

人脸识别中基于系数相似性的字典学习算法

施静兰¹ 常侃^{1,2} 张智勇¹ 覃团发^{1,2}

(广西大学计算机与电子信息学院 南宁 530004)¹

(广西高校多媒体通信与信息处理重点实验室(广西大学) 南宁 530004)²

摘要 在现有的基于稀疏表示分类算法的人脸识别中,使用通过稀疏学习得到的精简字典可以提高识别速度和精确度。metaface 学习(Metaface Learning, MFL)算法在字典学习过程中没有考虑同类样本稀疏编码系数之间具有相似性的特点。为了利用这一信息来提高字典的区分性,提出了一种基于系数相似性的 metaface 学习(Coefficient-Similarity-based Metaface Learning, CS-MFL)算法。CS-MFL 算法的学习过程中,在更新稀疏表示系数阶段加入同类训练样本稀疏编码系数相似的约束项。为了求解包含系数相似性约束的新的最优化问题,将目标函数中的两个 l_2 范数约束项进行合并,将原问题转化为典型 l_2-l_1 问题进行求解。在不同的人脸库上进行实验,结果表明,提出的 CS-MFL 算法能够获得比 MFL 算法更高的识别率,说明由 CS-MFL 算法学习得到的字典更高效且更具区分性。

关键词 稀疏表示,人脸识别,字典学习,稀疏编码

中图分类号 TP391.4 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2016.6.059

Coefficient-similarity-based Dictionary Learning Algorithm for Face Recognition

SHI Jing-lan¹ CHANG Kan^{1,2} ZHANG Zhi-yong¹ QIN Tuan-fa^{1,2}

(School of Computer and Electronic Information, Guangxi University, Nanning 530004, China)¹

(Guangxi Colleges and Universities Key Laboratory of Multimedia Communications and Information Processing, Guangxi University, Nanning 530004, China)²

Abstract Using a compact dictionary obtained by sparse learning could greatly improve the accuracy and speed up the procedure of classification for sparse representation-based face recognition method. However, the traditional metaface learning (MFL) method doesn't take into account the similarity among the training samples from the same person. In order to take the advantage of this prior information and make the learned dictionary more discriminative, an algorithm called coefficient-similarity-based metaface learning (CS-MFL) was proposed. In CS-MFL, the coefficient similarity is incorporated as a new constraint to the original objective function. To solve the new optimization problem, both l_2 norm-based constraints are combined, and the original problem becomes a typical l_2-l_1 problem. An experiment was carried out on different face databases, which shows that the proposed CS-MFL algorithm can achieve higher recognition rate than MFL algorithm, which demonstrates that the dictionary obtained by CS-MFL algorithm is more efficient and discriminative than that of the traditional MFL for face recognition application.

Keywords Sparse representation, Face recognition, Dictionary learning, Sparse coding

1 引言

当今信息化社会的快速发展使得人们对信息安全性的需求不断提高。人脸识别技术因直接友好、易操作以及安全性高等优点,已经成为计算机视觉应用领域中的研究热点,具有广阔的应用前景。传统的人脸识别方法主要有:基于几何特征的人脸识别方法、基于主成分分析(Principal Component Analysis, PCA)的人脸识别方法、基于神经网络的人脸识别方法以及基于支持向量机(Support Vector Machine, SVM)的

人脸识别方法等。近年来,伴随着压缩感知理论研究的兴起^[1,2],有关稀疏表示的研究引起了学术界的广泛关注。2009年, J. Wright 与 A. Yang 等人^[3]首次将稀疏表示理论应用于人脸识别,提出了基于稀疏表示分类(Sparse Representation-based Classification, SRC)的人脸识别方法。该方法使用原始训练样本作为字典,并利用测试样本在字典上的稀疏表示系数计算每个类别的表示误差,最终将测试样本判别为具有最小表示误差的类别。与传统的人脸识别方法相比, SRC 人脸识别方法能够有效地解决维数灾难问题与处理表情变化、光

到稿日期:2015-05-21 返修日期:2015-08-05 本文受国家自然科学基金资助项目(61401108, 61261023), 广西自然科学基金资助项目(2013GXNSFBA019272)资助。

施静兰(1990—),女,硕士生,主要研究方向为人脸识别、稀疏表示、图像处理等;常侃(1983—),男,博士,副教授,硕士生导师,CCF 会员,主要研究方向为压缩感知、稀疏表示、图像处理等, E-mail: pandack0619@163.com(通信作者);张智勇(1991—),男,硕士生,主要研究方向为稀疏表示、低秩矩阵及应用等;覃团发(1966—),男,博士,教授,主要研究方向为多媒体通信、无线传感器网络等。

