自适应多视角子空间聚类

唐启凡，张玉龙，何士豪，周志豪
（西安交通大学软件学院，710049，西安）

摘要：针对传统子空间聚类算法与超参数自适应算法结合时，面对特定多视角数据集的性能失效问题，提出了一种自适应多视角子空间聚类算法。基于不同视角数据点对之间相似度近似的原理，构建了表示矩阵和相似度矩阵在多视角范围内的相关性约束。通过设计的子空间聚类目标函数，在函数第1项和第2项构建了数据矩阵和表示矩阵在单个视角内的相关性约束，在第3项使用径向基函数构建单个视角内表示数据点对之间的相似度，并在与该点对所对应的连接图中的点做相似度差值处理，约束该值在多个视角内的一致性。使用块坐标下降法求解目标函数，根据线性最小二乘法原理自动求解超参数，在得到的连接图上应用谱聚类算法得到聚类分配结果。7个真实数据集上的实验结果表明：与无参数聚类算法COMIC相比，所提算法在4个数据集上性能分别提升了9.3%、6.8%、55.8%、15.39%，并解决了3个数据集中对比算法性能失效的问题；与7个手动调参算法相比，取得了2.15的平均排名。

关键词：多视角聚类；子空间聚类；超参数自适应
中图分类号：TP181 **文献标志码**：A
DOI：10.7652/xjtuxb202105012 **文章编号**：0253-987X(2021)05-0102-11

OSID码

Adaptive Multi-View Subspace Clustering

TANG Qifan, ZHANG Yulong, HE Shihao, ZHOU Zhihao
(School of Software Engineering, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: Aiming at the combination of traditional subspace clustering algorithm and super-parameter adaptive algorithm, an adaptive multi-view subspace clustering algorithm is proposed to deal with performance failure in specific data sets. According to the principle of similarity approximation between different view data point pairs, the correlation constraint of representation matrix and similarity matrix in multi-view range is constructed. Based on the designed subspace clustering objective function, the correlation between data matrix and representation matrix in a single view is constrained in the first and second terms of the function. In the third term, RBF radial basis function is used to construct the similarity between pairs of data points in a single view, and then the difference is processed at the points in the connection graph corresponding to the pairs of points. The consistency of the values in multiple views is constrained. To solve the above objective function, the block coordinate descent method is used, and the super-parameters are automatically solved by the principle of linear least square method. Based on the obtained connection graph, the spectral clustering algorithm is applied, and the cluster allocation result is finally obtained. Compared with COMIC, a nonparametric clustering

收稿日期：2020-11-18。 作者简介：唐启凡(1997—)，男，硕士生；张玉龙(通信作者)，男，讲师。 基金项目：国家自然科学基金资助项目(61573273)。
网络出版时间：2021-01-18 网络出版地址：http://kns.cnki.net/kcms/detail/61.1069.T.20210118.1136.002.html

method, the performance of the proposed algorithm on four data sets is improved by 9.3%, 6.8%, 55.8% and 15.39% respectively, and the problem of performance failure on three data sets is solved. In a comparison with the other seven manual parameter adjustment algorithms, the average ranking of the proposed algorithm is 2.15.

Keywords: multi-view clustering; subspace clustering; adaptive parameters tuning

近年来，随着计算机性能的大幅度提升和互联网通信速度与用户的大规模增长，海量的数据被提取出来。人们希望从这些海量的数据中找出有意义的结构信息，其中以多视角数据最为突出。举例说明多视角数据：①新闻既有视频形式也有音频形式；②网页既可以被 URL 地址所表示也可以被其中内容所表示；③雕像可以从不同侧面表示；④同一种语义有多种语言的表达方式。聚类算法作为一种行之有效的手段受到了研究人员广泛的关注，聚类的目的是将具有潜在相似结构的数据归为一类，从而为后续其他针对于数据的处理做准备。传统的单视角聚类算法在对待多视角数据时，会将每一个视角上的特征简单地连接在同一个矩阵中，并在该特征矩阵上直接运行聚类算法，例如 k-means、谱聚类等算法。但是，这样做有极大的缺点和限制：对于样本数较小的数据，如果将所有视角的特征简单连接起来会造成维度灾难，而且在连接时没有考虑不同视角数据之间的互补性和异构性，会造成多视角数据大量信息的缺失。由此，多视角聚类算法凭借其对于多视角数据独特的处理能力、良好的鲁棒性及泛化能力得到了研究人员越来越多的关注和研究。

从 21 世纪开始，有许多科研人员在多视角聚类领域内取得了不错的成果。为了挖掘数据的潜在结构，Elhamifar 等提出稀疏子空间聚类算法，该聚类算法的主要思想是高维数据通常能被低维的子空间数据所表示，可由公式 $\boldsymbol{X}=\boldsymbol{DC}$ 来表示，其中矩阵 $\boldsymbol{D}$ 为字典矩阵，矩阵 $\boldsymbol{C}$ 为权重矩阵[1]。通过对权重矩阵添加稀疏性约束，从而发掘出数据矩阵的低维表示。对学习出的权重矩阵应用传统的聚类算法，便可得到最终的结果。在 Elhamifar 等[1]的基础上，Liu 等提出了对权重矩阵添加低秩性约束，并在损失中加入噪声矩阵 $\boldsymbol{E}$，从而在挖掘出数据隐藏结构的同时去除数据矩阵中夹杂的噪声点、离群点以及损坏点[2]。通过对噪声矩阵约束不同的范数，从而达到对不同噪点的良好处理效果。在 Elhamifar 等[1]和 Liu 等[2]的研究中可以发现，现实中为了简化问题的处理，通常所使用的字典矩阵 $\boldsymbol{D}$ 都是原始数据矩阵 $\boldsymbol{X}$，这一做法的原理是数据的自表达性质，即类内数据点可以被类中其他所有数据点线性表示。Ji 等在传统深度学习网络 Auto-Encoder 的基础之上，通过利用数据的自表达性质，设计了神经网络的一个卷积层，使其夹在编码网络层和解码网络层中间[3]，开创了多视角聚类深度学习的先河。

这些算法取得了瞩目的成绩，但绝大多数受制于参数选择。为了选择更好的参数，这些算法都会利用归一化互信息评价指标来评价算法的性能，这又要求数据必须包含标签。但是，现实中的数据由于数据量过大，绝大多数情况下都很难对数据做出标注，例如道路视频监控数据、互联网舆情文本数据等。为此，许多免除参数选择的算法被提出[4-8]。这些算法利用计算得到的权重矩阵、相似度矩阵、指示矩阵来直接计算参数，在算法运行过程中不断更新参数，从而免除了人工选择参数。但是，这类算法依然存在缺陷，例如：在 Peng 等提出的算法中，由于数据表示矩阵 $\boldsymbol{Z}$ 和相似度矩阵 $\boldsymbol{S}$ 仅考虑了单个视角内的相关性，缺失了对于表示矩阵和相似度矩阵在多视角范围内的相关性约束，造成了算法在特定多视角数据集上的性能失效[8]；在 Nie 等提出的算法中，使用 RatioCut 切图算法，通过指示矩阵迭代更新各个聚类的权重，该切图算法使聚类之间相似性最小，但在考虑类内相似度时，仅将最大化类内数据点数作为目标，缺乏对于类内数据点之间相似度的深入挖掘，导致在大规模数据集上应用效果欠佳[4]。

