

基于特征提取矩阵的稀疏系数求解算法

李 伟 李开宇

(南京航空航天大学自动化学院 南京 211106)

摘要: 在压缩感知算法的基础上,提出了在字典学习算法过程中同时训练得到一个投影矩阵,通过该矩阵可以直接运算求解稀疏系数的方法。字典训练过程采用的是KSVD字典学习算法,并与传统的 L_1 范数求解算法进行比较,通过实验可知,该方法比传统利用贪婪法等 L_1 算法具有更加快速、识别率更高的特点,提出的算法通过矩阵运算可以直接求解出系数项,而后者则是一个NP问题,需要利用迭代算法来求解,这样对于大样本的测试来说提出的算法具有更好的应用空间,节约的时间将非常显著。

关键词: 特征提取;字典学习;特征矩阵;稀疏系数

中图分类号: TN391.4 **文献标识码:** A **国家标准学科分类代码:** 460.40

Solving algorithm of sparse coefficient based on feature extraction matrix

Li Wei Li Kaiyu

(College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211016, China)

Abstract: On the basis of the compression sensing algorithm, this paper proposes to train a projection matrix in the process of dictionary learning algorithm, through which the method can obtain the sparse coefficient directly. The dictionary training process is based on the KSVD dictionary learning algorithm and is compared with the traditional L_1 norm solving algorithm. It can be seen from the experiment that the method has more rapid and higher recognition rate than the traditional L_1 algorithm using greedy method. The algorithm can solve the coefficient term directly through the matrix operation, while the latter is an NP problem, which needs to be solved by the iterative algorithm. For the large sample test, the proposed algorithm has better application space, and the time of saving will very noticeable.

Keywords: feature extraction; dictionary learning; feature matrix; sparse coefficient

0 引 言

对基于压缩感知的人脸识别算法来说,识别过程概括来说可以分为3个阶段,第1阶段是人脸字典的构造阶段,即通过训练样本处理得到一个矩阵字典;第2阶段是字典表示阶段,即获得待识别的图像在第一阶段构造的字典上面的系数表示;第3阶段是识别阶段,通过第二步获得的表示结果去进行图像分类。本文研究的内容是第2阶段的相关研究,对于求取字典表示这一命题,其中一种经典的方法是通过 L_2 范数约束通过最小二乘法求解;另一种比较经典的方法是在建立优化表达式时将表达式调整为基于 L_0 范数的最优化方法;最后一种方法是在限制等距性质(restricted isometry property, RIP)条件下,将问题转换到 L_1 范数最小化问题^[1-7]。这些算法通常需要较多的迭代次数和较多的计算时间来实现求解,其中某些算法甚至无法求解高维数据优化问题,本文在KSVD算法字典训练基础

上训练一个投影矩阵,通过该矩阵进行运算即可以求解得出稀疏系数,可以大大减小运算复杂度和运行时间。

本文基于人脸数据库AR库进行研究,使用KSVD字典学习算法^[8-9]。本文提到的3种求解方法不管从结果表示的准确性上还是求解难度和使用频率上,使用最多的是 L_1 范数求解算法,所以本文提出的求解算法最终与 L_1 范数求解算法进行比较。

1 KSVD 算法

1.1 KSVD 算法简介

KSVD算法可以看作K-means的一种泛化形式,K-means算法中每个信号量只能用一个原子来近似表示,而KSVD中每个信号是用多个原子的线性组合来表示的。KSVD通过构建字典来对数据进行稀疏表示,经常用于图像压缩、编码、分类等应用^[10-11]。

KSVD算法用于解决如下目标函数:

$$\min_{\mathbf{D}, \mathbf{X}} \{ \|\mathbf{Y} - \mathbf{D}\mathbf{X}\|_F^2 \} \quad \text{subject to } \forall_i, \|\mathbf{X}_i\|_0 \leq T_0. \quad (1)$$

式中: \mathbf{Y} 表示训练样本, \mathbf{D} 为字典, \mathbf{X} 为稀疏系数, T_0 是稀疏阈值。KSVD 能够求出上述表达式的最优字典。

