

基于同类测试样本组的稀疏表示人脸识别

谢尚高,王丽平

(南京航空航天大学 理学院,江苏 南京 210016)

摘要:近年来的研究表明,稀疏表示分类(SRC)方法是一种有效的人脸识别方法。SRC 是单个样本基于向量 l_1 -范数正则化的最小二乘分类。但现实中常常存在着已知多个测试样本属于同一类的情况,无疑有利于分类,而基于 SRC 或其他单样本模型的方法却未能利用该信息。为利用类别标签信息,提出了一种新的鲁棒人脸识别方法。该方法基于同类测试样本组的稀疏表示分类(IGSRC),即将同类多个测试样本放至同组,采用矩阵 L_1 -范数正则化的最小二乘分类进行稀疏表示,将测试样本组判为类别中残差最小的标号。实验结果表明,相比于 SRC 与 IGSR 方法,所提出的方法不但能取得更高的人脸识别率(即使在每类别训练样本数较少、训练样本存在部分遮挡),而且具有更少的计算耗时。

关键词:类内测试样本组;稀疏表示;人脸识别;矩阵 L_1 -范数;多样本

中图分类号:TP301

文献标识码:A

文章编号:1673-629X(2017)08-0007-05

doi:10.3969/j.issn.1673-629X.2017.08.002

Sparse Representation Classification for Face Recognition with Intra-class Testing-sample Group

XIE Shang-gao, WANG Li-ping

(College of Science, Nanjing University of Aeronautics and Astronautics,
Nanjing 210016, China)

Abstract: Recent studies have shown that Sparse Representation Classification (SRC) is an effective method for face recognition. SRC is a least squares classification based on l_1 -norm regularized for a single testing-sample. However, in the case that multiple testing-samples are known to be the same class which is surely helpful in the classification, the common-class label information is not included in SRC or other single-sample models. Therefore, a novel robust face recognition method based on sparse representation classification is proposed which is on the basis of IGSR. Taking multiple intra-class testing-samples into the same group, it adopts the matrix L_1 -norm regularized least squares classification for sparse representation and judges the test sample group as the label with minimum error in classes. Experimental results show that compared with SRC and IGSR, the method proposed cannot only obtain better face recognition rate (even when the number of training samples per subject is small or training samples are partly occluded), also own less running time.

Key words: intra-class testing-samples; sparse representation; face recognition; matrix L_1 -norm; multiple samples

0 引言

人脸识别是指对未分类的人脸样本从众多人脸中判别出类别。人脸识别一直是计算机视觉、模式识别等领域的研究热点,主要集中在两部分:一是人脸特征的选取,二是分类器的构造。在人脸特征选取方面,早期的方法主要考虑人脸几何特征,后期则从统计学习的角度出发,有了更进一步的发展,包括基于主成分分析的特征脸^[1]、应用线性判别分析的 Fisher 脸^[2]、基于局部特征的提取方法(如 Gabor 小波法^[3-4]、LBP^[5]方法)等。在分类器方面,线性分类器使用最多,如最近

邻(Nearest Neighbor)^[6]、最近特征线(Nearest Feature Line)^[7]、最近子空间平面(Nearest Subspace Plane)^[8-10]等。

随着压缩感知理论^[11-13]的发展,J. Wright 等^[14]提出用于人脸识别的稀疏表示分类(Sparse Representation-based Classification, SRC)方法,并在数值实验上证明该方法的识别准确率和鲁棒性都有所提升。SRC 提出单个测试样本可以用所有类别的训练样本进行线性表示,它通过 l_1 -范数最小化自动寻找最稀疏的线性表示,大部分非零因子应集中在正确的类别上,从而

收稿日期:2016-08-03

修回日期:2016-11-08

网络出版时间:2017-07-05

基金项目:国家自然科学基金资助项目(11471159,61661136001);南京航空航天大学研究生创新开放基金(kfj20150706)

