

基于 Profiles 的 Fisher 判别约束字典学习算法

李争名^{1,2,3} 杨南粤¹ 岑健⁴

(1. 广东技术师范学院工业实训中心, 广州, 510665; 2. 哈尔滨工业大学深圳研究生院生物计算研究中心, 深圳, 518055; 3. 福建省信息处理与智能控制重点实验室(闽江学院), 福州, 350121; 4. 广东技术师范学院自动化学院, 广州, 510665)

摘要: 为了增强编码系数的判别性能, 提出编码系数矩阵行向量(Profiles)的 Fisher 判别字典(Profiles of fisher discriminative dictionary learning, PFDDL)学习算法。首先, 根据 Profiles 能反映原子在字典学习中的使用情况, 提出一种自适应的原子类标构造方法。然后, 利用 Profiles 与原子间的一一对应关系, 设计 Profiles 的 Fisher 判别准则作为判别式项, 使得同类原子对应 Profiles 的类内散度尽可能小, 不同类原子对应 Profiles 的类间散度尽可能大, 促使字典中的同类原子尽量表示同类训练样本, 提高编码系数的判别性能。在 3 个人脸和 1 个手写字体数据库上的实验结果表明, 提出的算法比其他稀疏编码和字典学习算法能取得更高的分类性能。

关键词: 字典学习; 稀疏表示; Fisher 判别; 协作表示

中图分类号: TN911.73 **文献标志码:** A

Fisher Discriminative Constraint Dictionary Learning Algorithm Based on Profiles

Li Zhengming^{1,2,3}, Yang Nanyue¹, Cen Jian⁴

(1. Industrial Training Center, Guangdong Polytechnic Normal University, Guangzhou, 510665, China; 2. Bio-Computing Research Center, Shenzhen Graduate School, Harbin Institute of Technology, Shenzhen, 518055, China; 3. Fujian Provincial Key Laboratory of Information Processing and Intelligent Control (Minjiang University), Fuzhou, 350121, China; 4. School of Automation, Guangdong Polytechnic Normal University, Guangzhou, 510665, China)

Abstract: To improve the discriminative ability of the coding coefficients, the Profiles (the line vectors of coding coefficients matrix) of Fisher discriminative dictionary learning (PFDDL) is proposed. Firstly, the Profiles can indicate the corresponding atoms which are used by the training samples to encode in the dictionary learning, and an adaptive method is proposed to construct the labels of atoms. Since there are one-to-one correspondences between the Profiles and atoms, then the Fisher discriminative criterion is imposed on the Profiles so that they have small within-class compactness but large between-class separability. Thus, it can encourage the atoms of the same class to reconstruct the training sample of the same class, and enhance the discriminative ability of the coding coefficients, then improve the performance of dictionary learning. Experimental results show that the PFDDL algorithm can achieve better classification performance than other sparse coding and dictionary learning algorithms on the three face and one handwriting databases.

Key words: dictionary learning; sparse representation; Fisher discriminative; collaborative representation

引言

基于稀疏表示^[1,2]和基于协作表示^[3]分类算法在图像分类和人脸识别等领域均取得了较好的性能。最近,大量研究证实从训练样本中学习的字典比原始训练样本具有更好的表示性能^[4-6]。因此,作为信号稀疏表示理论中的研究方向之一,判别字典学习算法受到研究者的广泛关注。

理想字典学习模型中,同类训练样本对应编码系数矩阵的列向量(编码系数)相似,因此,研究者提出许多利用编码系数结合其他特征构造判别式项的字典学习算法^[7-9]。由于类标在模式分类中起到十分重要的作用,Yang等^[8]利用训练样本的类标与编码系数构造基于Fisher准则约束的判别式项,并提出利用编码系数的Fisher准则构造判别式的字典学习算法(Fisher discrimination dictionary learning, FDDL)。Cai等^[9]提出基于编码系数和训练样本类标的支持向量指导判别字典学习算法。上述算法能够增强编码系数的判别性能,但是忽略了编码系数矩阵行向量特征。利用Profile能够表示字典中原子重构训练样本的使用情况,Lu等^[10]设计了一种自适应字典原子更新和选择算法。Sadeghi等^[11]首次给出了Profile的定义,并提出一种字典原子更新策略。在字典学习模型中原子与Profiles一一对应,而且Profiles能够衡量原子重构训练样本的贡献。为此,Li等^[12]把Profiles引入到判别式模型的设计中,提出基于原子局部特征和类标嵌入约束的判别字典学习算法。然而,该算法在原子类标嵌入约束项中只考虑同类原子对应Profiles间的关系,忽略了不同类原子对应Profiles间的关系。由于Fisher判别准则能够很好地反映样本间的判别性能,因此,如何利用Profiles间的相似性特征设计判别式约束模型,对增强编码系数的判别性具有一定的研究意义。

综上所述,为了增强编码系数的判别性,构造Profiles的Fisher判别准则作为判别式项的判别字典学习算法。提出的算法不仅能够减少同类原子对应Profiles的类内散度,而且能够增大不同类原子对应的Profiles的类间散度。由于Profiles能够表示字典中原子重构训练样本的使用情况,而且Profiles矩阵是编码系数矩阵的转置矩阵,因此,Profiles的Fisher判别准则能够促使字典中的同类原子尽量表示同类训练样本,提高编码系数的判别性能。

