

基于潜在特征空间的低秩表示算法

周翊航

广东工业大学计算机学院, 广东 广州

Email: 562220078@qq.com

收稿日期: 2021年3月28日; 录用日期: 2021年4月21日; 发布日期: 2021年4月28日

摘要

现有的大多低秩表示算法都是直接使用原始数据矩阵作为特征字典, 然而原始数据中的冗余特征和噪声信息很可能会导致算法效果不佳。针对这种情况, 本文提出一种基于潜在特征空间的低秩表示算法, 通过正交字典来学习原始数据的潜在表示, 然后利用潜在表示作为字典进行低秩表示学习, 从而避免原始数据中的不利影响。

关键词

潜在特征, 正交字典, 低秩表示

Low-Rank Representation Algorithm Based on Latent Feature Space

Yihang Zhou

School of Computers, Guangdong University of Technology, Guangzhou Guangdong

Email: 562220078@qq.com

Received: Mar. 28th, 2021; accepted: Apr. 21st, 2021; published: Apr. 28th, 2021

Abstract

Most of the existing low-rank representation algorithms directly use the original data matrix as the feature dictionary, but the redundant features and noise information in the original data may cause the algorithm to perform poorly. In response to this situation, a low-rank representation algorithm based on the latent feature space is proposed. The latent representation of the original data is learned through an orthogonal dictionary, and then the latent representation is used as a dictionary for low-rank representation learning, so as to avoid the adverse effects of the original data.

Keywords

Latent Features, Orthogonal Dictionary, Low-Rank Representation

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着社会经济的发展和科学技术的进步，人们获取信息的手段和渠道被大为拓宽。但是当我们在使用的特定数据时，往往会由于缺乏先验知识或者成本过高而难以进行人工标注。在这种数据没有标签的情况下，很自然地，我们会借助计算机来辅助完成一些任务。在机器学习中，这种根据类别未知的训练数据集来解决模式识别中的各种问题，称之为无监督学习。

“大数据”的迅猛发展推动了高维数据的处理需求，同时其也成为了无监督学习领域的研究热点之一。高维数据含有丰富的信息，对样本的描述更加充分，理论上这类高维数据应有助于机器学习到更好的模型。然而，高维的数据同样带来了算法复杂性的提升，不仅需要大量的空间来存储这些数据，而且还需要花费大量的时间来进行模型的训练，严重制约了算法的应用，特别是对于如视频跟踪等实时性要求比较严格的应用场景。此外，高维的数据尽管信息量丰富，但是却含有大量的冗余特征，甚至部分数据由于不规范的采集方式或恶劣的采集环境会出现大量的噪声或低分辨率等现象，给数据分析和模型学习带来了极大的挑战。如果直接对这样的大数据进行分析，不仅效率低下，而且由于冗余特征和噪声数据的存在难以学习到一个鲁棒的机器模型。为解决这一问题，人们在数据分析前对数据进行降维，提取出少量最有价值的紧凑信息进行分析，不仅节约了存储空间和处理时间，而且可以学习到鲁棒的模型。子空间学习作为一种经典的降维方法，由于其简单高效的特点而被广泛地研究使用。其主要含义是指通过寻找合适的投影，实现高维特征向低维子空间的映射，使得在降低维度的同时减少信息损耗。

近年来涌现了许多高效的子空间学习方法。其中低秩表示方法由于较低的计算复杂度和其出色的性能，而广受研究和关注。但是大多数低秩表示方法都是直接将原始数据作为特征字典来进行学习的，这样的实验结果很大程度上会受到原始数据中错误信息的影响。于是我们为了解决这个问题，提出基于潜在特征空间的低秩表示方法。

2. 算法提出

2.1. 相关算法

给定原始数据的特征矩阵 $X = [x_1, x_2, \dots, x_n] \in \mathcal{R}^{m \times n}$ ，其中列向量 x_i 表示一个样本。则低秩表示算法 (Low-Rank Representation, LRR) [1] 的模型可以用如下目标函数表示：

$$\min_{Z, E} \|Z\|_* + \lambda \|E\|_{2,1} \quad s.t. \quad X = XZ + E \quad (1)$$

