

具有判别能力的低秩投影字典对学习

邱立达 傅平 林南 张宁

闽江学院物理学与电子信息工程系, 福建 福州 350001

摘要 与传统字典学习算法相比,新的投影字典对学习(DPL)算法在字典学习过程中引入了投影字典,利用投影编码替代目标样本在字典上的稀疏编码,有效降低了模式识别算法的计算量,但是原始的DPL算法对遮挡和噪声干扰较为敏感。为了解决这一问题,提出了一种具有判别能力的低秩投影字典对学习(DLPL)算法,该算法在模型中增加了对字典的低秩约束并利用最小二乘估计法对标签样本投影编码的分类误差进行约束,待求字典和投影字典都具有封闭形式的解,通过交替优化方法进行快速求解。不同数据库中的实验结果表明,DLPL算法不仅可以改善字典在遮挡和噪声干扰下的性能,提高模式识别正确率,而且可有效缩短模型的训练时间和测试时间。

关键词 图像处理; 模式识别; 投影字典学习; 低秩约束; 编码特征; 交替优化

中图分类号 TP391.41 **文献标识码** A

doi: 10.3788/LOP53.111003

Discriminative Low-Rank Projection Dictionary Pair Learning

Qiu Lida Fu Ping Lin Nan Zhang Ning

Department of Physics and Electronic Information Engineering, Minjiang University, Fuzhou, Fujian 350001, China

Abstract Compared with the conventional dictionary learning algorithm, the novel projection dictionary pair learning (DPL) algorithm introduces a projection dictionary in the dictionary learning process and utilizes the projection coding to replace the sparse coding in the dictionary for object samples. The computational cost of the pattern recognition algorithm can be effectively reduced. However, the original DPL algorithm is sensitive to occlusion and noise interference. To solve this problem, a discriminative low-rank projection dictionary pair learning (DLPL) algorithm is proposed. A low-rank constraint is added to the dictionary in the model and the least squares estimation method is used to constrain the classification error of the projection coding for the labeled samples. The unknown dictionary and the projection dictionary with closed solutions can be solved quickly by the alternative optimization method. Experimental results in different databases show that the DLPL algorithm can not only improve the performance of dictionary under occlusion and noise interference and raise the pattern recognition rate, but also effectively reduce the training time and the test time for the model.

Key words image processing; pattern recognition; projection dictionary learning; low-rank constraint; coding feature; alternative optimization

OCIS codes 100.3008; 100.3010; 100.4994

1 引言

近年来,以人脸识别^[1-2]和目标分类^[3-4]为代表的模式识别技术在目标跟踪^[5-6]、行为检测^[7-8]等领域得到了广泛应用,这类方法的主要思想是利用目标样本构建字典,根据目标在字典上的不同表示决定其类别,如Wright等^[9]提出了稀疏表示分类(SRC)算法,利用已知类别的目标图像作为字典,将待测目标在字典上的稀疏编码作为分类特征;Naseem等^[10]提出了线性回归分类(LRC)算法,与SRC使用共享字典不同,LRC按

收稿日期: 2016-06-15; **收到修改稿日期:** 2016-07-29; **网络出版日期:** 2016-10-24

基金项目: 国家自然科学基金(51277091)、中国博士后科学基金(2013T60637)、福建省中青年骨干教师教育科研项目(JA15415)

作者简介: 邱立达(1984—),男,硕士,讲师,主要从事图像处理、模式识别和机器学习等方面的研究。

E-mail: qld.qq@163.com

照不同的目标类分别构建字典,将待测目标在各组字典上进行线性回归,通过最小二乘估计来判决目标类别。以 SRC 和 LRC 为代表的算法直接使用目标图像作为字典,未有效利用特征信息,当遭遇噪声和遮挡干扰时无法准确表示目标,为此研究人员提出了字典学习方法^[11-17],将字典作为待求参量进行训练学习。Jiang 等^[11]提出了标签常量-K 均值奇异值分解(LC-KSVD)算法,在字典学习模型中增加了分类误差约束,通过标签样本的连续正则化学习具有判别能力的过完备字典;Zhang 等^[12]将提出的 Fisher 判别字典学习(FDDL)算法引入基于 Fisher 准则的稀疏编码对字典进行判别约束;Li 等^[13]提出了低秩判别字典学习($D^2L^2R^2$)算法,通过对字典进行低秩约束提高其抗遮挡能力。上述方法均可提高字典性能,但是在传统字典学习模型中广泛采用基于 l_0 或 l_1 范数的稀疏约束(l_0 、 l_1 为范数正则项),计算量巨大,为此 Gu 等^[17]提出了一种新的投影字典对学习(DPL)算法。在 DPL 算法中,目标样本在字典上的编码系数由投影字典表示,使用投影编码替代传统的稀疏编码,有效降低了算法时间复杂度,但是当样本遭受严重噪声干扰时,字典性能受到较大影响。

针对上述问题,本文在 DPL 算法的基础上提出了一种具有判别能力的低秩投影字典对学习(DLPL)算法,DLPL 在模型中增加了针对字典的低秩约束和针对投影编码的分类误差约束,在优化过程中各待求参数具有封闭解。在多个数据库中的实验结果表明,DLPL 算法有效提高了字典的判别能力、抗遮挡和抗噪声能力,进一步减少了算法的运行时间。