作者简介:谢尚高(1989-),男,硕士研究生,研究方向为最优化算法及其应用;王丽平,副教授,研究方向为最优化理论与方法及应用。

网络出版地址: <http://kns.cnki.net/kcms/detail/61.1450.TP.20170705.1650.028.html>

自动满足寻找正确类别的过程。稀疏表示具有很好的鲁棒性^[15-16],后续对 SRC 算法的改进及拓展包括 CRC^[17]、ESRC^[18]、PCRC^[19]等。其中 CRC 认为 SRC 在识别中起主要作用的是协同表示,而非稀疏优化,所以 CRC 采用 l_2 -范数最小化取得了和 SRC 类似的表现,但识别速度更快;ESRC 则加强了对字典的学习,加入类内差异与原训练样本集组成新的训练字典。

SRC 及拓展方法都只考虑单个测试样本的表示与分类,而单个人脸样本因为表情、光照、遮挡等因素具有较大的随机性,有时同类之间也有较大的差异性,甚至存在类内差异大于类间差异的情况,而在许多应用场景中,同类的待识别样本往往有多个,单个样本逐一识别没有充分利用这些待识别样本属于同类这一先验知识。

针对该问题,考虑基于多样本的同类测试样本组的稀疏表示分类(Sparse Representation Classification with Intra-class testing-sample Group,IGSRC),将测试样本中已知属于同类的放至同组,对该组测试样本进行矩阵 L_1 -范数正则化的最小二乘(Frobenius 范数)表示,通过 ADMM(Alternating Direction Method of Multipliers)优化求解得表示系数后,计算对应的各个类别的重构残差,将该组测试样本判别为最小残差的类别,并通过实验进行验证。

1 稀疏表示分类

SRC 将人脸识别模型表述成如下 l_1 -范数最小化问题:

$$x^* = \arg \min_x \|y - Ax\|_2^2 + \lambda \|x\|_1 \in R^n \quad (1)$$

其中, $y \in R^d$ 表示待分类的单个测试样本; $A = [v_1, v_2, \dots, v_n] \in \mathbb{R}^{d \times n}$ 表示训练样本组成的集合; $v_i \in R^d$ 表示第 i 个训练样本; $x \in R^n$ 表示系数; $\|x\|_1 = \sum_{i=1}^n |x_i|$ 表示 l_1 -范数,其中 x_i 表示 x 的第 i 个分量; $\lambda (> 0)$ 表示正则化参数。

令 $\delta_j(x): R^n \rightarrow R^n$, 表示保持 x 对应于训练样本集 A 的第 j 类样本分量,将其他对应于非 j 类样本的分量置为 0,则 $\hat{y}_j = A\delta_j(x^*)$ 表示第 j 类训练样本生成的重构样本。令 $r_j(y) = \|y - \hat{y}_j\|_2$ 表示训练样本 y 的第 j 类重构残差,将测试样本 y 归类为最小残差的类别 j^* ,即:

$$j^* = \arg \min_j r_j(y), j = 1, 2, \dots, c \quad (2)$$

2 基于同类测试样本组的稀疏表示分类

给定测试样本集 $Y = [y_1, y_2, \dots, y_m] \in R^{d \times m}$, 将其

$[Y_1, Y_2, \dots, Y_g] \in R^{d \times m}$, g 表示组别数, $m = \sum_{i=1}^g m_i$, 其中 $Y_i = [y_1^i, y_2^i, \dots, y_{m_i}^i]$ 表示由 m_i 个样本组成的第 i 组同类测试样本集。给定 c 类训练样本集 $A = [A_1, A_2, \dots, A_c] = [v_{1,1}, v_{1,2}, \dots, v_{c,n_c}] \in R^{d \times n}$, 其中 $A_i = [A_{i,1}, A_{i,2}, \dots, A_{i,n_i}] \in R^{d \times n_i}$ 表示第 i 类训练样本集,因此得到基于同类测试样本组的稀疏表示模型:

$$X^* = \arg \min_X \|Y_i - AX\|_F^2 + \lambda \|X\|_1 \in R^{n \times m_i} \quad (3)$$

