

# 张量学习诱导的多视图谱聚类

陈曼笙<sup>1)</sup> 蔡晓莎<sup>2)</sup> 林家祺<sup>2)</sup> 王昌栋<sup>1,3)</sup> 黄 栋<sup>4)</sup> 赖剑煌<sup>1)</sup>

<sup>1)</sup>(中山大学计算机学院 广州 510006)

<sup>2)</sup>(中山大学数学学院(珠海) 广东 珠海 519000)

<sup>3)</sup>(广东省知识产权大数据重点实验室 广州 510006)

<sup>4)</sup>(华南农业大学数学与信息学院 广州 510642)

**摘 要** 现有的方法将通过张量奇异值分解(t-SVD)正则化的低秩表示应用到多视图子空间聚类中,取得了令人印象深刻的聚类性能.然而,它们都具有以下两个共同的缺点:(1)他们专注于探索样本之间的关系以构建表征,然后将其堆叠为张量,其计算复杂度至少为 $O(n^2 \log n)$ ;(2)他们总是直接在整合的表征上运行标准的谱聚类算法,而忽略了不同表征对最终聚类结果的先验知识.为了解决这些问题,本文提出了一种新颖的张量学习诱导的多视图谱聚类(TLIMSC)方法,其中同时探索了空间聚类结构和互补信息.具体来说,该方法将关联样本和簇关系的多视图谱嵌入表示堆叠成张量,计算复杂度最终变为 $O(n \log n)$ .然后,将学习到的带有不同自适应置信度的表征与最终的一致聚类结果联系起来.在五个数据集上的广泛实验证明了 TLIMSC 所具有的有效性和高效性.

**关键词** 多视图聚类;加权张量核范数;谱嵌入表征;自适应置信度

**中图法分类号** TP18 **DOI号** 10.11897/SP.J.1016.2024.00052

## Tensor Learning Induced Multi-View Spectral Clustering

CHEN Man-Sheng<sup>1)</sup> CAI Xiao-Sha<sup>2)</sup> LIN Jia-Qi<sup>2)</sup> WANG Chang-Dong<sup>1,3)</sup>

HUANG Dong<sup>4)</sup> LAI Jian-Huang<sup>1)</sup>

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

<sup>2)</sup>(School of Mathematics (Zhuhai), Sun Yat-sen University, Zhuhai, Guangdong 519000)

<sup>3)</sup>(Guangdong Provincial Key Laboratory of Intellectual Property and Big Data, Guangzhou 510006)

<sup>4)</sup>(College of Mathematics and Informatics, South China Agricultural University, Guangzhou 510642)

**Abstract** Low-rank representation coefficients regularized by the tensor-Singular Value Decomposition (t-SVD) scheme for multi-view subspace clustering have achieved impressive performances. However, all of them suffer from the following two common demerits. (1) They focus on exploring the relationships among samples to construct representations which are then stacked to be a tensor, whose computational complexity is at least  $O(n^2 \log n)$ ; (2) They always deploy the standard spectral clustering algorithm directly on the integrated representation, neglecting the prior knowledge of different representations towards the final results. To tackle these problems, we propose a novel Tensor Learning Induced Multi-view Spectral Clustering (TLIMSC) approach, where the spatial cluster structures and complementary information are simultaneously explored. Specifically, multi-view spectral embedding representations related from samples to clusters are focused to be stacked in a tensor, where the complexity finally becomes  $O(n \log n)$ . Later, a bridge would be

收稿日期:2022-11-28;在线发布日期:2023-07-19. 本课题得到国家自然科学基金(62276277)、广东省自然科学基金(2022B1515120059)、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, 主要研究方向为数据挖掘、多视图聚类. E-mail: chenmsh27@mail2.sysu.edu.cn. 蔡晓莎, 博士研究生, 主要研究方向为数据挖掘、多视图聚类. 林家祺, 博士研究生, 主要研究方向为数据挖掘、多视图聚类. 王昌栋(通信作者), 博士, 副教授, 主要研究方向为数据聚类、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, 主要研究方向为人工智能、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.

built to connect the learned representations carrying different adaptive confidences with the final consensus results. Extensive experiments on five datasets reveal the effectiveness and efficiency of TLIMSC.

**Keywords** multi-view clustering; weighted tensor nuclear norm; spectral embedding representations; adaptive confidences

## 1 引言

现普遍存在于大多数应用中的多视图数据,指的是从多个特征提取器或不同的域得到的数据<sup>[1-4]</sup>. 例如,可以使用不同种类的特征来描述图像,如颜色、边缘和纹理信息,每种特征可以视为一个特定的视图. 通过结合多个视图可进行数据分析,由此产生了多视图学习<sup>[5]</sup>. 多视图学习有两个至关重要的原则<sup>[6]</sup>,即共识原则和互补原则. 共识的目的是使多个不同视图之间的一致性最大化;互补则意味着数据的每个视图都包含一些其他视图所没有的信息.

多视图聚类是研究最广泛的多视图学习任务之一<sup>[7-12]</sup>. 通常多视图聚类有两个基本假设<sup>[13]</sup>,一方面,每个单一视图中的特征都足以捕获大部分聚类信息;另一方面,这些多视图特征中可能存在一些噪声或损坏,这可能导致一小部分数据点被错误聚类. 由于每个单独的视图都分别代表了数据的某种特征<sup>[14-16]</sup>,多视图聚类的关键挑战是如何整合多个视图并结合这些表征的潜在语义信息(如高阶信息)来探索数据内在的空间结构以获得更好的聚类性能. 处理多视图数据最简单的策略是将多个特征空间级联成一个新的单个特征表示,在此基础上采用单视图聚类方法(如谱聚类)来实现最终聚类. 但它并没有很好地考虑多个视图之间的互补信息,也没有发现潜在的语义信息<sup>[17]</sup>. 因此,如何在多视图聚类中很好地整合多个视图的信息仍然是一项具有挑战性的任务.

在过去几年中,研究人员致力于研究有效的多视图聚类方法,如协同正则化多视图聚类<sup>[18]</sup>、多视图  $k$ -均值聚类<sup>[19-20]</sup>、基于低秩表示的多视图聚类<sup>[21-23]</sup>等. 其中,基于低秩表示的多视图聚类由于性能优越,已成为多视图聚类方法的主流<sup>[3,13,24-25]</sup>. 新兴的基于张量奇异值分解(tensor Singular Value Decomposition, t-SVD)的张量核范数技术已被证明是  $l_1$  范数的有效凸松弛. 该技术可以挖掘输入的多视图数据的空间结构和高阶信息,从而进一步提高聚类

性能<sup>[13,26-30]</sup>. 例如,受低秩表征(Low-Rank Representation, LRR)方法<sup>[31]</sup>的启发,Zhang 等人<sup>[22]</sup>提出了低秩张量约束多视图子空间聚类(Low-rank Tensor Constrained Multi-view Subspace Clustering, LTMSC)方法,该方法侧重于最小化通过子空间表示构建的张量展开矩阵的核范数. 然而 LTMSC 的塔克张量分解不是塔克低秩的紧凸松弛,为此,Xie 等人<sup>[13]</sup>提出了基于 t-SVD 的多视图子空间聚类(t-SVD based Multi-view Subspace Clustering, tMSC)方法. 在 tMSC 的基础上,Gao 等人<sup>[28]</sup>又进一步提出了研究不同奇异值的不同正则化的多视图子空间聚类(t-SVD based Weighted Tensor Nuclear Norm Minimization, WTNNM)方法. 尽管这些方法取得了巨大的成功,但它们也有一些不足之处. 一方面,它们需要花费大量时间进行算法学习,其计算复杂度至少为  $O(n^2 \log n)$ ,其原因在于它们利用学习到的子空间表征构建张量,导致  $n$  个样本之间的关系都需要被计算. 另一方面,在张量学习之后,标准的谱聚类或  $k$ -均值算法被直接用于处理之前学习到的综合表征. 然而,这样的做法是有缺陷的,因为每个视图学习得到的表征对最终的聚类性能有不同的贡献,所以在学习过程中不应该被同等地对待.

为了解决上述挑战,本文提出了一种新颖的张量学习诱导的多视图谱聚类(Tensor Learning Induced Multi-view Spectral Clustering, TLIMSC)方法,可以同时挖掘空间聚类结构和互补信息. 具体来说,受锚点图的启发,TLIMSC 探讨了  $n$  个样本和  $c$  个簇之间的表征关系,并从多个视图中学习到了对应的谱嵌入矩阵. 然后,TLIMSC 利用多个谱嵌入表征来构建张量,通过基于 t-SVD 的加权张量核范数来发掘从样本到簇之间的高阶关联,其中该范数在张量学习中考虑了矩阵不同奇异值之间的先验知识. 此外,在获得具有高阶相关性的表征后,在最终的共识聚类结果和表征之间建立联系,自适应地学习不同表征对最终聚类结果的不同贡献. 因此,本文的主要目的是研究一种有效且高效的多视图聚类方法,不仅可以通过基于 t-SVD 的加权张量核范

数来捕获样本和簇之间的高阶关联,还考虑了多个表征对最终聚类结果的不同置信度.

本文的主要贡献总结如下:

(1) 本文不直接探索样本之间的关系,而是研究样本和簇之间的关系,即使用谱嵌入表征来构建张量,以此挖掘其中的高阶关联关系,其计算复杂度为  $O(n \log n)$ .

(2) 本文发掘了不同谱嵌入矩阵对最终共识聚类结果的不同贡献度,并在它们之间建立了相应的桥梁.

(3) 本文提出了一种交替最小化优化算法来解决所提出的 TLIMSC 方法. 在四个数据集上进行的大量实验表明,与最先进的多视图聚类方法相比,我们的方法具有优越性.

本文在第 2 节中简要介绍一些相关工作;第 3 节描述所提出的张量学习诱导的多视图谱聚类方法,其中包括文中使用的主要符号、所提出方法的优化算法和复杂性分析;在第 4 节中,记录四个数据集的实验结果;最后,本文在第 5 节中得出结论.

## 2 相关工作

近年来,通过探索多视图特征之间的关系来进行多视图聚类的方法被广泛研究<sup>[28,32-34]</sup>. 这些方法可以大致分为三大类<sup>[6]</sup>,即基于协同训练或协同正则化的方法、基于图的方法和基于子空间学习的方法.

对于基于协同训练或协同正则化的方法<sup>[18,32,35-37]</sup>,每个视图都有一个单独的学习器,且每个学习器的信息都被用来约束其他视图. 在文献<sup>[35]</sup>中,该工作在不同的视图中互换了分类信息. 受协同训练的启发,Kumar 等人<sup>[36]</sup>利用一个视图的谱嵌入来约束其他视图的邻接矩阵. 随后,Kumar 等人<sup>[18]</sup>提出了结合多个视图隐含图的谱聚类模型,以获得更好的聚类性能. Cao 等人<sup>[32]</sup>引入了希尔伯特-施密特独立准则(Hilbert Schmidt Independence Criterion, HSIC)作为正则化项来探索不同视图之间的互补信息. Cao 等人<sup>[37]</sup>考虑了视频人脸图像的成对约束和跨视图的一致性来对视频人脸图像进行聚类,其中一个共正则化项被用于规范不同视图之间的分歧. 通过 HSIC,Wang 等人<sup>[38]</sup>探索了视图之间的相似性,以便拥有与完整样本的最大依赖性.

基于图的方法通过视图融合的方式来发现多个视图之间的关系<sup>[17,21,39-40]</sup>. 基于马尔科夫链,Xia 等人<sup>[21]</sup>提出了一种鲁棒的多视图谱聚类(Robust

Multi-view Spectral Clustering, RMSC)方法. 该方法通过低秩和稀疏分解的共享低秩转移概率矩阵来处理与多视图相关的噪声. Zhan 等人<sup>[41]</sup>提出了一种用于多视图聚类的图学习方法,该方法首先用低秩约束来约束每个视图的初始图,然后通过优化过程将这些初始图组合得到一个全局图. Shu 和 Latecki<sup>[39]</sup>构建了跨视图张量积图来探索不同视图之间的高阶信息. Nie 等人<sup>[17]</sup>从多个原始输入图中发现了一个块对角矩阵,并为多视图聚类设计了一种新的权重学习策略. Wang 等人<sup>[40]</sup>将与单个视图相关的相似性诱导图、所有视图的统一图和聚类指示矩阵整合到一个统一的框架中. Nie 等人<sup>[42]</sup>设计了一个新颖的多视图学习模型,其中聚类/半监督分类以及局部结构学习是同时进行的. 在某种程度上,基于图的方法与基于多核学习的方法<sup>[43]</sup>相关,后者将每个视图视为给定的核矩阵,用一定的权重整合多个核,并进行数据划分.

