计算机科学与探索
Journal of Frontiers of Computer Science and Technology
1673-9418/2023/17(05)-1147-10
doi: 10.3778/j.issn.1673-9418.2109025

动态融合的多视图投影聚类算法

姜凯彬,周世兵+,钱雪忠,管娇娇
江南大学 人工智能与计算机学院,江苏 无锡 214122
+ 通信作者 E-mail: zshibing@jiangnan.edu.cn

摘　要:多视图聚类是一个日益受到关注的研究热点。现有的大多数多视图聚类方法通常先对数据进行图学习,再对融合得到的统一图进行聚类得到最终结果,这种图学习和图聚类的两步策略可能导致聚类结果具有随机性。此外,多视图数据本身存在不可避免的噪声并且各视图数据差异较大,在原始高维数据空间进行无效融合可能造成重要信息的损失,不同多视图数据也可能存在选择参数敏感的问题。为了解决上述问题,提出了一种动态融合的多视图投影聚类算法,将自适应降维图学习、无参数的自权重图融合和谱聚类整合在同一框架中,三个过程相互促进,联合优化投影矩阵、相似性矩阵、共识矩阵以及聚类标签。对动态融合过程中得到的共识矩阵的拉普拉斯矩阵施加秩约束,直接获得聚类结果。而且引入的启发式超参数会随着每次优化迭代自动调整。为了求解联合优化问题,设计了一种有效的交替迭代方法。在人工数据集和真实数据集上得到的实验结果表明该算法的优越性。

关键词:多视图聚类;投影降维;图融合;共识矩阵
文献标志码:A　　**中图分类号:**TP18

Dynamic-Fusion Multi-view Projection Clustering Algorithm

JIANG Kaibin, ZHOU Shibing+, QIAN Xuezhong, GUAN Jiaojiao
School of Artificial Intelligence and Computer Science, Jiangnan University, Wuxi, Jiangsu 214122, China

Abstract: Multi-view clustering is a hot research area, which has attracted increasing attention. Most existing multi-view clustering methods usually learn the data first, and then cluster the fused unified graph to get the final result. This two-step strategy of graph learning and graph clustering may lead to the randomness of clustering results. Besides, the inevitable noise of the data itself and the large differences among views, these invalid fusion methods in high-dimensional data space may cause important information loss, and different multi-view data may be sensitive to parameter selections. To solve the above problems, a multi-view projection clustering algorithm based on dynamic fusion is proposed, which integrates adaptive dimensionality reduction graph learning, self-weight fusion without parameters and spectral clustering in the same framework. The three processes promote each other and jointly optimize the projection matrix, similarity matrix, consensus matrix and clustering label. The Laplacian matrix of the best consensus matrix obtained by dynamic fusion is constrained by rank, and clustering results are obtained directly. Moreover, heuristic super-parameters are automatically adjusted with each optimization iteration. To solve the joint optimization problem, an effective alternative optimization method is designed. Experimental results on artificial datasets and real datasets show the superiority of the algorithm.

Key words: multi-view clustering; projection dimension reduction; graph fusion; consensus matrix

基金项目:国家自然科学基金(62076110);江苏省自然科学基金(BK20181341)。
This work was supported by the National Natural Science Foundation of China (62076110), and the Natural Science Foundation of Jiangsu Province (BK20181341).
收稿日期:2021-09-08　　**修回日期:**2021-10-25

聚类是机器学习中的一个重要课题，旨在发现数据的底层结构。在许多现实世界的应用程序中，数据通常是从不同的来源生成的。例如，信号可以从多个传感器获得，图像可以由多个特征描述符描述。所有这些都称为多视图数据[1]。因为不同视图可以捕捉不同数据视角，存在不同的统计特性，传统单视图的聚类方法无法充分反映多视图聚类结构的本质。因此通过集成异构和互补的信息的无监督多视图聚类方法开始变得流行。

简单地将所有特征直接连接成单个视图，然后对单个视图数据进行聚类，可能不会获得比单独使用单个视图的传统方法更好的性能。在过去十年中，许多考虑不同视图多样性和互补性的先进多视图聚类算法被提出，大致可分为多视图协同训练算法[2]、多视图多核聚类算法[3]、基于图的多视图聚类算法[4]、基于子空间的多视图聚类算法[5]。

在这些方法中，采用多视图谱聚类可以在任意形状的数据集上聚类并收敛得到全局最优解。因此，多视图谱聚类能更好地探索多视图数据的非线性结构，在实践中常优于其他多视图聚类方法。文献[2]开发了一种使用线性核来最小化不同谱嵌入之间的不一致并对多个分区进行正则化的方法。但这种方法不能区分不同视图的可靠性，容易被噪声较多的视图所干扰。为了区分不同视图对于聚类结果的影响，文献[6]提出了自适应加权的多视图谱聚类方法。文献[7]提出一种从多个视图中学习一个具有稀疏结构的一致相似矩阵，然后进行单独聚类的多视图谱聚类算法。为了避免额外的聚类步骤带来的不确定性，文献[8]提出一种基于非负矩阵分解的松弛全局相似性矩阵约束的多视图聚类算法。上述聚类方法成功解决了低维数据中的聚类问题。为了处理高维数据，文献[9]提出了两种无参数加权投影聚类方法，可同时进行结构图学习和数据降维。

虽然这些算法在不同的场景下取得了较好结果，但仍存在以下问题：(1)以往许多工作都集中在原始数据上，而忽略了高维数据引入的噪声和冗余信息。当高维数据直接用于聚类任务时，可能会导致重要信息的丢失以及算法性能的下降。(2)现有方法大都预先构造一个共识图，然后利用该固定的图执行聚类任务，这种两步的分离策略可能会造成聚类结果的次优解。(3)通过引入额外的超参数来解决不同视图在模型中产生的影响，可能会导致模型复杂度的提高。

为了解决上述问题，本文提出了一种动态融合的多视图投影聚类算法(dynamic-fusion multi-view projection clustering algorithm，DFMPC)，将自适应降维图学习、无参数的自权重图融合和谱聚类整合在同一框架中，联合优化投影矩阵、相似性矩阵、共识矩阵以及聚类标签。具体来说，首先将高维数据投影到低维子空间，建立不同视图的相似图。基于不同视图具有相同的底层聚类结构这一假设，在动态融合过程中，将共识矩阵与所有相似度矩阵对齐一致，自动学习各视图的权重并对得到的共识矩阵的拉普拉斯矩阵施加秩约束，直接获得聚类结果。与以往引入需要人工调整的超参数不同的是，本文引入的启发式超参数会随着每次优化迭代自动调整。此外，本文设计了一种有效的交替迭代方法来求解联合优化问题。在人工数据集和真实数据集上得到的实验结果表明该算法的优越性。

1 相关工作

本章介绍了多视图聚类方法的基本符号定义，并回顾了传统多视图聚类算法的基本形式。

1.1 符号与定义

定义一个具有 m 个视图、n 个样本的多视图数据集为 $\boldsymbol{X}=[\boldsymbol{X}^1,\boldsymbol{X}^2,\cdots,\boldsymbol{X}^m]\in\mathbb{R}^{d_v\times n}$，其中 d_v 代表第 v 个视图的维度。对于一个矩阵 $\boldsymbol{X}^v$ 来说，x_{ij}^v 和 $\boldsymbol{x}_j^v$ 分别表示矩阵第 i 行第 j 列元素以及第 j 列向量。矩阵 $\boldsymbol{X}$ 的Frobenius范数、迹和转置分别表示为 $\|\boldsymbol{X}\|_{\mathrm{F}}$、$\mathrm{tr}(\boldsymbol{X})$ 和 $\boldsymbol{X}^{\mathrm{T}}$。向量 $\boldsymbol{x}$ 的第2范数为 $\|\boldsymbol{x}\|_2$。此外，$\boldsymbol{I}$ 表示单位矩阵，$\boldsymbol{1}$ 表示元素全为1的列向量。