照变化、噪声污染以及局部遮挡等问题,具有良好的鲁棒性。但是,为了提高字典的表示能力,每个类别需要大量的训练样本,而 SRC 算法的复杂度与字典的维度紧密相关,字典的维度过高将会增加稀疏编码的难度,计算将会非常耗时。同时,为了保证字典原子能够合理地遍布于潜在的子空间,需要仔细选择训练样本集。此外,原始的训练样本具有大量的冗余信息和噪声干扰信息,这些信息会对识别造成消极影响。因此,需要采用精简且具备区分度的、通过稀疏学习方式得到的字典来提高识别精度和速度。

M. Aharon 等人^[4]提出的 K 奇异值分解(K-Singular Value Decomposition, K-SVD)算法是一种经典的高效字典学习算法,由该算法学习得出的字典能够精简地代表训练样本。文献^[5]在 K-SVD 算法的基础上结合主成分分析方法提出了 K 主成分分析字典学习方法,该方法对误差项进行 PCA 分解,提取其主成分作为字典中原子的更新,获得了较好的字典学习效果。但是它牺牲了效率,计算速度慢。文献^[6]在 K-SVD 算法的基础上加入分类误差约束项,提出了一种区分性 K-SVD (Discriminative K-SVD, D-KSVD)字典学习算法,使学习得到的字典同时具有识别能力。但是,当字典尺寸过小时,稀疏编码结果的区分性将无法得到保证。文献^[7]提出了一种标签一致的 K-SVD(Label Consistent K-SVD, LC-KSVD)字典学习算法,该算法强制了在稀疏编码结果中标签的一致,最终学习得到一个过完备字典和一个最佳线性分类器,进一步提高了识别性能。但是,该字典学习算法对训练样本是批量学习的,即全部的训练样本一次性学习完成,计算复杂度高,并且对动态变化的数据集的适应性差。文献^[8]提出利用非负矩阵因式分解从原始基数据中训练出一组“metagenes”,这些“metagenes”使得样本的类聚更加鲁棒。由此得到启发,文献^[9]提出一种 metaface 字典学习算法,即从原始训练样本中学习得到一个更加精简和鲁棒的基组,称为 metafaces,并用其代替原始训练样本作为字典来对输入的测试图像进行稀疏编码。每个类别的 metafaces 单独进行学习,各个类别的 metafaces 学习互不影响,降低了计算复杂度。由 MFL 算法学习得出的 metafaces 在人脸识别中更具代表性,且更有利于提高 l_1 范数最小化问题的求解速度。但是, MFL 算法在固定字典更新稀疏表示系数时,并没有考虑同类样本的稀疏编码系数具有相似性这一特点,所以字典的判别能力还不够强,识别率还有提升的空间。

为了弥补上述 metaface 学习算法的缺陷,提出一种基于系数相似性的 metaface 学习算法。在更新稀疏表示系数阶段,将同类样本中前后两个样本的稀疏编码系数的相似性作为新的约束项引入原稀疏表示问题,以降低类内系数分散性,并提高字典的区分性,进而提高人脸识别的准确率。

2 基于稀疏表示的人脸识别方法

2.1 SRC 人脸识别算法

信号稀疏表示的目的就是在给定的超完备字典中用尽可能少的原子来表示信号,以获得更为简洁的信号表示方式,从而更容易获取信号中所蕴含的信息,更方便进一步对信号进行加工处理,如压缩、编码等^[10]。

用 $\mathbf{A}_i = [v_{i,1}, v_{i,2}, \dots, v_{i,n}] \in \mathbb{R}^{m \times n_i}$ 表示第 i 类训练样本,则来自该类别的测试样本 $\mathbf{y} \in \mathbb{R}^m$ 可以近似表示为 \mathbf{A}_i 的线性组合:

$$\mathbf{y} = \sum_{j=1}^{n_i} \alpha_{i,j} \mathbf{v}_{i,j} = \mathbf{A}_i \boldsymbol{\alpha}_i \quad (1)$$

其中, $\boldsymbol{\alpha}_i = [\alpha_{i,1}, \alpha_{i,2}, \dots, \alpha_{i,n_i}]^T \in \mathbb{R}^{n_i}$ 表示系数。假设有来自 k 个类别的 n 个训练样本,记为: $\mathbf{A} = [\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_K] = [v_{1,1}, v_{1,2}, \dots, v_{k,n_k}]$, 则输入的测试图像 \mathbf{y} 可以表示为 $\mathbf{y} = \mathbf{A}\boldsymbol{\alpha}$, 或近似表示为 $\mathbf{y} \approx \mathbf{A}\boldsymbol{\alpha}$, 其中, $\boldsymbol{\alpha} = [\alpha_1, \dots, \alpha_i, \dots, \alpha_K]$ 。

给定一个类别待定的测试样本 $\mathbf{y} \in \mathbb{R}^m$ 和一个预定义的过完备字典 \mathbf{A} , SRC 人脸识别所要求解的即是 \mathbf{y} 在 \mathbf{A} 上的稀疏表示,然后利用稀疏表示系数计算每个类别的重构误差,最后将测试样本判别归类为具有最小误差的类别。SRC 算法^[3]概括如下。

