

基于多视图矩阵分解的聚类分析

张 祎¹ 孔祥维² 王振帆¹ 付海燕¹ 李明¹

摘 要 在计算机视觉和模式识别领域,随着多源信息越来越多,图像的描述方法也越来越丰富,多视图学习方法能更充分利用这种多源信息,进而提高聚类的准确率.因此,本文提出了两种基于多视图学习的方法:MultiGNMF 和 MultiGSemiNMF 方法.该方法是在矩阵分解的基础之上,结合以往多视图学习的框架准则,并利用了样本的局部结构形成的.MultiGNMF 和 MultiGSemiNMF 算法不仅能学习视图间的互补信息,同时能保持样本的空间结构.但是,MultiGNMF 算法只适用于非负的特征矩阵.因此,考虑到 SemiNMF 算法相对于 NMF 算法具有更大的扩展性,结合多视图学习的框架,本文又提出了多视图学习的 MultiGSemiNMF 算法.实验结果证实了这两种方法有较好的性能.

关键词 多视图学习, 聚类, 矩阵分解, 局部结构正则化

引用格式 张祎, 孔祥维, 王振帆, 付海燕, 李明. 基于多视图矩阵分解的聚类分析. 自动化学报, 2018, 44(12): 2160–2169

DOI 10.16383/j.aas.2018.c160636

Matrix Factorization for Multi-view Clustering

ZHANG Yi¹ KONG Xiang-Wei² WANG Zhen-Fan¹ FU Hai-Yan¹ LI Ming¹

Abstract In computer vision and pattern recognition fields, more and more data are represented by multiple views which describe different perspectives of the data. And multi-view learning methods are developed for utilizing the information sufficiently. In this paper, we propose two novel clustering methods called MultiGNMF and MultiGSemiNMF, respectively, which are based on multiview learning with local graph regularization, where the innerview relatedness between data is taken into consideration. However, MultiGNMF is based on NMF, which only applies to non-negative matrix. To eliminate this limit, we propose MultiGSemiNMF based on SemiNMF, which is also applicable for negative matrix. The experimental results demonstrate the effectiveness of our proposed methods.

Key words Multi-view learning, clustering, matrix factorization, local graph regularization

Citation Zhang Yi, Kong Xiang-Wei, Wang Zhen-Fan, Fu Hai-Yan, Li Ming. Matrix factorization for multi-view clustering. *Acta Automatica Sinica*, 2018, 44(12): 2160–2169

在计算机视觉和模式识别领域,输入数据的维数过高会增加计算的复杂度,使得对数据的处理变得困难.为了降低数据的维数,通常采用矩阵分解的方法.矩阵分解的方法将数据矩阵分解成多个矩阵的形式,其中一个矩阵可以很好地逼近原始矩阵,在保留原始信息的条件下,同时可以从高维到低维的映射,学习更好的特征描述方法.这样这种矩阵分解的方法有多种,如 SVD (Singular value decomposition)、QR 分解、NMF (Nonnegative matrix

factorization) 等.

NMF 可以将原始数据分解为基矩阵和系数矩阵的乘积.基矩阵的作用类似人脸的各个不同局部块的描述,系数矩阵为原始数据在这些基向量的线性组合的权重.对于给定的数据矩阵 X ,在学习到对应的基矩阵 U 之后,对应的系数矩阵 V 便可以描述原始矩阵.最近的研究表明,NMF 在人脑识别等领域已经超过基于 SVD 的方法^[1–3].

虽然单视图的 NMF 提高了原始数据的有效性,但是图像可以从不同的视图进行描述.这些视图可以是不同的数据集,也可以是侧重于不同角的特征提取方法,所以,这些视图具有多种特征:有的是视觉底层特征、有的是视觉美学特征,甚至有中层语义特征.这些特征往往可以互补地描述图像^[4],因此,多视图学习相比较于单视图的方法能更有效地利用视图之间的互补信息.而在实例检索^[5]任务上,Multi 的思想^[6]也被用于获得更加精确的检索效果.目前,利用多视图来进行聚类分析的方法主要有两种:多视图谱聚类和多视图 K 均值聚类.其中,多视图谱聚类有文献 [7–11] 等.虽然这些方法

收稿日期 2016-09-07 录用日期 2017-08-17
Manuscript received September 7, 2016; accepted August 17, 2017

国家自然科学基金 (61772111), 国家自然科学基金创新研究群体科学基金 (71421001) 资助

Supported by National Natural Science Foundation of China (61772111) and the Foundation for Innovative Research Groups of the NSFC (71421001)

本文责任编辑 王立威

Recommended by Associate Editor WANG Li-Wei

1. 大连理工大学信息与通信工程学院 大连 116024 2. 浙江大学数据科学与管理工程系 杭州 310058

1. School of Information and Communication Engineering, Dalian University of Technology, Dalian 116024 2. Department of Data Science and Engineering Management, Zhejiang University Hangzhou, 310058

性能的相比单视图的谱聚类有所提高,但是多视图谱聚类需要为每一个视图都计算亲和矩阵,视图越多,计算的复杂度也更大.相比之下,多视图 K 均值聚类不需要计算亲和矩阵,只需要利用原始特征即可.多视图 K 均值聚类又分为基于指示矩阵的方法和基于 NMF 矩阵分解的方法.基于指示矩阵的方法有 RMKMC^[12] 和 DEKM (Discriminatively embedded K-means)^[13] 等,这种方法是从样本角度来形成聚类质心,且这种方法的准确率依赖于指示矩阵的初始值.而基于 NMF 矩阵分解的方法是从维数的角度来进行降维,从而形成聚类质心.具体地,关于矩阵分解多视图学习方面的研究有很多:在文献 [14] 中, Kumar 等利用全局和局部 2 种视图的特征,提高了单视图 NMF 在人脸识别问题上的准确率;文献 [15] 利用稀疏编码框架来对多视图特征学习进行研究;文献 [16] 中, Akata 等提出 Collective-NMF 算法令多种特征数据共享相同的系数矩阵,这等同于先串联各种特征然后进行 NMF 分解.然而 Liu 等认为这种共享系数矩阵具有太强的约束,提出一种弱化的约束,旨在保证每种特征的一致性,即视角间的一致性保持^[17].

但是,以上这些方法没有考虑相同样本在不同视角中的空间结构关系,这种局部空间结构在半监督学习、流形学习等多个领域的方法中具有重要的意义.典型的计算空间局部结构约束的方法有 Locality preserving projection (LPP)^[18] 和 Spectral Regression^[19] 等. Cai 等在文献 [19] 的算法中,通过引入局部结构约束达到了大幅度提高实验准确率的效果.

因此,本文提出了一种基于局部结构约束的多视图特征学习方法,称之为 MultiGNMF.这种方法的主要目标是相似的数据在每个视图都有相近的相似性.因此,我们通过构建亲和矩阵来将每个视图的空间结构加入到目标函数中,并提出迭代规则来解决这个优化问题.然而,MultiGNMF 等多视图学习方法要求特征矩阵的值是非负的,而实际中并不能总是保证这一限制条件.为了消除这一限制条件,本文又提出了一种基于多视图学习的 MultiGSemiNMF 算法.

