

基于局部图拉普拉斯约束的鲁棒低秩表示聚类方法

李波¹ 卢春园¹ 冷成财¹ 金连宝¹

摘要 针对传统低秩表示聚类方法存在的稀疏性不足及噪声敏感等问题,提出了一种基于局部图拉普拉斯约束的鲁棒低秩表示聚类模型.一方面,通过加入图像数据局部相似性的约束,在保持表示矩阵分块对角的特性下,增强了其稀疏性;另一方面,从数据相关性的角度分析了低秩表示模型的聚类性质,通过采用鲁棒低秩表示模型,不仅降低了噪声的干扰,而且减弱了表示字典数据之间的线性相关性,从理论上保证了最终的邻接矩阵具有分块对角的良好聚类性质.与传统低秩表示方法相比,本文得到的表示矩阵既保证了分块性质,又更加稀疏,仿真实验结果表明聚类效果有明显提升.

关键词 低秩聚类, 数据聚类, 图拉普拉斯, 鲁棒主成分分析

引用格式 李波, 卢春园, 冷成财, 金连宝. 基于局部图拉普拉斯约束的鲁棒低秩表示聚类方法. 自动化学报, 2015, 41(11): 1971–1980

DOI 10.16383/j.aas.2015.c150031

Robust Low Rank Subspace Clustering Based on Local Graph Laplace Constraint

LI Bo¹ LU Chun-Yuan¹ LENG Cheng-Cai¹ JIN Lian-Bao¹

Abstract Low rank clustering is one of the state-of-art subspace clustering algorithm, but it suffers from dense adjacency map and noise. In this paper, we propose a robust low rank clustering algorithm based on the local graph laplace constraint, which enhances the sparsity of the adjacency matrix while maintain the clustering characteristic; on the other hand, we analyze the mechanics of clustering from the view of incoherence, and argue that the robust model proposed in this paper not only reduce the noise level, but also lower the coherence between data. Finally, experimental results show that our algorithm is more robust and more effective.

Key words Low rank clustering, data clustering, graph Laplace, robust principle component analysis (RPCA)

Citation Li Bo, Lu Chun-Yuan, Leng Cheng-Cai, Jin Lian-Bao. Robust low rank subspace clustering based on local graph Laplace constraint. *Acta Automatica Sinica*, 2015, 41(11): 1971–1980

随着信息科技的发展,人们每天所接触到和需要处理的图像数量呈现几何级的增长,但是海量的数据在带给人们各种便利的同时,也带来了极大的问题,人们容易迷失在浩瀚的数据中而难以找到自己真正需要的信息.面对如此大规模的图像资源,如何对它们进行有效组织,成为一个迫切需要解决的问题.图像聚类的目标就是要根据图像的特征或空间属性自动地将原本属于同一类的图像组织在一起.近年来,随着计算机视觉及机器学习技术的发展,大量的图像聚类算法被提出和研究,根据聚类的机理,

这些方法总体上可以分为两类:基于图像特征的局部聚类方法和基于图像空间的全局聚类方法.

基于局部的聚类算法^[1–3]首先需要提取图像的局部特征.常用的特征既有例如颜色、纹理及形状等的视觉特征,也有如尺度不变特征变换 (Scale-invariant feature transform, SIFT)^[4–5]、方向梯度直方图 (Histogram of oriented gradient, HOG)^[6–8]等构造式尺度旋转不变的特征.然后利用这些特征来描述图像内容,并进行图像聚类.局部算法充分利用了图像的几何结构和纹理信息,但随着聚类图像规模的不断增大,每幅图像一般需要采样几千到几万个特征,对存储空间与计算时间提出了更高的要求.另外,基于局部特征的聚类算法忽略了图像空间的全局属性,对噪声和特征的选取比较敏感.

与基于局部特征的聚类方法相比,全局聚类方法更侧重于图像空间的整体表示,对噪声更加鲁棒.基于图像空间的全局聚类方法假设图像数据集采样于若干互不相交的子空间中,每一个子空间代表着一

收稿日期 2015-02-04 录用日期 2015-07-16
Manuscript received February 4, 2015; accepted July 16, 2015
国家自然科学基金 (61262050, 61562062, 61363048, 61363049), 江西省自然科学基金 (20151BAB211006) 资助
Supported by National Natural Science Foundation of China (61262050, 61562062, 61363048, 61363049) and Natural Science Foundation of Jiangxi Province (20151BAB211006)
本文责任编辑 王亮
Recommended by Associate Editor WANG Liang
1. 南昌航空大学数学与信息科学学院 南昌 330063
1. School of Mathematics and Information Science, Nanchang Hangkong University, Nanchang 330063

类图像, 只要图像数据足够多到可以张成每个子空间, 就可以根据图像的空间属性自动地将原本属于同一子空间的图像组织在一起. 根据上述线性空间的假设, 每幅图像均可以表示为整个图像集合的线性组合, 即每幅图像对应着一个表示系数向量. 基于空间的图像聚类算法就是利用图像表示系数来描述一幅图像, 通过对表示系数施加不同的约束, 就可以得到不同的表示. 常见的全局聚类方法包括稀疏表示聚类 (Sparse representation clustering, SRC)^[9-11] 和低秩表示子空间聚类 (Low rank representation clustering, LRR)^[12-14]. 稀疏表示聚类算法是压缩感知理论的一个重要应用. 它考虑每一幅图像仅可以由图像集合的少数元素线性表示, 即最小化表示系数向量的 l_0 范数, 同时还需要满足全局线性表示的约束, 从而达到挖掘图像子空间结构的目标. 与局部聚类方法相比, 稀疏表示聚类算法对噪声更加鲁棒, 聚类效果也有了明显的提升. 但稀疏表示聚类方法仍存在很多问题. 一方面, 对每一幅图像该方法都需要重新计算一次 l_0 范数最小化问题, 当图像数据的规模比较大时, 该方法的计算复杂度和运行时间急速增长; 另一方面, 根据上述分析, 稀疏表示聚类方法是对每个图像数据进行独立的分析, 忽略了数据之间或图像子空间之间的结构属性, 如当来自同一个子空间 (同一类图像) 的数据具有较强的相关性或聚类性质时, l_0 稀疏表示方法最终仅会随机选取其中一个数据, 而其他数据的表示系数均为 0, 最终虽然得到了一个非常稀疏的表示, 却忽略了数据之间的相关性, 最终导致聚类性能的降低 (见图 1(a) 所示). 与稀疏表示聚类相比, 低秩表示子空间聚类方法通过采用全局表示矩阵的秩来作为约束, 有效地利用了数据之间的相关性及聚类性质, 最终得到的表示矩阵具有良好的分块对角的性质, 即来自同一子空间的图像他们的非零表示系数构成一个小的分块, 反之, 表示系数趋近于零 (如图 1(b) 所示). 低秩表示聚类方法虽然具有良好的分块聚类性质, 但只关注全局空间秩的约束, 却忽略了图像数据的局部性质, 使得最终的表示矩阵缺乏稀疏性, 如图 1(b), 在满足分块对角性质的同时, 仍存在大量稠密的非零系数, 这对聚类的效果都会产生影响. 另一方面, 在图像数据存在噪声的情况下, 低秩表示模型仍采用噪声图像集合作为表示字典, 这不可避免地会损害数据空间的内在结构.