其中， $Z \in \mathcal{R}^{m \times n}$ 为表示系数矩阵， E 为重构误差。 $\|\cdot\|_{2,1}$ 表示 $l_{2,1}$ 范数，其可以引导矩阵列稀疏。LRR 算法将原始特征矩阵作为字典，通过对原始特征矩阵进行重构，来学习一个具有低秩约束的表示系数矩阵。这样 Z 通常可以揭示样本点之间的相似性关系，从而用来做无监督的聚类分析。

2.2. 算法引入

在一般情况下，LRR 算法可以取得不错的聚类效果。然而在实际应用中，受到各种环境和人为等各种因素的影响，我们所得到的数据往往包含冗余特征和噪声信息。所以直接将原始特征矩阵作为字典，学习得到的表示系数矩阵通常会受到其中的不利因素的影响。为此，我们提出潜在特征空间学习方法，来减少这种负面影响：

$$X = PHX + E_1 \quad s.t. P^T P = I \quad (2)$$

其中 $P \in \mathcal{R}^{m \times k}$ 为正交字典， $H \in \mathcal{R}^{k \times m}$ 为转换矩阵， E_1 为重构误差。通过这种方式，我们可以利用原始数据重构，学习转换矩阵 H ，进而获得原始特征的潜在表示 HX 。这样可以使得潜在特征在显式地保留原始数据的主要信息同时还能过滤噪声信息 E_1 。

进一步，我们结合低秩表示方法，并将 HX 作为特征字典，可以得到基于潜在特征空间的低秩表示方法(Low-Rank Representation Algorithm Based on Latent Feature Space, LRRLFS)，其目标函数为：

$$\begin{aligned} \min_{P, H, Z, E_1, E_2} \|Z\|_* + \lambda_1 \|HX\|_F^2 + \lambda_2 (\|E_1\|_{2,1} + \|E_2\|_{2,1}) \\ s.t. X = PHX + E_1, HX = HXZ + E_2, P^T P = I \end{aligned} \quad (3)$$

其中 $\|HX\|_F^2$ 作为正则化项，一方面可以使得学习得到的潜在特征表示更为紧凑，另一方面可以在一定程度上防止模型过拟合。

2.3. 算法优化

显然目标函数(3)并不是一个凸优化问题。但是我们可以采用交替直接最小化(Alternating Direction Minimization, ADM) [2]技巧来对其进行优化。首先，引入辅助变量 J 和 K 来分别代替 Z 和 HX 。那么目标函数表达式变为如下形式：

$$\begin{aligned} \min_{P, H, Z, E_1, E_2, J, K} \|J\|_* + \lambda_1 \|HX\|_F^2 + \lambda_2 (\|E_1\|_{2,1} + \|E_2\|_{2,1}) \\ s.t. X = PK + E_1, HX = HXZ + E_2, P^T P = I, J = Z, K = HX \end{aligned} \quad (4)$$

接着，我们将上式(4)转化为增强拉格朗日问题(Augmented Lagrange Multiplier, ALM)：

$$\begin{aligned} \mathcal{L}(P, H, Z, E_1, E_2, J, K) \\ = \|J\|_* + \lambda_1 \|HX\|_F^2 + \lambda_2 (\|E_1\|_{2,1} + \|E_2\|_{2,1}) + \Phi(Y_1, X - PK - E_1) \\ + \Phi(Y_2, HX - HXZ - E_2) + \Phi(Y_3, J - Z) + \Phi(Y_4, K - HX) \\ s.t. P^T P = I \end{aligned} \quad (5)$$

其中 $\Phi(\cdot, \cdot)$ 的定义为 $\Phi(A, B) = \frac{\mu}{2} \|B\|_F^2 + Tr(A^T B)$ 。 A 表示拉格朗日乘数， $\mu > 0$ 为惩罚因子。我们可以将上式(5)分解为多个子问题，并通过固定其他变量来更新单个变量。这样多次迭代之后，便能得到模型的近似解。具体子问题如下：