算法 1 SRC 算法^[3]

1. 输入: k 个类别的训练样本 $\mathbf{A} = [\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_k] \in \mathbb{R}^{m \times n}$, 测试样本 $\mathbf{y} \in \mathbb{R}^m$, 以及容差 $\epsilon > 0$ 。
2. 正则化 \mathbf{A} 的列向量,使其具有单位 l_2 范数标准。
3. 求解 l_1 范数最小化问题:

$$\hat{\mathbf{x}}_1 = \arg \min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{s. t. } \mathbf{A}\mathbf{x} = \mathbf{y}$$

或求解:

$$\hat{\mathbf{x}}_1 = \arg \min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{s. t. } \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2 \leq \epsilon$$

4. 计算残差:

$$r_i(\mathbf{y}) = \|\mathbf{y} - \mathbf{A}\boldsymbol{\delta}_i(\hat{\mathbf{x}}_1)\|_2, i=1, \dots, k$$

其中, $\boldsymbol{\delta}_i(\hat{\mathbf{x}}_1) \in \mathbb{R}^n$ 是一个 $\hat{\mathbf{x}}_1$ 中除了与第 i 类相关的元素外,其余元素均为零的新矢量。

5. 输出: $\text{identity}(\mathbf{y}) = \arg \min_i r_i(\mathbf{y})$

2.2 metaface 字典学习算法

鉴于 metagenes 在基因数据分析^[8]和图像处理字典学习^[4,11,12]中的成功应用, M. Yang 和 L. Zhang 等人^[9]提出,从原始训练数据 \mathbf{A} 中学习出一组 metafaces,记为 \mathbf{D} ,并用 \mathbf{D} 代替 \mathbf{A} 进行 SRC 人脸识别。

用 $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] \in \mathbb{R}^{m \times n}$ 表示某个类别的训练样本, \mathbf{X} 中的每一列为一个样本向量。metaface 字典学习的目标即是从 \mathbf{X} 中学习出字典 $\mathbf{D} = [\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_p] \in \mathbb{R}^{m \times p}$, 其中 $p \leq n$, 并且要求其中的每一个 metaface \mathbf{d}_j ($j=1, 2, \dots, p$) 均为单位列向量,即 $\mathbf{d}_j^T \mathbf{d}_j = 1$ 。求解字典 \mathbf{D} 的目标函数如下:

$$J_{D,\Lambda} = \arg \min_{D,\Lambda} \{ \|\mathbf{X} - \mathbf{D}\boldsymbol{\Lambda}\|_F^2 + \lambda \|\boldsymbol{\Lambda}\|_1 \} \quad (2)$$

$$\text{s. t. } \mathbf{d}_j^T \mathbf{d}_j = 1, \forall j$$

其中,参数 λ 为权衡 F 范数项和 l_1 范数项的权重系数。 $\boldsymbol{\Lambda} \in \mathbb{R}^{p \times n}$ 为 \mathbf{X} 在 metafaces \mathbf{D} 上的表示矩阵。式(2)为一个 metafaces \mathbf{D} 和表示系数矩阵 $\boldsymbol{\Lambda}$ 的联合优化问题。对于此类多变量优化问题,可以通过交替迭代优化 \mathbf{D} 和 $\boldsymbol{\Lambda}$ 进行求解。

固定 \mathbf{D} 求解 $\boldsymbol{\Lambda}$ 时,目标函数简化为:

$$J_{\Lambda} = \arg \min_{\Lambda} \{ \|\mathbf{X} - \mathbf{D}\boldsymbol{\Lambda}\|_F^2 + \lambda \|\boldsymbol{\Lambda}\|_1 \} \quad (3)$$

该最小化问题可通过一些标准的凸优化方法求解,在文献^[9]中采用文献^[13]中的算法求解。

固定 $\boldsymbol{\Lambda}$ 更新 \mathbf{D} 时,目标函数简化为:

$$J_D = \arg \min_D \{ \|\mathbf{X} - \mathbf{D}\boldsymbol{\Lambda}\|_F^2 \} \quad \text{s. t. } \mathbf{d}_j^T \mathbf{d}_j = 1, \forall j \quad (4)$$

将 $\boldsymbol{\Lambda}$ 记为 $\boldsymbol{\Lambda} = [\boldsymbol{\beta}_1^T, \boldsymbol{\beta}_2^T, \dots, \boldsymbol{\beta}_p^T]^T$, 即 $\boldsymbol{\beta}_j$ ($j=1, 2, \dots, p$) 表示 $\boldsymbol{\Lambda}$ 中与 \mathbf{d}_j 相应的第 j 行。逐一更新每个 metaface 向量。当更新 \mathbf{d}_j 时, \mathbf{D} 中的其他列 \mathbf{d}_l ($l \neq j$) 固定不变。则式(4)中的 J_D 转化为:

$$J_{d_j} = \arg \min_{d_j} \{ \|\mathbf{X} - \sum_{l \neq j} \mathbf{d}_l \boldsymbol{\beta}_l - \mathbf{d}_j \boldsymbol{\beta}_j\|_F^2 \} \quad \text{s. t. } \mathbf{d}_j^T \mathbf{d}_j = 1 \quad (5)$$