其中, X 为表示系数; $\|Y_i - AX\|_F^2$ 为损失函数; $\|X\|_F = (\sum_{i,j} x_{ij}^2)^{\frac{1}{2}}$ 表示 Frobenius 范数, x_{ij} 表示矩阵 X 第 i 行第 j 列元素; $\|X\|_1 = \sum_{i,j} |x_{ij}|$ 表示矩阵 L_1 -范数; $\lambda > 0$ 为正则化参数,用于平衡矩阵 X 的稀疏度和损失函数。

2.1 系数表示

事实上,记 $X = [x_1, x_2, \dots, x_{m_i}]$, $X^* = [x_1^*, x_2^*, \dots, x_{m_i}^*]$, 将式(3)中 Y_i, X, X^* 分列展开,可得到:

$$\begin{aligned} [x_1^*, x_2^*, \dots, x_{m_i}^*] = \\ \arg \min_X \| [y_1^i, y_2^i, \dots, y_{m_i}^i] - A[x_1, x_2, \dots, x_{m_i}] \|_F^2 + \\ \lambda \| [x_1, x_2, \dots, x_{m_i}] \|_1 \in R^{n \times m_i} \end{aligned} \quad (4)$$

进一步展开,即:

$$\begin{aligned} [x_1^*, x_2^*, \dots, x_{m_i}^*] = \\ \arg \min_{x_1, x_2, \dots, x_{m_i}} ([\|y_1^i - Ax_1\|_2^2 + \lambda \|x_1\|_1] + \dots + \\ [\|y_{m_i}^i - Ax_{m_i}\|_2^2 + \lambda \|x_{m_i}\|_1]) \end{aligned} \quad (5)$$

于是可以拆分成如下 m_i 个子向量 l_1 -范数最小化进行:

$$x_k^* = \arg \min_{x_k} \|y_k^i - Ax_k\|_2^2 + \lambda \|x_k\|_1 \in R^n, k = 1, 2, \dots, m_i \quad (6)$$

对于式(6)的求解,可以选择常见的 l_1 -范数最小化方法,然后将各个向量优化的结果合并成矩阵,即为式(3)的求解 $X^* = [x_1^*, x_2^*, \dots, x_{m_i}^*]$ 。

2.2 分类

令 $\delta_j(X): R^{n \times m_i} \rightarrow R^{n \times m_i}$ 表示保持 X 的对应第 j 类的分量,将其余非 j 类的分量置为 0; 令 $\hat{Y}_j^i = A\delta_j(X^*)$ 表示 Y_i 的第 j 类训练样本生成的重构样本组, $r_j(Y_i) = \|Y_i - \hat{Y}_j^i\|_F$ 表示 Y_i 的第 j 类重构残差。将 Y_i 判别为最小残差的类别 j^* ,即:

$$j^* = \arg \min_j r_j(Y_i), j = 1, 2, \dots, c \quad (7)$$

总结以上过程,即为:

算法 1: 基于同类测试样本组的稀疏表示分类(IGSRC)。

(1) 输入: 给定同类测试样本组 $Y_i \in R^{d \times m_i}$, c 类训练样本集 $A = [v_{1,1}, v_{1,2}, \dots, v_{c,1}, \dots, v_{c,n_c}] \in R^{d \times n}$, 正则

