

基于 $L_{2,p}$ 矩阵范数稀疏表示的图像分类方法

时中荣¹, 王 胜², 刘传才¹

(1.南京理工大学 计算机科学与工程学院, 江苏 南京 210094;
2.河南大学 图像处理与模式识别研究所, 河南 开封 475004)

摘要: 为了提高基于稀疏表示分类算法的分类精度, 该文充分利用同类样本的非零系数高度集中的特点, 提出一种用 $L_{2,p}$ 矩阵范数进行稀疏约束的基于稀疏表示的分类方法。该算法的训练阶段, 构造的目标函数主要包括三个部分: 重构误差、稀疏矩阵类内一致性约束、稀疏矩阵类间不一致性约束, 其中的稀疏矩阵类内一致性约束用 $L_{2,p}$ 矩阵范数实现。该算法的测试阶段, 计算新样本的稀疏重构系数以用于分类。和传统的基于稀疏表示的分类方法比较, 该方法求稀疏重构系数时对样本不再单个处理, 而是对同类样本整体处理, 且充分利用同类样本的相似性和不同类样本的相异性, 提高了基于稀疏表示的图像分类方法的分类精度。实验结果表明: 该方法进一步提高了图像分类的准确率, 在 AR、Extended Yale B 和 Fifteen Scene Category 数据库上和基于稀疏表示的分类方法(Sparse representation based classification, SRC) 相比较, 识别率分别提高了 20.11%、20.88% 和 2.13%。

关键词: 图像分类; 稀疏表示; 稀疏分类; 矩阵范数; 稀疏编码; 字典学习; 稀疏正则项; 稀疏诱导范数

中图分类号: TP391.41 **文章编号:** 1005-9830(2017)01-0080-10

DOI: 10.14177/j.cnki.32-1397n.2017.41.01.011

Sparse representation via $L_{2,p}$ norm for image classification

Shi Zhongrong¹, Wang Sheng², Liu Chuancai¹

(1.School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China;

2.Institute of Image Processing and Pattern Recognition, Henan University, Kaifeng 475004, China)

Abstract: For the sparse representation-based classification method, since the non-zero elements of sparse coefficients with the same class are concentrated in a few rows, we propose to regularize the

收稿日期: 2016-12-13 修回日期: 2016-12-30

基金项目: 国家自然科学基金(61373063)

作者简介: 时中荣(1979-), 女, 博士生, 主要研究方向: 模式识别、稀疏分类, E-mail: shizrong@163.com; 通讯作者: 刘传才(1963-), 男, 博士, 教授, 博士生导师, 主要研究方向: 模式识别、图像处理, E-mail: chuancailiu@njust.edu.cn。

引文格式: 时中荣, 王胜, 刘传才. 基于 $L_{2,p}$ 矩阵范数稀疏表示的图像分类方法[J]. 南京理工大学学报, 2017, 41(1): 80-89.

投稿网址: <http://zxuebao.njust.edu.cn>

coefficient matrix using an $L_{2,p}$ matrix norm. In the training phase of the algorithm, the objective function consists of three parts: reconstruction error, sparse regularization, and inconsistency of reconstruction coefficients between different classes. The sparse regularization term is implemented by an $L_{2,p}$ matrix norm. In the test phase, the sparse reconstruction coefficient of a new sample is found using the dictionary learned in the training phase. Finally, the new sample is classified according to the sparse reconstruction coefficient. Compared with the traditional classification method based on sparse representation, the proposed method does not process a single sample to find its sparse reconstruction coefficient, but the whole sample matrix can be processed, this takes full advantage of the similarity among the same class. The experimental results show that this method can improve the accuracies of image classification 20.11%, 20.88%, and 2.13% compared with a baseline SRC (Sparse representation based classification) method in AR, Extended Yale B, and Fifteen Scene Category databases respectively. This method makes full use of the similarity of the same class and improves the accuracy of the image classification based on sparse representation.

Key words: image classification; sparse representation; sparse classification; matrix norm; sparse coding; dictionary learning; sparse regularization; sparse-inducing norm

在很多领域例如人脸识别^[1]、行为识别^[2]、图像检索^[3]、超分辨率图像重建^[4]和生物信息^[5]等领域,图像分类技术取得了显著的成果。图像分类技术通常包含两个主要步骤:特征抽取和分类。近 10 年来,基于稀疏表示的图像分类技术得到了长足的发展。Wright 等人提出的基于稀疏表示的分类算法^[1](Sparse representation based classification, SRC) 获得了广泛的关注。SRC 的主要思想是利用训练样本的一个线性组合拟合测试样本。因为全部训练样本构成了一个过完备的字典,所以这个线性组合的系数具有稀疏性,且非零系数主要集中在同类训练样本对应的位置上。SRC 算法的分类准则是根据这个稀疏系数计算各类的最小重构误差,取得最小重构误差的类别即可作为预测类别。SRC 没有一个显式的特征抽取过程,它用稀疏系数进行后续的分类操作,所以在基于稀疏表示的分类方法中关键是稀疏系数的计算。

SRC 直接利用训练样本做为过完备字典,然而很多研究^[6-10]表明,当字典大小相同时从训练样本中学习出一个鉴别字典会取得更好的分类效果。K-SVD 算法^[7]利用奇异值分解从训练样本中学习一个字典,使得字典学习方法有了进一步的发展。K-SVD 是一种无监督字典学习方法,得到的字典原子没有标定信息且被所有类别样本共享。D-KSVD 字典学习算法^[8]是对 K-SVD 的进一步扩展,是一种有监督的字典学习方法,其在学

习字典的同时可以同步学习一个线性分类器。LC-KSVD 算法^[9]在构造的目标函数上增加一个标定一致性限制项,这一方法不仅利用了样本的标定信息,且利用了字典原子的标定信息,取得了很好的分类效果。Yang 等人提出的 FDDL 字典学习方法^[10]不仅将 Fisher 线性鉴别准则作用在稀疏系数上以增加它的鉴别性,还将类别子字典引入字典学习方法,增加了分类精度。近年来很多基于稀疏表示的图像分类方法^[11-18]被提出,这些方法在各领域得到广泛应用且提高了图像的分类精度。