为此，本文提出自适应多视角子空间聚类算法 ASC，基于不同视角数据点对之间的相似度近似的原理，在多视角范围内约束表示矩阵和相似度矩阵间的相关性，构建了全新的损失函数，并使用块坐标下降法来优化，从而彻底解决了这类算法在部分数据集中失效的问题，提高了算法的性能。

1　相关工作

近几十年来，聚类算法的研究大致可以分为单视角聚类和多视角聚类两大类。本节挑选了其中最具代表性的算法进行简单总结，基本上概括了聚类算法的发展历程。

1.1 单视角聚类

在早期数据结构不复杂且数据量不庞大的背景下，聚类算法研究的普遍共识是同一类别的数据会围绕在一些标志性点的附近。所以，只要找出几个标志性的数据点，并根据其他数据点与标志点的相似度，便可以得到数据的分类情况。最具代表性的以这种思想为核心的算法为 k-means 算法。该算法首先根据聚类数来随机初始化标志性点，再将数据点分配到离它们最近的标志性点，最后根据每个聚类的重心更新标志点，循环往复直到标志点不再变化。与 k-means 不同，研究者基于概率密度函数和高斯混合模型，提出了最大期望算法。该算法将每一个类别都看作一个单高斯分布模型，在 E 步用高斯概率密度函数计算每个数据点所属的类别，在 M 步计算每个聚类的联合概率分布，优化聚类参数。这两种算法都是在原始数据矩阵上直接应用聚类算法，算法性能往往会受限于原始数据结构，而谱聚类算法首先对原始数据矩阵进行处理，挖掘出数据矩阵潜藏的结构，处理完成后再使用 k-means 等基础算法聚类。相比于 k-means 算法，谱聚类算法性能提高显著，尤其对于处理稀疏矩阵效果良好。

k-means、最大期望、谱聚类这 3 种算法假定所有数据点都处于同一维空间，对于高维数据，性能往往下降严重。为此，研究者提出了子空间聚类算法，将数据矩阵视为由多个维度子空间构成的公共空间，处于同一子空间中的数据点划分为一类。空间聚类算法中最具代表性的算法为 Elhamifar 等提出的稀疏子空间聚类[1]以及 Liu 等提出的低秩子空间聚类[2]，这两种子空间聚类算法分别对学习出的系数矩阵约束 L_1 范数和核范数的正则化项，最终所学习出的系数矩阵具有良好的块对角性质。文献[9]采用 L_2 范数作为正则化项，同样取得了不错的成果。在 Elhamifar 等[1]和 Liu 等[2]的基础上，Wang 等将低秩和稀疏的性质结合起来，使得学习出的系数矩阵同时具有稀疏与低秩两种特性[10]，从而提升了算法性能。

单视角聚类算法虽取得了长足发展，但随着其他各学科的发展，数据视角数越来越多，数据量也越来越大。因此，对聚类算法的要求也逐步提高，聚类结果既要保持多视角数据的一致性，也要挖掘出多视角数据的互补性。此外，单视角聚类算法还要面对包含大量噪声的数据，已逐步无法满足性能要求。

1.2 多视角聚类

与单视角聚类算法相比，多视角聚类算法通常可以挖掘出多视角数据中的隐藏信息与结构，并可以利用这些信息与结构显著地提升算法性能。多视角聚类算法通常分为基于图的多视角聚类、基于矩阵投影的多视角聚类以及多视角子空间聚类 3 类。

基于图的多视角聚类算法将数据点看作图中的结点，图中边上的数据为结点之间的相似度，常用的相似度度量算法有高斯核函数、$k$nn 算法等。构建好图之后再采用切分图的算法将图分为几类，最后应用单视角聚类算法即可[11-13]。Kumar 等将协同正则化算法与基于图的多视角聚类算法相结合，将成对视角的数据矩阵投影到嵌入空间中，使它们在该空间中的相关性达到最大[14]，但仅限于成对的视角，算法的性能随着视角数的提升而下降。基于矩阵投影的算法是将不同视角的数据投影到公共低维特征空间中，再进一步优化它们之间的相关性，最典型的算法为 Chaudhuri 等提出的算法，他们通过结合 CCA 算法来达到这一目的[15]。多视角子空间聚类算法与单视角子空间聚类算法本质相同，但加入了更多的正则化项来挖掘多视角数据潜在的互补信息与数据结构。Cao 等引入了希尔伯特施密特准则来衡量视角之间的多样性，从而挖掘出更多的互补性信息[16]。Zhang 等在子空间聚类的基础上引入了张量，基于张量对矩阵进行分解，并加入矩阵低秩约束，得到了较好的效果[17]。本文算法归属于多视角子空间聚类算法，在子空间聚类基本原理的基础上加入了多个行之有效的正则化项，同时根据数据结构寻找自适应超参数，既保持了多视角数据的一致性，也挖掘出了多视角数据之间的互补信息。

2 自适应多视角子空间聚类

2.1 损失函数

设矩阵 $\boldsymbol{X}_1,\cdots,\boldsymbol{X}_i,\cdots,\boldsymbol{X}_v$ 表示输入的多视角数据矩阵，$\boldsymbol{X}_i\in\mathbf{R}^{n\times d}$，符号 v、n、d 分别表示视角数、样本数和特征数。数据矩阵 $\boldsymbol{X}_i$ 既可以来自于人工数据集，也可以来自于真实数据集，例如，图像数据、文字媒体数据等。

本文提出的损失函数为

$$L = L_1^v + L_2 + L_3 \tag{1}$$

式中：L_1^v 用于约束每一个视角内的数据，主要任务是学习新的表示矩阵 $\boldsymbol{Z}$ 以及连接矩阵 $\boldsymbol{C}$；L_2 和 L_3 用于约束跨视角数据，主要任务是挖掘跨视角数据的互补信息，确保聚类分配一致性以及相似一致性。

L_1^v 公式为

$$L_1^v=\frac{1}{2}\sum_{i=1}^{n}\|\boldsymbol{x}_i^v-\boldsymbol{z}_i^v\|_2^2+\frac{\lambda^v}{2}\sum_{i,j\in\varepsilon}W_{i,j}^v\delta(\|\boldsymbol{z}_i^v-\boldsymbol{z}_j^v\|_2)\tag{2}$$

