

# 基于局部类内结构的鉴别性字典学习方法<sup>\*</sup>

陈子臻<sup>a</sup>, 胡高鹏<sup>a</sup>, 王晓明<sup>a,b</sup>, 黄增喜<sup>a</sup>, 杜亚军<sup>a</sup>

(西华大学 a. 计算机与软件工程学院; b. 机器人研究中心, 成都 610039)

**摘要:** 针对支持向量引导的字典学习(support vector guided dictionary learning, SVGDL)的鉴别约束项只体现了大间隔原理, 而没有很好地利用数据空间的内在结构信息的问题, 提出了一种新颖的鉴别性字典学习方法——基于局部类内结构的鉴别性字典学习方法。该方法结合了大间隔原理和局部 Fisher 线性鉴别分析作为鉴别约束条件来指导指点学习。通过建立一个局部类内散度矩阵来编码数据空间的分布结构, 增强了挖掘同类数据的空间局部结构的能力并进一步地表示了编码向量在数据空间中的局部相似性。为了评价文章方法在图像识别上的表现, 该方法在几个常见图像数据集上进行了实验。结果表明, 所提出的方法与大间隔方法相比, 在平均识别率上有着明显的提高。

**关键词:** 字典学习; 协作表示; 支持向量机; 局部 Fisher 鉴别分析; 图像识别

**中图分类号:** TP391      **doi:** 10.19734/j.issn.1001-3695.2019.11.0607

## Discriminative dictionary learning based on locality intra-class structure

Chen Ziliu<sup>a</sup>, Hu Gaopeng<sup>a</sup>, Wang Xiaoming<sup>a,b</sup>, Huang Zengxi<sup>a</sup>, Du Yajun<sup>a</sup>

(a. School of Computer & Software Engineer, b. Robotics Research Center, Xihua University, Chengdu 610039, China)

**Abstract:** Aiming at the limitation that the discriminative term of Support vector guided dictionary learning only embodies the max-margin principle as Support vector machine (SVM), and fails to utilize the intrinsic structure information of data space, this paper proposed a novel discriminative dictionary learning method called discriminative dictionary learning based on locality intra-class structure (LCSDDL). The proposed method combined max-margin principle with Local fisher discriminant analysis (LFDA) as the discriminative term to guide dictionary learning. This method constructed a local within-class scatter matrix to encode the local structure of data space, which enhances the ability of exploiting the local structure of same class data space and further reflects the local similarity of coding vectors in data space. In order to evaluate the performance of the proposed method for image recognition, the experiment carried on several common datasets. From the experimental result, the proposed method has an obvious improvement over the other competing methods.

**Key words:** dictionary learning; collaborative representation; support vector machine; local Fisher discriminant analysis; image recognition

## 0 引言

自从字典学习理论被提出以来, 由于其在实际任务<sup>[1-5]</sup>中的突出表现, 使得字典学习方法在计算机视觉和图像处理领域获得了的广泛关注。在图像分类中, 从一组给定的训练图像中学习鉴别信息并进一步完成分类任务的学习方法称为鉴别性字典学习(discriminative dictionary learning, DDL)。

从本质上讲, 鉴别性字典学习方法的目的是通过学习训练数据的鉴别信息来构造一个鉴别性字典。目前, 字典学习方法大体上主要分为两大类: 一种字典学习方法<sup>[6-8]</sup>是分别为不同类别的数据构造一个字典, 这种方式会将每一类别的数据特性直接作用于与之对应的子字典上。然而, 当训练数据量特别大以及数据类别众多的时候, 这些基于类别特性的字典学习方法并不能取得特别优秀的表现。另一种字典学习方法<sup>[9-15]</sup>是通过寻找一个最优的共享字典来提高字典的鉴别能力。通过在编码向量上引入预测损失来指导字典和分类器的学习过程, 这类方法可以得到一个最优的共享字典。基于这个思想, Zhang<sup>[9]</sup>和 Jiang<sup>[10]</sup>等人分别提出了鉴别性字典学习方法 discriminative K-SVD(D-KSVD)和 label consistent K-SVD(LC-KSVD), 用以改进无监督的 K-SVD<sup>[11]</sup>算法。此外,

Lian 等人<sup>[12]</sup>在字典学习的过程中引入了大间隔原理, 以此提出了一个新颖的图像分类框架。Wang<sup>[13]</sup>等人提出了一种基于大间隔的评价指标来建立 SRC 的决策超平面。另外, Yang 等人<sup>[6]</sup>提出了 Fisher 鉴别字典学习方法 (Fisher discrimination dictionary learning, FDDL), 该方法利用类别特性来构建结构性的字典, 并将 Fisher 鉴别准则作用在编码向量上以此来提高编码向量的鉴别能力。

此外, 因受到 Fisher 鉴别字典的权重分配策略的启发, Cai 等人<sup>[14]</sup>提出了支持向量引导的字典学习方法(support vector guided dictionary learning, SVGDL), 该方法克服了在 FDDL 中权重值已经事先固定的缺点, 并实现了权重值的自适应分配。文献[14]通过把所有的权重值参数化为函数, SVGDL 的鉴别项可以转换为二分类 SVM 的朗格朗日对偶问题。SVGDL 的核心思想在于采用 SVM 的目标函数作为鉴别性字典学习的鉴别项来指导字典的构建过程, 使得 SVGDL 在图像识别任务中取得了令人印象深刻的表现。

然而, SVGDL 仍然有不足之处, 与 SVM 一样, 其鉴别项忽略了数据空间的分布信息, 这在一定程度上限制了泛化性能的提高。相关研究如文献[16-20]说明了数据空间的分布结构在数据分析方面起到重要的作用。Stefanos 等人<sup>[18]</sup>利用

收稿日期: 2019-11-10; 修回日期: 2019-12-31      基金项目: 国家自然科学基金资助项目(61602390); 西华大学研究生创新基金项目(yccj2019095)

**作者简介:** 陈子臻(1996-), 男, 广东韶关人, 硕士研究生, 主要研究方向为模式识别、图像处理(623796034@qq.com); 胡高鹏(1994-), 男, 河南商丘人, 硕士研究生, 主要研究方向为模式识别、图像处理; 王晓明(1977-), 男, 四川简阳人, 副教授, 硕导, 博士, 主要研究方向为模式识别、机器学习、图像处理、计算机视觉; 黄增喜(1985-), 男, 广西百色人, 副教授, 硕导, 博士, 主要研究方向为深度学习、生物特征识别、图像处理; 杜亚军(1967-), 男, 四川巴中人, 硕导, 教授, 博士, 主要研究方向为数据库理论、智能网络。

Fisher 鉴别分析的最小类内散度的思想对 SVM 进行了改进, 提出了同时体现了空间分布信息和大间隔原理的最小类内支持向量机(minimum within-class scatter support vector machine, MCVSVM)。在这基础上, Xu 等人<sup>[20]</sup>将 MCVSVM 作为鉴别项对 SVGDL 进行改进, 提出了最小类内方差支持向量诱导的字典学习(minimum within-class scatter support vector guided dictionary learning, MCVGDL)。与 SVGDL 相比, MCVGDL 很好地利用了数据的分布信息。虽然融合了 Fisher 鉴别分析的 MCVGDL 取得了令人印象深刻的表现, 但是和 Fisher 一样, MCVGDL 在探索数据的分布信息时忽略了数据空间的局部结构。Masashi 等人<sup>[21]</sup>介绍了局部 Fisher 鉴别分析(local fisher discriminative analysis, LFDA), 在多模态数据集中 LFDA 很好的考虑了数据空间的局部结构。