文章主要结构分为 4 个部分:在第 1 节,简要回顾一下基于矩阵分解的特征学习方法,并将局部结构约束引入多视图学习框架中,提出基于局部结构约束的多视图 NMF 分解算法—MultiGNMF;在第 2 节,针对 MultiGNMF 等多视图学习方法只适用于非负矩阵的缺点,提出 MultiGSemiNMF 算法,并对其进行详细介绍;第 3 节对所提的算法进行实验验证和分析;最后,第 4 节对本文的工作进行总结.

1 基于非负矩阵分解的特征学习方法及其存在问题

在这一节,我们首先介绍基于 NMF 分解的单视图特征学习方法.然后,考虑到多视图特征比单视图特征能更好地描述图像,且分解后的系数矩阵应与原始特征据的空间结构相似,我们提出了基于局部结构约束的多视图特征学习方法—MultiGNMF.下面将分别介绍 NMF 和 MultiGNMF 这两种特征学习算法.

1.1 NMF 特征学习方法

下面介绍 NMF 的求解过程.

定义 $X = [X_{:,1}, \dots, X_{:,N}] \in \mathbf{R}^{M \times N}$ 为输入数据矩阵,每一列都描述的是一个数据, $U \in \mathbf{R}^{M \times K}$ 为基矩阵, $V \in \mathbf{R}^{N \times K}$ 为系数矩阵,即新的描述子,两者的乘积是对原始矩阵的逼近,如式 (1) 所示,其中 K 是一个变量,它决定了特征最终学习的特征维数.

$$X \approx UV^T, \quad U \geq 0, V \geq 0, X \geq 0 \quad (1)$$

如果从每一列的视图来观察这个表达式, $X_j \approx U(V_{j,:})^T = \sum_{k=1}^K U_{:,k} V_{j,k}$, $U_{:,k}$ 表示矩阵的第 k 列, $V_{j,:}$ 代表矩阵的第 j 行,我们称这两个向量为基向量和系数向量.这样可以很清晰地看到,一个样本数据 (X 的列向量) 是基矩阵的各个基向量 ($U_{:,k}$) 的线性累加了.度量矩阵分解的误差的方法有很多种,在欧氏空间中一般采用 l_F 范数来度量矩阵之间的距离.这样整个 NMF 求解便变成如下的有约束优化问题.

$$\min_{U,V} \|X - UV^T\|_F^2, \quad \text{s.t. } U \geq 0, V \geq 0 \quad (2)$$

采用 KKT (Karush-Kuhn-Tucher) 条件,分别引入拉格朗日因子 Ψ 和 Φ 将 $U \geq 0, V \geq 0$ 这两个约束条件变成无约束优化问题,如式 (3) 所示.

$$L = \|X - UV^T\|_F^2 + \text{tr}(\Psi U^T) + \text{tr}(\Phi V^T) \quad (3)$$

对于同时满足 U 和 V 的条件下对于式 (3) 这种无约束优化问题是非凸优化问题,没有精确的求解方法.文献 [19] 提出了乘子更新法则来迭代求解上述最小化约束问题.迭代过程如式 (4) 和 (5) 所示.

$$\begin{aligned} \frac{\partial L}{\partial U} = UV^T V - XV + \Psi &\Rightarrow \\ U_{j,k} = U_{j,k} \times \frac{(XV)_{j,k}}{(UV^T V)_{j,k}} &\quad (4) \end{aligned}$$

$$\frac{\partial L}{\partial V} = VU^T U - X^T U + \Phi \Rightarrow$$

$$V_{j,k} = V_{j,k} \times \frac{(X^T U)_{j,k}}{(VU^T U)_{j,k}} \quad (5)$$

1.2 MultiNMF 特征学习方法

NMF 是基于单视图的特征学习方法, 然而, 对同一个样本而言, 不同视图的多种特征往往可以互补地描述图像. 基于此, 文献 [17] 提出了对每种特征保持弱化的一致性约束的多视图学习方法—MultiNMF. 现将 MultiNMF 介绍如下. 定义 X 为 G 个视图的输入特征数据矩阵, $X^f = [X_{:,1}^f, \dots, X_{:,N}^f] \in \mathbf{R}^{M_f \times N}$, ($f \in [1, G]$) 为第 f 个视图的特征矩阵. 其中, M_f 为第 f 个视图的特征维数, N 为样本数据的数量. 对应地, 定义基矩阵 $U = U^1, \dots, U^G$ 和系数矩阵 $V = V^1, \dots, V^G$, $U^f \in \mathbf{R}^{M_f \times K}$, $V^f \in \mathbf{R}^{N \times K}$. 其中, K 为定义的新特征维数. Liu 在文献 [17] 中提出 Soft-regularization 的方法认为, 对每个视图映射后的特征 V^f 需要保持一致性的约束, 但是允许存在差异性, 差异的大小采用欧氏空间中的 l_F 范数来度量. 于是, 可以得到 Liu 的方法定义的损失函数如下式 (6) 所示.

$$\min_{U^f, V^f, V^*} \sum_{f=1}^G \|X^f - U^f V^f\|_F^2 + \lambda_f \|V^f - V^*\|_F^2$$

$$\text{s.t. } X^f \geq 0, U^f \geq 0, V^f \geq 0, V^* \geq 0 \quad (6)$$

$$\|U_{*,k}^f\|_1 = 1, k = 1, \dots, K, f = 1, \dots, G$$

方程的解需要采用乘子更新法则来进行迭代求得, 具体过程这里不再赘述.

1.3 MultiGNMF 特征学习方法

局部空间结构约束的思想主要是保持样本的局部空间结构不变 (近似不变), 这种局部空间结构在多个领域的方法中具有重要的意义. 而 MultiNMF 没有考虑相同样本的原始数据与降维之后数据的空间结构关系, 因此, 本文提出了基于 NMF 矩阵分解的局部结构正则化约束多视图学习方法, 称之为 MultiGNMF. 下面对 MultiGNMF 方法进行详细介绍.

1.3.1 局部结构正则化约束

局部空间结构近似不变的具体含义就是: 如果两个样本 x_i 和 x_j 在原始空间中相似, 那么我们认为它们在映射后的空间中 (分别用 $V_{i,\cdot}$ 和 $V_{j,\cdot}$ 表示) 也应该有近似的相似程度, 即原始数据与映射后的数据有相似的局部结构关系.