基于上述分析, 本文针对传统低秩表示聚类方法存在的稀疏性不足及噪声敏感等问题, 提出了一种基于局部图拉普拉斯约束的鲁棒低秩表示聚类模型. 一方面, 通过加入图像数据局部相似性的约束, 在保持表示矩阵分块对角的特性下, 增强了其稀疏

性; 另一方面, 针对噪声干扰, 本文采用了一种鲁棒的低秩表示模型, 并从另外一个新的角度对该模型进行了分析, 进一步揭示了该模型的聚类性质. 本文实验结果如图 1(c) 所示, 与传统低秩表示方法相比, 本文得到的表示矩阵更加稀疏, 分块效果更加明显.

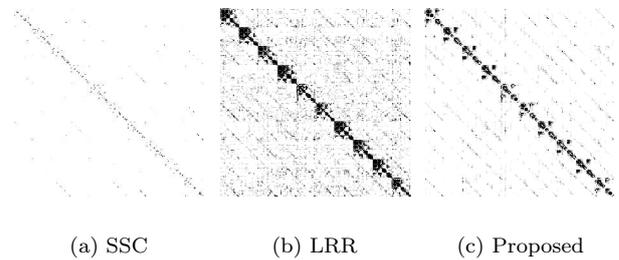


图 1 邻接矩阵

Fig. 1 The visualization of affinity matrix

本文主要分为五个部分: 第 1 节简要介绍了稀疏表示聚类模型和低秩表示聚类模型; 第 2 节是本文的主要工作, 提出了基于图拉普拉斯约束的鲁棒低秩表示聚类模型, 并对该模型进行了理论分析; 第 3 节针对本文提出模型建立了一种基于交替方向乘子方法的快速数值算法; 第 4 节是仿真实验及结果; 第 5 节对本文进行了总结.

1 预备知识

给定一组待聚类的图像数据, 首先将每一幅图像按照列排成一个列向量, 并作为矩阵 X 的列, 即 $X = [X_1, X_2, \dots, X_n] \in \mathbf{R}^{m \times n}$, $i = 1, 2, \dots, n$, 其中 m 表示每幅图像包含的像素数量, n 表示该组数据包含图像的数目. 我们假设所有图像数据都是采样于 C 个相互正交的子空间, 如 $[X_{i1}, X_{i2}, \dots, X_{ik}]$ 属于第 C_i 类图像所属的子空间 \mathcal{S}_i . 图像聚类的目标就是根据各种性质尽可能准确地把每幅图像 X_i 划分到它应属的类别 \mathcal{S}_i .

1.1 稀疏子空间聚类

稀疏子空间聚类^[9-10] 包含两步, 首先针对每一个图像数据 X_i , 通过求解稀疏表示问题获取其稀疏系数向量 Z_i :

$$\begin{aligned} \min_{Z_i} \|Z_i\|_0 \\ \text{s.t. } X_i = X^\dagger Z_i \end{aligned}$$

其中, $X^\dagger = [X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n]$. 但关于 l_0 范数最小化的稀疏求解问题是 NP-hard 的, 根据压缩感知理论的结论, 在满足一定条件下, 上述 l_0 优化问题可以等价于下述凸的 l_1 最小化问题:

$$\begin{aligned} & \min_{Z_i} \|Z_i\|_1 \\ & \text{s.t. } X_i = X^\dagger Z_i \end{aligned}$$

其次, 对每一幅图像 $X_i, i = 1, 2, \dots, n$ 求解出其稀疏系数 Z_i 后, 将这些系数向量按列排成邻接矩阵 $Z = [Z_1, Z_2, \dots, Z_n]$, 为满足非负性和对称性, 取 $Z' = |Z| + |Z^T|$, 其中 $|\cdot|$ 表示逐项取绝对值. 最后根据图的邻接矩阵 Z' 采用谱分割 (如 Normal cut^[15] 等) 的方法实现图像的聚类. 稀疏表示聚类方法是对每个图像数据进行独立的分析, 忽略了数据之间或图像子空间之间的结构属性.

1.2 低秩表示问题

与稀疏表示聚类方法相比, 低秩表示聚类^[12-13, 16] 方法更加全局, 充分考虑了图像数据之间的相互关系及结构属性. 假设图像数据采样于高维空间中的多个相互正交的子空间, 根据线性空间及线性表示模型的性质, 表示矩阵的秩描述了子空间的数目, 因此通过最小化该矩阵的秩就可以发掘数据空间的流形结构. 具体地, 低秩表示模型可以表示为

$$\begin{aligned} & \min_Z \text{rank}(Z) \\ & \text{s.t. } X = XZ \end{aligned}$$

由于秩是一个非凸的函数, 而且没有闭式的表达形式, 其优化问题非常困难. 相关研究表明, 矩阵的核范数, 即所有奇异值之和, 是矩阵秩的一个最小凸包络, 因此, 在满足一定条件的情况下, 关于秩的优化问题可以等价于关于下面核范数的凸优化问题:

$$\begin{aligned} & \min_Z \|Z\|_* \\ & \text{s.t. } X = XZ \end{aligned}$$

针对数据包含噪声的情况, 可以通过加入误差项 E 来增强鲁棒性.

$$\begin{aligned} & \min_{Z, E} \|Z\|_* + \lambda \|E\|_{2,1} \\ & \text{s.t. } X = XZ + E \end{aligned}$$

其中, $\|E\|_{2,1} = \sum_i \|E_i\|_2$, E_i 对应着矩阵 E 的第 i 列. 虽然通过对噪声的约束提升了鲁棒性, 但该模型仍然不合理, 因为直接采用带噪声的数据 X 本身作为表示字典会不可避免的破坏数据空间的结构. 因此, Wei 等^[17] 提出了一个改进的鲁棒低秩表示模型 (Minimal squared Frobenius norm representation, MSFNR), 首先对含噪声的原始数据 X 进行去噪,

得到一个相对干净的数据 D 作为字典, 通过 D 的自表示矩阵来挖掘内在的空间结构.

$$\begin{aligned} & \min_{D, Z, E} \|Z\|_* + \lambda \|E\|_{2,1} \\ & \text{s.t. } X = D + E, D = DZ \end{aligned} \quad (1)$$

通过采用去噪后的数据作为表示字典, 该鲁棒低秩表示模型的聚类效果取得了明显改善. 但去噪并不是提升低秩表示模型聚类效果的唯一原因, 本文将在下一部分进一步分析该模型影响聚类效果的另一个重要原因.