1) P 步骤：固定其他变量，更新 P 。

$$P^* = \arg \min_P \Phi(Y_1, X - PK - E_1) \quad s.t. P^T P = I \quad (6)$$

定理 1 对于最小化问题 $\min_C \|A - BC\|_F^2 \quad s.t. C^T C = CC^T = I$ ，其最优解为 $C^* = UV^T$ ，其中 U 和 V 分别表示对 $B^T A$ 进行奇异值分解(Singular Value Decomposition, SVD)分解后的左右奇异矩阵。

根据定理 1 [3]，我们将式(6)转化为类似的形式：

$$\begin{aligned}
P^* &= \arg \min_p \Phi(Y_1, X - PK - E_1) \\
&= \arg \min_p \frac{\mu}{2} \text{Tr}((X - PK - E_1)^T (X - PK - E_1)) + \text{Tr}((X - PK - E_1)^T Y_1) \\
&= \arg \min_p \frac{\mu}{2} \text{Tr}((X - PK - E_1 + Y_1/\mu)^T (X - PK - E_1 + Y_1/\mu)) \\
&= \arg \min_p \frac{\mu}{2} \|X - PK - E_1 + Y_1/\mu\|_F^2 \\
&= \arg \min_p \frac{\mu}{2} \|(X + Y_1/\mu - E_1) - PK\|_F^2 \\
&= \arg \min_p \frac{\mu}{2} \|(X + Y_1/\mu - E_1)^T - K^T P^T\|_F^2
\end{aligned} \tag{7}$$

然后通过 SVD 得到 $K(X + Y_1/\mu - E_1)^T = U\Sigma V^T$, 利用 U 和 V^T 来更新 P :

$$P^* = [UV^T]^T \tag{8}$$

2) H 步骤: 固定其他变量, 更新 H 。

$$H^* = \arg \min_H \Phi(Y_2, HX - HXZ - E_2) + \Phi(Y_4, K - HX) + \lambda_1 \|HX\|_F^2 \tag{9}$$

我们将上式(9)对 H 求偏导并设为零, 便可以得到其最小化问题的解:

$$H^* = [(\mu E_2 - Y_2)(I - Z)^T X^T + (\mu K + Y_4)X^T] [\mu X(2I - Z^T - Z + ZZ^T)X^T + 2\lambda_1 XX^T]^{-1} \tag{10}$$

3) Z 步骤: 固定其他变量, 更新 Z 。

$$Z^* = \arg \min_Z \Phi(Y_3, J - Z) + \Phi(Y_2, HX - HXZ - E_2) \tag{11}$$

我们将上式(11)对 Z 求偏导并设为零。便可以得到其最小化问题的解:

$$Z^* = (I + X^T H^T HX)^{-1} [(J + X^T H^T HX - X^T H^T E_2) + (Y_3 + X^T H^T Y_2)/\mu] \tag{12}$$

4) E 步骤: 固定其他变量, 更新 E 。

$$\begin{aligned}
E_1^* &= \arg \min_{E_1} \lambda_2 \|E_1\|_{2,1} + \Phi(Y_1, X - PK - E_1) \\
&= \arg \min_{E_1} \frac{\lambda_2}{\mu} \|E_1\|_{2,1} + \frac{1}{2} \|E_1 - (X - PK + Y_1/\mu)\|_F^2
\end{aligned} \tag{13}$$

$$\begin{aligned}
E_2^* &= \arg \min_{E_2} \lambda_2 \|E_2\|_{2,1} + \Phi(Y_2, HX - HXZ - E_2) \\
&= \arg \min_{E_2} \frac{\lambda_2}{\mu} \|E_2\|_{2,1} + \frac{1}{2} \|E_2 - (HX - HXZ + Y_2/\mu)\|_F^2
\end{aligned} \tag{14}$$

引理 1 给定 B , 则 $\min_A \alpha \|A\|_{2,1} + \frac{1}{2} \|A - B\|_F^2$ 的最优解 A^* 的第 i 列可以表示为:

$$[A^*]_{:,i} = \begin{cases} \frac{\| [B]_{:,i} \|_2 - \alpha}{\| [B]_{:,i} \|_2} B_{:,i}, & \text{if } \| [B]_{:,i} \|_2 > \alpha \\ 0, & \text{otherwise} \end{cases} \tag{15}$$