2 字典学习

对于模式分类问题,若训练集包含 K 类训练样本,提取 p 维样本特征向量并构造训练矩阵 $\mathbf{X} = [\mathbf{X}_1, \dots, \mathbf{X}_k, \dots, \mathbf{X}_K] \in \mathbf{R}_{p \times K_n}$, 其中 \mathbf{X}_k 为第 k 类样本的特征矩阵($\mathbf{X}_k \in \mathbf{R}_{p \times n}$), n 为每一类的样本数量。字典学习的目的是建立一个目标表示模型,通过训练集 \mathbf{X} 对模型进行训练,从而得到字典 $\mathbf{D} = [\mathbf{D}_1, \dots, \mathbf{D}_k, \dots, \mathbf{D}_K] \in \mathbf{R}_{p \times K_m}$ (m 为每一类字典中的字典原子数),利用 \mathbf{D} 线性表示目标,从而实现重构和模式分类。目前主要的字典学习方法的参数优化模型可表示为

$$\{\mathbf{D}^*, \mathbf{A}^*\} = \arg \min_{\mathbf{D}, \mathbf{A}} \|\mathbf{X} - \mathbf{D}\mathbf{A}\|_{\text{F}}^2 + \lambda \|\mathbf{A}\|_{\text{p}} + \Psi(\mathbf{D}, \mathbf{A}, \mathbf{Y}), \quad (1)$$

式中 \mathbf{Y} 为 \mathbf{X} 对应的分类标签, \mathbf{D} 为待求字典, \mathbf{A} 为编码系数矩阵, $\|\cdot\|_{\text{F}}$ 为矩阵的 F 范数, $\|\cdot\|_{\text{p}}$ 为 l_p 范数正则项, λ 为惩罚系数, $\Psi(\mathbf{D}, \mathbf{A}, \mathbf{Y})$ 为字典判别学习函数。(1)式中广泛使用 l_0 或 l_1 范数正则项对 \mathbf{A} 进行稀疏约束,使得字典对 \mathbf{X} 进行稀疏表示;但是模型在训练和测试阶段都需要对数据进行稀疏编码,计算量巨大,为此 Gu 等^[17]在新提出的 DPL 算法中引入投影字典矩阵 \mathbf{P} ($\mathbf{P} = [\mathbf{P}_1, \dots, \mathbf{P}_k, \dots, \mathbf{P}_K] \in \mathbf{R}_{K_m \times p}$)。根据 $\mathbf{A} = \mathbf{P}\mathbf{X}$, (1)式可改写为

$$\{\mathbf{D}^*, \mathbf{P}^*\} = \arg \min_{\mathbf{D}, \mathbf{P}} \|\mathbf{X} - \mathbf{D}\mathbf{P}\mathbf{X}\|_{\text{F}}^2 + \Psi(\mathbf{D}, \mathbf{P}, \mathbf{X}, \mathbf{Y}), \quad (2)$$

式中 \mathbf{P} 和 \mathbf{D} 构成字典对,分别用于生成 \mathbf{X} 的投影编码和 \mathbf{X} 的重构。在 DPL 算法中,假设 \mathbf{P}_k ($\mathbf{P}_k \in \mathbf{R}_{m \times p}$) 使得所有第 k 类样本在 \mathbf{D}_k ($\mathbf{D}_k \in \mathbf{R}_{p \times m}$) 上的重构误差最小,即 $\min_{\mathbf{D}, \mathbf{P}} \sum_{k=1}^K \|\mathbf{X}_k - \mathbf{D}_k \mathbf{P}_k \mathbf{X}_k\|_{\text{F}}^2$; 与此同时,要求第 i 类样本 ($i \neq k$) 在 \mathbf{D}_k 上的重构误差尽可能大,相应的投影编码尽可能小,即 $\mathbf{P}_k \mathbf{X}_i \approx \mathbf{0}, \forall i \neq k$ 。因此 DPL 算法的字典学习模型可表示为

$$\{\mathbf{D}^*, \mathbf{P}^*\} = \arg \min_{\mathbf{D}, \mathbf{P}} \sum_{k=1}^K \|\mathbf{X}_k - \mathbf{D}_k \mathbf{P}_k \mathbf{X}_k\|_{\text{F}}^2 + \lambda \|\mathbf{P}_k \bar{\mathbf{X}}_k\|_{\text{F}}^2, \|\mathbf{d}_i\|_2 \leq 1, \quad (3)$$

式中 $\bar{\mathbf{X}}_k$ 为 \mathbf{X}_k 在整个训练集 \mathbf{X} 中的补集, \mathbf{d}_i 为 \mathbf{D} 的第 i 个原子。

2.1 判别型低秩投影字典对学习

与传统方法相比, DPL 算法避免了对目标样本进行稀疏编码,大幅减少模型的训练时间和测试时间。但是,原始的 DPL 算法存在两个主要缺点:1) 未考虑不同类别样本投影编码 $\mathbf{P}_i \mathbf{X}_i$ 和 $\mathbf{P}_j \mathbf{X}_j$ ($\forall i \neq j$) 之间的关系,使得该算法针对样本编码特征的判别能力不足;2) 未考虑样本遭受严重噪声干扰或遮挡的情况,这时利用 DPL 算法学习到的字典 \mathbf{D} 无法准确表示目标,导致模式分类识别正确率大幅下降。

针对上述问题,提出了一种具有判别能力的低秩投影字典对学习算法, DLPL 算法在模型中引入基于标

签信息的分类误差约束,使得不同类样本 $\mathbf{X}_k (k=1,2,\dots,K)$ 通过子投影矩阵 \mathbf{P}_k 生成的投影编码 $\mathbf{P}_k \mathbf{X}_k$ 更具有区分度,即

$$\min_{\mathbf{W}, \mathbf{P}} \sum_{k=1}^K \|\mathbf{Y}_k - \mathbf{W}_k \mathbf{P}_k \mathbf{X}_k\|_{\text{F}}^2 + \lambda_4 \|\mathbf{W}_k\|_{\text{F}}^2, \quad (4)$$

式中 \mathbf{Y}_k 为 \mathbf{X}_k 对应的分类标签 ($\mathbf{Y}_k \in \mathbf{R}_{K \times n}$), \mathbf{W}_k 为待求分类权重 ($\mathbf{W}_k \in \mathbf{R}_{K \times m}$)。