因此, 受到 SVGDL 和 LFDA 的启发, 文中提出了一种新颖的字典学习算法——基于局部类内结构的鉴别性字典学习方法(discriminative dictionary learning based on locality intra-class structure, LCSDDL)。该方法利用了 LFDA 的基本思想改进了 SVGDL 的鉴别项, 使其在不同类别的数据中寻找最优决策超平面时还进一步考虑到数据空间的局部结构。不同于 SVGDL, 该方法最优决策超平面基于数据空间的局部结构, 同时结合了二次合页损失函数作为模型的鉴别项来指导字典学习。最后, 为了证明文中提出的方法比起其他的字典学习方法有着明显提高, 实验部分中将会把文中提出的算法与一些代表性的算法在几个常见数据集上作比较。

## 1 相关工作

在本节中, 首先简要的回顾 SVGDL, 然后对 LFDA 进行简要的介绍。文中假设  $X = [x_1, x_2, \dots, x_n] \in \mathbb{R}^{d \times n}$  是一组具有  $n$  个  $d$  维数据且具有 2 个类别的训练样本。  $y = [y_1, y_2, \dots, y_n]$  为与  $X$  对应的类标签向量, 且  $y_i = (1, -1)$ 。因为多分类问题可以采用一对多的策略解决, 因此为了方便起见, 文中只讨论二分类问题。

### 1.1 支持向量引导的字典学习方法

在稀疏表达中, 学习一个最优的字典的过程是一个严重的时间消耗问题。为了降低字典构建的时间成本, Zhang 等人<sup>[22]</sup>提出了一种基于协作表示的字典学习方法, 该方法结合了最小二乘问题且使用了  $L_2$  范数对编码向量进行了正则化。协作表示的模型可以表述为

$$\min_{D, P} \|X - DP\|_F^2 + \alpha \|P\|_2^2 \quad (1)$$

其中,  $D \in \mathbb{R}^{d \times K}$  是要学习的鉴别性字典,  $K$  为原子个数;  $P = [p_1, p_2, \dots, p_n] \in \mathbb{R}^{K \times n}$  是与输入数据  $X$  相关的编码矩阵;  $\|X - DP\|_F^2$  是重构误差项;  $\|P\|_2^2$  是作用于编码矩阵的  $L_2$  范数, 相比于稀疏表达, 其可以加快协作表示的运算速度同时保持精度;  $\alpha$  是正则化参数; 为了获得更具有鉴别能力的字典, 在字典学习模型中加入合适的鉴别项是必要的, 这种方法叫鉴别性字典学习, 其相关的模型可以表述如下:

$$\min_{D, P} \|X - DP\|_F^2 + \alpha \|P\|_2^2 + \beta \text{Disc}(P, y, D) \quad (2)$$

其中,  $\text{Disc}(P, y, D)$  是鉴别项;  $\alpha$  和  $\beta$  是一组平衡参数, 其用于协调  $L_2$  正则项与鉴别项之间的平衡关系。

文献[14]采用了支持向量机(SVM)的目标函数作为字典学习模型的鉴别项, 并推出了支持向量引导的字典学习模型式(SVGDL)。在二分类情况下, SVGDL 的模型定义如下:

$$\min_{D, P, u, b} \|X - DP\|_F^2 + \alpha \|P\|_2^2 + \beta \left( \|u\|_2^2 + \theta \sum_{i=1}^n \max(0, 1 - y_i (u^T p_i + b)) \right) \quad (3)$$

$$s.t. \|d_k\|_2 \leq \eta, \forall k \in \{1, 2, \dots, K\}$$

其中,  $u$  为 SVM 决策超平面的法向量,  $b$  是与其对应的偏置

值;  $\theta$  是一个惩罚参数, 其表示对于错误分类的惩罚程度;  $\eta$  是一个给定的常数。显然, SVGDL 同时结合了协作表示和 SVM。从本质上来说, SVGDL 利用 SVM 所体现的最大间隔原理来指导鉴别性字典的构建, 进而利用其中的基向量来获得原始数据的鉴别表示。

### 1.2 局部 Fisher 鉴别分析

局部 Fisher 鉴别准则结合了 Fisher 鉴别准则和流形学习中的局部保持投影(locality preserving projections, LPP)<sup>[23]</sup>的思想, 其使同类数据尽可能相近的同时也考虑了原始数据空间中的局部结构。根据文献[23], 局部类内散度矩阵可以定义如下:

$$\bar{S}_w = X(E - A)X^T = XLX^T \quad (4)$$

其中,  $E$  为  $n \times n$  的对角矩阵, 其每个对角元素等于矩阵  $A$  每一行(或每一列)的元素值之和, 即  $D_{i,i} = \sum_{j=1}^n A_{i,j}$ ; 矩阵  $A$  每个元素的定义的如下:

$$A_{i,j} = \begin{cases} G_{i,j}/n_c, & \text{if } y_i = y_j = c \\ 0, & \text{if } y_i \neq y_j \end{cases} \quad (5)$$

其中  $n_c$  为第  $c$  类样本的数量( $c=1, 2$ );  $G_{i,j}$  表示在亲密度矩阵  $G$  中数据样本  $x_i$  和  $x_j$  之间的权重; 根据文献[23], 矩阵  $G$  可以被简单地定义如下:

$$G_{i,j} = \begin{cases} 1, & \text{if } x_i \in nr(x_j) \text{ or } x_j \in nr(x_i) \\ 0, & \text{other} \end{cases} \quad (6)$$

其中,  $nr(x_i)$  为数据样本  $x_i$  在数据空间中的  $k$  个最近邻接点的集合。

## 2 LCSDDL 算法

SVGDL 的主要优点是在其利用最大间隔原理来指导鉴别性字典的学习, 从而使得其获得了良好的分类性能。然而 SVGDL 如 SVM 一样没有考虑到数据空间的分布信息。而局部 Fisher 鉴别准则充分考虑到数据的分布信息, 所以文中使用局部 Fisher 鉴别准则对基于大间隔原理的 SVGDL 模型的改进是可行的。此外, 与传统的 Fisher 鉴别准则相比, 局部 Fisher 鉴别准则同时反映了数据空间的分布信息和局部结构。这再次说明了可以使用局部 Fisher 鉴别准则改进 SVGDL 模型以建立更合理的最优分类超平面, 进一步提高分类器的泛化性能。

### 2.1 LCSDDL 算法的目标函数

该方法利用 LFDA 改进 SVM, 并将改进后的算法作为鉴别项替换式(3)中的鉴别条件, 以此指导字典和分类器的学习过程。为此, 文中提出了以下的模型:

$$\min_{D, P, u, b} \|X - DP\|_F^2 + \alpha \|P\|_2^2 + \beta \left( u^T \bar{S}_w u + \theta \sum_{i=1}^n \max(0, 1 - y_i (u^T p_i + b)) \right) \quad (7)$$

$$s.t. \|d_k\|_2 \leq \eta, \forall k \in \{1, 2, \dots, K\}$$

其中,  $\bar{S}_w$  为局部类内散度矩阵式(6)。与 SVGDL 相比, LCSDDL 可以看做一种结合了 SVM 与局部 Fisher 线性鉴别的字典学习方法。LCSDDL 基于数据空间的分布信息以及数据的局部结构来构建最优分类超平面, 故比 SVGDL 具有更好的泛化性能。