1 Profiles的定义

根据文献[10],理想情况下字典学习模型 $\mathbf{Y}=\mathbf{D}\times\mathbf{X}$ 可以写成图1。其中 $\mathbf{Y}=[\mathbf{y}_1,\mathbf{y}_2,\dots,\mathbf{y}_n]\in\mathbf{R}^{m\times n}$ 是训练样本集合, m 和 n 分别是训练样本的维数和个数。 $\mathbf{D}=[\mathbf{d}_1,\mathbf{d}_2,\dots,\mathbf{d}_k]\in\mathbf{R}^{m\times k}$ 表示学习得到的字典, k 是字典中原子的个数, $\mathbf{X}=[\mathbf{x}_1,\mathbf{x}_2,\dots,\mathbf{x}_n]\in\mathbf{R}^{k\times n}$ 表示编码系数矩阵, $\mathbf{x}_i=[x_{i1},x_{i2},\dots,x_{ik}]^T\in\mathbf{R}^k(1\leq i\leq n)$ 是第 i 个编码系数,表示字典中的原子重构训练样本的情况。

$$\begin{aligned} \mathbf{y}_1 &= d_{11}x_{11} + \dots + d_{k1}x_{k1} \\ \mathbf{y}_2 &= d_{12}x_{12} + \dots + d_{k2}x_{k2} \\ &\vdots \\ \mathbf{y}_n &= d_{1n}x_{1n} + \dots + d_{kn}x_{kn} \end{aligned}$$

图1 Profiles的定义

Fig.1 Definition of Profiles

根据文献[11],图1中每个长方形框表示一个Profile,并且与原子是一一对应关系。假设原子 d_j 对应的Profile为 $\mathbf{p}_j=[x_{j1},x_{j2},\dots,x_{jn}]^T(1\leq j\leq k)$,根据文献[11,12], \mathbf{p}_j 可以表示原子被训练样本使用的情况。如果 \mathbf{p}_j 为零矢量,则表明在字典学习中原子 d_j 没有被训练样本集合使用,如果 \mathbf{p}_j 为非零矢量,且 \mathbf{p}_j 中的 x_{j1} 和 x_{jn} 为非零元素,其余元素都为零,则 \mathbf{p}_j 表明训练样本 \mathbf{y}_1 和 \mathbf{y}_n 在字典学习中利用 \mathbf{p}_j 对应的原子 d_j ,而其他训练样本没有利用原子 d_j 对其进行重构。因此,编码系数中的元素能够反映原子重构训练样本的使用情况,而且其与训练样本一一对应;Profiles中的元素表示原子被训练样本集合使用的情况,而且其与原子一一对应。从图1可得,Profiles矩阵为 $\mathbf{P}=[\mathbf{p}_1,\mathbf{p}_2,\dots,\mathbf{p}_k]\in\mathbf{R}^{n\times k}$,而且 $\mathbf{P}=\mathbf{X}^T$ 是编码系数矩阵的转置矩阵。

2 基于Profiles的Fisher判别约束字典学习算法

2.1 Profiles的Fisher判别模型

根据文献[13],理想字典学习模型中同类训练样本应该由某些原子单独重构,并且这些原子应该与

重构的训练样本具有相同的类标。由于 Profiles 与原子是一一对应关系并且能表示原子被训练样本使用的情况,因此,可以利用 Profiles 自适应地分配原子类标。

(1) 假设原子 \mathbf{d}_j 对应的 Profile 为 \mathbf{p}_j , 求出 \mathbf{p}_j 中的最大元素 x_{ji} 。

(2) 由于 \mathbf{p}_j 中的最大元素 x_{ji} 表示训练样本 \mathbf{y}_i 使用原子 \mathbf{d}_j 的情况。根据图 1, 把训练样本 \mathbf{y}_i 的类标设为原子 \mathbf{d}_j 的类标。如果训练样本 \mathbf{y}_i 属于训练样本集合中的第 i 类, 则原子 \mathbf{d}_j 的类标可以定义为 $\mathbf{b}_j = [0, 0, \dots, 0, 1, 0, \dots, 0] \in \mathbf{R}^c$, 其中 \mathbf{b}_j 中的第 i 个元素为 1, 其余均为零。因此, 字典 \mathbf{D} 的类标矩阵 \mathbf{B} 可以定义为 $\mathbf{B} = [\mathbf{b}^1, \mathbf{b}^2, \dots, \mathbf{b}^k]^\top \in \mathbf{R}^{k \times c}$, c 为原子的类数。

根据文献[12], 同类原子对应的 Profiles 也相似。因此, 为了增强编码系数矩阵的判别性能, 根据原子与 Profiles 间的一一对应关系, 定义 Profiles 的类内散布矩阵 \mathbf{S}_W 和类间散布矩阵 \mathbf{S}_B 为

$$\mathbf{S}_W = \sum_{i=1}^k \sum_{\mathbf{p}_j \in P_i} (\mathbf{p}_i - \mathbf{v}_j) (\mathbf{p}_i - \mathbf{v}_j)^\top \quad (1)$$

$$\mathbf{S}_B = \sum_{i=1}^c \mathbf{n}_i (\mathbf{v}_i - \mathbf{v}) (\mathbf{v}_i - \mathbf{v})^\top \quad (2)$$

式中: \mathbf{p}_j 为第 j 类原子对应的 Profiles, \mathbf{v}_j 表示第 j 类原子对应 Profiles 的均值矢量, \mathbf{n}_j 表示第 j 类原子对应 Profiles 的数目, \mathbf{v} 为所有原子对应 Profiles 的均值矢量。

