

基于稀疏表示的高噪声人脸识别及算法优化

蔡体健^{1,2*}, 樊晓平^{1,3}, 刘遵雄²

(1. 中南大学 信息科学与工程学院, 长沙 410075; 2. 华东交通大学 信息工程学院, 南昌 330013;

3. 湖南财政经济学院 网络化系统研究所, 长沙 410205)

(* 通信作者电子邮箱 lao_cai68@126.com)

摘要:为提高基于稀疏表示人脸识别的速度和抗噪性能,研究了交叉花束(CAB)模型及压缩感知重构算法。针对重构算法中的大矩阵求逆,提出快速正交匹配追踪(FOMP)算法,可将运算量较高的矩阵求逆运算转变为轻量级向量矩阵运算。为增加高噪声图片的有效信息量,提出几种实用且有效的方法,并通过实验验证这些方法都能提高高噪声人脸识别率,可识别的噪声比例提高到75%,具有一定的实用价值。

关键词:压缩感知;稀疏表示;人脸识别;贪婪匹配追踪算法;过完备字典

中图分类号: TP391.4 **文献标志码:** A

Dense noise face recognition based on sparse representation and algorithm optimization

CAI Ti-jian^{1,2*}, FAN Xiao-ping^{1,3}, LIU Zun-xiong²

(1. School of Information Science and Engineering, Central South University, Changsha Hunan 410075, China;

2. School of Information Engineering, East China Jiaotong University, Nanchang Jiangxi 330013, China;

3. Institute of Networked Systems, Hunan University of Finance and Economics, Changsha Hunan 410205, China)

Abstract: To improve the speed and anti-noise performance of face recognition based on sparse representation, the Cross-And-Bouquet (CAB) model and Compressed Sensing (CS) reconstruction algorithm were studied. Concerning the large matrix inversion of reconstruction algorithm, a Fast Orthogonal Matching Pursuit (FOMP) algorithm was proposed. The proposed algorithm could convert the high complexity operations of matrix inversion into the lightweight operation of vector-matrix computation. To increase the amount of effective information in dense noise pictures, several practical and efficient methods were put forward. The experimental results verify that these methods can effectively improve the face recognition rate in dense noise cases, and identifiable noise ratio can reach up to 75%. These methods are of practical values.

Key words: Compressed Sensing (CS); sparse representation; face recognition; Orthogonal Matching Pursuit (OMP) algorithm; overcomplete dictionary

0 引言

人脸识别问题是一个经典的模式识别问题。近年来在压缩感知^[1-3]技术的影响下,人脸识别技术被归类到稀疏表示的技术框架下。基于稀疏表示的人脸识别是用已知分类的图片构造训练字典,再将待分类的图片表示成字典原子的线性组合,然后根据线性组合系数进行分类识别。其数学模型为: $Y = AX$, 其中 $Y \in \mathbf{R}^m$ 是 m 维待分类的图像; $A \in \mathbf{R}^{m \times n}$ 是由已知分类的若干人的 n 幅图像形成的训练字典,每幅图像由 m 个像素组成; $X \in \mathbf{R}^n$ 是 Y 在字典 A 下的表示系数。由于一般情况下,图像的维数较高,使得 $n \ll m$ (虽然潜在的情况是 $n \gg m$, 因为人的不同光照、不同角度的照片可以有无穷多张),从而破坏了方程的欠定性,为此需要使用一些降维技术对图像进行降维或者修改数学模型。目前稀疏表示的人脸识别常用的是交叉花束(Cross-And-Bouquet, CAB)模型^[4-5]:

$$Y = AX + E = [A \quad I] \begin{bmatrix} X \\ E \end{bmatrix} = BW \quad (1)$$

其中: $E \in \mathbf{R}^m$ 是未知的噪声干扰或遮挡, I 是单位矩阵, $B = [A \quad I] \in \mathbf{R}^{m \times (n+m)}$, $W = [X \quad E]^T \in \mathbf{R}^{n+m}$ 。

在 CAB 模型中,由于 $n + m > m$,使得 B 构成了过完备字典

典, $Y = BW$ 是欠定方程,因此可利用现有的压缩感知重构算法求解方程。

压缩感知重构算法主要包括最小 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^[6-7]、基于树型搜索的正交匹配追踪算法^[8]、正则化正交匹配追踪算法^[9-10]和压缩采样匹配追踪算法^[11]等。