根据以往文献的考察, 我们采用数据的亲和矩阵 W 来表征局部结构关系. 计算亲和矩阵的方法

很多, 在没有监督信息的前提下, 一般采用数据之间的欧氏距离构建亲和矩阵 W . 定义一个正整数变量 k , 一个样本 i 与其他样本的权重可以如此计算: 对所有样本与该样本的距离进行从小到大的排序, 如果样本 j 在前 k 个, 那么 $W_{ij} = 1$; 否则, $W_{ij} = 0$. 还有别的方法不采用 0/1 值作为权值, 而是直接用数据之间的核函数值作为权值. 典型的核函数有高斯核 $W_{ij} = e^{-\frac{\|x_i - x_j\|^2}{\sigma}}$, 线性核 $W_{ij} = x_i^T x_j$ 等. 不同的计算方法适应于不同的研究内容, 在文档特征中采用线性核表现更好, 在图像视觉特征中高斯核可能更适合. 在和别的方法比较的过程中, 我们采用比较通用的高斯核函数值作为权值, 从而可以得到原始数据的亲和矩阵 W . 但是, 这些方法的性能受参数的影响较大. 最近, Nie 等^[20] 提出了一种无参数构建亲和矩阵 W 的方法, 该方法不需要任何参数, 很好地解决了构图过程需要反复调节参数的问题. 但本文提出的 MultiGSemiNMF 和 MultiGNMF 算法直接应用无参数构图后缺乏普适性, 因此本文仍然使用高斯核函数来构建亲和矩阵, 在以后的工作将深入研究无参数构图和本文方法的结合. 对于经过映射后的数据 $V_{i,\cdot}$ 和 $V_{j,\cdot}$, 仍用欧氏距离来度量它们之间的相似程度, 如式 (7) 所示. 利用亲和矩阵 W , 我们便可以构造用于约束局部结构的平滑惩罚因子. 对于第 f 个视图的平滑惩罚因子 R^f , 公式如下:

$$D(V_{i,\cdot}, V_{j,\cdot}) = \|V_{j,\cdot}^f - V_{i,\cdot}^f\|^2 \quad (7)$$

$$R^f = \sum_{i,j=1}^N \|V_{j,\cdot}^f - V_{i,\cdot}^f\|^2 W_{ij}^f =$$

$$\text{tr}((V^f)^T D^f V^f) - \text{tr}((V^f)^T W^f V^f) =$$

$$\text{tr}((V^f)^T L^f V^f) \quad (8)$$

其中, $L = D - W$ 为拉普拉斯矩阵, D 是一个对角矩阵 $D_{(j,j)} = \sum_i W_{i,j}$.

1.3.2 目标方程的构建及求解

将平滑惩罚因子引入多视图特征学习 MultiNMF 的框架中, 我们得到 MultiGNMF 方法的目标函数如下:

$$\min_{U^f, V^f, V^*} \sum_{f=1}^G \|X^f - U^f (V^f)^T\|_F^2 + \lambda_f \|V^f - V^*\|_F^2 +$$

$$\mu \lambda_f \text{tr}((V^f)^T L^f V^f)$$

$$\text{s.t. } X^f \geq 0, U^f \geq 0, V^f \geq 0, V^* \geq 0 \quad (9)$$

$$\|U_{*,k}^f\|_1 = 1, k = 1, \dots, K, f = 1, \dots, G$$

Liu 在文献 [17] 中建议用对角阵 Q 归一化 U 和 V , 即 $UV^T = UQ^{-1}QV^T$. 其中, Q_f 定义如下:

$$Q_f := \text{diag} \left(\sum_{m=1}^M U_{m,1}^f, \dots, \sum_{m=1}^M U_{m,K}^f \right) \quad (10)$$

$\text{diag}(\cdot)$ 表示对角矩阵. 这样, 经过归一化后, $\|U_{*,i}\|_1 = 1$. 因此, 式 (9) 可以改写成:

$$\min_{U^f, V^f, V^*} \sum_{f=1}^G \|X^f - U^f(V^f)^T\|_F^2 + \lambda_f \|V^f Q^f - V^*\|_F^2 + \mu \lambda_f \text{Tr}((V^f)^T L^f V^f) \quad (11)$$

$$\text{s. t. } X^f \geq 0, U^f \geq 0, V^f \geq 0, V^* \geq 0$$

由此, 我们可以得到整体的损失函数定义如式 (12) 所示.

$$L_G = \sum_{f=1}^G \|X^f - U^f(V^f)^T\|_F^2 + \lambda_f \|V^f Q^f - V^*\|_F^2 + \mu \lambda_f \text{tr}((V^f)^T L^f V^f) \quad (12)$$

MultiGNMF 在 $U^f \geq 0, V^f \geq 0$ 的约束下最小化 L_G 是一个带约束的优化问题. 采用迭代的方法求解. 引入拉格朗日因子后, MultiGNMF 的第 f 种特征损失函数为式 (13).

$$L_1 = \|X - UV^T\|_F^2 + \lambda_f \|VQ - V^*\|_F^2 + \text{tr}(\Phi V^T) + \mu \lambda_f \text{tr}(V^T L V) + \text{tr}(\Psi U^T) \quad (13)$$

L_1 分别对 U 和 V 求导, 求导结果如下:

$$\frac{\partial L_1}{\partial U} = UV^T V - XV + \Psi + \lambda_f P \quad (14)$$

$$\frac{\partial L_1}{\partial V} = VU^T U - X^T U + \Phi + \lambda_f (V - V^* + \mu LV)$$

其中, $P := (\sum_{m=1}^M U_{m,k} \sum_{n=1}^N V_{n,k}^2 - \sum_{n=1}^N V_{n,k} V_{n,k}^*)$. 采用 KKT 条件 $\Psi_{j,k} U_{j,k} = 0$ 和 $\Phi_{j,k} V_{j,k} = 0$ 得到 U 和 V 的迭代式如下:

$$U_{j,k} \leftarrow U_{j,k} \times \frac{(XV)_{j,k} + \lambda_f \sum_{n=1}^N V_{n,k} V_{n,k}^*}{(UV^T V)_{j,k} + \lambda_f \sum_{n=1}^N V_{n,k}^2} \quad (15)$$

$$V_{j,k} \leftarrow V_{j,k} \times \frac{(X^T U + \lambda_f V^* + \mu \lambda_f W V)_{j,k}}{(V U^T U + \lambda_f V + \mu \lambda_f D V)_{j,k}}$$

根据文献 [17], 我们在每次迭代得到 U 和 V 之后, 按照式 (10) 进行归一化处理, 即:

$$U^f \leftarrow U^f (Q^f)^{-1}, V^f \leftarrow V^f Q^f \quad (16)$$

在优化 U 和 V 之后, 将它们视为常量, 对 L_1 求导得到 V^* 的更新表达式如下所示:

$$V^* = \frac{\sum_{f=1}^G \lambda_f V^f}{\sum_{f=1}^G \lambda_f} \quad (17)$$

NMF 和 MultiGNMF 都要求 $U \geq 0, V \geq 0, X \geq 0$. 这种非负性约束来源于在现实世界中大部分数据是非负的, 系数的累加也是非负的. 然而, 实际情况中, 我们提取到的图像特征往往存在负数, 这就使得以上两种特征学习方法存在局限性. 为了消除这个局限, 我们提出一种对负数特征也适用的多视图特征学习方法—MultiGSemiNMF, 我们将在下一节对其进行详细介绍.

2 MultiGSemiNMF 特征学习方法

在第 1 节中我们分析了各种基于 NMF 矩阵分解的特征学习方法的优越表现, 不过这些方法有一个重要的约束: 所有数据都必须是非负的. 在物理世界中可能大部分数据保持这个特性, 但是, 在图像处理中有些特征, 如由小波变换得到的三分法特征、结构特征等, 并没有保持非负性的条件, 如果强制向正数方向映射, 往往含有一定的失真. 这就使得基于 NMF 的特征学习方法有一定的局限性. 如何将 NMF 算法拓展到对负数矩阵也适用, 文献 [13] 中进行了详细探究. 其中一种方法是 SemiNMF, 下面我们对其进行介绍, 并由此引出本文所提算法—MultiGSemiNMF.