为了优化式(1,2), 假设字典中所有类原子个数相同, 并设为 f , 则字典中原子个数 $k = cf$ 。假设 Profiles 矩阵 $\mathbf{P} = [\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_c]$, $\mathbf{p}_j = [\mathbf{p}_{(j-1)f+1}, \mathbf{p}_{(j-1)f+2}, \dots, \mathbf{p}_{jf}]$ 是第 j 类原子对应的 Profiles。因此, 式(1)可以转换为

$$\mathbf{S}_W = \sum_{i=1}^c \sum_{j=1, \mathbf{p}_i \in P_j} (\mathbf{p}_j - \mathbf{v}_i) (\mathbf{p}_j - \mathbf{v}_i)^\top = (\mathbf{p}_1 - \mathbf{v}_1) (\mathbf{p}_1 - \mathbf{v}_1)^\top + \dots + (\mathbf{p}_f - \mathbf{v}_1) (\mathbf{p}_f - \mathbf{v}_1)^\top + \dots + (\mathbf{p}_{cf} - \mathbf{v}_c) (\mathbf{p}_{cf} - \mathbf{v}_c)^\top \quad (3)$$

式(3)中的第 j 类原子对应的 Profiles 散布矩阵为

$$\begin{aligned} & (\mathbf{p}_{(j-1)f+1} - \mathbf{v}_j) (\mathbf{p}_{(j-1)f+1} - \mathbf{v}_j)^\top + \dots + (\mathbf{p}_{jf} - \mathbf{v}_j) (\mathbf{p}_{jf} - \mathbf{v}_j)^\top = \\ & \mathbf{p}_{(j-1)f+1} (\mathbf{p}_{(j-1)f+1})^\top + \dots + \mathbf{p}_{jf} (\mathbf{p}_{jf})^\top + f \mathbf{v}_j \mathbf{v}_j^\top - \\ & 2 (\mathbf{p}_{(j-1)f+1} \mathbf{v}_j^\top + \dots + \mathbf{p}_{jf} \mathbf{v}_j^\top) \end{aligned} \quad (4)$$

因此, 把式(4)代入式(3), 则式(3)转变为

$$\mathbf{S}_W = \sum_{j=1}^k \mathbf{p}_j \mathbf{p}_j^\top + f \sum_{j=1}^c \mathbf{v}_j \mathbf{v}_j^\top - 2 \sum_{j=1}^c (\mathbf{p}_{(j-1)f+1} + \dots + \mathbf{p}_{jf}) \mathbf{v}_j^\top \quad (5)$$

由于 $\mathbf{v}_j = \frac{\mathbf{p}_{(j-1)f+1} + \dots + \mathbf{p}_{jf}}{f}$, 则式(5)转变为

$$\mathbf{S}_W = \sum_{j=1}^k \mathbf{p}_j \mathbf{p}_j^\top - \frac{1}{f} \sum_{j=1}^c (\mathbf{p}_{(j-1)f+1} + \dots + \mathbf{p}_{jf}) (\mathbf{p}_{(j-1)f+1} + \dots + \mathbf{p}_{jf})^\top \quad (6)$$

同时, 式(2)可转变为

$$\mathbf{S}_B = f \sum_{j=1}^c (\mathbf{v}_j - \mathbf{v}) (\mathbf{v}_j - \mathbf{v})^\top = f \sum_{j=1}^c \mathbf{v}_j \mathbf{v}_j^\top + f c \mathbf{v} \mathbf{v}^\top - 2 f \sum_{j=1}^c \mathbf{v}_j \mathbf{v}^\top \quad (7)$$

由于 $\mathbf{v} = \left(\frac{\mathbf{p}_1 + \dots + \mathbf{p}_f + \dots + \mathbf{p}_k}{k} \right) = \frac{1}{c} \sum_{j=1}^c \mathbf{v}_j$, 式(7)可转变为

$$\begin{aligned} \mathbf{S}_B &= f \sum_{j=1}^c \mathbf{v}_j (\mathbf{v}_j)^\top + k \mathbf{v} (\mathbf{v})^\top - 2 f \sum_{j=1}^c \mathbf{v}_j \mathbf{v}^\top = \\ & \frac{1}{f} \sum_{j=1}^c (\mathbf{p}_{(j-1)f+1} + \dots + \mathbf{p}_{jf}) (\mathbf{p}_{(j-1)f+1} + \dots + \mathbf{p}_{jf})^\top - \\ & \frac{1}{k} (\mathbf{p}_1 + \dots + \mathbf{p}_f + \dots + \mathbf{p}_k) (\mathbf{p}_1 + \dots + \mathbf{p}_f + \dots + \mathbf{p}_k)^\top \end{aligned} \quad (8)$$

为了提高编码系数的判别性能, Profiles 的 Fisher 判别准则定义为

$$\min(\text{Tr}(\mathbf{S}_W) - \text{Tr}(\mathbf{S}_B)) \quad (9)$$

式中 $\text{Tr}(\cdot)$ 表示对矩阵求迹。根据式(6,8),式(9)可以转换为

$$\begin{aligned} \min(\text{Tr}(\mathbf{S}_W) - \text{Tr}(\mathbf{S}_B)) &= \min\left(\text{Tr}\left(\sum_{j=1}^k \mathbf{p}_j \mathbf{p}_j^T\right)\right) + \\ \min\left(\frac{1}{k} \text{Tr}\left((\mathbf{p}_1 + \cdots + \mathbf{p}_f + \cdots + \mathbf{p}_k)(\mathbf{p}_1 + \cdots + \mathbf{p}_f + \cdots + \mathbf{p}_k)^T\right)\right) - \\ \min\left(\frac{2}{f} \text{Tr}\left(\sum_{j=1}^c (\mathbf{p}_{(j-1)f+1} + \cdots + \mathbf{p}_{jf})(\mathbf{p}_{(j-1)f+1} + \cdots + \mathbf{p}_{jf})^T\right)\right) \end{aligned} \quad (10)$$