基于所有的视图都是源于一个潜在子空间的假设,基于子空间学习的方法旨在发现共享的潜在子空间<sup>[3,22,24-25,28,32,44]</sup>,并在此基础上进行聚类. Zhang 等人<sup>[25]</sup>从多个视图中寻找潜在表征并进行数据重构. 进一步地,为了处理多个样本之间的一般相关性,Zhang 等人<sup>[3]</sup>设计了一个基于神经网络的更泛化的框架. Li 等人<sup>[33]</sup>通过将原始特征映射到一个核空间,从多个视图中发现了一个潜在的共同空间. Luo 等人<sup>[45]</sup>在多视图子空间聚类中同时学习了子空间表示的一致性和独特性. Tang 等人<sup>[46]</sup>通过同时考虑多样性正则化和低秩约束,研究了用于多视图子空间聚类的共享相似性表示. Zhou 等人<sup>[47]</sup>试图捕获来自多个视图的共享信息之间的相关性,同时通过多种信息恢复来自不同视图的不同属性. 在工作<sup>[48]</sup>中,Zhang 等人设计了一个基于深度矩阵分解和分块对齐的多视图聚类方法,其中每个视图对应的聚类表征是通过深度矩阵分解来得到的,继而通过整合多视图信息去得到最佳的聚类表征. 在基于 LRR<sup>[31]</sup>的子空间聚类方法的基础上,Zhang 等人<sup>[22]</sup>研究了一种多视图子空间方法,对张量施加了基于展开的低秩约束,而该张量是由不同视图的子空间表示所构建的. 然而,这种张量约束是对塔克秩的松弛近似,对一般张量缺乏明确的物理意义<sup>[28]</sup>. 因此,它不能彻底挖掘多视图之间的互补信息,这进一步促成了一种新的、紧凑的张量分解方案,即 t-SVD<sup>[13]</sup>.

依据文献<sup>[49-50]</sup>,通过 t-SVD 分解张量具有良好的理论特性和明确的物理意义. Xie 等人<sup>[13]</sup>提出了

一种基于 t-SVD 的多视图子空间聚类方法 (tMSC), 他们通过张量多秩来约束旋转子空间表示张量, 以探索高阶跨视图的相关性. 作为 RMSC<sup>[21]</sup> 的扩展, Wu 等人<sup>[26]</sup> 通过不同视图的多个状态转移矩阵来构建张量. 考虑到矩阵奇异值的先验知识, Gao 等人<sup>[28]</sup> 探索了基于 t-SVD 的加权张量核范数最小化 (WTNNM) 方法, 其中不同的矩阵奇异值以不同的置信度被收缩. Chen 等人<sup>[29]</sup> 提出在一个统一的框架内学习低秩表征张量和相似性矩阵, 该框架同时考虑了局部特征和多视图特征对最终相似性矩阵的不同置信度. Chen 等人<sup>[51]</sup> 联合研究核表示张量和相似性矩阵. 尽管上述基于张量核范数的多视图子空间方法取得了较为优异的性能, 但它们都集中在了刻画  $n$  个样本到  $n$  个样本关系的表征构建的张量学习中, 其计算复杂度至少为  $O(n^2 \log n)$ . 特别地, 除了基于 WTNNM 的方法<sup>[28]</sup>, 它们都不能很好地探索不同矩阵奇异值的先验知识. 此外, 它们都是直接利用谱聚类方法对整合的表征进行聚类, 没有考虑到不同表征的先验知识对最终聚类结果的影响.

### 3 方法

**符号系统.** 本文使用了书法体字母 (例如  $\mathcal{A}$ )、大写字母 (例如  $A$ ) 和粗体小写字母 (例如  $\mathbf{a}$ ) 分别表示张量、矩阵和向量. 矩阵  $A$  的核范数为  $\|A\|_* = \sum_i \sigma_i(A)$ , 其中  $\sigma_i(A)$  代表矩阵  $A$  中第  $i$  大奇异值. 矩阵  $A$  的 Frobenius 范数是  $\|A\|_F = \sqrt{\sum_{i,j} |A_{ij}|^2}$ . 对于三阶张量  $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ ,  $\mathcal{A}^{(k)}$  用来表示张量  $\mathcal{A}$  的第  $k$  个正面切片, 并且  $\tilde{\mathcal{A}} = \text{fft}(\mathcal{A}, [], 3)$  表示沿第三维度的快速傅立叶变换 (Fast Fourier Transformation, FFT). 类似地, 我们可以通过逆 FFT (Inverse Fast Fourier Transformation, IFFT) 从  $\tilde{\mathcal{A}}$  得到  $\mathcal{A}$ , 即  $\mathcal{A} = \text{ifft}(\tilde{\mathcal{A}}, [], 3)$ . 为了清晰起见, 我们将本文用到的主要符号总结在表 1 中.

表 1 基础符号和说明

符号	含义
$X = \{\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(V)}\}$	包含了 $V$ 个视图的数据集
$\mathbf{X}^{(v)} \in \mathbb{R}^{d_v \times n}$	第 $v$ 个视图的特征表达矩阵
$L^{(v)} \in \mathbb{R}^{n \times n}$	第 $v$ 个视图的拉普拉斯矩阵
$F^{(v)} \in \mathbb{R}^{n \times c}$	第 $v$ 个视图的谱嵌入矩阵
$F^* \in \mathbb{R}^{n \times c}$	共识谱嵌入矩阵
$\mathcal{F} \in \mathbb{R}^{n \times c \times V}$	多个谱嵌入矩阵构造的张量
$P \in \mathbb{R}^{n \times c}$	聚类指示矩阵
$R^{(v)} \in \mathbb{R}^{c \times c}$	第 $v$ 个视图的旋转矩阵
$\gamma$	权衡参数

(续 表)

符号	含义
$\alpha_v$	第 $v$ 个视图的权重值
$\omega \in \mathbb{R}^{1 \times V}$	加权张量核范数的权重参数
$Q \in \mathbb{R}^{n \times c \times V}$	拉格朗日乘数
$\rho > 0$	惩罚因子
$\mu$	收敛速度倍数
$I \in \mathbb{R}^{n \times n}$	单位矩阵
$d_v$	第 $v$ 个视图的特征维度
$V$	视图个数
$n$	样本量大小
$c$	簇个数

#### 3.1 动机

给定多视图数据集  $X = \{\mathbf{X}^{(v)} \in \mathbb{R}^{d_v \times n}\}_{v=1}^V$ , 它包含了  $n$  个样本和  $V$  个视图. 一般来说, 朴素的多视图谱聚类方法 (Naive Multi-view Spectral Clustering, NaMSC) 需要从相应的数据空间中学习拉普拉斯矩阵  $L^{(v)} \in \mathbb{R}^{n \times n}$ . 为了获得第  $v$  个视图的谱嵌入表征, 我们会有:

$$\begin{aligned} \min_{F^{(v)}} \sum_{v=1}^V \text{Tr}(F^{(v)\top} L^{(v)} F^{(v)}), \\ \text{s. t. } F^{(v)\top} F^{(v)} = I \end{aligned} \quad (1)$$

其中  $F^{(v)} \in \mathbb{R}^{n \times c}$  是视图特定的谱嵌入矩阵, 它由拉普拉斯矩阵  $L^{(v)}$  的前  $c$  小特征值对应的  $c$  个特征向量构成,  $c$  代表簇个数. 受共识原则的启发, 我们利用一种直接的整合方法将计算好的多个谱嵌入矩阵进行加和求平均, 即  $F^* = (\sum_{v=1}^V F^{(v)})/V$ . 此外, 一些方法还会考虑  $F^{(v)}$  的不同贡献<sup>[27]</sup>. 事实上, 为了获得一致的谱嵌入表征, 还有另一种常用的方法, 称为协同正则化多视图谱聚类 (Co-regularized multi-view spectral clustering, Co-reg)<sup>[18]</sup>, 其公式可以定义如下:

$$\begin{aligned} \max_{F^{(v)}, F^*} \sum_{v=1}^V \text{Tr}(F^{(v)\top} L^{(v)} F^{(v)}) + \\ \alpha_v \text{Tr}(F^{(v)} F^{(v)\top} F^* F^{*\top}), \\ \text{s. t. } F^{(v)\top} F^{(v)} = I, F^{*\top} F^* = I \end{aligned} \quad (2)$$

其中  $\alpha_v$  是第  $v$  个视图的权衡参数. 上述公式将多个视图特定的谱嵌入矩阵  $\{F^{(v)}\}$  共正则化为一致的  $F^*$ . 在获得共识  $F^*$  后, 我们可以对其使用  $k$ -均值得到最终的聚类结果. 一般来说, 我们通过不同的范数正则化项 (例如 trace 范数或 Frobenius 范数) 来计算视图间具体的最小化或错误表征<sup>[3, 24-25]</sup>, 但是它们都没有综合考虑具有底层关系的空间结构和多视图之间的互补信息, 无法很好地识别数据的重要特征. 因此, 尽管它们具有出色的性能, 但一方面, 它们仍然无法捕获多视图数据之间的高阶相关性, 从而忽略了一些更深层次的语义信息. 另一方面, 它们没

有考虑不同视图谱嵌入结构的不同置信度和最终共识聚类结果的关联,导致次优的聚类性能.

### 3.2 框架

为了解决上述问题,本文提出了一种新的张量学习诱导的多视图谱聚类(TLIMSC)方法来挖掘谱嵌入表征之间的高阶相关性,然后通过考虑带有不同置信度的视图相关的谱嵌入表征来获得最终的聚类结果.图1对本文所提出的 TLIMSC 方法流程进行了展示和说明.如图所示,该算法以从多个特征子集中获得的数据  $X^{(1)}, \dots, X^{(V)}$  作为输入.依据输入的数据,可以得到对应的不同视图的拉普拉斯矩阵  $L^{(1)}, \dots, L^{(V)}$ ,其中的局部性学习具体指的是本文通过构造  $k$  近邻图的方式来构造相似性矩阵,继而构造对应的拉普拉斯矩阵.与直接构造视图内各个样本点之间相似性的全局性学习不同,局部性学习只需通过  $k$  个近邻样本点之间的相似性学习来刻画相

似性矩阵.进一步地,可以得到相关的  $V$  个谱嵌入矩阵  $F^{(1)}, \dots, F^{(V)}$ .对于 TLIMSC 的第一部分,为了捕获不同视图中不同数据之间的高阶相关性,本文采用了张量构造技术,其中,张量  $\mathcal{F} \in \mathbb{R}^{n \times c \times V}$  由多个谱嵌入矩阵构造得到.具体来说,我们需要将该张量旋转变成  $\bar{\mathcal{F}} \in \mathbb{R}^{n \times c \times V}$ ,以便张量的学习更有效和高效,更详细地介绍将在下文引入.随后,我们可以通过基于 t-SVD 的加权张量核范数最小化来更新旋转后的张量  $\bar{\mathcal{F}}$ ,从而可以挖掘隐藏在多视图表示中的高阶相关性.特别地,通过加权张量核范数最小化,可以对  $\mathcal{S}$  的每个奇异值进行差异性正则化,并采用软阈值函数来收缩具有不同权重参数的不同奇异值.对于 TLIMSC 的第二部分,在获得具有高阶信息的谱嵌入矩阵后,我们可以通过桥接具有不同自适应置信度的不同谱嵌入矩阵来实现最终的一致聚类.

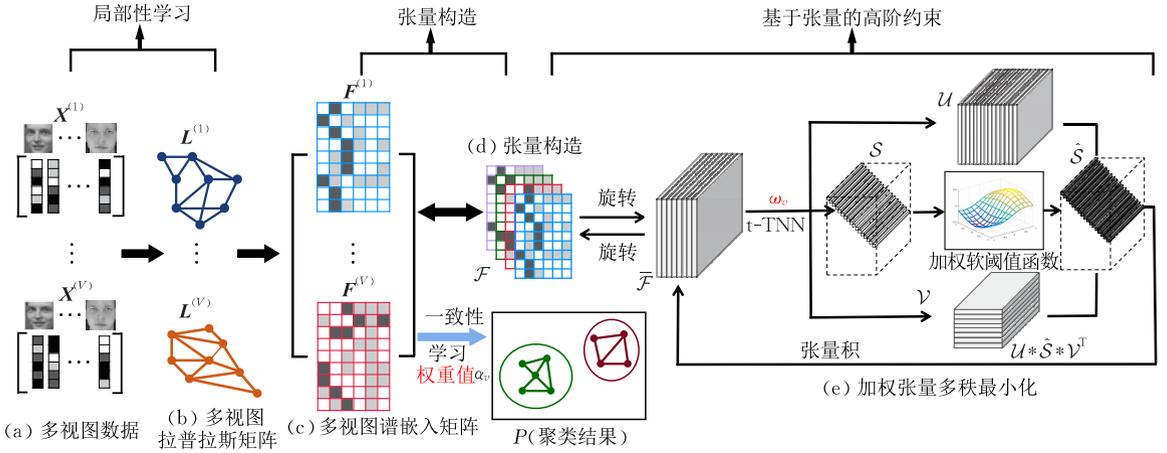


图1 TLIMSC 模型图示(给定多视图数据集  $X^{(1)}, \dots, X^{(V)}$  (子图(a)),所提出的 TLIMSC 方法旨在研究不同视图的多个谱嵌入矩阵  $F^{(1)}, \dots, F^{(V)}$  (子图(c)),通过构造张量  $\mathcal{F}$ ,挖掘视图之间的高阶相关性.为了更有效和高效, $\mathcal{F}$ 需要旋转为  $\bar{\mathcal{F}}$ ,继而  $\bar{\mathcal{F}}$ 通过基于 t-SVD 的加权张量多秩最小化来更新(子图(e)).最终,基于具有不同贡献  $\{\alpha_v\}$  的谱嵌入矩阵,得到最终的一致聚类结果)

首先,为了捕获谱嵌入矩阵内的高阶相关性,不同于 NaMSC 和 Co-reg 方法,本文所提出的 TLIMSC 方法旨在通过学习得到的谱嵌入构建一个三阶张量,

$$\min_{\mathcal{F}} \gamma \sum_{v=1}^V \text{Tr}(\mathbf{F}^{(v)\top} \mathbf{L}^{(v)} \mathbf{F}^{(v)}) + \Psi(\mathcal{F}),$$

$$\text{s. t. } \mathbf{F}^{(v)\top} \mathbf{F}^{(v)} = \mathbf{I}, \mathcal{F} = \Phi(\mathbf{F}^{(1)}, \mathbf{F}^{(2)}, \dots, \mathbf{F}^{(V)}) \quad (3)$$

其中  $\gamma$  是权衡参数.  $\Psi(\cdot)$  是关于所构造的张量  $\mathcal{F}$  的约束,  $\Phi(\cdot)$  表示通过将多个嵌入空间  $\mathbf{F}^{(v)}$  并入到一个三阶张量的方式来构造张量  $\mathcal{F}$ .