2.1 SemiNMF 特征学习方法

SemiNMF 中的数据与 NMF 中的数据一致, 但是, SemiNMF 算法对原始数据 X 和基矩阵 U 不带有非负性约束, 只需系数矩阵 $V \geq 0$. 由此, 我们可以得到 SemiNMF 的目标方程和损失函数分别如式 (18) 和 (19) 所示:

$$\min_{U, V} \|X - UV^T\|_F^2, \quad \text{s. t. } V \geq 0 \quad (18)$$

$$L_{semi} = \|X - UV^T\|_F^2 + \text{tr}(X X^T - 2X^T U V^T + V U^T U V^T) \quad (19)$$

SemiNMF 在 $V \geq 0$ 的约束下最小化 L_{semi} 是一个带约束的优化问题. 采用迭代的方法求解, 求解过程我们将在第 2.2 节具体介绍.

2.2 MultiGSemiNMF 特征学习方法

SemiNMF 这种方法具有很好的描述性, 同时相对 NMF 有较大使用范围和较高性能. 于是, 为了

克服基于 NMF 的各种特征学习方法只适用于非负数据的缺点, 本文提出基于 SemiNMF 的多视图特征学习算法, 我们称之为 MultiGSemiNMF. 下面将对这种方法做详细的介绍.

2.2.1 目标方程

多视图学习过程中我们需要遵守几个准则: 1) 每个视图学习的新特征需要保持一致性; 2) 在学习前的特征和学习后的特征对样本之间的局部结构进行度量, 需要保持这种结构的一致性. 于是, 我们得到弱化的一致性约束项 $\|V^f - V^*\|_F^2$ 和局部结构正则化约束项 $\text{tr}((V^f)^T L^f V^f)$. 我们将多视图学习的准则应用到 SemiNMF, 于是, 我们可以得到 MultiGSemiNMF 算法的目标函数如式 (20) 所示.

$$\min_{U^f, V^f, V^*} \sum_{f=1}^G \|X^f - U^f (V^f)^T\|_F^2 + \lambda_f \|V^f Q^f - V^*\|_F^2 + \mu \lambda_f \text{tr}((V^f)^T L^f V^f) \quad (20)$$

$$\text{s.t. } V^f \geq 0, V^* \geq 0$$

其中, Q^f 的引进是为了使方程满足 $\|U_{*,k}^f\|_1 = 1$ 的约束条件. 而 $\|U_{*,k}^f\|_1 = 1$ 的约束条件是为了每一种特征的数据归一化处理, 那么对应的系数矩阵也变得归一化, 具有比较性. 又因为 $\|X\| = \|\sum_j X_j\| \approx \sum_{k=1}^K \|U_{i,k} \sum_j V_{j,k}\| = \sum_{k=1}^K \|\sum_j V_{j,k}\| = \|V\|$. 所以, 在 U 归一化之后如果对 X 也进行归一化处理, V 便也达到归一化的目的. 而 X 为原始数据我们能更好的操作, 故在求解过程中只需要在每次迭代 U 之后对其进行归一化即可. 所以, 式 (20) 可以简写成 (21), 但是我们需要求解前对原始数据 X 进行归一化操作, 在每次迭代后对 U 进行归一化.

$$\min_{U^f, V^f, V^*} \sum_{f=1}^G \|X^f - U^f (V^f)^T\|_F^2 + \lambda_f \|V^f Q^f - V^*\|_F^2 + \mu \lambda_f \text{tr}((V^f)^T L^f V^f) \quad (21)$$

$$\text{s.t. } V^f \geq 0, V^* \geq 0$$

由此, 我们可以得到 MultiGSemiNMF 算法的整体损失函数如下所示:

$$L_{GSemi} = \sum_{f=1}^G \|X^f - U^f (V^f)^T\|_F^2 + \lambda_f \|V^f - V^*\|_F^2 + \mu \lambda_f \text{tr}((V^f)^T L^f V^f) \quad (22)$$

其中, $L = D - W$ 为拉普拉斯矩阵, D 是一个对角矩阵 $D_{(j,j)} = \sum_i W_{i,j}$, W 为样本在原始空间关系

的亲邻矩阵. 各变量的具体含义及计算方法详见第 1.3 节.

2.2.2 方程求解

MultiGSemiNMF 在 $V^f \geq 0$ 的约束下最小化 L_{GSemi} 是一个带约束的优化问题. 对于每个视图的特征, 引入拉格朗日因子后, MultiGSemiNMF 的第 f 个视图特征的损失函数为

$$L_2 = \|X - UV^T\|_F^2 + \lambda_f \|V - V^*\|_F^2 + \text{tr}(\Phi V^T) + \mu \lambda_f \text{tr}(V^T L V) \quad (23)$$

损失函数在变量 U 和 V 下不是一个凸优化问题没有精确的数值解, 采用迭代优化方法求解, 具体计算步骤如下:

1) 对 V 取随机矩阵或者采用 K-means 算法得到的系数矩阵.

2) 固定 V 更新 U . 将 L_2 中只有变量 U , 变成一个无约束优化问题, 对其求导 $\frac{\partial L_2}{\partial U} = UV^T V - XV = 0$, 获得 U 的解析解. 其中 $V^T V \in \mathbf{R}^{K \times K}$ 在一般情况下是一个半正定的矩阵, 在不可逆的情况下, 用伪逆矩阵来代替 (Matlab 中的 pinv 函数).

$$U = XV(V^T V)^{-1} \quad (24)$$

3) 固定 U 更新 V . 将 U 固定, L_2 中只有变量 V , 对其求导 $\frac{\partial L_2}{\partial V} = UV^T U - X^T U + \Phi + \lambda_f (V - V^* + \mu LV) = 0$, 采用 KKT 条件 $\Phi_{j,k} V_{j,k} = 0$ 获得式 (25) 的优化问题.

$$(UV^T U - X^T U + \lambda_f (V - V^* + \mu LV))_{ik} V_{ik} = \beta_{ik} V_{ik} = 0 \quad (25)$$

式 (25) 是一个典型的定点等式问题, 求解 $f(x) = 0$, 可以变成 $x = g(x)$ 的形式, 迭代式子 $x_{i+1} = g(x_i)$ 由于存在 $V \geq 0$ 的约束条件, 式 (25) 中的各项需要保持绝对的非负性. 所以, 我们采用拆分的方法. 通过拆分, $X^T U = (X^T U)^+ - (X^T U)^-$, $U^T U = (U^T U)^+ - (U^T U)^-$, 其中的拆分函数定义如式 (26) 所示. 这样式 (25) 便变成多个非负矩阵之间的线性组合.

$$A_{ik}^+ = \frac{|A_{ik}| + A_{ik}}{2}, A_{ik}^- = \frac{|A_{ik}| - A_{ik}}{2} \quad (26)$$

$$V_{ik} = V_{ik} \times$$

$$\sqrt{\frac{X^T U_{j,k}^+ + [V(U^T U)^-]_{j,k} + \lambda_f (\mu W V + V^*)_{j,k}}{X^T U_{j,k}^- + [V(U^T U)^+]_{j,k} + \lambda_f (\mu W V + V)_{j,k}}} \quad (27)$$

文献 [21] 提出一种满足定点等式的 $x = g(x)$ 形式, 将式 (27) 所示, 代入到式 (25) 中 $(VU^T U - X^T U + \lambda_f(V - V^* + \mu LV))_{ik} V_{ik} = 0$, 由于 KKT 条件约束 $\beta_{ik} V_{ik} = 0$, 当 $V_{ik} \neq 0$ 时, 必然存在 $(VU^T U - X^T U + \lambda_f(V - V^* + \mu LV))_{ik} = 0$, 式 (27) 满足式 (25); 反之, 当 $V_{ik} = 0$ 也满足. 综合考虑, 式 (27) 满足 KKT 条件. 迭代式 (27) 的收敛性证明方法可以详见文献 [21-22].