令 $\mathbf{Y} = \mathbf{X} - \sum_{i \neq j} \mathbf{d}_i \beta_i$, 使用拉格朗日算子, 然后对 \mathbf{d}_j 求偏导并令其等于 0, 可得式(5)的解为:

$$\mathbf{d}_j = \mathbf{Y} \beta_j^T / \|\mathbf{Y} \beta_j^T\|_2 \quad (6)$$

其中, $\|\cdot\|_2$ 为 l_2 范数。通过以上步骤便可更新全部的 \mathbf{d}_j , 由此更新整个 \mathbf{D} 。优化过程如算法 2。

算法 2 MFL 算法^[9]

1. 初始化 \mathbf{D} : 通过 l_2 范数将 \mathbf{D} 的每一列(每一个 metaface)初始化为一个随机向量。
2. 固定 \mathbf{D} 求解 \mathbf{A} : $\mathbf{J}_A = \arg \min_{\mathbf{A}} \{ \|\mathbf{X} - \mathbf{D}\mathbf{A}\|_F^2 + \lambda \|\mathbf{A}\|_1 \}$
3. 固定 \mathbf{A} 更新 \mathbf{D} : $\mathbf{d}_j = \mathbf{Y} \beta_j^T / \|\mathbf{Y} \beta_j^T\|_2$
对全部的 metafaces \mathbf{d}_j 逐一进行更新, 由此更新整个 \mathbf{D} 。
4. 返回步骤 2, 直到邻近的迭代中 $\mathbf{J}_{D,A}$ 的值足够接近, 或者已经达到迭代次数的最大值。
5. 输出字典 \mathbf{D} 。

3 基于系数相似性的字典学习算法

由 2.2 节可知, MFL 算法通过交替迭代更新稀疏表示系数 \mathbf{A} 和字典 \mathbf{D} , 每个类别的 metafaces 均单独训练, 最后用学习得到的所有类别的 metafaces 组成一个字典代替 SRC 中的训练样本集进行稀疏编码。MFL 算法在固定字典 \mathbf{D} 更新系数 \mathbf{A} 时, 并没有考虑同类样本之间的相关性, 学习得到的字典区分性不够高。为了利用同类样本稀疏编码系数之间的相似性信息来提高字典的区分性, 使学习得到的字典更具代表性, 提出一种新的字典学习算法——CS-MFL 算法。

通常而言, 不同人脸图像具有许多相似性, 即便是来自不同类别的人脸图像, 相似性仍然存在。假设有 A 与 B 两类人脸, 虽然根据 MFL 算法, 每个类别的 metafaces 是单独训练的, 但是仍有可能本应判成 A 类的某个测试图像能够被 B 类的 metafaces 稀疏表示。因此, 在人脸识别应用中, 字典的区分性尤为重要。如果在稀疏编码阶段, 将属于同一类别的人脸图像应具有相似的稀疏编码系数作为先验信息加以考虑, 降低字典的类内散布性, 那么学习得到的字典会更具有区分性。由此, 可降低不同类别的 metafaces 能够对同一测试图像进行稀疏表示的概率, 从而提高分类精度。所以, 基于此思想, 可以在式(3)的基础上将系数相似性作为约束项引入, 得到如下目标函数:

$$\mathbf{J}_A = \arg \min_{\mathbf{A}} \{ \|\mathbf{X} - \mathbf{D}\mathbf{A}\|_F^2 + \lambda_1 \|\mathbf{A}\|_1 + \lambda_2 \sum_{i=2}^n \|\mathbf{\Gamma}_{i-1} - \mathbf{\Gamma}_i\|_2^2 \} \quad (7)$$

其中, 参数 λ_2 权衡稀疏编码系数约束项在整个目标函数中的贡献度。 $\mathbf{\Gamma}_i$ 表示第 i 个训练样本 \mathbf{x}_i 在字典 \mathbf{D} 上的稀疏编码系数, 且有 $\mathbf{A} = [\mathbf{\Gamma}_1, \mathbf{\Gamma}_2, \dots, \mathbf{\Gamma}_n]$ 。

上述目标函数中, 约束了同类样本中前后相邻两个训练样本的稀疏编码系数相似, 因此相似性得以传递, 等效于约束了同类全部训练样本的稀疏编码系数均相似。通过约束同一类别的训练样本具有相似的稀疏编码系数, 可以降低字典的类内系数分散性, 进而提高字典的区分性。