根据引理 1, 我们可以依次更新 E_1 和 E_2 。

5) J 步骤: 固定其他变量, 更新 J 。

$$\begin{aligned}
J^* &= \arg \min_J \|J\|_* + \Phi(Y_3, J - Z) \\
&= \arg \min_J \|J\|_* + \frac{\mu}{2} \|J - (Z - Y_3/\mu)\|_F^2 \\
&= \arg \min_J \frac{1}{\mu} \|J\|_* + \frac{1}{2} \|J - (Z - Y_3/\mu)\|_F^2
\end{aligned} \tag{16}$$

我们可以利用奇异值阈值算法(Singular Value Thresholding, SVT) [4]算法计算最优解 J^* :

$$J^* = \Theta_{1/\mu}(Z - Y_3/\mu) \tag{17}$$

其中 Θ 表示 SVT 算法的收缩算子。

6) K 步骤: 固定其他变量, 更新 K 。

$$K^* = \arg \min_K \Phi(Y_4, K - HX) \tag{18}$$

我们将上式(18)对 K 求偏导并设为零, 便可以得到其最小化问题的解:

$$K^* = HX - Y_4/\mu \tag{19}$$

7) 更新拉格朗日乘子: 所有乘子均采用梯度下降法更新。

$$\begin{cases}
Y_1 = Y_1 + \frac{\partial \mathcal{L}}{\partial Y_1} = Y_1 + \mu(X - PK - E_1) \\
Y_2 = Y_2 + \frac{\partial \mathcal{L}}{\partial Y_2} = Y_2 + \mu(HX - HXZ - E_2) \\
Y_3 = Y_3 + \frac{\partial \mathcal{L}}{\partial Y_3} = Y_3 + \mu(J - Z) \\
Y_4 = Y_4 + \frac{\partial \mathcal{L}}{\partial Y_4} = Y_4 + \mu(K - HX)
\end{cases} \tag{20}$$

综上所述, 我们可以得到 LRRLFS 算法如下:

算法 1: LRRLFS 算法

输入: 原始特征矩阵 X , 超参 λ 和转换矩阵 H 的维度 k 。

初始化: $H = \arg \min_P \text{Tr}(P^T(-\Sigma)P)$, s.t. $P^T P = I$, 其中 Σ 表示协方差矩阵; $P = 0$; $E_1 = 0$; $E_2 = 0$; $J = Z = 0$; $K = HX$; $Y_1 = 0$; $Y_2 = 0$; $Y_3 = 0$; $Y_4 = 0$; $\rho = 1.09$; $\mu = 10^{-6}$; $\epsilon = 10^{-4}$; $\max_\mu = 10^6$ 。

while not covered **do**

1. 根据步骤 1~6, 更新 P, H, Z, E_1, E_2, J, K ;

2. 根据步骤 7, 更新 Y_1, Y_2, Y_3, Y_4 ;

3. 根据 $\mu = \min(\rho\mu, \max_\mu)$ 更新 μ ;

4. 检查收敛条件是否满足:

$$\|X - PK - E_1\|_\infty < \epsilon, \|HX - HXZ - E_2\|_\infty < \epsilon, \|J - Z\|_\infty < \epsilon, \|K - HX\|_\infty < \epsilon$$

end while

输出: P, H, Z, E_1 和 E_2 。

3. 实验分析

3.1. 数据集及其预处理

在聚类实验中使用的 4 个数据集介绍及其预处理方式如下:

Yale B 数据集[5]由 2414 张, 38 个类别的人脸图片组成。每一个类别大约包含 63 张不同光照条件下的正面人脸图片。超过一大半的图片都受到了不同程度的反射和阴影的损坏。我们将这些图片转换成灰度图片, 然后下采样为 32×32 像素的形式。最后将其转化为 1024 维的列向量并堆叠成特征矩阵。

PIE 数据集[6]包含 41,368 张, 68 个类别的人脸图片组成。每张图片拍摄于不同的情况之下, 其中包括 4 种不同表情, 13 种不同姿势, 43 种不同的光照条件。我们将这些图片转换成灰度图片, 然后下采样为 32×32 像素的形式。最后将其转化为 1024 维的列向量并堆叠成特征矩阵。

