

单步划分融合多视图子空间聚类算法^{*}

张 培, 祝 恩, 蔡志平⁺

国防科技大学 计算机学院, 长沙 410073

⁺ 通信作者 E-mail: zpcai@nudt.edu.cn

摘 要: 多视图子空间聚类方法因其可以揭示数据内在的低维结构而被广泛关注, 但大多数现有的多视图子空间聚类算法直接将多个来自原始数据的充满噪声的相似度矩阵进行融合, 并且通常是在得到一致的多视图表示之后再使用 K 均值算法聚类得到最终的结果, 这种将表示的学习过程和后续的聚类过程分离的两阶段算法可能会导致无法得到最优的聚类结果。为了解决这些问题, 本文提出一种单步划分融合多视图子空间聚类算法, 不是直接融合具有噪声和冗余信息的相似度矩阵, 而是从相似度矩阵中提取出更具有判别性信息的划分级信息进行融合。此外, 我们提出一个新的框架, 将表示学习、多视图信息融合以及最后的聚类过程整合在同一框架中, 这三个过程彼此促进, 好的聚类结果可以引导生成更好的多视图表示, 从而得到更好的聚类效果。本文提出一种有效的轮替优化算法来解决由此得到的优化问题。最后, 四个真实的基准数据集上得到的实验结果可以证明提出方法的有效性以及先进性。

关键词: 多视图聚类; 子空间学习; 数据融合; 划分融合

文献标志码: A **中图分类号:** TP301

张培, 祝恩, 蔡志平. 单步划分融合多视图子空间聚类算法[J]. 计算机科学与探索

ZHANG P, ZHU E, CAI Z P. One-stage Partition-fusion Multi-view Subspace Clustering Algorithm[J]. Journal of Frontiers of Computer Science and Technology

One-stage Partition-fusion Multi-view Subspace Clustering Algorithm^{*}

ZHANG Pei, ZHU En, CAI Zhiping⁺

School of Computer, National University of Defense Technology, Changsha 410073, China

Abstract: Multi-view subspace clustering has attracted increasing attention for revealing the inherent low-dimension structure of the data. Nevertheless, most existing methods directly fuse the multiple noisy affinity matrices from the original data, and commonly conduct clustering after obtaining a unified multi-view representation. Separating the representation learning from the clustering process could result in a suboptimal clustering result. To this end, this paper proposes a One-stage Partition-fusion Multi-view Subspace Clustering algorithm. Instead of directly fusing the noisy and redundant affinity matrices, this paper fuses the more discriminative partition-level information extracted from the affinity matrices. Moreover, we propose a new framework, integrating representation learning, multiple information fusion and final clustering process. The three sub-process promote each other to serve clustering best. The promising clustering results would lead to better representations and therefore better clustering perfor-

^{*}The National Natural Science Foundation of China under Grant Nos. 62072465, 61672528, 61773392 (国家自然科学基金).

mance. Consequently, this paper solves the resultant optimization problem through an alternative algorithm. Experiment results on four real-world benchmark datasets show the effectiveness and superior performance of the proposed method over the state-of-the-art approaches.

Key words: multi-view clustering; subspace clustering; data fusion; partition fusion

1 引言

传统的聚类方法通过计算数据特征之间的相似度来对样本进行聚类。然而仅使用单一的特征可能会存在一定的偏差或不足。随着信息技术、数据挖掘技术等的发展,数据可以被不同类型的特征描述。例如相同语义使用多语言形式表示^[1];生物特征识别中可以使用人脸、指纹、掌纹以及虹膜等特征来唯一的标识一个人;网页包含为文本、链接以及视频等多种模态的数据。这种使用不同的特征来描述的数据称为多视图数据。多视图数据不仅包含每个视图下特定信息,也包含视图之间的互相补充的信息^[2]。使用多视图数据进行聚类可以充分发挥各视图的优势、规避各自的局限,从而获得更好的聚类性能。

现有的多视图聚类算法可以被分为四类:直接连接的方法、多核聚类算法^[3]、基于图的多视图聚类算法^[4]以及多视图子空间聚类算法^[5]。多视图子空间聚类算法通常包含基于矩阵分解的方法和基于子空间的方法。其中基于矩阵分解的方法^{[6][7]}旨在从所有视图学习一个统一的特征表示。而基于自表达的多视图子空间聚类算法^{[8][9]}在很多应用中展现了其有效性和鲁棒性。其关键在于学习统一的子空间表示,从而揭示多个视图下理想的一致聚类结构。现有的多视图子空间聚类方法大多是将多个图或表示矩阵融合成一致的图或表示矩阵,即在相似度级别上融合多视图信息。如文献^[10]通过执行矩阵分解得到一致的稀疏子空间表示。类似的,文献^[11]得到低秩和稀疏约束下的一致相似度矩阵。也有一些方法首先学习一个潜在表示,然后从潜在空间进行子空间聚类。文献^[12]通过最大化视图之间的独立性来编码视图之间的互补信息,进而得到一致的表示。与之类似的是文献

^[13],通过假设所有的视图都源于一个潜在表示,通过对潜在表示进行子空间聚类得到子空间表示。

尽管这些方法在不同的场景下都取得了极大的成功,但仍存在以下问题:原始数据中通常包含噪声和一些冗余的信息,所以直接将从原始数据中学习一致的相似图矩阵或者表示矩阵融合可能会影响最终的聚类性能。此外,现有的方法大多为两阶段,即先进行相似度学习再应用谱聚类或者其他离散化方法得到最终的聚类结果。这样将表示学习的过程与最终的聚类任务分离的方法会导致学到并不是最适合聚类任务的表示,从而无法得到最优的聚类性能。