DLPL 算法在字典学习过程中引入了低秩约束。根据低秩学习理论^[18-19],同类样本在特征空间中往往呈现低维流形结构,具有线性相关性,因此学习得到的类子字典 \mathbf{D}_k 也应该具有低秩性;但是当样本受到严重遮挡或噪声干扰时,这种低秩性被破坏,使得字典退化,从而影响模型的识别率。在字典 \mathbf{D}_k 的学习过程中引入低秩约束以减少噪声和遮挡的干扰,即

$$\min_{\mathbf{D}, \mathbf{P}, \mathbf{E}} \sum_{k=1}^K \text{rank}(\mathbf{D}_k) + \lambda \|\mathbf{E}_k\|_0, \quad \mathbf{X}_k = \mathbf{D}_k \mathbf{P}_k \mathbf{X}_k + \mathbf{E}_k, \quad (5)$$

式中 \mathbf{E}_k 为未知的噪声矩阵 ($\mathbf{E}_k \in \mathbf{R}_{p \times n}$), $\text{rank}(\cdot)$ 为矩阵的秩。 $\text{rank}(\cdot)$ 和 l_0 范数的非凸非光滑性导致(5)式难以求解,因此将其松弛为如下凸优化问题,即

$$\min_{\mathbf{D}, \mathbf{P}, \mathbf{E}} \sum_{k=1}^K \|\mathbf{D}_k\|_* + \lambda \|\mathbf{E}_k\|_{\text{F}}^2, \quad \mathbf{X}_k = \mathbf{D}_k \mathbf{P}_k \mathbf{X}_k + \mathbf{E}_k, \quad (6)$$

式中 $\|\cdot\|_*$ 为矩阵核范数(矩阵秩的凸近似)。综合考虑(4)、(6)式,DLPL 算法的字典学习模型可表示为

$$\{\mathbf{D}^*, \mathbf{P}^*, \mathbf{W}^*, \mathbf{E}^*\} = \arg \min_{\mathbf{D}, \mathbf{P}, \mathbf{W}, \mathbf{E}} \sum_{k=1}^K \|\mathbf{X}_k - \mathbf{D}_k \mathbf{P}_k \mathbf{X}_k - \mathbf{E}_k\|_{\text{F}}^2 + \lambda_1 \|\mathbf{D}_k\|_* + \lambda_2 \|\mathbf{D}_k\|_{\text{F}}^2 + \lambda_3 \|\mathbf{E}_k\|_{\text{F}}^2 + \|\mathbf{Y}_k - \mathbf{W}_k \mathbf{P}_k \mathbf{X}_k\|_{\text{F}}^2 + \lambda_4 \|\mathbf{W}_k\|_{\text{F}}^2 + \lambda_5 \|\mathbf{P}_k \bar{\mathbf{X}}_k\|_{\text{F}}^2. \quad (7)$$

DLPL 算法无须对目标进行稀疏编码,通过字典低秩学习和针对投影编码的分类误差约束,提高字典的判别能力、抗遮挡和抗噪声能力。

2.2 DLPL 快速优化

DLPL 算法的目标函数通常为一个多变量非凸函数,如(7)式所示。为此引入一个中间变量矩阵 \mathbf{A} , $\mathbf{A} = [\mathbf{A}_1, \dots, \mathbf{A}_k, \dots, \mathbf{A}_K] \in \mathbf{R}_{m \times n}$, 将(7)式表示为

$$\{\mathbf{D}^*, \mathbf{P}^*, \mathbf{W}^*, \mathbf{E}^*, \mathbf{A}^*\} = \arg \min_{\mathbf{D}, \mathbf{P}, \mathbf{W}, \mathbf{E}} \sum_{k=1}^K \|\mathbf{X}_k - \mathbf{D}_k \mathbf{A}_k - \mathbf{E}_k\|_{\text{F}}^2 + \lambda_1 \|\mathbf{D}_k\|_* + \lambda_2 \|\mathbf{D}_k\|_{\text{F}}^2 + \lambda_3 \|\mathbf{E}_k\|_{\text{F}}^2 + \|\mathbf{Y}_k - \mathbf{W}_k \mathbf{A}_k\|_{\text{F}}^2 + \lambda_4 \|\mathbf{W}_k\|_{\text{F}}^2 + \lambda_5 \|\mathbf{P}_k \bar{\mathbf{X}}_k\|_{\text{F}}^2 + \tau \|\mathbf{P}_k \mathbf{X}_k - \mathbf{A}_k\|_{\text{F}}^2, \quad (8)$$

式中 τ 为惩罚系数。(8)式关于每一个待求变量都为凸函数(固定其他变量),存在封闭解,因此可通过变量交替优化对(8)式进行快速求解。具体步骤如下。

1) 输入:训练集 $\mathbf{X} = [\mathbf{X}_1, \dots, \mathbf{X}_K]$, 分类标签矩阵 $\mathbf{Y} = [\mathbf{Y}_1, \dots, \mathbf{Y}_K]$, $\lambda_1 \sim \lambda_5$ 为系数, m, n, p 为参数,随机初始化矩阵 $\mathbf{D}, \mathbf{P}, \mathbf{W} = [\mathbf{W}_1, \dots, \mathbf{W}_K]$ 和零矩阵 $\mathbf{E} = [\mathbf{E}_1, \dots, \mathbf{E}_K]$ 。