需要注意的是, 因为 $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] \in \mathbb{R}^{m \times n}$ 表示的是某个类别的训练样本, 所以式(7)是针对单个类别进行训练。因此, 不同类别的训练过程互不影响, 这就保证了编码系数相似性仅在同类样本之间进行传递, 并不会出现编码系数相似性跨类别进行传递的现象。

为了解决问题(7), 将其中的 $\|\mathbf{X} - \mathbf{D}\mathbf{A}\|_F^2 + \lambda_1 \|\mathbf{A}\|_1$ 写成向量形式, 得到下述等价问题:

$$\begin{aligned} J_{(\mathbf{\Gamma}_i)} = \arg \min_{\mathbf{\Gamma}_i} \{ & \sum_1^n \|\mathbf{x}_i - \mathbf{D}\mathbf{\Gamma}_i\|_2^2 + \lambda_1 \sum_1^n \|\mathbf{\Gamma}_i\|_1 + \\ & \lambda_2 \sum_{i=2}^n \|\mathbf{\Gamma}_{i-1} - \mathbf{\Gamma}_i\|_2^2 \} \quad (8) \end{aligned}$$

但是, 直接求解新的优化问题即式(8)的复杂度仍然较高, 因此将其进一步拆分为若干子问题进行求解, 其中, 第 $i(2 \leq i \leq n)$ 个子问题可以写为:

$$J_{\mathbf{\Gamma}_i} = \arg \min_{\mathbf{\Gamma}_i} \{ \|\mathbf{x}_i - \mathbf{D}\mathbf{\Gamma}_i\|_2^2 + \lambda_1 \|\mathbf{\Gamma}_i\|_1 + \lambda_2 \|\mathbf{\Gamma}_{i-1} - \mathbf{\Gamma}_i\|_2^2 \} \quad (9)$$

即逐个计算训练样本稀疏编码系数。对于第一个训练样本, 由于没有相邻样本系数可参考, 因此按照如下公式求解:

$$J_{\mathbf{\Gamma}_1} = \arg \min_{\mathbf{\Gamma}_1} \{ \|\mathbf{x}_1 - \mathbf{D}\mathbf{\Gamma}_1\|_2^2 + \lambda_1 \|\mathbf{\Gamma}_1\|_1 \} \quad (10)$$

由于 $\|\mathbf{\Gamma}_{i-1} - \mathbf{\Gamma}_i\|_2^2$ 约束了前后两个相邻训练样本的稀疏编码系数相似, 根据此相似性的传递性, $\mathbf{\Gamma}_i (i=2, 3, \dots, n)$ 均与 $\mathbf{\Gamma}_1$ 相似, 从而约束了同类全部训练样本的稀疏编码系数相似。

事实上, 将式(9)所代表的 $n-1$ 个子问题的目标函数与式(10)的目标函数相加, 可以得到问题(8)的目标函数。由此说明, 式(9)和式(10)是问题(8)的有效近似。在将式(8)拆成 n 个子问题后(如式(9)、式(10)所示), 当求解 $\mathbf{\Gamma}_i$ 时, 因为 $\mathbf{\Gamma}_{i-1}$ 结果已知, 所以将 $\mathbf{\Gamma}_{i-1}$ 视为固定值, 此时式(9)中的未知向量只有 $\mathbf{\Gamma}_i$, 求解难度大大降低。

为了求解式(9), 将其重写为:

$$\begin{aligned} J_{\mathbf{\Gamma}_i} = \arg \min_{\mathbf{\Gamma}_i} \{ & \|\mathbf{x}_i - \mathbf{D}\mathbf{\Gamma}_i\|_2^2 + \lambda_1 \|\mathbf{\Gamma}_i\|_1 + \lambda_2 \|\mathbf{\Gamma}_{i-1} - \\ & \mathbf{\Gamma}_i\|_2^2 \} \\ = \arg \min_{\mathbf{\Gamma}_i} \{ & \left\| \begin{pmatrix} \mathbf{x}_i \\ \sqrt{\lambda_2} \mathbf{\Gamma}_{i-1} \end{pmatrix} - \begin{pmatrix} \mathbf{D} \\ \sqrt{\lambda_2} \mathbf{I} \end{pmatrix} \mathbf{\Gamma}_i \right\|_2^2 + \lambda_1 \|\mathbf{\Gamma}_i\|_1 \} \quad (11) \end{aligned}$$

其中, \mathbf{I} 为 $m \times m$ 的单位矩阵, 令 $\mathbf{x}_{new} = (\mathbf{x}_i^T, \sqrt{\lambda_2} \mathbf{\Gamma}_{i-1}^T)^T$, $\mathbf{D}_{new} = (\mathbf{D}^T, \sqrt{\lambda_2} \mathbf{I}^T)^T$, 则式(11)等价于:

$$J_{\mathbf{\Gamma}_i} = \arg \min_{\mathbf{\Gamma}_i} \{ \|\mathbf{x}_{new} - \mathbf{D}_{new} \mathbf{\Gamma}_i\|_2^2 + \lambda_1 \|\mathbf{\Gamma}_i\|_1 \} \quad (12)$$

