

## 基于 Fisher 约束和字典对的图像分类

郭继昌\* 张帆 王楠  
(天津大学电子信息工程学院 天津 300072)

**摘要:** 基于稀疏表示的分类方法由于其所具有的简单性和有效性获得了研究者的广泛关注,然而如何建立字典原子与类别信息间的联系仍然是一个重要的问题,与此同时大部分稀疏表示分类方法都需要求解受范数约束的优化问题,使得分类任务的计算较复杂。为解决上述问题,该文提出一种新的基于 Fisher 约束的字典对学习新方法。新方法联合学习结构化综合字典和结构化解析字典,然后通过样本在解析字典上的映射直接求解稀疏系数矩阵;同时采用 Fisher 判别准则编码系数使系数具有一定的判别性。最后将新方法应用到图像分类中,实验结果表明新方法在提高分类准确率的同时还大大降低了计算复杂度,相较于现有方法具有更好的性能。

**关键词:** 图像分类; 稀疏表示; 字典对; Fisher 约束

中图分类号: TP391

文献标识码: A

文章编号: 1009-5896(2017)02-0270-08

DOI: 10.11999/JEIT160296

## Image Classification Based on Fisher Constraint and Dictionary Pair

GUO Jichang ZHANG Fan WANG Nan

(School of Electronic Information Engineering, Tianjin University, Tianjin 300072, China)

**Abstract:** Classification method based on sparse representation has won wide attention because of its simplicity and effectiveness, while how to adaptively build the relationship between dictionary atoms and class labels is still an important open question, at the same time most of the sparse representation classification methods need to solve a norm constraint optimization problem, which increases the computational complexity in the classification task. To address this issue, this paper proposes a novel Fisher constraint dictionary pair learning method to jointly learn a structured synthesis dictionary and a structured analysis dictionary, then directly obtains the sparse coefficient matrix by analysis dictionary. In this paper, the Fisher criterion is used to encode the coefficients. Finally the new method is applied to image classification task, the experimental results show that the new method not only improves the accuracy of classification but also greatly reduces the computational complexity. Compared with the existing methods, the new method has better performance.

**Key words:** Image classification; Sparse representation; Dictionary pair; Fisher constraint

### 1 引言

在对基于稀疏表示的分类方法进行研究的过  
程中, 研究者们发现字典性能的好坏直接关系到最后  
的分类结果<sup>[1-3]</sup>。字典最早用于解决信号重构的相  
关问题, 它旨在获得一个具有良好表示能力的字典  
而没有考虑字典的判别性, 因此并不利于分类问  
题<sup>[4]</sup>。近年来的研究表明判别性字典更加适合用于图  
像分类任务, 此类字典既要被用来对样本进行稀疏

编码, 又要被用于执行最后的分类判别。判别性字典大致可分为两类: 第 1 类是直接学习具有判别性的字典<sup>[5-7]</sup>, 文献[5]对不同类别的子字典添加低秩约束来学习一个具有判别性的字典, 文献[6]考虑不同类别子字典间的非一致性信息提出了结构不相干字典学习方法, 文献[7]提出同时学习一个所有类别的公共字典和多个特定类别的子字典的字典学习方法; 第 2 类是对稀疏系数添加约束项来提升系数的判别性, 进而获得具有判别性的字典<sup>[8,9]</sup>, 文献[8]在字典学习过程中对系数添加样本的类别信息提出了标签一致 K 奇异值分解(Label Consistent K-Singular Value Decomposition, LC-KSVD)方法, 文献[9]提出了对系数添加 Fisher 判别约束的字典学习方法。

上述字典学习方法均属于综合型字典学习。综合型字典学习方法一直是研究者关注的热点, 因此

收稿日期: 2016-03-31; 改回日期: 2016-07-25; 网络出版: 2016-10-09

\*通信作者: 郭继昌 jcguo@tju.edu.cn

基金项目: 国家 973 计划项目(2014CB 340400), 天津市自然科学基金(15JCYBJC15500)

Foundation Items: The National 973 Program of China (2014CB340400), The Natural Science Foundation of Tianjin (15JCYBJC15500)

基于综合模型的字典学习理论相对比较成熟<sup>[10]</sup>, 但是综合型字典学习方法的计算量都较大, 因此研究者们又提出了解析型字典学习方法<sup>[11,12]</sup>。解析模型中通过学习到的解析字典来获得样本的稀疏编码, 具有较高的编码效率。与综合模型相比, 在维数相同同时解析模型的子空间数较多, 有更灵活和更丰富的表示能力。

本文将解析字典学习模型和综合字典学习模型结合起来, 提出一种新的字典学习方法, 称之为基于 Fisher 判别约束的字典对学习方法。新方法一方面联合学习由结构化综合字典和结构化解析字典构成的字典对, 并对字典添加判别性约束使得学习到的字典对具有判别性; 另一方面通过样本在解析字典上的映射直接求解稀疏系数矩阵, 避免了对标准稀疏编码问题的求解, 同时采用 Fisher 判别准则编码系数使系数具有一定的判别性。最后将新模型应用到人脸图像库, 目标图像库和场景图像库上进行评估。

## 2 字典学习模型及字典对模型

### 2.1 综合型字典学习模型

综合型字典学习模型中  $\mathbf{Y}=[\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N] \in \mathbb{R}^{m \times N}$  是  $m$  维,  $N$  个训练样本的集合, 其优化模型可以表示为

$$J_{(D, X)} = \arg \min_{D, X} \left\{ \|\mathbf{Y} - \mathbf{D}\mathbf{X}\|_2^2 + \lambda \|\mathbf{X}\|_1 \right\},$$

$$\text{s. t. } \|\mathbf{d}_k\|_2 \leq 1, \forall k = 1, 2, \dots, t \quad (1)$$