在图像处理领域例如在人脸识别领域常常需要处理高维数据。在多数情况下,可以将不同类的数据视为是嵌入高维空间的低维流形,即同类数据聚类形成了低维的子空间。在稀疏表示领域,表现为同类样本的稀疏系数中非零元聚集在一起,聚集的位置是字典中的原子和样本来自同类的位置。

SRC 等方法主要利用 l_0 或者 l_1 范数做稀疏正则化。当字典中的原子来自不同类且形成组结构时, l_0 或者 l_1 范数不能充分满足这个组结构。组稀疏表示可以解决这一稀疏表示中的结构问题。GSC 算法^[19]把稀疏正则化的 l_1 范数扩展成 $l_{2,1}$ 范数,是对 SRC 算法的组稀疏表示的扩展。GSC 算法的字典根据原子的标定信息进行分组,测试样本用字典中尽量少的组来线性表示。稀疏表示问题已经从选择尽量少的原子转变为选择尽

量少的组。SRC 算法中字典原子是无标定信息的, GSC 算法中字典的原子按类别标签分组, 所以稀疏系数具有组结构。很多关于组稀疏表示的方法已经被提出且应用在不同的领域^[20-24]。

用于稀疏正则化的稀疏诱导范数, 包括 l_0 范数 l_1 范数 $l_{2,1}$ 范数, 这些都是应用于向量的向量范数。近些年, 在多视角学习、特征选择等领域的研究者提出用矩阵范数来做稀疏正则化。训练样本按列组成矩阵, 用矩阵范数来做稀疏约束。例如在特征选择领域 Nie 等人^[25]提出的用 $l_{2,1}$ 矩阵范数实现的鲁棒特征选择方法。同类数据按列组成矩阵, 矩阵的行向量对应某一特征。稀疏约束使它选择尽量少的非零行向量来实现特征选择的目的。文献[25]中的 $l_{2,1}$ 矩阵范数定义为行向量的 l_2 范数之和, 最小化 $l_{2,1}$ 范数的目的是选择尽量少的非零行向量。Peng^[26]和 Wang 等人^[27]提出用 $l_{2,p}$ 矩阵范数 ($0 < p < 1$) 做特征选择可以取得更稀疏的解。当 $0 < p < 1$ 时 $l_{2,p}$ 矩阵范数只是准范数, 不满足矩阵范数的三角不等式, 但本文中仍然称它为矩阵范数。

用向量范数做稀疏正则化时, 一次计算一个样本, 不能充分利用样本间的关系, 例如同类样本的结构相似性。用矩阵范数来做稀疏正则化, 则能利用这一结构信息。在本文中把同类样本的稀疏系数按列排列形成类稀疏系数矩阵, 此矩阵的非零元素主要集中于少数行向量, 这些行向量对应字典中和它同类的原子。本文用矩阵的 $l_{2,p}$ 范数 ($0 < p < 1$) 对这个矩阵做稀疏正则化, 约束此矩阵的非零元素主要集中于字典中同类原子对应的行向量上。这一正则项充分利用同类稀疏系数的结构相似性, 促使它们形成更相似的结构, 从而增加稀疏系数的鉴别信息, 提高后续分类算法的准确率。图 1 描述了理想状态下类稀疏系数矩阵的非零元位置。当每个子字典仅由本类样本组成时, 理想状态下非零元只聚集于和本类子字典对应的位置。用 $l_{2,p}$ 矩阵范数有利于实现这一理想结构。在增加同类样本稀疏系数的相似性同时, 本文亦提出增大不同类的类稀疏系数矩阵间的距离。

本文的贡献主要集中于以下 3 个方面:

(1) 为了充分利用样本的标定信息以利于后续的分类步骤, 训练样本按类别组成子矩阵, 系数也组成了类系数矩阵, 对这个类系数矩阵整体进行稀疏约束, 而不再是针对单个系数向量进行稀疏约束。

(2) 为了增大类系数矩阵的鉴别性, 稀疏约束项增加到两项, 即稀疏矩阵类内一致性约束和稀疏矩阵类间不一致性约束。前者可增加同类样本的相似性, 后者可减小不同类样本的相似性。

(3) 训练阶段的稀疏约束是矩阵范数, 但是测试阶段的稀疏约束是向量范数, 针对这个不一致性, 在测试阶段提出了和训练阶段等价的替代范数, 提高了分类算法的准确率。

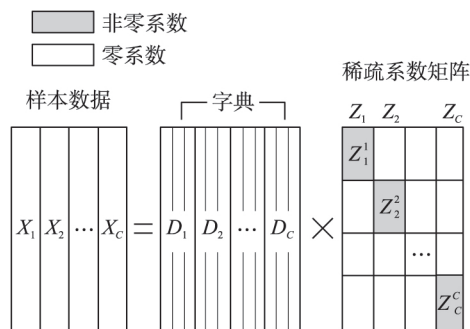


图 1 用块结构字典表示样本数据时稀疏系数非零元的理想分布

1 相关工作

假设有 C 类训练样本, 记为: $X = [X_1, X_2, \dots, X_C] \in \mathbf{R}^{m \times n}$, 其中第 i 类训练样本记为 $X_i = [x_{i,1}, x_{i,2}, \dots, x_{i,n_i}] \in \mathbf{R}^{m \times n_i}$, m 为特征空间的维数, n_i 为样本集 X_i 中的样本总数, 且满足 $n = \sum_{j=1}^C n_j$ 。字典记为 $D \in \mathbf{R}^{m \times K}$ (其中 $m \leq K$ 使得字典过完备)。给定一个新的测试样本 x , 为了判定它的类别需要计算它在字典 D 上的稀疏表示系数 z , 基于稀疏表示的分类方法关键是计算稀疏系数 z 的值。本节将简要介绍基于稀疏表示的分类方法 SRC 和关于 z 的一些稀疏正则化方法。

文中的符号表示: 加粗大写字母表示矩阵, 加粗小写字母表示向量, 非加粗字符是标量。例如, $D \in \mathbf{R}^{m \times K}$ 表示一个尺寸为 $m \times K$ 的矩阵, $D_{i,:}$ 和 $D_{:,j}$ 分别表示矩阵的第 i 行和第 j 列。