4) 交替迭代第 2) 步和第 3) 步, 直到损失函数值小于阈值或损失函数值变化小于阈值.

5) 在优化 U 和 V 之后, 将它们视为常量, 利用式 (17) 更新 V^* . MultiGSemiNMF 与 MultiGNMF 的不同之处在于缺少了 $U^f \geq 0$ 的约束, 因此不仅适用于特征矩阵非负的情况, 在特征矩阵中存在负数时, 也表现良好. 式 (28) 是各种变量的迭代方程.

$$U_{j,k} \leftarrow U_{j,k} \times \frac{(XV)_{j,k} + \lambda_f \sum_{n=1}^N V_{n,k} V_{n,k}^*}{(UV^T V)_{j,k} + \lambda_f \sum_{n=1}^N V_{n,k}^2}$$

$$V_{ik} \leftarrow V_{ik} \times \sqrt{\frac{X^T U_{j,k}^+ + [V(U^T U)^-]_{j,k} + \lambda_f(\mu W V + V^*)_{j,k}}{X^T U_{j,k}^- + [V(U^T U)^+]_{j,k} + \lambda_f(\mu W V + V)_{j,k}}}$$
(28)

3 实验设计与分析

为了验证 MultiGNMF 算法和 MultiGSemiNMF 算法的有效性, 我们在 4 个公共的图像数据库中和其他几种多特征学习算法比较图像聚类效果. 下面分别从实验设置、评估指标和实验结果及分析这三部分作详细介绍.

3.1 实验设置

3.1.1 数据库

实验中所用的数据库为 CMU PIE 人脸数据库、UCI 手写体数字图像数据库、3-Sources 文本数据库和 ORL 人脸库. 表 1 详细介绍这 4 个数据库的统计特性. 具体介绍如下:

CMU PIE 人脸数据库包含 41 368 张 32×32 像素的人脸图像, 这些数据是 68 个人按照指定的 13 种姿势角度和 43 种不同光照条件下采集人脸的图像. 在实验中, 我们从每个人的一个角度中随机选择 42 幅图像, 构成 2 856 幅人脸图像. 如果按照人脸作为聚类中心的话, 那么数据可以分成 68 个集群. 在本次实验中, 我们用图像的像素值 (二维图像按行展开) 和 HOG (Histogram of oriented gradient) 特征作为 CMU PIE 的两个视图.

UCI 手写体数字数据库有 UCI 大学构建数字 0~9 的手写体图像, 数据库包含 2 000 个样本. 我们利用手写体图像的低频傅里叶变换和原始像素值作为不同的视图.

3-Sources 文本数据库包含 BBC、Reuters 和 Guardian 三种流行的网上新闻资源. 其中, 有 169 条新闻被这三个新闻网报道过. 我们选取这 169 条新闻作为测试样本, 这三个新闻网分别作为三个视图来进行聚类分析. 这些新闻的主题包含商业、娱乐、健康、政治、运动和科技, 我们将其作为我们聚类的标签.

ORL (Olivetti research laboratory) 人脸库是英国剑桥大学 Olivetti 研究所制作的人脸数据库, 它共包含 400 张的人脸图像. 这些数据是 40 个人在不同的时间、变化的光线、面部表情 (张开/合拢眼睛、微笑/不微笑) 和面部细节 (戴眼镜/不戴眼镜) 下拍摄的. 所有的图像为实验者的正脸, 带有一定程度的朝上下左右的偏转或倾斜, 相似的黑暗同质背景. 所有图像的大小均为 28×23 像素. 在本次实验中, 我们用图像的像素值 (二维图像按行展开) 和低频傅里叶变换系数值作为 ORL 的两个视图.

表 1 4 个数据库的资料

Table 1 The information of four databases

数据库	数量	视图个数	聚类个数
CMU PIE	2 856	2	68
UCI Digit	2 000	2	10
3-Sources	169	3	6
ORL	400	2	40

3.1.2 对比算法

为了更全面地评估本文提出的算法, 我们比较了近期多种典型的多视图学习算法. 下面对它们进行描述.

1) 单视图算法 (BSV 和 WSV): 我们对每个视图利用 NMF 算法进行特征学习, 将学习的系数矩阵作为特征. 统计每个视图的效果, 我们取每个视图该算法对应的最好效果为 BSV 和最差效果为 WSV.

2) ConcatNMF: 这是一种前向融合的学习方法. 其算法可以简单理解为, 首先将每个视图的特征串联成一个向量 (矩阵) 作为数据新的特征, 然后利用非负矩阵分解算法进行特征学习, 所有算法流程和度量标准与单视图的方法一致.

3) ColNMF: 一种多视图学习方法, 采用一致性准则, 如式 (29) 所示, 所有的视图共享相同的系数矩阵. 与 ConcatNMF 类似, 不过每个视图添加了

权重和归一化约束.

$$\min_{U^f, V} \sum_{f=1}^G \|X^f - U^f(V)^T\|_F^2 \quad (29)$$

4) Co-reguSC: 一种改进的谱聚类算法, 在文献 [23] 中被 Kumar 等提出. 每个视图采用谱聚类学习作为基本的聚类学习算法, 与 MultiNMF 算法相似的是, 在视图之间采用弱一致性准则约束各个视图的映射矩阵之间的关系, 如式 (30) 所示. 在实验中我们采用高斯核函数, 该算法的详细介绍可以参考文献 [23].

$$\max_{U^v, U^*} \sum_{v=1}^m \text{tr}((U^v)^T L^v U^v) + \sum_v \alpha_v \text{tr}(U^v (U^v)^T U^* (U^*)^T) \quad (30)$$

5) MultiNMF: 本文算法的基准算法, 采用弱一致性约束的多视图 NMF 学习方法.

6) Sc-ML: 是一种改进的 Co-reguSC 算法. Dong 等在文献 [24] 中提出, 利用格拉斯曼流行距离中的一种投影距离来定义不同视图学习的子空间之间一致性准则.

7) MMSC: 是一种多模态的谱聚类方法. 不同的模态 (也就是图像特征) 共享一个相同的图拉普拉斯矩阵, 也就是最后的聚类指示矩阵 G . 该算法的详细介绍可以参考文献 [10].

$$\min_{G \geq 0, G^T G = I, G_i} \sum_i (\text{tr}(G^i)^T L^i G^i + \alpha \text{tr}(G - G^i)^T \times (G - G^i)) \quad (31)$$

8) AMGL: 是一种多视图谱聚类方法. 但是不同视图所占的权重是自动学得的, 不需要有额外的参数来指导训练. 该算法的详细介绍可以参考文献 [11].