2 局部图拉普拉斯约束的鲁棒低秩表示聚类算法

通过上述分析, 如图 1(b) 所示, 低秩表示聚类作为一种全局聚类方法, 在成功获取邻接矩阵分块对角性质的同时, 忽略了数据的局部相关性约束, 导致产生大量稠密的非零元素, 另一方面, 噪声和离群点对低秩表示聚类也会产生巨大影响. 针对上述两个问题, 本文提出了一种新的基于局部图拉普拉斯约束的低秩表示聚类算法, 首先通过拉普拉斯约束增强了邻接矩阵的稀疏性, 同时通过降低字典表示矩阵的噪声和相关性, 有效提高了聚类的性能和鲁棒性.

如果我们将每幅图像数据 X_i 理解为高维空间中的一个点, 并在任意两个点 X_i 和 X_j 之间建立一条边, 同时为每条边赋予一个权重 ω_{ij} , 这样就形成了一个由所有数据构成的完全无向图 G . 其中, 每条边的权重 ω_{ij} 描述了两幅图像 X_i 与 X_j 之间的相似性, 通常可采用关于欧氏距离的高斯函数来表示, 本文取 $\omega_{ij} = \exp^{-\|X_i - X_j\|^2 / \sigma^2}$. 由所有边的权重组成的矩阵记为 $W = \{\omega_{ij}\}$.

具体地, 本文提出的局部图拉普拉斯约束的鲁棒低秩表示聚类算法可以表示为

$$\begin{aligned} & \min_{D, Z, E} \text{rank}(Z) + \lambda \text{tr}(ZLZ^T) + \mu \|E\|_{2,1} \\ & \text{s.t. } X = D + E, D = DZ \end{aligned} \quad (2)$$

第 1 项 $\text{rank}(Z)$ 表示矩阵 Z 的秩, 通过对秩进行最小化可以充分挖掘数据空间的全局结构, 低秩表示理论表明当数据采样充分并没有噪声干扰的条件下, 最小化矩阵的秩得到的最优解 Z^* 具有分块对角的性质. 但 $\text{rank}(Z)$ 是非凸的, 本文采用 Z 的核范数 $\|Z\|_*$ 作为秩的凸逼近, 从而将非凸的问题 (2) 转换为一个等价的凸优化问题 (3).

$$\begin{aligned} \min_{D, Z, E} \|Z\|_* + \lambda \text{tr}(ZLZ^T) + \mu \|E\|_{2,1} \\ \text{s.t. } X = D + E, D = DZ \end{aligned} \quad (3)$$

第 2 项是局部图拉普拉斯约束函数, 其中矩阵 L 为图 G 的拉普拉斯矩阵, $L = D - W$, 其中 $D = \text{diag}(d)$, $d_i = \sum_j \omega_{ij}$, $i = 1, 2, \dots, n$. 根据拉普拉斯矩阵的定义, 该约束项可以重新表示为

$$\text{tr}(ZLZ^T) = \sum_i \sum_j \omega_{ij} \|Z_i - Z_j\|_2^2$$

其中, $\|\cdot\|_2$ 表示向量的 2-范数. 根据权重 ω_{ij} 的定义知, 两幅图像 X_i 与 X_j 距离越近, ω_{ij} 的取值越大. 因此, 要使得该约束项取得最小, 就要求两幅图像在表示空间中的系数 Z_i 与 Z_j 越接近, 即当 $\omega_{ij} \rightarrow \infty$ 时, $\|Z_i - Z_j\|_2 \rightarrow 0$. 局部图拉普拉斯项使得在原始数据空间中距离比较近的图像在表示空间中的表示系数向量仍然比较接近, 增强了传统低秩表示模型的局部性和稀疏性. λ 平衡了数据整体的秩约束与局部图拉普拉斯约束.

根据低秩表示模型的已有知识^[17], 无噪声的模型 $\min_Z \|Z\|_*$, s.t. $X = XZ$ 具有闭形式的解 $Z^* = VV^T$, 其中 V 为数据矩阵 X 的右奇异向量, 即 $[U, \Sigma, V] = \text{SVD}(X)$, 并且该问题的最小值恰为矩阵 X 的秩 $\text{rank}(X)$. 根据上述性质, 本文模型 (3) 可以拆分为两个子问题来解决.

定理 1. 假设 (D^*, Z^*, E^*) 是下述两个子问题的解:

$$\min_{D, E} \text{rank}(D) + \mu \|E\|_{2,1} \quad \text{s.t. } X = D + E \quad (4)$$

$$\min_Z \|Z\|_* + \lambda \text{tr}(ZLZ^T) \quad \text{s.t. } D = DZ \quad (5)$$

那么 (D^*, Z^*, E^*) 也是模型 (3) 的最优解.

定理 1 不仅使得原来的模型 (3) 计算变得简单, 而且更加容易揭示出该模型对效果提升的原因. 第一个子问题 (4) 可以看作是列稀疏约束的鲁棒主成分分析 (Robust principle component analysis, RPCA)^[18]. 正如鲁棒低秩表示模型的分析, 这一步 RPCA 的处理可以去除噪声, 利用较准确的数据 D 来作为低秩表示的字典可以提高对噪声的鲁棒性. 本文将从另外一个角度对子问题 (4) 进行分析. Candés 等^[18] 在讨论 RPCA 模型的理论性质时, 指出如果数据 X 采样于同一个子空间, RPCA 可以准确恢复出一个低秩矩阵 D 和一个稀疏误差矩阵 E 的充分条件是数据矩阵 D 要同时具有低秩结构和较强的不相关性; 低秩表示模型 LRR 将数据的采样空间进行了推广, 假设数据 X 采样于若干互不相交的子空间, 这时 LRR 如果要恢复出一个准确的分块

对角矩阵就必须满足字典表示矩阵 D 对应不同子空间的列向量具有较强的不相关性. 根据定理 1, LRR 可以分解为一个 RPCA 过程和一个局部约束的 LRR 问题, RPCA 子问题 (4) 求得一个关于原始数据 X 的低秩表示 D , 一方面降低了噪声干扰, 另一方面, 或许更重要的是, 低秩的约束使得子问题 (5) 的字典表示矩阵 D 符合了上述分析条件, 即使得 D 对应不同子空间的列向量具有了较强的不相关性^[19], 从而从本质上保证了最终的邻接矩阵 Z 具有分块对角的良好聚类性质. 本文将在仿真实验中给出相关实验比较与验证.