其中  $\mathbf{D}=[\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_t] \in \mathbb{R}^{m \times t}$  是包含  $t$  个字典原子的综合型字典, 训练样本  $\mathbf{Y}$  在综合字典  $\mathbf{D}$  下的系数矩阵表示为  $\mathbf{X}=[\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N] \in \mathbb{R}^{t \times N}$ 。图 1(a) 是单个样本情况下的综合型字典示意图。

### 2.2 解析型字典学习模型

解析型字典学习是综合型字典学习的对偶形式, 它的核心思想是通过学习到的解析字典直接获得样本在新映射空间的稀疏表示, 解析型字典如图 1(b) 所示。优化模型可以表示为式(2),  $\mathbf{P}=[\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_t]^T \in \mathbb{R}^{t \times m}$  为解析型字典,  $\varphi(\cdot)$  是度量稀疏度的函数。

$$J_{(P)} = \arg \min_P \varphi(\mathbf{P}\mathbf{Y}), \quad \text{s. t. } \varphi(\mathbf{P}\mathbf{Y}) = \|\mathbf{P}\mathbf{Y}\|_p \quad (2)$$

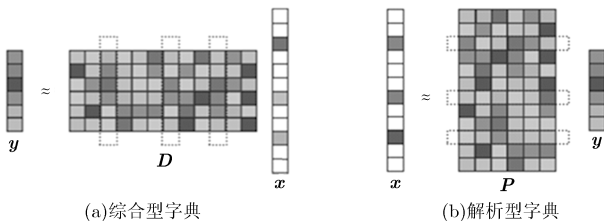


图1 字典模型示意图

### 2.3 字典对模型

基于解析型字典学习的稀疏表示研究处于起步阶段, 还没有较多地应用到稀疏表示分类问题中。比较有代表性的是文献[13]中提出的字典对学习(Dictionary Pair Learning, DPL)模型。DPL模型通过样本在解析字典上的映射直接求解稀疏系数矩阵, 避免了对受范数约束的标准稀疏编码问题的求解, 从而大大减少了分类任务的计算复杂度。设矩阵  $\mathbf{a} \in \mathbb{R}^{m \times n}, \mathbf{b} \in \mathbb{R}^{m \times p}, \mathbf{c} \in \mathbb{R}^{t \times n}$ , 本文中用  $[\mathbf{a}, \mathbf{b}] \in \mathbb{R}^{m \times (n+p)}$  表示将矩阵  $\mathbf{a}$  和  $\mathbf{b}$  按列进行合并, 用  $[\mathbf{a}; \mathbf{c}] \in \mathbb{R}^{(m+t) \times n}$  表示将矩阵  $\mathbf{a}$  和  $\mathbf{c}$  按行进行合并。DPL模型中用  $\mathbf{Y}=[\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_K] \in \mathbb{R}^{m \times Kn}$  表示  $K$  类训练样本, 其中每类训练样本的个数为  $n$ 。求解 DPL 模型得到综合字典  $\mathbf{D}=[\mathbf{D}_1, \mathbf{D}_2, \dots, \mathbf{D}_K] \in \mathbb{R}^{m \times Kt}$  和解析字典  $\mathbf{P}=[\mathbf{P}_1; \mathbf{P}_2; \dots; \mathbf{P}_K] \in \mathbb{R}^{Kt \times m}$ 。其中  $\{\mathbf{D}_i \in \mathbb{R}^{m \times t}, \mathbf{P}_i \in \mathbb{R}^{t \times m}\}$  是第  $i$  类样本对应的子字典对。在 DPL 模型中第  $i$  类的解析子字典  $\mathbf{P}_i$  对第  $j$  类样本的映射集合几乎为空集, 即  $\mathbf{P}_i \mathbf{Y}_j \approx 0, \forall i \neq j$ , 用字典对表示的最小化重构误差为

$$\min_{P, D} \sum_{i=1}^K \|\mathbf{Y}_i - \mathbf{D}_i \mathbf{P}_i \mathbf{Y}_i\|_2^2 \quad (3)$$

## 3 基于 Fisher 约束的字典对学习模型

DPL 模型利用解析字典大大减少了计算复杂度, 但是其并没有充分利用训练样本所具有类别信息, 同时在字典学习模型中也没有考虑不同类别样本对应的字典原子和系数矩阵间的相关性信息。为克服以上缺点, 本文提出了一种新的字典学习模型, 称之为基于 Fisher 约束的字典对学习(Fisher constraint Dictionary Pair Learning, FDPL)模型。该模型同时学习结构化综合字典和结构化解析字典, 并且在字典学习过程中对解析字典添加判别性约束来衡量解析子字典间存在的相关性信息, 同时对系数添加 Fisher 约束来衡量系数矩阵间的相关性信息。FDPL 模型的优化问题可以表示为

$$J_{(D, P, X)} = \arg \min_{D, P, X} \left\{ \sum_{i=1}^K \left( \phi(\mathbf{Y}_i, \mathbf{D}_i, \mathbf{P}_i, \mathbf{X}_i) + \tau h(\mathbf{D}_i, \mathbf{P}_i) + \omega f(\mathbf{X}_i) \right) \right\} \quad (4)$$

其中,  $\phi(\mathbf{Y}_i, \mathbf{D}_i, \mathbf{P}_i, \mathbf{X}_i)$  是误差项,  $h(\mathbf{D}_i, \mathbf{P}_i)$  是对字典添加的判别性约束项,  $f(\mathbf{X}_i)$  是对系数添加的判别性约束项,  $\tau$  和  $\omega$  是各项的权重系数, 下面讨论各项的设计问题。