3.2 评估指标

本文采用以往文献中的经典评估方法: AC (精确度) 和 NMI (归一化互信息) 来度量图像数据聚类的评估指标. 给定一幅图像, 定义为图像的标签类别, 为图像对应的聚类中心的类别标签, AC 可以用如下式 (31) 来定义. 其中 n 为所有的测试图像数量, 是一个脉冲函数, 函数中两个参数相同则返回 1, 否则返回 0. 是一种将聚类中心的标签映射到图像集已知的对等标签, 其中我们采用 Kuhn-Munkres^[25] 的映射方法.

$$AC = \frac{\sum_{i=1}^n \delta(\alpha_i, \text{map}(l_i))}{n} \quad (31)$$

给定两个数据集的所有聚类中心, 它们之间的互信息可以用式 (32) 计算.

$$MI(\mathcal{C}, \mathcal{C}') = \sum_{c_i \in \mathcal{C}, c'_j \in \mathcal{C}'} p(c_i, c'_j) \cdot \log_2 \frac{p(c_i, c'_j)}{p(c_i)p(c'_j)} \quad (32)$$

其中, $p(c_i)$ 和 $p(c'_j)$ 表示在所有测试图像中被分到各自对应聚类中心的概率, 在这里我们用数量比率代表概率, $p(c_i, c'_j)$ 表示一个图像被同时分到这两个聚类中心 c_i 和 c'_j 的概率, 同样我们用数量比代替概率 (注意在这里聚类算法计算得到的聚类中心的编号可以为任意的顺序). 在实验中, 第一个聚类中心集为图像的已知类别标签, 同一个标签的图像为一个聚类中心, 第二个聚类中心集为聚类算法得到的标签, 标签可以为任意的顺序标签. 归一化的互信息为互信息除以最大的聚类标签熵, 其中 $H(\mathcal{C}) = -\sum_i p(c_i) \log_2 p(c_i)$, $p(c_i)$ 指属于聚类中心 c_i 的图像数量比率.

$$NMI(\mathcal{C}, \mathcal{C}') = \frac{MI(\mathcal{C}, \mathcal{C}')}{\max(H(\mathcal{C}), H(\mathcal{C}'))} \quad (33)$$

3.3 实验结果与分析

3.3.1 聚类的准确性验证

在本文提出的方法需要计算各个数据之间的局部结构关系, 我们采用高斯核函数值作为权值, 即 $W_{ij} = e^{-\frac{\|x_i - x_j\|^2}{\sigma}}$. 变量 $k = 5$, 这种参数设置在非大量样本数据库中使用的较多. 在式 (12) 中所有的变量我们设定为 $\lambda_f = 0.01$ ($f = 1, \dots, G$), $\mu = 10$, 与算法 MultiNMF 和 MultiGNMF 保持一致.

在定义了参数之后, 计算每个视图图像之间的亲和矩阵, 每种方法我们都运行 20 次, 取所有次数运行结果的平均值和方差作为最终的算法效果评估值. 经过 20 次的实验运行, 我们统计了各种算法在 4 个数据库上的 AC 和 NMI 值, 如表 2 和表 3 所示.

在表 2 和表 3 中我们可以看到本文提出的算法在聚类分析中相比较其他多视图学习算法在准确度和归一化互信息两个指标下都有较好的表现. 同时我们注意到 Co-reguSC 和 SC-ML 算法也采用了局部结构约束的条件 (谱聚类算法是一种基于数据样本图结构的算法), MMSC 也采用了一致性约束, 本文算法同样超过该类算法. 在比较本文提出的算法 MultiGSemiNMF 和 MultiGNMF 算法中, 我们发现基于 MultiGSemiNMF 的算法相对于 MultiGNMF 算法聚类效果上也有提升, 即使所有特征并没有负数或零, 这证明在特征学习中 MultiGSemiNMF 算法比 MultiGNMF 算法在一些应用场景中具有更好的表现, 且 MultiGSemiNMF

消除了 MultiGNMF 算法要求所有特征非负的限制, 因此也有着更为广阔的应用平台.

表 2 不同方法在 4 个数据库中的 AC 值
Table 2 The AC values by different methods in four databases

算法	UCI Digit	CMU PIE	3-Sources	ORL
BSV	68.5 ± .05	55.2 ± .02	60.8 ± .01	51.8 ± .02
WSV	63.4 ± .04	47.6 ± .01	49.1 ± .03	42.8 ± .05
ConcatNMF	67.8 ± .06	51.5 ± .00	58.6 ± .03	66.8 ± .02
ColNMF	66.0 ± .05	56.3 ± .00	61.3 ± .02	66.3 ± .04
Co-reguSC	86.6 ± .00	59.5 ± .02	47.8 ± .01	70.5 ± .02
MultiNMF	88.1 ± .01	64.8 ± .02	68.4 ± .06	67.3 ± .03
SC-ML	88.1 ± .00	72.3 ± .00	54.0 ± .00	70.8 ± .00
MMSC	89.4 ± .38	61.6 ± .06	69.9 ± .04	70.7 ± .06
AMGL	82.6 ± .74	63.6 ± .19	68.3 ± .12	64.2 ± .10
MultiGNMF	95.1 ± .10	72.5 ± .02	73.4 ± .00	73.5 ± .00
MultiGSemiNMF	95.6 ± .02	77.2 ± .12	79.3 ± .00	76.2 ± .00

表 3 不同方法在 4 个数据库中的 NMI 值
Table 3 The NMI values by different methods in four databases

算法	UCI Digit	CMU PIE	3-Sources	ORL
BSV	63.4 ± .03	74.1 ± .00	53.0 ± .01	69.8 ± .01
WSV	60.3 ± .03	69.1 ± .02	44.1 ± .02	65.4 ± .04
ConcatNMF	60.3 ± .03	70.5 ± .00	51.7 ± .03	85.1 ± .01
ColNMF	62.1 ± .03	68.3 ± .00	55.2 ± .02	84.5 ± .03
Co-reguSC	77.0 ± .00	80.5 ± .01	41.4 ± .01	80.5 ± .01
MultiNMF	80.4 ± .01	82.2 ± .02	60.2 ± .06	87.6 ± .01
SC-ML	87.6 ± .00	85.1 ± .00	45.5 ± .00	85.3 ± .00
MMSC	88.0 ± .05	80.3 ± .01	66.3 ± .09	85.9 ± .01
AMGL	84.9 ± .33	81.8 ± .10	56.9 ± .11	81.9 ± .02
MultiGNMF	90.1 ± .04	90.2 ± .01	65.4 ± .00	88.5 ± .00
MultiGSemiNMF	91.2 ± .01	93.8 ± .05	73.4 ± .00	89.9 ± .00

3.3.2 参数对算法性能的影响

在式 (12) 中有 G 个参数变量 λ_f , 它的物理意义是不同视图对学习过程的重要性进行权衡. 如果我们具有一些先验知识, 带有噪声的视图特征可以适当减少权重, 被认为重要的视图如文本标签信息等可以适当增加权重. 在我们的实验中没有先验知识, 同时数据已经归一化处理, 所以所有的 λ_f 定义为相同的数. 但是 λ_f 的大小会影响聚类效果, λ_f 越小则对系数矩阵的一致性约束越松弛, 反之, λ_f 为无穷大的话所有系数矩阵为相同的数值.