2) 固定 $\{\mathbf{D}, \mathbf{P}, \mathbf{W}, \mathbf{E}\}$, 对 \mathbf{A} 进行更新,即

$$\mathbf{A}_k^* = \arg \min_{\mathbf{A}_k} \sum_{k=1}^K \|\mathbf{X}_k - \mathbf{D}_k \mathbf{A}_k - \mathbf{E}_k\|_{\text{F}}^2 + \|\mathbf{Y}_k - \mathbf{W}_k \mathbf{A}_k\|_{\text{F}}^2 + \tau \|\mathbf{P}_k \mathbf{X}_k - \mathbf{A}_k\|_{\text{F}}^2 = (\mathbf{D}_k^T \mathbf{D}_k + \mathbf{W}_k^T \mathbf{W}_k + \tau \mathbf{I})^{-1} [\mathbf{D}_k^T (\mathbf{X}_k - \mathbf{E}_k) + \mathbf{W}_k^T \mathbf{Y}_k + \tau \mathbf{P}_k \mathbf{X}_k], \quad (9)$$

式中 \mathbf{I} 为单位矩阵。

3) 固定 $\{\mathbf{A}, \mathbf{E}\}$, 对 $\mathbf{D}, \mathbf{P}, \mathbf{W}$ 进行更新,即

$$\mathbf{P}_k^* = \arg \min_{\mathbf{P}_k} \sum_{k=1}^K \lambda_5 \|\mathbf{P}_k \bar{\mathbf{X}}_k\|_{\text{F}}^2 + \tau \|\mathbf{P}_k \mathbf{X}_k - \mathbf{A}_k\|_{\text{F}}^2 = \tau \mathbf{A}_k \mathbf{X}_k^T (\tau \mathbf{X}_k \mathbf{X}_k^T + \lambda_5 \bar{\mathbf{X}}_k \bar{\mathbf{X}}_k^T + \gamma \mathbf{I})^{-1}, \quad (10)$$

式中 γ 为一个保证矩阵可逆的极小常数, $\gamma = 10e^{-4}$;

$$\mathbf{D}_k^* = \arg \min_{\mathbf{D}_k} \sum_{k=1}^K \|\mathbf{X}_k - \mathbf{D}_k \mathbf{A}_k - \mathbf{E}_k\|_{\text{F}}^2 + \lambda_1 \|\mathbf{D}_k\|_* + \lambda_2 \|\mathbf{D}_k\|_{\text{F}}^2 = \mathbf{U} \mathbf{S}_{\lambda_1/2} [\boldsymbol{\Sigma}] \mathbf{V}^T, \quad (11)$$

式中 svd 为对矩阵进行奇异值分解, $\{\mathbf{U}, \boldsymbol{\Sigma}, \mathbf{V}\} = \text{svd}[(\mathbf{X}_k - \mathbf{E}_k) \mathbf{A}_k^T (\mathbf{A}_k \mathbf{A}_k^T + \lambda_2 \mathbf{I})^{-1}]$, 软阈值 $\mathbf{S}_{\epsilon}[\mathbf{X}] = \text{sign}(\mathbf{X}) \max(|\mathbf{X}| - \epsilon, 0)$;

$$\mathbf{W}^* = \arg \min_{\mathbf{W}} \sum_{k=1}^K \|\mathbf{Y}_k - \mathbf{W}_k \mathbf{A}_k\|_{\text{F}}^2 + \lambda_4 \|\mathbf{W}_k\|_{\text{F}}^2 = \mathbf{Y}_k \mathbf{A}_k^{\text{T}} (\mathbf{A}_k \mathbf{A}_k^{\text{T}} + \lambda_4 \mathbf{I})^{-1}. \quad (12)$$

4) 固定 $\{\mathbf{D}, \mathbf{P}, \mathbf{W}, \mathbf{A}\}$, 对 \mathbf{E} 进行更新, 即

$$\mathbf{E}_k^* = \arg \min_{\mathbf{E}_k} \sum_{k=1}^K \|\mathbf{X}_k - \mathbf{D}_k \mathbf{A}_k - \mathbf{E}_k\|_{\text{F}}^2 + \lambda_3 \|\mathbf{E}_k\|_{\text{F}}^2 = (1 + \lambda_3)^{-1} (\mathbf{X}_k - \mathbf{D}_k \mathbf{A}_k). \quad (13)$$

5) 判断 $\{\mathbf{D}, \mathbf{P}, \mathbf{W}\}$ 是否收敛, 若收敛则结束迭代, 否则跳转至步骤 2)。

6) 输出: 字典对 $\{\mathbf{D}^*, \mathbf{P}^*\}$ 和分类权值矩阵 \mathbf{W}^* 。

2.3 模式分类识别

假设第 k 类的测试目标样本为 \mathbf{y} , 对应的标签向量为 $\mathbf{Y} (\mathbf{Y} \in \mathbf{R}_K)$, 则 \mathbf{y} 在子字典 \mathbf{D}_k 上的重构误差应该小于在其他类子字典 $\mathbf{D}_i (\forall i \neq k)$ 上的重构误差, 即

$$\|\mathbf{y} - \mathbf{D}_k \mathbf{P}_k \mathbf{y}\|_{\text{F}}^2 \leq \|\mathbf{y} - \mathbf{D}_i \mathbf{P}_i \mathbf{y}\|_{\text{F}}^2, \forall i \neq k. \quad (14)$$

\mathbf{y} 在 \mathbf{W}_k 上的投影编码分类误差也应该小于在 \mathbf{W}_i 上的误差, 即

$$\|\mathbf{Y} - \mathbf{W}_k \mathbf{P}_k \mathbf{y}\|_{\text{F}}^2 \leq \|\mathbf{Y} - \mathbf{W}_i \mathbf{P}_i \mathbf{y}\|_{\text{F}}^2, \forall i \neq k. \quad (15)$$

综合考虑(14)、(15)式, 则样本 \mathbf{y} 的所属类别可以表示为

$$\text{identity}(\mathbf{y}) = \arg \min_i \|\mathbf{Y} - \mathbf{W}_i \mathbf{P}_i \mathbf{y}\|_{\text{F}}^2 + \|\mathbf{y} - \mathbf{D}_i \mathbf{P}_i \mathbf{y}\|_{\text{F}}^2. \quad (16)$$