1.1 SRC 方法

SRC 是 Wright 等人提出的一种基于稀疏表示的分类方法, 它直接利用训练样本作为稀疏表示的字典, 不涉及字典的学习过程。SRC 方法主要包含以下两个步骤。首先给定一个测试样本 x 在训练样本组成的字典 $X = [X_1, X_2, \dots, X_C]$ 上它被稀疏表示为

$$\hat{z} = \operatorname{argmin}_z \|x - Xz\|_2^2 + \gamma \|z\|_1 \quad (1)$$

再根据每类子字典的重构残差做分类

$$\operatorname{identity}(x) = \operatorname{argmin}_i \|x - X_i \hat{z}_i\|_2$$

where $\hat{z} = [\hat{z}_1, \hat{z}_2, \dots, \hat{z}_C]^T$ (2)

由上述算法步骤可见, SRC 先寻找测试样本的稀疏表示, 再计算此稀疏表示在各类子字典上的残差, 残差最小的类即是它的类别。测试样本稀疏表示的非零系数比较集中, 主要集中在和它来自相同类的子字典对应的位置上, 所以用同类子字典拟合时的残差最小, 找到这个最小残差可以判断类别。

1.2 稀疏正则项

SRC 算法中的稀疏正则项采用的是 l_1 范数。对于任意向量 $z \in \mathbf{R}^K$, 用作稀疏正则化的 l_0 范数, l_1 范数, l_p 范数定义为

$$\|z\|_0 = \sum_{z_i \neq 0} |z_i| \quad (3)$$

$$\|z\|_1 = \sum_{i=1}^K |z_i| \quad (4)$$

$$\|z\|_p = \left(\sum_{i=1}^K |z_i|^p \right)^{\frac{1}{p}} \quad (5)$$

式中: $p \in (0, 1)$ 。 l_0 范数和 l_p 范数 ($0 < p < 1$) 都是准范数, 在稀疏表示的文献中一般仍称它们为范数。

为了适应向量的组稀疏结构, 向量的 $l_{2,1}$ 范数定义如下, 向量 $z \in \mathbf{R}^K$ 被分成 C 个不重叠的块, 记为 $z = [z_1, z_2, \dots, z_C] \in \mathbf{R}^K$, 则它的 $l_{2,1}$ 范数可以定义为

$$\|z\|_{2,1} = \sum_{i=1}^C \|z_i\|_2 \quad (6)$$

为了解决矩阵内在行向量上的稀疏性, 本文引入可做稀疏正则化的矩阵范数。对于任意矩阵 $Z \in \mathbf{R}^{K \times n}$, 可将矩阵的 $l_{2,0}$ 范数和 $l_{2,1}$ 范数定义为

$$\|Z\|_{2,0} = \sum_{\|Z_{i \cdot}\|_2 \neq 0} (\|Z_{i \cdot}\|_2)^0 \quad (7)$$

$$\|Z\|_{2,1} = \sum_{i=1}^K \|Z_{i \cdot}\|_2 \quad (8)$$

式中: $Z_{i \cdot}$ 表示矩阵的第 i 行。在特征选择领域, Peng^[26] 和 Wang 等人^[27] 提出可以泛化矩阵的 $l_{2,1}$ 范数, 从而得到矩阵的 $l_{2,p}$ 范数

$$\|Z\|_{2,p} = \left(\sum_{i=1}^K \|Z_{i \cdot}\|_2^p \right)^{\frac{1}{p}} \quad (9)$$

式中: $p \in (0, 1)$ 。相较于 $l_{2,p}$ 矩阵范数, $l_{2,0}$ 矩阵范数的解是 NP 问题, 且 $l_{2,1}$ 矩阵范数的解不够稀疏, 所以本文中用 $l_{2,p}$ 范数做矩阵的稀疏正则项, 提出

了基于 $l_{2,p}$ 矩阵范数稀疏表示的图像分类算法。

2 基于 $l_{2,p}$ 矩阵范数稀疏表示的分类方法

2.1 模型

$D \in \mathbf{R}^{m \times K}$ 表示一个 $m \times K$ 的字典, 有 C 类训练样本, 第 i 类训练样本记作 $X_i \in \mathbf{R}^{m \times n_i}$, 它的稀疏系数矩阵记为 $Z_i \in \mathbf{R}^{K \times n_i}$ 。

与 SRC 类似, 模型所构建的目标函数中用最小化误差平方和来做数据的拟合。因为同类样本的稀疏系数非零项集中于少数行向量, 所以第二项即稀疏矩阵类内一致性约束用 $l_{2,p}$ 矩阵范数实现, 模型表示为

$$\operatorname{argmin}_{Z_i} \|X_i - DZ_i\|_F^2 + \lambda_1 \|Z_i\|_{2,p} \quad (10)$$

因为

$$\operatorname{argmin}_{Z_i} \|Z_i\|_{2,p} = \operatorname{argmin}_{Z_i} \|Z_i\|_{2,p}^p \quad (11)$$

为了后续计算方便, 模型改写为

$$\operatorname{argmin}_{Z_i} \|X_i - DZ_i\|_F^2 + \lambda_1 \|Z_i\|_{2,p}^p \quad (12)$$

为了减小不同类稀疏系数的相似性, 在式 (12) 上增加第三项即稀疏矩阵类间不一致性约束。本文提出的基于 $l_{2,p}$ 矩阵范数稀疏表示的分类 (Sparse representation based classification via $L_{2,p}$ norm, SRC_L2p) 模型定义为

$$\operatorname{argmin}_{Z_i} \|X_i - DZ_i\|_F^2 + \lambda_1 \|Z_i\|_{2,p}^p + \lambda_2 \sum_{j=1, j \neq i}^C \|Z_i Z_j^T\|_F^2 \quad (13)$$

式中: 第一项为重构误差, 第二项为稀疏矩阵类内一致性约束, 第三项为稀疏矩阵类间不一致性约束。其中第二项由于 $l_{2,p}$ 范数的定义, 使得同类样本稀疏系数的非零项集中于少数相同行, 这一约束加强了同类样本稀疏系数之间的结构相似性; 第三项是不同类样本稀疏系数不一致性约束, 当不同类样本稀疏系数的非零项的位置完全错开没有重叠项时, 第三项的值达到最小, 这项约束使得不同类样本稀疏系数减小了结构相似性。第二项和第三项约束使得稀疏系数结构在类内更相似, 在类间更相异, 更有益于后续的分类操作。第二项选择 $l_{2,p}$ 矩阵范数 ($0 < p < 1$) 是因为同类别样本的稀疏系数的非零项集中于少数行向量位置, 且 $l_{2,p}$ 范数能获得比 $l_{2,1}$ 范数更稀疏的解。