1.2 KSVD 算法步骤

KSVD 算法是一种经典的字典训练算法, 依据误差最小原则, 对误差项进行奇异值分解(SVD), 选择使误差最小的分解项作为更新的字典原子和对应的原子系数, 经过不断的迭代从而得到优化的解。给定训练数据后一次找到全局最优的字典为 NP 问题, 只能逐步逼近最优解。构造 D 算法分两步: 稀疏表示和字典。KSVD 的求解是一个迭代过程。首先, 假设字典 \mathbf{D} 是固定的, 用 MP、OMP、BP 等算法, 可以得到字典 \mathbf{D} 上 \mathbf{Y} 的稀疏表示的系数矩阵 \mathbf{X} , 然后让 \mathbf{X} 固定, 根据 \mathbf{X} 更新字典 \mathbf{D} , 如此循环直到收敛为止^[12-13]。其具体步骤如下。

1) 初始化: 根据字典设置的尺寸以及训练样本初始化字典, 并将其归一化。设置初始迭代次数为 1。

2) 求稀疏系数: 利用贪婪法求取每一列训练样本 y_i 的稀疏系数 x_i , 其表达式如式(2)。

$$\min_{x_i} \{ \|y_i - \mathbf{D}x_i\|_2^2 \} \quad \text{subject to } \|\mathbf{X}_i\|_0 \leq T_0$$

$$i = 1, 2, \dots, N \quad (2)$$

3) 更新字典: 将字典进行一列一列的更新, 当更新某一系列时, 其他的列保持不变。

$$\|\mathbf{Y} - \mathbf{D}\mathbf{X}\|_F^2 = \left\| \mathbf{Y} - \sum_{j=1}^K d_j \mathbf{X}_T^j \right\|_F^2 =$$

$$\left\| \left(\mathbf{Y} - \sum_{j \neq k} d_j \mathbf{X}_T^j \right) - d_k \mathbf{X}_T^k \right\|_F^2 = \|\mathbf{E}_K - d_k \mathbf{X}_T^k\|_F^2 \quad (3)$$

然而, 用 SVD 分解求上面的列字典, 不满足稀疏约束条件。为了解决此问题, 定义

$$\omega_k = \{ i \mid 1 \leq i \leq K, \mathbf{x}_T^k(i) \neq 0 \} \quad (4)$$

定义 Ω_k 的尺寸为 $N \times \omega_k$, 其中 $(\omega_k(i), i)$ 的值为 1, 其他地方的值为 0。令 $\mathbf{E}_k^R = \mathbf{E}_K \Omega_k$, $\mathbf{x}_R^k = \mathbf{x}_T^k \Omega_k$ 。则有如下方程:

$$\|\mathbf{E}_k^R \Omega_k - d_k \mathbf{x}_R^k \Omega_k\|_F^2 = \|\mathbf{E}_k^R - d_k \mathbf{x}_R^k\|_F^2 \quad (5)$$

利用 SVD 分解, 能够求得最优的 d_k 。

4) 迭代次数加 1, 判断是否满足迭代停止条件, 若不足, 重复步骤 2)~4)。反之, 结束训练。

2 基于特征提取矩阵求解稀疏系数算法

2.1 特征提取矩阵原理

Zhang 等人^[14]的实验表明对于人脸图像协同表示比稀疏表示更为重要, 因此, 在测试时本文提出了利用训练时得到的投影矩阵 \mathbf{P} 来求稀疏系数的方法, 它比利用贪婪法等 L_1 算法更加快速, 对于大样本测试这将节约很多时间。为了在训练字典时得到投影矩阵 \mathbf{P} , 本文在目标模型上加了如下表达式:

$$\|\mathbf{P}\mathbf{Y} - \mathbf{X}\|_F^2 \quad (6)$$

式中: \mathbf{Y} 为训练样本, \mathbf{X} 为稀疏系数。在训练阶段, 字典是

为了对人脸图像稀疏系数的结构性进行约束, 而投影矩阵则是为了提取样本的稀疏系数, 实际上特征提取矩阵相当于字典的逆矩阵。在测试阶段用特征提取矩阵能够大大加快算法的运算速度, 节约时间。

2.2 特征提取矩阵与 L_1 算法比较

特征提取矩阵是利用训练得到的矩阵 \mathbf{P} 乘以测试样本得到稀疏系数的, 而 L_1 算的是利用训练得到字典来求稀疏系数的^[15], 前者可以直接求解, 相当于矩阵的乘法, 后者则是一个 NP 问题, 需要利用迭代法来求解, 这样会加大其运算的时间, 不利于在实际中的运用。