2.4 计算复杂度分析

在模型训练阶段, 原始的 DPL 算法对参数 \mathbf{A}_k 、 \mathbf{P}_k 和 \mathbf{D}_k 进行交替更新, 在每次迭代中, 时间计算复杂度分别为 $\mathcal{O}(m p n + m^3 + m^2 n)$ 、 $\mathcal{O}(m n p + m p^2)$ 和 $\mathcal{O}[W(p m n + m^3 + m^2 p + p^2 m)]$, 其中 W 为更新 \mathbf{D}_k 时交替方向乘子(ADMM)算法所需迭代次数 ($5 < W < 20$)。与 DPL 算法相似, DLPL 算法也在每次迭代中对待求参数 \mathbf{A}_k 、 \mathbf{P}_k 、 \mathbf{D}_k 、 \mathbf{W}_k 和 \mathbf{E}_k 进行交替更新, 相应的时间计算复杂度分别为 $\mathcal{O}(m p n + m^3 + m^2 n)$ 、 $\mathcal{O}(m n p + m p^2)$ 、 $\mathcal{O}(p m n + m^3 + m^2 p + p^2 m)$ 、 $\mathcal{O}(K m n + m^3 + m^2 n)$ 和 $\mathcal{O}(m p n)$ 。一般情况下, p 远大于 m 和 n , 因此 DPL 和 DLPL 算法的计算量主要集中在对字典 \mathbf{P}_k 和 \mathbf{D}_k 的更新上, DLPL 算法对 \mathbf{D}_k 的更新具有封闭形式, 计算复杂度低于 DPL 算法, 使得模型训练时间进一步减少。在测试阶段, DPL 和 DLPL 算法的时间计算复杂度都为 $\mathcal{O}(K m p)$ 。

3 实验结果与分析

为了验证 DLPL 算法的有效性, 在多个数据库上进行模式分类实验, 包括人脸数据库 Extend Yale B 和 AR, 目标分类数据库 COIL100 和 Caltech101, 行为识别数据库 UCF50 以及手写数字数据库 MNIST。将 SRC^[9]、LRC^[10]、DPL^[17]、具有判别能力的分析字典学习方法 (DADL)^[16]、LC-KSVD^[11]、FDDL^[12] 和 $\text{D}^2\text{L}^2\text{R}^2$ ^[13] 7 种模式分类识别方法与 DLPL 算法进行比较。DLPL 算法采取和文献[16-17]类似的 10 倍交叉检验方法为各个数据库设置实验参数: 对于 Extended Yale B 数据库, $\lambda_1 = 1, \lambda_2 = 0.001, \lambda_3 = 5, \lambda_4 = 0.001, \lambda_5 = 0.003, \tau = 0.1$; 对于 AR 数据库 $\lambda_1 = 1, \lambda_2 = 0.001, \lambda_3 = 5, \lambda_4 = 0.001, \lambda_5 = 0.005, \tau = 0.1$; 对于 COIL100 数据库, $\lambda_1 = 1, \lambda_2 = 0.0001, \lambda_3 = 5, \lambda_4 = 0.0001, \lambda_5 = 0.0005, \tau = 0.1$; 对于 Caltech101 数据库, $\lambda_1 = 1, \lambda_2 = 0.0001, \lambda_3 = 5, \lambda_4 = 0.0001, \lambda_5 = 0.0004, \tau = 0.1$; 对于 UCF50 数据库, $\lambda_1 = 1, \lambda_2 = 0.01, \lambda_3 = 5, \lambda_4 = 0.01, \lambda_5 = 0.01, \tau = 0.05$; 对于 MNIST 数据库, $\lambda_1 = 1, \lambda_2 = 0.001, \lambda_3 = 5, \lambda_4 = 0.001, \lambda_5 = 0.005, \tau = 0.1$ 。其余对比算法根据原文最佳参数值进行设置。

Extended Yale B 数据库: 该数据库包括各种光照条件下 38 人的 2414 幅脸部图像, 每幅图像的尺寸为 192 pixel \times 168 pixel。挑选每人 32 幅图像用于训练, 其余图像用于测试, 利用随机矩阵采样^[9]对每幅图像提取 504 维特征向量。为了评估算法的抗噪声性能, 实验中对样本添加随机像素遮挡, 不同信噪比(SNR)的样本图像如图 1 所示, 设信噪比的值为 R_{SNR} 。各算法在不同信噪比样本下的分类识别率如表 1 所示, 训练时间和测试时间如表 2 所示。DLPL 算法和 DPL 算法在不同信噪比下的重构误差如图 2、3 所示。由于在模型中加入了字典的低秩约束, DLPL 算法可有效减小噪声对字典学习的影响, 在样本遭受严重噪声干扰时仍具有较高的识别正确率。另一方面, DLPL 算法的待求参数具有封闭解形式, 大幅缩短了模型的训练时间和测试时间。



图1 不同信噪比下 Extended Yale B 数据库中的样本图像。

(a) 原始图像 ($R_{\text{SNR}} > 100$ dB); (b) $R_{\text{SNR}} = 5$ dB; (c) $R_{\text{SNR}} = 2$ dB; (d) $R_{\text{SNR}} = 1$ dB

Fig. 1 Sample images with different SNRs in Extended Yale B database.