定理 2. 假设待聚类数据 X 充分采样于 k 个相互正交的子空间 \mathcal{S}_i , $i = 1, 2, \dots, k$, 且 (D^*, Z^*, E^*) 是本文模型的最优解:

$$\min_{D, Z, E} \text{rank}(Z) + \lambda \text{tr}(ZLZ^T) + \mu \|E\|_{2,1}$$

$$\text{s.t. } X = D + E, D = DZ$$

那么邻接矩阵 Z^* 将具有分块对角的结构

$$\begin{vmatrix} Z_1 & 0 & 0 & \cdots & 0 \\ 0 & Z_2 & 0 & \cdots & 0 \\ 0 & 0 & Z_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & Z_k \end{vmatrix}$$

证明. 假设 Z^* 是本文模型的最优解, 定义

$$\tilde{Z}_{ij} = \begin{cases} Z_{ij}^*, & \text{如果 } X_i \text{ 和 } X_j \text{ 属于同一个子空间} \\ 0, & \text{否则} \end{cases}$$

根据子空间正交性的假设, \tilde{Z} 也是一个可行解. 令 $C = Z^* - \tilde{Z} \succeq 0$, 即 $Z^* = \tilde{Z} + C$. 由于 $C \succeq 0$, 易得 $\text{tr}(\tilde{Z}L\tilde{Z}^T) \leq \text{tr}(Z^*LZ^{*T}) = \text{tr}((\tilde{Z} + C)L(\tilde{Z}^T + C))$. 同时, 根据范数性质知, $\|\tilde{Z}\|_* \leq \|Z^*\|_*$, 因此 \tilde{Z} 也是该模型的一个最优解. 但由于本文模型是严格凸的, 即仅有唯一的最优解, 因此, $\|Z^*\|_* + \text{tr}(Z^*LZ^{*T}) = \|\tilde{Z}\|_* + \text{tr}(\tilde{Z}L\tilde{Z}^T)$. 从而, $\text{tr}(\tilde{Z}L\tilde{Z}^T) \geq \text{tr}(Z^*LZ^{*T})$. 综合上述分析, 可得 $\text{tr}(\tilde{Z}L\tilde{Z}^T) = \text{tr}(Z^*LZ^{*T})$, $C = 0$, 从而 $Z^* = \tilde{Z}$, 即具有相同的分块对角结构. \square

在求得邻接矩阵 Z 后, 最终图像的聚类结构可通过谱聚类方法来获得, 本文采用 Normal cut 方法. 首先为满足对称性, 取 $\tilde{Z} = |Z| + |Z^T|$, 其中 $|\cdot|$ 表示逐项取绝对值, 然后再利用 Normalized cuts 算法^[15] 获取最终的聚类结果.

3 数值优化算法

针对本文提出的基于局部图拉普拉斯约束的低秩表示聚类模型 (3), 本节设计了一种基于交替方向乘子方法 (Alternative direction method of multipliers, ADMM)^[20-21] 的快速算法. 根据定理 1, 本文模型 (3) 可以转化为两个子问题 (4) 和 (5) 的求解. 首先, 关于 RPCA 子问题 (4) 的计算已经比较成熟, 本文直接采用已有算法进行计算; 问题 (5) 是一个典型的凸优化问题, 但核范数 $\|\cdot\|_*$ 虽然是凸的, 但不可微, 为克服此问题, 本文引入一个辅助变量 P , 那么模型 (5) 可以等价如下述优化函数

$$\min_{P,Z} \|P\|_* + \lambda \text{tr}(ZLZ^T) \quad (6)$$

$$\text{s.t. } D = DZ, Z = P \quad (7)$$

上述模型的增广 Lagrange 函数可以表示为

$$\begin{aligned} L(P, Z, L_{1,2}) = & \|P\|_* + \lambda \text{tr}(ZLZ^T) + \langle L_1, D - DZ \rangle + \\ & \langle L_2, Z - P \rangle + \frac{\beta_1}{2} \|D - DZ\|_2^2 + \frac{\beta_2}{2} \|Z - P\|_2^2 \end{aligned}$$

其中, L_1, L_2 为拉格朗日乘子, β_1, β_2 为增广项参数.

显然增广函数 $L(\cdot)$ 对于每一个优化变量都是可分离的, 那么就可以在优化每一个变量的同时固定其他变量, 实现交替优化. 关于变量 Z , 其优化问题是一个光滑优化, 即

$$\begin{aligned} Z^{k+1} = & \min_Z \lambda \text{tr}(ZLZ^T) + \\ & \langle L_1^k, D - DZ \rangle + \langle L_2^k, Z - P^k \rangle + \\ & \frac{\beta_1}{2} \|D - DZ\|_2^2 + \frac{\beta_2}{2} \|Z - P^k\|_2^2 \quad (8) \end{aligned}$$

问题 (7) 最终可归结为一个典型的 Sylvester 方程. 关于变量 P ,

$$P^{k+1} = \min_P \frac{1}{\beta_2} \|P\|_* + \frac{1}{2} \left\| P - \frac{L_2^k}{\beta_2} - Z^{k+1} \right\|_2^2 \quad (9)$$

该问题有闭形式的解, 即奇异值阈值算法 (Singular value thresholding, SVT)^[22]. 首先对 $L_2^k/\beta_2 + Z^{k+1}$ 做 SVD 分解, 即 $[U, \Sigma, V] = \text{SVD}(L_2^k/\beta_2 + Z^{k+1})$, 然后对奇异值 Σ 进行阈值化处理, $S_\tau(\Sigma) = \text{diag}((\sigma_i - \tau)_+)$, σ_i 是 Σ 的主对角线元素, τ 是阈值, 这里取 $1/\beta_2$. $S_\tau(\Sigma)$ 表示如果 Σ 对角线元素比 τ 大则保持不变; 若比 τ 小则取 0. 最终 P 的最优解可以表示为 $P^* = U \times S_\tau(\Sigma)V^T$.

算法 1. 基于局部图拉普拉斯约束的鲁棒低秩表示聚类算法

步骤 1. 输入已有待聚类图像数据 X , 初始化变量 $Z^0, P^0, L_1^0, L_2^0, \beta_1, \beta_2$.

步骤 2. 迭代

1) 根据式 (4) 计算更新 (D^{k+1}, E^{k+1}) ;

2) 针对子问题 (5) 利用 ADMM 算法:

a) Z : 通过 (7) 求解 Z^{k+1} ;

b) P : 利用奇异值阈值算法解决 (8), 求解 $P^{k+1} = U \times S_{1/\beta_2}(\Sigma)V^T$;

c) 乘子变量及增广系数的更新:

$$L_1^{k+1} = L_1^k + \beta_1(D^{k+1} - D^{k+1}Z^{k+1})$$

$$L_2^{k+1} = L_2^k + \beta_2(Z^{k+1} - P^{k+1})$$

$$\beta_1 = \frac{\beta_1}{2}$$

$$\beta_2 = \frac{\beta_2}{2}$$

3) 如果 $\|Z^{k+1} - Z^k\| < \text{Tolerance}$, 停止.