在本文中,我们采用加权张量核范数来捕获视图之间的高阶相关性,其中加权张量核范数的一般定义可见于定义 1.特别地,考虑到不同奇异值的不同重要性,采用了具有不同权重参数的软阈值函数

来收缩学习到的不同奇异值.因此,模型可以重新表述为

$$\min_{\mathcal{F}} \gamma \underbrace{\sum_{v=1}^V \text{Tr}(\mathbf{F}^{(v)\top} \mathbf{L}^{(v)} \mathbf{F}^{(v)})}_{\text{高阶相关性}} + \|\mathcal{F}\|_{\omega, \otimes},$$

$$\text{s. t. } \mathbf{F}^{(v)\top} \mathbf{F}^{(v)} = \mathbf{I}, \mathcal{F} = \Phi(\mathbf{F}^{(1)}, \mathbf{F}^{(2)}, \dots, \mathbf{F}^{(V)}) \quad (4)$$

定义 1<sup>[28]</sup>. 给定张量  $B \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ ,基于张量奇异值分解的加权张量核范数可以定义为

$$\|\mathcal{B}\|_{\omega, \otimes} = \sum_{j=1}^{n_3} \sum_{i=1}^{\min(n_1, n_2)} \omega_i \sigma_i(\tilde{B}^{(j)}) \quad (5)$$

其中  $\sigma_i(\tilde{B}^{(j)})$  是  $\tilde{B}^{(j)}$  的第  $i$  大奇异值,  $\omega_i$  是对应的第  $i$  个加权系数值,并且  $\omega_i$  维度与  $\min(n_1, n_2)$  的取值相关.

在进一步详细的计算之前,我们需要对构建的张量进行旋转,将张量的维度从  $n \times c \times V$  变为  $c \times V \times n$ ,其变换可见于图 1 中的旋转模块.同时,我们也很容易有

$$\Phi_{(v)}^{-1}(\mathcal{F}) = \mathbf{F}^{(v)} \quad (6)$$

其中  $\Phi^{-1}(\cdot)$  表示  $\Phi(\cdot)$  的逆运算,其下标  $(v)$  表示从张量中提取的第  $v$  个正面切片.在本文中,由于  $n_1 = c, n_2 = V, n_3 = n$  且  $c > V$ ,因而会有  $\min(c, V) = V$ ,也即是,  $\omega_v$  维度的选择和视图的个数相关.此外,在张量上的旋转是至关重要的.一方面,因为 t-TNN 在  $\tilde{\mathcal{F}}$  的每个正面切片上进行 SVD,使得旋转后的张量可以挖掘更多关于视图间(不仅是视图内)的高阶相关性.另一方面,旋转后的计算成本将显著降低.

值得注意的是,当张量的维度从  $n \times c \times V$  变成  $c \times V \times n$ ,其正面切片的维度则变为了  $c \times V$ .由于 t-TNN 是在每个正面切片上进行 SVD 的,那在张量学习的过程中考虑的是不同视图间不同簇的互补关系学习,以便学习得到不同视图间共识的簇关系.此外,张量学习本身刻画了比二阶矩阵学习更高阶的关系(也即是三阶关系).因此,所提出的 TLIMSC 方法能够捕捉到视图之间的互补性信息和高阶信息.

在对式(4)中的模型进行计算后,我们可以获得多个谱嵌入矩阵,这些矩阵对多视图数据与相应簇之间的高阶相关性进行编码.为了得到最终的聚类结果,我们通常会在这些谱嵌入空间上执行  $k$ -均值算法.尽管如此,它并没有考虑多个谱嵌入矩阵对最终的聚类结果的不同贡献度,这可能导致次优的聚类性能.与之前直接使用  $k$ -均值的多视图聚类方法不同,本文考虑了从不同权重的多个谱嵌入矩阵中学习最终的一致聚类.

依据文献[41, 52],当理想的相似性矩阵存在时,聚类指示矩阵的列向量可以张成对应的谱嵌入空间.具体来说,在一个理想的相似性矩阵中,如果样本点属于同一个簇,则它们是连接在一起的.否则,它们将不会有边连接,这表明理想的相似性矩阵是  $c$ -连通的.同样地,相应的邻接矩阵及其拉普拉斯矩阵都是数据样本点经适当重新排列后的分块对角矩阵.由文献[5, 53]启发,我们可以推断出理想拉普拉斯矩阵前  $c$  小特征值对应的  $c$  个特征向量构成的特征空间是由聚类指示矩阵的列向量所张成的.

事实上,理想的相似性矩阵对应的谱嵌入矩阵  $\mathbf{F}^{(v)}$  不一定等于指示矩阵  $\mathbf{P}$ .受文献[41, 52]的启发,  $\mathbf{F}^{(v)}$  的列是构成  $\mathbf{P}$  的正交基,我们可以通过  $\mathbf{F}^{(v)}$  的旋转来恢复指示矩阵  $\mathbf{P}$ ,即

$$\begin{aligned} \min_{\mathbf{P}, \mathbf{R}^{(v)}} \sum_{v=1} \|\mathbf{P} - \mathbf{F}^{(v)} \mathbf{R}^{(v)}\|_F^2, \\ \text{s. t. } \mathbf{P} \in \text{Ind}, \mathbf{R}^{(v)\top} \mathbf{R}^{(v)} = \mathbf{I} \end{aligned} \quad (7)$$

此外,式(7)中的模型没有考虑不同视图对共识指示矩阵  $\mathbf{P}$  的不同贡献,容易导致聚类性能的不稳定.因此,本文引入了一种关于多视图自适应加权的策略,其公式可以定义为

$$\begin{aligned} \min_{\mathbf{P}, \mathbf{R}^{(v)}} \underbrace{\sum_{v=1} \frac{1}{\alpha_v} \|\mathbf{P} - \mathbf{F}^{(v)} \mathbf{R}^{(v)}\|_F^2}_{\text{加权一致性}}, \\ \text{s. t. } \mathbf{P} \in \text{Ind}, \mathbf{R}^{(v)\top} \mathbf{R}^{(v)} = \mathbf{I} \end{aligned} \quad (8)$$

其中  $\alpha_v$  是以下形式的目标值,

$$\alpha_v \stackrel{\text{def}}{=} \|\mathbf{P} - \mathbf{F}^{(v)} \mathbf{R}^{(v)}\|_F \quad (9)$$

可以明显看出,上述式(9)中的  $\alpha_v$  相对于  $\mathbf{P}$  和  $\mathbf{R}^{(v)}$  独立的,当有确切的  $\mathbf{P}$  和  $\mathbf{R}^{(v)}$  时,我们可以设它为常数.

为了直接比较我们的共识聚类学习部分和  $k$ -均值算法的运用,在本文中,我们没有将以上两个部分放在一个统一的框架中.特别地,虽然它们不是在一个框架中一起学习的,但是通过有效地学习视图间的高阶相关性,并由不同置信度的嵌入表征来得到一致的聚类结果,我们也可以获得优异的实验结果,这可以在后续的实验部分进一步证明.

### 3.3 优化

本文采用交替最小化策略来求解所提出的 TLIMSC 方法,即更新一个特定变量时固定其他变量.下面我们先解决 TLIMSC 第一部分中关于高阶相关性学习,然后介绍如何解决第二部分关于最终聚类结果的学习.

#### 3.3.1 高阶相关性学习

由增强拉格朗日乘数(Augmented Lagrangian Multiplier, ALM)诱导,我们进一步运用 ALM 与替代方向最小化(Alternative Direction Minimizing, ADM)优化算法.为了使  $\mathcal{F}$  可分离,采用变量拆分技术,引入一个辅助张量变量  $\mathcal{J} \in \mathbb{R}^{n \times c \times V}$  来替换  $\mathcal{F}$ .因此,我们可以将式(4)中的模型重新表述为以下优化问题:

$$\begin{aligned} \min_{\mathbf{F}^{(v)}, \mathcal{J}} \gamma \sum_{v=1} \text{Tr}(\mathbf{F}^{(v)\top} \mathbf{L}^{(v)} \mathbf{F}^{(v)}) + \|\mathcal{J}\|_{\omega, \otimes}, \\ \text{s. t. } \mathbf{F}^{(v)\top} \mathbf{F}^{(v)} = \mathbf{I}, \mathcal{F} = \mathcal{J} \end{aligned} \quad (10)$$

进一步地,上述模型相应的增广拉格朗日函数可表述为

$$\begin{aligned} \mathcal{L}(\mathcal{F}, \mathcal{J}) = \gamma \sum_{v=1} \text{Tr}(\mathbf{F}^{(v)\top} \mathbf{L}^{(v)} \mathbf{F}^{(v)}) + \|\mathcal{J}\|_{\omega, \otimes} + \\ \frac{\rho}{2} \left\| \mathcal{J} - \left( \mathcal{F} + \frac{\mathcal{Q}}{\rho} \right) \right\|_F^2, \\ \text{s. t. } \mathbf{F}^{(v)\top} \mathbf{F}^{(v)} = \mathbf{I} \end{aligned} \quad (11)$$

其中  $\mathcal{Q} \in \mathbb{R}^{n \times c \times V}$  是拉格朗日乘数,  $\rho > 0$  是惩罚因子. 借助于交替最小化策略, 我们可以将式(11)中的问题分解为以下三个子问题, 然后交替优化每个变量.

**$\mathcal{J}$ 子问题.** 通过固定其他变量,  $\mathcal{J}$  可以通过解决以下问题来更新,

$$\min_{\mathcal{J}} \frac{1}{\rho} \|\mathcal{J}\|_{\omega, \otimes} + \frac{1}{2} \left\| \mathcal{J} - \left( \mathcal{F} + \frac{\mathcal{Q}}{\rho} \right) \right\|_F^2 \quad (12)$$

为了求解上述加权张量核范数最小化问题, 需引入以下两个定理.

**定理 1**<sup>[54]</sup>. 给定  $Y \in \mathbb{R}^{m \times n}$  的 SVD 形式  $Y = U_Y * D_Y * V_Y^T$ ,  $\tau > 0$ ,  $l = \min(m, n)$ , 我们可以将标准加权核范数最小化问题定义为

$$\arg \min_X \frac{1}{2} \|X - Y\|_F^2 + \tau \|X\|_{\omega, *}, \quad (13)$$

则我们可以得到模型在式(13)中的最优解,

$$X^* = \Gamma_{\tau * \omega}[Y] = U_Y * P_{\tau * \omega}(Y) * V_Y^T \quad (14)$$

其中我们有  $P_{\tau * \omega}(Y) = \text{diag}(\xi_1, \xi_2, \dots, \xi_l)$ ,  $\xi_i = \text{sign}(\sigma_i(Y)) \max(\sigma_i(Y) - \tau * \omega_i, 0)$ .

**定理 2**<sup>[28]</sup>. 给定  $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ ,  $l = \min(n_1, n_2)$ , 我们有  $\mathcal{A} = \mathcal{U} * \mathcal{S} * \mathcal{V}^T$ . 针对模型

$$\arg \min_{\mathcal{X}} \frac{1}{2} \|\mathcal{X} - \mathcal{A}\|_F^2 + \tau \|\mathcal{X}\|_{\omega, \otimes} \quad (15)$$

我们可以定义其最优解为

$$\mathcal{X}^* = \Gamma_{\tau * \omega}(\mathcal{A}) = \mathcal{U} * \text{ifft}(P_{\tau * \omega}(\tilde{\mathcal{A}})) * \mathcal{V}^T \quad (16)$$

其中  $\tilde{\mathcal{A}} = \text{fft}(\mathcal{A}, [ , ], 3)$ ,  $P_{\tau * \omega}(\tilde{\mathcal{A}})$  是一个张量, 其中  $P_{\tau * \omega}(\tilde{\mathcal{A}}^{(i)})$  表示  $P_{\tau * \omega}(\tilde{\mathcal{A}})$  中的第  $i$  个正面切片.