1.2 多视图聚类

给定一个具有 m 个视图、n 个样本的多视图数据集 $\{\boldsymbol{X}^v\}_{v=1}^m=\{\boldsymbol{x}_1^v,\boldsymbol{x}_2^v,\cdots,\boldsymbol{x}_n^v\}_{v=1}^m$。聚类算法寻求最优划分，即同一组数据点相似性较大，而不同组数据点相似性较小。在多视图聚类中，将数据矩阵 $\boldsymbol{X}^v$ 转化为相似矩阵 $\boldsymbol{S}^v$ 最常用的是 k 最近邻图。不同数据点间的相似度通常用高斯核函数 $\boldsymbol{S}_{ij}^v=\exp\left(-\dfrac{\|\boldsymbol{x}_i^v-\boldsymbol{x}_j^v\|_2^2}{\sigma^2}\right)$ 表示，其中 σ 是控制邻域大小的超参数。若 $\boldsymbol{x}_i^v$ 属于 $\boldsymbol{x}_j^v$ 的 k 近邻，则这两个数据点相连接，其相似度由高斯核函数获得；反之，相似度为0。文献[10]中的研究发现稀疏表示对噪声和异常值具有鲁棒性，使用一种稀疏表示方法来构造相似度矩阵。具体表示如下：

$$\begin{cases}\min\limits_{S^v}\sum\limits_{v=1}^{m}\left(\sum\limits_{i,j=1}^{n}\|\boldsymbol{x}_i^v-\boldsymbol{x}_j^v\|s_{ij}^v+\beta\sum\limits_{i=1}^{n}\|\boldsymbol{s}_i^v\|_2^2\right)\\ \text{s.t.}\forall v,s_{ii}^v=0,s_{ij}^v\geqslant 0,\mathbf{1}^{\mathrm{T}}\boldsymbol{s}_i^v=1\end{cases}\tag{1}$$

其中，$\{\boldsymbol{S}^v\}_{v=1}^m=\{\boldsymbol{s}_1^v,\boldsymbol{s}_2^v,\cdots,\boldsymbol{s}_n^v\}_{v=1}^m$ 是每个原始视图的相似度矩阵表示，$\beta\sum\limits_{v=1}^{m}\sum\limits_{i=1}^{n}\|\boldsymbol{s}_i^v\|_2^2$ 是正则化项，避免得到平凡解，其中 β 是调整参数。归一化 $\mathbf{1}^{\mathrm{T}}\boldsymbol{s}_i^v=1$ 相当于相似度矩阵 $\boldsymbol{S}$ 上的稀疏约束，保证了正则化项的恒定。大部分多视图聚类方法通过对得到的相似度矩阵进行处理，得到一致表示的共识矩阵 $\boldsymbol{A}$，然后通过谱聚类得到谱嵌入矩阵。具体如式(2)所示：

$$\min_{\boldsymbol{F}}\operatorname{tr}(\boldsymbol{F}^{\mathrm{T}}\boldsymbol{L}\boldsymbol{F}),\ \text{s.t.}\boldsymbol{F}^{\mathrm{T}}\boldsymbol{F}=\boldsymbol{I}\tag{2}$$

其中，共识矩阵 $\boldsymbol{A}$ 的拉普拉斯矩阵 $\boldsymbol{L}$ 定义为 $\boldsymbol{L}=\boldsymbol{D}-(\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{A})/2$，度矩阵 $\boldsymbol{D}$ 是矩阵 $\boldsymbol{A}$ 的对角矩阵，并且它的第 i 个对角元素表示为 $d_{ii}=\sum\limits_j(a_{ij}+a_{ji})/2$。$\boldsymbol{F}\in\mathbb{R}^{n\times c}$ 是原始数据的谱嵌入矩阵，c 是聚类中心数目。通过计算矩阵 $\boldsymbol{A}$ 的 c 个最大特征值对应的 c 个特征向量得到谱嵌入矩阵 $\boldsymbol{F}$ 的解，即离散标签的松弛解。最终的聚类标签结果可通过对谱嵌入进行 K 均值[11-12]或谱旋转得到。

2 算法模型

本章在上文介绍的基础上，提出了本文的算法模型，并给出了模型的优化求解算法。

2.1 动态融合的多视图投影聚类算法

鉴于子空间学习在高维数据处理中的优越性，本文在前文介绍的式(1)的基础上，将原始数据 $\boldsymbol{X}^v$ 投影到低维子空间，从具有尽可能少的噪声和冗余的低维子空间中学习相似度矩阵。其对应的低维子空间的自适应图学习优化目标如式(3)所示：

$$\begin{cases}\min\limits_{S^v,\boldsymbol{W}^v}\sum\limits_{v=1}^{m}\left(\sum\limits_{i,j=1}^{n}\|(\boldsymbol{W}^v)^{\mathrm{T}}\boldsymbol{x}_i^v-(\boldsymbol{W}^v)^{\mathrm{T}}\boldsymbol{x}_j^v\|_2^2s_{ij}^v+\beta\sum\limits_{i=1}^{n}\|\boldsymbol{s}_i^v\|_2^2\right)\\ \text{s.t.}\forall v,s_{ii}^v=0,s_{ij}^v\geqslant 0,\mathbf{1}^{\mathrm{T}}\boldsymbol{s}_i^v=1,(\boldsymbol{W}^v)^{\mathrm{T}}\boldsymbol{X}^v(\boldsymbol{X}^v)^{\mathrm{T}}\boldsymbol{W}^v=\boldsymbol{I}\end{cases}\tag{3}$$

其中，$\boldsymbol{W}^v\in\mathbb{R}^{d_v\times d_v'}$ 是投影变换矩阵，d_v' 表示低维子空间的特征维数，$\{\boldsymbol{S}^v\}_{v=1}^m$ 是每个原始视图的相似度矩阵表示，β 是启发式超参数，$\beta\sum\limits_{v=1}^{m}\sum\limits_{i=1}^{n}\|\boldsymbol{s}_i^v\|_2^2$ 是正则化项。$\mathbf{1}^{\mathrm{T}}\boldsymbol{s}_i^v=1$ 为归一化项。$(\boldsymbol{W}^v)^{\mathrm{T}}\boldsymbol{X}^v(\boldsymbol{X}^v)^{\mathrm{T}}\boldsymbol{W}^v$ 将正交约束应用于散射矩阵进行低维子空间学习，保留了数据的有效信息并缓解了维数灾难问题。

在多视图学习中，考虑到每个视图的相似度图都是共识图的扰动，以及为了避免低质量视图的影响，让共识图更好地捕捉隐藏在多视图数据中的真实样本相似性，本文通过衡量不同视图的重要性来输出最终的共识表示，并期望低维子空间中的共识矩阵 $\boldsymbol{A}$ 可以将所有相似度矩阵 $\boldsymbol{S}$ 一致对齐。在此基础上，自权重融合机制如式(4)所示：

$$\begin{cases}\min\limits_{\boldsymbol{A}}\sum\limits_{v=1}^{m}\alpha_v\|\boldsymbol{A}-\boldsymbol{S}^v\|_{\mathrm{F}}^2\\ \text{s.t.}\forall i,a_{ij}\geqslant 0,\mathbf{1}^{\mathrm{T}}\boldsymbol{a}_i=1\end{cases}\tag{4}$$

其中，权重 α_v 代表不同视图的重要性，可以采用反向距离加权方案[6]得到权重 α_v，具体公式如下：

$$\alpha_v=\frac{1}{2\sqrt{\|\boldsymbol{A}-\boldsymbol{S}^v\|_{\mathrm{F}}^2}}\tag{5}$$