步骤 3. 令 $\tilde{Z} = |Z| + |Z^T|$, 利用 Normal cut 算法得到最终聚类结果.

4 仿真实验

为验证本文算法有效性, 我们分别在仿真数据和标准数据库上进行了实验. 标准数据库分别包括: Hopkins155 运动分割数据库、USPS 手写字体数据库和 YaleB 人脸数据库. 本文的实验条件为 Intel I5-2450M 2.5 GHz CPU, NVIDIA 610 M 显卡, 4 GB DDR3 内存, 以及 Matlab 2012a. 使用的聚类误差率为: 误差率 = 错误聚类样本数目/当前所取样本总数目. 本文方法的参数如表 1 所示. 其中, Tolerance 是终止参数, β_1, β_2 是拉格朗日系数, λ 是局部约束和秩的约束之间的平衡参数, λ_{err} 是 RPCA 的参数.

表 1 本文方法的参数

Table 1 The parameters of this method

参数	Hopkins155	USPS	YaleB
Tolerance	0.0001	0.0001	0.0001
β_1	0.01	1	0.1
β_2	1	1000	1.7
λ	1	0.1	1
λ_{err}	10	10	3

本文将实验结果与稀疏子空间聚类 SSC^[10]、低秩表示模型 LRR^[13]、最小二乘回归模型 LSR^[23]、LSA^[24]、MSFNR^[17] 及鲁棒的潜在低秩表示

RLLRR^[25] 等在当前聚类问题中较为优秀的方法进行比较.

4.1 仿真数据

首先在构造的仿真数据上进行实验. 本文在 20 维空间中随机选取 4 个子空间, 每个子空间的维数是 4 维, 每个子空间随机的选取 100 组数据. 为降低所构造子空间数据之间的相关性, 本文首先随机取一个 20×100 的矩阵 P_1 , 其秩为 4, 然后其余的样本点由 $P_{i+1} = T_i P_i$, $i = 1, 2, 3$ 得到, 其中 T_i 是一个随机矩阵. 由此我们得到一个包含 4 个子空间样本的 20×400 的数据矩阵. 按照上述构造方法, 本文重复随机实验 20 次, 三种方法的聚类精度如表 1 所示, 均可达到 98% 以上的准确率, 我们的方法更达到了 99% 以上的准确率. 每种算法的迭代次数如表 2 所示, 其中 LSR、LSA 及 SSC 三种方法均具有闭形式的解, 因此不需要迭代, RLLRR 算法采用了文献 [25] 所提供代码的默认值.

表 2 七种算法平均错误率的比较 (%) 及其迭代次数

Table 2 Comparison of the average error rate (%) and the number of iterations of the seven algorithms

	LRR	LSR	SSC	LSA	MSFNR	RLLRR	本文方法
错误率	1.8	1.5	0.5	1.4	0.9	0.25	0.2
迭代次数	506	-	-	-	640	1000	5

4.2 Hopkins155 数据库

本文在 Hopkins155 运动分割数据库上进行对比实验. 运动分割是指从视频序列中截取一系列 2D 的点运动轨迹, 这些运动轨迹是不同的刚性运动所形成的. 这个数据库¹ 包含了 155 个序列, 每个序列是两个或三个类别. 每一个序列, 这些 2D 运动轨迹都是在运动中被提取出来的, 图 2 展示了一些 Hopkins155 数据库中带特征点的样本图片.

表 3 中对比分析了本文方法和 SSC、LSR、LRR、LSA、MSFNR 及 RLLRR 七种方法在 Hopkins155 数据库上的结果. 可以看出本文方法具有较好的表现. 不论是二分类、三分类还是全部的数据, 本文方法的均值都是最低的, 说明本文方法在这 155 组数据上均取得了最好的聚类结果.

为了更清晰地描述实验对比结果, 我们将所有的 155 组数据按照二分类和三分类分成两组, 图 3 和图 4 分别展示了几种算法在两组数据上的聚类误差曲线图.



图 2 Hopkins155 中的带特征点的运动物体的特征轨迹
Fig. 2 Feature trajectories of moving objects with feature points in Hopkins155 database

表 3 Hopkins155 数据库上的聚类错误率 (%) 比较
Table 3 Comparison of clustering error rate (%) on the Hopkins155 database

方法	LRR	SSC	LSR	LSA	MSFNR	RLLRR	本文方法
二分类							
平均值	19.15	2.45	25.32	3.45	9.79	4.01	2.38
中值	14.72	0.00	25.12	0.59	1.74	0.66	0.39
三分类							
平均值	26.07	7.89	29.87	9.79	11.65	7.57	7.44
中值	24.53	4.19	28.66	3.05	2.20	0.85	2.61
全部							
平均值	20.71	3.68	26.35	4.88	10.21	4.82	3.52
中值	20.33	0.00	27.02	0.90	1.83	0.80	0.57

从图像上可以看出, 在二分类的情况下本文方法要优于其他几类方法, 虽然没有达到百分之百, 但误差率也基本接近于零. 但在三分类的情况下虽然本文算法的误差均值要低于其他算法, 但在最后的一些序列数据的分类上稳定性不足. 通过图 3 和图 4 我们可以发现, 与其他几种方法相比, SSC 算法和 RLLRR 算法的误差率与本文算法最为接近. 究其原因, 系本文方法在鲁棒低秩表示模型的基础上加入了局部流形结构的约束, 在保证模型分块对角性质的基础上, 增强了邻接矩阵的稀疏性, 从而提升了数据的鉴别性, 提高了聚类的准确率, 如图 1 所示. 根据上述分析, 局部性可导致稀疏性, 但反之在一般情况下不成立, 即稀疏不一定会增强数据的局部性或鉴别性. 但对于 Hopkins 数据库, 由于类别数较少 (2 或 3), 即子空间的个数较少, 稀疏性在一定程度上就会增强了数据的局部性, 因而 SSC 算法可以