1 快速的正交匹配追踪算法

最近的理论研究^[12]表明,对于 OMP 算法,只要测量矩阵满足参数为 $(1 + k, \delta)$ 的限制等容条件,其中 $\delta = (1 + \sqrt{k})^{-1}$, 则对任意稀疏度为 k 的向量都可以获得重建精度范围为 $\|X - \hat{X}\|_2 \leq O(\sqrt{k}) \|X - X_k\|_2 + O(\|X - X_k\|_1) + O(\sqrt{k}) \|E\|_2$ 的结果, X_k 代表 X 的 k 稀疏截断, E 代表测量噪声。文献[13]指出,一般情况下,OMP 算法在重构精度、收敛性和稳定性方面优于其他算法。

OMP 算法具体的操作步骤是:在每一次的迭代过程中,

收稿日期:2012-01-13;修回日期:2012-03-03。 基金项目:国家自然科学基金资助项目(60870010,61065003)。

作者简介:蔡体健(1968-),女,湖南长沙人,副教授,博士,主要研究方向:压缩感知、图像处理;樊晓平(1961-),男,浙江绍兴人,教授,博士生导师,博士,主要研究方向:无线传感器网络、智能信息处理、系统与控制;刘遵雄(1967-),男,江西瑞昌人,教授,博士,主要研究方向:机器学习、数据挖掘。

从完备原子库 A 里选择与信号最匹配的原子 A_q 来进行稀疏逼近, 并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, 然后不断迭代选出与信号余量最为匹配的原子。OMP 算法的计算量主要集中在正交化过程中的矩阵求逆, 文献[14]利用矩阵 Cholesky 或 QR 分解来减少矩阵求逆计算量, 在此利用递增求逆的方法来加速矩阵求逆。在正交化的公式 $X = (A_\varphi^T A_\varphi)^{-1} A_\varphi^T Y$ 中, A_φ 是由 A_q 组成的子字典, 由于 A_φ 在每次迭代中只增加一个原子, 因此对 $(A_\varphi^T A_\varphi)^{-1}$ 的求逆, 可以用它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}_φ , 则更新后的子字典 $A_\varphi = [\tilde{A}_\varphi \ A_q]$ 。令 $\tilde{F} = \tilde{A}_\varphi^T \tilde{A}_\varphi$, 则可以推得

$$F = A_\varphi^T A_\varphi = \begin{bmatrix} \tilde{A}_\varphi^T \tilde{A}_\varphi & \tilde{A}_\varphi^T A_q \\ A_q^T \tilde{A}_\varphi & A_q^T A_q \end{bmatrix} = \begin{bmatrix} \tilde{F} & V \\ V^T & c \end{bmatrix} \quad (2)$$

其中: $V = \tilde{A}_\varphi^T A_q$, $c = A_q^T A_q$ 。

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

$$\begin{cases} w = 1/(c - V^T \tilde{F}^{-1} V) \\ Q = -\tilde{F}^{-1} Vw \\ P = \tilde{F}^{-1} (I - VQ^T) \end{cases} \quad (3)$$

其中: c 和 w 都是标量, A_q 是一列向量, V 和 Q 也是列向量, \tilde{F}^{-1} 是上一次迭代的逆矩阵, 是个已知矩阵, 因此 F^{-1} 矩阵求逆变成了向量矩阵的轻量级运算。

FOMP 算法的具体操作步骤如下:

输入: 字典 A , 列信号 Y , 误差容限 ε 或稀疏阈值 k ;

输出: 稀疏的系数 X 。

- 1) 初始化: 设置残差 $R := Y$, 系数 $X := 0$, $A_\varphi := []$, $n := 1$;
- 2) do while(迭代条件成立)
- 3) $q := \{i | \max |A^T R|\}$
- 4) If $n > 1$ then
 计算 c, V, w, Q, P , 然后计算 $F^{-1} = \begin{bmatrix} P & Q \\ Q^T & w \end{bmatrix}$
 Else
 $F^{-1} = (A_\varphi^T A_\varphi)^{-1}$
 End if
- 5) $A_\varphi := [A_\varphi \ A_q]$
- 6) $X = F^{-1} A_\varphi^T Y$
- 7) $R := Y - A_\varphi X$
- 8) end do