### 3.1 误差项的设计

FDPL 模型中对训练样本和字典均采用结构化表示方式。结构化表示方式下解析字典和系数的优化问题以及综合字典和系数的优化问题分别表示为

$$\left. \begin{aligned} J_{(P,X)} &= \arg \min_{P,X} \sum_{i=1}^K \left\| P_i Y_i - X_i \right\|_2^2, \\ J_{(D,X)} &= \arg \min_{D,X} \sum_{i=1}^K \left\| Y_i - D_i X_i \right\|_2^2 \end{aligned} \right\} \quad (5)$$

因此 FDPL 模型的重构误差项和编码误差项为

$$\phi(Y_i, D_i, P_i, X_i) = \|Y_i - D_i X_i\|_2^2 + \lambda \|P_i Y_i - X_i\|_2^2 \quad (6)$$

### 3.2 字典判别性约束项的设计

FDPL 模型希望第  $i$  类的解析子字典  $P_i$  对第  $i$  类的训练样本有很好的表示能力, 而对第  $j$  类样本的表示能力较差, 即  $P_i Y_i \gg P_i Y_j, \forall i \neq j$ 。因此可以用  $P_i \bar{Y}_i$  来衡量解析子字典间的相关性, 即将最小化  $\|P_i \bar{Y}_i\|_2^2$  作为解析字典的判别性约束项, 其中  $\bar{Y}_i$  是训练样本  $Y$  去掉第  $i$  类样本后的剩余矩阵。与此同时限制综合字典  $D_i$  中每个原子的能量避免  $P_i$  为零, 使 FDPL 模型更加稳定。

### 3.3 系数判别性约束项的设计

文献[9]中指出 Fisher 约束可以用来衡量样本间的相似性信息, 具体是通过最小化系数的类内离散度和最大化系数的类间离散度来实现的。设稀疏系数矩阵为  $X = [X_1, \dots, X_i, \dots, X_K]$ , 其中  $X_i = [X_i^1, \dots, X_i^j, \dots, X_i^K]$  为第  $i$  类样本对应的系数矩阵,  $X_i^j$  表示第  $j$  类字典对第  $i$  类样本的编码矩阵。 $X$  的类内离散度、类间离散度为

$$\left. \begin{aligned} S_W(X) &= \sum_{i=1}^K \sum_{x_p \in X_i} (x_p - m_i)(x_p - m_i)^T \\ S_B(X) &= \sum_{i=1}^K n(m_i - m)(m_i - m)^T \end{aligned} \right\} \quad (7)$$

对于每类的系数矩阵  $X_i$ , 其 Fisher 约束项可以表示为

$$\begin{aligned} f(X_i) &= \text{tr}(S_W(X_i)) - \text{tr}(S_B(X_i)) + \eta \|X_i\|_2^2 \\ &= \|X_i - M_i\|_2^2 - \sum_{i=1}^K \|M_i - M\|_2^2 + \eta \|X_i\|_2^2 \end{aligned} \quad (8)$$

设单位列向量  $\mathbf{1} \in \mathbb{R}^{t \times 1}$ , 则  $m_i = X_i \cdot \mathbf{1} / n$ ;  $m = \sum_{i=1}^K m_i / K$ ;  $M_i = m_i \cdot \mathbf{1}^T$ ;  $M = m \cdot \mathbf{1}^T$ 。将上述各项综合起来得到 FDPL 模型的优化公式, 其中  $\lambda$ ,  $\tau$ ,  $\omega$  和  $\eta$  是各项的权重系数。

$$\left. \begin{aligned} J_{(D,P,X)} &= \arg \min_{D,P,X} \left\{ \sum_{i=1}^K \left( \|Y_i - D_i X_i\|_2^2 \right. \right. \\ &\quad \left. \left. + \lambda \|P_i Y_i - X_i\|_2^2 + \tau \|P_i \bar{Y}_i\|_2^2 \right) \right\} \\ &\quad + \omega \left( \|P_i Y_i - M_i\|_2^2 - \sum_{i=1}^K \|M_i - M\|_2^2 \right. \\ &\quad \left. + \eta \|P_i Y_i\|_2^2 \right), \\ \text{s.t. } \|d_k\|_2^2 &\leq 1, \forall k = 1, 2, \dots, t \end{aligned} \right\} \quad (9)$$

### 3.4 模型求解

求解 FDPL 模型可以采用迭代优化的方法, 将优化求解分为固定字典对  $\{D, P\}$  更新系数  $X$  和固定系数  $X$  更新字典对  $\{D, P\}$  两个子问题。

(1) 首先固定字典对  $\{D, P\}$  更新系数  $X$ , 并且在求解  $X_i$  时保持所有的  $X_j (i \neq j)$  均固定。

$$\begin{aligned} J_{(X_i)} &= \arg \min_{X_i} \left\{ \sum_{i=1}^K \left( \|Y_i - D_i X_i\|_2^2 + \lambda \|P_i Y_i - X_i\|_2^2 \right) \right\} \\ &\quad + \omega \left( \|P_i Y_i - M_i\|_2^2 - \sum_{i=1}^K \|M_i - M\|_2^2 \right) \end{aligned} \quad (10)$$

式(10)为标准最小二乘问题, 求其闭式解得到系数  $X_i^*$ , 其中单位矩阵  $I \in \mathbb{R}^{t \times t}$ 。

$$\begin{aligned} X_i^* &= \left[ D_i^T D_i + \left( \lambda - \frac{\omega}{nK^2} + \frac{2\omega}{nK} \right) I \right]^{-1} \\ &\quad \times \left[ D_i^T D_i + \lambda P_i Y_i + \frac{\omega}{n} P_i Y_i \cdot \mathbf{1} \cdot \mathbf{1}^T \right. \\ &\quad \left. + \omega \left( \frac{1}{nK^2} - \frac{1}{nK} \right) \sum_{i=1}^K m_i \cdot \mathbf{1}^T \right] \end{aligned} \quad (11)$$