该问题为典型的 l_2-l_1 形式的最小化问题。为与 MFL 算法比较, 同样采用文献[13]中的算法求解。

需要注意的是, 原式(8)中的 $\mathbf{\Gamma}_i$ 是与 $\mathbf{\Gamma}_{i-1}$ 相关的, 将式(8)拆分成 $n-1$ 个子问题(即式(9))和 1 个子问题(即式(10))是一种求解的近似。由于子问题和原始问题并不等效, 因此求解结果不一定为全局最优点, 反而很有可能陷入局部最优。但是, 这类迭代求解方法在优化问题中经常使用, 例如著名的 K-SVD 算法^[4], 将原始问题拆分成字典更新问题和稀疏系数编码问题迭代进行; 在课题组提出的其他算法中^[15,16], 也应用了此迭代求解方法, 证明了其有效性。这类方法的好处是能够将多变量优化问题转换为单变量优化问题处理, 这样就大大降低了求解的难度和计算的复杂度。本文的实验结果表明, 采用所提出的 CS-MFL 算法可以提升识别率, 间接说明了上述求解算法的有效性。

综上, 将提出的 CS-MFL 算法概括为算法 3。在 CS-MFL 算法中, 迭代更新字典 \mathbf{D} 和稀疏编码系数 \mathbf{A} 。在固定 \mathbf{D} 求解 \mathbf{A} 时, 利用同类训练样本稀疏编码系数的相似性, 针对每个训练样本, 求解问题(9)或(10)。在固定 \mathbf{A} 更新 \mathbf{D} 时, 按照文献[9]所述的方式进行。

算法 3 CS-MFL 算法

1. 初始化 \mathbf{D} : 通过 l_2 范数将 \mathbf{D} 的每一列(每一个 metaface)初始化为一个

个随机向量。

2. 固定 D 求解 A :

依次求解每个训练样本 x_i 的稀疏编码系数 Γ_i , 其中, $i=1, 2, \dots, n$ 。

若 $i=1$, 则

$$J_{\Gamma_1} = \arg \min_{\Gamma_1} \{ \|x_1 - D\Gamma_1\|_2^2 + \lambda \|\Gamma_1\|_1 \}$$

若 $1 < i \leq n$, 则

$$J_{\Gamma_i} = \arg \min_{\Gamma_i} \{ \|x_i - D\Gamma_i\|_2^2 + \lambda_1 \|\Gamma_i\|_1 + \lambda_2 \|\Gamma_{i-1} - \Gamma_i\|_2^2 \}$$

3. 固定 A 更新 D ; $d_j = Y\beta_j^T / \|Y\beta_j^T\|_2$

对全部的 metafaces d_j 逐一进行更新, 由此更新整个 D 。

4. 返回步骤 2, 直到邻近的迭代中 $J_{D,A}$ 的值足够接近, 或者已经达到迭代次数的最大值。

5. 输出字典 D 。

相较于 MFL 算法, CS-MFL 算法改进了系数更新方式, 将目标函数转化为典型的最小化问题后, 依旧可以通过一些标准的凸优化方法求解。此外, CS-MFL 算法并未改变字典更新方式, 因此 CS-MFL 算法的复杂度与 MFL 算法的相近。

CS-MFL 算法通过迭代更新稀疏表示系数 A 和字典 D 。实验中, 利用 CS-MFL 算法依次单独训练每个类别的 metafaces, 最后将学习得到的所有类别的 metafaces 组成一个字典代替原始的训练样本集进行 SRC 人脸识别。即分类过程如算法 1 的步骤 3—步骤 5 所述, 首先求解测试样本在由训练得到的新字典上的稀疏编码系数, 然后利用该系数计算每个类别的表示误差, 最后将测试样本判别为具有最小表示误差的类别。

4 实验结果与分析

为了验证所提 CS-MFL 算法的有效性, 分别在 Extended Yale B、AR 和 ORL 3 个人脸库上进行实验, 并与 SRC、MFL 算法进行比较。实验中, 首先采用 Eigenfaces^[14] 对人脸图像进行降维。由于约束了同类训练样本中前后两个相邻样本的稀疏编码系数相似, 因此第一个训练样本的选择是关键。实验中对于每一个类别, 选择正面无遮挡、无表情变化、光照正常的人脸图像作为第一个训练样本。

为保证相同的实验条件, CS-MFL 算法中的 λ_1 与 MFL 中的 λ 取相同值, 即 $\lambda_1 = \lambda = 0.005, \lambda_2 = 1$ 。接下来分别对 3 个人脸库的实验结果进行分析。