因为 $l_{2,p}$ 矩阵范数是非光滑和非凸的, 所以式 (13) 中寻找稀疏系数的最优解是一个困难的问题。

2.2 优化

2.2.1 稀疏编码

在文献 [26] 中, Peng 等人提出 $l_{2,p}$ 范数等价式 (14)

$$\|A\|_{2,p}^p = \|\Sigma A\|_F^2 \quad (14)$$

式中: Σ 为一个对角矩阵, 它的第 i 个对角线元素为 $\Sigma_{ii} = \frac{1}{\|A_{i\cdot}\|_2^{1-p/2}}$ 。因此, 可将目标函数改写为

$$\begin{aligned} \|X_i - DZ_i\|_F^2 + \lambda_1 \|\Sigma_i Z_i\|_F^2 + \lambda_2 \sum_{j=1, j \neq i}^C \|Z_i Z_j^T\|_F^2 = \\ \text{trace} \left[(X_i - DZ_i)^T (X_i - DZ_i) + \lambda_1 (\Sigma_i Z_i)^T \Sigma_i Z_i + \right. \\ \left. \cdots + \lambda_2 \sum_{j=1, j \neq i}^C (Z_i Z_j^T)^T Z_i Z_j^T \right] \end{aligned} \quad (15)$$

式中

$$\Sigma_i = \text{diag} \left(\frac{1}{\|(Z_i)_{1\cdot}\|_2^{1-2/p}}, \cdots, \frac{1}{\|(Z_i)_{K\cdot}\|_2^{1-2/p}} \right)$$

式 (15) 是关于 Σ_i 和 Z_i 的函数, 优化算法采用交替迭代优化的方法: 先固定 Σ_i 求 Z_i , 再固定 Z_i 求 Σ_i 。

当已知 Σ_i 时, 令式 (15) 关于 Z_i 的导数为零, 可得到

$$(D^T D + \lambda_1 \Sigma_i \Sigma_i + \lambda_2 \sum_{j=1, j \neq i}^C Z_j Z_j^T) Z_i - D^T X_i = 0 \quad (16)$$

整理后, 获得稀疏系数 Z_i 的表达式

$$Z_i = (D^T D + \lambda_1 \Sigma_i \Sigma_i + \lambda_2 \sum_{j=1, j \neq i}^C Z_j Z_j^T)^{-1} D^T X_i \quad (17)$$

当固定 Z_i 时, Σ_i 的值如前文中它的定义。

式 (17) 可以得到稀疏编码的闭合解, 和凸优化、匹配追踪算法等稀疏编码算法相比较, SRC_L2p 算法更快速有效, 可以得到稳定的解。

2.2.2 字典学习

在本文的基于 $l_{2,p}$ 矩阵范数稀疏表示的分类方法中, 字典是未知的, 需要从训练样本中学习得到。已知稀疏系数矩阵 $Z = [Z_1 \ Z_2 \ \cdots \ Z_C]$ 学习模型中的字典 D 时, 由式 (13) 可得关于字典的目标函数

$$\begin{aligned} D = \argmin_D \|X - DZ\|_F^2 \\ \|D_{\cdot i}\|_2 \leq 1 \end{aligned} \quad (18)$$

式中: $\|D_{\cdot i}\|_2 \leq 1$ 是为了避免稀疏系数的平凡解。式 (18) 可以用常用的字典学习方法求解, 例如拉格朗日对偶法^[28] 或者 K-SVD 方法^[7]。

2.3 分类算法

在 SRC_L2p 算法的训练阶段, 有两个未知变量即字典和稀疏系数。和常用的字典学习方法类

似, 算法采用固定字典求稀疏系数, 再固定稀疏系数求字典的循环优化算法。求解稀疏系数矩阵时, 需要按类别求解 C 个类的稀疏系数矩阵。字典的学习可以采用拉格朗日对偶法。

在 SRC_L2p 算法的测试阶段, 因为测试样本的类别未知, 不能像训练阶段分类别划分训练样本形成矩阵且用矩阵的 $l_{2,p}$ 范数做稀疏正则化。为了和训练阶段保持一致性, 本文提出利用训练阶段学习得到的对角矩阵 Σ_i ($i=1, 2, \cdots, C$) 对测试样本 x 进行编码。把测试样本分别编码进 C 个类, 则它在 C 个类的稀疏重构系数由下式可得

$$\begin{aligned} \hat{z}_i = \argmin_{z_i} \|x - Dz^i\|_2^2 + \lambda_1 \|\Sigma_i z^i\|_2^2 \\ i=1, 2, \cdots, C \end{aligned} \quad (19)$$

最终分类的判断准则如下

$$\begin{aligned} \text{identity}(x) = \argmin_i \|x - D\hat{z}_i\|_2^2 + \lambda_1 \|\Sigma_i \hat{z}_i\|_2^2 \\ i=1, 2, \cdots, C \end{aligned} \quad (20)$$

即具有最小重构误差和稀疏性的类别即是测试样本 x 的类别。

SRC_L2p 算法如算法 1 所示。在实验中, 发现此算法能很快收敛到一个局部极小值解。

算法 1 基于矩阵范数稀疏表示的分类 (SRC_L2p) 算法。

输入: 训练样本集 $X = [X_1 \ X_2 \ \cdots \ X_C] \in \mathbf{R}^{m \times n}$, 测试样本 x 。

输出: 字典 D , Σ_i ($i=1, 2, \cdots, C$), 测试样本 x 的类别。

步骤 1 初始化字典 D 为训练样本。

步骤 2 字典学习

Repeat

稀疏编码:

Repeat

(1) 分类别计算

$$\Sigma_i = \text{diag} \left(\frac{1}{\|(Z_i)_{1\cdot}\|_2^{1-2/p}}, \cdots, \frac{1}{\|(Z_i)_{K\cdot}\|_2^{1-2/p}} \right) \\ i=1, 2, \cdots, C。$$