本文提出了一种单步划分融合多视图子空间聚类算法(One-stage Partition-fusion Multi-view Subspace Clustering, OP-MVSC),将表示学习、划分融合以及聚类过程集成在一个框架中。在这个框架中,我们联合优化每个视图的相似性矩阵、划分矩阵以及聚类标签。具体来说,本文首先在每个视图上建立各自的子空间表示。不同的视图之间聚类结构应该是类似的,基于这一假设,我们将每个视图的划分矩阵融合为一致的划分矩阵。这种划分级别的融合可以避免相似度级别融合中引入的原始数据中噪声和冗余。此外,通过对一致的划分施以旋转矩阵可以直接得到聚类结果,避免使用额外的谱聚类或者其他离散化方法获得聚类标签。通过这三个子过程(表示学习、划分融合、谱旋转)的交互,它们之间互相促进,从而获得更好的聚类结果。除此之外,我们提出一种有效的算法来解决由此产生的优化问题。并在四个广泛使用的多视图数据集上与一些先进的方法进行了对比实验,证明了所提方法的有效性和先进性。

本文组织结构如下:第1节介绍文中要用到的符号与定义;第2节详细描述提出的方法及其优化

算法；第3节给出实验结果和实验分析；最后进行总结和展望。

2 相关工作

2.1 符号与定义

本文使用加粗的大写斜体字母表示矩阵。对于矩阵 \mathbf{A} , $\mathbf{A}_{i,:}$ 表示第 i 行, $\mathbf{A}_{:,j}$ 表示矩阵 \mathbf{A} 第 i 行、第 j 列处的元素。矩阵 \mathbf{A} 的 Frobenius 范数表示为 $\|\mathbf{A}\|_F$, 向量的 2 范数表示为 $\|\mathbf{A}_{i,:}\|_2$ 。矩阵的迹、转置分别表示为 $\text{Tr}(\mathbf{A})$ 及 \mathbf{A}^T 。

定义一个有 m 个视图, n 个样本的数据集为 $\mathbf{A} = [\mathbf{A}^1, \mathbf{A}^2, \dots, \mathbf{A}^m] \in \mathbb{R}^{d \times n}$, 其中, $d = \sum_{v=1}^m d_v$ 表示第 v 个视图中特征的维度。第 v 个视图的数据则表示为 $\mathbf{A}^v \in \mathbb{R}^{d_v \times n}$ 。此外使用粗体的 $\mathbf{1}$ 表示元素全为 1 的列向量, 使用 \mathbf{I}_k 表示 $k \times k$ 维的单位矩阵。

2.2 多视图子空间聚类相关工作

对于一组数据来说, 其数据点通常分布在潜在的低维子空间中而不是均匀分布在整个空间^[10]。因此数据点可以通过低维的子空间来表示。在获得数据的子空间结构后, 可以在子空间中进行聚类来代替在整个空间中聚类。给定 n 个样本的数据 $\mathbf{A} = \{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\} \in \mathbb{R}^{d \times n}$, 其中某个样本 \mathbf{a}_i 可以使用与它位于同一个子空间的其他样本的线性组合表示。选取 \mathbf{A} 中的 d 个样本构成字典 \mathbf{D} , 那么 \mathbf{A} 中每一个样本都可以表示为 \mathbf{D} 中数据点的线性组合。随着自表达性质^[14]的兴起, 可以使用整个数据矩阵作为字典来表示数据本身, 如公式(1)所示,

$$\min_{\mathbf{S}} \|\mathbf{A} - \mathbf{AS}\|_F^2 \quad \text{s.t. } 0 \leq S_{i,j} \leq 1, \mathbf{S}^T \mathbf{1} = \mathbf{1}, \text{diag}(\mathbf{S}) = 0, \quad (1)$$

其中 $\mathbf{S} = \{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_n\} \in \mathbb{R}^{n \times n}$ 是数据点的线性组合系数矩阵, 称为子空间表示, 每个 \mathbf{s}_i 是原始数据 \mathbf{a}_i 的子空间表示。其中, 约束 $\text{diag}(\mathbf{S}) = 0$ 要求 \mathbf{S} 的对角线元素为 0, 避免数据点由自己来表示。扩展到多视图场景, 范式如下:

$$\min_{\mathbf{S}} \sum_{v=1}^m \|\mathbf{A}^v - \mathbf{A}^v \mathbf{S}^v\|_F^2 + \lambda \Omega(\mathbf{S}^v) + \Phi(\mathbf{S}^v, \mathbf{S}^*) \quad (2)$$

$$\text{s.t. } 0 \leq S_{i,j}^v \leq 1, (\mathbf{S}^v)^T \mathbf{1} = \mathbf{1}, \text{diag}(\mathbf{S}^v) = 0,$$

其中 \mathbf{A}^v 表示第 v 个视图的数据, \mathbf{S}^v 表示第 v 个视图的子空间表示。 $\Omega(\cdot)$ 为某种正则化函数, $\Phi(\cdot)$ 为某种可将多个视图的子空间表示融合为一致表示的策略。 $\lambda > 0$ 是平衡参数。使用不同的正则化函数和不同的一致策略, 组成了不同的多视图子空间聚类方法。其中大部分方法在获得一致的表示 \mathbf{S}^* 后, 通过 $\mathbf{W} = 1/2(|\mathbf{S}^*| + |\mathbf{S}^{*T}|)$ 得到相似度矩阵, 然后进行谱分解得到划分矩阵 \mathbf{F} , 再对 \mathbf{F} 使用某种聚类算法 (通常是 K-means) 得到最终的聚类标签。