因此, 根据定理 2, 我们可以很容易地得到模型在式(12)中的解:

$$\mathcal{J}^* = \Gamma_{\frac{1}{\rho} * \omega} \left[ \mathcal{F} + \frac{1}{\rho} \mathcal{Q} \right] \quad (17)$$

**$F^{(v)}$ 子问题.** 通过固定其他变量,  $F^{(v)}$  可以得到更新, 式(11)中的问题变为

$$\begin{aligned} \min_{F^{(v)}} \gamma \sum_{v=1} Tr(\mathbf{F}^{(v)T} \mathbf{L}^{(v)} \mathbf{F}^{(v)}) + \\ \frac{\rho}{2} \left\| \mathcal{J} - \left( \mathcal{F} + \frac{\mathcal{Q}}{\rho} \right) \right\|_F^2, \\ \text{s. t. } \mathbf{F}^{(v)T} \mathbf{F}^{(v)} = \mathbf{I} \end{aligned} \quad (18)$$

具体来说, 对于每个视图, 式(18)中的问题可以重写为

$$\begin{aligned} \min_{F^{(v)}} \gamma Tr(\mathbf{F}^{(v)T} \mathbf{L}^{(v)} \mathbf{F}^{(v)}) + \\ \frac{\rho}{2} \left\| J^{(v)} - \left( \mathbf{F}^{(v)} + \frac{\mathbf{Q}^{(v)}}{\rho} \right) \right\|_F^2, \\ \text{s. t. } \mathbf{F}^{(v)T} \mathbf{F}^{(v)} = \mathbf{I} \end{aligned} \quad (19)$$

上述式(19)中的模型可以进一步简化为

$$\begin{aligned} \min_{F^{(v)T} \mathbf{F}^{(v)} = \mathbf{I}} Tr(\mathbf{F}^{(v)T} (\gamma \mathbf{L}^{(v)} - \lambda \mathbf{I}) \mathbf{F}^{(v)}) - \\ Tr \left( \mathbf{F}^{(v)T} \rho \left( J^{(v)} - \frac{\mathbf{Q}^{(v)}}{\rho} \right) \right), \\ \Leftrightarrow \max_{F^{(v)T} \mathbf{F}^{(v)} = \mathbf{I}} Tr(\mathbf{F}^{(v)T} (\lambda \mathbf{I} - \gamma \mathbf{L}^{(v)}) \mathbf{F}^{(v)}) + \\ 2Tr \left( \mathbf{F}^{(v)T} \frac{\rho}{2} \left( J^{(v)} - \frac{\mathbf{Q}^{(v)}}{\rho} \right) \right), \\ \Leftrightarrow \max_{F^{(v)T} \mathbf{F}^{(v)} = \mathbf{I}} Tr(\mathbf{F}^{(v)T} \mathbf{B}^{(v)} \mathbf{F}^{(v)}) + 2Tr(\mathbf{F}^{(v)T} \mathbf{C}^{(v)}) \end{aligned} \quad (20)$$

其中  $\mathbf{B}^{(v)} = \lambda \mathbf{I} - \gamma \mathbf{L}^{(v)} \in \mathbb{R}^{n \times n}$  和  $\mathbf{C}^{(v)} = \frac{\rho}{2} \left( J^{(v)} - \frac{\mathbf{Q}^{(v)}}{\rho} \right) \in \mathbb{R}^{n \times c}$ .  $\lambda$  表示一个常数, 以确保矩阵  $\mathbf{B}^{(v)}$  是正定的.  $\mathbf{I} \in \mathbb{R}^{n \times n}$  是一个单位矩阵.

为了得到上述问题的解  $F^{(v)*}$ , 我们需要引入以下定理.

**定理 3**<sup>[27]</sup>. 考虑以下的问题:

$$\max_{F^T \mathbf{F} = \mathbf{I}} Tr(\mathbf{F}^T \mathbf{B} \mathbf{F}) + 2Tr(\mathbf{F}^T \mathbf{C}) \quad (21)$$

我们可以得到解  $F^* = \mathbf{U} \mathbf{V}^T$ , 其中  $\mathbf{U} \mathbf{S} \mathbf{V}^T$  是  $\mathbf{M} = \mathbf{B} \mathbf{F} + \mathbf{C}$  的紧凑 SVD 形式.

证明. 首先引入一个对称的拉格朗日乘子  $\mathbf{\Pi}$ , 然后式(21)中的模型可以重新表述为

$$\begin{aligned} \mathcal{L}(\mathbf{F}, \mathbf{\Pi}) = Tr(\mathbf{F}^T \mathbf{B} \mathbf{F}) + 2Tr(\mathbf{F}^T \mathbf{C}) - \\ Tr(\mathbf{\Pi}(\mathbf{F}^T \mathbf{F} - \mathbf{I})) \end{aligned} \quad (22)$$

为了获得  $F^*$ , 我们将  $\mathcal{L}(\mathbf{F}, \mathbf{\Pi})$  关于  $\mathbf{F}$  的导数设置为零, 即

$$2\mathbf{B} \mathbf{F} + 2\mathbf{C} - 2\mathbf{F} \mathbf{\Pi} = 0 \quad (23)$$

然后式(23)变成

$$\mathbf{F} \mathbf{\Pi} = \mathbf{B} \mathbf{F} + \mathbf{C} = \mathbf{M} \quad (24)$$

由于  $\mathbf{M} \in \mathbb{R}^{n \times c}$  的紧凑 SVD 形式, 我们有  $\mathbf{M} = \mathbf{U} \mathbf{S} \mathbf{V}^T$ , 其中  $\mathbf{U} \in \mathbb{R}^{n \times c}$ ,  $\mathbf{S} \in \mathbb{R}^{c \times c}$  和  $\mathbf{V} \in \mathbb{R}^{c \times c}$ .

为了得到  $\mathbf{\Pi}$ , 我们进行以下转换:

$$\begin{aligned} (\mathbf{F} \mathbf{\Pi})^T (\mathbf{F} \mathbf{\Pi}) &= \mathbf{M}^T \mathbf{M} \\ \Leftrightarrow \mathbf{\Pi}^T \mathbf{F}^T \mathbf{F} \mathbf{\Pi} &= (\mathbf{U} \mathbf{S} \mathbf{V}^T)^T (\mathbf{U} \mathbf{S} \mathbf{V}^T) \\ \Leftrightarrow \mathbf{\Pi}^T \mathbf{\Pi} &= \mathbf{V} \mathbf{S}^2 \mathbf{V}^T \\ \Leftrightarrow \mathbf{\Pi}^2 &= \mathbf{V} \mathbf{S}^2 \mathbf{V}^T \\ \Leftrightarrow \mathbf{\Pi} &= \mathbf{V} \mathbf{S} \mathbf{V}^T \end{aligned} \quad (25)$$

之后, 我们可以将  $\mathbf{\Pi} = \mathbf{V} \mathbf{S} \mathbf{V}^T$  代入式(24)中, 即

$$\mathbf{F} \mathbf{\Pi} = \mathbf{F} (\mathbf{V} \mathbf{S} \mathbf{V}^T) = \mathbf{M} = \mathbf{U} \mathbf{S} \mathbf{V}^T \quad (26)$$

因此, 式(22)中问题的解  $F^*$  由如下公式可以计算得到,

$$F^* = \mathbf{U} \mathbf{V}^T \quad (27)$$

证毕.

相应地, 由于定理 3, 式(19)中模型的最优解

$F^{(v)*}$  可以定义为

$$F^{(v)*} = \mathbf{U}^{(v)} \mathbf{V}^{(v)T} \quad (28)$$

其中  $\mathbf{U}^{(v)} \mathbf{S}^{(v)} \mathbf{V}^{(v)\top}$  是  $\mathbf{M}^{(v)} = \mathbf{B}^{(v)} \mathbf{F}^{(v)} + \mathbf{C}^{(v)}$  的紧凑 SVD 形式.

**Q 和  $\rho$  子问题.** 拉格朗日乘子  $Q$  和惩罚因子  $\rho$  可以通过以下方式更新,

$$\begin{aligned} Q &= Q + \rho(\mathcal{F} - \mathcal{J}), \\ \rho &= \mu * \rho \end{aligned} \quad (29)$$

其中  $\mu > 1$  用于提高收敛速度<sup>[55]</sup>. 我们在算法 1 中总结了式(4)的求解过程.

**算法 1.** 学习视图特定的谱嵌入空间  $\{\mathbf{F}^{(v)}\}$ .

输入: 多视图数据集  $X = \{\mathbf{X}^{(v)} \in \mathbb{R}^{d_v \times n}\}_{v=1}^V$ , 权衡参数  $\gamma > 0$ , 权重向量  $\omega > 0$  和聚类个数  $c$

输出: 多个谱嵌入空间  $\{\mathbf{F}^{(v)}\}$

1. 用 knn 初始化  $\mathbf{W}^{(v)}$ , 计算  $\mathbf{L}^{(v)}$  和  $\mathbf{F}^{(v)}$ ;
2. 初始化  $Q = \mathcal{J} = 0$ ;
3. WHILE 不收敛 DO
4. 通过解决式(17)来更新  $\mathcal{J}$ ;
5. FOR  $\forall v = 1, \dots, V$  DO
6.  $\mathbf{U}^{(v)} \mathbf{S}^{(v)} \mathbf{V}^{(v)\top} = \text{SVD}(\mathbf{B}^{(v)} \mathbf{F}^{(v)} + \mathbf{C}^{(v)})$ ;
7. 通过  $\mathbf{F}^{(v)} = \mathbf{U}^{(v)} \mathbf{V}^{(v)\top}$  更新  $\mathbf{F}^{(v)}$ ;
8. END FOR
9. 通过  $Q = Q + \rho(\mathcal{F} - \mathcal{J})$  更新  $Q$ ;
10. 通过  $\rho = \mu\rho$  更新  $\rho$ ;
11. END WHILE

### 3.3.2 最终聚类结果学习

为了解决式(8)中的问题,我们也采用了交替优化策略.

**P 子问题.** 通过固定除  $\mathbf{P}$  之外的其他变量,式(8)相当于求解下面的问题:

$$\begin{aligned} \min_{\mathbf{P}} \sum_{v=1}^V \frac{1}{\alpha_v} \|\mathbf{P} - \mathbf{F}^{(v)} \mathbf{R}^{(v)}\|_F^2, \\ \text{s. t. } \mathbf{P} \in \text{Ind} \end{aligned} \quad (30)$$

在简单地展开上述等式的  $v$  个平方范数之后,我们有

$$\begin{aligned} \min_{\mathbf{P} \in \text{Ind}} \sum_{v=1}^V \frac{1}{\alpha_v} (\|\mathbf{P}\|_F^2 + \|\mathbf{F}^{(v)} \mathbf{R}^{(v)}\|_F^2) - \\ \sum_{v=1}^V \frac{2}{\alpha_v} \text{Tr}(\mathbf{P}^\top \mathbf{F}^{(v)} \mathbf{R}^{(v)}) \end{aligned} \quad (31)$$

此外,由于  $\mathbf{P} \in \text{Ind}$  和  $\mathbf{F}^{(v)\top} \mathbf{F}^{(v)} = \mathbf{I}$  的约束,式(31)中的问题可以改写为

$$\min_{\mathbf{P} \in \text{Ind}} \sum_{v=1}^V \frac{n+c}{\alpha_v} - 2 \text{Tr} \left( \mathbf{P}^\top \left( \sum_{v=1}^V \frac{\mathbf{F}^{(v)} \mathbf{R}^{(v)}}{\alpha_v} \right) \right) \quad (32)$$

因此,为了获得最优解  $\mathbf{P}^*$ ,我们需要解决如下问题:

$$\max_{\mathbf{P} \in \text{Ind}} \text{Tr} \left( \mathbf{P}^\top \left( \sum_{v=1}^V \frac{\mathbf{F}^{(v)} \mathbf{R}^{(v)}}{\alpha_v} \right) \right) \quad (33)$$

其中它有一个闭合解,即

$$P_{ij}^+ = \begin{cases} 1, & j = \arg \max_c \left[ \sum_{v=1}^V \frac{\mathbf{F}^{(v)} \mathbf{R}^{(v)}}{\alpha_v} \right]_c \\ 0 & \text{其他} \end{cases} \quad (34)$$

**$\mathbf{R}^{(v)}$  子问题.** 通过固定其他变量,我们可以通过以下问题更新  $\mathbf{R}^{(v)}$ ,

$$\begin{aligned} \min_{\mathbf{R}^{(v)}} \sum_{v=1}^V \frac{1}{\alpha_i} \|\mathbf{P} - \mathbf{F}^{(v)} \mathbf{R}^{(v)}\|_F^2, \\ \text{s. t. } \mathbf{R}^{(v)\top} \mathbf{R}^{(v)} = \mathbf{I} \end{aligned} \quad (35)$$

显然,式(35)中的第  $v$  个子问题可以独立求解(其中  $\frac{1}{\alpha_i}$  为常数),即

$$\begin{aligned} \min_{\mathbf{R}^{(v)}} \|\mathbf{P} - \mathbf{F}^{(v)} \mathbf{R}^{(v)}\|_F^2, \\ \text{s. t. } \mathbf{R}^{(v)\top} \mathbf{R}^{(v)} = \mathbf{I} \end{aligned} \quad (36)$$

为了计算上述问题的最优解  $\mathbf{R}^{(v)*}$ ,本文提供了以下的定理.