此时动态融合得到的共识矩阵 $\boldsymbol{A}$ 并不能直接获得最终聚类结果，还需要额外的聚类步骤。由文献[13]可知，$\boldsymbol{A}$ 的拉普拉斯矩阵 $\boldsymbol{L}_{\boldsymbol{A}}$ 的特征值0的重数 r 等于 $\boldsymbol{A}$ 的图中连通分量数。为了使动态融合得到共识矩阵的数据点精确聚集成 c 个簇，即 $rank(\boldsymbol{L}_{\boldsymbol{A}})=n-c$，希望对共识矩阵 $\boldsymbol{A}$ 的图拉普拉斯矩阵施加秩约束，使 $\boldsymbol{A}$ 达到理想的效果。但由于 $\boldsymbol{L}_{\boldsymbol{A}}$ 依赖于目标矩阵 $\boldsymbol{A}$ 并且 $rank(\boldsymbol{L}_{\boldsymbol{A}})=n-c$ 也是非线性的，直接引入秩约束会使得优化问题难以求解。根据Ky Fan's定理[14]，可以得到式(6)：

$$\sum_{i=1}^{k}\sigma_i(\boldsymbol{L}_{\boldsymbol{A}})=\min_{\boldsymbol{F},\boldsymbol{F}^{\mathrm{T}}\boldsymbol{F}=\boldsymbol{I}}\operatorname{tr}(\boldsymbol{F}^{\mathrm{T}}\boldsymbol{L}_{\boldsymbol{A}}\boldsymbol{F})\tag{6}$$

其中，$\sigma_i(\boldsymbol{L}_{\boldsymbol{A}})$ 表示 $\boldsymbol{L}_{\boldsymbol{A}}$ 第 i 小的特征值。很明显，$\boldsymbol{L}_{\boldsymbol{A}}$ 前 i 小的特征值均为0，即 $\sum\limits_{i=1}^{k}\sigma_i(\boldsymbol{L}_{\boldsymbol{A}})=0$，使得 $\boldsymbol{L}_{\boldsymbol{A}}$ 的秩为 $n-c$，共识矩阵 $\boldsymbol{A}$ 的连通分量数等于簇个数，直接得到聚类结果。

最后，整合式(3)、式(4)、式(6)，得到动态融合的多视图投影聚类算法模型的目标函数：

$$\begin{cases}\min\limits_{S^v,\boldsymbol{W}^v,\boldsymbol{A},\boldsymbol{F}}\sum\limits_{v=1}^{m}\sum\limits_{i,j=1}^{n}\|(\boldsymbol{W}^v)^{\mathrm{T}}\boldsymbol{x}_i^v-(\boldsymbol{W}^v)^{\mathrm{T}}\boldsymbol{x}_j^v\|_2^2s_{ij}^v+\beta\sum\limits_{v=1}^{m}\sum\limits_{i=1}^{n}\|\boldsymbol{s}_i^v\|_2^2+\\ \qquad\sum\limits_{v=1}^{m}\alpha_v\|\boldsymbol{A}-\boldsymbol{S}^v\|_{\mathrm{F}}^2+2\lambda\operatorname{tr}(\boldsymbol{F}^{\mathrm{T}}\boldsymbol{L}_{\boldsymbol{A}}\boldsymbol{F})\\ \text{s.t.}\forall v,s_{ii}^v=0,s_{ij}^v\geqslant 0,\mathbf{1}^{\mathrm{T}}\boldsymbol{s}_i^v=1,\\ \qquad a_{ij}\geqslant 0,\mathbf{1}^{\mathrm{T}}\boldsymbol{a}_i=1,\boldsymbol{F}^{\mathrm{T}}\boldsymbol{F}=\boldsymbol{I}\\ \qquad(\boldsymbol{W}^v)^{\mathrm{T}}\boldsymbol{X}^v(\boldsymbol{X}^v)^{\mathrm{T}}\boldsymbol{W}^v=\boldsymbol{I}\end{cases}\tag{7}$$

通过这种方式，可以同时研究投影矩阵、相似矩阵、共识矩阵、权系数和谱嵌入矩阵，在统一的优化框架下联合学习自适应图、自权重图融合和聚类标签。具体来说，$(\boldsymbol{W}^v)^{\mathrm{T}}\boldsymbol{X}^v(\boldsymbol{X}^v)^{\mathrm{T}}\boldsymbol{W}^v$ 将正交约束应用于散

射矩阵进行低维子空间学习，建立不同视图的相似图 $\boldsymbol{S}$，缓解了维数灾难，保留了数据有效信息。融合过程中共识矩阵 $\boldsymbol{A}$ 自动学习各 $\boldsymbol{S}$ 视图的权重，学习到的 $\boldsymbol{A}$ 返回更新各视图的相似矩阵 $\boldsymbol{S}$。共识矩阵的拉普拉斯矩阵 $\boldsymbol{L}_{\boldsymbol{A}}$ 上的秩约束也适用于约束共识矩阵中连通分量的个数（等于所需的簇数 c），直接获得聚类结果。

2.2 模型优化

2.2.1 固定 $\boldsymbol{W}^v$、$\boldsymbol{A}$、$\boldsymbol{F}$，优化 $\boldsymbol{S}^v$

关于初始化相似度矩阵，去掉其他无关项，每个视图初始化相似度矩阵优化式可以转化为：

$$\begin{cases}\min\limits_{\boldsymbol{s}_i^v}\sum\limits_{j=1}^{n}\|(\boldsymbol{W}^v)^{\mathrm{T}}\boldsymbol{x}_i^v-(\boldsymbol{W}^v)^{\mathrm{T}}\boldsymbol{x}_j^v\|_2^2 s_{ij}^v+\beta\|\boldsymbol{s}_i^v\|_2^2\\ \text{s.t.}\forall v, s_{ii}^v=0, s_{ij}^v\geqslant 0, \boldsymbol{1}^{\mathrm{T}}\boldsymbol{s}_i^v=1\end{cases}\tag{8}$$

简单起见，定义 $\boldsymbol{g}_i$ 是一个向量并且令其第 j 个元素 $g_{ij}=\|(\boldsymbol{W}^v)^{\mathrm{T}}\boldsymbol{x}_i^v-(\boldsymbol{W}^v)^{\mathrm{T}}\boldsymbol{x}_j^v\|_2^2$，则式（8）可以化简为：

$$\min_{\boldsymbol{s}_i^v}\frac{1}{2}\left\|\boldsymbol{s}_i^v+\frac{\boldsymbol{g}_i}{2\beta}\right\|_2^2\tag{9}$$

由式（9）得到目标函数的拉格朗日函数：

$$L(\boldsymbol{s}_i^v,\eta,\xi)=\frac{1}{2}\left\|\boldsymbol{s}_i^v+\frac{\boldsymbol{g}_i}{2\beta}\right\|_2^2-\eta(\boldsymbol{1}^{\mathrm{T}}\boldsymbol{s}_i^v-1)-\xi\boldsymbol{s}_i^v\tag{10}$$

其中，η 和 ξ 是拉格朗日乘子。根据 KKT（Karush-Kuhn-Tucker）条件[15]，求解得到式（10）中初始化相似度矩阵 $\boldsymbol{S}^v$ 的最优解为：

$$s_{ij}=\left(-\frac{g_{ij}}{2\beta}+\eta\right)_+\tag{11}$$

然而，关于迭代优化过程中的相似度矩阵，去掉其他无关项，优化式可以转化为：

$$\begin{cases}\min\limits_{\boldsymbol{s}_i^v}\sum\limits_{j=1}^{n}\|(\boldsymbol{W}^v)^{\mathrm{T}}\boldsymbol{x}_i^v-(\boldsymbol{W}^v)^{\mathrm{T}}\boldsymbol{x}_j^v\|_2^2 s_{ij}^v+\beta\|\boldsymbol{s}_i^v\|_2^2+\alpha_v\|\boldsymbol{u}_i-\boldsymbol{s}_i^v\|_2^2\\ \text{s.t.}\forall v, s_{ii}^v=0, s_{ij}^v\geqslant 0, \boldsymbol{1}^{\mathrm{T}}\boldsymbol{s}_i^v=1\end{cases}\tag{12}$$