为了比较两种处理方法的效果, 选取了分别利用两种方法在 Extended Yale B 库中求取稀疏系数, 训练样本与测试样本的实验结果如图 1 所示。

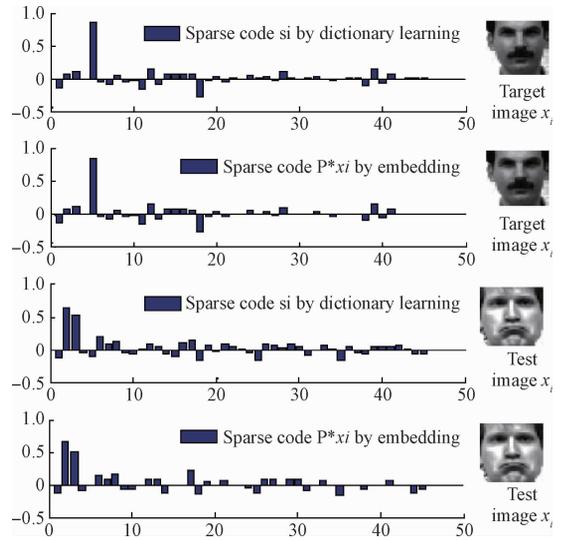


图 1 测试样本两种方法得到的稀疏系数

由图 1 可知, 采用特征提取矩阵和 L_1 算法得到稀疏表示稀疏大致是一致的, 而且采用特征提取矩阵得到稀疏系数没有忽略到稀疏表示中的主要成分, 一些次要部分被去除了, 这更有利于后期对于图像的分类。再加上采用特征提取矩阵求稀疏系数的运算量很小, 因此其更适合在实际中应用。

2.3 特征提取矩阵模型与求解

本次模型中, 加入了训练图像的标签 \mathbf{H} 以及判别稀疏编码标签 \mathbf{Q} ^[16]。假设有 9 个训练样本, 其中分为 3 类, 训练样本中, 第 1、2、3 列属于第 1 类, 第 4、5、6 列属于第 3 类, 第 7、8、9 列属于第 3 类。字典的尺寸为 6, 其中字典的第 1、2 列属于第 1 类, 第 3、4 列属于 2 类, 第 5、6 列属于第 3 类。则它们的 \mathbf{H} 和 \mathbf{Q} 如下所示:

$$\mathbf{H} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix} \quad (7)$$

$$Q = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix} \quad (8)$$

针对上述描述,本文提出了以下的人脸识别模型:

$$\langle D, X, W, P \rangle = \underset{D, X, W, P}{\operatorname{argmin}} \|Y - DX\|_F^2 + \alpha \|Q - AX\|_F^2 + \gamma \|PY - X\|_F^2 + \beta \|H - WPY\|_F^2 \text{ Sub } j \|x_i\|_0 \leq T_1, i \in \{j \mid j = 1, 2, \dots, N\} \quad (9)$$

式中: Y 为训练样本, D 为字典, X 为稀疏系数。 Q 为判别稀疏编码, A 为线性变换矩阵, P 为投影矩阵, H 为训练图像的标签, W 为分类器参数。 α, β, γ 为各项所占权重。

本节模型的求解步骤分为 2 步。1) 保持特征提取矩阵和分类器参数不变, 求解字典以及稀疏系数。2) 根据步骤 1) 求得的字典和稀疏系数, 更新特征提取矩阵以及分类器参数。其具体算法如下:

1) 保持 P, W 不变, 求解 D, X 。表达式如下:

$$\langle D, A, X \rangle = \underset{D, X, A}{\operatorname{argmin}} \|Y - DX\|_F^2 + \alpha \|Q - AX\|_F^2 + \gamma \|PY - X\|_F^2 \text{ Sub } j \text{ to } \|x_i\|_0 \leq T_1, i \in \{j \mid j = 1, 2, \dots, N\} \quad (10)$$

式(10)可以利用 KSVD 算法进行求解, 得到 D_{new}, X 。在 D_{new} 中得到字典 D 和 A , 并进行归一化。

2) 保持 D, X 不变, 求取 P, W 。

$$\langle P, W \rangle = \underset{P, W}{\operatorname{argmin}} \|PY - X\|_F^2 + \beta \|H - WPY\|_F^2 \quad (11)$$

其中, 先保持 W 不变, 求取 P :

$$P = (\gamma I^K + \beta W^T W)^{-1} \times (\gamma X Y^T + \beta W^T H Y^T) \times (Y Y^T + \lambda_3 I^T)^{-1} \quad (12)$$