然而真实世界的的数据会有噪声、异常点或冗余信息, 这可能导致相似度矩阵质量较差, 直接融合这样充满噪声和冗余的多视图相似信息会导致较差的聚类结果。且后续得到聚类标签的过程与前面的相似度学习的过程分离, 这也导致了无法得到最优的聚类性能。为了解决这些问题, 我们提出一种新的单步划分融合多视图子空间聚类算法。

3 算法

本节介绍我们提出了单步划分融合多视图子空间聚类方法, 并给出相应的优化算法。

3.1 单步划分融合多视图子空间聚类算法

根据前文介绍的多视图子空间聚类范式(2), 我们可以得到每个视图对应的子空间表示:

$$\min_{\mathbf{S}^v} \sum_{v=1}^m \|\mathbf{A}^v - \mathbf{A}^v \mathbf{S}^v\|_F^2 + \lambda \|\mathbf{S}^v\|_F^2 \quad (3)$$

$$\text{s.t. } 0 \leq S_{i,j}^v \leq 1, (\mathbf{S}^v)^T \mathbf{1} = \mathbf{1}, \text{diag}(\mathbf{S}) = 0.$$

一个理想相似度矩阵中应具有如下性质: 相似度矩阵中的连通分量的个数等于聚类的簇的个数。根据文献^[15]可知, 相似度矩阵中连通分量的个数等于对应拉普拉斯矩阵中零特征值的重数。因此, 我们希望相似度矩阵对应的拉普拉斯矩阵的零特征值数等于聚类的簇的个数 k , 即 $\text{rank}(\mathbf{L}^v) = n - k$ 。自然地, 我们希望将这个秩约束加入到公式(3)中, 使得 \mathbf{S}^v 具有所希望的理想性

质。然而直接使用 $\text{rank}(\mathbf{L}^v) = n - k$ 的秩约束会使得该优化问题难以解决。根据 Ky Fan 定理^[16]，我们可以得到 $\sum_{i=1}^k \sigma_i(\mathbf{L}^v) = \min \text{Tr}((\mathbf{F}^v)^T \mathbf{L}^v \mathbf{F}^v)$ ，其中 $\sigma_i(\mathbf{L}^v)$ 是 \mathbf{L}^v 的第 i 小的特征值。很明显， \mathbf{L}^v 的前 k 小的特征值均为 0，即 $\sum_{i=1}^k \sigma_i(\mathbf{L}^v) = 0$ ，将使得 \mathbf{L}^v 的秩为 $n - k$ 。因此，问题可以自然转化为如下形式：

$$\min_{\mathbf{S}^v, \mathbf{F}^v} \sum_{v=1}^m \left(\|\mathbf{A}^v - \mathbf{A}^v \mathbf{S}^v\|_F^2 + \lambda \|\mathbf{S}^v\|_F^2 + \text{Tr}((\mathbf{F}^v)^T \mathbf{L}^v \mathbf{F}^v) \right) \quad (4)$$

$$\text{s.t. } 0 \leq \mathbf{S}_{i,j}^v \leq 1, (\mathbf{S}^v)^T \mathbf{1} = \mathbf{1}, \text{diag}(\mathbf{S}) = 0, (\mathbf{F}^v)^T \mathbf{F}^v = \mathbf{I}_k,$$

其中 $\mathbf{L}^v = \mathbf{D}^v - 1/2(\mathbf{S}^v + (\mathbf{S}^v)^T)$ ， \mathbf{D}^v 是 \mathbf{S}^v 的度矩阵，其第 i 个对角元素为 $\mathbf{D}_{i,i}^v = \sum_{j=1}^n \mathbf{S}_{i,j}^v$ 。

在公式(4)中，每个视图可以得到各自的划分矩阵。在多视图聚类中，基于不同视图之间的聚类结构应该是类似的假设，即无论在哪个视图，相似的样本都应该被聚在同一个簇中。因此我们将每个视图的划分 \mathbf{F}^v 对齐，形式化如下：

$$\min_{\mathbf{F}^*, \mathbf{F}^v} \sum_{v=1}^m \|\mathbf{F}^v - \mathbf{F}^*\|_F^2 \quad (5)$$

$$\text{s.t. } (\mathbf{F}^v)^T \mathbf{F}^v = \mathbf{I}_k, (\mathbf{F}^*)^T \mathbf{F}^* = \mathbf{I}_k.$$

为了建立划分矩阵与最后的聚类标签之间的联系，我们引入旋转矩阵 $\mathbf{P} \in \mathbb{R}^{k \times k}$ ：

$$\min_{\mathbf{Y}, \mathbf{P}} \sum_{i=1}^n \sum_{c=1}^k (y_{ic})^\gamma \|t_c - \mathbf{F}_{i,:}^* \mathbf{P}\|_2^2 \quad (6)$$

$$\text{s.t. } y_{ic} \geq 0, \mathbf{Y}_{i,:} \mathbf{1} = 1, (\mathbf{P})^T \mathbf{P} = \mathbf{I}_k.$$