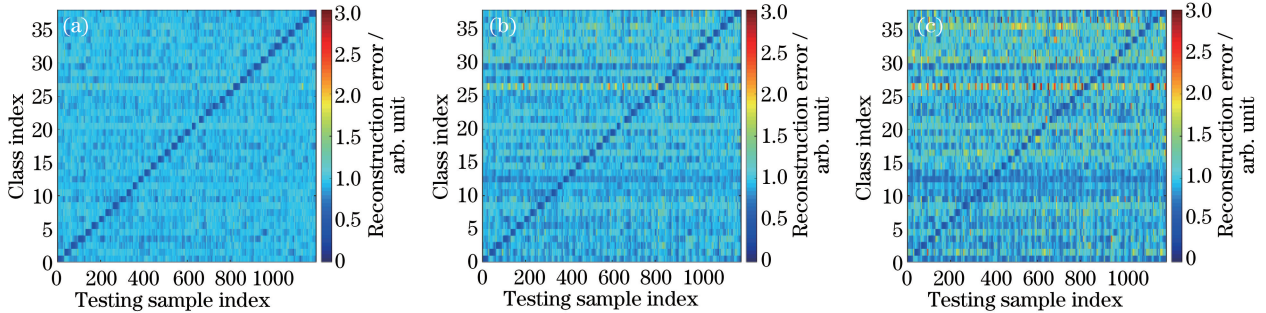
(a) Original images ($R_{\text{SNR}} > 100$ dB); (b) $R_{\text{SNR}} = 5$ dB; (c) $R_{\text{SNR}} = 2$ dB; (d) $R_{\text{SNR}} = 1$ dB图2 不同信噪比下 DLPL 算法在 Extended Yale B 数据库中的重构误差。(a) $R_{\text{SNR}} = 5$ dB; (b) $R_{\text{SNR}} = 2$ dB; (c) $R_{\text{SNR}} = 1$ dB

Fig. 2 Reconstruction error of DLPL algorithm in Extended Yale B database with different SNRs.

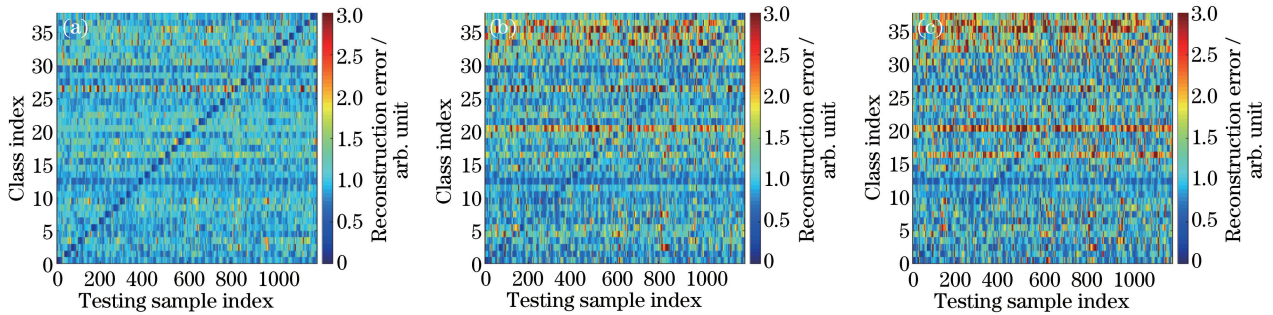
(a) $R_{\text{SNR}} = 5$ dB; (b) $R_{\text{SNR}} = 2$ dB; (c) $R_{\text{SNR}} = 1$ dB图3 不同信噪比下 DPL 算法在 Extended Yale B 数据库中的重构误差。(a) $R_{\text{SNR}} = 5$ dB; (b) $R_{\text{SNR}} = 2$ dB; (c) $R_{\text{SNR}} = 1$ dB

Fig. 3 Reconstruction error of DPL algorithm in Extended Yale B database with different SNRs.

(a) $R_{\text{SNR}} = 5$ dB; (b) $R_{\text{SNR}} = 2$ dB; (c) $R_{\text{SNR}} = 1$ dB

表1 不同信噪比下各算法在 Extended Yale B 数据库中的识别率

Table 1 Recognition rate of different algorithms in Extended Yale B database with different SNRs

SNR /dB	Recognition rate /%							
	DPL	DADL	LC-KSVD	FDDL	$D^2L^2R^2$	LRC	SRC	DLPL
≥ 100	97.5	97.6	96.7	96.7	95.5	96.5	96.4	98.5
5	62.6	63.4	62.7	45.8	84.2	66.6	43.9	85.1
2	44.9	45.3	43.1	25.3	72.1	47.2	32.5	74.4
1	31.8	32.7	30.2	19.3	60.2	31.6	18.8	62.3

表2 各算法在 Extended Yale B 数据库中的运行时间

Table 2 Running time of different algorithms in Extended Yale B database

Algorithm	DPL	DADL	LC-KSVD	FDDL	$D^2L^2R^2$	LRC	SRC	DLPL
Training time	4610	3620	416250	6601260	835710	No need	No need	2980
Testing time	0.17	0.15	0.52	1510.31	523.19	1.22	21.83	0.21

AR 数据库: 该数据库包括不同光照条件、面部表情和遮挡状态下的 126 人的脸部图像, 每幅图像的尺寸为 $43 \text{ pixel} \times 60 \text{ pixel}$, 如图 4(a) 所示, 从中挑选 50 位男性和 50 位女性, 选取每人 7 幅图像用于训练, 每人 7 幅图像用于测试。利用 3 种特征进行实验, 原始图像大小为 $43 \text{ pixel} \times 60 \text{ pixel}$ (feature 1), 裁剪后图像大小为 $27 \text{ pixel} \times 20 \text{ pixel}$ (feature 2), 通过随机矩阵采样^[9] 提取 540 维特征向量 (feature 3)。使用 540 维特

征向量时,模型所需的训练时间和测试时间如表 3 所示,各算法的模式分类识别率如图 5 所示。可以看出,与其他算法相比,DLPL 算法对特征类型的变化更不敏感,在不同特征下都具有较高的识别正确率且波动较小,算法模型具有较高的稳定性。