### 2.2 基于局部类内结构的鉴别性字典学习算法的求解

本节将会详细地讨论如何优化 LCSDDL 算法模型式(7)。实际上, 同时更新所有的变量来求解 LCSDDL 模型是相当困难的。借鉴文献[14]的更新策略, 文中采用交替优化的策略来求模型的局部最优解。每次迭代更新的过程可分为如下三部分: a)假设字典矩阵  $D$  和编码矩阵  $P$  是固定的常量, 然后求解分类器  $\langle u, b \rangle$ ; b)假设字典矩阵  $D$  和分类器  $\langle u, b \rangle$  是固定的常量, 然后求解编码矩阵  $P$ ; c)假定编码向量  $P$  和分类器  $\langle u, b \rangle$  是固定常量, 然后更新字典矩阵  $D$ 。

更新分类器  $\langle u, b \rangle$

当字典矩阵  $D$  和编码矩阵  $P$  被固定时, 模型式(7)实际上可以被看做是求解二分类 SVM 的基本问题。为了更新分类器  $\langle u, b \rangle$ , 定义了如下模型:

$$\min_{D, P, u, b} \|X - DP\|_F^2 + \alpha \|P\|_2^2 + \beta \left( u^T \bar{S}_w u + \theta \sum_{i=1}^n \max(0, 1 - y_i (u^T p_i + b)) \right) \quad (8)$$

*s.t.*  $\|d_k\|^2 \leq \eta, \forall k \in \{1, 2, \dots, K\}$

显然, 与  $L_2$ -SVM<sup>[24]</sup>相似, 该模型式(8)采用了可导的二次损失函数。受到文献[24]与[25]的启发, 模型式(8)将采用梯度下降的优化方法来求解。为了方便起见, 将模型式(8)化简如下:

$$\mathcal{L} = \min_{u, b} u^T \bar{S}_w u + \theta \sum_{i=1}^n \tau_i (1 - y_i (u^T p_i + b))^2 \quad (9)$$

其中,  $\tau_i$  被定义如下:

$$\tau_i = \begin{cases} 1 & \text{if } 1 - y_i (u^T p_i + b) > 0 \\ 0 & \text{other} \end{cases} \quad (10)$$

为了采用梯度下降的优化方法, 首先对模型式(9)分别求  $u$  和  $b$  的偏导数。所以  $\partial \mathcal{L} / \partial u$  和  $\partial \mathcal{L} / \partial b$  为

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial u} &= 2\bar{S}_w u - 2\theta \sum_{i=1}^n \tau_i y_i (1 - y_i (u^T p_i + b)) p_i \\ \frac{\partial \mathcal{L}}{\partial b} &= -2\theta \sum_{i=1}^n \tau_i y_i (1 - y_i (u^T p_i + b)) \end{aligned} \quad (11)$$

因此, 分类器  $\langle u, b \rangle$  的梯度优化过程可被定义如下:

$$u^{t+1} = u^t - \rho * \frac{\partial \mathcal{L}}{\partial u}, \quad b^{t+1} = b^t - \rho * \frac{\partial \mathcal{L}}{\partial b} \quad (12)$$

其中,  $t$  是优化过程中当前迭代的次数;  $\rho$  是与学习率有关的固定常数。

综上所述, 算法一中概述了采用梯度下降的方法更新分类器  $\langle u, b \rangle$  的主要步骤。

算法 1 采用梯度下降的方法求解分类器  $\langle u, b \rangle$

输入: 带有类标  $y$  的训练集  $X$ , 初始化的分类器  $\langle u, b \rangle$ , 最大迭代次数  $T$  以及收敛阈值

输出: 分类器  $\langle u, b \rangle$

- 1 初始化并固定字典  $D$  和编码矩阵  $P$
- 2 根据式(4), 利用输入的训练集  $X$  构造局部类内散度矩阵  $\bar{S}_w$
- 3 Repeat
- 4 根据输入的样本判断  $\tau_i$  的值
- 5 根据式(11)计算  $u$  和  $b$  的梯度, 并根据式(12)迭代更新  $u$  和  $b$
- 6 Until ( $u$  收敛 || 迭代次数达到  $T$ )
- 7 结束。

### 2.2.1 更新编码矩阵 $P$

本节将会讨论如何更新编码矩阵  $P = [p_1, p_2, \dots, p_n]$ 。首先, 先固定字典矩阵  $D$  以及分类器  $\langle u, b \rangle$ 。根据文献[17]和[26]的更新方法, 在每次迭代过程中逐个更新编码向量  $p_i$ , 同时保持其他的编码向量不变。故模型式(8)可以定义如下:

$$\min_{p_i} \mathcal{L}(p_i) = \mathcal{L}_1(p_i) + \mathcal{L}_2(p_i) \quad (13)$$

其中,

$$\begin{aligned} \mathcal{L}_1(p_i) &= \sum_{j=1}^n \|x_j - Dp_j\|_F^2 + \alpha \|p_i\|_2^2 \\ \mathcal{L}_2(p_i) &= \beta \sum_{i,j=1}^n u^T p_i L_{ij} p_j^T u + \beta \theta \sum_{i=1}^n \max(0, 1 - y_i (u^T p_i + b))^2 \end{aligned} \quad (14)$$

根据式(10),  $\mathcal{L}_2(p_i)$  在形式上可以化为

$$\mathcal{L}_2(p_i) = \beta \sum_{i,j=1}^n u^T p_i L_{ij} p_j^T u + \beta \theta \sum_{i=1}^n \tau_i (1 - y_i (u^T p_i + b))^2 \quad (15)$$

为了获得  $p_i$  的最优解,  $\mathcal{L}_1(p_i)$  和  $\mathcal{L}_2(p_i)$  的梯度表示如下:

$$\begin{aligned} \frac{\partial \mathcal{L}_1(p_i)}{\partial p_i} &= D^T D p_i + D^T x_i + \alpha p_i \\ \frac{\partial \mathcal{L}_2(p_i)}{\partial p_i} &= \beta p_i ((L_{ii} + \theta \tau_i) u u^T) - \beta \left( \sum_{j \neq i}^n u u^T L_{ij} p_j - \theta \tau_i u (y_i - b) \right) \end{aligned} \quad (16)$$

根据式(13)和(16), 该模型式(13)的导数如下:

$$\frac{\partial \mathcal{L}(p_i)}{\partial p_i} = \frac{\partial \mathcal{L}_1(p_i)}{\partial p_i} + \frac{\partial \mathcal{L}_2(p_i)}{\partial p_i} \quad (17)$$

进而, 模型式(13)的 Hessian 矩阵为

$$\frac{\partial^2 \mathcal{L}(p_i)}{\partial p_i^2} = D^T D + \alpha I + \beta ((L_{ii} + \theta \tau_i) u u^T) \quad (18)$$

其中,  $I$  为单位矩阵。由式(18)可见,  $\partial^2 \mathcal{L}(p_i) / \partial p_i^2 > 0$ , 故该模型的 Hessian 矩阵是半正定的, 所以该模型式(13)为凸函数。通过求该模型的鞍点, 即令  $\partial \mathcal{L}(p_i) / \partial p_i = 0$ , 便可以得到模型的最优解:

$$p_i = Z_i^{-1} T_i \quad (19)$$

其中,

$$\begin{aligned} Z_i &= D^T D + \alpha I + \beta (L_{ii} + \theta \tau_i) u u^T \\ T_i &= D^T x_i - \beta \left( \sum_{j \neq i}^n u u^T L_{ij} p_j - \theta \tau_i u (y_i - b) \right) \end{aligned} \quad (20)$$