**定理 4.** 给定问题  $\min_{\mathbf{R}^\top \mathbf{R} = \mathbf{I}} \|\mathbf{H} - \mathbf{G}\mathbf{R}\|_F^2$ , 我们可以得到  $\mathbf{R}$  的闭式解,即  $\mathbf{R}^* = \bar{\mathbf{U}} \bar{\mathbf{V}}^\top$ , 其中  $\bar{\mathbf{U}}$  和  $\bar{\mathbf{V}}$  分别为  $\mathbf{G}^\top \mathbf{H}$  的 SVD 左右奇异值.

因此,根据定理 4,我们可以得到式(36)中问题的最优解,

$$\mathbf{R}^{(v)} = \bar{\mathbf{U}}^{(v)} \bar{\mathbf{V}}^{(v)\top} \quad (37)$$

其中  $\bar{\mathbf{U}}^{(v)} \bar{\mathbf{S}}^{(v)} \bar{\mathbf{V}}^{(v)\top}$  是  $\mathbf{F}^{(v)\top}$  的 SVD 形式.

**$\alpha_v$  子问题.** 通过固定除  $\alpha_v$  之外的其他变量,该优化问题相当于处理式(9)中的模型,可以由已经计算得到的  $\mathbf{P}$  和  $\mathbf{R}^{(v)}$  求解. 在算法 2 中我们总结了式(8)的求解过程.

**算法 2.** 学习最终聚类结果  $\mathbf{P}$ .

输入: 多个谱嵌入矩阵  $\{\mathbf{F}^{(v)} \in \mathbb{R}^{n \times c}\}_{v=1}^V$

输出: 最终聚类结果  $\mathbf{P}$

1. 令  $[\alpha_v \in \mathbb{R}^{V \times 1}]_{v=1}^V = 1, \{\mathbf{R}^{(v)} \in \mathbb{R}^{n \times c}\}_{v=1}^V = \mathbf{I}$ ;
2. WHILE 不收敛 DO
3. 通过式(34)更新  $\mathbf{P}$ ;
4. 通过解决式(37)来更新  $\mathbf{R}^{(v)}$ ;
5. 通过式(9)来更新  $\alpha_v$ ;
6. END WHILE

### 3.4 复杂度分析

在这一小节中,我们给出了这两个优化算法的计算复杂度分析. 对于算法 1,在初始化步骤中,我们使用了 knn 方法初始化相似性矩阵  $\mathbf{W}^{(v)}$ ,其复杂度为  $O(nc)$ ,然后我们使用标准的谱聚类方法来计算每个  $\mathbf{F}^{(v)}$ ,这需要  $O(nc^2V)$  的复杂度. 在算法的优化迭代中,首先解决了  $\mathcal{J}$  子问题,其中维度为  $c \times V \times n$  的张量的 3D FFT 及 3D IFFT 和维度为  $c \times V$

的矩阵在傅立叶域中的  $n$  个 SVD 的过程为主要的计算. 因此, 在每次迭代中更新  $\mathcal{J}$  的计算复杂度为  $O(2ncV\log(ncV) + cV^2n)$ , 其中  $O(2ncV\log(ncV))$  是对于 FFT 和 IFFT 的计算成本,  $O(cV^2n)$  是维度为  $cV \times n$  的矩阵 SVD 分解的成本. 对于更新  $\mathbf{F}^{(v)}$ , 依据定理 3, 每次迭代的计算成本为  $O(nc^2V)$ . 对于乘子  $\mathcal{Q}$ , 我们只需要  $O(V)$  的复杂度. 因此, 总的计算复杂度为  $O((2nc^2V + 2ncV\log(ncV) + ncV^2 + nc + V)t_F)$ , 其中  $t_F$  表示算法 1 的迭代次数, 并且  $n \gg c, n \gg V$ . 对于算法 2, 更新  $\mathbf{P}$  时, 式(34)的计算成本为  $O(nc^2V + nc)$ , 其中  $O(nc^2V)$  用于构造公式  $\sum_{v=1} \frac{\mathbf{F}^{(v)} \mathbf{R}^{(v)}}{\alpha_v}$ ,  $O(nc)$  用于执行  $\arg \max$  运算符. 对于  $\mathbf{R}^{(v)}$  子问题, 矩阵乘法需要  $O(nc^2V)$ . 对于求解  $\alpha_v$ , 式(9)的计算成本仅需  $O(ncV)$ . 因此, 整个计算复杂度为  $O((2nc^2V + ncV + nc)t_P)$ , 其中  $t_P$  为算法 2 的迭代次数. 总体而言, 所提出的 TLIMSC 方法总共需要  $O((2nc^2V + 2ncV\log(ncV) + ncV^2 + nc + V)t_F + (2nc^2V + ncV + nc)t_P)$  的计算复杂度, 由于  $t_F$  和  $t_P$  通常很小(在我们的实验中小于 20), 因此算法的总计复杂度接近于  $O(n \log n)$ .

### 3.5 收敛性分析

在这一小节中, 为了对所提出的 TLIMSC 方法的收敛性进行分析, 我们将分别分析其中两部分算法的收敛性. 对于第一部分的算法, 需要优化更新的两个变量分别是  $\mathbf{F}^{(v)}$  和  $\mathcal{J}$ . 文献[56]指出当不精确的增广拉格朗日乘子法至多有两个变量模块时, 其收敛性是可以确保的, 因而也推导出了第一部分算法理论上的收敛性. 进一步地, 在几个真实世界数据集上的经验收敛曲线证据证明了该算法在学习具有更高阶关系的谱嵌入空间时具有强大而稳定的收敛性. 对于第二部分的算法, 它的收敛性证明如下. 具体来说, 从这个算法中我们可以看出, 对应三个变量的三种更新规则都不会使相应的目标函数值增加. 进一步地, 由于自适应加权策略的存在, 该目标函数可以相应地简化为以下问题:

$$\arg \min_{\mathbf{P} \in \text{Ind}, \mathbf{R}^{(v)} \mathbf{R}^{(v)\top} = \mathbf{I}} \sum_{v=1} \|\mathbf{P} - \mathbf{F}^{(v)} \mathbf{R}^{(v)}\|_F \quad (38)$$

上述式(38)有下界, 所以我们可以从理论上推导出第二部分算法的收敛性[52], 并且几个真实世界数据集上的经验收敛曲线证据也验证了其快速的收敛行为.

因此, 我们所提出的 TLIMSC 方法的收敛性是可以得到保证的.

## 4 实 验

在本节中, 为了验证所提出方法的有效性, 我们在五个真实世界的数据集上进行了大量实验, 在聚类性能和运行时间方面将所提出的方法与现有最先进的方法进行比较. 同时, 还综合地记录了消融实验、参数分析和收敛性分析的结果. 所有的实验均在酷睿 2 四核 1.7 GHz 和 8 GB 内存的机器上使用 Matlab R2016a 实现. TLIMSC 方法的源代码和所用真实世界数据集可以从 <https://pan.baidu.com/s/1eTtDMcxV3bdTpbk5q6XkDA> 下载, 其中的提取码为 TLIM.

### 4.1 实验设置

#### 4.1.1 数据集

我们在五个真实世界的数据集上进行了实验, 以便对所提出的 TLIMSC 方法进行全面评估. 我们将这五个数据集的统计信息总结在表 2 中.

表 2 五个真实世界数据的统计信息

数据集	样本量	视图个数	类个数
ORL	400	3	40
UCI	2000	3	10
NUS	2400	6	12
Scene-15	4485	3	15
Hdigit	10 000	2	10

#### 4.1.2 对比方法

以下最先进的聚类方法被用于与我们提出的 TLIMSC 方法进行比较: 分别应用于每个视图的标准谱聚类方法 (SC)[57]、应用于所有视图级联特征的标准谱聚类方法 (FeaCon)、协同训练多视图谱聚类方法 (Co-trained Multi-view Spectral Clustering, CoTr)[36]、RMSC[21]、一致性与独特性多视图子空间聚类方法 (Consistent and Specific Multi-view Subspace Clustering, CSMSC)[45]、LTMSC[22]、潜在多视图子空间聚类方法 (Latent Multi-view Subspace Clustering, LMSC)[25]、基于样本感知相似性的多视图子空间聚类方法 (Multi-view Subspace Clustering with Intactness-Aware Similarity, MSCIAS)[38]、基于自适应近邻的多视图学习方法 (Multi-view Learning with Adaptive Neighbors, MLAN)[42]、基于图的多视图聚类方法 (Graph-based Multi-view Clustering, GMC)[40]、WTNNM[28]、tMSC[13]、用于多视图谱聚类的基本张量学习方法 (Essential Tensor Learning for Multiview Spectral Clustering, ETLMSC)[26]、用于多视图内在子空间聚类的局部

结构低秩张量近似方法(Low-rank Tensor Approximation with Local Structure for Multi-view Intrinsic Subspace Clustering, LTALS)<sup>[58]</sup>.

对于所有方法,包括所提出的 TLIMSC 方法和上述 14 种方法,每种方法都进行了 20 次实验,并报告了算法的平均性能.同时,大多数方法需要  $k$ -均值算法才能实现最终的聚类性能,因此我们将  $k$ -均值算法运行 10 次以减轻每次实验中随机初始化的影响.对于我们的方法,具体来说,权衡参数  $\gamma$  在  $[0.2, 3]$  的范围内调整,不同视图对应的奇异值权重  $\omega$  在  $(0, 70]$  内调整.由于不同视图的属性不同,ALM 函数优化的惩罚因子  $\rho$  选取范围为  $\{0.003, 0.005, 0.008, 0.009, 0.03, 0.05, 0.08, 0.09, 0.3, 0.5, 0.8, 0.9\}$ ,相应的收敛提升参数  $\mu (> 1)$  在  $[1.5 : 0.5 : 8]$  范围内选取.

#### 4.1.3 评价指标

对于评价指标,我们运用了四种广泛使用的指标来综合评估聚类性能,即准确性(Accuracy, ACC)、归一化互信息(Normalized Mutual Information, NMI)、F 分数(Fscore)和调整等级指数(Adjusted Rank Index, ARI).对于上述所有的评估措施,较高的值表示更好的聚类性能<sup>[59]</sup>.由于每个评估指标都会惩罚或偏向于聚类结果中的特定属性,因此我们可以通过四种不同的指标报告最终的聚类结果来获得更全面的评估.

#### 4.2 对比实验

表 3 至表 7 分别报告了在五个真实世界的数据集上采用不同聚类方法得到的关于 ACC、NMI、Fscore 和 ARI 的聚类结果,其中我们将所有方法运

表 4 UCI 数据集上的平均性能

方法	数据集 UCI			
	ACC	NMI	Fscore	ARI
SC1	0.6176	0.5858	0.5063	0.4510
SC2	0.6847	0.5874	0.5542	0.5044
SC3	0.5468	0.4897	0.4271	0.3628
FeaCon	0.5919	0.5767	0.5152	0.4603
CoTr	0.8401	0.7961	0.7799	0.7549
RMSC	0.8599	0.8225	0.8006	0.7779
CSMSC	0.8826	0.7877	0.7841	0.7600
LTMSC	0.8004	0.7689	0.7489	0.7206
LMSC	0.8566	0.7837	0.7628	0.7364
MSCIAS	0.9766	0.9469	0.9533	0.9482
MLAN	0.9680	0.9258	0.9370	0.9301
GMC	0.7355	0.8153	0.7134	0.6779
WTNNM	0.9965	0.9904	0.9930	0.9922
tMSC	0.9960	0.9891	0.9920	0.9911
ETLMSC	0.9339	0.9610	0.9393	0.9321
LTALS	0.9950	0.9866	0.9900	0.9889
TLIMSC	<b>0.9990</b>	<b>0.9973</b>	<b>0.9980</b>	<b>0.9978</b>

表 5 NUS 数据集上的平均性能

方法	数据集 NUS			
	ACC	NMI	Fscore	ARI
SC1	0.1601	0.0392	0.1150	0.0149
SC2	0.1542	0.0386	0.1186	0.0118
SC3	0.1406	0.0316	0.0983	0.0115
SC4	0.1543	0.0404	0.1048	0.0156
SC5	0.1400	0.0353	0.1115	0.0110
SC6	0.2384	0.1134	0.1394	0.0584
FeaCon	0.2438	0.1181	0.1427	0.0619
CoTr	0.3122	0.1782	0.1821	0.1057
RMSC	0.2496	0.1240	0.1460	0.0676
CSMSC	0.2597	0.1400	0.1677	0.0866
LTMSC	0.2473	0.1245	0.1543	0.0735
LMSC	0.2597	0.1283	0.1519	0.0746
MSCIAS	0.2230	0.1187	0.1420	0.0550
MLAN	0.2552	0.1592	0.1775	0.0444
GMC	0.1650	0.1224	0.1592	0.0124
WTNNM	0.5219	0.3638	0.3505	0.2909
tMSC	0.4686	0.3434	0.3226	0.2597
ETLMSC	0.8357	0.7728	0.7491	0.7259
LTALS	0.1604	0.0713	0.1701	0.0243
TLIMSC	<b>0.9758</b>	<b>0.9542</b>	<b>0.9527</b>	<b>0.9484</b>