式中：$\boldsymbol{z}_i^v$ 为学习出的第 v 视角中数据点 $\boldsymbol{x}_i$ 的表示点；$\delta(\cdot)$为施加于正则化项的惩罚函数；λ^v 为平衡前后两项的超参数；ε 为数据点 $\boldsymbol{x}_i$ 的前 m 个最近邻数据点集合；$W_{i,j}^v$是第 v 个视角中权重矩阵的第 i 行第 j 列项。为了提高算法的鲁棒性，本文主要采用 m-$k$nn 算法来计算 $W_{i,j}^v$，而没有使用传统的 $k$nn 算法，其中 m 设置为 10，也就是取前 10 个最近邻数据点，并采用欧式距离来衡量数据点之间的相似度。$W_{i,j}^v$在本文中的主要作用是对数据点筛选，减少计算量，使算法运行在本就具有一定相似性的数据集合中。

为了挖掘数据矩阵 $\boldsymbol{X}$ 的潜在结构，本文重构矩阵 $\boldsymbol{X}$ 为新的表示矩阵 $\boldsymbol{Z}$，因此矩阵 $\boldsymbol{X}$ 与 $\boldsymbol{Z}$ 之间具有天然的相关性，本文在式(2)中的第一项对数据点 $\boldsymbol{x}_i$ 与 $\boldsymbol{z}_i$ 之间的欧几里得距离进行约束。针对式(2)第 2 项，本文采用文献[2,18-19]的思想，在迭代的过程中，属于同一类别数据点的表示点 $\boldsymbol{z}_i$ 会逐渐靠近，并最终坍缩为少数几个标志性点。ASC 算法基本原理如图 1 所示。首先，对原始图片分别提取局部二值模式(LBP)、方向梯度直方图(HOG)等多种特征，代表多个视角的数据。在得到视角数据后，将其输入到 ASC 算法当中。这里对数据做了简化处理，图 1 中仅假设包含三角和圆两种类别。迭代一开始，表示矩阵 $\boldsymbol{Z}$ 与原始数据矩阵 $\boldsymbol{X}$ 一致，当算法迭代 100 次后，相同类别的表示数据点 $\boldsymbol{z}_i$ 会逐渐靠近。同时，处于不同视角相同类别的数据点，它们之间的表示点 $\boldsymbol{z}_i$ 也在逐渐靠近，反之逐渐变远，即不同视角中相同数据点对之间的相关性应该相同。

为了实现这一过程，本文在式(2)第二项针对一对表示点 $\boldsymbol{z}_i$ 与 $\boldsymbol{z}_j$ 的欧几里得距离添加惩罚函数 $\delta(\cdot)$。理想状态下，$\delta(\cdot)$应为 l_0 范数，但由于 l_0 范数是非凸的，造成问题无解，因此通常都松弛到 l_1 范数。然而，来自于真实数据集的矩阵会包含大量的噪点，通过这种数据矩阵计算出的最近邻数据点集合 ε 会包含虚假的数据点，也就是并无相关性的点对。l_1 范数是数据绝对值的和，在优化的过程中，噪点并不会随着迭代而被排除，从而造成算法性能的下降。所以，本文中选择使用 l_2 范数来进行约束，式(2)转换为

$$L_1^v=\frac{1}{2}\sum_{i=1}^{n}\|\boldsymbol{x}_i^v-\boldsymbol{z}_i^v\|_2^2+\frac{\lambda^v}{2}\sum_{i,j\in\varepsilon}W_{i,j}^v\|\boldsymbol{z}_i^v-\boldsymbol{z}_j^v\|_2^2\tag{3}$$

传统的子空间聚类算法仅局限于式(3)，基于 Black 等[20]的线过程思想，引入辅助连接矩阵 $\boldsymbol{C}$，为式(2)中第二项引入系数 $C_{i,j}^v$，在加强表示点对之间相关性的同时，学习出一个连接矩阵 $\boldsymbol{C}$。此时，式(2)变为

$$L_1^v=\frac{1}{2}\sum_{i=1}^{n}\|\boldsymbol{x}_i^v-\boldsymbol{z}_i^v\|_2^2+\frac{\lambda^v}{2}\sum_{i,j\in\varepsilon}W_{i,j}^v(C_{i,j}^v\|\boldsymbol{z}_i^v-\boldsymbol{z}_j^v\|_2^2+\mu(C_{i,j}^v-1)^2)\tag{4}$$

式(4)中最后一项 $\mu(C_{i,j}^v-1)^2$ 有以下作用：①惩罚连接矩阵中第 i 行第 j 列项被忽略的情况，防止有连接在迭代过程中被遗漏；②当点对之间连接还未建立时，该项趋近于 μ，能够提高损失，当点对之间连接已经建立时，该项趋近于 0，能够降低损失。

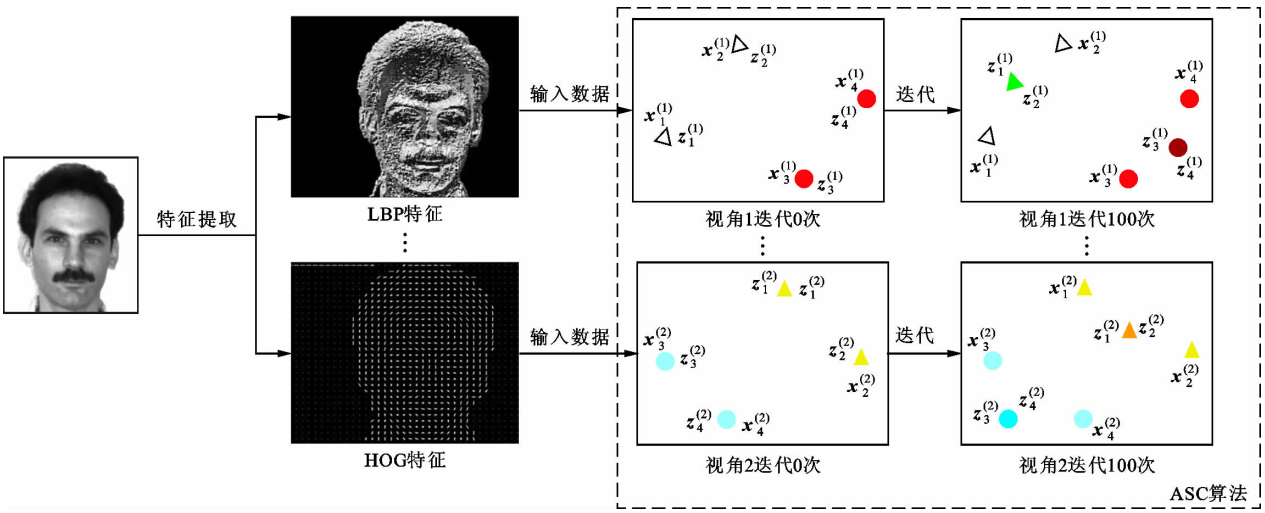

图 1　ASC 算法基本原理

Fig. 1　Basic principle of ASC algorithm

在有了针对于单个视角内数据的约束 L_1^v 后，还必须利用多视角数据的互补信息。L_2 与 L_3 用于建立跨视角的约束条件，实现聚类分配一致性和相似度一致性，公式为

$$L_2 = \frac{1}{2}\sum_{i,j\in\varepsilon}\sum_{v\neq k}(C_{i,j}^v - C_{i,j}^k)^2 \tag{5}$$

$$L_3 = \sum_{v\neq k}(\varphi(\boldsymbol{Z}^v,\boldsymbol{C}^v) - \varphi(\boldsymbol{Z}^k,\boldsymbol{C}^k))^2 \tag{6}$$