同理，按照求解式（8）的方法求解式（12），从而得到相似度矩阵 $\boldsymbol{S}^v$ 的最优解为：

$$s_{ij}=\left(\frac{-\frac{g_{ij}}{2}+\alpha_v a_{ij}+\beta\eta}{\alpha_v+\beta}\right)_+\tag{13}$$

2.2.2 固定 $\boldsymbol{S}^v$、$\boldsymbol{A}$、$\boldsymbol{F}$，优化 $\boldsymbol{W}^v$

去掉其他无关项，关于 $\boldsymbol{W}^v$ 优化式可以转化为：

$$\begin{cases}\min\limits_{\boldsymbol{W}^v}\sum\limits_{i,j=1}^{n}\|(\boldsymbol{W}^v)^{\mathrm{T}}\boldsymbol{x}_i^v-(\boldsymbol{W}^v)^{\mathrm{T}}\boldsymbol{x}_j^v\|\\ \text{s.t.}(\boldsymbol{W}^v)^{\mathrm{T}}\boldsymbol{X}^v(\boldsymbol{X}^v)^{\mathrm{T}}\boldsymbol{W}^v=\boldsymbol{I}\end{cases}\tag{14}$$

如果将函数值 $\boldsymbol{x}_i^v\boldsymbol{W}^v$ 视为一个节点 i 的值，式（14）可以进一步转化为：

$$\begin{cases}\min\limits_{\boldsymbol{W}^v}\operatorname{tr}((\boldsymbol{W}^v)^{\mathrm{T}}\boldsymbol{X}^v\boldsymbol{L}_*^v(\boldsymbol{X}^v)^{\mathrm{T}}\boldsymbol{W}^v)\\ \text{s.t.}(\boldsymbol{W}^v)^{\mathrm{T}}\boldsymbol{X}^v(\boldsymbol{X}^v)^{\mathrm{T}}\boldsymbol{W}^v=\boldsymbol{I}\end{cases}\tag{15}$$

其中，$\boldsymbol{L}_*^v=\boldsymbol{D}_*^v-\boldsymbol{S}^v$，度矩阵 $\boldsymbol{D}_*^v$ 是矩阵 $\boldsymbol{S}^v$ 的对角矩阵，$\boldsymbol{W}^v$ 的最优解由 $(\boldsymbol{X}^v(\boldsymbol{X}^v)^{\mathrm{T}})^{-1}\boldsymbol{X}^v\boldsymbol{L}_*^v(\boldsymbol{X}^v)^{\mathrm{T}}$ 的 c 个最小非零特征值对应的 d 个特征向量得到。

2.2.3 固定 $\boldsymbol{S}^v$、$\boldsymbol{W}^v$、$\boldsymbol{F}$，优化 $\boldsymbol{A}$

去掉其他无关项，关于 $\boldsymbol{A}$ 优化式可以转化为：

$$\begin{cases}\min\limits_{\boldsymbol{A}}\sum\limits_{v=1}^{m}\sum\limits_{i,j=1}^{n}\alpha_v\|\boldsymbol{A}-\boldsymbol{S}^v\|_{\mathrm{F}}^2+2\lambda\operatorname{tr}(\boldsymbol{F}^{\mathrm{T}}\boldsymbol{L}_{\boldsymbol{A}}\boldsymbol{F})\\ \text{s.t.}\forall i, a_{ij}\geqslant 0, \boldsymbol{1}^{\mathrm{T}}\boldsymbol{a}_i=1\end{cases}\tag{16}$$

为了加速计算，文献[16]中提出的有效迭代算法使 $\boldsymbol{A}$ 完全稀疏。定义 $\boldsymbol{v}_i$ 是一个向量并且令其第 j 个元素 $v_{ij}=\|\boldsymbol{f}_i-\boldsymbol{f}_j\|_2^2$，$\boldsymbol{A}$ 的最优解可化简为：

$$\boldsymbol{a}_i=\frac{\sum\limits_v\alpha_v\boldsymbol{s}_i^v-\frac{\lambda\boldsymbol{v}_i}{2}}{\sum\limits_v\alpha_v}\tag{17}$$

2.2.4 固定 $\boldsymbol{S}^v$、$\boldsymbol{W}^v$、$\boldsymbol{A}$，优化 $\boldsymbol{F}$

去掉其他无关项，关于 $\boldsymbol{F}$ 优化式可以转化为：

$$\min_{\boldsymbol{F}}\operatorname{tr}(\boldsymbol{F}^{\mathrm{T}}\boldsymbol{L}_{\boldsymbol{A}}\boldsymbol{F}),\ \text{s.t.}\boldsymbol{F}^{\mathrm{T}}\boldsymbol{F}=\boldsymbol{I}\tag{18}$$

$\boldsymbol{F}$ 的最优解通过对 $\boldsymbol{L}_{\boldsymbol{A}}$ 进行特征值分解，选择 c 个最小非零特征值对应的 c 个特征向量得到。

2.2.5 优化 α_v、β、λ

对于权重系数 α_v 的更新，在式（5）中已经给出，它依赖于 $\boldsymbol{A}$ 和 $\boldsymbol{S}^v$。而传统方法的参数 β、λ 值可能从 0 到无穷大，很难调整，因此本文以启发式的方法进行更新优化[17]，避免了选择参数的不确定性，能较好地保留数据结构信息且可以得到更好的性能。对于参数 β，在不失普遍性的情况下，假设式（13）中 $g_{i1},g_{i2},\cdots,g_{in}$ 从小到大排列，若限制 $\boldsymbol{s}_i$ 有 k 个非零项，可以得到 $s_{ik}>0$ 和 $s_{i,k+1}=0$。此外，还要对式（13）施加 $\boldsymbol{1}^{\mathrm{T}}\boldsymbol{s}_i^v=1$ 正则化约束。为了得到具有 k 个非零值的最优稀疏 $\boldsymbol{s}_i$，考虑数据的局部性，β 可设置为：

$$\beta=\frac{kg_{i,k+1}-\sum\limits_{h=1}^{k}g_{ih}}{2}-k\alpha_v a_{i,k+1}-\alpha_v\tag{19}$$

而 λ 的初始值设为 1，该值在每次迭代中自动调整。若在迭代过程中 $\boldsymbol{A}$ 的连通分量数小于簇的个数 c，则 λ 变为两倍；反之，则 λ 减少一半。

2.3 时间复杂度分析

整个算法流程如算法1所示。在整个优化过程中，所提出的DFMPC算法的时间复杂度主要由三部分组成：相似图构造、图融合和谱聚类。其对应的复杂度分别为 $O(mnd+4nd^2+d^3)$ 、$O(n^3+mn)$ 以及 $O(cn^2)$ 。其中 m 代表视图数，d 代表特征数，c 代表聚类数。由于 $n \gg m$ 并且 $n \gg c$ ，DFMPC算法的时间复杂度可以表示为 $O(4nd^2+d^3+n^3)$ 。

算法1 DFMPC

输入：具有 m 个视图的多视图数据集 $\{\boldsymbol{X}^v\}_{v=1}^m$ ，聚类个数 c，初始启发式超参数 λ 。

输出：共识矩阵 $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ 及聚类结果。

1. 初始化：通过式(11)来初始化 $\boldsymbol{S}^v$ ，随机初始化 $\boldsymbol{W}^v$ ，初始化 $\alpha_v = 1/m$ ，$\boldsymbol{A} = \alpha_v \boldsymbol{S}^v$ 以及 $\boldsymbol{F}$ 。