- 化参数 λ 。
- (2)单位化 \mathbf{A} 、 \mathbf{Y}_i 的列向量,使其具有 l_2 -范数单位长度。
- (3)求解以下矩阵 L_1 -范数最小化问题:
- $$\mathbf{X}^* = \arg \min_{\mathbf{X}} \|\mathbf{Y}_i - \mathbf{A}\mathbf{X}\|_F^2 + \lambda \|\mathbf{X}\|_1 \quad (8)$$
- (4)计算残差 $r_j(\mathbf{Y}_i) = \|\mathbf{Y}_i - \mathbf{A}\delta_j(\mathbf{X}^*)\|_F, j = 1, 2, \dots, c$ 。
- (5)输出:将 \mathbf{Y} 判别为类别 $j^* = \operatorname{argmin}_j r_j(\mathbf{Y}_i), j = 1, 2, \dots, c$ 。

3 数值实验

利用公开数据库对算法 IGSRC 和 SRC 进行比较,选择各自实验效果最佳的参数 λ ,报告实验结果。对于 SRC 模型的 l_1 -范数最小化,文献[14]给出的算法^[20-21]是化为二阶锥规划利用牛顿迭代求解(下面简称 L1QC),同时采用速度更快的 l_1 -范数最小化算法-ADMM^[22],所以同时比较上述两种优化方法得到的结

果,对应的 SRC 模型分别称之为 SRC(L1QC)、SRC(ADMM),而对于算法 IGSRC,采用 ADMM 算法求解式(6)。

3.1 无遮挡人脸识别

3.1.1 AR 数据库

选择 AR 数据库^[23]常用的子版本,包含 50 个男性和 50 个女性在光照变化、表情变化、有无遮挡下的每人 26 张 165×120 维面部图像,将其分成两部分:session1 和 session2,每部分每人包含 7 张无遮挡图像、3 张墨镜遮挡、3 张围巾遮挡,共计 13 张图像,如图 1 所示。记数据库中第 1 个男性为类别 1,第 2 个男性为类别 2……第 1 个女性为类别 51,依此类推,共 100 类。选择每人只有光照、表情变化的 14 张图像进行无遮挡数值实验,即将来自 session1 的前 7 张作为训练样本,session2 的前 7 张作为测试样本。将每张图像转化成灰度矩阵,同文献[14]的采样方式,将图像裁剪成 4 个特征维数:30、54、130、540,对应于采样率 1/24、1/18、1/12、1/6,并下拉成向量,进行向量 l_2 -范数单位化。

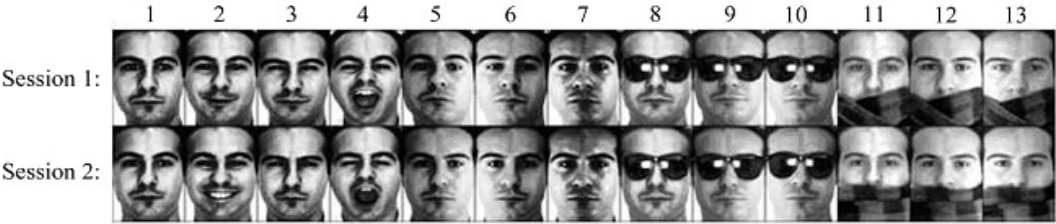


图 1 AR 数据库单个人的所有样本

实验结果如图 2(a)所示。

由图 2(a)可见,IGSRC 的识别率在 4 个特征维数下都比对应的 SRC(L1QC)、SRC(ADMM)更好,尤其是在较低维 30、54、130 的情况下,比 SRC(L1QC)和 SRC(ADMM)中的识别率较高者分别高出 9.43%、6.43%、7%。而在 540 维时,取 $\lambda = 10^{-3}$,IGSRC 取得了最高的识别率 96%,而相应的 SRC(ADMM),当 $\lambda = 8 \times 10^{-5}$ 时取得它最好的识别率 93%(当 $\lambda = 10^{-3}$ 时识别率为 91.57%)。

3.1.2 Extended Yale B 数据库

Extended Yale B 数据库^[24]包含 38 个人每人约 64 张共计 2 414 张人脸图像。图像是在实验室控制下的光照变化得到的,像素是 192×168 维。同样,将每张图像转化为灰度矩阵,同文献[14]的采样方式,将图像裁剪成 30、56、120、504 维,并下拉成列向量和进行 l_2 -范数单位化,随机取每个人约一半的图像作为训练集,剩下的为测试集,得到的结果如图 2(b)所示。