式中

$$\varphi(\boldsymbol{Z}^v,\boldsymbol{C}^v) = \sum_{i,j\in\varepsilon}\exp\left(-\frac{\| \boldsymbol{z}_i^v - \boldsymbol{z}_j^v \|_2^2}{2\sigma^2}\right) - C_{i,j}^v \tag{7}$$

L_2 与 L_3 都基于同一种思想：不同视角中相同数据点对(i,j)之间的相关性应该是近似的。对于 L_2，如果在视角 1 中 $C_{i,j}^1=1$，而在视角 2 中 $C_{i,j}^2=0$，则说明这两个视角分别学习出的连接矩阵 $\boldsymbol{C}$ 是不正确的，从而造成损失增大。对于 L_3，在迭代过程中，$C_{i,j}^v$ 不会直接等于 1 或者 0，而是趋向于 1 或者 0，数据点对(i,j)越接近，则 $C_{i,j}^v$ 越趋近于 1，否则趋近于 0。因此，$C_{i,j}^v$ 便有了衡量数据点对(i,j)之间相似度的属性。同样，在迭代过程中，当数据点对(i,j)属于同一个聚类时，它们的表示数据点 $\boldsymbol{z}_i$ 与 $\boldsymbol{z}_j$ 会越来越近，此时可利用高斯核函数对它们之间的相似度进行度量。表示数据点 $\boldsymbol{z}_i$ 与 $\boldsymbol{z}_j$ 之间利用高斯核函数得到的相似度与 $C_{i,j}^v$ 正好互相映射，其单调性是相同的，且都处于$(0,1)$之间。

与 L_2 类似，在不同视角中相同数据点对(i,j)之间的高斯相似度与 $C_{i,j}^v$ 的差应该是近似的，从而得到 L_3 函数。综上，本文得到最终的损失函数

$$\frac{1}{2}\sum_{i=1}^{n}\| \boldsymbol{x}_i^v - \boldsymbol{z}_i^v \|_2^2 + \frac{\lambda^v}{2}\sum_{i,j\in\varepsilon}W_{i,j}^v(C_{i,j}^v\| \boldsymbol{z}_i^v - \boldsymbol{z}_j^v \|_2^2 + \mu(C_{i,j}^v-1)^2) + \frac{1}{2}\sum_{i,j\in\varepsilon}\sum_{v\neq k}(C_{i,j}^v - C_{i,j}^k)^2 + \sum_{v\neq k}(\varphi(\boldsymbol{Z}^v,\boldsymbol{C}^v) - \varphi(\boldsymbol{Z}^k,\boldsymbol{C}^k))^2 \tag{8}$$

2.2 优化

根据 2.1 小节的分析可知，式(8)是关于变量 $\boldsymbol{z}^v$、$\boldsymbol{C}^v$ 的多元凸函数。对于多元凸函数，采用交替最小化策略来进行优化。当变量 $\boldsymbol{Z}^v$ 固定时，对 $C_{i,j}^v$ 求导，并使其公式为 0，以此求解 $C_{i,j}^v$。当变量 $C_{i,j}^v$ 固定时，求导后问题转变为线性最小二乘问题，$\boldsymbol{Z}^v$ 同样可求解。在收敛性方面，该优化方法为传统的块坐标下降法，已被证明是可以收敛的。

当 $\boldsymbol{Z}^v$ 固定时，去掉式(8)中的无关项，并求导为 0，可得

$$C_{i,j}^v = \frac{\mu^v + 2\sum_{v\neq k}C_{i,j}^k}{\mu^v + 2(v-1) + \lambda^v W_{i,j}^v\| \boldsymbol{z}_i^v - \boldsymbol{z}_j^v \|_2^2} \tag{9}$$

由 $v-1=\sum_{v\neq k}C_{i,j}^k$[8] 可得：当表示点 $\boldsymbol{z}_i$ 与 $\boldsymbol{z}_j$ 具有相似的潜在结构时，$C_{i,j}^v$ 趋近于 1；反之，$C_{i,j}^v$ 趋近于 0。

当 $C_{i,j}^v$ 固定时，去掉式(8)中的无关项，可得

$$\frac{1}{2}\sum_{i=1}^{n}\| \boldsymbol{x}_i^v - \boldsymbol{z}_i^v \|_2^2 + \frac{\lambda^v}{2}\sum_{i,j\in\varepsilon}W_{i,j}^v C_{i,j}^v\| \boldsymbol{Z}^v(\boldsymbol{e}_i^v - \boldsymbol{e}_j^v) \|_2^2 + \sum_{v\neq k}\sum_{i,j\in\varepsilon}[\| \boldsymbol{Z}^v(\boldsymbol{e}_i^v - \boldsymbol{e}_j^v) \|_2^2 - \boldsymbol{Z}^k(\boldsymbol{e}_i^k - \boldsymbol{e}_j^k) \|_2^2]^2 \tag{10}$$

式中 $\boldsymbol{e}_i^v$ 为指示向量，其中第 i 个元素为 1。式(10)针对变量 $\boldsymbol{Z}^v$ 求导为 0，可得

$$\boldsymbol{X}^v = \boldsymbol{Z}^v\{\boldsymbol{I} + (\boldsymbol{e}_i^v - \boldsymbol{e}_j^v)(\boldsymbol{e}_i^v - \boldsymbol{e}_j^v)^{\mathrm{T}}[\lambda^v\sum_{i,j\in\varepsilon}W_{i,j}^v C_{i,j}^v + 2((v-1)\| \boldsymbol{z}_i^v - \boldsymbol{z}_j^v \|_2^2 - \sum_{v\neq k}\| \boldsymbol{z}_i^k - \boldsymbol{z}_j^k \|_2^2)]\} \tag{11}$$

将式(11)改写为 $\boldsymbol{X}=\boldsymbol{Z}\boldsymbol{M}$，可以看出，式(11)的解法与线性最小二乘法问题解法相同。将式(11)中的一部分提取出来设为 $\boldsymbol{\Theta}$

$$\boldsymbol{\Theta} = (\boldsymbol{e}_i^v - \boldsymbol{e}_j^v)(\boldsymbol{e}_i^v - \boldsymbol{e}_j^v)^{\mathrm{T}}[\lambda^v\sum_{i,j\in\varepsilon}W_{i,j}^v C_{i,j}^v + 2((v-1)\| \boldsymbol{z}_i^v - \boldsymbol{z}_j^v \|_2^2 - \sum_{v\neq k}\| \boldsymbol{z}_i^k - \boldsymbol{z}_j^k \|_2^2)] \tag{12}$$