表 3 ORL 数据集上的平均性能

方法	数据集 ORL			
	ACC	NMI	Fscore	ARI
SC1	0.6500	0.7983	0.5285	0.5169
SC2	0.7744	0.8911	0.7120	0.7049
SC3	0.7049	0.8421	0.6116	0.6021
FeaCon	0.6523	0.8014	0.5359	0.5246
CoTr	0.7539	0.8813	0.6880	0.6802
RMSC	0.7583	0.8841	0.6981	0.6907
CSMSC	0.8166	0.9173	0.7741	0.7685
LTMSC	0.7913	0.9027	0.7393	0.7329
LMSC	0.8013	0.9066	0.7455	0.7392
MSCIAS	0.8724	0.9310	0.8244	0.8202
MLAN	0.7275	0.8384	0.5097	0.4946
GMC	0.6325	0.8571	0.3599	0.3367
WTNNM	0.9776	0.9934	0.9774	0.9769
tMSC	0.9741	0.9925	0.9742	0.9736
ETLMSC	0.7172	0.8579	0.6403	0.6313
LTALS	0.9925	0.9952	0.9859	0.9855
TLIMSC	<b>1.0000</b>	<b>1.0000</b>	<b>1.0000</b>	<b>1.0000</b>

表 6 Scene-15 数据集上的平均性能

方法	数据集 Scene-15			
	ACC	NMI	Fscore	ARI
SC1	0.4256	0.4050	0.3179	0.2673
SC2	0.3435	0.3515	0.2511	0.5000
SC3	0.3227	0.2910	0.2145	0.1557
FeaCon	0.3205	0.2918	0.2141	0.1552
CoTr	0.5848	0.5513	0.4594	0.4193
RMSC	0.4473	0.4212	0.3417	0.2931
CSMSC	0.5927	0.5684	0.4742	0.4339
LTMSC	0.5657	0.5710	0.4643	0.4234
LMSC	0.5234	0.5356	0.4236	0.3768
MSCIAS	0.5675	0.6016	0.4688	0.4261
MLAN	0.4442	0.4724	0.2837	0.1986
GMC	0.3810	0.5186	0.2809	0.1905
WTNNM	0.9021	0.9266	0.8981	0.8905
tMSC	0.9013	0.9264	0.8974	0.8897
ETLMSC	0.8302	0.8591	0.8091	0.7944
LTALS	0.8067	0.8333	0.7775	0.7610
TLIMSC	<b>0.9688</b>	<b>0.9522</b>	<b>0.9432</b>	<b>0.9390</b>

表 7 Hdigit 数据集上的平均性能

方法	数据集 Hdigit			
	ACC	NMI	Fscore	ARI
SC1	0.5264	0.4685	0.4128	0.3459
SC2	0.4856	0.4555	0.3912	0.3227
FeaCon	0.7562	0.6714	0.6482	0.6086
CoTr	0.9102	0.8263	0.8452	0.8288
RMSC	0.7292	0.6720	0.6392	0.5986
CSMSC	0.8348	0.7389	0.7311	0.7010
LTMSC	0.7825	0.6616	0.6490	0.6098
LMSC	0.8020	0.7966	0.7588	0.7309
MSCIAS	0.7287	0.8314	0.7638	0.7340
MLAN	0.7106	0.8374	0.7620	0.7317
GMC	0.9981	0.9939	0.9962	0.9958
WTNNM	0.9989	0.9964	0.9978	0.9976
tMSC	0.9972	0.9913	0.9944	0.9938
ETLMSC	0.9178	0.9629	0.9328	0.9247
LTALS	0.9994	0.9980	0.9988	0.9987
TLIMSC	<b>1.0000</b>	<b>1.0000</b>	<b>1.0000</b>	<b>1.0000</b>

行 20 次并报告它们的平均结果。对于每个数据集，我们以粗体突出显示关于不同度量的最佳性能，“SC1”表示谱聚类算法是在数据集的第一个视图上运行的，“SC2”、“SC3”等也类似。

从五个表中可以观察到，本文所提出的 TLIMSC 方法在所有测试数据集上实现了最佳聚类性能。例如，在 ORL 数据集上，在 ACC 和 NMI 方面，相对于第二好的方法 LTALS 的性能改进分别为 0.75% 和 0.48%。在 NUS 数据集上，TLIMSC 在 ACC 和 NMI 方面分别实现了 14.01% 和 18.14% 的改进，显著优于第二好的 ETLMSC。特别地，可以观察到基于奇异值分解的张量核范数的方法，即 TLIMSC、LTALS、ETLMSC、tMSC 和 WTNNM，通常比其他对比方法能够获得更好的聚类结果，这表明张量核范数的引入有效地捕获了多视图数据之间的高阶相关性，并且对于聚类至关重要。尽管如此，由于 TLIMSC 能够发掘谱嵌入之间的高阶相关性，并通过具有不同置信度的多个谱嵌入获得了最终的聚类结果，因此比其他方法表现得更好。

对于每个基准数据集，本文利用谱聚类算法(SC)来揭示不同视图的不同聚类能力，这反映了在学习时对不同视图进行加权的重要性。因此，TLIMSC 方法考虑了不同的谱嵌入表征对一致聚类结果的不同贡献。同时，在通过张量探索高阶相关性的过程中，多个不同视图对应的奇异值被不平等地对待。同样地，WTNNM 方法也优于 tMSC，因为后者忽略了多个子空间表示的奇异值的不同置信度。尽管如此，TLIMSC 在所有基准数据集上的表现都优于 WTNNM。依据这些实验结果，可以证明我们提出方法中的两种加权策略是相当有效的。

值得注意的是，LTMSC 是基于展开的张量核范数的方法，它是塔克秩的松弛替代。由于松弛性，LTMSC 方法在大多数基准数据集上获得的聚类性能比其他基于张量的方法差得多。例如，在 MSRCv1 数据集上，所提出的 TLIMSC 方法在 ACC 和 NMI 方面优于 LTMSC，分别提高了 18.6% 和 26.79% 的聚类性能，这验证了基于张量奇异值分解的张量核范数的有效性。依据这五个表，可以看出 MSCIAS 总体上表现良好，尤其是在 ORL 和 UCI 数据集上，这显示了基于完整空间学习技术的完整感知相似性的优势。此外，CSMSC 也取得了不错的结果，但在发现不同数据点之间更深层次的相关性和语义信息方面无法取得进一步的突破。因此，所提出的 TLIMSC 方法通过严格的张量分解发现了多视图数据点之间的高阶相关性，并更全面地考虑了多个谱嵌入之间的关系，从而获得最终一致的聚类结果。

### 4.3 消融实验

为了进一步分析，我们对 TLIMSC 进行了消融实验，包括张量的作用以及在谱嵌入和最终一致聚类之间架起桥梁的策略。从以上对比结果可以看出，仅考虑低秩张量子空间表示的方法(如 WTNNM、tMSC、ETLMSC)无法获得进一步满意的聚类结果。此外，它们都没有考虑到谱嵌入与最终结果之间的相关性。为了解决这些问题，本文提出了 TLIMSC 方法来捕获谱嵌入之间的高阶相关性，并通过具有不同自适应置信度的多个谱嵌入来实现最终的聚类性能。为了分别研究 TLIMSC 这两个部分的贡献，我们通过执行两个测试来进行进一步的实验。具体来说，对于第一个测试，我们将舍弃张量核范数项，仅留下第一项关于标准谱嵌入的学习。而对于第二个测试，我们学习到具有高阶相关性的谱嵌入之后，在累加整合的共识谱嵌入上执行  $k$ -均值算法，以此获得最终结果，这意味着每个视图的谱嵌入表征对于最终的聚类结果具有相同的置信度。

为了符号简单起见，我们将第一个测试表示为 TLIMSC-w1，类似地，将第二个测试表示为 TLIMSC-w2。这两个测试是在所有基准数据集上进行的，其在 NMI 方面的对比结果见于表 8。从表 8 中可以看出，与 TLIMSC-w1 和 TLIMSC-w2 相比，TLIMSC 在所有基准数据集中都实现了更优的聚类性能。因此，消融实验验证了通过张量捕获谱嵌入表征之间高阶相关性的重要性，以及每个视图的谱嵌入对实现最终一致聚类结果有不同等置信度的必要性。

表 8 TLIMSC 及其变体的比较结果(NMI)

变体	ORL	UCI	NUS	Scene-15	Hdigit
TLIMSC	<b>1.0000</b>	<b>0.9973</b>	<b>0.9542</b>	<b>0.9522</b>	<b>1.0000</b>
TLIMSC-w1	0.8764	0.9434	0.1156	0.5097	0.7792
TLIMSC-w2	0.9310	0.9887	0.7758	0.7769	0.9975

#### 4.4 运行时间对比

在本节中,我们记录了所有方法所消耗的平均运行时间(以 s 为单位),如表 9 所示.从表 9 中可以看出,除了  $SC_{best}$ (在数据集最好的视图上执行 SC 所得到的结果),具有计算复杂度为  $O(n \log n)$  的 TLIMSC 总是在四个基准数据集上运行最快,共同排名第一.相比之下,尽管 WTNNM 性能优越,它通常是 13 种比较方法中运行最慢的.除了 ETLMSC 之外,基于低秩张量的对比方法,即 LTMSC、WTNNM 和 tMSC,都需要较多的运行时间.在某种程度上,ETLMSC 改善了算法的执行时间,但它的性能表现通常不如 WTNNM 和 tMSC,这也表明了它们各有优缺点.综上所述,本文所提出的 TLIMSC 方法不

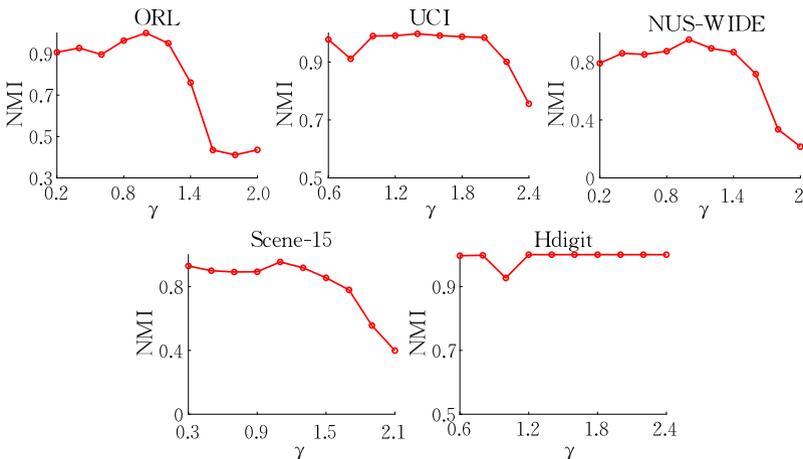
表 9 对比结果以平均运行时间表示 (单位:s)

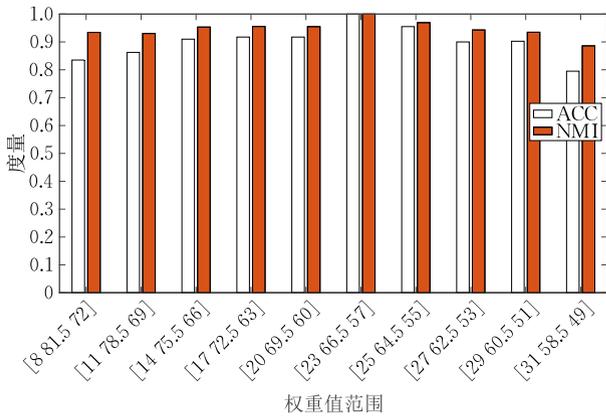
数据集	ORL	UCI	NUS	Scene-15	Hdigit
$SC_{best}$	0.75	<b>1.86</b>	<b>3.91</b>	22.98	107.05
CoTr	31.26	59.07	155.58	260.07	666.70
RMSC	4.96	97.19	171.19	1142.61	11705.72
CSMSC	18.13	61.91	138.67	763.66	622.29
LTMSC	30.92	206.37	647.51	1856.81	11095.46
LMSC	17.15	135.47	225.44	1383.54	36205.90
MSCIAS	26.01	32.26	48.36	170.76	523.04
MLAN	1.68	20.54	38.41	157.12	833.88
GMC	3.22	36.06	77.96	217.00	1185.08
WTNNM	61.41	248.35	838.35	2943.89	31523.49
tMSC	26.55	124.04	426.22	1197.46	10474.91
ETLMSC	3.68	2.92	288.38	570.35	7697.43
LTALS	24.43	375.75	1398.80	3999.00	19657.00
TLIMSC	<b>0.51</b>	2.23	5.30	<b>7.79</b>	<b>20.28</b>

仅实现了最佳的聚类性能,而且与大多数对比方法相比消耗了最少的计算时间,这进一步证明了捕获多个谱嵌入之间高阶相关性的有效性.