(2) 其次固定系数  $X$  更新字典对  $\{D, P\}$ , 并且更新第  $i$  类的字典对时第  $j$  类 ( $i \neq j$ ) 的字典对保持不变。

$$\left. \begin{aligned} J_{(P)} &= \arg \min_P \left\{ \sum_{i=1}^K \left( \lambda \|P_i Y_i - X_i\|_2^2 + \tau \|P_i \bar{Y}_i\|_2^2 \right) \right\} \\ &\quad + \omega \left( \|P_i Y_i - M_i\|_2^2 + \eta \|P_i Y_i\|_2^2 \right) \\ J_{(D)} &= \arg \min_D \sum_{i=1}^K \|Y_i - D_i X_i\|_2^2, \\ \text{s.t. } \|d_k\|_2^2 &\leq 1, \forall k = 1, 2, \dots, t \end{aligned} \right\} \quad (12)$$

其中  $P$  可以通过求闭式解进行更新:

$$\begin{aligned} P_i^* &= (\lambda X_i Y_i^T + \omega M_i Y_i^T) \\ &\quad \cdot [(\lambda + \omega(1 + \eta)) Y_i Y_i^T + \tau \bar{Y}_i \bar{Y}_i^T]^{-1} \end{aligned} \quad (13)$$

更新  $D$  时采用交替方向乘法 (Alternating Direction Method of Multipliers, ADMM)<sup>[14]</sup>, 相较于对子字典中每个字典原子进行逐个更新的方法, ADMM 方法可以对每个类别的子字典进行整体更新, 效率更高。具体是通过引入一个辅助矩阵  $S$  实现的, 此时式(12)的第2项变为

$$\left. \begin{aligned} J_{(D,S)} &= \arg \min_{D,S} \sum_{i=1}^K \|Y_i - D_i X_i\|_2^2 \\ \text{s.t. } D &= S \|s_k\|_2^2 \leq 1, \forall k = 1, 2, \dots, t \end{aligned} \right\} \quad (14)$$

根据 ADMM 方法对第  $i$  类子字典进行整体更新的过程为

$$\left. \begin{aligned} D_i^{(u+1)} &= (\rho (S_i^{(u)} - T_i^{(u)}) + Y_i X_i^T) (\rho I + X_i X_i^T)^{-1} \\ S_i^{(u+1)} &= \text{normal}(D_i^{(u+1)} + T_i^{(u)}) \\ T_i^{(u+1)} &= T_i^{(u)} + D_i^{(u+1)} - S_i^{(u+1)} \end{aligned} \right\} \quad (15)$$

式(15)中  $u$  是 ADMM 方法在更新子字典  $D_i$  时的迭代次数, normal 表示对求得的矩阵进行列归一化来满足  $\|s_k\|_2 \leq 1$ 。  $T_i$  是 ADMM 方法求解过程中产生的中间变量。 FDPL 模型的字典学习过程如表 1 所示。

表 1 FDPL模型的字典学习过程

输入: 训练样本 $\mathbf{Y} = [\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_K]$ , 参数 $\lambda, \tau, \omega, \eta$ , 最大迭代次数 $c$ 。
初始化: 随机生成初始化综合字典 $\mathbf{D}$ 和解析字典 $\mathbf{P}$ , 根据初始化字典初始系数 $\mathbf{X}$ 。
执行循环(循环所有的类别 $i = 1, 2, \dots, K$ ):
(1)更新系数 $\mathbf{X}_i^{c+1} \leftarrow \mathbf{P}_i^c, \mathbf{D}_i^c$ , 根据式(11);
(2)更新解析字典 $\mathbf{P}_i^{c+1} \leftarrow \mathbf{X}_i^{c+1}, \mathbf{D}_i^c$ , 根据式(13);
(3)更新解析字典 $\mathbf{D}_i^{c+1} \leftarrow \mathbf{X}_i^{c+1}$ , 根据式(15)。直到目标函数收敛或达到最大迭代次数, 循环结束。
输出: 综合字典 $\mathbf{D} = [\mathbf{D}_1, \mathbf{D}_2, \dots, \mathbf{D}_K]$ 和解析字典 $\mathbf{P} = [\mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_K]$ 。

### 3.5 分类准则

本文采用两种分类准则, 第 1 种是将样本的重构误差作为分类准则, 在这种分类准则下直接利用训练阶段求得的解析字典来获得测试样本的稀疏编码; 第 2 种是同时考虑样本的重构误差和编码误差进行分类, 此时根据训练阶段获得的综合字典通过求解标准稀疏表示问题获得测试样本的稀疏编码, 再与解析字典获得的稀疏编码相比较作为编码误差。

(1)根据重构误差进行分类: 通过求解 FDPL 模型得到每个类别对应的字典对  $\{\mathbf{D}_i^*, \mathbf{P}_i^*\}$ , 解析型子字典  $\mathbf{P}_i^*$  只对它同类别的样本具有很好的表示能力, 同时综合型子字典  $\mathbf{D}_i^*$  可以根据编码系数  $\mathbf{P}_i^* \mathbf{Y}_i$  对第  $i$  类的样本进行重构, 此时的重构误差  $\|\mathbf{Y}_i - \mathbf{D}_i^* \mathbf{P}_i^* \mathbf{Y}_i\|_2^2$  较小, 但当  $i \neq j$  时  $\|\mathbf{Y}_j - \mathbf{D}_i^* \mathbf{P}_i^* \mathbf{Y}_j\|_2^2$  的值较大。因此在测试阶段如果待测试样本  $\mathbf{y}$  属于第  $i$  类则  $\|\mathbf{y} - \mathbf{D}_i^* \mathbf{P}_i^* \mathbf{y}\|_2^2 \ll \|\mathbf{y} - \mathbf{D}_j^* \mathbf{P}_j^* \mathbf{y}\|_2^2 (i \neq j)$ , 显然根据重构误差可以用来确定测试样本的类别  $L(\mathbf{y})$ :