COIL100 数据库:该目标分类识别数据库包含 100 个不同的目标类,如图 4(b)所示。为了全面评价算法性能,选取不同数量的目标类进行 5 组实验,每组实验中随机抽取 10 类样本图像构成训练集,剩余类样本图像构成测试集,每组实验重复 20 次,将平均值作为最终实验结果,相应的分类识别率如图 6 所示。可以看出,针对不同数量的目标类,DLPL 算法取得了最好的识别正确率,表明 DLPL 算法不仅适用于人脸识别,同样适用于多目标分类快速识别。

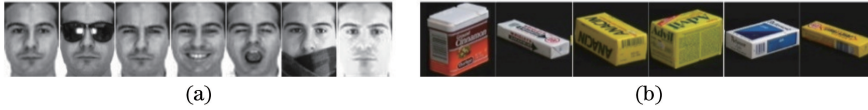


图 4 数据库中的样本图像。(a) AR 数据库;(b) COIL100 数据库

Fig. 4 Sample images in databases. (a) AR database; (b) COIL100 database

表 3 540 维特征向量下各算法在 AR 数据库中的运行时间

Table 3 Running time of different algorithms in AR database with 540-dimensional feature vector ms

Algorithm	DPL	DADL	LC-KSVD	FDDL	D ² L ² R ²	LRC	SRC	DLPL
Training time	9810	7130	1275110	51489410	1239710	No need	No need	5220
Testing time	0.27	0.22	0.63	1950.48	672.19	1.78	30.78	0.31

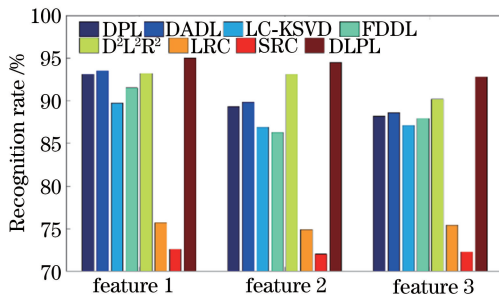


图 5 不同特征下各算法在 AR 数据库中的识别率
Fig. 5 Recognition rate of different algorithms in AR database with different features

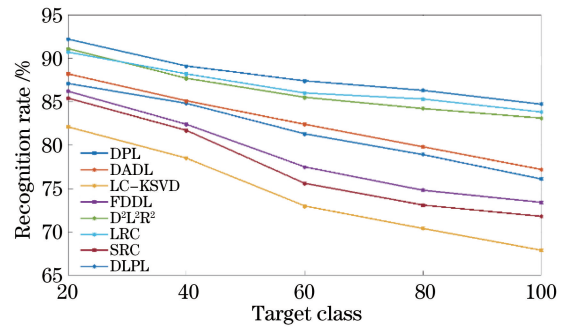


图 6 各算法在 COIL100 数据库上的识别率
Fig. 6 Recognition rate of different algorithms in COIL100 database

Caltech101 数据库:该目标分类识别数据库包括 102 类的 9144 幅图像,每一类包含 31~800 幅图像样本,从每一类选取 30 幅图像作为训练样本,其余图像作为测试样本,对样本进行高维空间金字塔匹配特征提取^[20],接着对特征进行主成分分析降维后得到 3000 维特征向量,各算法在 Caltech101 数据库中的模式分类识别率、训练时间和测试时间分别如表 4~6 所示。

UCF50 数据库:该行为识别数据库包含 50 个类别,共 6680 个人类行为视频,每一类视频分为 25 组,同组视频的观测对象、观测视角相同,具有相似的背景,利用 5 倍数据分割生成用于交叉验证的训练样本集和测试样本集并提取样本行为特征^[21]。

MNIST 数据库:该手写数字库包含 10 个数字类,共 6×10^4 幅训练图像和 1×10^4 幅测试图像,所有图像的尺寸均为 28 pixel \times 28 pixel,从每类中选取 200 幅训练图像构成训练集,选取 200 幅测试图像构成测试集。

表 4 各算法在 3 种数据库中的识别率

Table 4 Recognition rate of different algorithms in three databases

Database	Recognition rate / %							
	DPL	DADL	LC-KSVD	FDDL	D ² L ² R ²	LRC	SRC	DLPL
Caltech101	73.9	74.6	73.6	73.2	71.9	72.5	70.7	76.3
UCF50	77.4	77.8	70.1	76.5	75.3	76.9	75.0	79.7
MNIST	83.7	85.1	83.1	84.8	84.2	82.7	78.8	86.2

表 5 各算法在三种数据库中的训练时间
Table 5 Training time of different algorithms in three database

Database	Training time /s							
	DPL	DADL	LC-KSVD	FDDL	$D^2L^2R^2$	LRC	SRC	DLPL
Caltech-101	105.21	78.45	10080.63	82473.67	27557.41	No need	No need	51.66
UCF-50	620.45	330.41	19383.22	461250.71	51230.75	No need	No need	95.65
MNIST	11.26	8.55	3320.47	70782.33	6477.62	No need	No need	5.61

表 6 各算法在三种数据库中的测试时间
Table 6 Testing time of different algorithms in three database

Database	Testing time /ms							
	DPL	DADL	LC-KSVD	FDDL	$D^2L^2R^2$	LRC	SRC	DLPL
Caltech101	1.15	1.08	4.05	11750.77	3787.26	19.72	656.48	1.23
UCF50	1.82	1.51	107.55	82153.49	18562.78	177.69	6501.33	1.91
MNIST	0.47	0.35	0.91	3120.22	1075.12	3.72	57.25	0.55

DLPL 算法兼顾了投影编码的重构误差和分类误差,使得字典更具判别能力,在 Caltech101、UCF50 和 MNIST 数据库中都取得了最高的识别正确率;在 DLPL 算法的优化过程中,各待求参数具有封闭解,大幅缩短了模型的训练时间和测试时间,适用于实时性应用^[22-24]。