式中 γ 为模糊系数。 t_c 为 $1 \times k$ 维的向量，其中只有第 c 个元素为 1 其余元素为 0。 $\mathbf{F}_{i,:}^*$ 为 \mathbf{F}^* 的第 i 行，对应第 i 个样本的一致表示。 y_{ic} 表示第 i 个样本属于第 c 个簇的概率。 \mathbf{P} 则建立了标签 \mathbf{Y} 和一致表示 \mathbf{F}^* 之间合理的关系。如果样本 i 在位置 c 显示出突出的结构，则对应 y_{ic} 有一个较大的概率值，表示样本 i 较大概率属于第 c 个簇。

将公式(4)(5)(6)整合在一起，我们可以得到整个框架：

$$\min_{\mathbf{S}^v, \mathbf{F}^v, \mathbf{F}^*, \mathbf{Y}, \mathbf{P}} \sum_{v=1}^m \left(\|\mathbf{A}^v - \mathbf{A}^v \mathbf{S}^v\|_F^2 + \lambda \|\mathbf{S}^v\|_F^2 + \text{Tr}((\mathbf{F}^v)^T \mathbf{L}^v \mathbf{F}^v) \right) + \sum_{v=1}^m \|\mathbf{F}^v - \mathbf{F}^*\|_F^2 + \sum_{i=1}^n \sum_{c=1}^k (y_{ic})^\gamma \|t_c - \mathbf{F}_{i,:}^* \mathbf{P}\|_2^2 \quad (7)$$

$$\text{s.t. } 0 \leq \mathbf{S}_{i,j}^v \leq 1, (\mathbf{S}^v)^T \mathbf{1} = \mathbf{1}, (\mathbf{F}^v)^T \mathbf{F}^v = \mathbf{I}_k,$$

$$(\mathbf{F}^*)^T \mathbf{F}^* = \mathbf{I}_k, \mathbf{P}^T \mathbf{P} = \mathbf{I}_k, y_{ic} \geq 0, \mathbf{Y}_{i,:} \mathbf{1} = 1$$

通过这种方式，相似度矩阵、一致划分及标签矩阵得以在一个框架中联合优化，并且这三个过程可以互相促进从而获得更好的聚类结果。

3.2 优化过程

3.2.1 固定 $\mathbf{F}^v, \mathbf{F}^*, \mathbf{P}$ 优化 \mathbf{S}^v 。

忽略无关项，目标函数的优化式可转化为：

$$\min_{\mathbf{S}^v} \|\mathbf{A}^v - \mathbf{A}^v \mathbf{S}^v\|_F^2 + \lambda \|\mathbf{S}^v\|_F^2 + \text{Tr}((\mathbf{F}^v)^T \mathbf{L}^v \mathbf{F}^v) \quad (8)$$

由于每个视图都需要进行子空间表示的建立，因此可以对每个视图的 \mathbf{S}^v 进行分别优化。将 \mathbf{S}^v 按列展开可以得到如下向量形式：

$$\min_{\mathbf{S}_{:,i}} \|\mathbf{A}_{:,i} - \mathbf{A} \mathbf{S}_{:,i}\|^2 + \lambda \mathbf{S}_{:,i}^T \mathbf{S}_{:,i} + \frac{1}{2} \mathbf{Q}_{i,:} \mathbf{S}_{:,i},$$

其中， $\mathbf{Q}_{i,j} = \|\mathbf{F}_{i,:} - \mathbf{F}_{j,:}\|^2$ ， $\mathbf{F}_{i,:}$ 为对应 \mathbf{F} 的第 i 行。对 $\mathbf{S}_{:,i}$ 求导并令导数为零可以得到 $\mathbf{S}_{:,i}$ 的闭式解：

$$\mathbf{S}_{:,i} = (\mathbf{X}^T \mathbf{X} + \lambda \mathbf{I})^{-1} (\mathbf{X}^T \mathbf{X}_{:,i} - \frac{1}{4} \mathbf{Q}_{i,:}^T) \quad (9)$$

3.2.2 固定 $\mathbf{S}^v, \mathbf{F}^v, \mathbf{P}, \mathbf{Y}$ ，优化 \mathbf{F}^* 。

去掉其他无关项，关于 \mathbf{F}^* 的优化式可转化为：

$$\min_{\mathbf{F}^*} \sum_{v=1}^m \|\mathbf{F}^v - \mathbf{F}^*\|_F^2 + \sum_{i=1}^n \sum_{c=1}^k (y_{ic})^\gamma \|t_c - \mathbf{F}_{i,:}^* \mathbf{P}\|_2^2 \quad (10)$$

$$\text{s.t. } (\mathbf{F}^*)^T \mathbf{F}^* = \mathbf{I}_k$$

定理 1： 求解问题 $\min \sum_{i=1}^n \sum_{c=1}^k (y_{ic})^\gamma \|t_c - \mathbf{F}_{i,:}^* \mathbf{P}\|_2^2$ 等价于 $\max \text{Tr}(\mathbf{P}^T (\mathbf{F}^*)^T \mathbf{G})$ ，其中 $\mathbf{G}_{i,:} = \sum_{c=1}^k (y_{ic})^\gamma t_c$ 。