为了更好地优化式(10),定义 k 阶矩阵 $\mathbf{A} = \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ 1 & 1 & \cdots & 1 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \in \mathbf{R}^{k \times k}$, \mathbf{A} 中的每个元素均为 1。定

义 k 阶矩阵 $\mathbf{G} = \frac{1}{f} \mathbf{U} \mathbf{U}^T$, $\mathbf{U} = \mathbf{B}(\mathbf{B}^T \mathbf{B})^{-\frac{1}{2}}$ 。则式(10)可以转换为

$$\begin{aligned} \min(\text{Tr}(\mathbf{S}_W) - \text{Tr}(\mathbf{S}_B)) &= \\ \min(\text{Tr}(\mathbf{X}^T \mathbf{X})) + \min\left(\frac{1}{k} \text{Tr}(\mathbf{X}^T \mathbf{A} \mathbf{X})\right) - \min\left(\frac{2}{f} \text{Tr}(\mathbf{X}^T \mathbf{G} \mathbf{X})\right) &= \min(\text{Tr}(\mathbf{X}^T \mathbf{L} \mathbf{X})) \end{aligned} \quad (11)$$

式中 $\mathbf{L} = \mathbf{I} + \mathbf{A} - 2\mathbf{G}$, \mathbf{I} 是单位矩阵。

2.2 PFDDL 算法的目标函数

为了鼓励同类原子对应的 Profiles 比不同类原子对应的 Profiles 有更强的相似性,促使字典中的同类原子尽量表示同类训练样本,提高编码系数的判别性能,本文提出的 PFDDL 算法的目标函数为

$$\min_{D, X, L} (\|\mathbf{Y} - \mathbf{D} \mathbf{X}\|_2^2 + \alpha \text{Tr}(\mathbf{X}^T \mathbf{L} \mathbf{X}) + \beta \|\mathbf{X}\|_2) \quad \text{s. t. } \|\mathbf{d}_j\|_2 \leq 1, j = 1, 2, \dots, k \quad (12)$$

式中: $\|\mathbf{Y} - \mathbf{D} \mathbf{X}\|_2^2$ 表示字典的重构性能, $\|\mathbf{X}\|_2$ 是编码系数的正则项, α 和 β 是参数, $\text{Tr}(\mathbf{X}^T \mathbf{L} \mathbf{X})$ 表示 Profiles 的 Fisher 判别准则。

PFDDL 算法与现有判别字典学习算法的区别主要体现在以下 3 个方面:(1)PFDDL 算法把对编码系数的约束转换为对 Profiles 的约束,现有判别字典学习算法利用训练样本与编码系数一一对应的关系构造判别式约束项,而 PFDDL 算法利用原子与 Profiles 的一一对应关系构造判别式约束项;(2)PFDDL 算法利用 Profiles 构造自适应的原子类标,并在此基础上利用同类原子和不同类原子对应的 Profiles 间的关系构造判别式约束项。由于 Profiles 在字典学习过程中能够自适应更新,PFDDL 算法构造的判别式项具有一定的自适应性。现有的判别字典学习算法利用训练样本特征结合编码系数设计判别式约束项,在字典学习过程中,训练样本的结构特征不能自适应更新;(3)目前判别字典学习算法大多利用 l_1 范数或 l_0 范数约束编码系数,算法的计算复杂度较高,而 PFDDL 算法利用 l_2 范数对编码系数约束,其与判别式项结合可以直接求导,减少算法的计算复杂度。

PFDDL 算法和 FDDL 算法都利用 Fisher 判别准则构造判别式约束项提高编码系数的判别性。FDDL 算法利用编码系数与训练样本间的关系设计编码系数的 Fisher 判别准则提高编码系数间的相似性和差异性。利用原子与 Profiles 间的关系,PFDDL 算法首先为每个原子自适应地分配类标,并在此基础上构造 Profiles 的 Fisher 判别准则,促使同类原子对应的 Profiles 尽可能地相似,不同类原子对应的 Profiles 间的差异尽可能地大。由于 Profiles 能反映训练样本在字典学习过程中原子的使用情况,并且原子的类标是根据 Profiles 中的元素对应的训练样本的类标一致,因此,Profiles 的 Fisher 判别准则能促使同类的原子尽可能地重构某一类训练样本,增强编码系数矩阵的判别性能。此外,由于 Profiles 矩

阵是编码系数矩阵的转置矩阵,PFDDL 算法把对编码系数的约束转换为对 Profiles 的约束,并结合编码系数矩阵的 l_2 范数约束,使得目标函数可以交替直接求导,减少算法的计算复杂度。

3 PFDDL 算法的求解及分类方法

PFDDL 算法的目标函数可以采取交替求导的方法进行求解。首先,利用基于 K 均值的奇异值分解(K-means-based singular value decomposition, K-SVD)^[14] 算法初始化字典 \mathbf{D} 和编码系数矩阵 \mathbf{X} ; 然后,利用编码系数矩阵的转置矩阵,也即是 Profiles 矩阵,求出原子的类标矩阵 \mathbf{B} ,在利用原子类标矩阵 \mathbf{B} 计算矩阵 \mathbf{G} 并设置矩阵 \mathbf{A} ;更新目标函数直到求解出最优解。

3.1 字典 \mathbf{D} 的求解

假设 PFDDL 算法的目标函数中的编码系数矩阵 \mathbf{X} 和矩阵 \mathbf{L} 为常量,则式(12)可以转换为

$$\min_{\mathbf{D}} (\|\mathbf{Y} - \mathbf{D}\mathbf{X}\|_2^2) \quad \text{s. t. } \|\mathbf{d}_j\|_2^2 \leq 1, j=1, 2, \dots, k \quad (13)$$