2. 重复执行

3. 通过式(13)来优化 $\boldsymbol{S}^v$

4. 通过式(15)来优化 $\boldsymbol{W}^v$

5. 通过式(5)来优化 α_v

6. 通过式(17)来优化 $\boldsymbol{A}$

7. 通过式(18)来优化 $\boldsymbol{F}$

8. 直到算法满足收敛条件

9. 返回具有精确 c 个连通分量的共识矩阵 $\boldsymbol{A}$ 及聚类结果

3 实验研究

本文在2个人工数据集、8个图像和文档真实数据集上进行了实验，同时与目前已有的一些先进算法进行比较，验证本文提出的DFMPC算法的优越性。

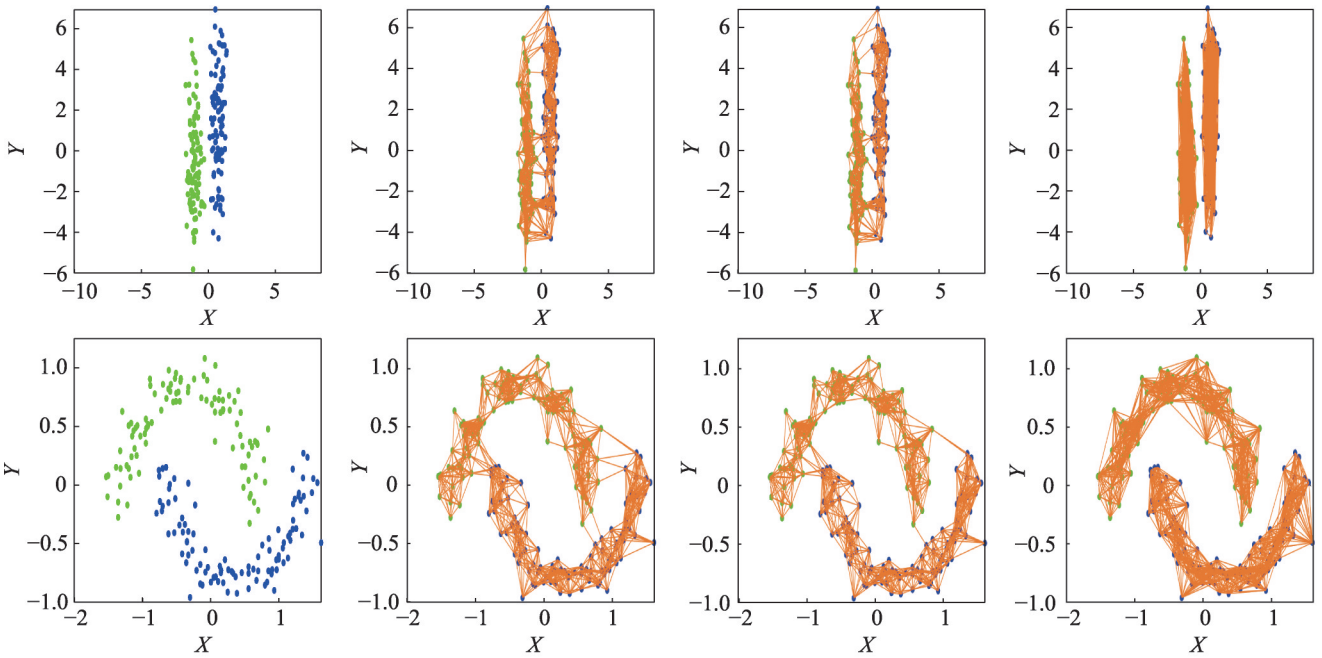
3.1 人工数据集实验

本节选取两组人工数据集RandomGaussian数据集和TwoMoon数据集进行实验，其中两组人工数据集介绍如下：

(1)RandomGaussian数据集：两个视图都由中心分别为(0, 1)和(−1, 0)的二维两高斯分布数据组成，每个类有100个样本点，每个类的协方差矩阵为 $\begin{bmatrix} 0.1 & 0 \\ 0 & 5.0 \end{bmatrix}$ 。

(2)TwoMoon数据集：两个视图分别通过添加0.12%的随机高斯噪声叠加弯月的形状生成，每个类有100个样本点。

由于页面空间限制，本文只展示DFMPC在每个人工数据集上单个视图的聚类过程。图1显示了RandomGaussian数据集和TwoMoon数据集的第一个视图的聚类过程。图1(b)展示出不同视图初始的相似矩阵构造图，可以看到两个类簇是有连接的。图1(c)展示出不同视图迭代学习后的相似矩阵构造图。可以看出一些有噪声的相似度边缘被删除，而



(a) 原始样本 (a) Original data
(b) 初始相似矩阵构造图 (b) Constructed graph with initialized similarity matrix
(c) 学习的相似矩阵构造图 (c) Constructed graph with learned similarity matrix
(d) 共识矩阵构造图 (d) Constructed graph with consensus matrix

图1 DFMPC在人工数据集上的聚类结果

Fig.1 Clustering results of DFMPC on artificial datasets

一些可信的相似度边缘被加强，这表明学习得到的共识矩阵可以有效改进相似度矩阵。但两个类簇此时并未完全分离。图1(d)展示出不同视图最终学习到的共识矩阵构造图。由于本文算法可以利用不同视图的互补信息，对不同视图中难以分离的点进行更为有效的分离从而得到完全分离的两个类簇。

总的来说，不同视图的相似矩阵和自加权得到的共识矩阵的学习可以相互加强，彼此促进，形成一种动态融合的增强效果，进而得到一个拥有最佳的底层聚类结构的共识矩阵。

3.2 真实数据集实验

为了全面评估DFMPC算法对不同的多视图数据集的兼容性和有效性，本节在表1所示的8个真实数据集上，与7种现有算法进行了实验比较。

8个多视图数据集分别是：来源于文献[6]的Caltech101-7、Caltech101-20，来源于Clément Grimal（http://lig-membres.imag.fr/grimal/data.html）的Newsgroups（NGs）以及来源于文献[8]的BBCSport、3sources、UCI、MSRCv1和ORL。

用于比较的7种多视图谱聚类算法分别是：SC_best、Co-reg（co-regularized）[2]、SwMC（self-weighted multiview clustering）[6]、SwMPC（self-weighted multi-view projected clustering）[9]、MCLES（multi-view clustering in latent embedding space）[8]、MSC_IAS（multi-view subspace clustering with intactness-aware similarity）[10]和GFSC（multi-graph fusion for multi-view spectral clustering）[18]。其中SC_best方法是将本文算法涉及到的谱聚类算法在多视图数据集的每个单视图上执行，然后选取最佳单个视图结果。

对比方法参数根据原始论文中的建议进行调整，以产生最佳结果。在实验中，算法精度为20次运行结果的平均值和标准差。采用了4种常用的评价指标，即准确度（accuracy，ACC）、归一化互信息（normalized mutual information，NMI）、调整兰德指数（adjusted Rand index，ARI）和纯度（purity，PUR）。对于这些度量，值越高表示聚类性能越好。表2至表5报告了8个真实数据集的详细聚类结果，粗体值代表最佳性能，其中→0→0表示均值和标准差非常接近于0。

在某些情况下，基于单视图基线的SC_best方法比一些多视图基线方法稍好，这表明探索多视图数据仍然需要良好的多视图聚类技术。相比之下，本文提出的DFMPC方法表现出更好的性能。这主要是因为DFMPC采用了动态图融合和自加权策略，学