$$e_{i,1} = \|\mathbf{y} - \mathbf{D}_i^* \mathbf{P}_i^* \mathbf{y}\|_2^2 \quad (16)$$

$$L(\mathbf{y}) = \arg \min_i \{e_{i,1}\} (i = 1, 2, \dots, K) \quad (17)$$

(2)根据重构误差和编码误差进行分类: 由综合字典  $\mathbf{D}$  求测试样本  $\mathbf{y}$  的稀疏系数  $\mathbf{x}^* = [\mathbf{x}_1^*, \mathbf{x}_2^*, \dots, \mathbf{x}_K^*] \in \mathbb{R}^{K \times 1}$ , 其中  $\mathbf{x}_i^* \in \mathbb{R}^{1 \times 1}$  是在综合型子字典  $\mathbf{D}_i^*$  下求得的系数。令  $\mathbf{x}_i = [0, \dots, \mathbf{x}_i^*, \dots, 0] \in \mathbb{R}^{K \times 1}$ , 如果

待测试样本  $\mathbf{y}$  属于第  $i$  类则  $\mathbf{P}_i^* \mathbf{y} \gg \mathbf{P}_j^* \mathbf{y} (i \neq j)$ , 因此由解析字典  $\mathbf{P}$  得到的编码系数  $\mathbf{P} \mathbf{y}$  将近似于  $\mathbf{x}_i$ , 即  $\mathbf{P} \mathbf{y} \approx \mathbf{x}_i$ 。显然编码误差也可以用来确定测试样本的类别:

$$e_{i,2} = \omega_1 \|\mathbf{y} - \mathbf{D}_i^* \mathbf{P}_i^* \mathbf{y}\|_2^2 + \omega_2 \|\mathbf{P} \mathbf{y} - \mathbf{x}_i\|_2^2 + \omega_3 \|\mathbf{x}_i^* - \mathbf{m}_i\|_2^2 \quad (18)$$

$$L(\mathbf{y}) = \arg \min_i \{e_{i,2}\} (i = 1, 2, \dots, K) \quad (19)$$

$\omega_1, \omega_2$  和  $\omega_3$  是各项的权重系数,  $\|\mathbf{x}_i^* - \mathbf{m}_i\|_2^2$  是系数向量  $\mathbf{x}_i^*$  和第  $i$  类的均值向量  $\mathbf{m}_i$  之间的距离。

## 4 实验及分析

为了验证 FDPL 算法的性能, 本文不仅比较 FDPL 算法在两种分类准则下的性能, 而且将 FDPL 方法与已有的效果较好的字典学习方法进行比较, 例如 LC-KSVD<sup>[8]</sup>, Fisher 判别约束字典学习 (Fisher Discrimination Dictionary Learning, FDDL)<sup>[9]</sup>、支持向量引导的字典学习 (Support Vector Guided Dictionary Learning, SVGDL)<sup>[15]</sup> 和 DPL<sup>[13]</sup>。衡量分类模型性能好坏时既要考虑其分类准确率又要考虑其分类效率, 因此实验中采用两种衡量指标: 一是分类准确率, 二是分类复杂度。

### 4.1 图像库

Extended YaleB 图像库包含 38 人共 2414 张图像, 每人随机选取 32 张图像作为训练集, 其余作为测试集, 图像特征采用 504 维的随机脸。AR 图像库包含 126 人共 4000 多张图像, 选取 50 个男性和 50 个女性共计 2600 张图像来进行实验, 每人随机选取 20 张图像作为训练集其余作为测试集, 图像特征采用 540 维的随机脸。Caltech-101 图像库共有 102 类 9144 张图像, 每类图像个数从 31 到 800 不等, 每类随机选取 30 张图像作为训练集其余作为测试集。Scene15 图像库共有 15 类 4485 张图像, 每类图像个数从 200 到 400 不等, 每类随机选取 100 张图像作为训练集其余作为测试集。在 Caltech-101 和 Scene15 图像库上首先按  $16 \times 16$  的图像块、6 像素的步长提取图像的原始稠密 SIFT 特征; 然后对原始特征进行空间金字塔最大池化, 得到 SIFT 池化特征; 通过 K-means 方法对池化特征进行稀疏编码, 对每幅图像的所有稀疏编码运用空间金字塔最大化池方法; 最后通过 PCA 降维得到 3000 维的含有空间信息的 SIFT 特征。

### 4.2 参数设置

$\lambda, \tau, \omega, \eta$  是算法设计过程中涉及到的参数,  $\lambda$  是编码误差项的系数;  $\tau$  是解析字典约束项的系数, 用于控制解析字典  $\mathbf{P}$  的判别力;  $\omega$  是系数判别性约

束项的系数； $\eta$  是系数判别性约束项的内部系数。实验中采用十倍交叉验证的方法来确定各项的值。FDPL 方法在各个图像库上的参数设置如表 2 所示。

表 2 FDPL 算法的主要参数

参数	Extended YaleB	AR	Caltech-101	Scene15
$\lambda$	0.050	0.050	0.010	0.010
$\tau$	0.003	0.005	0.001	0.001
$\omega$	0.001	0.001	0.001	0.001
$\eta$	0.005	0.005	0.010	0.010

