

文章编号:1005-0523(2012)01-0010-05

稀疏表示的人脸识别及其优化算法

郑轶¹, 蔡体健^{2,3}

(1. 华东交通大学轨道交通学院, 江西 南昌 330013; 2. 华东交通大学信息工程学院, 江西 南昌 330013; 3. 中南大学信息科学与工程学院, 湖南 长沙 410075)

摘要:稀疏表示的本质就是稀疏正规化约束下的信号分解。提出一种改进的正交匹配追踪算法,使运算量较高的矩阵求逆运算转变为轻量级的向量运算或向量与矩阵的运算,可以加快逆矩阵和大矩阵乘积的求解。将此算法应用于稀疏表示的人脸识别,探讨并验证了稀疏阈值的设置和训练字典的选择对人脸识别率和识别速度的影响。

关键词:压缩感知;人脸识别;稀疏表示;快速正交匹配追踪算法;特征提取

中图分类号:TP391.4

文献标识码:A

人脸识别问题是一个经典的模式识别问题。近年来,受到压缩感知理论的启发,基于稀疏表示的人脸识别技术得到了广泛研究。基于稀疏表示的人脸识别是利用训练图片构造字典,再通过求解一个欠定方程来求得测试图片的最稀疏线性组合系数,然后根据这些系数来对图像进行识别分类。

稀疏表示的人脸识别问题表示成数学形式为: $Y=AX$, 其中 $Y \in R^m$ 是 m 维自然信号, $A \in R^{m \times n}$ 是预定义的基(又称为字典), $X \in R^n$ 是自然信号在预定义基下的 n 维稀疏表示。在已知原始信号的基础上,求解其在预定义基下的稀疏表示,是一个稀疏编码问题,可有以下两种求解方式^[1-3]:

稀疏正规化约束下的稀疏编码:

$$\hat{X} = \arg \min \|Y - AX\|_2^2 \quad \text{s.t. } \|X\|_0 \leq K \quad (1)$$

误差约束下的稀疏编码:

$$\hat{X} = \arg \min \|X\|_0 \quad \text{s.t. } \|Y - AX\|_2^2 \leq \varepsilon \quad (2)$$

式中: $\hat{X}$ 是原始信号 Y 在预定义基下的稀疏表示系数; ε 是误差容限; K 是稀疏阈值; $\|\cdot\|_0$ 表示 l_0 范数,即列数向量中非0元素的个数。

稀疏编码与信号的压缩感知重构具有相同的数学形式,其主要的求解算法包括最小 l_0 范数法、贪婪迭代匹配追踪系列算法等。其中,匹配追踪类方法为其近似求解提供了有力工具,在稳定性和运行速度方面具有一定的优势。目前常用的匹配追踪类算法包括:正交匹配追踪(orthogonal matching pursuit, OMP)算法^[4-5],基于树型搜索的正交匹配追踪算法^[6],正则化正交匹配追踪算法^[7-8],压缩采样匹配追踪算法^[9]等。

1 正交匹配追踪算法

正交匹配追踪思想^[4-5]本质上是来自于 K “稀疏”,就是从完备字典的 N 个原子中寻找 K 个关键分量,这 K 个关键分量系数的绝对值应该比其它 $N-K$ 个分量大得多。算法在每一次的迭代过程中,从完备原子库里选择与信号最匹配的原子来进行稀疏逼近并求出余量,然后不断迭代选出与信号余量最为匹配的原子。为了减少迭代次数,算法通过递归对已选择原子集合进行正交化以保证迭代的最优性。具体见算法

收稿日期:2011-12-26

基金项目:国家自然科学基金项目(61165007);华东交通大学科研项目(09111004)

作者简介:郑轶(1970—),男,副教授,硕士,研究方向为机械工程与信息控制及处理。

1,其中第5步是贪心选择原子的步骤, q 是所选择原子的索引;第6步是将所选择的原子 A_q 添加到子空间 A_ϕ , A_ϕ 是 A 的子矩阵;第7步是对已选择的原子进行正交化,并计算信号的表示系数 $X=(A_\phi^T A_\phi)^{-1} A_\phi^T Y$,在此需要计算 $(A_\phi^T A_\phi)^{-1}$ 逆矩阵,矩阵的求逆运算是重量级的运算,往往需要对矩阵进行三角分解或QR分解;第8步是计算最小残差 R 。