易得 $\boldsymbol{\Theta}$ 为拉普拉斯矩阵，根据拉普拉斯矩阵性质，可得到 $\boldsymbol{M}$ 为对称半正定矩阵，因此在 $\boldsymbol{X}^v$ 已知的情况下，便可根据式(11)对 $\boldsymbol{Z}^v$ 进行求解。

2.3 自适应参数选择

式(8)共包含 λ 和 μ 两个超参数。根据文献[18]，当损失函数中数据点 $\boldsymbol{x}_i$ 与表示点 $\boldsymbol{z}_i$ 之间为线性最小二乘法问题时，为了平衡 L_1^v 函数中前后两项之间的关系，设超参数 λ^v 为

$$\lambda^v = \frac{\| \boldsymbol{X} \|_2}{\| \boldsymbol{\Theta} \|_2} \tag{13}$$

通过施加数据矩阵 $\boldsymbol{X}$ 与 $\boldsymbol{\Theta}$ 之间最大奇异值的比例系数，达到前后两项之间的平衡。每一次迭代都仅改变 $\boldsymbol{\Theta}$，数据矩阵 $\boldsymbol{X}$ 不变，因此本文只需在预训练中计算好 $\| \boldsymbol{X} \|_2$ 即可。对于参数 μ，本文初始化 μ 为 θ^2，其中 θ 为矩阵 $\boldsymbol{W}$ 中最大边的长度。ASC 算法伪代码如下。

输入：具有 v 个视角的数据矩阵 $\boldsymbol{X}_1,\cdots,\boldsymbol{X}_i,\cdots,\boldsymbol{X}_v$

1 预训练权重矩阵 $\boldsymbol{W}$ 以及 $\| \boldsymbol{X} \|_2$

2 初始化连接矩阵 $\boldsymbol{C}$ 为单位矩阵，表示矩阵 $\boldsymbol{Z}$ 为数据矩阵 $\boldsymbol{X}$，根据 2.3 节对参数 λ 和 μ 进行计算

3　repeat
4　　利用式(9)更新 $C_{i,j}^v$
5　　利用式(11)更新 $\boldsymbol{Z}^v$
6　　根据式(13)更新 λ
7　until 满足收敛条件或达到最大迭代数
输出：通过在得到的连接矩阵 $\boldsymbol{C}$ 上应用谱聚类算法，得到最终聚类分配结果

3　实　验

3.1　数据集

为了验证算法的鲁棒性及有效性，本文实验所采用的 7 种数据及均来自于现实世界，并对个别数据集做了相应的调整。

MSRCV1[21]包含 8 种类别的 240 张图片。本文挑选了 7 种类别，分别为树、建筑、飞机、牛、人脸、汽车和自行车。同时，对这些图片采用 6 种特征进行描述，即有 CENT、CMT、GIST、HOG、LBP、SIFT 共 6 个视角。

BBCSport[22]包含从 BBC Sport 网站中收集的来自 5 个顶尖领域的 544 篇体育新闻，由两种视角构成。

ORL[23]包含 40 种不同的对象，每种对象包含 10 张不同的特征图像，有 Intensity、BP、Gabor 共 3 种类型的特征被选取。

Caltech101[24]包含从 20 种类别中挑选出的 2 386张图片，每张图片由 Gabor、WM、CENT、HOG、GIST、LBP 共 6 种不同的特征构成。

LandUse-21[25]包含 21 种类别的 100 张河流领域卫星图片，有 GIST、PHOG、LBP 共 3 种特征。

Movies617[26]包含从 IMDb 网站收集的 17 种类别共 617 场电影数据。由两种视角构成，这两种视角分别为1 878个关键字及1 398个演员，每个关键字至少用于 2 部电影，一个演员至少出演 3 部电影。

Scene-15[27]包含4 485张室内和室外场景图片。图片的特征与 Caltech101 数据集一致。

3.2　评价指标

本文采用归一化互信息指标(N)、聚类性能评估指标(V)、预测精确度指标(A)从聚类划分的准确度、聚类性能以及预测精度 3 个方面对算法进行评价，这 3 种评价指标都是值越大越好。N 的计算公式为

$$N(\Omega,C)=\frac{I(\Omega;C)}{(H(\Omega)+H(C))/2} \tag{14}$$

式中：I 为互信息；H 为熵。V 的计算公式为

$$V=\frac{2(1-H(C\mid K))(1-H(K\mid C)H^{-1}(C))}{2-(H(C\mid K)+H(K\mid C))^{-1}(C)} \tag{15}$$

式中：$H(C\mid K)$是给定划分条件下类别划分的熵，$H(C\mid K)=-\sum_{C=1}^{|C|}\sum_{k=1}^{|K|}\frac{n_{c,k}}{n}\log\frac{n_{c,k}}{n_k}$，$n$ 表示样本总数，n_k 表示类别 k 下的样本数，$n_{c,k}$表示类别 c 中被划分为类别 k 的实例数；$H(C)$是类别划分熵，$H(C)=-\sum_{C=1}^{|C|}\frac{n_c}{n}\log\frac{n_c}{n}$，$n_c$ 表示类别 c 下的样本数。A 的计算公式为

$$A=\frac{n_{\text{correct}}}{n} \tag{16}$$

式中 n_{correct}表示聚类划分结果中符合预期的样本数。

3.3　实验设置

对比算法均来自本领域非常经典且具有代表性的算法，其中单视角聚类算法 2 种，多视角聚类算法 5 种，自适应多视角聚类算法 1 种。对于单视角聚类算法，取它们在多个视角中表现最好的那个视角，给 k-means 算法设置对应的聚类数，LRR 算法的超参数根据 Liu 等[2]的设置，在[3,5]内使用最小二乘法逼近 10 次，取最好的结果。对于有参数的多视角聚类算法，首先根据它们各自对应文献中推荐的超参数进行取值，再根据经验对超参数逼近 10 次，取结果最好的那个参数。对于自适应的多视角聚类方法 COMIC，使用文献[8]中推荐的超参数。

3.4　实验结果

每种算法在每个数据集上均运行 30 次，指标采用“平均值±标准差”的格式进行展示。综合排名是本文算法与其他算法在同一数据集上相同评价指标下的排序，1 表示效果最好。实验结果如表 1～7 所示，其中 ASC 为本文算法。

由表 1～7 可知：相比于无参数多视角聚类算法 COMIC，本文算法在数据集 BBCSport、ORL、Movies617 数据集中不会失效，在 MSRCV1 数据集中综合性能提升 9.3%(N、V、A 这 3 个指标提升百分比的平均值)，在 Caltech101 数据集中综合性能提升 6.8%，在 LandUse-21 数据集中综合性能提升 55.8%，在 Scene-15 数据集中综合性能提升 15.39%；相比有参数多视角聚类算法，本文算法在 MSRCV1、Caltech 101、LandUse-21、Movies617 数据集中均取得了前 2 名的成绩，在其他数据集中平均排名 2.15。