下面我们在两个数据库 CMU PIE 和 UCI Digit 上用 MultiNMF、MultiGNMF 和 MultiSemiNMF 三个算法做聚类分析, 准确率 AC 作为指标度量不同的 λ_f 值对实验的影响. 我们设定 λ_f 为

0.001, 0.005, 0.002, 0.01, 0.02, 0.05, 0.1, 图 1 和图 2 描述了对应的聚类效果.

从图 1 和图 2 中我们可以看到, 在 MultiNMF、MultiGNMF 和 MultiSemiNMF 三个算法中, MultiGSemiNMF 算法的准确率最高, 且受参数影响很小, 在较小或者较大的值均能有稳定的表现. MultiGNMF 算法最不稳定, 但是在大部分实验中仍能超过基准算法 MultiNMF. 三种算法均在 $\lambda = 0.01$ 时有较高的表现.

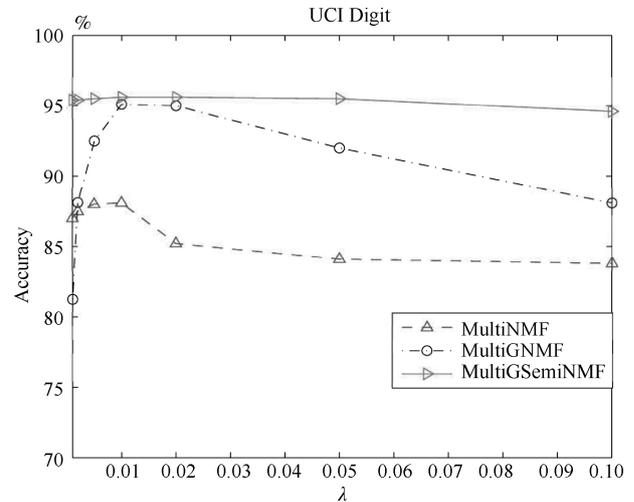


图 1 在 UCI Digit 数据库中参数 λ 对本文算法的影响
Fig. 1 The influences of λ on UCI Digit database

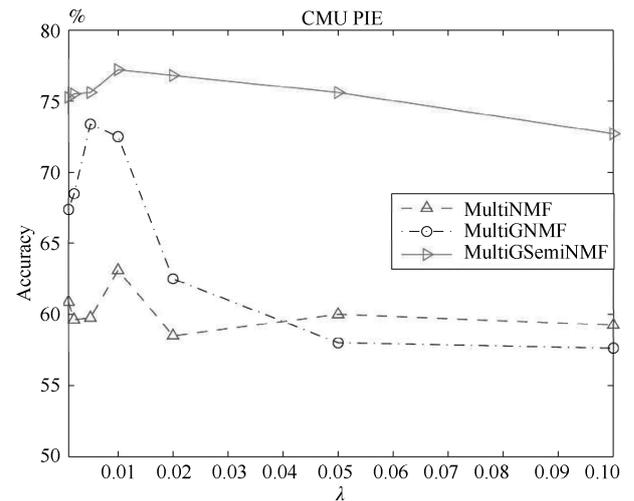


图 2 在 CMU PIE 数据库中参数 λ 对本文算法的影响
Fig. 2 The influences of λ on CMU PIE database

本文提出的算法 MultiGSemiNMF 和 MultiGNMF 需要构建一个样本亲和矩阵, 在损失函数其中有一个参数 k , k 值越大样本的局部结构约束越多, 反之局部结构约束越少. 不同的 k 值对聚类效果有一定影响, 在与其他算法的比较中, 我们采用其他算法一样的 k 值, 在这里我们另外分析 k 值对本

文提出算法的影响. 在 UCI Digit 数据总库, 我们选取 $k = 5, 10, 15, 20$ 构建亲和矩阵 ($\lambda = 0.01$), 用聚类的准确率 AC 来评估算法效果, 如图 3 所示.

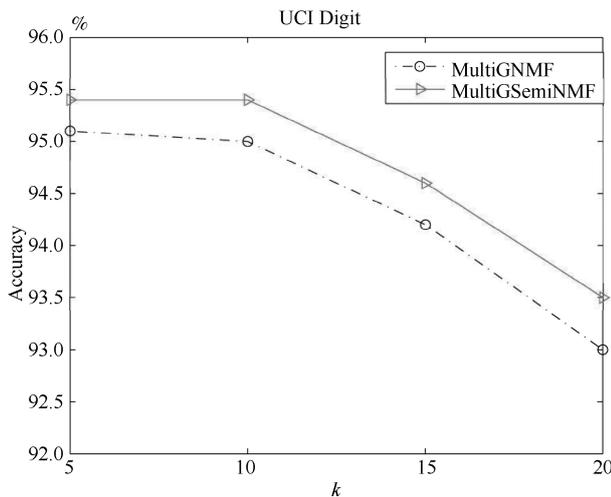


图 3 在 UCI Digit 中参数 k 对本文提出算法的影响

Fig. 3 The influences of k on UCI Digit database

从图 3 我们看到参数 k 对本文算法有一定的影响, k 值越大虽然能更充分地保留样本之间的局部空间结构关系, 但是从图中的结果来看聚类准确度随着 k 的变大有降低的趋势. 过度的保留空间结构关系并不能提升算法的效果, 反而可能产生过度拟合的副作用, 因为样本的结构关系是基于原始特征的距离计算得到, 而原始特征并不能充分描述样本信息, 其中或多或少含有冗余和噪声. 同时, 也可以看出, 无论参数 k 取何值时, MultiGSemiNMF 算法的性能都优于 MultiGNMF 算法.

4 总结

本文提出了两种多视图学习的方法: MultiGNMF 算法和 MultiGSemiNMF 算法. MultiGNMF 和 MultiGSemiNMF 算法都是基于样本局部结构空间约束的非负矩阵分解多视图学习方法.

本文首先介绍了一种单视图学习方法: NMF 矩阵分解. 然后, NMF 算法的基础之上, 在以往多视图学习的框架准则下, 本文提出了基于样本局部结构空间约束的非负矩阵分解多视图学习方法 MultiGNMF. 但是, MultiGNMF 方法只适用于非负的特征矩阵. MultiGSemiNMF 算法则不限于此.

为了验证本文提出的多视图学习算法的性能, 我们在公有的图像数据库中做聚类分析. 实验中和以往的算法比较, 实验结果表明本文提出的算法相对于其他基于矩阵分解的多视图学习方法有更好的表现. 同时实验中分析了算法中的参数变化对算法性能的影响, 实验结果表明 MultiGSemiNMF 对参

数变化具有很好的鲁棒性. 在未来的工作中, 我们将探索一种新的基于局部结构约束的多视图学习方法.