根据文献[15],其可以利用拉格朗日函数进行求解,于是式(13)转换为

$$g(\lambda) = \inf (\|\mathbf{Y} - \mathbf{D}\mathbf{X}\|_2^2 + \sum_{j=1}^k \lambda_j (\|\mathbf{d}_j\|_2^2 - 1)) \quad (14)$$

式中, $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, \dots, \lambda_j, \dots, \lambda_k]$, $\lambda_j (j=1, 2, \dots, k)$ 是拉格朗日乘子。为了求解式(14),定义对角矩阵 $\mathbf{A} \in \mathbf{R}^{k \times k}$, 其中 $A_{jj} = \lambda_j$ 表示对角元素。根据文献[15],式(14)可转变为

$$F(\mathbf{D}, \boldsymbol{\lambda}) = \|\mathbf{Y} - \mathbf{D}\mathbf{X}\|_2^2 + \text{Tr}(\mathbf{D}^T \mathbf{D} \mathbf{A}) - \text{Tr}(\mathbf{A}) \quad (15)$$

为了获得最优的字典 \mathbf{D} ,对式(15)求一阶导数,并令其等于零可得

$$\mathbf{D}^* = \mathbf{Y}\mathbf{X}^T (\mathbf{X}\mathbf{X}^T + \mathbf{A})^{-1} \quad (16)$$

为了减少计算复杂度,利用 $\gamma \mathbf{I}$ 代替 \mathbf{A} 获得最优的字典 \mathbf{D}^* 为

$$\mathbf{D}^* = \mathbf{Y}\mathbf{X}^T (\mathbf{X}\mathbf{X}^T + \gamma \mathbf{I})^{-1} \quad (17)$$

式中: γ 是参数, \mathbf{I} 是单位矩阵。最优 \mathbf{A} 的求解方法见文献[15]。

3.2 编码系数矩阵 \mathbf{X} 和矩阵 \mathbf{L} 的求解

假设 PFDDL 算法的目标函数中字典 \mathbf{D} 和矩阵 \mathbf{L} 为常量,则式(12)可以转换为

$$\min_{\mathbf{X}} (\|\mathbf{Y} - \mathbf{D}\mathbf{X}\|_2^2 + \alpha \text{Tr}(\mathbf{X}^T \mathbf{L} \mathbf{X}) + \beta \|\mathbf{X}\|_2^2) \quad (18)$$

式(18)可以直接求导得到编码系数矩阵 \mathbf{X} 。因此,对式(18)求一阶导数并令其为零,可得

$$-\mathbf{D}^T \mathbf{Y} + \mathbf{D}^T \mathbf{D} \mathbf{X} + \alpha \mathbf{L} \mathbf{X} + \beta \mathbf{X} = 0 \quad (19)$$

根据式(19),可得编码系数矩阵 \mathbf{X}^* 的解

$$\mathbf{X}^* = (\mathbf{D}^T \mathbf{D} + \alpha \mathbf{L} + \beta \mathbf{I})^{-1} \mathbf{D}^T \mathbf{Y} \quad (20)$$

一旦获得编码系数矩阵 \mathbf{X} ,利用 Profiles 更新原子的类标矩阵 \mathbf{B} ,并利用 $\mathbf{U} = \mathbf{B}(\mathbf{B}^T \mathbf{B})^{-\frac{1}{2}}$ 计算 $\mathbf{G} = \frac{1}{f} \mathbf{U} \mathbf{U}^T$ 。然后根据 $\mathbf{L} = \mathbf{I} + \mathbf{A} - 2\mathbf{G}$,更新矩阵 \mathbf{L} 。

3.3 分类方法

判别字典学习算法常利用训练样本类标和编码系数设计分类器,然后结合测试样本的稀疏表示系数进行分类,典型算法见文献[12]。由于 PFDDL 算法中原子具有类标信息。因此,利用 FDDL 算法中的全局分类方法作为 PFDDL 算法的分类方法。根据文献[8],对于测试样本 \mathbf{y} 和字典 \mathbf{D} ,利用式(21)获得测试样本的稀疏表示系数为

$$\hat{\mathbf{a}} = \min_{\mathbf{a}} (\|\mathbf{y} - \mathbf{D}\mathbf{a}\|_2^2 + \eta \|\mathbf{a}\|_1) \quad (21)$$

式中 η 是参数。假设 $\hat{\mathbf{a}} = [\hat{a}_1, \hat{a}_2, \dots, \hat{a}_c]$, \hat{a}_i 是第 i 类原子对应的稀疏表示系数。如果测试样本 \mathbf{y} 属于

第 i 类样本, 则 $\|\mathbf{y} - \mathbf{D}_i \hat{\mathbf{a}}_i\|_2$ 应该小, 而 $\|\mathbf{y} - \mathbf{D}_j \hat{\mathbf{a}}_j\|_2 (j \neq i)$ 应该较大。此外, 稀疏表示系数 $\hat{\mathbf{a}}$ 与第 i 类编码系数间的距离小, 而与其他类编码系数间的距离大。因此, 可以利用式(22)对测试样本进行分类。

$$e_i = \|\mathbf{y} - \mathbf{D}_i \hat{\mathbf{a}}_i\|_2 + \omega \|\hat{\mathbf{a}} - \mathbf{s}_i\|_2 \quad (22)$$

式中 ω 是参数, \mathbf{s}_i 是第 i 类训练样本对应的编码系数的均值矢量。 e_i 是第 i 类字典对测试样本的重构误差, 测试样本 \mathbf{y} 的类标分配到产生最小误差 e_i 的字典 \mathbf{D}_i 对应的类。

PFDDL 算法

输入: 训练样本集合 \mathbf{Y} 和测试样本 \mathbf{y} , 参数 $\alpha, \beta, \gamma, \eta, \omega$ 和迭代次数