¹<http://www.vision.jhu.edu/data/hopkins155>

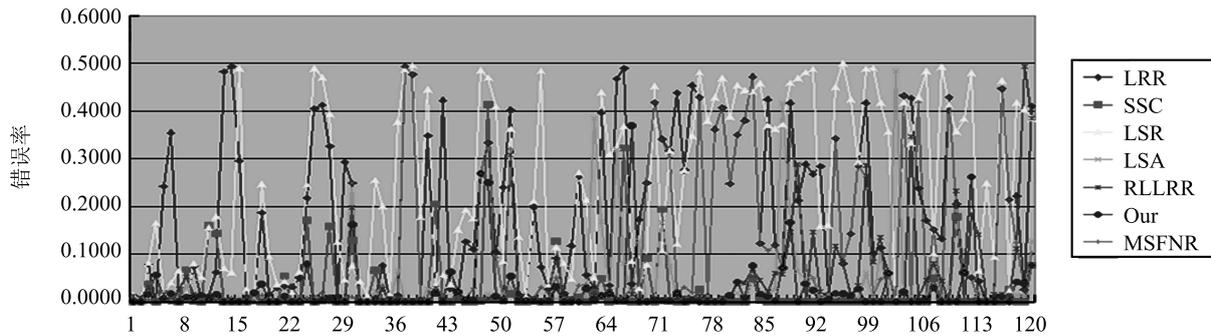


图3 Hopkins155 数据库二分类部分数据比较

Fig. 3 Comparison of two part data on the Hopkins155 database

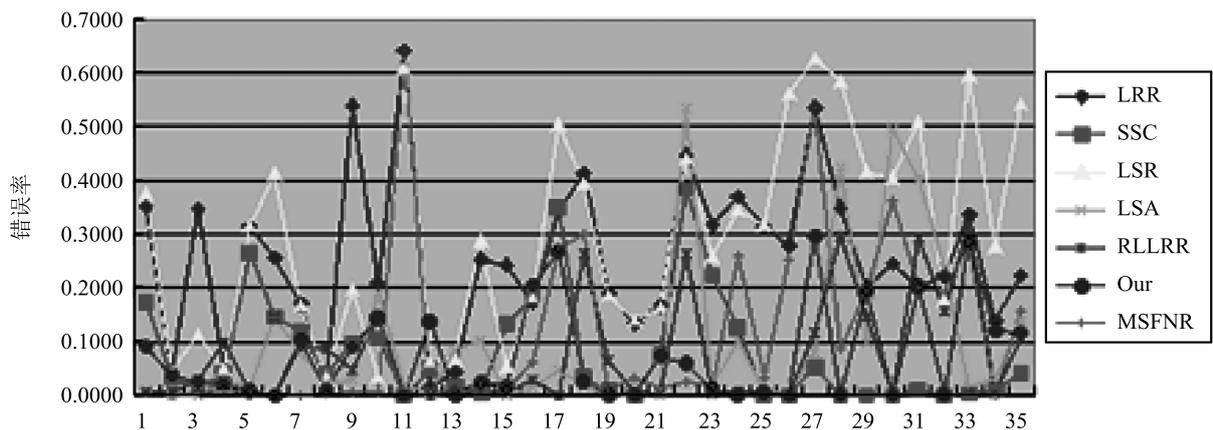


图4 Hopkins155 数据库三分类数据比较

Fig. 4 Comparison of three part data on the Hopkins155 database

得到较高的聚类准确率. RLLRR 算法是关于 LLRR (Latent low rank representation)^[26] 的改进, 针对 LLRR 模型解不唯一的问题, RLLRR 算法选取了解集中最稀疏的解, 从而在保证聚类性质的前提下一定程度上提高了聚类矩阵的稀疏性, 提升了聚类的准确率.

4.3 USPS 数据库

USPS 数据实验库包含 0~9 十个数字共 9 298 个图像, 每一个图像的大小是 16 像素 × 16 像素. 在这个实验中, 每一个图像拉成一个 256 维的列向量.

在 USPS 数据库上, 为保证实验结果分析真实有效, 针对每个数据库, 本文综合考虑不同类别情况下的聚类误差, 对于 $k = 2, 3, \dots, 9$, 我们随机选取 k 类样本, 对每一类随机选取 100 幅图像作为训练集合, 然后进行聚类实验, 统计其误差率. 对每个数据库进行重复实验 10 次, 取其误差的平均值作为最终的误差率. 图 5 展示了 USPS 数据库上的样本图片.

首先, 我们在无噪声的原始数据库上进行实验, 最终每次实验的误差率如表 4 所示. 实验结果表明, 本文算法由于引入了局部图拉普拉斯约束, 在针对类间距离较大的训练集上由于很好的保持了类内距离, 有效提高了聚类的准确率.

其次, 我们对数据库加入一定的随机遮挡, 针对遮挡数据分别进行了同样的实验, 各类算法平均误差率 ER 曲线如图 6 所示. 可以看出, 当训练数据包含遮挡时, LRR 及 LSR 算法的准确率下降明显, 而本文方法对噪声鲁棒性更强, 进一步佐证了本文的分析.

4.4 YaleB 数据库

扩展的 YaleB 数据库包含 38 个人的在不同光照条件下的 64 幅人脸图像, 如图 7 所示. 其中每个人脸的图像可以看成是一个子空间. 由于其中有一半的图像存在较大的阴影是有噪音的, 所以可以认为这些图像遭到了破坏. 这个数据库共包含 2 414 幅图像, 每一个图像的大小是 32 像素 × 32 像素. 在这

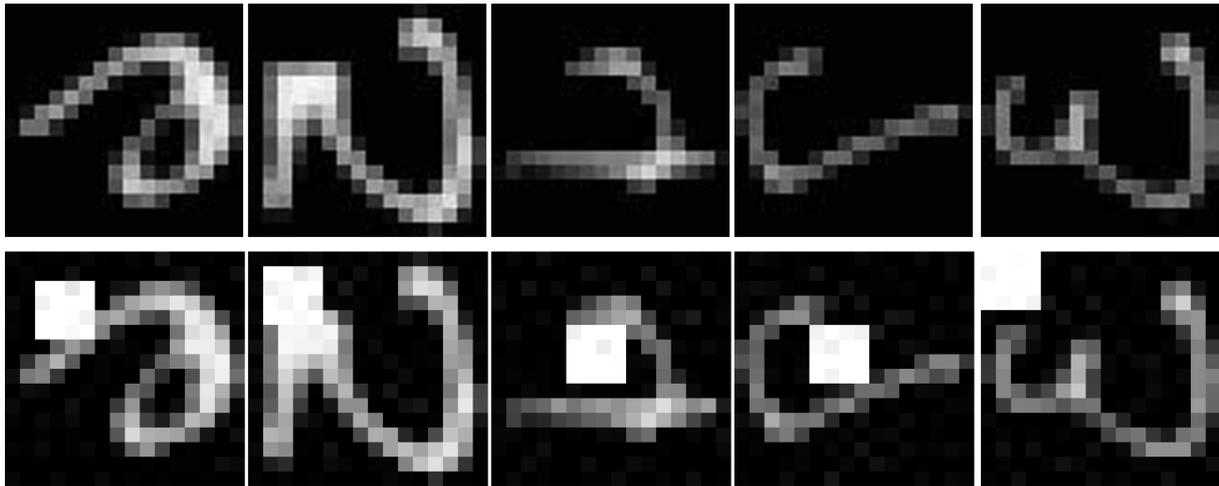


图 5 USPS 数据库中随机抽取五幅图像及加遮挡图像 (第 1 行为原始数据, 第 2 行为带遮挡数据)

Fig. 5 Five images randomly selected from the USPS database and occluded (The first row are the original images and the second row are the occluded images.)