算法1 正交匹配追踪算法

第1步 输入:字典 A ,列信号 Y ,误差容限 ε 或稀疏阈值 K

第2步 输出:稀疏的系数 X

第3步 初始化:设置残差 $R:=Y$,系数 $X:=0$, $A_\phi:=[]$

第4步 do while(迭代条件成立)

第5步 $q:=\max\{k\|\mathbf{A}_k^T R\|\}$

第6步 $A_\phi:=\begin{bmatrix} A_\phi & A_q \end{bmatrix}$

第7步 $X=(A_\phi^T A_\phi)^{-1} A_\phi^T Y$

第8步 $R:=Y-A_\phi X$

第9步 end do

2 改进的正交匹配追踪快速算法

正交匹配追踪算法的计算量主要集中在 $(A_\phi^T A_\phi)^{-1}$ 矩阵求逆,而 A_ϕ 在每次迭代中只是增加一列,因此对 $(A_\phi^T A_\phi)^{-1}$ 的求逆,完全可以用它的上一次迭代的逆矩阵推算得来。假设未更新前的子字典为 $\tilde{A}_\phi$,则更新后的子字典 $A_\phi=\begin{bmatrix} \tilde{A}_\phi & A_q \end{bmatrix}$ 。令 $\tilde{F}=\tilde{A}_\phi^T \tilde{A}_\phi$,则可以推得:

$$F=A_\phi^T A_\phi=\begin{bmatrix} \tilde{A}_\phi^T \tilde{A}_\phi & \tilde{A}_\phi^T A_q \\ A_q^T \tilde{A}_\phi & A_q^T A_q \end{bmatrix}=\begin{bmatrix} \tilde{F} & V \\ V^T & c \end{bmatrix} \quad (3)$$

式中: $V=\tilde{A}_\phi^T A_q$; $c=A_q^T A_q$ 。

由于 F 是对称正定矩阵,因此可令 $F^{-1}=\begin{bmatrix} X & Y \\ Y^T & w \end{bmatrix}$,又由于 $FF^{-1}=I$ (I 为单位矩阵),经过运算可得到:

$$w=1/(c-V^T \tilde{F}^{-1} V); Y=-\tilde{F}^{-1} Vw; X=\tilde{F}^{-1}(I-VY^T) \quad (4)$$

由于每次迭代都只是增加一列,因此在上式中 c 和 w 都是标量, V 和 Y 是向量,而 $\tilde{F}^{-1}$ 是上一次迭代的逆矩阵,是个已知矩阵,因此 F^{-1} 的求解变成了向量或向量矩阵的轻量级运算。

为了进一步减少正交匹配追踪算法的计算量,可以跳过残差 R 的计算,直接计算 $A^T R$,并充分利用上一次迭代的结果进行递推。其演算过程如下:

$$A^T R=A^T(Y-A_\phi X)=A^T(Y-A_\phi(A_\phi^T A_\phi)^{-1} A_\phi^T Y)=A^T Y-A^T A_\phi F^{-1} A_\phi^T Y \quad (5)$$

令 $\tilde{G}=A^T \tilde{A}_\phi$, $\tilde{Q}=\tilde{A}_\phi^T Y$,则 $G=\begin{bmatrix} \tilde{G} & A^T A_q \end{bmatrix}$, $Q=\begin{bmatrix} \tilde{Q}^T & (A_q^T Y)^T \end{bmatrix}$, G 和 Q 的值都可以从上一次迭代的值递推得到, $A^T R=A^T Y-GF^{-1}Q=A^T Y-GX$ 。需要注意的是每一次迭代过程中字典 A 也需要更新,即将上次选中的原子删除。同样 $\tilde{G}$ 也必需做同样的更新。改进的快速正交匹配追踪快速算法(fast orthogonal matching pursuit, FOMP)见算法2:

算法2 改进的快速正交匹配追踪快速算法

第1步 输入:字典 A ,列信号 Y ,误差容限 ε 或稀疏阈值 K

第2步 输出:稀疏的系数 X

第3步 初始化:设置残差 $R:=Y$, 系数 $X:=0$, $A_\phi:=[]$, $G:=[]$, $Q:=[]$, $n:=1$, product $0=A^T Y$

第4步 do while(迭代条件成立)

第5步 $q:=\max_k |A_k^T R|$

第6步 If $n>1$ then

第7步 计算 c, V, w, Y, X , 然后计算 $F^{-1}=\begin{bmatrix} X & Y \\ Y^T & w \end{bmatrix}$

第8步 Else

第9步 $F^{-1}=(A_q^T A_q)^{-1}$

第10步 End if

第11步 $A_\phi:=\begin{bmatrix} A_\phi & A_q \end{bmatrix}$

第12步 $G=\begin{bmatrix} \tilde{G} & A^T A_q \end{bmatrix}$, $Q=\begin{bmatrix} \tilde{Q} \\ A_q^T Y \end{bmatrix}$

第13步 $X=F^{-1}Q$

第14步 $A^T R=\text{product}0-GX$

第15步 更新 product0和 G

第16步 end do

3 在稀疏表示的人脸识别中的应用

基于稀疏表示的人脸识别利用测试人身份的稀疏性来进行人脸识别^[10]。数学表示形式为: $Y=AX$, 其中 $Y \in R^m$ 是测试人脸图像, $A \in R^{m \times n}$ 是用已标识的若干个人的 n 幅人脸图像形成训练字典, 每幅图像有 m 个像素点。 $X=[0, \dots, 0, X_i^T, 0, \dots, 0]^T \in R^n$ 是稀疏系数, 除了与测试图像对应的系数为非0, X 的其他值大多为0。由于一般情况下 $n \ll m$ (虽然潜在的情况是 $n \gg m$), A 是非完备字典, 高维的图像数据破坏了方程的欠定性, 无法利用现有的稀疏编码算法求解, 并且在现实中, 噪声或遮挡不可避免, 为此改为用 $Y=AX+e$ 来描述, 其中 $e \in R^n$ 是 n 维未知噪声, 再将此式变换一下:

$$Y=[A \ I] \begin{bmatrix} X \\ e \end{bmatrix} = BW \quad (6)$$

式中: I 是单位矩阵, $B=[A \ I] \in R^{m \times (n+m)}$; $W=\begin{bmatrix} X \\ e \end{bmatrix} \in R^{n+m}$, 此时 $n+m > m$, 因此基于稀疏表示的人脸识别问题变成求解以下最优化问题, 可以用以上的FOMP算法来计算测试样本的表示系数 X 。

$$\hat{W}_0 = \arg \min \|W\|_0 \quad \text{s.t. } Y=BW \quad (7)$$

人脸识别问题实际上是一个测试样本的归类问题。获得了测试样本的稀疏表示 X 后, 由于 X 中的非0系数是与字典中属于某一样本类的原子相关, 根据这些非0系数就可以简单快速判断测试样本所属类别。但是由于不同样本类的同一光照角度的图像很相似, 容易产生误判, 另外噪声的存在使得许多小的非0项和多个类别相关, 因此我们需要设计一个更精确的分类器。具体做法是: 将 X 中某样本类的系数保留, 其他样本类的系数清0, 然后计算测试信号的非线性逼近误差值, 最后比较所有样本类的误差值, 最小的非线性逼近误差所对应的类即是测试样本归属的类。

4 实验与参数设置

以 Yale B 人脸库^[11] (the yale face database B) 做人脸识别实验, 选用10个人某种姿势下的64种不同光照条件下的640幅图片, 每幅人脸图片大小为 84×96 像素, 对图片进行了直方图均衡化及归一化处理, 实验中选取每个人的一部分图片形成训练字典, 随机选取其他的图片做测试样本。所选用的机器是 Dell 笔记本电脑, Genuine Intel CPU 1.66 GHz, 1 G 内存。