证明： 易得 $\min \sum_{i=1}^n \sum_{c=1}^k (y_{ic})^\gamma \|t_c - \mathbf{F}_{i,:}^* \mathbf{P}\|_2^2$ 等价于 $\max \sum_{i=1}^n \sum_{c=1}^k (y_{ic})^\gamma \text{Tr}(\mathbf{P}^T (\mathbf{F}_{i,:}^*)^T t_c)$ ，按元素展开可得 $\text{Tr}(\mathbf{P}^T (\mathbf{F}^*)^T t_c) = \mathbf{P}_{c,:}^T (\mathbf{F}^*)_{:,i}^T$ 。由 $\mathbf{G}_{i,:} = \sum_{c=1}^k (y_{ic})^\gamma t_c$ ，那么 $\sum_{i=1}^n \sum_{c=1}^k (y_{ic})^\gamma \mathbf{P}_{c,:}^T (\mathbf{F}^*)_{:,i}^T = \text{Tr}(\mathbf{P}^T (\mathbf{F}^*)^T \mathbf{G})$ 。从而得到 $\min \sum_{i=1}^n \sum_{c=1}^k (y_{ic})^\gamma \|t_c - \mathbf{F}_{i,:}^* \mathbf{P}\|_2^2$ 等价于 $\max \text{Tr}(\mathbf{P}^T (\mathbf{F}^*)^T \mathbf{G})$ ，其中 $\mathbf{G} \in \mathbb{R}^{n \times k}$ ，第 i 行为 $\mathbf{G}_{i,:} = \sum_{c=1}^k (y_{ic})^\gamma t_c$ 。

定理 2： 若矩阵 \mathbf{A} 的秩为 k 的奇异值分解 (SVD) 为 $\mathbf{A} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T$ ，则问题

$$\max_X \text{Tr}(X^T A) \quad s.t. X^T X = I \quad (11)$$

最优解为 $X = UV^T$ 。

证明： $\text{Tr}(X^T A) = \text{Tr}(X^T U \Sigma V^T) = \text{Tr}(V^T X^T U \Sigma)$ ，令 $H = V^T X^T U$ ，根据西矩阵的性质以及约束条件可得 $HH^T = V^T X^T U U^T X V = I$ 。因此 $\text{Tr}(V^T X^T U \Sigma) = \text{Tr}(H \Sigma) \leq \text{Tr}(I \Sigma) = \sum_{i=1}^k \sigma_i$ 。 σ_i 为 Σ 对角线上的第 i 个元素。则当 $H \Sigma = I \Sigma$ 即 $V^T X^T U = I$ 时，问题(11)可得到 X 的最优解 $X = UV^T$ 。

根据公式(10)及定理 1，令 $N = \sum_{v=1}^m F^v + G P^T$ ，则问题(10)等价于 $\max_{F^*} \text{Tr}((F^*)^T N) \quad s.t. (F^*)^T F^* = I_k$ 。根据定理 2，最优的 F^* 可通过对 N 奇异值分解得到。

3.2.3 固定 S^v, F^*, P , 优化 F^v 。

关于 F^v 的优化式可以转化为：

$$\begin{aligned} \min_{F^v} \quad & \text{Tr}((F^v)^T L F^v) - 2\text{Tr}((F^v)^T F^*) \\ s.t. \quad & (F^v)^T F^v = I_k \end{aligned} \quad (12)$$

令 λ_{\max} 为 L 的最大特征值，优化式放松为：

$$\begin{aligned} \max_{F^v} \quad & \text{Tr}((F^v)^T (\lambda_{\max} I - L) F^v) + 2\text{Tr}((F^v)^T F^*) \\ s.t. \quad & (F^v)^T F^v = I_k, (F^*)^T F^* = I_k \end{aligned}$$

该问题可以通过文献^[17]中的方法来解决，具体算法流程如下：

算法 1 更新 F^v

- 1: 重复执行
- 2: $M = 2(\lambda_{\max} I - L) F^v + 2F^*$
- 3: 解 $\max \text{Tr}((F^v)^T M) \quad s.t. (F^v)^T F^v = I_k$ 更新 F^v
- 4: 对 M 执行 SVD 得到 $M = U_1 \Sigma_1 V_1^T$
- 5: $F^v = U_1 V_1^T$
- 6: 直至算法收敛

3.2.4 固定 S^v, F^v, F^*, Y , 优化 P 。

忽略其他无关项， P 的优化式可以转化为：

$$\begin{aligned} \min_P \quad & \sum_{i=1}^n \sum_{c=1}^k (y_{ic})^\gamma \|t_c - F_{i,:}^* P\|_2^2 \\ s.t. \quad & P^T P = I_k \end{aligned} \quad (13)$$

该式等价于 $\max_P \text{Tr}(P^T (F^*)^T G) \quad s.t. P^T P = I_k$ 。对 $(F^*)^T G$ 进行奇异值分解即可得到 P 的最优解。

3.2.5 固定 S^v, F^v, F^*, P 优化 Y 。

忽略其他无关项，关于 Y 的优化可以转化为：

$$\begin{aligned} \min_Y \quad & \sum_{i=1}^n \sum_{c=1}^k (y_{ic})^\gamma \|t_c - F_{i,:}^* P\|_2^2 \\ s.t. \quad & y_{ic} \geq 0, Y_{i,:} \mathbf{1}_k = 1 \end{aligned} \quad (14)$$