表1 数据集介绍

Table 1 Introduction of datasets

视图	Caltech101-7	3sources	NGs	MSRCv1	BBCSport	ORL	Caltech101-20	UCI
1	Gabor(48)	Guardian(3 560)	M1(2 000)	CM(24)	Seg1(3 183)	GIST(512)	Gabor(48)	PC(216)
2	WM(40)	Reuters(3 631)	M2(2 000)	HOG(576)	Seg2(3 203)	LBP(59)	WM(40)	FC(76)
3	CENTR(254)	BBC(3 068)	M3(2 000)	GIST(512)		HOG(864)	CENTR(254)	K-LC(64)
4	HOG(1 984)			LBP(256)		CENTR(254)	HOG(1 984)	MORPH(6)
5	GIST(512)			CENTR(254)			GIST(512)	PA(240)
6	LBP(928)						LBP(928)	ZM(47)
样本数	1 474	169	500	210	544	400	2 386	2 000
类数	7	6	5	7	5	40	20	10

表2 不同算法在数据集上的ACC

Table 2 ACC of different algorithms on datasets 单位：%

Dataset	SC_best	Co-reg	SwMC	SwMPC	MSC_IAS	MCLES	GFSC	DFMPC
Caltech101-7	45.54±1.09	44.64±3.61	67.98±0.00	61.51±0.00	44.59±3.08	63.91±2.54	55.24±2.66	**69.95±0.00**
3sources	66.01±0.30	57.98±5.65	69.10±0.00	65.80±0.00	55.95±5.92	63.31±2.33	50.30±3.11	**77.51±0.00**
NGs	75.66±0.15	21.60±1.79	97.40±0.00	96.41±0.00	77.49±5.09	94.80±1.89	35.91±8.31	**98.40±0.00**
MSRCv1	67.45±1.24	72.19±4.63	74.67±0.00	74.07±0.00	68.45±1.00	70.00±2.53	70.07±5.24	**86.67±0.00**
BBCSport	57.54±0.00	38.97±2.02	62.50±0.00	77.68±0.00	61.06±6.06	80.13±4.13	55.07±2.91	**81.25±0.00**
ORL	65.75±1.99	74.57±0.17	81.75±0.00	78.70±0.00	83.59±2.17	78.25±1.51	60.61±4.06	**83.75±0.00**
Caltech101-20	24.89±0.46	39.31±2.97	60.81±0.00	52.51±0.00	41.70±2.63	58.24±3.52	48.23±4.12	**67.85±0.00**
UCI	75.78±1.96	20.17±3.22	81.85±0.00	83.56±0.00	75.52±3.17	81.63±2.77	83.75±6.34	**85.55±0.00**

表3 不同算法在数据集上的NMI

Table 3 NMI of different algorithms on datasets 单位：%

Dataset	SC_best	Co-reg	SwMC	SwMPC	MSC_IAS	MCLES	GFSC	DFMPC
Caltech101-7	17.11±0.31	45.73±1.77	57.78±0.00	53.38±0.00	22.63±2.25	54.66±4.27	48.28±2.95	**65.74±0.00**
3sources	56.10±0.61	51.59±1.43	62.98±0.00	65.21±0.00	43.62±3.96	58.90±8.12	40.06±2.52	**69.24±0.00**
NGs	57.93±0.21	36.13±1.57	91.47±0.00	90.60±0.00	55.06±5.11	84.19±1.62	20.96±9.48	**94.61±0.00**
MSRCv1	56.33±0.46	66.73±2.74	71.69±0.00	68.22±0.00	67.69±0.69	60.76±3.26	60.62±3.60	**77.36±0.00**
BBCSport	48.08±0.00	14.07±0.55	61.22±0.00	66.38±0.00	45.35±5.10	**76.62±5.32**	30.52±3.16	75.95±0.00
ORL	82.34±0.71	89.36±0.74	90.59±0.00	89.35±0.00	93.82±0.77	87.53±2.10	79.43±1.93	**93.88±0.00**
Caltech101-20	27.76±0.51	56.27±1.31	60.40±0.00	58.09±0.00	45.44±6.33	46.86±2.43	56.51±4.24	**64.45±0.00**
UCI	79.00±1.40	32.95±2.73	86.70±0.00	80.31±0.00	53.61±2.21	76.63±3.86	83.22±2.84	**91.03±0.00**

表4 不同算法在数据集上的ARI

Table 4 ARI of different algorithms on datasets 单位：%

Dataset	SC_best	Co-reg	SwMC	SwMPC	MSC_IAS	MCLES	GFSC	DFMPC
Caltech101-7	13.29±0.65	29.94±2.43	55.82±0.00	48.76±0.00	→0→0	48.62±3.81	35.12±3.30	**59.01±0.00**
3sources	53.21±0.56	34.79±3.12	**64.33±0.00**	51.06±0.00	33.67±6.25	36.64±4.28	15.83±4.73	57.12±0.00
NGs	54.92±0.22	→0→0	93.56±0.00	92.61±0.00	54.92±6.26	87.45±2.14	12.44±8.07	**96.04±0.00**
MSRCv1	46.53±0.52	54.78±7.18	63.89±0.00	59.79±0.00	54.10±0.81	52.45±3.53	50.46±5.05	**71.88±0.00**
BBCSport	39.99±0.00	12.56±1.03	52.09±0.00	63.64±0.00	38.58±6.33	66.27±3.51	21.64±5.88	**68.08±0.00**
ORL	53.08±1.98	71.39±3.51	74.96±0.00	71.56±0.00	**80.62±1.64**	69.49±2.52	46.95±4.69	76.32±0.00
Caltech101-20	9.18±0.38	29.16±3.07	**51.70±0.00**	42.68±0.00	20.76±5.52	38.54±2.78	28.11±6.69	41.10±0.00
UCI	68.04±2.19	26.35±2.73	79.60±0.00	78.66±0.00	50.44±2.83	74.81±3.68	77.45±5.72	**83.61±0.00**

表5 不同算法在数据集上的PUR

Table 5 PUR of different algorithms on datasets 单位：%

Dataset	SC_best	Co-reg	SwMC	SwMPC	MSC_IAS	MCLES	GFSC	DFMPC
Caltech101-7	63.59±0.31	50.27±1.73	88.43±0.00	69.03±0.00	71.18±3.12	87.72±1.34	65.43±2.55	**88.81±0.00**
3sources	73.73±0.30	66.86±5.24	74.21±0.00	73.28±0.00	61.48±5.79	71.01±2.85	67.22±5.33	**81.07±0.00**
NGs	75.66±0.15	96.00±0.08	97.40±0.00	97.33±0.00	78.64±4.39	94.80±1.21	88.20±6.42	**98.40±0.00**
MSRCv1	67.57±1.27	75.08±3.85	78.48±0.00	77.21±0.00	75.45±0.96	71.43±1.43	72.79±3.66	**86.67±0.00**
BBCSport	65.07±0.02	47.79±2.96	69.85±0.00	80.15±0.00	70.52±4.81	**85.33±1.54**	76.30±3.36	84.92±0.00
ORL	68.77±1.57	82.52±1.66	85.25±0.00	83.05±0.00	**89.51±0.78**	80.75±1.33	77.09±2.29	86.75±0.00
Caltech101-20	55.89±1.08	45.34±1.78	75.52±0.00	59.47±0.00	49.72±2.04	70.38±2.10	62.78±3.22	**76.28±0.00**
UCI	79.15±1.96	87.49±1.01	85.65±0.00	85.48±0.00	78.72±3.29	84.38±1.87	83.10±3.01	**88.05±0.00**

习到了更精确的一致特征表示。

DFMPC在ACC、NMI、ARI和PUR四个指标上都明显优于Co-reg经典协同训练多视图方法。相比之下，DFMPC能更好区分不同视图的可靠性。

基于图的SwMC方法在3sources和Caltech101-20数据集的ARI指标分别比DFMPC高7.21个百分点和10.6个百分点，但在其他数据集中DFMPC表现更好。这表明通过联合学习单个图和统一图，DFMPC可以融合所有视图学习更好的一致特征表示。