算法中, 步骤3) 是贪心选择原子的步骤, 是在字典 A 的 i 列中选择与残差 R 的内积最大的原子, 并将其索引传送给 q ; 步骤4) 利用式(2)和(3)求解逆矩阵; 步骤5) 是将所选择的原子 A_q 添加到子字典 A_φ ; 步骤6) 是将信号正交投影到已选择的原子张成的子空间, 得到信号的表示系数 X ; 步骤7) 是计算最小残差 R 。此过程不断迭代, 直到残差足够小, 或残差的差值足够小时, 则退出迭代, 输出表示系数矩阵 X 。

FOMP 算法可提高 OMP 算法的运行速度, 在重构性能保证方面与 OMP 算法相同。

2 高噪声人脸识别

在人脸识别的研究中, 不可避免地要考虑噪声的问题。

已有研究表明, CAB 模型能完全正确地识别含 60% 噪声的图片^[15]; Wright 等^[16]从理论和实验上进一步证明, 当人脸图片的分辨率无穷增大, 字典 A 的各原子方差足够小, 表示系数 X 的支撑集无穷减小趋近于 1 时, 利用 CAB 模型纠正的错误率将无穷趋近于 100%。

这是一种理想的状态, 实际上随着噪声的增加, 图片中的有效信息减少, 很难得到表示系数的支撑集趋近于 1 的字典。但是对于高噪声图片, 可以采用一些直观而有效的方法来补偿信息量, 例如提高图片的分辨率^[16]; 增加字典中原子数, 使其尽可能地囊括人脸的各种变化; 或者采用多测试样本, 以弥补单样本的信息不足。

现实中的图像往往不是孤立的, 比如在视频中所截取的图像中, 每个样本类会有一系列图像, 即使是静态图像, 也会有多个样本。因此, 很容易为每个样本类采集多个样本, 形成测试集, 然后将测试集中的每个测试样本应用于 CAB 模型, 利用压缩感知重构算法求解每个测试样本的表示系数矩阵, 再比较各表示系数矩阵的稀疏性来进行分类识别。

信号的稀疏性度量通常有以下几种方法^[17]: l_0 范数、 l_p 范数、稀疏因子 (Sparseness Factor, SF) 和非线性逼近误差。当信号中含有噪声时, l_0 范数并不十分有效, 实际的信号表示方法中, 很少表示能使许多系数严格地等于 0。 l_p 范数度量系数非零个数和信号重建 l_2 误差之间的均衡, 当 $0 < p < 1$ 时, l_p 可以度量稀疏, l_p 范数定义如下:

$$\|X\|_p = (\sum_i X_i^p)^{1/p} \quad (4)$$

稀疏因子是等于表示系数中绝对值大于某一给定阈值的系数个数除以系数总个数。稀疏因子与所设定的阈值有关。非线性逼近误差也是衡量某种变换的稀疏性及压缩性能的重要指标, 其计算公式为: $\varepsilon = \|Y - BW\|_2$, 非线性逼近误差 ε 是残差的 l_2 范数, 在同样系数个数的条件下, 误差越小, 说明能量越集中, 稀疏性越好。

实验中, 采取 l_p 范数和非线性逼近误差相结合的方法来衡量表示系数的稀疏性。在单样本分类时, 将表示系数 X 中某样本类对应的系数保留下来, 其他样本类对应的系数清 0, 然后计算非线性逼近误差, 并比较所有样本类的非线性逼近误差, 得到最小误差的类。

在多样本分类时, 计算各样本的各类别累计非线性逼近误差, 根据累计最小误差判断样本类别。当误差的差别较小时, 则根据 l_p ($p = 0.5$) 范数判断表示系数的稀疏性, 然后选择稀疏性最高的样本类别。

3 仿真和实验

实验选用了扩展的 Yale B 人脸库^[18]中 30 人 64 种不同光照条件下的 1920 幅图片, 在每个人的 64 张图片中, 选择若干极端光照图片生成训练字典, 随机抽取不在字典中的其他图片作为测试样本, 对所有图片进行了直方图均衡化及归一化处理, 所选用的机器是 Dell 笔记本电脑, Genuine Intel CPU 1.66 GHz, 1 GB 内存。