(2) 分类别计算稀疏系数矩阵

$$Z_i = (D^T D + \lambda_1 \Sigma_i \Sigma_i + \lambda_2 \sum_{j=1, j \neq i}^C Z_j Z_j^T)^{-1} D^T X_i \\ i=1, 2, \cdots, C。$$

Until convergence

字典更新:

用拉格朗日对偶法^[28] 学习字典。

Until convergence

步骤 3 对于测试样本 x , 分别计算它在 C 个类的稀疏系数 $\hat{z}_i = \argmin_{z_i} \|x - Dz^i\|_2^2 + \lambda_1 \|\Sigma_i z^i\|_2^2$ ($i=1, 2, \cdots, C$)。

步骤 4 计算 x 的类别标签 $\text{identity}(x) = \argmin_i \|x - D\hat{z}_i\|_2^2 + \lambda_1 \|\Sigma_i \hat{z}_i\|_2^2$ ($i=1, 2, \cdots, C$)。

3 实验与结果

3.1 实验与结果

本文用两个不同的图像分类任务即人脸识别和景物分类来评估提出的 SRC_L2p 算法。人脸识别用 AR 和 Extended Yale B 数据库,景物分类用 Fifteen Scene Category 数据库。三个数据库的部分图像分别如图 2、图 3 和图 4 所示。在所有实验里,随机选择的训练样本作为初始化的字典。



图 2 AR 数据库



图 3 Extended Yale B 数据库



图 4 Fifteen Scene Category 数据库

AR 数据库包含超过 4 000 幅人脸图像,库中有 126 人,每人有 26 幅图像。人脸图像如图 2 所示,AR 数据库主要展示了人脸的表情、光照、遮挡的变化。实验中采用通常的 AR 数据库设置,选择一个 AR 的子集作为实验数据集,子集包含 50 位男性和 50 位女性的人脸图像。对于每个人,随机选择 20 幅图像做训练,余下的 6 幅图像做测试。在实验准备阶段 AR 人脸图像用 Randomface^[1]方法投影到 540 维的低维空间。学习到的字典包含 500 个原子,每个人对应 5 个原子。

Extended Yale B 数据库包含有 2432 幅人脸图像,共有 38 个人,每人 64 幅人脸图像。如图 3 所示,数据库中的人脸图像是在各种光照条件下获得,变化的光照和表情是这个数据库的主要挑战。和 AR 数据库类似,本文也用 Randomface^[1]方法把每幅人脸图像投影到 132 维的低维空间。每个人随机选择一半图像即 32 幅做训练,余下图像做测试。学习到的字典包含 342 个原子,每人对应 9 个原子。

Fifteen Scene Category 数据库包含 15 个类别的景物图像,包括高速公路、森林、高楼、卧室等。图 4 展示了 15 类的每类中一幅图像。每类图像包含 200 到 400 幅不等的图像。在实验准备阶段,采用了 Jiang^[9]等抽取的 3000 维 Spatial Pyramid features 特征作为实验数据。每类随机选择 100 幅图像做训练,余下的做测试。学习的字典有 450 个原子,每类对应 30 个原子。

本文的 SRC_L2p 算法和 SRC^[1]、K_SVD^[7]、D_KSVD^[8]、LC_KSVD^[9]、FDDL^[10]和 SVGDL^[29]等经典的基于稀疏表示的分类算法进行比较,实验结果都是在 6 次随机地划分训练和测试样本上取平均。三个数据库上的实验结果如表 1 所示。可以发现 SRC_L2p 算法和其他算法比较,在识别率上有显著的提升。

表 1 AR、Extended Yale B、Fifteen Scene Category 三个数据库上的识别率

算法	AR/%	Extended Yale B/%	Fifteen Scene Category/%
SRC	77.34	74.47	96.21
K-SVD	89.41	78.25	93.60
D-KSVD	90.60	80.59	96.11
LC-KSVD1	91.78	81.41	94.05
LC-KSVD2	92.50	82.50	97.01
FDDL	93.01	84.62	97.04
SVGDL	95.80	90.40	98.01
SRC_L2p	97.45	95.35	98.34

3.2 收敛性

SRC_L2p 算法循环优化两个变量:稀疏系数矩阵 Z 和字典 D 。Peng 等人在文献[26]中证明了在式(14)中用 F 范数代替 $L_{2,p}$ 范数时算法的收敛性,所以 SRC_L2p 算法在求解稀疏系数矩阵 Z 时可以得到局部最小解。在更新字典时,拉格朗日对偶法^[28]可以保证目标函数的值是逐步递减的。综上所述, SRC_L2p 算法是收敛的。如图 5

(ac) 所示, 它们分别记录了式(13)中目标函数值在三个数据库上随着迭代次数增加而逐步递减的过程, 可以看出 SRC_L2p 算法是收敛的。

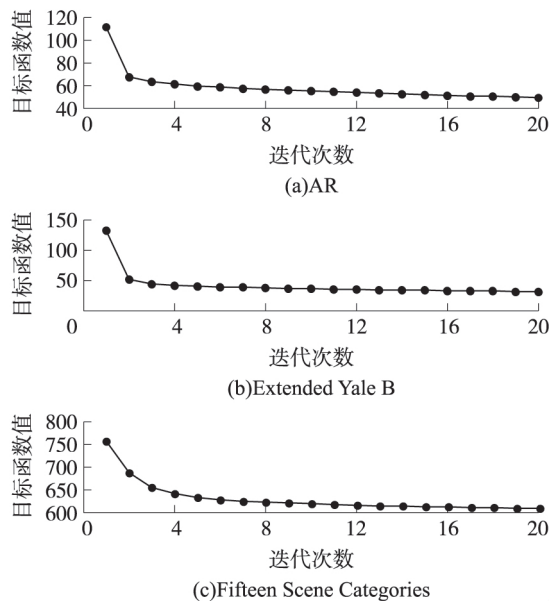


图 5 SRC_L2p 算法在数据库上的收敛性

3.3 参数分析

SRC_L2p 算法有三个参数 λ_1 、 λ_2 和 p , 其中 λ_1 和 λ_2 分别调节稀疏矩阵类内一致性约束项和稀疏矩阵类间不一致性约束项, p 的取值范围是 $p \in (0, 1)$, p 的值越靠近 0, 系数矩阵越稀疏。