多数真实数据集上的结果表明，DFMPC算法优于MSC_IAS、SwMPC、MCLES和GFSC多视图子空间聚类方法。这是因为DFMPC对高维数据投影降维，并用非参数自加权方法提高了聚类效果。

3.3 收敛性分析

为了验证DFMPC算法的收敛性，图2显示了8个真实数据集的目标函数值的收敛曲线，x轴和y轴分别表示迭代次数和相应的目标值。目标函数值与迭代次数成反比，目标函数值在10次迭代内急剧下降达到最小值并趋于稳定。这表明DFMPC收敛效率很高。

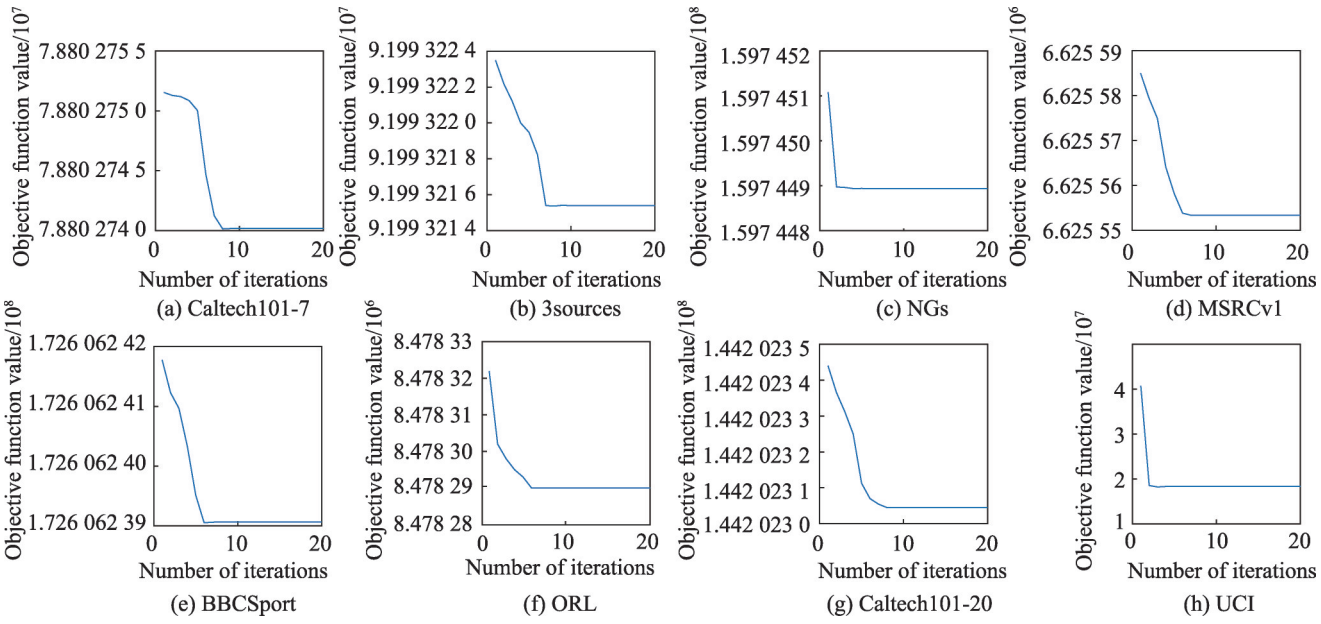


图2 DFMPC在8个数据集上的收敛曲线

Fig.2 Convergence curves of DFMPC on 8 datasets

3.4 鲁棒性分析

本节通过向不同维度真实数据集添加均值为0，方差0到0.5（步长为0.05）不同强度的随机噪声来验证算法的鲁棒性。由于页面空间限制，只展示了4种效果较好的多视图聚类算法在ORL数据集的鲁棒性。如图3所示，随着噪声的增加，所有算法的性能都会下降，但DFMPC的性能在绝大多数情况下领先其他算法并且算法性能增益更为显著。当随机噪声方差从0增到0.5时，DFMPC与SwMC的结果相比，在ACC、NMI、ARI和PUR的性能增益从2.00、3.29、1.36、1.50个百分点提高到14.00、4.55、26.50、15.25个百分点。与MCLES算法相比，DFMPC的4个指标性能增益从5.50、6.35、6.83、6.00个百分点提高到26.00、16.21、34.78、26.69个百分点。与SwMPC算法相比，DFMPC的4个指标性能增益从5.05、4.53、4.76、3.70个百分点提高到17.27、13.22、28.27、18.50个百分点。这些结果表明，当数据集含有噪声时，正交约束应用于散射矩阵进行低维子空间学习的DFMPC方法能够抑制噪声并保持良好的聚类性能，具有良好的鲁棒性。

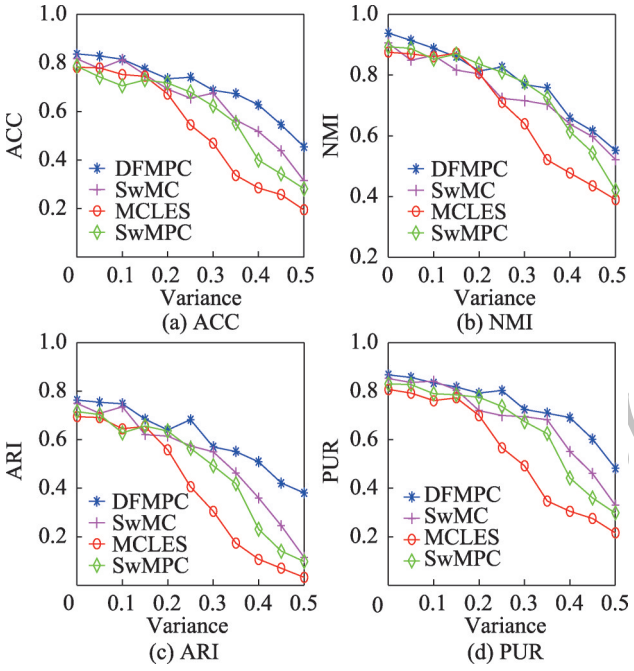


图3 ORL数据集上4种多视图方法的鲁棒性比较

Fig.3 Robustness comparison of 4 multi-view algorithms on ORL dataset

3.5 可视化分析

为了更直观地观察DFMPC算法的聚类性能，本文采用一种非线性降维的*t*-分布式随机邻近嵌入（t-distributed stochastic neighbor embedding，t-SNE）算法[19]，将每个视图的原始特征和由DFMPC得到的一致相似特征映射到2D空间，使样本点在2D空间中被可视化。由于页面空间限制，本文只提供了3sources和NGs数据集的可视化结果，具体如图4所示，其中不同的颜色表示不同的类别。本文方法可以清楚地揭示底层的聚类结构，即动态融合得到的一致相似表示比每个视图的原始特征更能体现良好的聚类结构。这进一步证实了DFMPC算法的有效性。