3.1 算法的比较

为了对比经典的 OMP 算法与 FOMP 算法, 实验中为每样本类选择 15 个光照图片, 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 100 幅人脸的平均识别时间,所得到的曲线图如图 1 所示。由图 1 可知,FOMP 算法对 100 幅人脸的平均识别时间少于 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多达 14 s。

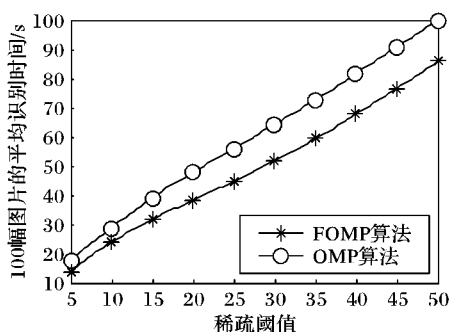


图1 FOMP 算法与 OMP 算法运行速度的对比

3.2 字典中的原子数目对识别率的影响

在单测试样本的实验中,为了提高表示系数 X 的稀疏性,对训练字典进行了学习,使字典尽可能地囊括各种光照角度的样本。具体做法是:在人脸库中寻找与字典中已有原子相关性的累计和最小的图片,然后排序添加入字典。调整字典中的原子数目,观察不同噪声下测试样本的识别率。图 2 是设置稀疏阈值为 40,在 10% 噪声、55% 噪声和 65% 噪声三种情况下,人脸识别率随训练字典中每个样本类的原子个数变化的曲线。由图 2 可知,随着字典中的原子数增加,识别率也随着提高;但是,当原子数达到一定值时(每样本类 20 个原子),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,这已远远超过了人类肉眼的识别能力。另外随着原子个数增加,人脸识别率出现有规律的波动,可能的原因是由于字典中增加了光照角度,而某些光照角度非常容易引起人脸识别的误差。因此并不是字典越大,识别率越高。

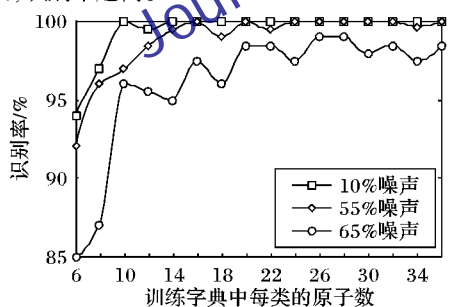


图2 各噪声比例下识别率随训练字典大小的变化曲线

3.3 测试样本数目对识别率的影响

为了检验测试样本数目对人脸识别率的影响,在同样的实验条件(图片分辨率 96×84 ,稀疏阈值 $k = 100$,字典中每样本类 20 个原子)下,通过改变测试样本的数目,测试不同噪声下的人脸识别率,测试结果如图 3 所示。由图 3 可知,当测试样本数为 1 时,能 100% 识别的图像噪声比例最高为 55%;当增加测试样本数后,此比例可达到 75%。图 4 是在不同噪声比例时,人脸识别率随着测试样本数改变的曲线。当测试样本数超过 10 以上时,噪声比例为 75% 的测试样本仍可以被 100% 正确地识别。

3.4 分辨率对识别率的影响

为了检验测试样本的分辨率对人脸识别率的影响,在同

样的实验条件(稀疏阈值 $k = 100$,字典中每样本类 20 个原子,测试样本数目为 10)下,通过改变测试样本的分辨率,测试在不同噪声下的人脸识别率,测试结果如图 5 所示。由图 5 可知,测试样本的分辨率越高,其信息量越大,人脸识别率就越高。