图 68 分别展示了在 AR、Extended Yale B 和 Fifteen Scene Category 数据库上三个参数对于分类精度的影响。每幅图中的 (af) 分别展示了当 p 的值取 0.1、0.3、0.5、0.7、0.9 和 1.0 时 λ_1 和 λ_2 对识别率的影响。图中的圆点分别记录了在每个 p 取值下的最高识别率。例如在图 6(a) 中圆点表明: 在 AR 数据库上, 在 $p=0.1$ $\lambda_1=0.001$ 和 $\lambda_2=0.01$ 时, 取得最高识别率 98.667%。综合图 68 中得到: AR 数据库中的最高识别率是 98.833% 在 $p=0.7$ $\lambda_1=0.1$ 和 $\lambda_2=0.1$ 时取得; Extended Yale B 数据库中的最高识别率是 96.628% 在 $p=0.3$ 或者 $p=0.5$ $\lambda_1=0.1$ 和 $\lambda_2=0.01$ 时取得; Fifteen Scene Category 数据库中的最高识别率是 98.796% 在 $p=0.1$ $\lambda_1=0.01$ 和 $\lambda_2=0.1$ 时取得。

为了取得这些最高识别率, 三个数据库上的参数值都是不同的。在实验里为了减少参数设置的复杂性, 可以寻找共同的参数设置。观察图 68 可得到, 识别率对参数 λ_1 和 λ_2 的取值更敏感。为了避免识别率过大波动, 选取对 λ_1 和 λ_2 取值变化不太敏感且有较高识别率的区域。经过比较, 设置了共同的参数值: $\lambda_1=0.1$ 、 $\lambda_2=0.1$ 和 $p=0.7$ 。例如在表 1 中, 三个数据库上的实验都使用了这组共同的参数值, 仍然取得了比其他稀疏表示分类方法更高的识别率。