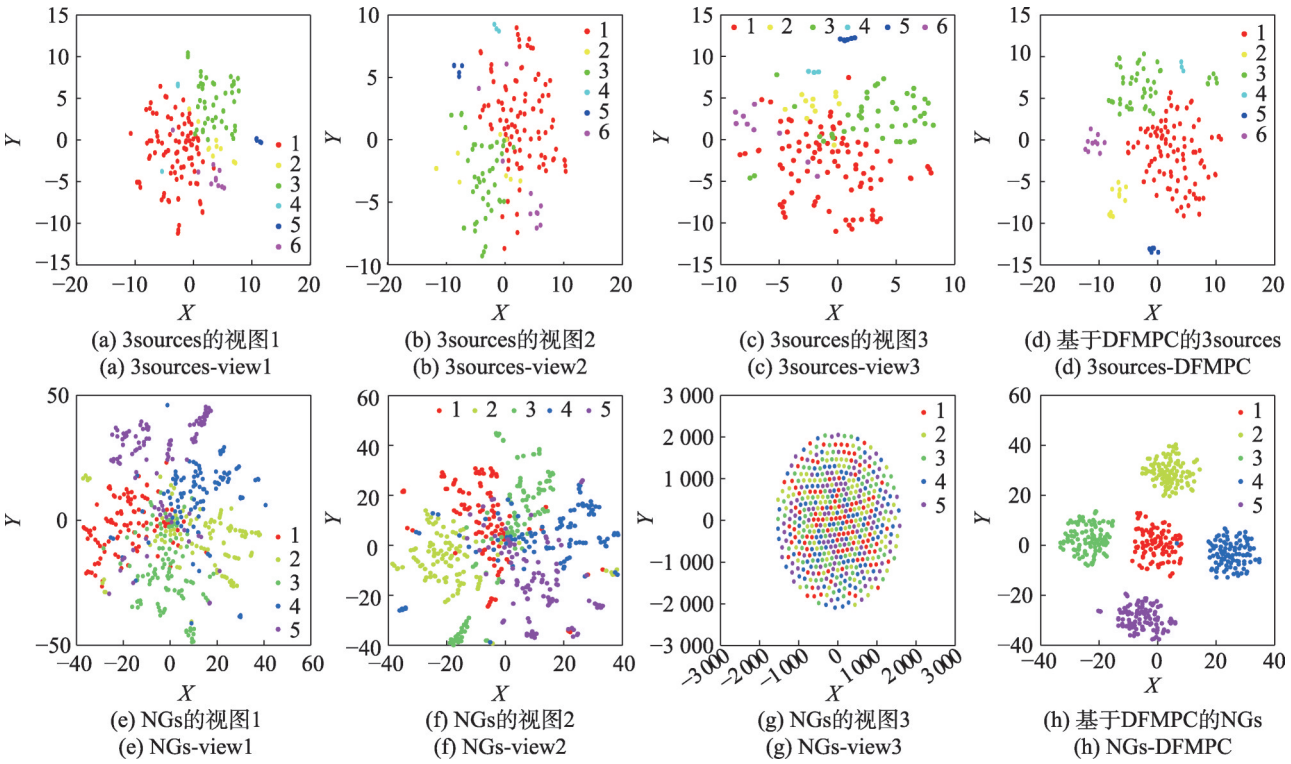


图4 *t*-SNE在3sources和NGs数据集上的可视化结果

Fig.4 Visualization results of *t*-SNE on 3sources and NGs datasets

4 结束语

本文提出了一种新的动态融合的多视图投影聚类算法，在统一的优化框架下联合学习自适应图、无参数的自权重图融合和精确的聚类标签。该算法可以同时研究投影矩阵、相似矩阵、共识矩阵和聚类标签，在低维空间上得到清楚的底层聚类结构。最后，直接从动态融合得到的最佳共识矩阵得到聚类结果。在人工数据集和真实数据集上的实验证明了本文算法的有效性和良好性能。

参考文献：

[1] YANG Y, WANG H. Multi-view clustering: a survey[J]. Big Data Mining and Analytics, 2018, 1(2): 83-107.

[2] KUMAR A, RAI P, DAUME H. Co-regularized multi-view spectral clustering[C]//Advances in Neural Information Processing Systems 24, Granada, Dec 12-14, 2011: 1413-1421.

[3] 张炜, 邓赵红, 王士同. 基于核诱导的不完整多视角聚类[J]. 计算机科学与探索, 2021, 15(2): 284-293.
ZHANG W, DENG Z H, WANG S T. Kernel-induced incomplete multi-view clustering[J]. Journal of Frontiers of Computer Science and Technology, 2021, 15(2): 284-293.

[4] NIE F P, CAI G H, LI X L. Multi-view clustering and semi-supervised classification with adaptive neighbours[C]//Proceedings of the 31st AAAI Conference on Artificial Intelligence, San Francisco, Feb 4-9, 2017. Menlo Park: AAAI, 2017: 2408-2414.

[5] 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, 2021, 15(6): 1062-1073.
FAN R D, HOU C P. Robust auto-weighted multiview subspace clustering[J]. Journal of Frontiers of Computer Science and Technology, 2021, 15(6): 1062-1073.

[6] NIE F P, LI J, LI X L. Self-weighted multiview clustering with multiple graphs[C]//Proceedings of the 26th International Joint Conference on Artificial Intelligence, Melbourne, Aug 19-25, 2017: 2564-2570.

[7] ZHAN K, ZHANG C, GUAN J, et al. Graph learning for multiview clustering[J]. IEEE Transactions on Cybernetics, 2017, 48(10): 2887-2895.

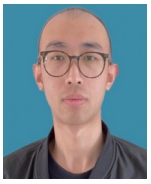
[8] CHEN M S, HUANG L, WANG C D, et al. Relaxed multi-view clustering in latent embedding space[J]. Information Fusion, 2021, 68: 8-21.

[9] WANG R, NIE F P, 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.

[10] WANG X B, LEI Z, GUO X J, et al. Multi-view subspace clustering with intactness-aware similarity[J]. Pattern Recognition, 2019, 88: 50-63.

[11] XIA S, PENG D, MENG D, et al. A fast adaptive k-means with no bounds[J]. IEEE Transactions on Pattern Analysis

and Machine Intelligence, 2022, 44(1): 87-99.
[12] DING Y, ZHAO Y, SHEN X, et al. Yinyang K-means: a drop-in replacement of the classic K-means with consistent speed-up[C]//Proceedings of the 32nd International Conference on Machine Learning, Lille, Jul 6-11, 2015: 579-587.
[13] MOHAR B, ALAVI Y, CHARTRAND G, et al. The Laplacian spectrum of graphs[J]. Graph Theory, Combinatorics and Applications, 1991, 18(7): 871-898.
[14] 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): 652-655.
[15] VANDENBERGHE L, BOYD S. Convex optimization[M]. Cambridge: Cambridge University Press, 2004: 146-159.
[16] DUCHI J C, SHALEV-SHWARTZ S, SINGER Y, et al. Efficient projections onto the l1-ball for learning in high dimensions[C]//Proceedings of the 25th International Conference on Machine Learning, Helsinki, Jun 5-9, 2008. New York: ACM, 2008: 272-279.
[17] NIE F P, WANG X Q, HUANG H. Clustering and projected clustering with adaptive neighbors[C]//Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, Aug 24-27, 2014. New York: ACM, 2014: 977-986.
[18] KANG Z, SHI G, HUANG S, et al. Multi-graph fusion for multi-view spectral clustering[J]. Knowledge-Based Systems, 2020, 189: 105102.
[19] VAN DER MAATEN L, HINTON G. Visualizing data using t-SNE[J]. Journal of Machine Learning Research, 2008, 9 (11): 2579-2605.



姜凯彬(1998—),男,山东烟台人,硕士研究生,CCF学生会员,主要研究方向为人工智能、机器学习。
JIANG Kaibin, born in 1998, M.S. candidate, student member of CCF. His research interests include artificial intelligence and machine learning.

周世兵(1972—),男,江苏盐城人,博士,副教授,主要研究方向为模式识别、人工智能。
ZHOU Shibing, born in 1972, Ph.D., associate professor. His research interests include pattern recognition and artificial intelligence.

钱雪忠(1967—),男,江苏无锡人,硕士,副教授,CCF会员,主要研究方向为数据挖掘、机器学习、人工智能。
QIAN Xuezhong, born in 1967, M.S., associate professor, member of CCF. His research interests include data mining, machine learning and artificial intelligence.

管娇娇(1995—),女,河南信阳人,硕士研究生,CCF学生会员,主要研究方向为模式识别、机器学习。
GUAN Jiaojiao, born in 1995, M.S. candidate, student member of CCF. Her research interests include pattern recognition and machine learning.