AR 数据集[7]包含超过 4000 张, 126 个类别的人脸彩色图片。这些图片都是正面姿势拍摄, 但是有不同的面部表情, 光照条件和遮挡物(围巾和墨镜)。我们将这些图片转换成灰度图片, 然后下采样为 40×50 像素的形式。最后将其转化为 2000 维的列向量并堆叠成特征矩阵。

COIL20 数据集[8]包括 1440 张 20 个类别的图片。其中每个类别一共有 72 张拍摄角度相差 5 度的 32×32 像素的灰度图片。我们将其转化为 1024 维的列向量并堆叠成特征矩阵。

对于数量为 c 的不同聚类簇设定下, 我们直接选取前 c 个类作为数据子集, 并重复实验 30 次取平均作为最终结果。部分数据集图片示例如图 1 所示。



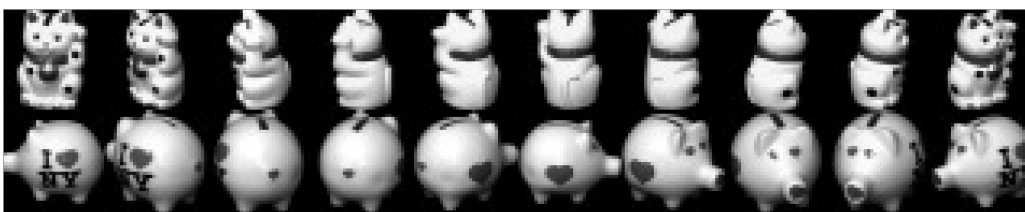
(a) Yale B 数据集



(b) PIE 数据集



(c) AR 数据集



(d) COIL20 数据集

Figure 1. Image examples of different data sets

图 1. 不同数据集的部分图片示例

3.2. 对比实验

我们采用 Accuracy (ACC) [9] 度量来衡量实验结果的有效性。给定 x_i ，令 g_i 和 l_i 分别表示对应的真实标签和预测标签。ACC 的定义为：

$$\text{ACC} = \frac{\sum_{i=1}^n \delta(g_i, \text{map}(l_i))}{n} \quad (21)$$

其中， n 表示样本数。当且仅当 $p = q$ 时， $\delta(p, q) = 1$ 。否则 $\delta(p, q) = 0$ 。 $\text{map}(l_i)$ 为置换映射函数，其可以将每个的预测标签 l_i 映射到对应的数据集标签中。一般采用 Kuh-Munkres Algorithm [10] 来寻找最佳的置换映射函数。

对比实验结果如表 1~4 所示。从中不难发现，LRRIFS 算法相较于现有的基于低秩表示的算法有着更高效的聚类准确性。这是因为 LRRIFS 算法通过潜在特征表示，过滤了原始数据中的冗余特征和噪声信息。然后以其作为低秩表示学习的特征字典，获得具有清晰聚类结构的表示系数矩阵。

Table 1. ACC of different methods on Yale B data set

表 1. Yale B 数据机上不同算法的 ACC

c	LRR	LatRR	NSLLRR	DGLRR	NNDGLRR	LRPP_GRR	PLrSC	LDRGC	LRRIFS
#2	99.22	99.22	98.20	99.22	98.16	98.44	99.22	80.47	97.51
#8	78.68	80.00	76.36	78.65	80.65	83.98	82.85	37.31	91.42
#14	78.39	77.20	75.86	80.40	77.96	88.55	81.31	31.18	90.30
#20	74.70	73.08	73.60	77.81	79.38	71.87	79.28	34.39	86.71
#26	76.07	74.93	72.69	76.67	75.55	72.90	76.98	32.38	75.14
#32	74.70	75.99	69.43	74.92	75.82	66.40	74.96	33.55	78.48
#38	71.30	73.69	71.22	71.14	74.87	66.20	74.30	33.06	75.34
Avg.	79.01	79.16	76.77	79.83	80.34	78.33	81.27	40.33	84.99

