineering, 2020, 32(6): 1116-1129.

[21] KANG Z, SHI G, HUANG S, et al. Multi-graph fusion for multi-view spectral clustering[J]. Knowledge-Based Systems, 2020, 189: 105102.

[22] ZHANG C, HU Q, FU H, et al. Latent multi-view subspace clustering[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017: 4279-4287.

[23] 陈梅,马学艳,钱罗雄,等.基于多级自表示约束的不完备多视图聚类[J].控制与决策,2025,40(2): 645-654.
CHEN Mei, MA Xueyan, QIAN Luoxiong, et al. Incomplete multi-view clustering based on multi-level self-representation constraints[J]. Control and Decision, 2025, 40(2): 645-654.

[24] 吴杰,万源,刘秋杰.一致块对角和限定的多视角子空间聚类算法[J].计算机科学,2025(4): 138-146.
WU Jie, WAN Yuan, LIU Qiujie. Multi-view subspace clustering algorithm with consistent block diagonal and boundary[J]. Computer Science, 2025(4): 138-146.

[25] NG A Y, JORDAN M I, WEISS Y. On spectral clustering: Analysis and an algorithm[C]//Proceedings of the 14th International Conference on Neural Information Processing Systems. Cambridge: MIT Press, 2002: 849-856.

[26] LIU G C, YAN S C. Latent low-rank representation for subspace segmentation and feature extraction[C]//Proceedings of 2011

IEEE 13th International Conference on Computer Vision. [S.l.]: IEEE, 2011:1615-1622.

[27] LIU G C, LIN Z C, YU Y. Robust subspace segmentation by low-rank representation[C]//Proceedings of the 27th International Conference on Machine Learning. Haifa: OMNI Press, 2010: 663-670.

[28] XIA G Y, SUN H J, FENG L, et al. Human motion segmentation via robust kernel sparse subspace clustering[J]. IEEE Transactions on Image Processing, 2018, 27(1): 105-135.

[29] JI P, ZHONG Y R, LI H D, et al. Null space clustering with applications to motion segmentation and face clustering[C]// Proceedings of the 2014 IEEE International Conference on Image Processing. Piscataway: IEEE, 2014: 283-287.

[30] MA Y, DERKSEN H, HONG W, et al. Segmentation of multivariate mixed data via lossy data coding and compression[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 29(9):1546-1562.

[31] ZHANG P, LIU X, XIONG J, et al. Consensus one-step multi-view subspace clustering[J]. IEEE Transactions on Knowledge and Data Engineering, 2020, 34(10): 4676-4689.

[32] 陈梦瑶,陈秀宏.多样性特征增强与低秩表示的多视图聚类算法[J].小型微型计算机系统,2025, 45(6): 1391-1399.
CHEN Mengyao, CHEN Xiuhong. Multi-view clustering algorithm based on diversity feature enhancement and low-rank representation[J]. Journal of Chinese Computer Systems, 2025, 45(6): 1391-1399.

[33] HŸTCH M, HAWKES P W. Morphological image operators[M]. [S.l.]: Academic Press, 2020.

[34] CHEN M S, HUANG L, WANG C D. et al. Multi-view clustering in latent embedding space[J]. Proceedings of AAAI, 2020, 34(4): 3513-3520.

[35] ZHANG F, WANG W, HUANG J, et al. RIP-based performance guarantee for low-tubal-rank tensor recovery[J]. Journal of Computational and Applied Mathematics, 2020, 374: 112767.

[36] WANG H, YANG Y, LIU B. GMC: Graph-based multi-view clustering[J]. IEEE Transactions on Knowledge and Data Engineering,2020, 32(6): 1116-1129.

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

[38] XIE Y, TAO D, ZHANG W, et al. On unifying multi-view self-representations for clustering by tensor multi-rank minimization[J]. International Journal of Computer Vision, 2018, 126(11): 1157-1179.

[39] WANG R, NIE F, WANG Z, et al. Parameter-free weighted multi-view projected clustering with structured graph learning[J]. IEEE Transactions on Knowledge and Data Engineering, 2019, 32(10): 2014-2025.

[40] WU J, LIN Z, ZHA H. Essential tensor learning for multi-view spectral clustering[J]. IEEE Transactions on Image Processing, 2019, 28(12): 5910-5922.