1) 两种算法的对比。为了对比经典的OMP算法与改进的OMP算法的优劣,我们为每样本类选择15个光照图片,选择的原则是尽量考虑各种光照角度的图片,从而构建一个大小为 8064×8214 的训练字典。随机选择不在字典中的300幅人脸图片用于测试,设置稀疏阈值从5变化到50,在相同条件下应用这两种算法进行人脸识别实验,分别记录下两种算法平均100幅人脸识别时间以及对应的识别率。所得到的曲线图如图1所示。

由图1可知,改进的FOMP算法100幅人脸识别时间少于经典OMP算法,并且随着稀疏阈值的增加,两种算法的人脸识别时间的差距加大,当稀疏阈值 $K=50$ 时,两种算法的100幅人脸识别时间相差达到14秒多。

2) 影响识别率的主要因素。在实验中我们发现增加稀疏阈值,相应地人脸识别率随着提高。图2是设置字典大小为 8064×8214 (每样本类15张图片),并给测试图片增加不同比例椒盐噪声的情况下,人脸识别率也随稀疏阈值的变化曲线,由图2可知,在稀疏阈值较小时,增加稀疏阈值可明显提高识别率,但稀疏阈值增加到一定值时(如 $K>15$),识别率趋于稳定,分析其原因可知字典中每样本类的原子数目决定了稀疏阈值的设置。

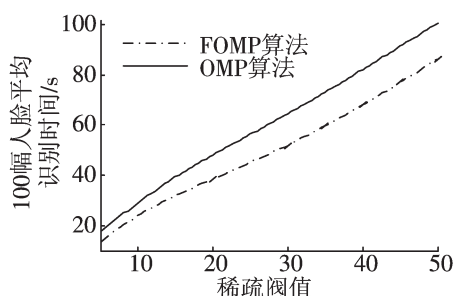


图1 FOMP算法与OMP算法运行速度的对比
Fig.1 The speed comparison between the OMP algorithm and the FOMP algorithm

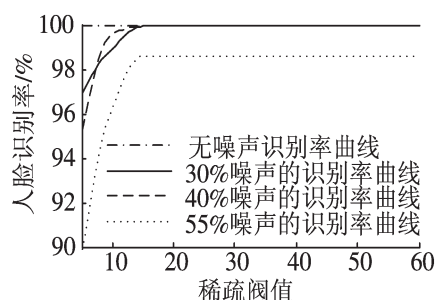


图2 不同噪声下人脸识别率随稀疏阈值的变化情况
Fig.2 Face recognition rate changes with the sparse threshold under different noise

为了进一步提高人脸识别率,需对训练字典进行调整。为每个样本类提供各种角度的光照图片,特别是极端光照条件下的人脸图片,能明显提高人脸识别率,见图3(a)。图4是设置稀疏阈值为20,分别为测试人脸添加10%,30%,55%的椒盐噪声的情况下,人脸识别率随训练字典变化的曲线。由图4可知,在字典大小达到一定值时(每样本类20幅图片),人脸识别率基本稳定下来,添加了55%椒盐噪声的测试样本的识别率仍可达到99.5%以上,这远远超过了人类肉眼的识别能力。图3(b)是为测试图像分别添加10%,30%,55%的椒盐噪声后的图片。从图3可知,噪声达30%及以上的图片,人类肉眼已很难识别了。

由图3和图4可知,在噪声较小的情况下(30%以下噪声),每样本类的光照图片只需20张,也就是说在光照变动不超过 30° 的情况下,人脸图片是可以完全正确识别的,Yale B人脸库光照的左右变动范围为 $-130^\circ \sim 130^\circ$,上下变动范围为 $-40^\circ \sim 90^\circ$,光照角度并不全,如果补全数据,性能将可以进一步得到提高。

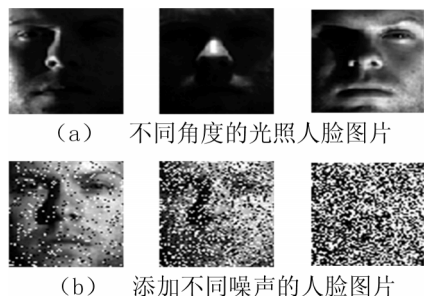


图3 训练图片和测试图片
Fig.3 training picture and testing picture

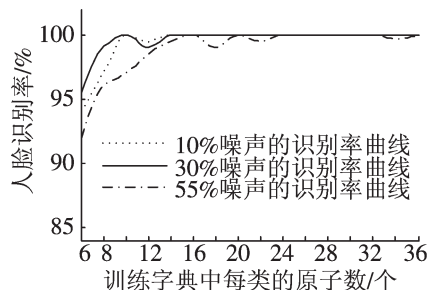


图4 不同噪声下人脸识别率随训练字典大小的变化情况
Fig.4 Face recognition rate changes with the size of dictionary under different noise