### 2.2.2 更新字典矩阵 $D$

本节将就如何更新字典矩阵  $D$  进行详细的讨论。在每次迭代更新  $D$  的过程中, 当固定编码矩阵  $P$  以及分类器  $\langle u, b \rangle$  时, 只有重构误差项涉及到了  $D$ 。因此, 可以定义优化模型如下:

$$\min_D \|X - DP\|_F^2 \quad (21)$$

*s.t.*  $\|d_k\|^2 \leq \eta, \forall k \in \{1, 2, \dots, K\}$

根据文献[27]的优化策略, 通过把模型式(21)转换为其拉格朗日对偶问题, 通过求其对偶优化问题可以得到字典  $D$  的最优解。模型式(21)的对偶形式如下:

$$\mathcal{J}(D, \lambda) = \text{Tr}((X - DP)^T (X - DP)) + \sum_{k=1}^K \lambda_k (\|d_k\|^2 - \eta) \quad (22)$$

其中,  $\lambda = [\lambda_1, \lambda_2, \dots, \lambda_K]$  是与模型式(21)中不等式约束相对应的拉格朗日乘子。对式(22)求与  $D$  有关的偏导数, 并设该偏导数为 0, 可以得到:

$$D^T = X P^T (P P^T + A)^{-1} \quad (23)$$

其中,  $A$  是  $K \times K$  的对角矩阵, 且  $A = \text{diag}(\lambda)$ 。通过将(23)代入到式(22)中, 式(22)可以化为如下形式:

$$D^T = X P^T (P P^T + A)^{-1} \quad (24)$$

通过使用拟牛顿法<sup>[28]</sup>求得(24)的最优解  $\lambda$ 。最后, 将  $\lambda$  代入式(23)中, 更新后的字典矩阵可以表示为

$$D^T = X P^T (P P^T + \text{diag}(A))^{-1} \quad (25)$$

### 2.3 分类方法

综合以上求解模型的步骤, 算法二中总结了文中模型更新的主要过程。当模型训练完成后, 可以获得学习后的字典  $D$  以及分类器  $\langle u, b \rangle$ 。随后可通过以下步骤进行分类任务, 首先将输入的测试数据  $x$  通过矩阵  $V$  投影为编码向量  $p = Vx$ 。其中矩阵  $V$  可以表述如下:

$$V = (D^T D + \alpha I)^{-1} D^T \quad (26)$$

使用学习得到的分类器  $\langle u, b \rangle$ , 可以通过如下式子预测  $x$  的类标  $y$ 。

$$y = \text{argmax } u^T p + b \quad (27)$$

算法 2 本文提出的模型的主要算法流程

输入: 带有类标  $y$  的训练集  $X$ , 初始化的分类器  $\langle u, b \rangle$ , 最大迭代次数  $T$  以及收敛阈值

输出: 分类器  $\langle u, b \rangle$ , 字典矩阵  $D$

- 1 初始化并固定字典  $D$  和编码矩阵  $P$ , 采用交替优化的方式更新模型
- 2 固定字典  $D$  以及编码矩阵  $P$ , 根据算法一的流程更新分类器  $\langle u, b \rangle$
- 3 Repeat
- 4 固定分类器  $\langle u, b \rangle$  以及字典  $D$ , 根据式(12)判断  $\tau_i$  的值, 进而根据式(21)更新编码向量  $p_i$
- 5 Until (逐行更新完编码矩阵  $P$ )
- 6 固定编码矩阵  $P$  以及分类器  $\langle u, b \rangle$ , 通过求解式(23)的对偶问题

得到字典  $D$

7 如果相邻的两次迭代的目标函数值的差小于收敛阈值, 则迭代过程终止, 否则返回第 2 步继续执行, 直到收敛或者迭代次数达到  $T$

8 结束。

### 2.4 目标函数优化过程的收敛性分析

根据文献[14]可知, 字典学习模型往往是高度非凸的, 目标函数通常难以获得全局最优解。借鉴文献[14], 如上述几个小节中所述, 针对目标函数的字典矩阵  $D$ , 编码矩阵  $P$  以及分类器  $\langle u, b \rangle$  采用多轮的交替更新方式以得到目标函数的局部最小值。假设  $f(D_m^t, P_m^t, u^t, b^t)$  为目标函数式(8)的最小值, 同时令  $f(D_m^t, P_m^t, u^t, b^t)$  为目标函数在完成第  $t$  轮交替更新后的最小值。在  $t+1$  轮交替更新过程中, 算法流程中固定第  $t$  轮得到字典矩阵  $D_m^t$  以及编码矩阵  $P_m^t$ 。分类器的更新可通过求解最小化问题式(9), 因此整体的目标函数值在这一过程中将会下降, 且有如下不等式:

$$f(D_m^t, P_m^t, u^{t+1}, b^{t+1}) \leq f(D_m^t, P_m^t, u^t, b^t) \quad (28)$$

当更新编码矩阵  $P_m^{t+1}$  时, 通过求解仅与编码矩阵更新有关的优化模型式(13), 可以得到:

$$f(D_m^t, P_m^{t+1}, u^{t+1}, b^{t+1}) \leq f(D_m^t, P_m^t, u^{t+1}, b^{t+1}) \leq f(D_m^t, P_m^t, u^t, b^t) \quad (29)$$

同理, 在更新字典矩阵  $D_m^{t+1}$  时, 通过求解式(21)的拉格朗日对偶问题, 可以有下述不等式:

$$f(D_m^{t+1}, P_m^{t+1}, u^{t+1}, b^{t+1}) \leq f(D_m^t, P_m^{t+1}, u^{t+1}, b^{t+1}) \quad (30)$$

故,  $f(D_m^{t+1}, P_m^{t+1}, u^{t+1}, b^{t+1}) \leq f(D_m^t, P_m^t, u^t, b^t)$  (31)

不等式(32)直观地表明了目标函数值在交替迭代更新的过程中是非增的。同时, 由目标函数式(7)可知, 在目标函数中各项的值均大于等于 0, 故目标函数有下界。所以随着对目标函数的分类器, 编码矩阵以及字典矩阵多轮的交替迭代更新, 目标函数值最终将收敛。

由上述推论可知, 文中提出的模型将会求解得到局部最优解甚至有可能全局最优解。此外, 为了进一步验证文中提出的算法的目标函数值收敛性, 实验部分将会给出文中提出

的算法的目标函数值收敛曲线。

## 3 实验设计与分析

为了对文中提出的鉴别性字典学习模型进行评价, 在本节中, 首先将就模型中涉及到的参数对实验结果的影响进行讨论。此外, 本节将会对文中提出的模型在 7 个数据集上进行综合性的仿真实验。为了保证实验公平, 文中所有的实验平台为 MATLAB R2016b 应用软件, 且文中所有的比较算法的 MATLAB 代码和相关参数设定均来自相关论文。

### 3.1 本文模型的相关参数对识别精度的影响

本节中, 将会对模型涉及到的相关参数进行讨论。与其他的鉴别字典学习模型一样, 本文提出的模型也涉及到了平衡参数  $\alpha$  和  $\beta$ 。此外, 本文针对 MCVSVM 的 Fisher 准则使用了 LFDA 进行了改进, 所以引进了 LFDA 的相关参数——最近邻接点参数  $k$ 。