输出: 测试样本 \mathbf{y} 的类标

(1) 利用 K-SVD 算法初始化字典 \mathbf{D} 和编码系数矩阵 \mathbf{X} 。

(2) 利用 Profiles 分配字典 \mathbf{D} 中原子的类标, 获得字典 \mathbf{D} 的类标矩阵 \mathbf{B} , 并计算矩阵 $\mathbf{U} = \mathbf{B}(\mathbf{B}^T \mathbf{B})^{-\frac{1}{2}}$ 。

(3) 利用矩阵 \mathbf{U} , 根据 $\mathbf{G} = \frac{1}{f} \mathbf{U} \mathbf{U}^T$ 计算矩阵 \mathbf{G} 。

(4) 利用矩阵 \mathbf{A} 和 \mathbf{G} , 根据 $\mathbf{L} = \mathbf{I} + \mathbf{A} - 2\mathbf{G}$ 计算矩阵 \mathbf{L} 。

(5) For $i = 1 : \xi$

 利用式(17)计算字典 \mathbf{D}^i ;

 利用式(20)计算编码系数矩阵 \mathbf{X}^i ;

 利用步骤(2—4)更新矩阵 \mathbf{L}^i ;

End

(6) 获得最优的字典 $\mathbf{D} = \mathbf{D}^\xi$ 和编码系数矩阵 $\mathbf{X} = \mathbf{X}^\xi$ 。

(7) 针对测试样本 \mathbf{y} , 利用式(21)获得稀疏表示系数, 并提取每类原子对应的稀疏表示系数。

(8) 利用式(22)计算重构误差, 测试样本的类标分配到产生最小误差的字典对应的类。

4 实验及结果分析

为了验证 PFDDL 算法的性能, 本节给出其与基于稀疏表示的分类(Sparse representation based classification, SRC)^[1], 基于协同表示的分类(Collaborative representation based classification, CRC)^[3], FDDL^[8], 支持向量指导字典学习(Support vector guided dictionary learning, SVGDL)^[9]和局部约束和类标嵌入字典学习(Locality constrained and label embedding dictionary learning, LCLE-DL)^[12]算法在 AR^[16], LFW^[17]和 FERET^[18]人脸数据库和 USPS^[14]手写字数据库中的实验结果。对于 SRC 算法, 利用 l_1 方法获得测试样本的表示系数。

为了选择合适的参数, 随机选择 LFW 数据库中每人的 5 幅图像为训练集合, 剩下的图像作为测试集合。在 PFDDL 算法中, 参数的范围设定为 $10^{-6}, 10^{-5}, \dots, 10^5$ 。对每个参数进行遍历, 选择分类性能较好的参数作为 PFDDL 算法的参数。当 $\alpha = 10^{-5}$, $\beta = 10^{-2}$ 和 $\gamma = 10^{-2}$ 时, PFDDL 算法在 LFW 上述数据集中取得较好的分类性能。为了增强算法的鲁棒性, 在本文所有实验中, PFDDL 算法的参数均设定为 $\alpha = 10^{-5}$, $\beta = 10^{-2}$ 和 $\gamma = 10^{-2}$ 。另外, 最大迭代次数 $\xi = 20$, ω 和 η 与 FDDL 算法中参数相同, 即 $\omega = 0.5$ 和 $\eta = 0.005$ 。

4.1 AR 数据库中的实验结果

AR 数据库包括超过 4 000 幅 126 人的彩色人脸图, 每幅图像具有不同的表情、亮度和遮挡。根据文献[19], 选择 120 个人的采集于两个不同阶段的图像作为实验所用集合。此外, 每幅图像调整为 40

像素 \times 50 像素大小的灰度图像,部分图像如图 2 所示。在本实验中,选取第 1 阶段的图像为训练集合,第 2 阶段图像为测试样本。原子大小等于训练样本个数,也即 1 560。实验结果如表 1 所示。



图 2 AR 人脸数据库中的部分图像

Fig. 2 Some samples from AR face database

表 1 AR 数据库中的识别率,训练字典的时间和分类 1 个样本的平均时间

Tab. 1 Average recognition rates, computing time for training a dictionary and average time of classifying test sample on AR database

算法	CRC	SRC	FDDL	SVGDL	LCLE-DL	PFDDL
识别率/%	71.3	70.6	74.7	73.6	73.5	77.6
训练时间/s			$3e+4$	$7e+3$	100	70
测试时间/s	$1.3e-3$	3.1	7.9	$1.1e-4$	$4e-3$	7.8

4.2 LFW 人脸数据库中的实验结果

LFW 数据库包括 13 000 幅来自于互联网的 1 680 个人的 人脸图像。根据文献[12,19],从 LFW 数据的裁剪版本(LFW crop)中选择每类图像个数为 11~20 的人脸图像作为本实验所用集合。该集合包括 86 人的 1 215 幅图像,并把每幅图像的大小调整为 32 像素 \times 32 像素。LFW Crop 中的部分图像如图 3 所示。在本实验中,训练集合由随机选择的每类 5 幅图像组成,剩下的图像组成测试样本。原子大小等于训练样本个数,也即 430。此外,随机运行 10 次实验,将其平均值作为实验结果。表 2 给出了提出的算法和 5 个对比算法的实验结果。



图 3 LFW Crop 数据库中的部分图像

Fig. 3 Some samples from LFW Crop database

表 2 LFW Crop 数据库中的识别率,训练字典的时间和分类 1 个样本的平均时间

Tab. 2 Average recognition rates, computing time for training a dictionary and average time of classifying test sample on LFW Crop database