Table 2. ACC of different methods on PIE data set

表 2. PIE 数据集上不同算法的 ACC

c	LRR	LatRR	NSLLRR	DGLRR	NNDGLRR	LRPP_GRR	PLrSC	LDRGC	LRRIFS
#4	100	100	100	100	100	100	100	100	100
#12	89.68	88.49	76.98	91.67	78.17	96.43	74.05	80.16	97.32
#20	81.19	88.57	86.19	92.62	87.86	95.71	76.64	85.48	96.24
#28	77.21	73.13	86.91	92.69	85.03	92.35	78.78	85.54	93.12
#36	68.92	68.25	76.59	87.83	85.32	92.06	82.84	84.13	93.45
#44	71.86	74.68	80.2	84.42	87.12	93.72	89.15	83.87	94.36
#52	71.25	74.08	74.18	88	82.69	97.25	91.38	85.81	98.22
#60	65.87	69.76	78.73	85.56	85.24	97.62	88.53	85.48	98.63
#68	65.27	64.29	85.08	88.1	82.7	97.91	87.13	83.54	97.98
Avg.	76.81	77.92	82.76	90.1	86.01	95.89	85.38	86.06	96.59

Table 3. ACC of different methods on AR data set
表 3. AR 数据集上不同算法的 ACC

c	LRR	LatRR	NSLLRR	DGLRR	NNDGLRR	LRPP_GRR	PLrSC	LDRGC	LRRLFS
#15	76.51	77.21	74.87	78.46	77.95	73.59	77.4	60.51	80.26
#30	74.29	76.86	71.92	74.17	77.41	73.33	73.88	55.77	77.62
#45	77.14	77.16	76.15	77.72	75.48	76.75	75.8	55.04	78.21
#60	78.72	77	74.74	78.12	75.74	75.00	76.08	59.74	79.36
#75	77.53	77.13	74.24	77.94	74.69	75.08	73.92	56.13	78.67
#90	78.33	76.08	75.47	77.12	74.23	74.62	72.7	57.01	78.94
#105	76.08	73.58	73.52	76.67	72.88	73.26	75.27	51.9	77.54
#120	76.09	70.04	72.08	73.79	70.29	73.11	72.65	49.49	74.17
Avg.	76.836	75.63	74.12	76.74	74.83	74.34	74.71	55.70	78.10

Table 4. ACC of different methods on COIL20 data set
表 4. COIL20 数据集上不同算法的 ACC

c	LRR	LatRR	NSLLRR	DGLRR	NNDGLRR	LRPP_GRR	PLrSC	LDRGC	LRRLFS
#4	68.4	62.85	63.89	69.44	81.6	96.53	89.17	84.72	98.15
#6	76.85	75.93	75.93	77.78	58.8	90.97	90.25	80.35	92.36
#8	74.65	82.12	81.94	78.3	73.96	77.26	89.41	75.35	93.41
#10	62.92	68.33	68.89	65.83	75.42	80.97	81.14	80.28	91.64
#12	68.17	68.75	59.95	62.38	69.33	83.91	83.7	83.56	92.75
#14	66.77	46.53	51.29	54.46	70.44	86.01	84.54	85.91	90.34
#16	65.71	60.93	62.59	67.71	73.52	87.76	81.87	87.67	91.74
#18	65.43	48.77	59.18	63.89	66.9	89.12	79.29	88.81	91.14
#20	64.58	60.28	61.88	70.21	74.93	83.75	77.02	79.24	90.47
Avg.	68.16	63.83	65.06	67.78	71.66	86.25	84.04	82.87	92.44

3.3. 重构误差分析

为了进一步直观地展示 LRRLFS 方法在过滤冗余特征和噪声信息时的有效性，我们选取部分 Yale B 数据集的图片进行可视化处理。结果如图 2 所示。我们可以发现，潜在特征表示保留了大部分原始图片信息，而大多数的阴影噪声得到了过滤。