下列实验采用 COIL 20 物体数据集<sup>[29]</sup>作为实验数据。COIL 20 数据集含有 20 个不同的类别的物体共 1440 张图片, 其中每个类别共包含 72 张在不同角度下拍摄的图片。为了实验公平, 该数据集中所有图片的尺寸统一预处理为  $32 \times 32$  像素, 并在每一类中随机选取 20 张图片作为训练数据, 其余的作为测试数据。在下列实验中记录的实验精度均来自 20 次重复的图像选取过程并进行实验的平均识别精度。

a) 研究平衡参数  $\alpha$  和  $\beta$  的选择对模型精度的影响。首先, 实验中分别设定  $\alpha$  和  $\beta$  的选择范围为  $\{1e-7, 1e-6, 1e-5, 1e-4, 5e-4, 1e-3, 2e-3, 5e-3, 1e-2, 1.5e-2\}$  和  $\{1e-8, 1e-7, 1e-6, 2e-6, 1e-5, 2e-5, 1e-4, 2e-4, 1e-3, 2e-3\}$ 。表 1 显示了模型识别精度在不同组合的  $\alpha$  和  $\beta$  下的变化。在绝大部分情况下, 识别精度均保持在 90%以上并在  $\alpha=2e-3$  和  $\beta=1e-6$  时识别精度达到 97.04%, 这说明了本文的模型对于平衡参数的变化在大体上并不敏感。但是, 当  $\beta > 2e-6$  时, 识别精度明显有所降低, 这表明了在实验中  $\beta$  的取值范围应当设定于  $(1e-8 \leq \beta \leq 2e-6)$  中。

表 1 本文模型中平衡系数  $\alpha$  和  $\beta$  对识别精度的影响

Tab. 1 Experimental results of different combinations of  $\alpha$  and  $\beta$  %

参数设置	$\beta=1e-8$	$\beta=1e-7$	$\beta=1e-6$	$\beta=2e-6$	$\beta=1e-5$	$\beta=2e-5$	$\beta=1e-4$	$\beta=2e-4$	$\beta=1e-3$
$\alpha=1e-7$	96.49	96.33	96.17	95.85	93.35	93.64	90.88	91.78	90.56
$\alpha=1e-6$	96.49	96.26	96.55	95.91	93.25	93.03	90.98	92.39	90.50
$\alpha=1e-5$	96.30	96.26	96.36	96.07	92.51	92.64	91.30	90.21	89.73
$\alpha=1e-4$	96.15	96.23	96.33	95.91	93.12	93.09	92.64	92.55	89.85
$\alpha=5e-4$	96.65	96.58	96.71	96.23	93.89	94.05	91.78	90.56	90.59
$\alpha=1e-3$	96.68	96.78	96.62	96.58	94.63	93.57	92.10	91.33	91.17
$\alpha=2e-3$	96.58	96.55	97.04	96.49	94.08	94.50	93.83	91.07	91.81
$\alpha=5e-3$	96.58	96.58	96.71	96.75	94.44	93.83	91.36	92.93	91.07
$\alpha=1e-2$	96.65	96.52	96.84	96.42	95.53	95.43	92.39	91.39	90.59

b) 研究最近  $k$  个最近邻接点的选择对模型精度的影响。首先, 实验设定  $k$  的选择范围为  $\{2, 3, 5, 7, 9, 12, 15, 20\}$ 。此外, 为了验证局部 Fisher 准则改进的 MCVSVM 作为 SVGDL 的鉴别项的有效性, 实验中引入了传统 MCVSVM 作为 SVGDL 的鉴别项的字典学习方法 MCVGDL<sup>[20]</sup>。图 1 中显示了最近  $k$  个邻接点的变化对于本文模型在识别精度上的影响。根据图 1 可知, 本文的模型和 MCVGDL 均展现了比 SVGDL 更好的性能, 这说明了数据空间分布信息对于鉴别信息的挖掘具有重要作用。此外, 与 MCVGDL 相比, 本文的模型进一步的考虑的数据分布的局部结构并在分类性能上获得了显著的提升。这说明了在本文提出的方法中利用局部 Fisher 准则改进后的 SVM 目标函数作为 SVGDL 的鉴别项是有效的。

### 3.2 实验设置及结果

本节中, 文中提出的模型将会在人脸数据集, 物体数据集以及手写数字数据集共 7 个标准数据集上与 9 个具有代表性的机器学习算法以及字典学习算法进行对比实验。其中, 这些算法有 FDDL<sup>[6]</sup>, SVGDL<sup>[14]</sup>, DLSPC<sup>[27]</sup>, ASF-SRC<sup>[30]</sup>, MMSC<sup>[31]</sup>, LC-MMSC<sup>[32]</sup>, MDDL<sup>[33]</sup>, BDLRR<sup>[34]</sup>, MCVGDL<sup>[20]</sup>。数据集和实验设置如下:

a) PIE 人脸数据集<sup>[35]</sup>由卡耐基梅隆大学的 3D 实验室收集。该数据集包含了 68 个人的人脸图像, 图片数量超过 40000。精确识别该数据集的主要难点在于其中的人脸图像均拍摄于不断变化的光照条件以及丰富的人物表情。实验中, 分别从每一类别中选择 5, 10, 15, 20 张图片作为训练样本, 其余部分作为测试样本。为了降低运算成本, 所有的样本均裁剪为

32×32 大小。对于本文提出的模型, 实验中固定字典原子数  $K=340$  并设置最近邻接点参数  $k=3$ 。表 2 列出了重复 20 次实验后的平均识别精度。通过表 2 可以看出, 本文提出的算法相比于其他算法取得了最高的识别精度。与 SVGDL 相比, 本文的算法提高了 3%~9%。与 MCVGDL 相比, 在样本数量为 5 和 10 的情况下, 识别精度没有明显的提高, 但是随着训练样本数量的增加, 本文提出的模型明显表现了更高的识别精度和更优秀的分类性能。

b)AR 人脸数据集<sup>[36]</sup>由 120 个类别大约 4000 张人脸图像组成。与 PIE 数据集相同, 该数据集丰富的人脸表情以及不同的光照强度提升了识别 AR 数据集的难度。在实验中, 选择了 100 个类别共 2600 张图像(每类 26 张)作为实验数据。并从每一类中选择 2, 3, 5, 7 张图像作为训练样本, 其余的数据作为测试样本。对于文中提出的模型, 固定字典原子数  $K=400$  以及设置最近邻接点参数  $k=2$ 。表 3 列出了本文算法和其余 9 个比较算法经过 20 次运行后的平均识别精度和标准差。从表 3 可以看出, LCSDDL 相比于其他算法至少提高了 1%~2%, 平均提高了 4%~5%。这说明了本文提出的模型在多类别的人脸数据集中仍然有着比其他比较算法更高的分类性能。

c)extend Yale B 人脸数据集<sup>[37]</sup>共含有 38 个类别 2414 张人脸图像(每个类别分别有 64 张实验图像)。该数据集的包含了不同光照强度以及多变的人物表情的人脸图像。为了减少运算成本, 该实验把每张图片裁剪为 32×32 像素(原图为 192×168 像素)。此外, 分别从每一类中选择 5, 10, 15, 20 张人脸图像作为训练样本, 其他的图片作为测试样本。设置字典原子数  $K=114$  和最近邻接点参数  $k=3$ 。表 4 列出了个算法在运行 20 次后的平均识别精度和标准差。从表 4 可以知道, 在训练样本数量为 10 的情况下, LC-MMSC 的识别精度略微高于本文的算法, 但是随着样本数量的增加, 与其他算法比较, 本文的模型仍然获得了最优的分类性能。

d) COIL20 物体数据集由 20 个类别共 1440 张的物体图像组成, 每个类别均包含 72 张从不同角度拍摄的图片。COIL100 物体数据集<sup>[38]</sup>由 100 个类别共 7200 张物体图像组成, 其中每个类别的共 72 张拍摄至不同角度的图片。为了降低运算成本, 每一类均裁剪为 32×32 的图像。对于 COIL20 数据集, 在实验中, 分别从每一类中随机选择了 5, 10, 15, 20 张图片作为训练样本, 剩余部分作为测试样本。同时, 设置字典原子数  $K$  以及最近邻接点参数  $k$  分别为 360 以及 3。对于 COIL100 数据集, 除了分别为每一类别随机挑选 10, 15, 20, 30 张图片作为训练样本, 其余的实验设置均与 COIL20 的实验设置一致。表 5 和 6 分别展示了本文的方法以及其余的比较算法在 COIL20 和 COIL100 数据集上的运行