[41] GAO Q, XIA W, WAN Z, et al. Tensor-SVD based graph learning for multi-view subspace clustering[C]//Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence. [S.l.]: AAAI, 2020: 2153-2159.

[42] NIE F, CAI G, LI X. Multi-view clustering and semi-supervised classification with adaptive neighbours[C]//Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence. [New York]: AAAI, 2017: 2408-2414.

[43] LIN Z, CHEN M, MA Y. The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices[J]. arXiv preprint arXiv:1009.5055.

作者简介:

史德胜(1997-),男,硕士研究生,研究方向:大数据分析与处理,E-mail: 1222045606@njupt.edu.cn。

徐鹤(1985-),**通信作者**,男,博士,教授,研究方向:人工智能与大数据、物联网技术,E-mail: xuhe@njupt.edu.cn。

李鹏(1979-),男,博士,教授,博士生导师,研究方向:高性能计算、大数据、人工智能和网络安全,E-mail: lipeng@njupt.edu.cn。

(编辑:夏道家)

Multi-view Subspace Clustering Method Based on Tensor Low-Rank Learning

SHI Desheng[1], XU He[1,2*], LI Peng[1,2]

(1. School of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing 210023, China;2. Jiangsu HPC and Intelligent Processing Engineer Research Center, Nanjing 210023, China)

Abstract: Multi-view clustering is a powerful technique for improving analytical performance by fusing complementary multi-source information. However, there are deficient in two ways: It neglects the strong inherent correlation between representation tensors and affinity matrices, and the separate two-step strategy of representation learning and clustering leads to lack of association between these processes, rendering inefficient in handling missing data, noise and outliers in multi-view data processing. In order to address these issues, this paper proposes a multi-view sub-space clustering method based on tensor low-rank learning. A methodology is put forward for the analysis of high-order correlations among data points and the identification of the intrinsic structure of the data. The method involves the introduction of a high-order tensor constraint based on low-rank representation (LRR) and the adoption of tensor nuclear norm minimization (TNNM) based on tensor singular value decomposition (t-SVD). This approach facilitates the transformation of the original non-convex optimization problem into a solvable convex one. The application of an adaptive weighted Schatten-p norm has been utilized to capture the inherent differences between singular values, with the assistance of their prior information. Spectral clustering has been integrated into a unified framework for the purpose of optimizing the affinity matrix, with a view to more effectively characterizing clustering structures. The inexact augmented Lagrange multiplier (ALM) method has been utilized to decompose the model into four solvable sub-problems for the purpose of efficient optimization. Comprehensive experiments are conducted on six benchmark datasets spanning facial images, news stories, handwritten digits and general objects, with systematic optimization of key parameters to ensure reliability. The findings demonstrate that the proposed method exhibits a substantial enhancement in performance when compared to four contemporary algorithms, namely t-SVD-MSC, ETLMSC, WTNNM and MLAN. The proposed method demonstrated an accuracy of 0.981 on the Yale dataset, 0.995 on the UCI-Digits dataset, and 0.971 on the Scene-15 dataset. The proposed method effectively increases the robustness of the affinity matrix against noise and outliers. It accurately extracts the intrinsic subspace structure of multi-view data and demonstrates excellent practical performance and strong generalization ability in the analysis of high-dimensional and incomplete multi-view data.

Highlights

1. This paper proposes a method of representing data as a high-order, tensor-constrained, low-rank matrix, which can be used to analyse high-order correlations between data points. It can also effectively enhance the robustness of affinity matrices against complex noise and various outliers in multi-view data.
2. This paper uses T-SVD-based TNNM and an adaptive weighted Schatten-p norm to optimize affinity matrices. This approach captures the inherent differences between singular values and exploits their key structural information in tensor data.
3. This paper proposes a unified tensor low-rank learning framework that integrates spectral clustering techniques to optimize the fusion similarity matrix. The method demonstrates strong clustering performance across six classic benchmark datasets spanning four data types, with significant improvements over previous algorithms in evaluation metrics.

Key words: multi-view clustering; tensor low-rank; subspace clustering; high-order correlation; tensor decomposition

Foundation item: National Key Research and Development Program of China (No.2019YFB2103003).
Received: 2024-09-20; **Revised:** 2024-12-12
***Corresponding author, E-mail:** xuhe@njupt.edu.cn.