在 BBCSport、ORL 和 Scene-15 这 3 个小规模数据集中，本文算法性能与手动调参聚类算法性能

表 1 MSRCV1 数据集对比实验结果

Table 1 Contrast of experimental results on MSRCV1 dataset

算法类型	算法	N	V	A
单视角聚类	k-means	0.500 3±0.012	0.493 3±0.033	0.523 1±0.049
	LRR[2]	0.581 4±0.007	0.581 2±0.002	0.628 0±0.018
有参数的多视角聚类	LMSC[25]	0.590 7±0.004	0.590 7±0.010	0.666 6±0.028
	RMSC[22]	0.562 3±0.025	0.562 3±0.025	0.669 0±0.049
	CSMC[28]	0.594 3±0.012	0.594 1±0.116	0.653 9±0.065
	LTMSC[17]	0.655 6±0.004	0.657 7±0.004	0.742 8±0.000
	Co-REG[14]	0.485 7±0.003	0.488 8±0.012	0.533 9±0.000
自适应多视角聚类	COMIC[8]	0.638 9±0.003	0.612 5±0.026	0.714 2±0.001
	ASC	0.643 9±0.001	0.630 9±0.002	0.688 2±0.031
ASC 算法综合排名		2	2	3

表 2 BBCSport 数据集对比实验结果

Table 2 Contrast experimental results on BBCSport dataset

算法类型	算法	N	V	A
单视角聚类	k-means	0.294 7±0.032	0.282 9±0.021	0.542 8±0.005
	LRR[2]	0.674 3±0.018	0.674 3±0.018	0.802 5±0.005
有参数的多视角聚类	LMSC[25]	0.709 2±0.006	0.709 1±0.025	0.795 2±0.054
	RMSC[22]	0.724 2±0.062	0.724 2±0.036	0.811 7±0.002
	CSMC[28]	0.742 4±0.000	0.724 5±0.000	0.846 6±0.000
	LTMSC[17]	0.726 0±0.006	0.732 0±0.003	0.787 0±0.071
	Co-REG[14]	0.737 0±0.003	0.728 0±0.017	0.746 0±0.006
自适应多视角聚类	COMIC[8]	0.000 0±0.000	0.000 0±0.000	0.000 0±0.000
	ASC	0.725 1±0.001	0.719 3±0.013	0.792 2±0.000
ASC 算法综合排名		3	4	4

表 3 ORL 数据集对比实验结果

Table 3 Contrast experimental results on ORL dataset

算法类型	算法	N	V	A
单视角聚类	k-means	0.799 2±0.004	0.799 1±0.044	0.647 5±0.008
	LRR[2]	0.744 9±0.005	0.744 6±0.007	0.702 5±0.003
有参数的多视角聚类	LMSC[25]	0.813 2±0.022	0.813 1±0.018	0.817 5±0.022
	RMSC[22]	0.772 0±0.012	0.721 8±0.023	0.623 0±0.006
	CSMC[28]	0.808 6±0.009	0.808 6±0.009	0.704 9±0.027
	LTMSC[17]	0.815 2±0.013	0.815 2±0.013	0.704 4±0.016
	Co-REG[14]	0.743 0±0.008	0.745 0±0.004	0.725 0±0.012
自适应多视角聚类	COMIC[8]	0.000 0±0.000	0.000 0±0.000	0.000 0±0.000
	ASC	0.793 2±0.004	0.785 8±0.003	0.806 2±0.011
ASC 算法综合排名		4	5	2

有差距，原因主要有两点：①在小规模数据集中，手动调参聚类算法可以在较短时间内找到准确的超参数，而在大规模数据集中，由于可选参数范围很大，即使经过较长时间的调参，也很难达到较好的性能；

②在小规模数据集中，超参数需要更多轮算法迭代才能获得性能更优的值，而本文算法收敛过快，导致超参数计算不够精准。

综合来看，本文算法在大规模数据集 Caltech 101、

表 4　Caltech101 数据集对比实验结果

Table 4 Contrast experimental results on Caltech101 dataset

算法类型	算法	N	V	A
单视角聚类	k-means	0.634 6±0.005	0.634 6±0.005	0.436 7±0.061
	LRR[2]	0.512 7±0.004	0.510 2±0.007	0.391 1±0.002
有参数的多视角聚类	LMSC[25]	0.620 4±0.017	0.616 9±0.016	0.438 3±0.027
	RMSC[22]	0.410 2±0.003	0.408 8±0.021	0.332 3±0.019
	CSMC[28]	0.653 8±0.016	0.651 6±0.002	0.544 2±0.004
	LTMSC[17]	0.620 9±0.005	0.617 6±0.003	0.458 3±0.005
	Co-REG[14]	0.566 4±0.018	0.565 3±0.007	0.613 4±0.001
自适应多视角聚类	COMIC[8]	0.689 6±0.002	0.683 8±0.016	0.724 4±0.066
	ASC	0.742 5±0.027	0.743 4±0.036	0.756 2±0.002
ASC 算法综合排名		1	2	1

表 5　LandUse-21 数据集对比实验结果

Table 5 Contrast experimental results on LandUse21 dataset

算法类型	算法	N	V	A
单视角聚类	k-means	0.250 8±0.072	0.249 2±0.015	0.176 1±0.015
	LRR[2]	0.341 8±0.009	0.341 4±0.047	0.272 8±0.009
有参数的多视角聚类	LMSC[25]	0.345 6±0.012	0.345 6±0.002	0.305 7±0.044
	RMSC[22]	0.235 9±0.006	0.235 8±0.006	0.201 8±0.007
	CSMC[28]	0.350 5±0.003	0.350 7±0.028	0.316 2±0.000
	LTMSC[17]	0.349 9±0.002	0.349 8±0.038	0.316 1±0.003
	Co-REG[14]	0.305 4±0.002	0.301 3±0.010	0.313 3±0.002
自适应多视角聚类	COMIC[8]	0.432 9±0.005	0.388 9±0.005	0.412 2±0.001
	ASC	0.632 8±0.006	0.597 4±0.007	0.689 3±0.006
ASC 算法综合排名		1	1	1

表 6　Movies617 数据集对比实验结果

Table 6 Contrast experimental results on Movies617 dataset

算法类型	算法	N	V	A
单视角聚类	k-means	0.134 8±0.017	0.104 2±0.009	0.129 6±0.011
	LRR[2]	0.268 2±0.019	0.268 2±0.000	0.267 4±0.018
有参数的多视角聚类	LMSC[25]	0.265 9±0.022	0.265 9±0.021	0.252 8±0.061
	RMSC[22]	0.179 7±0.013	0.179 6±0.003	0.199 3±0.006
	CSMC[28]	0.292 5±0.009	0.292 6±0.008	0.294 2±0.010
	LTMSC[17]	0.275 8±0.007	0.275 6±0.007	0.269 9±0.009
	Co-REG[14]	0.260 0±0.001	0.259 9±0.006	0.257 4±0.002
自适应多视角聚类	COMIC[8]	0.000 0±0.000	0.000 0±0.000	0.000 0±0.000
	ASC	0.402 9±0.004	0.392 5±0.000	0.502 2±0.000
ASC 算法综合排名		1	1	1