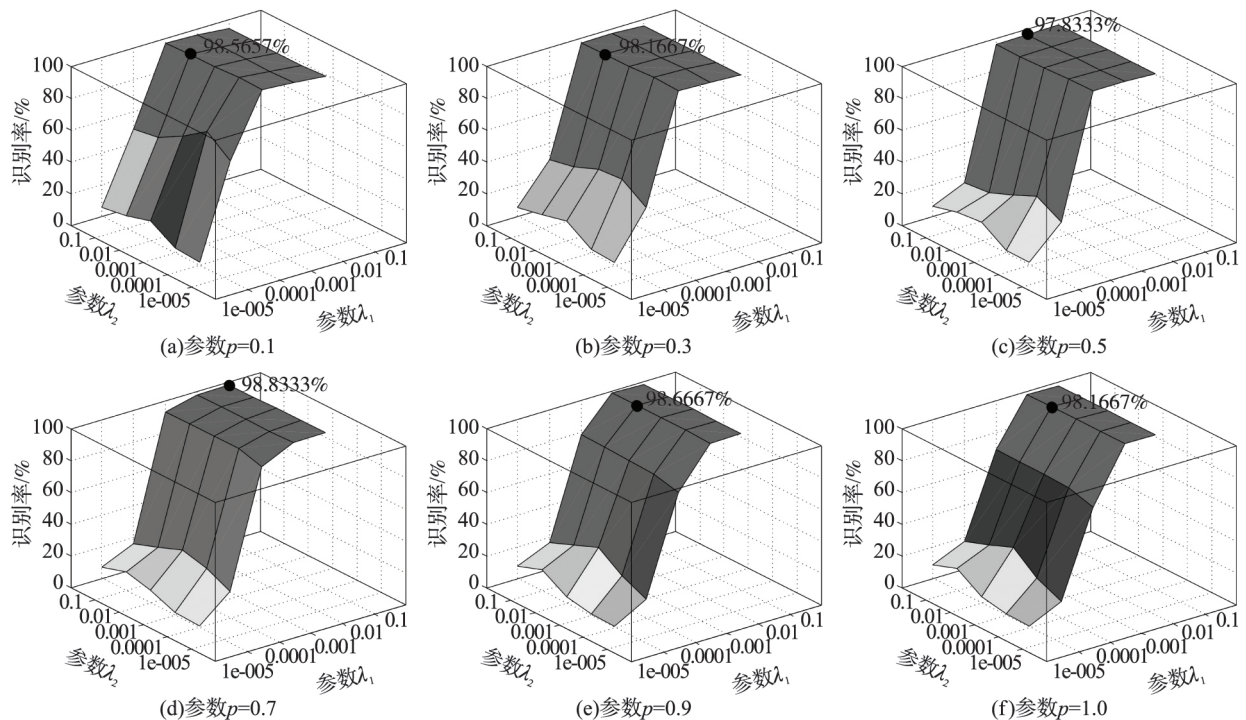
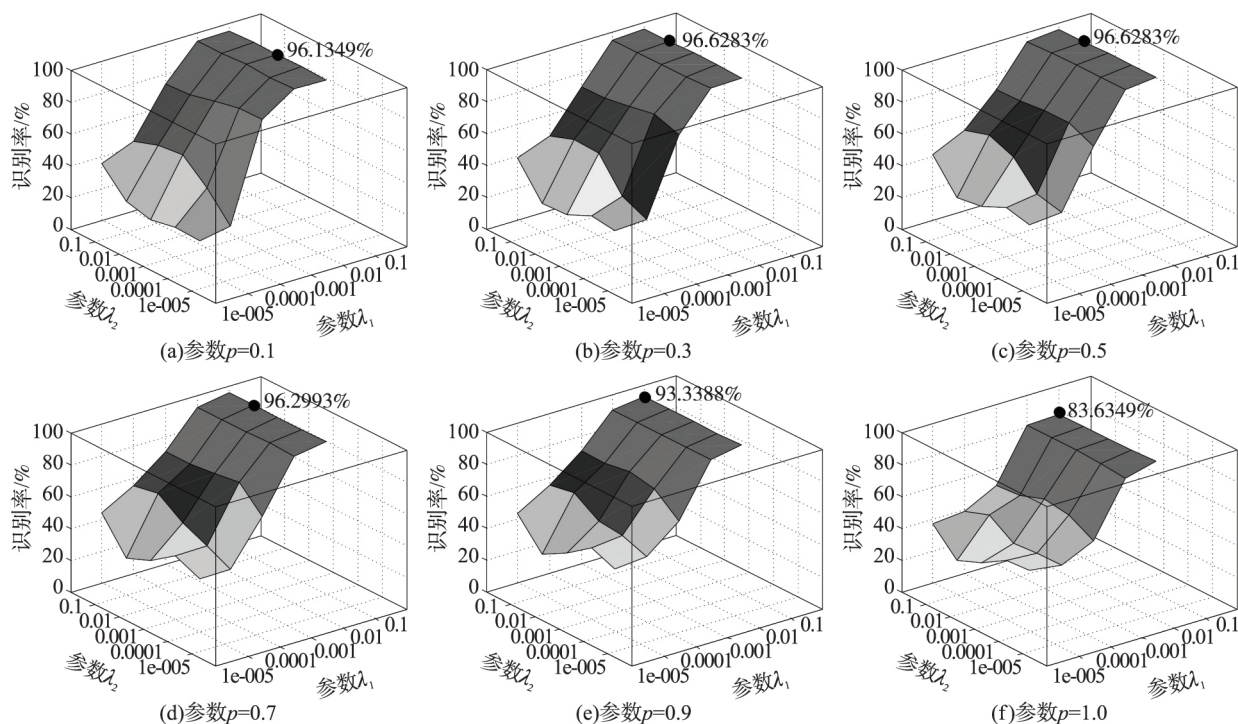
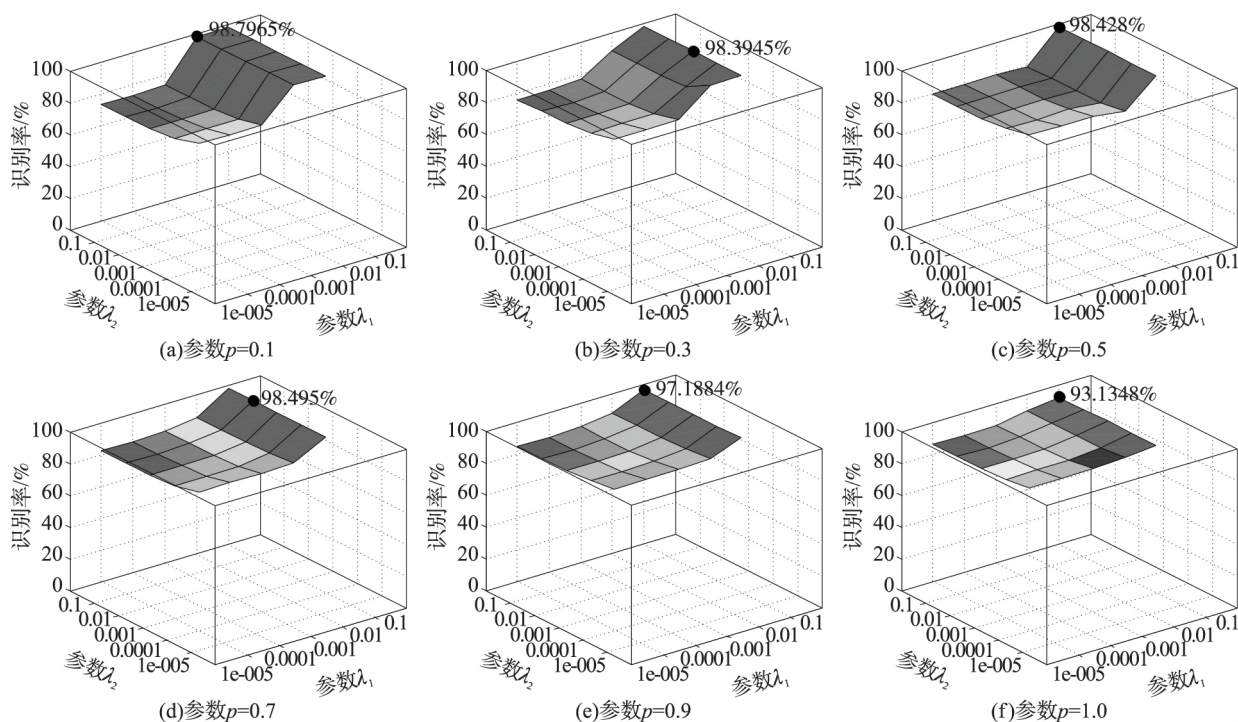


图 6 参数 λ_1 、 λ_2 和 p 在 AR 数据库上对识别率的影响

图 7 参数 λ_1 、 λ_2 和 p 在 Extended Yale B 数据库上对识别率的影响图 8 参数 λ_1 、 λ_2 和 p 在 Fifteen Scene Category 数据库上对识别率的影响

当 $p=1.0$ 时 $l_{2,p}$ 范数退化为 $l_{2,1}$ 范数, 也即是在图 6(f)、图 7(f) 和图 8(f) 中记录了用 $l_{2,1}$ 矩阵范数稀疏正则化的识别率。在三个数据库上 $l_{2,1}$ 范数的最高识别率分别是: 98.167%、83.635% 和 93.135% 比 $l_{2,p}$ 范数上的识别率都要低。最终的实验结果验证了 SRC_L2p 算法对于分类任务的

有效性。

4 结论

本文提出了基于 $l_{2,p}$ 矩阵范数稀疏表示的分类算法。此算法训练阶段的数据按类组成类矩

阵,用矩阵范数做稀疏约束。算法用 $l_{2,p}$ 矩阵范数做同类样本的稀疏约束,使得同类样本的非零项聚集在一起,增强了同类样本间的相似性。算法对不同类样本做不一致性约束,减少了不同类样本间的相似性。在算法测试阶段提出新的分类准则,避免了训练阶段和测试阶段求解稀疏系数方法时的一致,取得了更好的分类结果。实验结果表明提出的 SRC_L2p 算法在 AR、Extended Yale B 和 Fifteen Scene Category 三个数据库上取得了理想的识别精度,比常用的基于稀疏表示的分类方法 SRC 分别增加了 20.11%、20.88%、2.13% 的准确率。