令 $e_{i,c} = \|t_c - F_{i,:}^* P\|_2$ 。当 $\gamma = 1$ 时，可以直接得到对应位置的聚类标签： $y_{ic} = \arg \min_j e_{ij}$ 。当 $\gamma > 1$ 时，通过其拉普拉斯函数对 y_{ic} 求导并令导数等于零，可得到：

$$y_{ic} = \frac{(e_{ic})^{\frac{1}{1-\gamma}}}{\sum_{c=1}^k (e_{ic})^{\frac{1}{1-\gamma}}}$$

3.3 时间复杂性分析

整个算法流程如算法 2 所示。在整个优化过程中，算法 OP-MVSC 的计算复杂性可以分为三个子阶段：相似图构建、划分融合以及谱旋转。其对应的时间复杂度分别为 $O(mn^3 + (m+1)nk^3)$ ， $O(k^3)$ 以及 $O(nk^2)$ 。因此，算法 2 的整体时间复杂度为 $O(T(mn^3 + mnk^3 + k^3))$ ，其中 T 为迭代次数。

算法 2 OP-MVSC

输入： 具有 m 个视图的多视图数据 $\{A^v\}_{v=1}^m$ ，聚类个数 k ，超参数 λ 、 γ 。

输出： 聚类标签矩阵 Y 。

初始化： 初始化 F^v 、随机正交初始化 F^* 、初始化 P 为 $k \times k$ 维的单位矩阵、初始化 Y 为每行只有一个位置为 1 其余为 0 的 $n \times k$ 维的矩阵

- 1: 重复执行
- 2: 通过式(9)来优化 S^v 。
- 3: 通过解式(10)来优化 F^* 。
- 4: 通过解式(12)来优化 F^v 。
- 5: 通过解式(13)来优化 P 。
- 6: 通过解式(14)来优化 Y 。
- 7: 直至算法收敛。
- 8: 返回聚类标签矩阵 Y 。

4 实验及结果分析

4.1 数据集介绍

我们在四个广泛使用的公开数据集上进行了实验。如表 1 所示，**Wikipedia Articles** 是一个广泛使用的跨模检索数据集，包含属于 10 个类的

693 个样本。将从图像中提取 128 维的 SIFT 特征和从文本的隐含狄利克雷分布模型中提取的 10 维特征组成两个视图组成多视图数据。**MSRC-v1** 是一个场景识别数据集，其中有 8 个类，每个类有 30 个样本。我们从中挑选了 7 个类（树、汽车、人脸、奶牛、自行车、建筑和飞机）共 210 张图像。从每张图像中提取多种特征组成多视图数据集 **Caltech7** 和 **Caltech20** 来源于 101 类的图像数据集，我们选择其中广泛使用的 7 个类和 20 个类，提取不同特征构成多视图数据集。

Table 1 Introduction of Datasets

表 1 数据介绍

数据集	样本个数	视图个数	簇个数
Wikipedia Articles	693	2	10
MSRC-v1	210	6	7
Caltech7	1474	6	7
Caltech20	2386	6	20

4.2 对比方法介绍

我们将提出的 OP-MVSC 与下面的方法进行对比。其中直接连接特征方法（Direct）作为一种基准方法，直接将多个视图的特征拼接在一起然后执行 K-means 得到最终的结果。

Co-reg(Co-regularized)^[18]方法使用协同正则化方法使得不同视图的划分之间达成一致，本文仅对比其中的“成对”方法。**LMSC**(Latent multi-view subspace clustering)^[13]不是在原始特征上而是在共同的潜在表示上进行数据的重建，从而得到共同的潜在子空间表示，进而执行谱聚类得到最终结果。**RMKMC**(Multi-View K-Means Clustering on Big Data)^[19]通过获得共同聚类指示矩阵并引入 $\ell_{2,1}$ 范数，对输入数据的异常值具有鲁棒性。**mPAC**(multiple partition aligned clustering)^[20]方法为每个划分矩阵分配不同的旋转矩阵得到统一的聚类指示矩阵。**FMR**(Flexible multi-view representation learning for subspace clustering)^[12]方法通过希尔伯特-史密斯独立性准则（Hilbert-Schmidt Independence Criterion, HSIC）构造包含互补和一致信息的潜在表示，然后对潜在表示进行数据重建得到子空间表示。上述的方

法，我们均按照论文中推荐的参数范围进行网格搜索并选取最好的结果。

本文使用准确率（Accuracy, ACC）和 F-score 两个指标来衡量聚类效果。其定义如下：定义 TP(True Positive)、TN(True Negative)、FP(False Positive)、FN(False Negative)分别为判断正样本为正、判断负样本为负、判断正样本为负以及判断负样本为正。准确率定义为模型判断正确的样本占总样本的比例，表示如下：

$$ACC = \frac{TP+TN}{TP+TN+FP+FN}$$

F-score 是精确率（Precision）和召回率（Recall）的加权调和平均值，定义如下：

$$F\text{-score} = (1 + \beta^2) \cdot \frac{\text{Precision} \cdot \text{Recall}}{\beta^2 \cdot \text{Precision} + \text{Recall}}$$

本文取 β 为 1，即同等地看待精确率和召回率的重要性。其中精确率和召回率的定义分别为 $\text{Precision} = TP / (TP + FP)$ ， $\text{Recall} = TP / (TP + FN)$ 。

4.3 模型分析

本文中有两个超参数 $\{\lambda, \gamma\}$ ， λ 从 $\{10, 20, 30\}$ 中选取， γ 从 1.2 至 1.8 的范围内以步长为 0.2 进行选择。以 Caltech7 为例，图 1 显示了不同的超参数对聚类性能的影响。可以观察到在 $\lambda \in \{10, 20\}$ 且 $\gamma \in \{1.6, 1.8\}$ 的范围内聚类效果相当稳定，同样的结果在其他数据集上也有体现。