然后保持 P 不变, 求 W :

$$W = H Y^T P^T (P Y Y^T P^T + \lambda_4 I^K)^{-1} \quad (13)$$

对于测试阶段的算法, 假设测试样本为 testdata , 测试样本的标签为 H_{test} , 根据训练样本得到的特征提取矩阵 P 以及分类器参数 W 进行分类, 具体步骤如下:

$$X = P \times \text{testdata};$$

$$H = W \times X \quad (14)$$

首先根据特征提取矩阵 P 与测试样本得到稀疏系数, 然后再根据稀疏系数以及分类器参数得到测试样本的标签, 测试标签中那个位置的参数最大则它就属于那一类, 然后在于实际预设的测试样本的类别在比较, 看它们是否一致。根据一致的数目与实际的测试样本数, 得到最终的识别率。

3 实验及结果分析

本次实验在 AR 人脸数据库以及 Extend Yale b 库上进行, 并对实验结果进行了分析。其中 Extended Yale B 主要反映的是光照的变化, AR 反映了表情变化, 光照变化以

及遮挡。实验测试设置为 $\alpha = 0.6, \beta = 1.5, \gamma = 0.4$ 。为了说明本文提出算法的有效性, 将把本文算法与 SRC 算法, KSVD 算法, DKSVD 算法, LCKSVD 算法进行比较^[17-19]。

3.1 AR 人脸数据库实验结果及分析

AR 人脸数据库(图 2)包括 126 个人的共计 4 000 幅人脸图像, 每个人有拍摄于 2 个时期的共计 26 幅图像, 每个时期 13 幅。其中有 3 幅围了围巾, 3 幅带了墨镜, 其他的图像或是人脸表情不同或是光照不同, 每幅图片的尺寸为 165×120 。本次实验中, 为了提高运算的速度, 利用随机投影的方式, 将样本变成 540 维的矩阵。

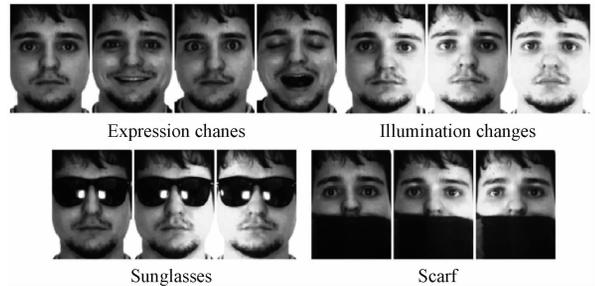


图 2 AR 数据库中一个人的样本图像

本次实验, 选取了 100 个人, 每人 26 幅人脸图像共计 2 600 幅, 其中男士 50 人, 女士 50 人作为样本。每一类中, 选取 20 幅作为训练样本, 用来得到字典以及投影矩阵。余下的样本作为测试样本。字典的尺寸设置成 500, 相当于每一类只选取了 5 个样本作为初始样本, 然后进行训练。AR 人脸数据库识别率如表 1 所示, 稀疏数据解时间如表 2 所示。

表 1 AR 人脸数据库的识别率

算法	SRC	KSVD	DKSVD	LCKSVD	本文
识别率/%	66.5	86.5	88.8	93.7	97.5

表 2 AR 人脸数据库稀疏系数求解时间

算法	系数求解时间/s
LKCSVD	0.281
本文	0.031

由表 1 和 2 可得, 在使用同一样本数据以及字典规模一样的情况下, 本文提出的算法的识别率明显高于其他四种比较主流的算法。其次, 在实际实验时, 发现本文算法的运行速度比其他的要快, 因为本文算法在测试阶段用投影矩阵求取稀疏系数, 而不是采用一般的 L_1 求解方法, 这就节约了很多时间。

3.2 Extend Yale B 实验结果及分析

Extended Yale B 库(图 3)包含了 38 个人的共计 2 414 幅正面人脸图像, 它们在不同的光照条件下拍摄的, 每个人

大概有 64 张图片,并且每幅人脸图像的尺寸为 192×168 。本次实验中,为了提高运算的速度,利用随机投影的方式,将样本变成 504 维的矩阵。



图 3 Extended-Yale-B 数据库中一个人的样本图像

本次实验,对于每一类人脸,选取其中 32 幅作为训练样本,余下的作为测试样本,字典的尺寸设置为 570,相当于每一类训练样本中去 15 个初始字典,然后进行训练。Extend-Yale-B 人脸数据库的识别如表 3 所示,稀疏稀疏求解时间如表 4 所示。