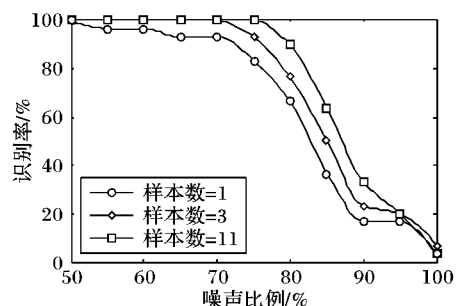


图3 各样本数目下识别率随噪声比例的变化曲线

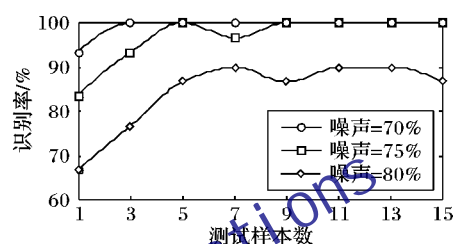


图4 各噪声比例下识别率随测试样本数改变的曲线

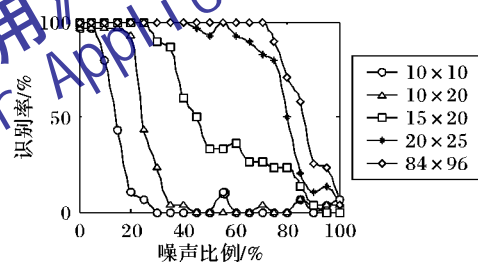


图5 各分辨率下识别率随噪声比例的变化曲线

4 结语

我们在对高噪声等信息量较少的图片进行肉眼识别时,经常需要从多个样本中获得信息才能得到比较完整的认识;或者通过提高图片的分辨率,增加描述特征等手段来得到更多的细节信息。为此,在稀疏表示的人脸识别中,提出了几种有效补偿高噪声图片信息量的方法,包括增加训练字典的原子数,增加测试样本数目,提高测试图片的分辨率,并通过实验证明了可 100% 识别的图片含噪比例达到 75%。针对 OMP 算法运算量大的问题,提出了一种改进的 FOMP 算法,可以将重量级的矩阵求逆运算转变为轻量级的向量矩阵运算,从而提高了算法的运算速度。但 FOMP 算法并没有改进 OMP 算法的局部最优问题,原子一旦选定后就不会退出,目前已有这方面的改进算法^[19]。今后的研究将在运算速度和精度方面进一步改进压缩感知的重构算法,可靠识别的图片含噪比例将进一步提高。

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] 石光明,刘丹华,高大化,等. 压缩感知理论及其研究进展[J]. 电子学报,2009,37(5): 1070-1081.

(下转第 2319 页)

由表2可知,当训练样本数为5时,TS+PCA方法的识别率为92.5%,略低于改进的M2DPCA方法,但高于另外三种方法;当训练样本数为7,测试样本数为4时,TS+PCA方法的识别率为95%,高于改进的M2DPCA方法,同时高于另外三种方法。

3 结语

本文介绍了一种新型的基于纹理的人脸识别方法,利用纹理分析技术提取人脸图像的TS特征,保障了特征的稳定高效性;PCA算法对高维的特征降维,解决了维数过高的问题;KNN分类器完成人脸识别,识别率高于PCA算法和LBP算法。然而,光照变化会对识别率产生影响,如何提高该方法在复杂情况下的识别率是进一步研究的重点。

参考文献:

- [1] 薛冰,郭晓松,蒲鹏程. 人脸识别技术综述[J]. 四川兵工学报, 2010, 31(7): 119-121.
- [2] SAMALIA F. Face recognition using hidden Markov models[D]. Cambridge: University of Cambridge, 1994.
- [3] 张练兴, 罗烈. 置信度判别嵌入式隐马尔可夫模型人脸识别[J]. 计算机应用研究, 2010, 27(5): 1987-1990.
- [4] SIGARI M H. Best wavelength selection for Gabor wavelet using GA for EBCG algorithm [C]// ICMV 2007: International Conference on Machine Vision. Piscataway: IEEE, 2007: 35-39.
- [5] KIRBY M, SIROVICH L. Application of the Karhunen-Loeve procedure for the characterization of human faces[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1990, 12(1): 103-108.
- [6] NICHOLL P, AMIRA A. DWT/PCA face recognition using automatic coefficient selection[C]// Proceedings of the 4th IEEE International Workshop on Electronic Design. Washington, DC: IEEE Computer Society, 2008: 390-393.
- [7] ZHU YULIAN, LIU JUN, CHEN SONGCAN. Semi-random subspace method for face recognition [J]. Image and Vision Computing, 2009, 27(9): 1358-1370.
- [8] HYVÄRINEN A, KARHUNEN J, OJA E. Independent component analysis [M]. New York: John Wiley & Sons, 2001: 147-164.
- [9] 黄璞, 陈才扣. 基于局部人脸图像的ICA人脸识别方法[J]. 计算机工程与设计, 2010, 31(11): 2550-2553.
- [10] HE D-C, WANG L. Texture Unit, texture spectrum, and texture analysis [J]. IEEE Transactions on Geoscience and Remote Sensing, 1990, 28(1): 509-513.
- [11] 陆春君, 陈松灿, 谭晓阳. 基于纹理频谱子集的纹理识别[J]. 江南大学学报: 自然科学版, 2007, 6(6): 753-757.
- [12] BARCELO A, MONTSENY E, SOBREVILLA P. Fuzzy texture unit and fuzzy texture spectrum for texture characterization[J]. Fuzzy Sets and Systems: Image Processing, 2007, 158(3), 239-252.
- [13] HUNG C-C, PHAM M, ARASTEH S, *et al.* Image texture classification using texture spectrum and local binary pattern [C]// IGARSS 2006: IEEE International Conference on Geoscience and Remote Sensing Symposium. Piscataway: IEEE, 2006: 2750-2753.
- [14] AHONEN T, HADID A, PETIKAINEN M. Face description with local binary patterns: application to face recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(12): 2037-2041.
- [15] 张岩, 武玉强. 改进的模块2DPCA人脸识别算法[J]. 计算机工程, 2011, 37(7): 228-230.
- [16] 杨颖娟. 基于PCA算法和小波包变换的人脸识别技术[J]. 微电子学与计算机, 2011, 28(1): 92-94.
- [17] DO T T, GAN L, NGUYEN N, *et al.* Sparsity adaptive matching pursuit algorithm for practical compressed sensing [C]// 42nd Asilomar Conference on Signals, Systems and Computers. Piscataway: IEEE, 2008: 581-587.
- [18] MO QUN, SHEN YI. Remark on the restricted isometry property in orthogonal matching pursuit[EB/OL]. [2012-01-18]. http://www.paper.edu.cn/index.php/default/en_releasepaper/content/4408549.
- [19] MALEH R. Improved RIP analysis of orthogonal matching pursuit [EB/OL]. [2012-01-18]. <http://arxiv.org/abs/1102.4311>.
- [20] 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): 267-270.
- [21] CANDÈS E J, TAO T. Decoding by linear programming[J]. IEEE Transactions on Information Theory, 2005, 51(12): 4203-4215.
- [22] WRIGHT J, MA Y. Dense error correction via ℓ_1 -minimization [C]// ICASSP 2009: IEEE International Conference on Acoustics, Speech and Signal Processing, 2009: 3033-3036.
- [23] 邓承志. 图像的稀疏表示及其应用研究[D]. 武汉: 华中科技大学, 2008.
- [24] GEORGIADIS 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 Transactions on Pattern Analysis and Machine Intelligence, 2001, 23(6): 643-660.
- [25] 方红, 杨海蓉. 贪婪算法与压缩感知理论[J]. 自动化学报, 2011, 37(12): 1413-1421.

(上接第2315页)

- [4] YANG A Y, ZHOU Z, MA Y, *et al.* Towards a robust face recognition system using compressive sensing [C]// Proceedings of Interspeech: 11th Annual Conference of the International Speech Communication Association. Makuhari, Chiba: [s. n.], 2010: 2250-2253.
- [5] WRIGHT J, GANESH A, YANG A Y, *et al.* Robust face recognition via sparse representation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(2): 210-227.
- [6] 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.
- [7] SHA W. Orthogonal matching pursuit algorithm for compressive sensing [CP/OL]. [2011-09-01]. <http://www.eec.hku.hk/~wsha/Freecode/freecode.htm>.
- [8] KARABULUT G Z, MOURA L, PANARIO D, *et al.* Flexible tree-search based orthogonal matching pursuit algorithm [C]// ICASSP '05: IEEE International Conference on Acoustics, Speech, and Signal Processing. Piscataway: IEEE, 2005: 673-676.
- [9] NEEDELL D, VERSHYNIN R. 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.
- [10] 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, 32(11): 2713-2717.