表 7 Scene-15 数据集对比实验结果

Table 7 Contrast experimental results on Scene-15 dataset

算法类型	算法	N	V	A
单视角聚类	k-means	0.361 4±0.021	0.290 1±0.074	0.163 6±0.002
	LRR[2]	0.489 7±0.014	0.489 7±0.001	0.492 7±0.008
有参数的多视角聚类	LMSC[25]	0.529 6±0.018	0.529 4±0.034	0.511 7±0.007
	RMSC[22]	0.422 1±0.013	0.422 1±0.002	0.487 1±0.036
	CSMC[28]	0.556 0±0.014	0.556 0±0.012	0.506 8±0.006
	LTMSC[17]	0.516 2±0.036	0.516 1±0.001	0.522 2±0.000
	Co-REG[14]	0.479 1±0.036	0.480 3±0.001	0.512 6±0.004
自适应多视角聚类	COMIC[8]	0.465 9±0.001	0.388 9±0.000	0.455 6±0.091
	ASC	0.495 8±0.007	0.495 6±0.004	0.511 9±0.001
ASC 算法综合排名		3	3	2

LandUse-21 以及噪声数据集 Movies617 表现突出，取得了首位排名的好成绩，并在全部 7 个数据集中相对于无参数算法性能均有提升。由此可见，本文算法性能优异。

3.5 时间复杂度分析

在 Caltech101 数据集上，根据经验对超参数 λ 和 μ 进行人工调参，结果如图 2 所示。

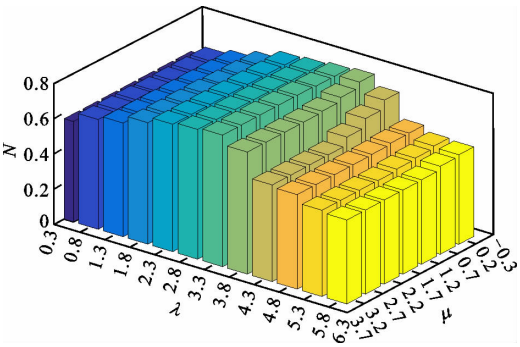

图 2 时间复杂度分析实验结果

Fig. 2 Time complexity analysis of experimental results

共进行 80 次手动调参，耗时 14 h 18 min，超参数 λ 和 μ 最终收敛在 3.8 和 0.2 附近，并得到归一化互信息评价指标最大值为0.751 3。本文自适应超参数算法耗时 11 min 54 s，与手动调参相比性能差值仅为 1.18%。由此可知，本文算法所得超参数基本发挥出算法的最大性能，且在时间和人力成本上有着明显优势。

3.6 收敛性分析

在 Caltech101、LandUse-21 这两个大规模数据集上进行收敛性分析实验，结果如图 3 和图 4 所示。可以看出，即使是在大规模数据集中，本文算法依然能够快速收敛，在 Caltech101 数据集中，迭代 38 次后基本收敛，在 LandUse-21 数据集中，迭代 26 次后基本收敛。这一实验结果证明了本文算法为块坐标下降法属类的基本判断，具有优良的收敛性能。

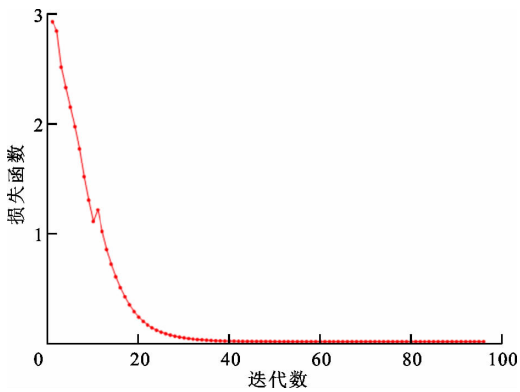

图 3 Caltech101 数据集上的收敛性分析

Fig. 3 Convergence analysis on Caltech101 dataset

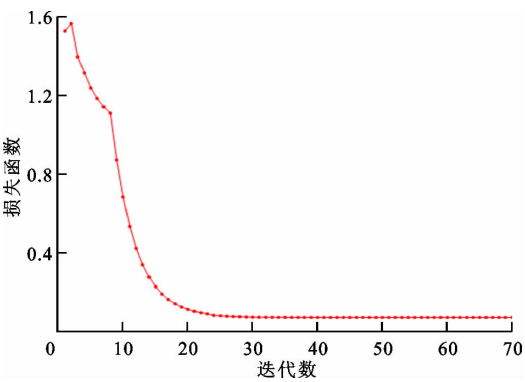

图 4 LandUse-21 数据集上的收敛性分析

Fig. 4 Convergence analysis on LandUse-21 dataset

4 总 结

高维数据矩阵通常包含多个低维度子空间，本

文认为处于同一子空间中的数据是属于同一类的数据。基于此，本文利用子空间进行聚类，提出了自适应多视角子空间聚类算法，结论如下。

(1)来自于现实世界的数据通常都是无标签数据，超参数调优往往是聚类算法难以应用于现实的障碍。本文算法将超参数自适应算法与子空间聚类算法相结合，在保证算法性能的同时极大地降低了通常算法调优过程中所耗费的人力物力成本，使子空间聚类算法应用于现实无标签数据成为可能。

(2)相比于多视角聚类领域中已经提出的无参数聚类算法，本文在前人的基础上进行改进，成功解决了其在特定数据集上算法性能失效的问题，并在多个大规模数据集中取得了良好的性能提升。

(3)本文算法在小规模数据集中与有参数聚类算法存在性能差距，这也是无参数聚类算法的普遍缺陷，后续研究可考虑在小规模数据集的性能上做出提升。

参考文献：

[1] ELHAMIFAR E, VIDAL R. Sparse subspace clustering: algorithm, theory, and applications [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(11): 2765-2781.

[2] LIU Guangcan, LIN Zhouchen, YAN Shuicheng, et al. Robust recovery of subspace structures by low-rank representation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(1): 171-184.

[3] JI Pan, ZHANG Tong, LI Hongdong, et al. Deep subspace clustering networks [C]// Proceedings of the 31st Annual Conference on Advances in Neural Information Processing Systems. Vancouver, Canada: NIPS, 2017: 24-33.

[4] NIE Feiping, LI Jing, LI Xuelong. Parameter-free auto-weighted multiple graph learning: a framework for multiview clustering and semi-supervised classification [C]// Proceedings of the 2016 International Joint Conference on Artificial Intelligence (IJCAI). Sacramento, California, USA: IJCAI, 2016: 1881-1887.