20 次后的平均实验结果。通过表 5 和 6 可以看出在多类别的物体识别数据集中, 本文的方法相较于 SVGDL 和 MCVGDL 提高了 2%~3%, 说明了体现大间隔原理的同时兼顾数据分布的局部结构信息的思想对 SVGDL 的改进是有效的。

e) CalTech101 物体数据集<sup>[39]</sup>由 101 个类别共 9144 张图片组成。实验中, 分别从每个类别中随机选择 15, 20, 25, 30 张不同的图片作为训练样本, 其余图片作为测试样本。同时, 设置字典原子数  $K$  以及最近邻接点参数  $k$  分别为 505 以及 5。实验结果展示在表 7 中, 表 7 表明了本文的算法获得了优于其他算法的识别精度, 相比于 SVGDL 和 MCVGDL 提高了 1%~2%。这说明了通过考虑数据空间的局部结构以及利用 SVM 的大间隔理论, 可以有效地提高字典学习模型的重构能力和分类器的泛化性能。

f)USPS 手写数据集<sup>[40]</sup>由 10 个类别(数字 0-9)共 9298 张图片组成。虽然 USPS 只有 10 个类别, 但是识别该数据集的挑战主要在于每个类别的数字均由简单的笔画构造以及丰富多样的字迹组成。实验中随机选取了 30, 60, 100, 200 张不同的手写数字图片作为训练样本, 其余的图片作为测试样本。同时, 字典原子数  $K$  以及最近邻接点参数  $k$  分别设置为  $K=200$  以及  $k=10$ 。本文的方法以及作为对比的方法在运行 20 次后得到的平均实验精度和标准差展示在表 8 中, 相比于其他算法, 本文的方法平均提高了 2%~5%。同时与 SVGDL 和 MCVGDL 相比, 文中的提出的方法有着显著的提高, 说明了采用局部 Fisher 鉴别准则的局部类内散度矩阵来编码数据空间的局部结构可以增加 SVGDL 的鉴别能力, 提高模型的泛化性能。

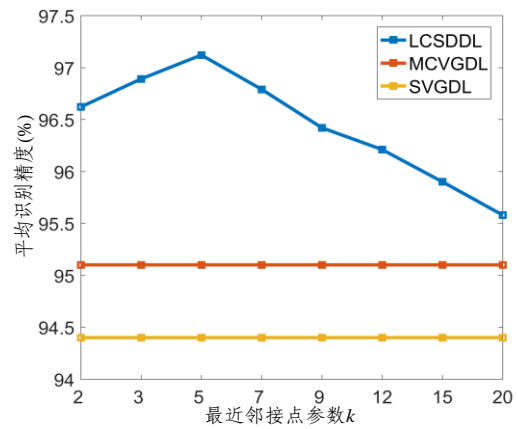


图 1 邻近点参数 k 的变化对实验精度的影响

Fig. 1 The influence of the nearest neighborhood parameter k on experimental accuracy

表 2 PIE 数据集的平均识别率和标准差

Tab. 2 The experimental results of PIE database /%

Accuracy	FDDL	SVGDL	ASF-SRC	DLSPC	BDLRR	MDDL	MMSC	LC-MMSC	MCVGDL	LCSDDL
PIE(5)	71.22±1.92	63.56±3.41	72.14±1.78	61.58±3.25	70.24±1.38	71.48±2.65	65.26±1.41	65.78±1.82	72.45±3.12	<b>73.24±2.17</b>
PIE(10)	85.82±2.41	79.14±3.63	86.67±2.16	76.46±2.68	85.47±2.31	87.16±2.18	81.94±2.16	82.40±3.12	87.44±3.54	<b>88.46±2.74</b>
PIE(15)	88.25±3.41	81.56±3.81	90.68±1.96	83.86±1.58	90.18±1.76	91.16±1.08	85.46±1.81	85.28±1.34	91.57±2.91	<b>92.70±1.56</b>
PIE(20)	91.55±3.24	84.19±3.29	91.56±2.48	90.45±2.50	91.81±1.78	92.25±2.50	89.23±2.21	90.64±2.67	92.13±1.17	<b>93.78±2.11</b>

表 3 AR 数据集的平均识别率和标准差

Tab. 3 The experimental results of AR database /%

Accuracy	FDDL	SVGDL	ASF-SRC	DLSPC	BDLRR	MDDL	MMSC	LC-MMSC	MCVGDL	LCSDDL
AR(2)	89.58±3.25	89.17±2.60	90.78±2.12	66.18±1.78	87.65±2.12	91.58±1.48	89.17±2.60	90.56±1.45	<b>92.88±1.56</b>	92.22±2.31
AR(3)	94.92±0.81	94.61±1.92	93.75±1.42	70.15±2.11	90.75±1.82	94.55±2.40	93.81±1.02	94.37±1.98	94.82±1.64	<b>95.85±1.80</b>
AR(5)	94.71±2.41	94.75±2.41	96.21±1.34	86.57±1.43	96.21±1.64	96.97±2.33	95.75±2.11	97.01±1.01	97.21±0.88	<b>98.43±0.76</b>
AR(7)	96.53±1.87	96.43±2.51	97.68±1.29	92.56±1.18	97.68±1.49	98.46±0.28	96.83±1.81	97.87±0.21	98.20±0.34	<b>99.12±0.26</b>

表 4 Extend Yale B 数据集的平均识别率和标准差  
Tab. 4 The experimental results of Extend Yale B database /%

Accuracy	FDDL	SVGDL	ASF-SRC	DLSPC	BDLRR	MDDL	MMSC	LC-MMSC	MCVGDL	LCSDDL
EYB (5)	63.45±1.83	76.75±2.21	64.35±2.35	71.74±1.70	77.15±1.45	78.74±2.74	76.85±1.21	78.84±1.46	77.42±3.31	<b>79.03±3.54</b>
EYB (10)	87.59±1.20	88.75±2.12	78.44±1.24	85.26±1.35	90.14±1.04	91.56±3.15	90.75±2.62	<b>91.70±1.81</b>	89.72±1.80	91.22±2.07
EYB (15)	91.76±2.10	91.48±1.09	84.25±1.96	90.36±1.22	92.25±1.61	93.86±1.32	93.23±1.53	93.46±1.66	92.55±1.37	<b>94.12±1.56</b>
EYB (20)	93.18±1.75	95.65±1.63	91.53±0.81	92.45±1.60	94.73±0.61	97.15±2.60	96.15±2.03	97.42±0.87	96.51±0.97	<b>98.17±0.56</b>