(1) Extended Yale B Database. Extended Yale B 人脸库包含 38 个人的 2414 张正面的人脸图像, 在多方面的实验室控制光照条件下拍摄而得。Extended Yale B 人脸库中部分人脸图像如图 1 所示。对于每个类别, 随机选择一半的人脸图像作为训练样本(每类 32 张人脸图像), 剩下的作为测试样本。实验中字典原子个数 p 与文献[9]中相同, 取 $p=18$, 即每类包含 18 个 metafaces。



图 1 Extended Yale B 人脸库中的部分人脸图像

图 2 为 Extended Yale B 人脸库中 3 种算法的识别率和特征维度的关系曲线图。从图 2 中可以看出, 提出的 CS-MFL 算法性能均优于 MFL 算法和 SRC 算法, 识别率最高, SRC 算法识别率最低。与 MFL 算法相比, 在维度较高的情况下, CS-MFL 算法获得识别率的提升比在维度低时所获得识别率的提升明显; 当维度为 300 时, CS-MFL 算法识别率的

提升最大。表 1 列出了 3 种算法所获得的最高识别率及其相应的特征维度。在使用更少字典原子的情况下, MFL 算法与 CS-MFL 算法均能够获得比 SRC 算法更高的识别率, 验证了字典学习的有效性。而 CS-MFL 算法在和 MFL 算法使用相同数量字典原子的情况下, 获得了更高的识别率, 说明由 CS-MFL 算法学习出的字典更具区分性。

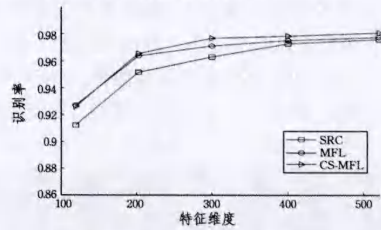


图 2 Extended Yale B 人脸库各算法识别率

表 1 Extended Yale B 人脸库中各算法最高识别率及其相应特征维度比较

算法	SRC	MFL	CS-MFL
识别率(%)	97.62	97.78	98.11
维度	520	520	520

(2) AR Database. AR 人脸库包含来自 126 个人超过 4000 张的正面人脸图像。每个人有 26 张人脸图像, 分别在 2 个不同时期不同光照变化、表情变化以及面部遮挡条件下拍摄采集而得。AR 人脸库中部分人脸图像如图 3 所示。实验中, 选择一组包含 50 个男性对象和 50 个女性对象的人脸子集。对于每个对象, 选择来自第一时期中的 7 张仅有光照变化和表情变化的人脸图像作为训练样本, 另外来自第二时期中的 7 张仅有光照变化和表情变化的人脸图像作为测试样本。由于每个对象所包含的训练样本个数非常有限(每类仅含 7 张人脸图像), 因此实验中字典原子个数取 $p=7$, 即每类包含 7 个 metafaces。



图 3 AR 人脸库中的部分人脸图像

图 4 为 AR 人脸库中 3 种算法识别率和特征维度的关系曲线图。从图 4 中可以看出, 在字典原子个数相同的情况下, CS-MFL 算法性能最佳, MFL 算法次之, SRC 算法最低。与 MFL 算法相比, 在维度较高时, CS-MFL 算法识别率提升较大; 同样在维度为 300 时, CS-MFL 算法识别率提升最大。表 2 列出了 3 种算法所获得的最高识别率及其相应的特征维度。由此可以看出, 由 CS-MFL 算法学习而得的 metafaces 比由 MFL 算法学习得到的 metafaces 更具代表性, 有效地提高了识别率。

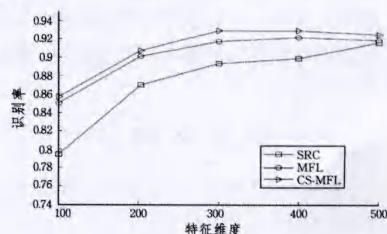


图 4 AR 人脸库各算法识别率

表2 AR人脸库中各算法最高识别率及其相应特征维度比较

算法	SRC	MFL	CS-MFL
识别率(%)	91.57	92.14	92.86
维度	500	400	300

(3) ORL Database. ORL人脸库包含40个人的400张人脸图像,分别在不同的光照条件下拍摄而得,并且有不同程度的脸部表情和脸部细节变化,人脸姿态深度旋转和平面旋转可达 20° 。ORL人脸库中部分人脸图像如图5所示。实验中,对于每个类别,随机选择6张人脸图像作为训练样本,剩下的4张作为测试图像。由于每类训练样本的数量较少,因此实验中字典原子个数取 $p=6$,即每类包含6个metafaces。



图5 ORL人脸库中的部分人脸图像

图6为ORL人脸库中3种算法识别率和特征维度的关系曲线图。从图6中可以看出,在字典原子个数相同的情况下,CS-MFL算法获得的识别率最高。与MFL算法相比,在维度为80时,CS-MFL算法识别率提升最大。表3列出了3种算法所获得的最高识别率及其相应的特征维度。由此可以看出,CS-MFL算法的分类结果更精确,验证了CS-MFL算法的有效性。