References

- 1 Lee D D, Seung S. Learning the parts of objects by non-negative matrix factorization. *Nature*, 1999, **401**(6755): 788–791
- 2 Gausier E, Goutte C. Relation between PLSA and NMF and implications. In: Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. Salvador, Brazil: ACM, 2005. 601–602
- 3 Huang J, Nie F P, Huang H, Ding C. Robust manifold non-negative matrix factorization. *ACM Transactions on Knowledge Discovery from Data*, 2014, **8**(3): Article No. 11
- 4 Zhao S C, Yao H X, Yang Y. Affective image retrieval via multi-graph learning. In: Proceedings of the 22nd ACM International Conference on Multimedia. Orlando, Florida, USA: ACM, 2014. 1025–1028
- 5 Zheng L, Yang Y, Tian Q. SIFT meets CNN: a decade survey of instance retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, DOI: 10.1109/TPAMI.2017.2709749
- 6 Zheng L, Wang S J, Wang J J, Tian Q. Accurate image search with multi-scale contextual evidences. *International Journal of Computer Vision*, 2016, **120**(1): 1–13
- 7 Cao X C, Zhang C Q, Fu H Z, Liu S, Zhang H. Diversity-induced multi-view subspace clustering. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, Massachusetts, USA: IEEE, 2015. 586–594
- 8 Guo D Y, Zhang J, Liu X W, Cui Y, Zhao C X. Multiple kernel learning based multi-view spectral clustering. In: Proceedings of the 22nd International Conference on Pattern Recognition. Columbus, Ohio, USA: IEEE, 2014. 3774–3779
- 9 Wang H X, Weng C Q, Yuan J S. Multi-feature spectral clustering with minimax optimization. In: Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, Ohio, USA: IEEE, 2014. 4106–4113
- 10 Cai X, Nie F P, Huang H, Kamangar F. Heterogeneous image feature integration via multi-modal spectral clustering. In: Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition. Colorado, USA: IEEE, 2011. 1977–1984
- 11 Nie F P, Li J, Li X L. Parameter-free auto-weighted multiple graph learning: a framework for multiview clustering and semi-supervised classification. In: Proceedings of the 25th International Joint Conference on Artificial Intelligence. New York, USA: AAAI, 2016. 1881–1887
- 12 Cai X, Nie F P, Huang H. Multi-view K-means clustering on big data. In: Proceedings of the 23rd International Joint Conference on Artificial Intelligence. New York, USA: AAAI, 2013. 2598–2604.
- 13 Xu J L, Han J W, Nie F P. Discriminatively embedded k-means for multi-view clustering. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, Nevada, USA: IEEE, 2016. 5356–5364

- 14 Kumar A, Rai P, Daume III H. Co-regularized multi-view spectral clustering. In: Proceedings of the 24th International Conference on Neural Information Processing Systems. Granada, Spain: Curran Associates Inc., 2011. 1413–1421
- 15 Kong S, Wang X K, Wang D H, Wu F. Multiple feature fusion for face recognition. In: Proceedings of the 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition. Shanghai, China: IEEE, 2013. 1–7
- 16 Akata Z, Thureau C, Bauckhage C. Non-negative matrix factorization in multimodality data for segmentation and label prediction. In: Proceedings of the 16th Computer Vision Winter Workshop. Mitterberg, Austria: Verlag der Technischen Universität Graz, 2011.
- 17 Liu J L, Wang C, Gao J, Han J W. Multi-view clustering via joint nonnegative matrix factorization. In: Proceedings of the 2013 SIAM International Conference on Data Mining. Texas, USA: SIAM, 2013. 252–260
- 18 He X F, Yan S C, Hu Y X, Niyogi P, Zhang H J. Face recognition using Laplacianfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, **27**(3): 328–340
- 19 Cai D, He X F, Han J W, Huang T S. Graph regularized non-negative matrix factorization for data representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, **33**(8): 1548–1560
- 20 Nie F P, Wang X Q, Jordan M I, Huang H. The constrained Laplacian rank algorithm for graph-based clustering. In: Proceedings of the 30th AAAI Conference on Artificial Intelligence. Phoenix, Arizona, USA: AAAI, 2016. 1969–1976
- 21 Li S Z, Hou X W, Zhang H J, Cheng Q S. Learning spatially localized, parts-based representation. In: Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Kauai, HI, USA: IEEE, 2001, 1: I-207–I-212
- 22 Ding C H Q, Li T, Jordan M I. Convex and semi-nonnegative matrix factorizations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, **32**(1): 45–55
- 23 Kumar A, Rai P, Daume III H. Co-regularized multi-view spectral clustering. In: Proceedings of the 24th International Conference on Neural Information Processing Systems. Granada, Spain: Curran Associates Inc., 2011. 1413–1421
- 24 Dong X W, Frossard P, Vandergheynst P, Nefedov N. Clustering on multi-layer graphs via subspace analysis on grassmann manifolds. *IEEE Transactions on Signal Processing*, 2014, **62**(4): 905–918
- 25 Lovász L, Plummer M D. *Matching Theory*. Providence, R.I.: American Mathematical Society, 2009.

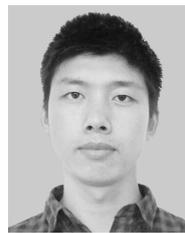


张祎 大连理工大学硕士研究生。主要研究方向为多媒体信息安全和计算机视觉。E-mail: yiz@mail.dlut.edu.cn
(ZHANG Yi Master student at the School of Information and Communication Engineering, Dalian University of Technology. Her research interest covers multimedia information security and computer vision.)



孔祥维 浙江大学数据科学与管理工程系教授。2003 年获得管理科学与工程专业博士学位。2006~2007 年美国普渡大学访问学者。主要研究方向为数字图像处理, 图像检索, 计算机视觉, 多媒体信息安全和数字媒体取证。本文通信作者。E-mail: kongxiangwei@zju.edu.cn

(KONG Xiang-Wei Professor in the Department of Data Science and Engineering Management, Zhejiang University, China. She received her Ph.D. degree in management science and engineering, Dalian University of Technology, China, in 2003. From 2006 to 2007, she was a visiting researcher in the Department of Computer Science at Purdue University, USA. Her research interest covers digital image processing, image retrieval, computer vision, multimedia information security, and digital media forensics. Corresponding author of this paper.)



王振帆 大连理工大学硕士研究生。主要研究方向为计算机视觉。E-mail: zfwang@mail.dlut.edu.cn

(WANG Zhen-Fan Master student at the School of Information and Communication Engineering, Dalian University of Technology. His main research interest is computer vision.)



付海燕 大连理工大学信息与通信工程学院副教授。2014 年获得大连理工大学信息与信号处理博士学位。主要研究方向为图像检索和计算机视觉。E-mail: fuhy@dlut.edu.cn

(FU Hai-Yan Associate professor at the School of Information and Communication Engineering, Dalian University of Technology. She received her Ph.D. from Dalian University of Technology in 2014. Her research interest covers image retrieval and computer vision.)



李明 大连理工大学信息与通信工程学院副教授。2010 年获得纽约州立大学电子工程系博士学位。主要研究方向为通信理论和信号处理。E-mail: mli@dlut.edu.cn

(LI Ming Associate professor at the School of Information and Communication Engineering, Dalian University of Technology, China. He received his Ph.D. in electrical engineering from the State University of New York, USA in 2010. His research interest covers communication theory and signal processing.)