表 5 COIL20 数据集的平均识别率和标准差  
Tab. 5 The experimental results of COIL20 database /%

Accuracy	FDDL	SVGDL	DLSPC	BDLRR	MDDL	MMSC	LC-MMSC	MCVGDL	LCSDDL
COIL20 (5)	81.42±2.75	80.42±2.75	81.86±1.67	81.35±1.73	81.72±2.15	82.36±2.15	82.56±3.55	81.47±2.75	<b>83.46±3.43</b>
COIL20 (10)	90.15±1.56	90.05±1.52	90.66±2.18	89.96±2.48	90.97±2.06	90.05±2.14	90.11±3.21	90.72±4.21	<b>92.29±1.76</b>
COIL20 (15)	93.46±2.14	92.63±1.85	94.05±1.59	92.62±2.15	93.66±1.74	92.42±1.45	92.71±1.45	92.34±1.89	<b>94.51±2.13</b>
COIL20 (20)	95.18±1.62	94.45±1.08	95.87±2.56	93.52±1.28	95.79±1.60	93.51±1.22	94.21±1.44	95.12±1.28	<b>97.16±1.45</b>

表 6 COIL100 数据集的平均识别率和标准差  
Tab. 6 The experimental results of COIL100 database /%

Accuracy	FDDL	SVGDL	DLSPC	BDLRR	MDDL	MMSC	LC-MMSC	MCVGDL	LCSDDL
COIL100(10)	76.45±0.97	76.48±1.93	76.82±2.45	77.54±1.66	77.60±1.24	74.48±1.63	75.41±2.15	77.61±3.12	<b>78.90±1.47</b>
COIL100(15)	81.73±1.67	81.76±2.66	80.34±1.64	82.15±2.56	81.90±2.16	78.56±1.65	80.69±3.17	<b>82.37±3.44</b>	82.33±2.48
COIL100(20)	82.83±3.47	82.12±1.67	82.76±2.31	84.44±1.46	83.73±1.66	80.52±2.37	82.67±1.58	83.35±3.74	<b>84.89±1.93</b>
COIL100(30)	89.82±2.66	90.12±1.55	88.60±1.58	91.90±1.57	91.82±1.46	86.32±1.43	89.70±2.97	90.23±1.87	<b>92.12±0.89</b>

表 7 CalTech101 数据集的平均识别率和标准差  
Tab. 7 The experimental results of caltech101 database /%

Accuracy	FDDL	SVGDL	DLSPC	BDLRR	MDDL	MMSC	LC-MMSC	MCVGDL	LCSDDL
CAL101 (10)	65.61±1.70	68.32±1.85	68.72±2.10	64.72±2.71	69.20±2.51	65.42±2.15	69.45±2.13	68.98±3.45	<b>69.55±3.45</b>
CAL101 (20)	68.65±2.12	72.12±1.95	73.09±1.87	71.87±1.62	73.65±1.12	69.42±1.65	71.88±1.59	73.24±2.14	<b>74.35±3.45</b>
CAL101 (25)	71.66±3.28	75.33±1.66	75.62±1.89	74.18±1.47	75.86±2.48	72.43±2.06	74.45±1.77	75.20±1.66	<b>76.44±2.14</b>
CAL101 (30)	73.62±3.26	76.44±2.65	76.87±1.55	76.21±1.54	78.02±1.29	74.84±1.35	76.17±0.29	76.86±1.83	<b>78.18±1.56</b>

表 8 USPS 数据集的平均识别率和标准差  
Tab. 8 The experimental results of USPS database /%

Accuracy	FDDL	SVGDL	DLSPC	BDLRR	MDDL	MMSC	LC-MMSC	MCVGDL	LCSDDL
USPS (30)	84.95±2.57	86.64±1.44	85.86±2.62	87.72±2.76	88.95±1.53	86.44±2.24	87.96±1.89	88.45±1.78	<b>90.15±2.01</b>
USPS (60)	87.78±1.73	88.69±2.65	88.05±1.94	90.27±1.56	91.78±2.73	88.41±1.45	90.87±2.15	91.21±2.35	<b>92.02±1.82</b>
USPS (100)	90.84±1.25	90.87±1.97	91.24±1.66	91.94±2.43	93.96±2.75	91.50±2.07	92.47±1.69	93.42±1.88	<b>94.58±3.41</b>
USPS (200)	92.08±3.45	92.32±1.53	92.68±1.98	94.23±1.86	95.88±2.44	92.92±1.83	94.57±0.87	95.63±1.32	<b>97.23±2.13</b>

### 3.3 优化模型的收敛性实验

为了可视化本文提出的模型在优化过程中的收敛结果, 本节给出了目标函数在 PIE 数据集上的目标函数的数值变化曲线图。从图 2 中可以发现, 目标函数的数值在大约 10 次迭代后收敛到极小值, 这说明了本文模型的最小值在优化过程中是收敛的。此外, 比起 SVGDL, 本文的模型与 MCVGDL 收敛速度更快且有着更加小的目标函数值。这证明了在数据分析过程中, 考虑数据的分布信息对于 SVGDL 的提高有着明显的作用。同时, 与 MCVGDL 相比, 文中的模型明显有着更小的目标函数值。这再一次证明了利用数据的分布信息且考虑数据空间的局部结构对能够有效地进一步提高 SVGDL 的性能。

## 4 结束语

SVGDL 仅仅体现了大间隔思想, 并没有进一步考虑到数据的分布信息以及局部结构, 因此限制了分类器的泛化性能的提高。故而, 本文借鉴局部 Fisher 鉴别准则的思想, 利用了局部类内散度矩阵来编码数据分布结构改进了原始的 SVGDL 字典学习模型。在此基础上提出了基于局部结构的字典学习方法, 并利用交替优化策略来优化鉴别性字典以及分类器。

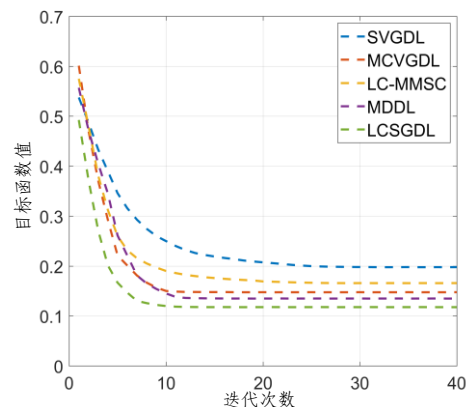


图 2 文中的模型在 PIE 数据集的收敛曲线

Fig. 2 The convergence curve of models on PIE database

然而, 作为一个新提出的算法, 本文的字典学习方法仍然有改进的空间。本文方法的求解方式是采用交替优化的方式来获取局部最优解, 这在一定程度上限制了本文方法的分类性能, 因此研究一个更有效的求解方法去获得全局最优解将会是本文的未来的工作之一。此外, 把本文提出的算法与神经网络结合也是本文未来的研究方向之一。