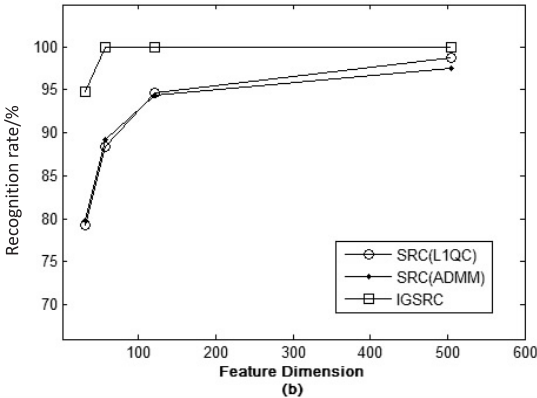
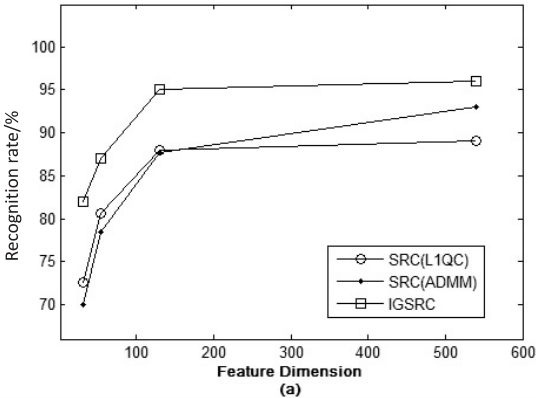


图 2 IGSRC 和 SRC 在 AR 与 Extended Yale B 上的识别率

由图 2(b)可见,IGSRC 的识别率仍然在所有维数中表现最好,尤其在低维(30、56 维)时超过 SRC 方法 10%,在 56 维及以上取得了 100% 的识别率。因为 Extended Yale B 较 AR 样本类别数少且样本个数多,类似的维数下识别率比 AR 高也是意料之中。

3.2 更少训练样本数的人脸识别

减少训练集每类样本数量以测试算法的性能。在 AR 数据库选择无遮挡人脸图像,训练样本从之前的每个人 7 张图像以 1 张递减逐步减少到 1 张,而测试集与之前一致,仍为每人 7 张图像。所有实验图像都

表 1 在 AR 数据库上不同的训练样本数下的识别率

算法	6	5	4	3	2	1
SRC(ADMM)	90.97%±2.25%	87.86%±2.71%	83.24%±4.65%	79.07%±4.38%	66.46%±11.10%	46.56%±14.48%
IGSRC	96%±0%	96%±0.85%	95%±0.86%	94.7%±0.70%	90.9%±1.67%	66.3%±25.79%

3.3 有遮挡人脸识别

为了测试 IGSRC 的鲁棒性,将 AR 数据库中包含遮挡的人脸图像纳入进行数值实验。(1)全混合:选择 session1 中 100 个人每人 13 张图像(包含 6 张遮挡图像,其中 3 张佩戴墨镜、3 张佩戴围巾,见图 1)作为训练集,选择 session2 中所有 13 图像(包含 6 张遮挡)作为测试集。(2)半混合:选择 session1 的 100 个人每人 7 张无遮挡图像作为训练集,session2 中每人所有 13 张图像(包含 6 张遮挡)作为测试集。(3)无混合:选择 session1 的 100 个人每人 7 张无遮挡图像作为训练集,session2 中每人 6 张遮挡图像作为测试集。这三种选择使得识别任务的挑战性逐步增大,训练集中从所有样本到仅有无遮挡图像,测试集从所有样本到仅有遮挡图像。所有实验图像都采样成 540 维,其他处理方式同 3.1.1。具体的数值实验结果如表 2 所示。