实验中的最大迭代次数  $c$  是根据目标函数的收敛性曲线来选取的。FDPL 模型的目标函数是关于  $\mathbf{X}$  和  $\{\mathbf{D}, \mathbf{P}\}$  的双凸问题，即固定  $\mathbf{X}$  时是关于  $\{\mathbf{D}, \mathbf{P}\}$  的凸函数，固定  $\{\mathbf{D}, \mathbf{P}\}$  时是关于  $\mathbf{X}$  的凸函数。对于双凸问题的收敛性已在文献[16]中进行了详细讨论。本文的 FDPL 优化算法是一种交替凸搜索算法，因此可以通过更新  $\mathbf{X}, \mathbf{D}, \mathbf{P}$  获得最优解，并且目标函数值具有下界 0。图 2 是在 AR 人脸图像库上的收敛曲线，从图中可以看出目标函数值下降很快，在第 4 次迭代后已经很小。在其它图像库上的实验也表明 FDPL 算法的收敛性很好，第 20 次迭代前目标函数值已经很小，因此在实验中将迭代次数设置为 20。

#### 4.3 FDPL 方法在不同分类准则下的分类性能对比

3.5 节对两种分类准则进行了具体描述，第 1 种分类准则直接利用解析字典求解稀疏系数，第 2 种分类准则在求编码误差时仍然需要求解标准稀疏编码问题来获得待测样本在综合字典下的稀疏系数。

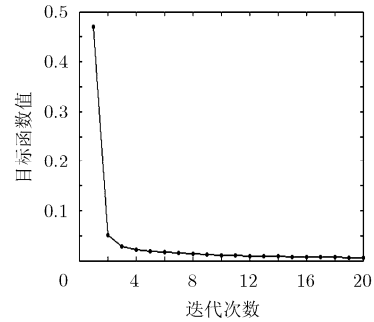


图 2 AR 图像库上目标函数的收敛性曲线

首先对比表 3 中的分类准确率可以看出，第 2 种分类准则下的分类准确率较高但提升不是很明显；然后对比分类用时可以看出，第 1 种分类准则在分类用时方面具有明显优势。衡量分类模型性能好坏时既要考虑其分类准确率又要考虑其分类效率，综合评定可知第 1 种分类准则的优势更为明显，也说明了采用样本在解析字典上的投影直接获得稀疏系数矩阵的方法具有高效性和可行性。

#### 4.4 字典原子个数不同时分类性能对比

已有研究表明，字典学习过程中每类选取的字典原子个数直接关系到分类性能的好坏，因此将其设置为实验过程中的变量。LC-KSVD 和 SVGDL 算法中限制总的字典原子个数  $Kt$  不大于选取的训练样本个数  $Kn$ ，在 AR 图像库上每类选取的训练样本个数为 20，因此在字典原子个数为 30 时没有对应的数据。改变字典原子个数时的分类准确率如表 4 所示，黑体标注为相同字典原子个数下的最高分类准确率。表 5 是各种方法保持硬件条件一致的情况下的分类用时比较。

表 3 FDPL 方法在不同分类准则下的分类准确率(%)和分类用时(s)

图像库	字典原子个数	准确率 ( $e_{i,1}$ )	准确率 ( $e_{i,2}$ )	训练用时 ( $e_{i,1}$ )	训练用时 ( $e_{i,2}$ )	分类用时 ( $e_{i,1}$ )	分类用时 ( $e_{i,2}$ )
Extended YaleB	10	94.98	95.17	25.02	26.43	0.31	1138.01
	20	96.28	96.68	26.08	27.77	0.33	2060.59
	30	96.49	97.27	32.25	32.96	0.33	3217.47
AR	10	98.23	98.47	106.27	105.33	0.26	748.56
	20	98.54	98.81	107.98	110.28	0.30	1849.46
	30	98.50	99.00	109.45	113.56	0.30	2602.37
Caltech-101	10	74.25	74.22	4277.09	4275.13	25.02	28057.86
	20	74.42	74.87	4624.72	4542.56	26.13	56259.43
	30	74.53	75.01	4980.18	4980.18	28.29	70398.28
Scene15	10	98.01	98.11	585.42	587.31	1.4	1107.83
	30	98.27	98.66	566.91	563.57	1.54	2652.45
	50	98.45	98.75	623.61	643.85	1.75	3686.68

表 4 不同原子个数下的分类准确率(%)

图像库	字典原子个数	LC-KSVD	FDDL	SVGDL	DPL	FDPL ( $\epsilon_{i,1}$ )	FDPL ( $\epsilon_{i,2}$ )
Extended YaleB	10	91.32	93.37	93.67	94.37	94.98	<b>95.17</b>
	20	93.78	95.22	95.53	96.24	96.28	<b>96.68</b>
	30	95.18	95.24	95.41	96.44	96.49	<b>97.27</b>
AR	10	94.52	93.37	97.23	97.17	98.23	<b>98.47</b>
	20	97.74	94.79	95.74	97.33	98.54	<b>98.81</b>
	30	-	93.42	-	97.33	98.50	<b>99.00</b>
Caltech-101	10	61.52	73.34	73.91	73.16	<b>74.25</b>	74.22
	20	69.91	73.36	72.35	73.27	74.42	<b>74.87</b>
	30	73.41	73.43	73.91	73.37	74.53	<b>75.01</b>
Scene15	10	92.64	97.76	97.81	97.31	98.01	<b>98.11</b>
	30	94.54	97.84	98.22	97.74	98.27	<b>98.66</b>
	50	96.94	97.76	98.13	97.74	98.45	<b>98.75</b>

表 5 不同原子个数下的分类用时(s)