4 结 论

针对原始的 DPL 算法对噪声敏感、字典区分能力不强等问题,提出了一种具有判别能力的 DLPL 算法。DLPL 算法在模型中增加了对字典的低秩特征约束和投影编码的分类误差约束,有效降低了噪声干扰对字典学习的影响并提高了字典的判别能力。在训练阶段,DLPL 算法对字典的更新具有封闭形式,避免了传统方法中耗时的 ADMM 迭代求解过程,大幅降低时间计算复杂度。实验结果表明,DLPL 算法对噪声不敏感,当样本遭受严重噪声干扰时,可保持较高的分类识别正确率。与 DPL 等多种算法相比,DLPL 算法在人脸识别、目标分类和行为识别等多种数据库中都具有最高的识别正确率和最少的训练时间,在对实时性要求较高的场合具有重要的应用价值。

参 考 文 献

- 1 Ma L, Wang C, Xiao B, *et al.* Sparse representation for face recognition based on discriminative low-rank dictionary learning[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2012: 2586-2593.
- 2 Yang M, Zhang L, Yang J, *et al.* Regularized robust coding for face recognition[J]. IEEE Transactions on Image Processing, 2013, 22(5): 1753-1766.
- 3 Liu B D, Wang Y X, Zhang Y J, *et al.* Learning dictionary on manifolds for image classification[J]. Pattern Recognition, 2013, 46(7): 1879-1890.
- 4 Li Z Z, Chen J, Hou Q, *et al.* Sparse representation for infrared dim target detection via a discriminative over-complete dictionary learned online[J]. Sensors, 2014, 14(6): 9451-9470.
- 5 Wang N, Wang J, Yeung D Y. Online robust non-negative dictionary learning for visual tracking[C]. IEEE International Conference on Computer Vision, 2013: 657-664.
- 6 Bai T, Li Y F, Zhou X. Learning local appearances with sparse representation for robust and fast visual tracking[J]. IEEE Transactions on Cybernetics, 2015, 45(4): 663-675.
- 7 Li C, Han Z, Ye Q, *et al.* Visual abnormal behavior detection based on trajectory sparse reconstruction analysis[J]. Neurocomputing, 2013, 119: 94-100.
- 8 Zhou Tongchi, Cheng Xu, Wu Zhenyang. Action recognition using hierarchically tree-structured dictionary encoding[J]. Journal of Image and Graphics, 2014, 19(7): 1054-1061.
周同驰, 程 旭, 吴镇扬. 分层树结构字典编码的行为识别[J]. 中国图象图形学报, 2014, 19(7): 1054-1061.
- 9 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.
- 10 Naseem I, Togneri R, Bennamoun M. Linear regression for face recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(11): 2106-2112.
- 11 Jiang Z, Lin Z, 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.
- 12 Zhang D, Zhang L, Feng X, *et al.* Fisher discrimination dictionary learning for sparse representation[C]. IEEE International Conference on Computer Vision, 2011: 543-550.
- 13 Li L, Li S, Fu Y. Learning low-rank and discriminative dictionary for image classification[J]. Image and Vision Computing, 2014, 32(10): 814-823.
- 14 Zhang H, Zhang Y, Huang T S. Simultaneous discriminative projection and dictionary learning for sparse representation based classification[J]. Pattern Recognition, 2013, 46(1): 346-354.
- 15 Gao S, Yu Y, Cheng Y. Linear subspace learning based on a learned discriminative dictionary for sparse coding[C]. International Conference on Computer Vision Theory and Applications, 2013, 1: 530-538.
- 16 Guo J, Guo Y, Kong X, *et al.* Discriminative analysis dictionary learning[C]. AAAI Conference on Artificial Intelligence, 2016: 1617-1623.
- 17 Gu S, Zhang L, Zuo W, *et al.* Projective dictionary pair learning for pattern classification[J]. Advances in Neural Information Processing Systems, 2014, 1: 793-801.
- 18 Wright J, Ganesh A, Rao S, *et al.* Robust principal component analysis: exact recovery of corrupted low-rank matrices via convex optimization[C]. Advances in Neural Information Processing Systems, 2009: 2080-2088.
- 19 Liu G, Lin Z, Yan S, *et al.* Robust recovery of subspace structures by low-rank representation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(1): 171-184.
- 20 Lazebnik S, Schmid C, Ponce J. Beyond bags of features: spatial pyramid matching for recognizing natural scene categories[C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2006: 2169-2178.
- 21 Sadanand S, Corso J. Action bank: a high-level representation of activity in video[C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2012: 1234-1241.
- 22 Zhou Liangji, Li Qingwu, Huo Guanying, *et al.* Face image recognition method based on the NSCT and bionic pattern[J]. Laser & Optoelectronics Progress, 2015, 52(3): 031001.
周亮基, 李庆武, 霍冠英, 等. 基于 NSCT 和仿生模式的人脸图像识别方法[J]. 激光与光电子学进展, 2015, 52(3): 031001.
- 23 Cai Jiabin, Feng Guocan, Tang Xin, *et al.* Human action recognition by learning pose dictionary[J]. Acta Optica Sinica, 2014, 34(12): 1215002.
蔡加欣, 冯国灿, 汤鑫, 等. 基于姿势字典学习的人体行为识别[J]. 光学学报, 2014, 34(12): 1215002.
- 24 Guo Guangming. Research on large-scale face recognition using opto-electronic hybrid matched filtering correlator[J]. Chinese J Lasers, 2013, 40(8): 0809003.
郭广明. 光电混合匹配滤波相关器的大规模人脸识别研究[J]. 中国激光, 2013, 40(8): 0809003.