表 3 Extend-Yale-B 人脸数据库的识别率

算法	SRC	KSVD	DKSVD	LKCSVD	文本
识别率/%	80.5	93.1	94.1	95	95.2

表 4 Extend Yale B 人脸数据库稀疏系数求解时间

算法	系数求解时间/s
LKCSVD	0.592
Our Method	0.062

由表 3、4 可知,在不同的人脸数据库中,本文提出的方法在识别率和求解时间上都能取得有效的提高,对于在使用同一样本数据以及字典规模一样的情况下,本文提出的算法较其他四种算法有一些提高,并且系数求解运算速度快,因此适合在实际中进行应用,有一定的实用价值。综上所述,本文的算法适合在实际中应用。

4 结 论

本文在基于 KSVD 算法的人脸识别过程中引入了特征提取矩阵求取稀疏系数的算法,该方法区别于传统的 L_1 范数求解,不需要进行迭代运算处理,进行矩阵运算即可以算出对应的稀疏系数。本文通过对传统 L_1 范数求解和特征提取矩阵求解稀疏系数的实验可以得出如下一些结论:

- 1) 采用特征提取矩阵和 L_1 算法得到的稀疏系数基本一致,前者在次要成分上占的比例更少,即表示效果更好;
- 2) 采用特征提取矩阵进行稀疏系数求解相对于传统范数求解方法,在识别率时间上有很大的提高。
- 3) 采用特征提取矩阵求取系数的方法在识别率上相对于传统的范数求解方法也能够取得提高。

参考文献

[1] 周巍. L_1 范数最小化算法及应用[D]. 广州:华南理

工大学, 2013.

- [2] 刘杰, 李昆仑. 快速 L_1 范数最小化算法的性能分析和比较[J]. 电脑知识与技术, 2011, 7(19): 4641-4643.
- [3] 孙宁, 郭行, 刘佑鑫, 等. 非限制环境下的低秩协同人脸性别识别研究[J]. 电子测量与仪器学报, 2016, 30(11): 1790-1800.
- [4] 王东霞, 刘秋菊, 刘书伦. 一种 $L_{(2,1)}$ 范数最小化问题的算法研究[J]. 福州大学学报, 2013(1): 12-14.
- [5] 刘勇为. 关于混合 L_{1-1_2} 范数最小问题的算法研究[D]. 海口:海南师范大学, 2014.
- [6] 刘三明. 求解 L_0 -范数最小化问题的一种新方法[J]. 上海电机学院学报, 2012, 15(5): 345-350.
- [7] 孔繁镛, 王丹丹, 沈秋. L_1 - L_2 范数联合约束的鲁棒目标跟踪[J]. 仪器仪表学报, 2016, 37(3): 690-697.
- [8] RUBINSTEIN R, PELEG T, 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.
- [9] ZHANG Q, LI B. Discriminative K-SVD for dictionary learning in face recognition[C]. Computer Vision and Pattern Recognition, IEEE, 2010: 2691-2698.
- [10] 孔舒亚, 叶伟, 劳国超. 基于压缩感知的 SAR 宽带干扰抑制方法[J]. 电子测量技术, 2016, 39(11): 160-163.
- [11] ZHANG Z, LIANG Y, BAI L, et al. Discriminative sparse representation for face recognition [J]. Multimedia Tools and Applications, 2016, 75(7): 3973-3992.
- [12] 翟雪含, 朱卫平, 康彬. 结合 KSVD 和分类稀疏表示的图像压缩感知[J]. 计算机工程与应用, 2015, 51(6): 193-198.
- [13] 张敏. 基于 KSVD 的双字典学习算法及其在图像处理中的应用[D]. 北京:北京工业大学, 2015.
- [14] ZHANG Z, LI F, CHOW T W S, 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): 1-1.
- [15] 张宏星, 邹刚, 赵键, 等. 基于 Gabor 特征与协同表示的人脸识别算法[J]. 计算机工程与设计, 2014, 35(2): 666-670.
- [16] SHI Z, LU Y. An efficient initialization method for D-KSVD algorithm for image classification [C]. International Congress on Image and Signal Processing, IEEE, 2013: 1029-1034.

- [17] 胡文强, 齐向阳. 基于分布式压缩感知的双通道 SAR GMTI[J]. 国外电子测量技术, 2016, 35(7): 17-22.
- [18] 封睿, 李小霞. 基于 GLC-KSVD 的稀疏表示人脸识别算法[J]. 四川兵工学报, 2014, 35(4):88-92.
- [19] JIANG Z, LIN Z, DAVIS L S. Label consistent K-SVD: Learning a discriminative dictionary for

recognition [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2013, 35(11): 2651-2664.

作者简介

李伟,工学硕士,主要研究方向为计算机测控。
E-mail:654020341@qq.com