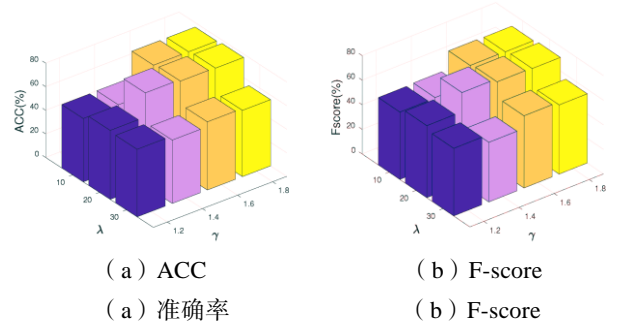


Fig.1 The sensitivity of parameters on Caltech7

图 1 数据集 Caltech7 上的参数敏感性

图 2 为在 MSRC-v1 上随着迭代次数的增加聚类性能变化的曲线，可以看出，算法在有限的迭代次数内可以达到最优的性能并保持稳定。

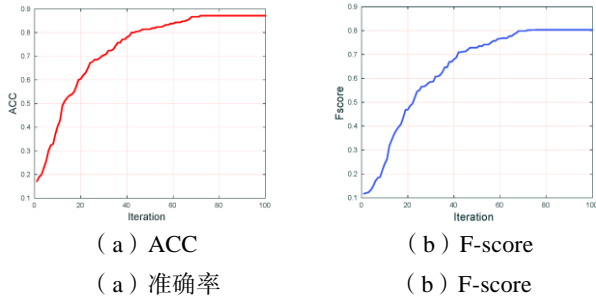


Fig.2 Clustering performance during iterations

图 2 迭代过程中的聚类性能

这也证明了从一致的表示中得到的聚类标签对后续优化中表示的学习起到的正向引导作用，进而提高了聚类性能。在不断的迭代中互相促进，从而达到更好的聚类性能。

4.4 实验结果分析

对比算法以及我们的方法在四种真实数据集的结果如表 2 和表 3 所示。其中最好的结果用红色标注，次好结果用蓝色标注。

如表 2 所示，在大多数情况下，直接连接特征的方法表现出最差的性能，这表明了多视图聚类算法在探索不同视图之间互补信息的有效性。我们提出的方法在不同数据集下的 ACC 均为最高，分别超过次好的方法 2.16%、1.9%、7.26% 以及 0.54%。从表 3 中可以看到，在 F-score 指标下，我们的方法在 Wikipedia Article、MSRC-v1 和 Caltech7 上都取得了最佳的性能，分别超过次优 2%、6.48%、6.92%；仅在 Caltech20 上略低于 mPAC 方法。

Table 2 The comparison of different algorithms under four datasets in terms of ACC

表 2 对比算法在不同数据集下的 ACC

方法	Wikipedia	MSRC-v1	Caltech7	Caltech20
Direct	0.5592	0.4541	0.5010	0.3165
Co-reg ^[18]	0.5623	0.6933	0.4323	0.3801
LMSC ^[13]	0.5931	0.7476	0.5217	0.5088
RMKMC ^[19]	0.5657	0.7095	0.3725	0.4116
mPAC ^[20]	0.5426	0.8143	0.7164	0.5478
FMR ^[12]	0.5758	0.8524	0.4423	0.3864
Ours	0.6147	0.8714	0.7890	0.5532

Table 3 The comparison of different algorithms under four datasets in terms of F-score

表 3 对比算法在不同数据集下的 F-score

方法	Wikipedia	MSRC-v1	Caltech7	Caltech20
Direct	0.4753	0.3856	0.4939	0.2782
Co-reg ^[18]	0.4593	0.5866	0.4634	0.3386
LMSC ^[13]	0.5014	0.6384	0.5382	0.4347
RMKMC ^[19]	0.4794	0.5894	0.4447	0.3737
mPAC ^[20]	0.4344	0.7320	0.6725	0.5179
FMR ^[12]	0.4838	0.7381	0.4849	0.3399
Ours	0.5214	0.8029	0.7417	0.4996

相比于其他先进的方法，上述实验结果很好的说明了我们提出的 OP-MVSC 方法的有效性。我们将其归因于以下两点：1) OP-MVSC 联合优化子空间表示、划分矩阵以及聚类标签。当在某一轮迭代中产生更加准确的聚类标签时，我们在下一次迭代中充分利用这种高质量的标签来引导视图表示和划分的学习。通过这样在算法迭代中彼此促进，进而可以改进聚类性能。2) 本文在划分级对多视图信息进行融合，相对于相似度融合，这种划分级的融合具有更少的噪声和冗余信息，从而得到更好的聚类性能。

5 结束语

本文提出了一种新颖的单步划分融合多视图子空间聚类算法。与现有的基于相似度融合多视图信息的方法不同，本文从更富含聚类结构信息的划分级别进行融合，并引入旋转矩阵直接得到聚类结果，避免了两阶段算法分离表示学习和聚类过程的缺陷。直接将表示学习、划分融合和聚类过程统一在一个框架中，使得表示和聚类结果在迭代过程中互相促进，从而达到理想的聚类效果。多个多视图数据集上的实验结果验证了该方法的有效性和优越性。在未来的工作中，我们将考虑大规模多视图聚类问题以及自适应融合的方法。