图像库	字典原子个数	LC-KSVD	FDDL	SVGDL	DPL	FDPL ( $\epsilon_{i,1}$ )	FDPL ( $\epsilon_{i,2}$ )
Extended YaleB	10	29.51	3446.90	213.78	28.08	25.33	1164.44
	20	69.01	6029.52	943.18	29.18	26.41	2088.36
	30	138.64	10422.85	2267.74	30.97	32.58	3250.43
AR	10	86.83	3752.31	4409.23	100.64	106.53	853.89
	20	475.46	7324.16	14435.25	102.24	108.28	1959.74
	30	-	12564.71	-	106.71	109.75	2715.93
Caltech-101	10	215.23	46772.37	8644.76	4189.74	4302.11	32332.99
	20	511.73	91735.91	23595.60	4615.96	4650.85	60801.99
	30	801.34	124304.00	71675.16	4786.12	5008.47	75373.57
Scene15	10	59.73	34017.51	67.50	560.77	586.82	1695.14
	30	191.22	95168.72	498.50	563.69	568.45	3216.02
	50	356.67	161681.10	1467.37	620.32	625.36	4330.53

对比表中每列的数据, 为了说明本文算法的有效性首先与 LC-KSVD, SVGDL 算法进行对比, 通过 LC-KSVD 和 SVGDL 算法获得的是所有类别通用的综合字典, 并且在字典学习过程中没有考虑字典原子间的相关性信息, 因此采用 LC-KSVD 和 SVGDL 算法获得的字典的判别性能较差, 进而造成分类准确率较低。为了进一步论证本文算法的有效性, 将与 FDDL 和 DPL 算法进行对比, FDDL 算法通过对系数添加 Fisher 约束来获得具有判别性的字典, 但该方法没有考虑字典原子间的相关性信息, DPL 算法联合学习综合字典和解析字典, 但是该方法并没有考虑系数间的相关性信息, 因此采用 FDDL, DPL 算法进行图像分类的效果也较差。通过对比发现 FDPL 算法具有较高的分类准确率。

4.3 节中通过对比 FDPL 方法在两种分类准则下的分类用时, 可知第 2 种分类准则在时间消耗上并不具有优势, 因此在与其它方法进行对比时主要考虑第 1 种分类准则下的分类用时。与其它方法相比 FDPL 方法分类用时明显较少并且受字典原子个数的影响较小。分析上述结果原因主要有以下两点: 在训练阶段进行迭代更新时 FDPL 方法需要求解的是标准最小二乘问题, 该问题可以通过解析方法求其闭式解, 从而避免了处理 FDDL, LC-KSVD 和 SVGDL 方法中受范数约束的稀疏编码问题; 在分类阶段 FDPL 方法通过样本在解析字典上的映射直接求解稀疏系数矩阵, 也避免了对标准稀疏编码问题的求解, 因此用时较短。

然而通过运行时间分析复杂度并不可靠, 接下

来将从理论上对 FDPL 方法的计算复杂度进行分析。采用迭代更新的方法更新  $\mathbf{X}_i, \mathbf{P}_i, \mathbf{D}_i$  的时间复杂度分别为  $O(mnt + t^3 + nt^2)$ ,  $O(mnt + m^3 + m^2n + m^2t)$ ,  $O(U(mnt + t^3 + mt^2 + m^2t))$ ,  $U$  是采用 ADMM 方法更新字典  $\mathbf{D}$  时的迭代次数, 通常  $U \leq 20$ 。在实际应用中每类训练样本个数  $n$  和字典原子个数  $t$  都远远小于特征维数  $m$ , 因此训练阶段的计算复杂度主要是由更新  $\mathbf{P}_i$  引起的。第 1 种分类准则下重构误差项的计算复杂度为  $O(mt)$ , 分类阶段总的复杂度为  $O(Kmt)$ 。相较于求解标准稀疏编码问题, FDPL 方法具有较小的计算复杂度且受字典原子个数和训练样本个数影响较小。

#### 4.5 稀疏性分析

将实验过程中得到的样本的系数矩阵提取出来进行稀疏性分析, 以 Extended YaleB 为例。图 3 为在每类字典原子个数为 10 的情况下的稀疏性分析。

从图 3 可知第  $i$  类的样本只在第  $i$  类的字典处具有较高的稀疏度值, 验证了解析子字典  $\mathbf{P}_i$  对同类训练样本有很好的表示能力, 而对其它类的训练样本

表示能力则较差, 因此可以利用综合型子字典  $\mathbf{D}_i^*$  根据编码系数  $\mathbf{P}_i^* \mathbf{Y}_i$  对第  $i$  类样本进行重构, 此时重构误差较小, 同时验证了重构误差可以用来确定测试样本的类别。

#### 5 结束语

本文提出了基于 Fisher 约束和字典对学习的图像分类方法, FDPL 方法一方面既考虑字典判别性又考虑系数判别性, 使得所得字典和系数均具有良好的判别能力; 另一方面通过样本在解析字典上的映射直接求解稀疏系数矩阵, 避免了对标准稀疏编码问题的求解, 减少了计算复杂度。在多个图像库上对 FDPL 方法进行了实验验证并与现有效果较好的方法进行了比较, 实验结果表明 FDPL 方法克服了当下主流字典学习算法中计算复杂、训练速度慢的缺点, 在总体上缩短了图像分类时间, 并且具有更高的分类准确率。本文并未考虑图像特征表示这一因素对分类性能的影响, 采用了已有的提取图像 SIFT 特征的方法, 因此在未来的研究中将会寻求更好的图像特征表示方法, 然后与本文方法相结合。