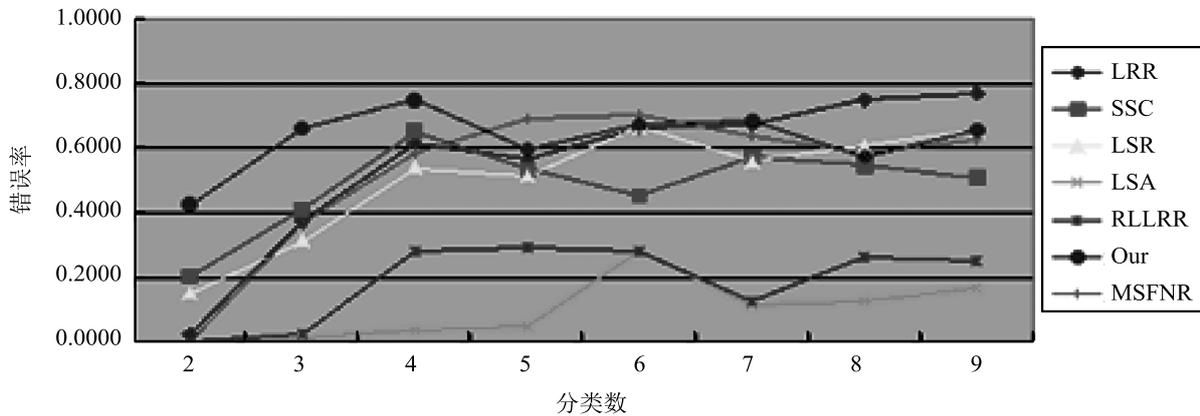


图 6 在 USPS 数据库上加上随机的 4×4 的遮挡

Fig. 6 Add a random 4×4 block on the USPS database

表 4 USPS 数据库上的错误率 (%) 比较
Table 4 Comparison of clustering error rate (%) on the USPS database

类别数	LRR	SSC	LSR	LSA	MSFNRR	RLLRR	本文方法
2 聚类	3.50	0.50	23.50	0.50	1.50	0.50	0.50
3 聚类	33.67	18.33	17.33	2.00	3.00	2.00	2.00
4 聚类	29.50	25.75	23.75	3.50	32.00	3.75	1.75
5 聚类	21.40	28.60	25.60	5.20	24.00	7.20	4.20
6 聚类	34.67	33.67	28.83	29.00	35.00	29.83	27.33
7 聚类	10.86	39.00	40.86	22.29	27.57	20.57	11.29
8 聚类	25.75	43.00	45.75	21.50	39.50	28.13	12.88
9 聚类	25.89	49.33	43.11	31.44	36.89	31.56	25.00
均值	23.16	26.46	31.09	14.43	24.93	15.41	10.62



图 7 YaleB 中的人脸图像示例

Fig. 7 The face image sample in YaleB database

个实验中, 每一个图像拉成一个 1024 维的列向量.

在这个数据库上我们分别选取其中的五分类、七分类、九分类图像作为实验组进行实验, 其中每组随机进行 10 次实验并求其平均值作为最后的实验结果, 其结果如表 5 所示.

表 5 YaleB 数据库上的错误率 (%) 比较
Table 5 Comparison of clustering error rate (%) on the YaleB database

类别数	LRR	SSC	LSR	LSA	MSFNR	RLLRR	本文方法
5 聚类	8.75	9.69	67.20	19.38	65.63	4.06	5.00
7 聚类	27.23	12.95	67.43	26.79	71.42	6.14	11.38
9 聚类	39.58	28.65	76.44	32.12	73.26	8.51	18.92

可以看出, 在 YaleB 数据库上 RLLRR 算法取得了最好的聚类效果. 由于 YaleB 数据库包含有大量不同姿态、光照及表情下的数据, 而每人仅有 64 幅图像, 因此传统聚类方法如 LRR、SSC 等无法通过全局表示下的低秩或稀疏约束很好的挖掘其空间结构, 尽管本文算法加入了局部的图拉普拉斯约束, 取得了一定的效果, 但仍不理想. 与之相比, RLLRR 算法通过考虑了潜在的因素影响, 有效的弥补了样本不足等问题, 提高了聚类的准确率.

5 结论

本文在考虑了局部性对聚类的重要性的基础上提出了一种基于局部图拉普拉斯约束的鲁棒低秩表示聚类模型. 通过加入局部图约束及有效降低了表示字典的线性相关性, 有效提高了聚类的准确率和鲁棒性. 本文从理论上给出了本文模型的分块对角性质, 保证了良好的聚类能力; 另外, 针对该模型, 本文提出了一种基于 ADMM 的快速优化算法. 最终, 本文在仿真数据及 3 个标准的数据库上进行了实验, 并与相关方法进行了比较分析, 验证了本文算法的有效性.

References

- Liu Ping-Ping, Zhao Hong-Wei, Zang Xue-Bai, Dai Jin-Bo. A fast local feature description algorithm. *Acta Automatica Sinica*, 2010, **36**(1): 40–45
(刘萍萍, 赵宏伟, 臧雪柏, 戴金波. 一种快速局部特征描述算法. 自动化学报, 2010, **36**(1): 40–45)
- Yu Wang-Sheng, Tian Xiao-Hua, Hou Zhi-Qiang, Huang An-Qi, Liu Xiang. Visual tracking algorithm based on feature matching of key regions. *Acta Electronica Sinica*, 2014, **42**(11): 2150–2156
(余旺盛, 田孝华, 侯志强, 黄安奇, 刘翔. 基于关键区域特征匹配的视觉跟踪算法. 电子学报, 2014, **42**(11): 2150–2156)
- Zeng Hui, Mu Zhi-Chun, Wang Xiu-Qing. A robust method for local image feature region description. *Acta Automatica Sinica*, 2011, **37**(6): 658–664
(曾慧, 穆志纯, 王秀青. 一种鲁棒的图像局部特征区域的描述方法. 自动化学报, 2011, **37**(6): 658–664)
- Chen Shu-Rong, Li Bo, Dong Rong, Chen Qi-Mei. Contourlet-SIFT feature matching algorithm. *Journal of*