## 参考文献:

- [1] 汤镇宇, 孟凡荣, 王志晓. 基于稀疏表示的快速 L2-范数人脸识别方法 [J]. 计算机应用研究, 2016, 33 (9): 2831-2836. (Tang Zhenyu, Meng Fanrong, Wang Zhixiao. Fast face recognition with regularized least square via sparse representation [J]. Application Research of Computers, 2016, 33 (9): 2831-2836.)
- [2] Lu Jiwen, Wang Gang, Zhou Jie, *et al.* Simultaneous feature and dictionary learning for image set based face recognition [J]. IEEE Trans on Image Processing, 2017: 1-1.
- [3] 王晓明, 黄凤, 刘少鹏, 等. 改进的单幅图像自学习超分辨率重建方法 [J]. 计算机应用研究, 2019, 36 (08): 2534-2538. (Wang Xiaomng, Huang Feng, Liu Shaopeng, *et al.* Improved super-resolution reconstruction method for self-learning of single image [J]. Application Research of Computers, 2019, 36 (08): 2534-2538.)
- [4] 任越美, 张艳宁, 魏巍, 等. 基于稀疏表示和词袋模型的高光谱图像分类 [J]. 计算机科学, 2014, 41 (10): 113-116. (Ren Yuemei, Zhang Yanning, Wei Wei, *et al.* Classification of hyperspectral image based on sparse representation and bag of words [J]. Computer Science, 2014, 41 (10): 113-116.)
- [5] 焦莉娟, 王文剑, 赵青杉, 等. 近邻局部 OMP 稀疏表示图像去噪 [J]. 中国图像图形学报, 2017 (11). (Jiao Lijuan, Wang Wenjian, Zhao Qingshan, *et al.* Image denoising based on sparse representation of neighbor local OMP [J]. Journal of Images and Graphics, 2017 (11).)
- [6] 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.
- [7] Wang Wei, Yan Yan, Winkler S, *et al.* Category specific dictionary learning for attribute specific feature selection [J]. IEEE Trans on Image Processing, 2016, 25 (3): 1465-1478.
- [8] Ramirez I, Sprechmann P, Sapiro G. Classification and clustering via dictionary learning with structured incoherence and shared features [C]// Proc of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2010: 3501-3508.
- [9] Zhang Qiang, Li Baoxin. Discriminative K-SVD for dictionary learning in face recognition [C]// Proc of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2010: 2691-2698.
- [10] Jiang Zhuolin, Lin Zhe, Davis L S. Label consistent K-SVD: learning a discriminative dictionary for recognition [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2013, 35 (11): 2651-2664.
- [11] Aharon M, Elad M, Bruckstein A M, *et al.* K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation [J]. IEEE Trans on Signal Processing, 2006, 54 (11): 4311-4322.
- [12] Lian Xiaochen, Li Zhiwei, Lu Baoliang, *et al.* Max-margin dictionary learning for multiclass image categorization [C]// Proc of the 11th European Conference on Computer Vision. 2010: 157-170.
- [13] Wang Zhaowen, Yang Jianchao, Nasrabadi N, *et al.* A max-margin perspective on sparse representation-based classification [C]// Proc of IEEE International Conference on Computer Vision. 2013: 1217-1224
- [14] Cai Sijia, Zuo Wangmeng, Zhang Lei, *et al.* Support vector guided dictionary learning [C]// Proc of European Conference on Computer Vision. 2014: 624-639.
- [15] Li Zhengming, Lai Zhihui, Xu Yong, *et al.* A locality-constrained and label embedding dictionary learning algorithm for image classification [J]. IEEE Trans on Neural Networks and Learning Systems, 2015: 1-16.
- [16] Wang Xiaoming, Wang Shitong, Huang Zengxi, *et al.* Structure regularized sparse coding [J]. Knowledge-Based Systems, 2019, 174: 87-102.
- [17] Zheng Miao, Bu Jiajun, Chen Chun, *et al.* Graph regularized sparse coding for image representation [J]. IEEE Trans on Image Processing, 2011, 20 (5): 1327-1336.
- [18] Zafeiriou S, Tefas A, Pitas I. Minimum class variance support vector machines [J]. IEEE Trans on Image Processing, 2007, 16 (10): 2551-2564.
- [19] Wang Xiaoming, Chung Fulai, Wang Shitong. On minimum class locality preserving variance support vector machine [J]. Pattern Recognition, 2010, 43 (8): 2753-2762.
- [20] 王晓明, 徐涛, 冉彪. 最小类内方差支持向量诱导的字典学习 [J/OL]. 计算机工程: 1-10. (2019-06-04) [2019-10-11]. <http://kns.cnki.net/kcms/detail/31.1289.tp.20190531.1707.008.html>. (Wang Xiaoming, Xu Tao, Ran Biao. Minimum class variance support vector guided dictionary learning [J/OL]. Computer Engineering: 1-10. (2019-06-04) [2019-12-27]. <http://kns.cnki.net/kcms/detail/31.1289.tp.20190531.1707.008.html>.)
- [21] Sugiyama M, Ide T, Nakajima S, *et al.* Semi-supervised local fisher discriminant analysis for dimensionality reduction [J]. Machine Learning, 2010, 78 (1): 35-61.
- [22] Zhang Lei, Yang Meng, Feng Xiangchu. Sparse representation or collaborative representation: which helps face recognition? [C]// Proc of IEEE International Conference on Computer Vision. 2011: 471-478.
- [23] He Xiaofei, Niyogi P. Locality preserving projections [C]// Proc of Neural Information Processing Systems. 2003: 153-160.
- [24] Chapelle O. Training a Support vector machine in the primal [J]. Neural Computation, 2007, 19 (5): 1155-1178.
- [25] Yang Jianchao, Yu Kai, Gong Yihong, *et al.* Linear spatial pyramid matching using sparse coding for image classification [C]// Proc of Computer Vision and Pattern Recognition. 2009: 1794-1801.
- [26] Lee H, Battle A, Raina R, *et al.* Efficient sparse coding algorithms [C]// Proc of Neural Information Processing Systems. 2006: 801-808.
- [27] Wang Donghui, Kong Shu. A classification-oriented dictionary learning model: Explicitly learning the particularity and commonality across categories [J]. Pattern Recognition, 2014, 47 (2): 885-898.
- [28] Boyd S, Parikh N, Chu E K, *et al.* Distributed optimization and statistical learning via the alternating direction method of multipliers [J]. Foundations & Trends in Machine Learning, 2010, 3 (1): 1-122.
- [29] Rate C, Retrieval C. Columbia object image library (COIL-20) [J]. Computer, 2011.
- [30] Xu Yong, Zhang Zheng, Lu Guangming, *et al.* Approximately symmetrical face images for image preprocessing in face recognition and sparse representation based classification [J]. Pattern Recognition, 2016, 54 (54): 68-82.
- [31] Liu Chienliang, Hsiao W, Xiao Bin, *et al.* Maximum-margin sparse coding [J]. Neurocomputing, 2017: 340-350.
- [32] Hsiao W, Liu Chienliang, Wu Weiliang, *et al.* Locality-constrained max-margin sparse coding [J]. Pattern Recognition, 2017: 285-295.
- [33] Song Jianqiang, Xie Xuemei, Shi Guangming, *et al.* Multi-layer discriminative dictionary learning with locality constraint for image classification [J]. Pattern Recognition, 2019: 135-146.
- [34] Zhang Zheng, Xu Yong, Shao Ling, *et al.* Discriminative Block-Diagonal Representation Learning for Image Recognition [J]. IEEE Trans on Neural Networks, 2018, 29 (7): 3111-3125.
- [35] Sim T, Baker S, Bsat M. The CMU pose, illumination, and expression database [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2003, 25 (12): 1615-1618.
- [36] Aleix M, Benavente R. The ar face database [J]. Cvc Technical Report, 1998, 24.

- [37] Lee K, Member S, Ho J, *et al.* Acquiring linear subspaces for face recognition under variable lighting [M]. IEEE Computer Society, 2005.
- [38] Murase H, Nayar S K. Visual learning and recognition of 3D objects from appearance [J]. International Journal of Computer Vision, 1995, 14 (1): 5-24.
- [39] Li Feifei, Fergus R, Perona P. Learning generative visual models from few training examples: an incremental bayesian approach tested on 101 object categories [C]// Proc of Computer Vision and Pattern Recognition. 2004: 178-178.
- [40] Hull J J. A database for handwritten text recognition research [J]. IEEE Trans on Pattern Analysis & Machine Intelligence, 2002, 16 (5): 550-554.