算法	CRC	SRC	FDDL	SVGDL	LCLE-DL	PFDDL
识别率/%	24.6	23.5	24.2	22.8	25.2	29.1
训练时间/s			642	176	4	4
测试时间/s	$3.4e-3$	0.5	0.9	$8.5e-5$	$8.9e-4$	0.9

4.3 FERET 人脸数据库中的实验结果

在本实验中,根据文献[20],选择 200 个人的 1 400 幅图片作为数据集合,并把每幅图像的大小调

整为 40 像素×40 像素。图 4 显示了 FERET 中的部分图像。

在本实验中,训练样本集合由每人的 3 幅图像组成,测试样本集合由剩余的图像组成。在 FERET 数据库中每类有 7 幅图像,则训练样本和测试样本可划分为 $C_7^3 = 35$ 种不同的组合对。原子大小等于训练样本个数,也即 600。表 3 显示了 PFDDL 算法和 5 个对比算法的平均实验结果。

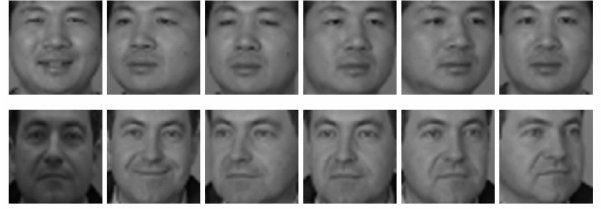


图 4 FERET 人脸数据库中的部分图像

Fig. 4 Some samples from FERET face database

表 3 FERET 数据库中的识别率,训练字典的时间和分类 1 个样本的平均时间

Tab. 3 Average recognition rates, computing time for training a dictionary and average time of classifying test sample on FERET database

算法	CRC	SRC	FDDL	SVGDL	LCLE-DL	PFDDL
识别率/%	53.6	56	62.8	58.4	57.6	64.1
训练时间/s			$6.6e+3$	639	12	11
测试时间/s	$7.5e-3$	8	2.2	$2.2e-4$	$1.3e-3$	2.2

4.4 USPS 数据库中的实验结果

USPS 手写字体数据库共计 10 类 9 298 幅手写图像。在 USPS 手写字体数据库中,每类字体的数目从 708~1 553 不等。根据文献 [15],每幅图像大小调整为 16 像素×16 像素。USPS 手写字体数据库中的部分图像如图 5 所示。在本实验中,选择每类手写字体的前 50 个样本图像为训练样本,剩余的手写字体图像为测试样本。表 4 显示了提出的算法和 5 个对比算法的实验结果。



图 5 USPS 数据库中的部分图像

Fig. 5 Some samples from USPS database

表 4 USPS 数据库中的识别率,训练字典的时间和分类 1 个样本的平均时间

Tab. 4 Average recognition rates, computing time for training a dictionary and average time of classifying test sample on USPS database

算法	CRC	SRC	FDDL	SVGDL	LCLE-DL	PFDDL
识别率/%	89.5	90.1	90.2	88.6	90.0	91.8
训练时间/s			86	70	12	9
测试时间/s	$9.2e-4$	0.8	$2.5e-1$	$1.7e-5$	$1.5e-3$	$2.5e-1$

4.5 实验结果分析

根据上述实验结果,可以得出以下结论:

(1)从表 2,3,4 中可以看出,PFDDL 算法的识别率比 FDDL 算法高 2.9%,4.1%,1.3%和 1.6%。虽然 FDDL 算法和 PFDDL 算法都是利用 Fisher 准则构造判别式约束项,PFDDL 算法首先利用 Profiles 自适应地分配原子类标,然后构造基于 Profiles 的判别式约束项。PFDDL 算法不仅能够促使同类原子对应 Profiles 的类内散度矩阵最小化和不同类原子对应 Profiles 的类间散度矩阵最大化,增强编码系数矩阵的判别性。而且,PFDDL 算法构造的判别式约束项还能表示训练集合利用原子的信息,促使字典

中的同类原子尽量表示同类训练样本,增强字典的判别性能。此外,PFDDL 算法把对编码系数的约束转换为对 Profiles 的约束,并自适应地分配原子类标。但是 FDDL 算法利用训练样本类标和编码系数构造判别式约束项,该约束项不能自适应性更新。因此,PFDDL 算法比 FDDL 算法能够更好地反映原子被训练样本使用的情况,提高了 PFDDL 算法的分类性能。

(2)从表 1—4 中可以看出,PFDDL 算法的识别率比 LCLE-DL 和 SVGDL 算法均高。主要原因是 LCLE-DL 算法仅仅考虑了同类原子对应 Profiles 间的关系,而忽略了不同类原子对应 Profiles 间的关系。SVGDL 算法也仅仅考虑了同类训练样本对应编码系数的权重约束。它们都在一定程度上降低了编码系数矩阵的判别性。PFDDL 算法利用 Fisher 准则构造 Profiles 的判别约束项能够克服 LCLE-DL 和 SVGDL 算法存在的问题。

(3)当原子个数等于训练样本个数时,表 1—4 表明 PFDDL 算法取得比 CRC 和 SRC 算法更高的识别率。主要原因是利用学习得到的字典比直接利用原始的训练样本具有更好的分类性能。

(4)从表 1—4 中可以看出,PFDDL 算法训练字典的时间比 FDDL 和 SVGDL 算法都小,与 LCLE-DL 算法几乎相等。主要原因是 PFDDL 和 LCLE-DL 算法都采用 l_2 范数约束,目标函数可以直接求导,减少算法的计算复杂度。由于使用相同的分类方法,PFDDL 算法与 FDDL 算法的分类时间几乎相等,比 SVGDL 和 LCLE-DL 算法都高。