在 3 种混合选择中,IGSRC 依然比 SRC 在识别率上的表现更好,在“全混合”和“半混合”中,IGSRC 基本不受遮挡影响,识别率高达 96% 以上。而在条件最严苛的“无混合”中,IGSRC 的识别率为 66%,仍然远高于 SRC(36.17%)。说明在训练样本和测试样本中存在相同“类型”(同含有遮挡或无遮挡)时,IGSRC 的鲁棒性较高。

表 2 在 AR 数据库有遮挡下的识别率 %

算法	全混合	半混合	无混合
SRC(ADMM)	90.31	66.86	36.17
IGSRC	97	96	66

3.4 时间对比

SRC 是对单个测试样本逐一进行系数表示、分类,IGSRC 是基于同类测试样本组进行整体表示、整体分类,如果训练样本数量和同组测试样本数量越多,IGSRC 理应有更快的识别(系数表示+分类)时间。在

采样成 540 维,其他处理方式同 3.1.1。

表 1 显示了当每个人的训练样本数为 1~6 时 10 次随机置换训练样本后识别率的均值和相对标准差。在每人训练样本数为 2 个及以上,IGSRC 都取得了较高的识别率均值(90% 以上),相对标准差只有对应的 SRC 方法的 1/6~1/3,SRC 在训练样本数低于 4 时识别率均值已不足 80%。表明 IGSRC 较 SRC 的识别率和稳定性更佳。而在每人训练样本数为 1 时,两种方法表现都欠佳。

Intel(R) Core(TM) i5-2400 3.10 GHz CPU, 4 GB RAM 的个人电脑 Matlab R2012b 上,按 3.1.1 与 3.1.2 无遮挡的选择方式(其中仅选择 AR 数据库图像维数为 540 维,Extended Yale B 图像维数为 504 维),计算 IGSRC 和 SRC 对所有测试样本(AR 数据库 700 个,Extended Yale B 约 1 200 个)的识别时间,实验重复进行 10 次,结果取均值与相对标准差,如表 3 所示。

表 3 在 AR 和 Extended Yale B 上对所有测试样本的识别时间 s

算法	AR 数据库	Extended Yale B 数据库
SRC(ADMM)	35.22%±0.20%	49.11%±9.71%
IGSRC	32.47%±0.34%	38.58%±1.01%

由表 3 可见,IGSRC 比 SRC 具有更佳的识别时间及稳定性,同时在 Extended Yale B 中由于测试样本组中包含的样本数(约 32 个)比 AR 测试样本组包含样本数(7 个)更多,较 SRC 的时间优势体现得更明显,对约 1 200 个测试样本识别时间提升 10 s 左右。

4 结束语

SRC 及其改进方法基本上是基于单样本进行表示分类。而在许多应用场景中,同类的测试样本往往有多个。为充分利用这一信息,采用 IGSRC 对人脸进行识别,将已知同类测试样本分至一组,通过矩阵 L_1 -范数整体系数表示和整体分类,使得一些单样本的类内较大差异被类内较小差异的样本所弥补,因而 IGSRC 算法获得了比基于向量 l_1 -范数的 SRC 更高的识别率、鲁棒性和更低的相对标准差。同时,IGSRC 较 SRC 计算耗时更低。但是,IGSRC 对测试样本组内的测试样本“质量”要求较高,若算法中类内差异较大的样本造成的影响大于类内较小差异的样本,则整个测试样本组将被判错,但能容纳类内差异较大样本也是

IGSRC 的优势,如何平衡二者的影响,是进一步研究的方向。

参考文献:

- [1] Turk M A, Pentland A P. Face recognition using eigenface [C]//Proceedings of IEEE conference on computer vision & pattern recognition. [s. l.]:IEEE,2011.
- [2] Belhumeur P N, Hespanha J P, Kriegman D J. Eigenfaces vs. fisherfaces: recognition using class specific linear projection [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 1997, 19(7):711-720.
- [3] Lades M, Vorbruggen J C, Buhmann J, et al. Distortion invariant object recognition in the dynamic link architecture [J]. IEEE Transactions on Computers, 1993, 42(3):300-311.
- [4] Liu C. Capitalize on dimensionality increasing techniques for improving face recognition grand challenge performance [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2006, 28(5):725-737.
- [5] Ahonen T, Hadid A, Pietikäinen M. Face recognition with local binary patterns [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2006, 28(12):2037-2041.
- [6] Duda R O, Hart P E, Stork D G. Pattern classification [M]. [s. l.]:John Wiley & Sons,2012.
- [7] Li S Z, Lu J. Face recognition using the nearest feature line method [J]. IEEE Transactions on Neural Networks, 1999, 10(2):439-443.
- [8] Chien J T, Wu C C. Discriminant waveletfaces and nearest feature classifiers for face recognition [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2002, 24(12):1644-1649.
- [9] Ho J, Yang M H, Lim J, et al. Clustering appearances of objects under varying illumination conditions [C]//IEEE computer society conference on computer vision and pattern recognition. [s. l.]:IEEE,2003.
- [10] Lee K C, Ho J, Kriegman D J. Acquiring linear subspaces for face recognition under variable lighting [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2005, 27(5):684-698.
- [11] Donoho D L. For most large underdetermined systems of linear equations the minimal L1-norm solution is also the sparsest solution [J]. Communications on Pure & Applied Mathematics, 2007, 59(6):797-829.
- [12] Candès E J, Romberg J K, Tao T. Stable signal recovery from incomplete and inaccurate measurements [J]. Communications on Pure & Applied Mathematics, 2006, 59(8):1207-1223.
- [13] Candès E J, Tao T. Near-optimal signal recovery from random projections: universal encoding strategies? [J]. IEEE Transactions on Information Theory, 2007, 52(12):5406-5425.
- [14] Wright J, Yang A Y, Ganesh A, et al. Robust face recognition via sparse representation [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2008, 31(2):210-227.
- [15] 闫德勤, 刘胜蓝, 李燕燕. 一种基于稀疏嵌入分析的降维方法 [J]. 自动化学报, 2011, 37(11):1306-1312.
- [16] 胡正平, 宋淑芬. 基于类别相关近邻子空间的最大似然稀疏表示鲁棒图像识别算法 [J]. 自动化学报, 2012, 38(9):1420-1427.
- [17] Zhang L, Yang M, Feng X. Sparse representation or collaborative representation: which helps face recognition? [C]//International conference on computer vision. [s. l.]:IEEE Computer Society, 2011:471-478.
- [18] Deng W, Hu J, Guo J. Extended SRC: undersampled face recognition via intraclass variant dictionary [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2012, 34(9):1864-1870.
- [19] Cai S, Zhang L, Zuo W, et al. A probabilistic collaborative representation based approach for pattern classification [C]//IEEE conference on computer vision and pattern recognition. [s. l.]:IEEE,2016.
- [20] Boyd S, Vandenberghe L. Convex optimization [M]. [s. l.]:Cambridge Univ. Press,2004.
- [21] Candès E, Romberg J. l1-magic: recovery of sparse signals via convex programming [R]. [s. l.]:[s. n.],2005.
- [22] Boyd S, Parikh N, Chu E, et al. Distributed optimization and statistical learning via the alternating direction method of multipliers [J]. Foundations & Trends® in Machine Learning, 2011, 3(1):1-122.
- [23] Martinez A R, Benavente R. The AR face database [R]. [s. l.]:[s. n.],1998.
- [24] Georgiades A S, Belhumeur P N, Kriegman D J. From few to many: illumination cone models for face recognition under variable lighting and pose [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2001, 23(6):643-660.