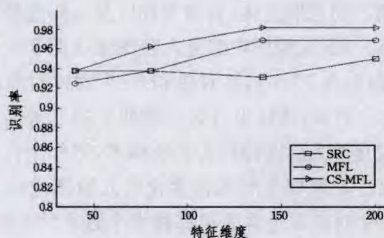


图6 ORL人脸库各算法识别率

表3 ORL人脸库中各算法最高识别率及其相应特征维度比较

算法	SRC	MFL	CS-MFL
识别率(%)	95	96.88	98.13
维度	200	200	200

结束语 为了利用同类训练样本稀疏编码系数的相似性信息来提高字典的区分性,提出了一种新颖的系数相似的metaface字典学习算法,并将其应用于人脸识别。CS-MFL算法在固定字典更新系数阶段加入同类训练样本稀疏编码系数相似的约束项,约束了同类训练样本中前后两个相邻样本的稀疏编码系数相似,降低了类内散布,提高了字典的区分性。在Extended Yale B、AR以及ORL 3个不同人脸库中进行实验,结果验证了CS-MFL算法能够取得比SRC算法和MFL算法更高的识别率,说明了该算法的有效性。未来的工作将从以下两方面展开:1)借鉴文献[17,18]中优化问题的求解方法来求解本文问题(7),以获取全局最优解;2)研究如何自适应约束同类训练样本稀疏编码系数相似,使算法更稳定,提高算法性能。

参考文献

[1] Donoho D. Compressed sensing [J]. IEEE Transaction on Information Theory, 2006, 52(4): 1289-1306
 [2] Candes E, Romberg J, Tao T. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency in-

formation [J]. IEEE Transactions on Information Theory, 2006, 52(2): 489-509
 [3] Wright J, Yang Y, Ganesh A, et al. Robust face recognition via sparse representation [J]. Pattern analysis and machine intelligence, 2009, 31(2): 210-227
 [4] Aharon M, Elad M, Bruckstein A. K-SVD: An algorithm for designing over-complete dictionaries for sparse representation [J]. IEEE Transactions on Signal Processing, 2006, 54(1): 4311-4322
 [5] Yu F P, Feng Y Q, Fan C L, et al. Dictionary learning based on principle component analysis [J]. Control and Decision, 2013, 28(7): 1109-1112 (in Chinese)
 余付平,冯有前,范成礼,等.基于主成分分析的字典学习[J].控制与决策, 2013, 28(7): 1109-1112
 [6] Zhang Q, Li B X. Discriminative K-SVD for dictionary learning in face recognition [C] // IEEE Conference on Computer Vision & Pattern Recognition, 2010: 2691-2698
 [7] Jiang Z L, 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
 [8] Brunet J P, Tamayo P, Golub T R, et al. Metagenes and molecular pattern discovery using matrix factorization [J]. Proceedings of the National Academy of Sciences, 2004, 101(12): 4164-4169
 [9] Yang H, Zhang L, Yang J, et al. Metaface learning for sparse representation based face recognition [C] // IEEE International Conference on Image Processing ICIP, 2010: 1601-1604
 [10] 郭金库,刘光斌,余志勇,等.稀疏信号表示理论及其应用[M].北京:科学出版社, 2013
 [11] Rubinstein R, Bruckstein A M, Elad M. Dictionaries for sparse representation modeling [J]. Proceedings of IEEE, 2010, 98(6): 1045-1057
 [12] Olshausen B A, Field D J. Emergence of simple-cell receptive field properties by learning a sparse code for natural images [J]. Nature, 1996, 381(6583): 607-609
 [13] Kim S J, Koh K, Lustig M, et al. A method for large-scale l1-regularized least squares [J]. IEEE Journal on Selected Topics in Signal Processing, 2007, 1(4): 606-617
 [14] Turk M, Pentland A. Eigenfaces for Recognition [J]. Journal of Cognitive Neuroscience, 1991, 3(1): 71-86
 [15] Chang K, Qin T F, Tang Z H. Reconstruction Algorithm for Compressed Sensing of Video Based on Joint Total Variation Minimization [J]. Chinese Journal of Electronics, 2014, 42(12): 2415-2421 (in Chinese)
 常侃,覃团发,唐振华.基于联合总变分最小化的视频压缩感知重建算法[J].电子学报, 2014, 42(12): 2415-2421
 [16] Chang K, Qin T F, Tang Z H, et al. Reconstruction of Compressed-sensed Video Using Compound Regularization [C] // Proceedings of the IEEE International Conference on Multimedia and Expo (ICME), July 2014: 229-234
 [17] Zheng M, Bu J, Chen C, et al. Graph regularized sparse coding for image representation [J]. IEEE Transactions on Image Processing a Publication of the IEEE Signal Processing Society, 2011, 20(5): 1327-1336
 [18] Cai D, He X, Han J, et al. Graph regularized nonnegative matrix factorization for data representation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(8): 1548-1560