References :

- [1] YANG Y, WANG H. Multi-view clustering: A survey[J]. Big Data Mining and Analytics, 2018, 1(2): 83-107.
- [2] HE X M. A Survey of Multi-view Clustering Algorithms [J]. Software Guide, 2019,18(04) 79-81,86.
何雪梅.多视图聚类算法综述[J].软件导刊,2019,18(04) 79-81,86.
- [3] GONEN M, ALPAYDINI E. Multiple kernel learning algorithms[J]. Journal of machine learning research, 2011, 12(64): 2211-2268.
- [4] WANG H, YANG Y, LIU B. GMC: Graph-based multi-view clustering[J]. IEEE Transactions on Knowledge and Data Engineering, 2019.
- [5] GAO H, NIE F, LI X, et al. Multi-view subspace clustering[C]. Proceedings of the IEEE international conference on computer vision. 2015: 4238-4246.
- [6] HE M J. Research on the multi-view clustering algorithms based on NMF[D].Chengdu: Southwest Jiaotong University, 2017
何梦娇. 基于非负矩阵分解的多视图聚类研究[D]. 成都: 西南交通大学, 2017
- [7] HUANH S D. Research and Application of Matrix Factorization Algorithms for Multi-source Heterogeneous Data[D], Chengdu: University of Electronic Science and Technology of China, 2019.
黄树东. 面向多源异构数据的矩阵分解算法研究及应用 [D].电子科技大学,2019.
- [8] FAN R D, HOU C P, Robust Auto-Weighted Multi-View Subspace Clustering[J]. Journal of Frontiers of Computer Science and Technology.
范瑞东, 侯臣平. 鲁棒自加权的多视图子空间聚类[J]. 计算机科学与探索.
- [9] LIU J H, WANG Y, HE X L, Improved multi-view subspace clustering with diversify driven[J]. Computer Era, 2020(09): 91-94.
刘金花,王洋,贺潇磊.改进的多样性驱动的多视图子空间聚类算法[J].计算机时代,2020(09):91-94.
- [10] GUO Y, Convex subspace representation learning from multi-view data[C]. Proceedings of the 27th AAAI Conference on Artificial Intelligence, AAAI 2013, pp. 387–393, 2013.
- [11] BRBIC M and IVICA. K, Multi-view low-rank sparse subspace clustering[J]. Pattern Recognition, vol. 73, pp. 247–258, 2018.
- [12] LI R H, ZHANG C, HU Q, et al. Flexible multi-view representation learning for subspace clustering[C]. IJCAI, vol. 2019-Augus, pp. 2916–2922, 2019.
- [13] 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. 2017:4279-4287.
- [14] ELHAMIFAR, E., & Vidal, R. (2013). Sparse subspace clustering: Algorithm, theory, and applications[J]. IEEE transactions on pattern analysis and machine intelligence, 35(11), 2765-2781.
- [15] MOHAR B, ALAVI Y, CHARTRAND G, et al. The Laplacian spectrum of graphs[J]. Graph Theory, Combinatorics and Applications, vol. 2, p. 5364, 1991.
- [16] FAN K. On a theorem of Weyl concerning eigenvalues of linear transformations I[J]. Proceedings of the National Academy of Sciences of the United States of America, 1949, 35(11): 652.
- [17] NIE F, WANG X, and HUANG H, Clustering and projected clustering with adaptive neighbors[C]. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, no. December, pp. 977–986, 2014.
- [18] KUMAR A, RAI P, DAUME H. Co-regularized multi-view spectral clustering[C]. Advances in neural information processing systems. 2011: 1413-1421.
- [19] CAI X, NIE F, HUANG H. Multi-View K-Means Clustering on Big Data[C]. IJCAI, vol. 31, no. 5, pp. 724-730, 2013.
- [20] KANG Z, GUO Z, HUANG S, et al. Multiple Partitions Aligned Clustering[C]. IJCAI, pp. 2701-2707, 2019.



ZHANG Pei was born in 1996. She received the B.S. degree from Yunan university in 2018. Now, she is a M.S. candidate at National University of Defense Technology. Her current research interests include multi-view learning, multi-kernel clustering, deep clustering, etc.

张培（1996-），女，山东省菏泽市人，2018 年于云南大学获得学士学位，目前为国防科技大学计算学院硕士研究生，主要研究领域为多视图学习、多核聚类以及深度聚类等。



ZHU En was born in 1976. He is a professor and Ph.D. supervisor at National University of Defense Technology, the senior member of CCF. He has published over 160 academic papers on his research interests include clustering, anomaly detection, computer vision, medical image analysis, etc.

祝恩（1976-），男，湖南益阳人，国防科学技术大学计算机学院教授、博士生导师，CCF 高级会员，理论计算机科学技术专委会副主任。在包括聚类算法、异常检测、计算机视觉、医学图像分析等主要研究领域发表论文 160 余篇。



CAI Zhiping was born in 1975. He is a professor and Ph.D. supervisor at National University of Defense Technology, the senior member of CCF. His research interests include network security, big data and dispersed computing. He has published over 150 papers with more than 2000 citations.

蔡志平（1975-），男，湖南益阳人，国防科学技术大学计算机学院教授、博士生导师，CCF 高级会员，理论计算机科学技术专委会秘书长。主要研究领域为网络安全，大数据，分散计算，在所在研究领域发表了超过 150 篇学术论文。