#### 4.5 参数分析

在 TLIMSC 中,有四个需要调整的权衡参数  $\gamma$ 、 $\omega$ 、 $\rho$  和  $\mu$ .为了研究这四个参数对不同数据集的影响,在本小节中,我们进行了参数分析.具体来说,对于参数  $\gamma$ ,在图 2 中,我们展示了在五个数据集中不同  $\gamma$  参数值对于 NMI 的影响.从图 2 可以看出,在所有的基准数据集上, $\gamma$  的选择对于模型的聚类效果起了很关键的作用.当模型取得最佳效果时, $\gamma$  的取值范围是在  $[0.8, 1.2]$  区间,这主要是由于所提出的谱嵌入矩阵学习项和加权张量核范数项对于模型学习是同等重要的,因而权衡参数  $\gamma$  的取值在 1.0 附近选取可以让两项学习得较好.对于不同的数据集, $\gamma$  具体的最佳选取也会所差异.对于加权系数  $\omega$ ,我们以第一个基准数据集,即 ORL 为例来进行参数分析,其结果报告于图 3.此图记录了不同权重值下 NMI 的结果.我们可以观察到,在给定的参数范围内,加权系数的取值对模型的聚类性能影响不大,而合适的加权系数可以实现和真实聚类标签匹配的聚类性能.例如,当 ORL 数据集的三个视图对应的参数  $\omega$  分别调整为  $[23, 66.5, 57]$  时,我们的方法可以达到更好的性能.对于参数  $\rho$  和  $\mu$ ,在图 4 中,我们展示了在五个数据集中不同取值的  $\rho$  和  $\mu$  对于 NMI 的影响.从图 4 中可以观察到,在 NUS 数据集上,不同  $\rho$  和  $\mu$  的取值会对模型的聚类效果影响较大,这主要是由于 NUS 数据集较稀疏的属性导致的.以往的研究常常忽略讨论这两参数不同

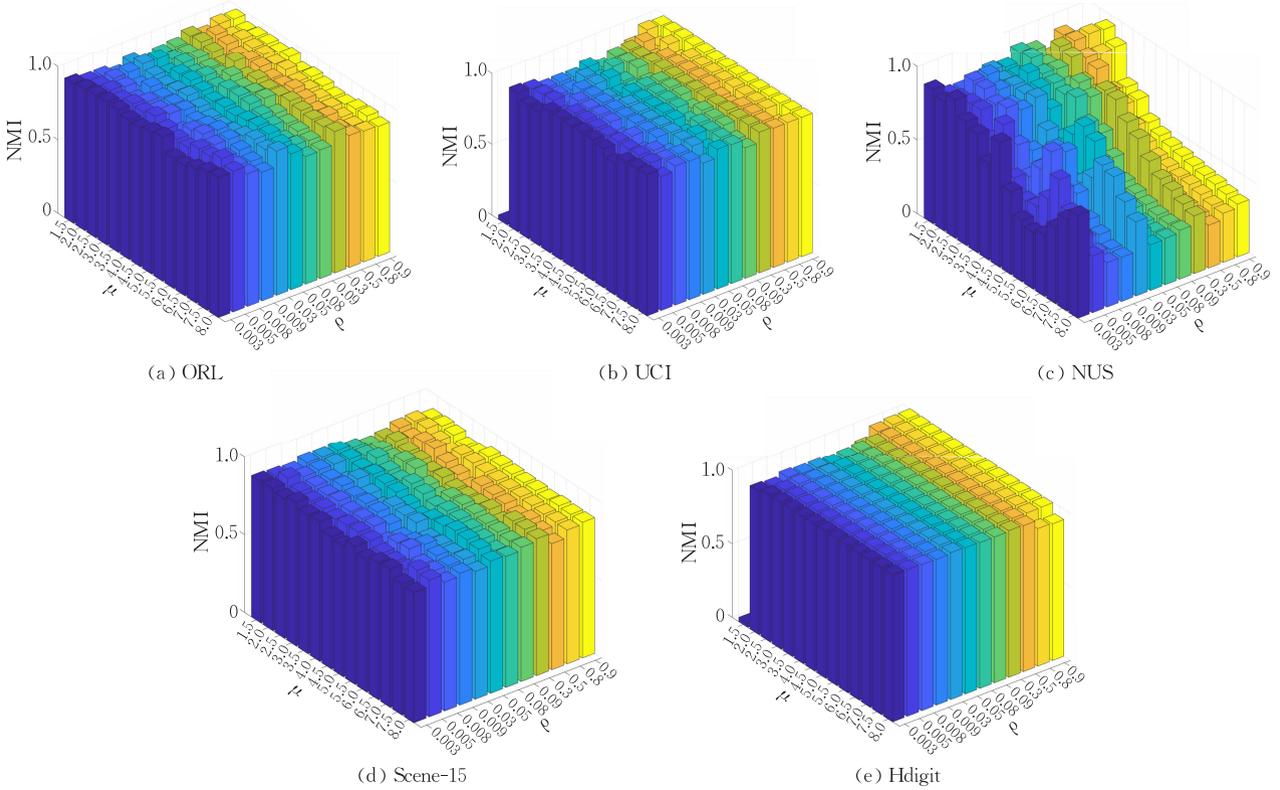
图 2  $\gamma$  在五个数据集上的参数分析

图 3  $\omega$  在 ORL 数据集三个视图上的参数分析

取值对聚类效果的影响,一般是直接默认其为定值.在该研究工作中,我们依据不同数据集的特定属性对参数值的选择进行了调整.为了清晰起见,我们将不同参数在五个基准数据集上设置的默认值归纳于表 10 中.

表 10 四个参数在五个基准数据集上的默认值

参数	ORL	UCI	NUS	Scene-15	Hdigit
$\gamma$	1	3	1	1.1	3
$\omega$	[23, 66.5, 57]	[12, 47, 45]	[15, 12, 5, 9, 8, 8]	[14, 30, 30]	[20, 8]
$\rho$	0.05	0.003	0.3	0.03	0.003
$\mu$	2	3.5	2.5	7.5	3.5

图 4  $\rho$  和  $\mu$  在五个数据集上的参数分析

#### 4.6 收敛性分析

在本小节中,我们对四个基准数据集进行了收敛性分析,以验证 TLIMSC 的收敛性,其中涉及了 TLIMSC 的两个部分.具体来说,对于 TLIMSC 的第一部分,我们记录了每次迭代步骤对应的变量误差  $\sum_{v=1}^V \|F_{(k+1)}^{(v)} - J_{(k+1)}^{(v)}\|_{\infty}$ ,其结果绘制在图 5 中.在图 5 中,横轴表示迭代次数,纵轴表示上述的变量误差.我们可以观察到,误差最终可以在 10 次迭代内趋于零,这意味着 TLIMSC 的第一部分能够在几次迭代后收敛.对于 TLIMSC 的第二部分,我们记录了每次

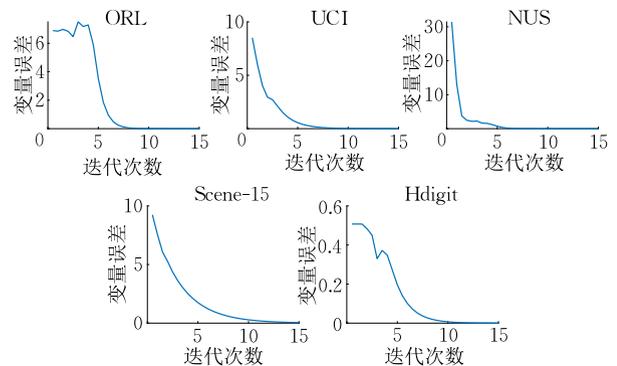


图 5 TLIMSC 第一部分模型的收敛性分析

迭代对应的目标函数值  $\sum_{v=1}^V \|P_{(k+1)} - F_{(k+1)}^{(v)} R_{(k+1)}^{(v)}\|_F$ ,

其结果绘制在图 6 中. 在图 6 中, 类似地, 横轴和纵轴分别表示迭代次数和目标函数值. 可以看出, 目标值在 5 次迭代内就可以得到一个稳定的值, 这说明 TLIMSC 的第二部分也可以在仅仅几次迭代后收敛. 综上所述, 所提出的 TLIMSC 方法可以保证在几次迭代后收敛.

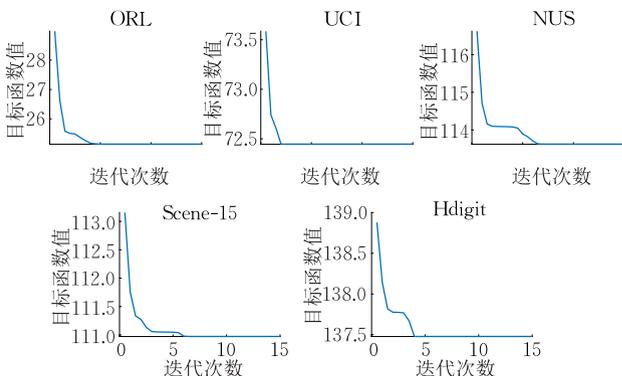


图 6 TLIMSC 第二部分模型的收敛性分析

## 5 总结与展望

本文提出了一种新颖的张量学习诱导的多视图谱聚类(TLIMSC)方法. 受基于 t-SVD 加权张量核范数约束方法的启发, 本文从关联样本和簇关系的多个谱嵌入表征堆叠的张量中恢复了多视图之间的高阶相关性. 借助张量旋转, 本文可以更有效地优化所提出的 TLIMSC 方法, 将其计算复杂度降低为  $O(n \log n)$ , 并且可以很好地研究跨视图之间的主要组成成分. 在获得谱嵌入表征后, 我们将带有自适应置信度的多个谱嵌入表征与最终的一致聚类结果巧妙地关联起来. 与现有最先进的多视图聚类方法相比, 在四个数据集上的大量实验结果证明了所提出的 TLIMSC 方法的有效性和高效性. 对于未来的工作, 我们计划利用张量学习来处理不完整多视图聚类中的挑战性问题.

## 参 考 文 献

- [1] Wang Chang-Dong, Lai Jian-Huang, Yu P S. Multi-view clustering based on belief propagation. *IEEE Transactions on Knowledge and Data Engineering*, 2016, 28(4): 1007-1021
- [2] Tao Hong, Hou Chenping, Yi Dongyun, Zhu Jubo. Multiview classification with cohesion and diversity. *IEEE Transactions on Cybernetics*, 2020, 50(5): 2124-2137
- [3] Zhang Changqing, Fu Huazhu, Hu Qinghua, et al. Generalized latent multi-view subspace clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(1): 86-99
- [4] Liu Bao-Yu, Huang Ling, Wang Chang-Dong, et al. Multiview consensus proximity learning for clustering. *IEEE Transactions on Knowledge and Data Engineering*, 2022, 34(7): 3405-3417
- [5] Chen Mansheng, Huang Ling, Wang Chang-Dong, Huang Dong. Multi-view clustering in latent embedding space// *Proceedings of the 34th AAAI Conference on Artificial Intelligence*. New York, USA, 2020: 3513-3520
- [6] Chen Mansheng, Lin Jia-Qi, Li Xiang-Long, et al. Representation learning in multiview clustering: A literature review. *Data Science and Engineering*, 2022, 7(3): 225-241
- [7] Hou Chenping, Nie Feiping, Tao Hong, Yi Dongyun. Multi-view unsupervised feature selection with adaptive similarity and view weight. *IEEE Transactions on Knowledge and Data Engineering*, 2017, 29(9): 1998-2011
- [8] Liu Xinwang, Zhu Xinzong, Li Miaomiao, et al. Late fusion incomplete multi-view clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019, 41(10): 2410-2423
- [9] Zhou Sihang, Liu Xinwang, Li Miaomiao, et al. Multiple kernel clustering with neighbor-kernel subspace segmentation. *IEEE Transactions on Neural Networks and Learning Systems*, 2020, 31(4): 1351-1362
- [10] Xu Jie, Ren Yazhou, Li Guofeng, et al. Deep embedded multi-view clustering with collaborative training. *Information Sciences*, 2021, 573: 279-290
- [11] Xu Jie, Ren Yazhou, Tang Huayi, et al. Multi-VAE: Learning disentangled view-common and view-peculiar visual representations for multiview clustering// *Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision*. Montreal, Canada, 2021: 9214-9223
- [12] Xu Jie, Tang Huayi, Ren Yazhou, et al. Multilevel feature learning for contrastive multi-view clustering// *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. New Orleans, USA, 2022: 16030-16039
- [13] Xie Yuan, Tao Dacheng, Zhang Wensheng, et al. On unifying multi-view self-representations for clustering by tensor multi-rank minimization. *International Journal of Computer Vision*, 2018, 126(11): 1157-1179
- [14] Liu Bao-Yu, Huang Ling, Wang Chang-Dong, et al. Adaptively weighted multiview proximity learning for clustering. *IEEE Transactions on Cybernetics*, 2021, 51(3): 1571-1585
- [15] Zhan Kun, Shi Jinhui, Wang Jing, Tian Feng. Graph-regularized concept factorization for multi-view document clustering. *Journal of Visual Communication and Image Representation*, 2017, 48: 411-418
- [16] Huang Shudong, Kang Zhao, Tsang I W, Xu Zenglin. Auto-weighted multi-view clustering via kernelized graph learning. *Pattern Recognition*, 2019, 88: 174-184