Electronics and Information Technology, 2013, **35**(5): 1215–1221

(陈抒蓉, 李勃, 董蓉, 陈启美. Contourlet-SIFT 特征匹配算法. 电子与信息学报, 2013, **35**(5): 1215–1221)

- Lin Hai-Feng, Ma Yu-Feng, Song Tao. Research on object tracking algorithm based on SIFT. *Acta Automatica Sinica*, 2010, **36**(8): 1204–1208
(蔺海峰, 马宇峰, 宋涛. 基于 SIFT 特征目标跟踪算法研究. 自动化学报, 2010, **36**(8): 1204–1208)
- Ye Lin, Chen Yue-Lin, Lin Jing-Liang. Pedestrian fast detection based on histograms of oriented gradient. *Computer Engineering*, 2010, **36**(22): 206–210
(叶林, 陈岳林, 林景亮. 基于 HOG 的行人快速检测. 计算机工程, 2010, **36**(22): 206–210)
- Guo Jin-Xin, Chen Wei. Face recognition based on HOG multi-feature fusion and random forest. *Computer Science*, 2013, **40**(10): 279–282
(郭金鑫, 陈玮. 基于 HOG 多特征融合与随机森林的人脸识别. 计算机科学, 2013, **40**(10): 279–282)
- Tong Ying. Facial expression recognition algorithm based on spatial multi-scaled HOG feature. *Computer Engineering and Design*, 2014, **35**(11): 3918–3922, 3979
(童莹. 基于空间多尺度 HOG 特征的人脸表情识别方法. 计算机工程与设计, 2014, **35**(11): 3918–3922, 3979)
- Wright J, Yang A Y, Ganesh A, Sastry S S, Ma Y. Robust face recognition via sparse representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009, **31**(2): 210–227
- Elhamifar E, Vidal R. Sparse subspace clustering: algorithm, theory, and applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, **35**(11): 2765–2781
- Yang Rong-Gen, Ren Ming-Wu, Yang Jing-Yu. Sparse representation based face recognition algorithm. *Computer Science*, 2010, **37**(9): 267–269, 278
(杨荣根, 任明武, 杨静宇. 基于稀疏表示的人脸识别方法. 计算机科学, 2010, **37**(9): 267–269, 278)
- Vidal R, Favaro P. Low rank subspace clustering (LRSC). *Pattern Recognition Letters*, 2014, **43**: 47–61
- Liu G C, Lin Z C, Yan S C, Sun J, Yu Y, Ma Y. Robust recovery of subspace structures by low-rank representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, **35**(1): 171–184
- Shi Jia-Rong, Zheng Xiu-Yun, Wei Zong-Tian, Yang Wei. Survey on algorithms of low-rank matrix recovery. *Application Research of Computers*, 2013, **30**(6): 1601–1605
(史加荣, 郑秀云, 魏宗田, 杨威. 低秩矩阵恢复算法综述. 计算机应用研究, 2013, **30**(6): 1601–1605)
- Shi J, Malik J. Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, **22**(8): 888–905
- Peng Yi-Gang, Suo Jin-Li, Dai Qiong-Hai, Xu Wen-Li. From compressed sensing to low-rank matrix recovery: theory and applications. *Acta Automatica Sinica*, 2013, **39**(7): 981–994
(彭义刚, 索津莉, 戴琼海, 徐文立. 从压缩传感到低秩矩阵恢复: 理论与应用. 自动化学报, 2013, **39**(7): 981–994)

- 17 Wei S M, Yu Y Z. Subspace segmentation with a minimal squared frobenius norm representation. In: Proceedings of the 21st International Conference on Pattern Recognition (ICPR). Tsukuba: IEEE, 2012. 3509–3512
- 18 Candés E, Li X D, Ma Y, Wright J. Robust principal component analysis? *Journal of the ACM*, 2011, **58**(3): Article No. 11
- 19 Liu G C, Li P. Recovery of coherent data via low-rank dictionary pursuit. *Advances in Neural Information Processing Systems*, 2014, **27**: 1206–1214
- 20 Boyd S, Parikh N, Chu E, Peleato B, Eckstein J. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning*, 2011, **3**(1): 1–122
- 21 Parikh N, Boyd S. Proximal algorithms. *Foundations and Trends in Optimization*, 2014, **1**(3): 123–231
- 22 Cai J F, Candés E J, Shen Z W. A singular value thresholding algorithm for matrix completion. *SIAM Journal on Optimization*, 2010, **20**(4): 1956–1982
- 23 Lu C Y, Min H, Zhao Z Q, Zhu L, Huang D S, Yan S C. Robust and efficient subspace segmentation via least squares regression. *Computer Vision — ECCV 2012*. Berlin Heidelberg: Springer, 2012. 347–360
- 24 Yan J Y, Pollefeys M. A general framework for motion segmentation: independent, articulated, rigid, non-rigid, degenerate and non-degenerate. *Computer Vision — ECCV 2006*. Berlin, Heidelberg: Springer, 2006. 94–106
- 25 Zhang H Y, Lin Z C, Zhang C, Gao J B. Robust latent low rank representation for subspace clustering. *Neurocomputing*, 2014, **145**: 369–373
- 26 Liu G C, Yan S C. Latent low-rank representation for subspace segmentation and feature extraction. In: Proceedings of the 2011 IEEE International Conference on Computer Vision (ICCV). Barcelona: IEEE, 2011. 1615–1622



李波 南昌航空大学数学与信息科学学院副教授. 主要研究方向为计算机图形学及图像处理. 本文通信作者.

E-mail: libo@nchu.edu.cn

(**LI Bo** Associate professor at the School of Mathematics and Information Science, Nanchang Hangkong University. His research interest covers computer graphics and image processing. Corresponding author of this paper.)



卢春园 南昌航空大学数学与信息科学学院硕士研究生. 主要研究方向为图像处理及数据分析.

E-mail: lun_chunyuan@sina.com

(**LU Chun-Yuan** Master student at the School of Mathematics and Information Science, Nanchang Hangkong University. Her research interest covers image processing and data analysis.)



冷成财 南昌航空大学数学与信息科学学院讲师. 主要研究方向为医学图像处理. E-mail: 70587@nchu.edu.cn

(**LENG Cheng-Cai** Lecturer at the School of Mathematics and Information Science, Nanchang Hangkong University. His research interest covers medical image processing.)



金连宝 南昌航空大学数学与信息科学学院硕士研究生. 主要研究方向为图像处理及数据分析.

E-mail: jin.lianbao@sina.com

(**JIN Lian-Bao** Master student at the School of Mathematics and Information Science, Nanchang Hangkong University. His research interest covers image processing and data analysis.)