5 结论

研究提出了一种改进的FOMP算法,此算法是将重量级的矩阵求逆或大矩阵乘积运算转变为轻量级的向量运算或向量矩阵运算,从而提高了算法的运算速度。此算法可广泛应用于稀疏分解或压缩重构,特别适合于人脸识别以及类似的高维数据的分类。将此算法用于人脸识别实验,分析了影响人脸识别率的因素,通过实验验证了在光照角度的变动不超过 30° 的条件下,可以实现人脸的完全正确识别。如果进一步考虑人脸的表情、姿态等的变化,需要对训练字典的选择做进一步研究。

参考文献:

- [1] 戴琼海,付长军,季向阳. 压缩感知研究[J]. 计算机学报, 2011, 34(3): 425-434.
- [2] 李树涛,魏丹. 压缩传感综述[J]. 自动化学报, 2009, 35(11): 1369-1377.
- [3] 杨荣根,任明武,杨静宇. 基于稀疏表示的人脸识别方法[J]. 计算机科学, 2010, 37(9): 266-269.
- [4] TROPP J A, GILBERT A C. Signal recovery from random measurements via orthogonal matching pursuit [J]. IEEE Transactions on Information Theory, 2007, 53(12): 4655-4665.
- [5] WEI E, SHA I. Orthogonal matching pursuit algorithm for compressive sensing[EB/OL]. (2008-08-01)[2011-09-01]http://www.eee.hku.hk/~wsha/Freecode/freecode.htm.
- [6] KARABULUT G Z, MOURA L, PANARIO D, et al. Flexible tree search based orthogonal matching pursuit algorithm[C]// IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005: 673-676.
- [7] DEANNA N, ROMAN V. Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit [J]. IEEE journal of selected topics in signal processing, 2010, 4(2): 310-316.
- [8] 刘亚新,赵瑞珍,胡绍海,等. 用于压缩感知信号重建的正则化自适应匹配追踪算法[J]. 电子与信息学报, 2010, 82(11): 2713-2717.
- [9] DO T T, L GAN, NGUYEN N, TRAN T D. Sparsity adaptive matching pursuit algorithm for practical compressed sensing [C]//Asilomar Conference on Signals, Systems and Computers, 2008: 581-587.
- [10] WRIGHT J, GANESH A, YANG A, et al. Robust face recognition via sparse representation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(2): 210-227.
- [11] GEORGHIADES A S, BELHUMEUR P N, KRIEGMAN D J. From few to many: illumination cone models for face recognition under variable lighting and pose [J]. IEEE Trans Pattern Anal Mach Intelligence, 2001, 23(6): 643-660.

Face Recognition Based on Sparse Representation and Optimization Algorithms

Zheng Yi¹, Cai Tijian^{2,3}

(1. School of Railway Tracks and Transportation, East China Jiaotong University, Nanchang 330013, China; 2. School of Information Engineering, East China Jiaotong University, Nanchang 330013, China; 3. School of Information and Science Engineering, Central South University, Changsha 410075, China)

Abstract: The essence of sparse representation is the signal decomposition under the constraint of sparse regularized. This paper proposes an improved orthogonal matching pursuit algorithm, which would convert the high complexity operations of matrix into the lightweight operation of vector or vector-matrix computing. This improved method could accelerate the solution of inverse matrix and large matrix multiplication. This paper applies the algorithm to face recognition based on sparse representation, exploring and validating the effect of the sparse valve setting and training of the dictionary on face recognition accuracy and recognition speed.

Key words: compressed sensing; face recognition; sparse representation; orthogonal matching pursuit algorithm; feature extraction