参考文献:

- [1] Wright J, Yang A Y, Ganesh A, et al. Robust face recognition via sparse representation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence 2009, 31(2): 210–227.
- [2] Poppe R. A survey on vision-based human action recognition [J]. Image and Vision Computing 2010, 28(6): 976–990.
- [3] Datta R, Joshi D, Li J, et al. Image retrieval: ideas, influences, and trends of the new age [J]. ACM Computing Surveys 2008, 40(2): 5.
- [4] 刘梓, 宋晓宁, 於东军, 等. 基于多成分字典和稀疏表示的超分辨率重建算法 [J]. 南京理工大学学报: 自然科学版 2014, 38(1): 1–5.
Liu Zi, Song Xiaoning, Yu Dongjun, et al. Super-resolution reconstruction algorithm based on multi-component dictionary and sparse representation [J]. Journal of Nanjing University of Science and Technology, 2014, 38(1): 1–5.
- [5] Beaulieu-Jones B K, Greene C S. Semi-supervised learning of the electronic health record with denoising autoencoders for phenotype stratification [J]. bioRxiv: 039800 2016.
- [6] Mairal J, Bach F, Ponce J. Task-driven dictionary learning [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence 2012, 34(4): 791–804.
- [7] Aharon M, Elad M, Bruckstein A. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation [J]. IEEE Transactions on Signal Processing 2006, 54(11): 4311–4322.
- [8] Zhang Qiang, Li Baoxin. Discriminative K-SVD for dictionary learning in face recognition [C]//2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). San Francisco, USA: IEEE, 2010: 2691–2698.
- [9] Jiang Zhuolin, Lin Zhe, Davis L S. Label consistent K-SVD: Learning a discriminative dictionary for recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence 2013, 35(11): 2651–2664.
- [10] Yang Meng, Zhang Lei, Feng Xiangchu, et al. Sparse representation based fisher discrimination dictionary learning for image classification [J]. International Journal of Computer Vision 2014, 109(3): 209–232.
- [11] Yang Jianchao, Yu Kai, Gong Yihong, et al. Linear spatial pyramid matching using sparse coding for image classification [C]//2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Miami, USA: IEEE 2009: 1794–1801.
- [12] Do T H, Tabbone S, Terrades O R. Sparse representation over learned dictionary for symbol recognition [J]. Signal Processing 2016, 125: 36–47.
- [13] Lu Na, Miao Hongyu. Structure constrained nonnegative matrix factorization for pattern clustering and classification [J]. Neurocomputing 2016, 171: 400–411.
- [14] Lu Zhenyu, Zhang Linghua. Face recognition algorithm based on discriminative dictionary learning and sparse representation [J]. Neurocomputing, 2016, 174: 749–755.
- [15] Wu Fei, Jing Xiaoyuan, You Xinge, et al. Multi-view low-rank dictionary learning for image classification [J]. Pattern Recognition 2016, 50: 143–154.
- [16] Yuan Li, Liu Wei, Li Yang. Non-negative dictionary based sparse representation classification for ear recognition with occlusion [J]. Neurocomputing, 2016, 171: 540–550.
- [17] 宋晓宁, 薛益时, 於东军, 等. 一种优化恢复压缩传感矩阵的模糊自适应方法 [J]. 南京理工大学学报, 2013, 37(4): 479–485.
Song Xiaoning, Xue Yishi, Yu Dongjun, et al. Fuzzy adaptive method for optimizing recovery of compressive sensing matrix [J]. Journal of Nanjing University of Science and Technology 2013, 37(4): 479–485.
- [18] 张少辉, 王逸冉. 用于图像识别的稀疏高斯编码 [J]. 南京理工大学学报 2016, 40(1): 61–66.
Zhang Shaohui, Wang Yiran. Sparse Gaussian coding for image recognition [J]. Journal of Nanjing University of Science and Technology 2016, 40(1): 61–66.
- [19] Majumdar A, Ward R K. Classification via group sparsity promoting regularization [C]//IEEE International Conference on Acoustics Speech and Signal Processing. Taipei, Taiwan, China: IEEE 2009: 861–864.

- [20] Bach F ,Jenatton R ,Mairal J ,et al.Structured sparsity through convex optimization [J]. Statistical Science , 2012 27(4) : 450–468.
- [21] Bengio S ,Pereira F ,Singer Y ,et al. Group sparse coding [C]//Advances in Neural Information Processing Systems. Whistler , Canada: MIT Press , 2009: 82–89.
- [22] Chi Y T ,Ali M ,Rajwade A ,et al.Block and group regularized sparse modeling for dictionary learning [C]// 2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) . Oregon ,Portland: IEEE ,2013: 377–382.
- [23] Lei Hao ,Mei Kuizhi ,Zheng Nanning ,et al. Learning group-based dictionaries for discriminative image representation [J]. Pattern Recognition ,2014 ,47 (2) : 899–913.
- [24] Liu Baodi ,Wang Yuxiong ,Shen Bin ,et al. Blockwise coordinate descent schemes for efficient and effective dictionary learning [J]. Neurocomputing ,2016 ,178: 25–35.
- [25] Nie Feiping ,Huang Heng ,Cai Xiao ,et al.Efficient and robust feature selection via joint ℓ_2 , ℓ_1 -norms minimization [C]//Advances in Neural Information Processing Systems. Vancouver ,Canada: MIT Press ,2010: 1813–1821.
- [26] Peng Hanyang ,Fan Yong. Direct $\ell_{2,p}$ -norm learning for feature selection [J].arXiv preprint arXiv: 1504.00430 2015.
- [27] Wang Liping ,Chen Songcan.Joint representation classification for collective face recognition [J]. arXiv preprint arXiv: 1505.04617 2015.
- [28] Lee H ,Battle A ,Raina R ,et al.Efficient sparse coding algorithms [C]//Advances in Neural Information Processing Systems. Vancouver ,Canada: MIT Press ,2006: 801–808.
- [29] Cai Sijia ,Zuo Wangmeng ,Zhang Lei ,et al ,Support vector guided dictionary learning [C]//European Conference on Computer Vision. Zurich: Springer ,2014: 624–639.