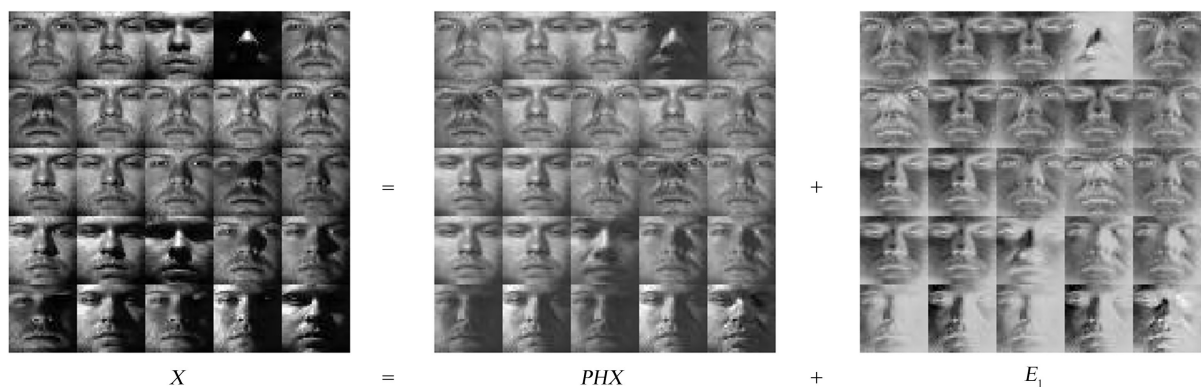


Figure 2. Reconstruction error on the Yale B data set
图 2. Yale B 数据集上的重构误差

4. 结论与展望

针对现有低秩表示方法的直接使用原始数据作为特征字典所带来的问题，我们提出了基于潜在特征空间的低秩表示方法。其在一定程度上避免原始数据中冗余特征和噪声信息对学习表示系数矩阵的不利影响，从而提升聚类性能。在未来工作中，我们会考虑将本方法扩展到多视图聚类领域。

参考文献

- [1] Liu, G., Lin, Z., Yan, S., Sun, J., Yu, Y. and Ma, Y. (2013) Robust Recovery of Subspace Structures by Low-Rank Representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **35**, 171-184. <https://doi.org/10.1109/TPAMI.2012.88>
- [2] Lin, Z., Chen, M. and Ma, Y. (2010) The Augmented Lagrange Multiplier Method for Exact Recovery of Corrupted Low-Rank Matrices. *Mathematical Programming*, **9**, arXiv:1009.5055.
- [3] Wahba, G. (1965) A Least Squares Estimate of Satellite Attitude. *Siam Review*, **7**, 409-409. <https://doi.org/10.1137/1007077>
- [4] Cai, J.F., Candès, E.J. and Shen, Z.W. (2010) A Singular Value Thresholding Algorithm for Matrix Completion. *SIAM Journal on Optimization*, **20**, 1956-1982. <https://doi.org/10.1137/080738970>
- [5] Georgiades, A.S., Belhumeur, P.N. and Kriegman, D.J. (2001) From Few to Many: Illumination Cone Models for Face Recognition under Variable Lighting and Pose. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **23**, 643-660. <https://doi.org/10.1109/34.927464>
- [6] Sim, T., Baker, S. and Bsat, M. (2003) The CMU Pose, Illumination, and Expression Database. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **25**, 1615-1618. <https://doi.org/10.1109/TPAMI.2003.1251154>
- [7] Martínez, A. and Benavente, R. (1998) The AR Face Database. *Technical Report*. Computer Vision Center, Number 24, Bellaterra. <http://www.cat.uab.cat/Public/Publications/1998/MaB1998>
- [8] Nene, S.A., Nayar, S.K. and Murase, H. (1996) Columbia Object Image Library (Coil-20). *Technical Report CUCS-005-96*, February 1996.
- [9] He, X., Cai, D. and Han, J. (2005) Document Clustering Using Locality Preserving Indexing. *IEEE Transactions on Knowledge and Data Engineering*, **17**, 1624-1637. <https://doi.org/10.1109/TKDE.2005.198>
- [10] Wessel, W. (1988) Lovász, L.; Plummer, M. D., Matching Theory. Budapest, Akadémiai Kiadó 1986. XXXIII, 544 S., Ft 680,—. ISBN 9630541688. *Journal of Applied Mathematics and Mechanics*, **68**, 146-146. <https://onlinelibrary.wiley.com/>
<https://doi.org/10.1002/zamm.19880680310>