- [17] Nie Feiping, Li Jing, Li Xuelong. Selfweighted multiview clustering with multiple graphs//Proceedings of the 26th International Joint Conference on Artificial Intelligence. Melbourne, Australia, 2017; 2564-2570
- [18] Kumar A, Rai P, Daumé III H. Co-regularized multi-view spectral clustering//Advances in Neural Information Processing Systems 24: 25th Annual Conference on Neural Information Processing Systems 2011. Granada, Spain, 2011; 1413-1421
- [19] Cai Xiao, Nie Feiping, Huang Heng. Multiview  $k$ -means clustering on big data//Proceedings of the 23rd International Joint Conference on Artificial Intelligence. Beijing, China, 2013; 2598-2604
- [20] Ren Zhenwen, Sun Quansen. Simultaneous global and local graph structure preserving for multiple kernel clustering. IEEE Transactions on Neural Networks and Learning Systems, 2021, 32(5): 1839-1851
- [21] Xia Rongkai, Pan Yan, Du Lei, Yin Jian. Robust multi-view spectral clustering via lowrank and sparse decomposition//Proceedings of the 28th AAAI Conference on Artificial Intelligence. Québec City, Canada, 2014; 2149-2155
- [22] Zhang Changqing, Fu Huazhu, Liu Si, et al. Low-rank tensor constrained multiview subspace clustering//Proceedings of the IEEE International Conference on Computer Vision. Santiago, Chile, 2015; 1582-1590
- [23] Zhang Xiaoqian, Sun Huaijiang, Liu Zhigui, et al. Robust low-rank kernel multi-view subspace clustering based on the Schatten  $p$ -norm and correntropy. Information Sciences, 2019, 477: 430-447
- [24] Gao Hongchang, Nie Feiping, Li Xuelong, Huang Heng. Multi-view subspace clustering//Proceedings of the IEEE International Conference on Computer Vision. Santiago, Chile, 2015; 4238-4246
- [25] Zhang Changqing, Hu Qinghua, Fu Huazhu, et al. Latent multi-view subspace clustering//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA, 2017; 4279-4287
- [26] Wu Jianlong, Lin Zhouchen, Zha Hongbin. Essential tensor learning for multi-view spectral clustering. IEEE Transactions on Image Processing, 2019, 28(12): 5910-5922
- [27] Xu Huiling, Zhang Xiangdong, Xia Wei, et al. Low-rank tensor constrained co-regularized multi-view spectral clustering. Neural Networks, 2020, 132: 245-252
- [28] Gao Quanxue, Xia Wei, Wan Zhizhen, et al. Tensor-SVD based graph learning for multi-view subspace clustering//Proceedings of the 34th AAAI Conference on Artificial Intelligence. New York, USA, 2020; 3930-3937
- [29] Chen Yongyong, Xiao Xiaolin, Zhou Yicong. Multi-view subspace clustering via simultaneously learning the representation tensor and affinity matrix. Pattern Recognition, 2020, 106: 107441
- [30] Chen Man-Sheng, Wang Chang-Dong, Lai Jian-Huang. Low-rank tensor based proximity learning for multi-view clustering. IEEE Transactions on Knowledge and Data Engineering, 2023, 35(5): 5076-5090
- [31] Liu Guangcan, Lin Zhouchen, Yan Shuicheng, et al. Robust recovery of subspace structures by low-rank representation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(1): 171-184
- [32] Cao Xiaochun, Zhang Changqing, Fu Huazhu, et al. Diversity-induced multi-view subspace clustering//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA, 2015; 586-594
- [33] Li Ruihuang, Zhang Changqing, Hu Qinghua, et al. Flexible multi-view representation learning for subspace clustering//Proceedings of the 28th International Joint Conference on Artificial Intelligence. Macao, China, 2019; 2916-2922
- [34] Chen Mansheng, Huang Ling, Wang Chang-Dong, et al. Relaxed multi-view clustering in latent embedding space. Information Fusion, 2021, 68: 8-21
- [35] Bickel S, Scheffer T. Multiview clustering//Proceedings of the 4th IEEE International Conference on Data Mining. Brighton, UK, 2004; 19-26
- [36] Kumar A, Daumé III H. A co-training approach for multi-view spectral clustering//Proceedings of the 28th International Conference on Machine Learning. Bellevue, USA, 2011; 393-400
- [37] Cao Xiaochun, Zhang Changqing, Zhou Chengju, et al. Constrained multi-view video face clustering. IEEE Transactions on Image Processing, 2015, 24(11): 4381-4393
- [38] Wang Xiaobo, Lei Zhen, Guo Xiaojie, et al. Multiview subspace clustering with intactness-aware similarity. Pattern Recognition, 2019, 88: 50-63
- [39] Shu Le, Latecki L J. Integration of single-view graphs with diffusion of tensor product graphs for multi-view spectral clustering//Proceedings of the 7th Asian Conference on Machine Learning. Hong Kong, 2015; 362-377
- [40] Wang Hao, Yang Yan, Liu Bing. GMC: Graphbased multi-view clustering. IEEE Transactions on Knowledge and Data Engineering, 2020, 32(6): 1116-1129
- [41] Zhan Kun, Zhang Changqing, Guan Junpeng, Wang Junsheng. Graph learning for multiview clustering. IEEE Transactions on Cybernetics, 2018, 48(10): 2887-2895
- [42] Nie Feiping, Cai Guohao, Li Jing, Li Xuelong. Auto-weighted multi-view learning for image clustering and semi-supervised classification. IEEE Transactions on Image Processing, 2018, 27(3): 1501-1511
- [43] Tzortzis G, Likas A. Kernelbased weighted multi-view clustering//Proceedings of the 12th IEEE International Conference on Data Mining. Brussels, Belgium, 2012; 675-684
- [44] Wang X, Guo X, Lei Z, et al. Exclusivity-consistency regularized multi-view subspace clustering//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA, 2017; 1-9
- [45] Luo Shirui, Zhang Changqing, Zhang Wei, Cao Xiaochun. Consistent and specific multiview subspace clustering//Proceedings of the 32nd AAAI Conference on Artificial Intelligence. New Orleans, USA, 2018; 3730-3737

- [46] Tang Chang, Zhu Xinzhong, Liu Xinwang, et al. Learning a joint affinity graph for multiview subspace clustering. *IEEE Transactions on Multimedia*, 2019, 21(7): 1724-1736
- [47] Zhou Tao, Zhang Changqing, Peng Xi, et al. Dual shared-specific multiview subspace clustering. *IEEE Transactions on Cybernetics*, 2020, 50(8): 3517-3530
- [48] Zhang Chen, Wang Siwei, Liu Jiyuan, et al. Multi-view clustering via deep matrix factorization and partition alignment // *Proceedings of the ACM Multimedia Conference*. Virtual Event, China, 2021: 4156-4164
- [49] Kilmer M E, Martin C D. Factorization strategies for third-order tensors. *Linear Algebra and Its Applications*, 2011, 435(3): 641-658
- [50] Kilmer M E, Karen B S, Hao Ning, Randy H C. Third-order tensors as operators on matrices: A theoretical and computational framework with applications in imaging. *SIAM Journal on Matrix Analysis and Applications*, 2013, 34(1): 148-172
- [51] Chen Yongyong, Xiao Xiaolin, Zhou Yicong. Jointly learning kernel representation tensor and affinity matrix for multi-view clustering. *IEEE Transactions on Multimedia*, 2020, 22(8): 1985-1997
- [52] Nie Feiping, Tian Lai, Li Xuelong. Multiview clustering via adaptively weighted procrustes // *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. London, UK, 2018: 2022-2030
- [53] Huang Jin, Nie Feiping, Huang Heng. Spectral rotation versus  $k$ -means in spectral clustering // *Proceedings of the 27th AAAI Conference on Artificial Intelligence*. Bellevue, USA, 2013: 431-437
- [54] Chen Kun, Dong Hongbo, Chan Kung-Sik. Reduced rank regression via adaptive nuclear norm penalization. *Biometrika*, 2013, 100(4): 901-920
- [55] Chen Yongyong, Wang Shuqin, Zheng Fangying, Cen Yigang. Graph-regularized least squares regression for multi-view subspace clustering. *Knowledge-Based Systems*, 2020, 194: 105482
- [56] Lin Zhouchen, Chen Minming, Ma Yi. The augmented Lagrange multiplier method for exact recovery of corrupted low-rank matrices. *CoRR abs/1009.5055*, 2010
- [57] Ng A Y, Jordan M I, Weiss Y. On spectral clustering: Analysis and an algorithm // *Advances in Neural Information Processing Systems 14 Neural Information Processing Systems: Natural and Synthetic*. Vancouver, Canada, 2002: 849-856
- [58] Fu Lele, Yang Jinghua, Chen Chuan, Zhang Chuanfu. Low-rank tensor approximation with local structure for multi-view intrinsic subspace clustering. *Information Sciences*, 2022, 606: 877-891
- [59] Wang Chang-Dong, Chen Mansheng, Huang Ling, et al. Smoothness regularized multiview subspace clustering with kernel learning. *IEEE Transactions on Neural Networks and Learning Systems*, 2021, 32(11): 5047-5060

**CHEN Man-Sheng**, Ph. D. candidate.

Her research interests include data mining and multi-view clustering.



include data mining and multi-view clustering.

**WANG Chang-Dong**, Ph. D., associate professor. His research interests include data clustering, network analysis, recommendation algorithm and big data information security.

**HUANG Dong**, Ph. D., associate professor. His research interests include artificial intelligence, big data analysis, deep learning, graph neural network.

**LAI Jian-Huang**, Ph. D., professor. His research interests include biometric identification, digital image processing, pattern recognition and machine learning.

**CAI Xiao-Sha**, Ph. D. candidate. Her research interests include data mining and multi-view clustering.

**LIN Jia-Qi**, Ph. D. candidate. His research interests

## Background

Multi-view data generated from diverse feature extractors or collected from different domains become extremely common in many real-world applications, which induces the appearance of multi-view learning by combining these multiple views for data analysis. Multi-view clustering is one of the most widely studied multi-view learning tasks. The key challenge of multi-view clustering is how to integrate multiple views and combine the underlying semantic information (e. g. the higher-order information) of these representations to explore the spatial

structure for better clustering performances. In the past few years, considerable efforts have been made in developing effective multi-view clustering methods, such as co-regularized multi-view clustering, multi-view  $k$ -means clustering, low-rank representation-based multi-view clustering and so on. In particular, due to the superior performances, the low-rank representation-based multi-view clustering methods become the mainstream. Motivated by the impressive results of low-rank representation (LRR), Zhang et al. firstly developed a

low-rank tensor constrained multi-view subspace clustering (LTMSC). To avoid the Tucker tensor decomposition of LTMSC, t-SVD based multi-view subspace clustering (tMSC) was proposed. Based on the tMSC method, WTNM for multi-view subspace clustering further studied the distinct regularization of different singular values.

Despite the great success of these methods, they have some shortcomings. On the one hand, they spend much time for algorithm learning, with their computation complexity being at least  $O(n^2 \log n)$ , and the blame lies in the insistence of tensor constructed by the learned subspace representation, in which relationships among  $n$  samples are under study. On the other hand, after the tensor learning, the standard spectral clustering or  $k$ -means algorithm is directly employed on the integrated learned representation, being common but deficient. Specifically, each learned representation of individual view makes different contributions towards the final clustering performances, which should be treated unequally during the learning procedure. To address the aforementioned challenges, we propose a novel Tensor Learning Induced Multi-view Spectral Clustering (TLIMSC) method, where the spatial cluster structures and complementary information are simultaneously explored. To be specific, inspired by the idea of anchor graph, the relationships of the learned representations

between  $n$  samples and  $c$  clusters are explored, and the spectral embedding matrices learned from multiple views are considered. Then, we utilize multiple spectral embedding representations to construct the tensor, and the higher-order correlations from the samples to the clusters would be discovered by the t-SVD based weighted tensor nuclear norm, which takes the salient cues between different singular values of the matrix into account during the tensor learning. Besides, after obtaining the improved representations with higher-order correlations, the bridge between the final consensus clustering results and the representations would be built, in which different contributions of different representations towards the final clustering results can be adaptively learned. Consequently, the main purpose of this paper is to study an effective and efficient multi-view clustering methods so as to not only capture the higher-order correlations between the samples and the clusters via t-SVD based weighted tensor nuclear norm, but also consider different confidences of multiple representations towards the final clustering results.

This work was supported by the National Natural Science Foundation of China (No. 62276277), the Guangdong Basic and Applied Basic Research Foundation(No. 2022B1515120059), and the Guangdong Provincial Key Laboratory of Intellectual Property and Big Data (No. 2018B030322016).