5 结束语

为了增强字典和编码系数的判别性,构造一种新的基于 Profiles 的 Fisher 判别约束项,把基于编码系数的约束转换为对基于 Profiles 的约束,促使同类原子对应 Profiles 的类间散度最大化和不同类原子对应 Profiles 的类内散度最小化。由于同时考虑了原子对应 Profiles 的类内散度和类间散度,PFDDL 算法能够较好地反映训练样本使用原子的情况,促使同类的原子尽可能地重构某一类训练样本,增强字典的判别性能。另外,利用 l_2 范数对编码系数进行约束,使得 PFDDL 算法具有闭性解,能有效地减少算法的复杂度。在 AR, LFW, FERET 和 USPS 这 4 个数据库中,实验结果表明 PFDDL 算法的识别率不仅比直接利用原始训练的 SRC 和 CRC 算法高,而且比 FDDL, SVGDL 和 LCLE-DL 算法高。

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] 胡昭华, 赵孝嘉, 徐玉伟, 等. 一种新的低秩分块稀疏表示的人脸识别算法[J]. 数据采集与处理, 2015, 30(5): 1109-1120.
Hu Zhaohua, Zhao Xiaolei, Xu Yuwei, et al. Face recognition algorithm based on novel low-rank and block-based sparse representation[J]. Journal of Data Acquisition and Processing, 2015, 30(5): 1109-1120.
- [3] Zhang Lei, Yang Meng, Feng Xiangchu. Sparse representation or collaborative representation: Which helps face recognition? [C]//IEEE International Conference on Computer Vision. Spain: IEEE, 2011: 471-478.
- [4] Fawzi A, Davies M, Frossard P. Dictionary learning for fast classification based on soft-thresholding[J]. International Journal of Computer Vision, 2015, 114(2): 306-321.
- [5] 郭继昌, 金仰亨嘉. 一种基于字典学习的压缩感知视频编解码模型[J]. 数据采集与处理, 2015, 30(1): 59-67.
Guo Jichang, Jin Maohengjia. Dictionary learning-based compressive video sensing codec model[J]. Journal of Data Acquisition and Processing, 2015, 30(1): 59-67.
- [6] Shrivastava A, Patel V M, Chellappa R. Non-linear dictionary learning with partially labeled data[J]. Pattern Recognition, 2015, 48(11): 3283-3292.
- [7] Zhang Zhao, Li Fanzhang, Tommy W S C, et al. Sparse codes auto-extractor for classification: A joint embedding and dictionary learning framework for representation[J]. IEEE Transactions on Signal Processing, 2016, 64(14): 3790-3805.
- [8] Yang Meng, Zhang Lei, Feng Xiangchu, et al. Sparse representation based fisher discrimination dictionary learning for image

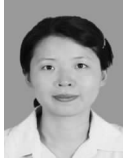
classification[J]. *International Journal of Computer Vision*, 2014, 109(3): 209-232.

- [9] Cai Sijia, Zuo Wangmeng, Zhang Lei, et al. Support vector guided dictionary learning[C]//European Conference on Computer Vision. Switzerland: Springer, 2014: 624-639.
- [10] Lu Cewu, Shi Jianping, Jia Jiaya. Scale adaptive dictionary learning[J]. *IEEE Transactions on Image Processing*, 2014, 23(2): 837-847.
- [11] Sadeghi M, Babaie-Zadeh M, Jutten C. Learning overcomplete dictionaries based on atom-by-atom updating[J]. *IEEE Transactions on Signal Processing*, 2014, 62(4): 883-891.
- [12] Li Zhengming, Lai Zhihui, Xu Yong, et al. A locality-constrained and label embedding dictionary learning algorithm for image classification[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2017, 28(2): 278-293.
- [13] Zhang Yangmuzi, Jiang Zhuolin, Davis L S. Learning structured low-rank representations for image classification[C]//IEEE Conference on Computer Vision and Pattern Recognition. Portland: IEEE, 2013: 676-683.
- [14] Aharon M, Bruckstein M E. K-SVD: An algorithm for designing of over-complete dictionaries for sparse representation[J]. *IEEE Transactions on Signal Processing*, 2006, 54(11): 4311-4322.
- [15] Zheng Miao, Bu Jiajun, Chen Chun, et al. Graph regularized sparse coding for image representation[J]. *IEEE Transactions on Image Processing*, 2011, 20(5): 1327-1336.
- [16] Martinez A M, Benavente R. The AR face database[R]. CVC Technical Report #24, Columbus, Ohio, USA: The Ohio State University 1998.
- [17] Huang G B, Ramesh M, Berg T, et al. Labeled faces in the wild: A database for studying face recognition in unconstrained environments[R]. University of Massachusetts, Amherst, Technical Report 07-49, 2007.
- [18] Phillips P J, Hyeonjoon M, Rizvi S A, et al. The FERET evaluation methodology for face recognition algorithms[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 22(10): 1090-1104.
- [19] Xu Yong, Li Zhengming, Zhang Bo, et al. Sample diversity, representation effectiveness and robust dictionary learning for face recognition[J]. *Information Sciences*, 2017, 375: 171-182.
- [20] Xu Yong, Fang Xiaozhao, You Jane, et al. Noise-free representation based classification and face recognition experiments [J]. *Neurocomputing*, 2015, 147(1): 307-314.

作者简介:



李争名(1982-),男,博士,高级实验师,研究方向:模式识别, E-mail: lizhengming2004@126.com。



杨南粤(1977-),女,硕士,讲师,研究方向:图像处理与虚拟现实。



岑健(1967-),女,博士,教授,研究方向:故障诊断、建筑电气智能化与节能。

(编辑:陈 珺)