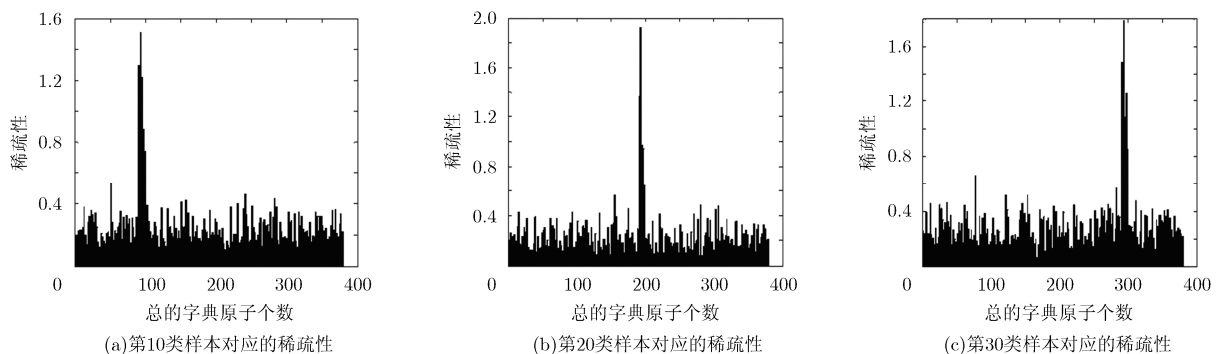


图 3 稀疏性分析

#### 参考文献

- [1] RUBINSTEIN R, BRUCKSTEIN A, ELAD M, *et al.* Dictionaries for sparse representation modeling[J]. *Proceedings of the IEEE*, 2010, 98(6): 1045-1057. doi: 10.1109/JPROC.2010.2040551.
- [2] GAO Shenghua, TSANG I, and MA Yi. Learning category-specific dictionary and shared dictionary for fine-grained image categorization[J]. *IEEE Transactions on Image Processing*, 2014, 23(2): 623-634. doi: 10.1109/TIP.2013.2290593.
- [3] 宋相法, 焦李成. 基于稀疏编码和集成学习的多示例多标记图像分类方法[J]. *电子与信息学报*, 2013, 35(3): 622-626. doi: 10.3724/SP.J.1146.2012.01218.
- SONG Xiangfa and JIAO Licheng. A multi-instance multi-label image classification method based on sparse coding and ensemble learning[J]. *Journal of Electronics & Information Technology*, 2013, 35(3): 622-626. doi: 10.3724/SP.J.1146.2012.01218.
- [4] AHARON M, ELAD M, and BRUCKSTEIN A. K-SVD: an algorithm for designing of overcomplete dictionaries for sparse representation[J]. *IEEE Transactions on Signal Processing*, 2006, 54(11): 4311-4322. doi: 10.1109/TSP.2006.881199.
- [5] MA Long, WANG Chunheng, XIAO Baihua, *et al.* Sparse representation for face recognition based on discriminative low-rank dictionary learning[C]. *IEEE Conference on Computer Vision and Pattern Recognition*, Providence, USA, 2012: 2586-2593.
- [6] RAMIREZ I, SPRECHMANN P, SAPIRO G, *et al.*

- Classification and clustering via dictionary learning with structured incoherence and shared features[C]. IEEE Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA, 2010: 3501–3508.
- [7] NING Zhou and FAN Jianping. Jointly learning visually correlated dictionaries for large-scale visual recognition applications[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014, 36(4): 715–730. doi: 10.1109/TPAMI.2013.189.
- [8] JIANG Zhuolin, LIN Zhe, LARRY S, *et al.* Label consistent K-SVD: Learning a discriminative dictionary for recognition [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 35(11): 2651–2664. doi: 10.1109/TPAMI.2013.88.
- [9] 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. doi: 10.1007/s11263-014-0722-8.
- [10] 练秋生, 石保顺, 陈书贞. 字典学习模型、算法及其应用研究进展[J]. *自动化学报*, 2015, 41(2): 240–260. doi: 10.16383/j.aas.2015.c140252.
- LIAN Qiusheng, SHI Baoshun, and CHEN Shuzhen. Research advances on dictionary learning models, algorithms and applications[J]. *Acta Automatica Sinica*, 2015, 41(2): 240–260. doi: 10.16383/j.aas.2015.c140252.
- [11] RUBINSTEIN R, PELEG T, and ELAD M. Analysis K-SVD: A dictionary-learning algorithm for the analysis sparse model[J]. *IEEE Transactions on Signal Processing*, 2013, 61(3): 661–677. doi: 10.1109/TSP.2012.2226445.
- [12] CHEN Yunjin, RANFTL R, and POCK T. Insights into analysis operator learning: From patch-based sparse models to higher order MRFs[J]. *IEEE Transactions on Image Processing*, 2014, 23(3): 1060–1072. doi: 10.1109/TIP.2014.2299065.
- [13] GU Shuhang, ZHANG Lei, ZUO Wangmeng, *et al.* Projective dictionary pair learning for pattern classification[C]. Advances in Neural Information Processing System, Vancouver, BC, Canada, 2014, 1: 793–801.
- [14] RAKOTOMAMONJY A. Applying alternating direction method of multipliers for constrained dictionary learning[J]. *Neurocomputing*, 2013, 61(3): 126–136. doi: 10.1016/j.neucom.2012.10.024.
- [15] CAI Sijia, ZUO Wangmeng, ZHANG Lei, *et al.* Support vector guided dictionary learning[C]. European Conference on Computer Vision, Zurich, Switzerland, 2014, 8692: 624–639.
- [16] GORSKI J, PFEUFFER F, and KLAMROTH K. Biconvex sets and optimization with biconvex functions: a survey and extensions[J]. *Mathematical Methods of Operations Research*, 2007, 66(3): 373–407. doi: 10.1007/s00186-007-0161-1.
- 郭继昌: 男, 1966 年生, 教授, 研究方向为数字图像处理、语音信号处理等。
- 张帆: 女, 1993 年生, 硕士生, 研究方向为图像分类。
- 王楠: 男, 1990 年生, 硕士生, 研究方向为图像分类。