[5] REN Pengzhen, XIAO Yun, XU Pengfei, et al. Robust auto-weighted multi-view clustering [C]// Proceedings of the 2018 International Joint Conference on Artificial Intelligence (IJCAI). Sacramento, California, USA: IJCAI, 2018: 2644-2650.

[6] HUANG Shudong, KANG Zhao, XU Zenglin. Self-weighted multi-view clustering with soft capped norm [J]. Knowledge-Based Systems, 2018, 158: 1-8.

[7] 夏菁，丁世飞. 基于低秩稀疏约束的自权重多视角子空间聚类 [J]. 南京大学学报(自然科学)，2020，56(6)：862-869.
XIA Jing, DING Shifei. Self-weighted multi-view subspace clustering with low-rank sparse constrain [J]. Journal of Nanjing University (Natural Science), 2020, 56(6): 862-869.

[8] PENG X, HUANG Z, LÜ J, et al. COMIC: multi-view clustering without parameter selection [C]// Proceedings of the 36th International Conference on Machine Learning. Princeton, NJ, USA: IMLS, 2019: 8925-8934.

[9] JI Pan, SALZMANN M, LI Hongdong. Efficient dense subspace clustering [C] // Proceedings of the IEEE Winter Conference on Applications of Computer Vision. Piscataway, NJ, USA: IEEE, 2014: 461-468.

[10] WANG Yuxiang, XU Huan, LENG Chenlei. Provable subspace clustering: when LRR meets SSC [J]. IEEE Transactions on Information Theory, 2013, 65(9): 5406-5432.

[11] NIE F, ZHU W, LI X. Unsupervised large graph embedding [C]// Proceedings of the 31st AAAI Conference on Artificial Intelligence. Palo Alto, CA, USA: AAAI, 2017: 2422-2428.

[12] NIE Feiping, ZHANG Rui, LI Xuelong. A generalized power iteration method for solving quadratic problem on the Stiefel manifold [J]. Science China: Information Sciences, 2017, 60(11): 112101.

[13] ZHANG Rui, NIE Feiping, LI Xuelong. Self-weighted spectral clustering with parameter-free constraint [J]. Neurocomputing, 2017, 241: 164-170.

[14] KUMAR A, RAI P, DAUME H. Co-regularized multi-view spectral clustering [C/OL] // Proceedings of the 25th Annual Conference on Neural Information Processing Systems. Vancouver, Canada: NIPS, 2011. [2020-10-01]. https://www.researchgate.net/profile/Abhishek_Kumar64/publication/228469445_Co-regularized_Multi-view_Spectral_Clustering/links/0046353a42e9190dce000000/Co-regularized-Multi-view-Spectral-Clustering.pdf.

[15] CHAUDHURI K, KAKADE S M, LIVESCU K, et al. Multi-view clustering via canonical correlation analysis [C]// Proceedings of the 26th Annual International Conference on Machine Learning. Princeton, NJ, USA: IMLS, 2009: 129-136.

[16] CAO Xiaochun, ZHANG Changqing, FU Huazhu, et

al. Diversity-induced multi-view subspace clustering [C]// Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2015: 586-594.

[17] ZHANG Changqing, FU Huazhu, LIU Si, et al. Low-rank tensor constrained multiview subspace clustering [C]// Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV). Piscataway, NJ, USA: IEEE, 2015: 1582-1590.

[18] SHAH S A, KOLTUN V. Robust continuous clustering [J]. Proceedings of the National Academy of Sciences of the United States of America, 2017, 114 (37): 9814-9819.

[19] ZHANG Zheng, LIU Li, SHEN Fumin, et al. Binary multi-view clustering [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 41(7): 1774-1782.

[20] BLACK M J, RANGARAJAN A. On the unification of line processes, outlier rejection, and robust statistics with applications in early vision [J]. International Journal of Computer Vision, 1996, 19(1): 57-91.

[21] XU Jinglin, HAN Junwei, NIE Feiping. Discriminatively embedded *K*-means for multi-view clustering [C]// Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2016: 5356-5364.

[22] XIA Rongkai, PAN Yan, DU Lei, et al. Robust multi-view spectral clustering via low-rank and sparse decomposition [C]// Proceedings of the 28th AAAI Conference on Artificial Intelligence. Palo Alto, CA, USA: AAAI, 2014: 2149-2155.

[23] VIDAL R, TRON R, HARTLEY R. Multiframe motion segmentation with missing data using PowerFactorization and GPCA [J]. International Journal of Computer Vision, 2008, 79(1): 85-105.

[24] LI Yeqing, NIE Feiping, HUANG Heng, et al. Large-scale multi-view spectral clustering via bipartite graph [C]// Proceedings of the 29th AAAI Conference on Artificial Intelligence. Palo Alto, CA, USA: AAAI, 2015: 2750-2756.

[25] ZHANG Changqing, HU Qinghua, FU Huazhu, et al. Latent multi-view subspace clustering [C]// Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2017: 4333-4341.

[26] CLÉMENT G. Free dataset [DS/OL]. [2020-10-20]. http://membres-lig.imag.fr/grimal/data.html.

[27] ZHAO Handong, DING Zhengming. Multi-view clustering via deep matrix factorization [C]// Proceedings of the 31st AAAI Conference on Artificial Intelligence. Palo Alto, CA, USA: AAAI, 2017: 2921-2927.

[28] LUO Shirui, ZHANG Changqing, ZHANG Wei, et al. Consistent and specific multi-view subspace clustering [C]// Proceedings of the 32nd AAAI Conference on Artificial Intelligence. Palo Alto, CA, USA: AAAI, 2018: 3730-3737.

[本刊相关文献链接]

李敬曼,朱磊,张博,等.一种带约束搜索窗的非局部平均相干斑抑制算法.2020,54(10):54-62.[doi:10.7652/xjtuxb202010007]

张克旭,杜昌旺,赵浩淇,等.针对局灶性癫痫患者的脑电微状态分析.2020,54(9):157-163.[doi:10.7652/xjtuxb202009018]

黄鹤,李昕芮,吴琨,等.引入改进飞蛾扑火的*K*均值交叉迭代聚类算法.2020,54(9):32-39.[doi:10.7652/xjtuxb202009003]

程士卿,郝问裕,李晨,等.低秩张量分解的多视角谱聚类算法.2020,54(3):119-125+133.[doi:10.7652/xjtuxb202003015]

樊红卫,邵偲洁,张旭辉,等.一种对称极坐标图像模糊*C*均值聚类的电主轴失衡故障诊断方法.2019,53(12):57-62+86.[doi:10.7652/xjtuxb201912008]

苏亚丽,吴健行,惠维,等.一种用于视觉颜色特征分类的脉冲神经网络.2019,53(10):115-121.[doi:10.7652/xjtuxb201910016]

(编辑 陶晴)