

文章编号:1004-4213(2011)02-0316-5

基于聚类的自适应图像稀疏表示算法及其应用*

徐健^{1,2}, 常志国^{3,4}

(1 西安邮电学院 通信与信息工程学院, 西安 710121)

(2 西安交通大学 电子与信息工程学院, 西安 710049)

(3 长安大学 信息工程学院, 西安 710064)

(4 陕西省道路交通检测与装备工程技术研究中心, 西安 710064)

摘要:提出了一种针对一类图像进行稀疏表示的字典训练方法,并证明了该算法的收敛性.该算法的几何解释是,以最少的超平面来逼近样本所在的一小块球冠.算法流程为聚类每一步迭代所产生的余项,将聚类中心作为新的字典原子,令字典能够更适应于样本的稀疏表示.该算法与传统的字典训练方法相比具有适应性强,对训练样本规模和字典规模要求低,收敛速度快,算法复杂度低等特点.利用该算法训练得到的字典用于压缩感知、图像去噪等实验表明,该字典具有很好的效果.

关键词:稀疏表示;聚类;压缩感知;字典;原子;稀疏度

中图分类号:TP391.4

文献标识码:A

doi:10.3788/gzxb20114002.0316

0 引言

压缩感知是一个目前比较热门的学科.压缩感知理论证明:高分辨率的信号或者图像可以由比奈奎斯特理论所规定低的多的采样点数恢复^[1-2].相关研究中,矩阵的稀疏表示是一个重要内容.

目前关于稀疏表示的研究有两个方向:第一个方向是事先定义好字典数学表达式的表示方法^[3].例如曲波(curvelets)变换、轮廓波(contourlet)变换、楔形波(wedgelet)变换、带状波(bandlet)等.这些变换都拥有快速算法,因此被广泛应用于图像的压缩、去噪和修补.但是这些算法对图像的几何特征过于依赖.对于拥有多种几何特征的自然图像,事先定义数学表达式的稀疏表示方法并不理想.还有一种是基于学习的表示方法.例如:核奇异值分解(Kernel Singular Value Decomposition, K-SVD)^[4],核特征值分解(Kernel Eigen Value Decomposition, KEVD)^[5]和核模糊代码本估计(Kernel Fuzzy Codebook Estimation, KFCE)^[6].这些算法都没有事先定义字典的解析式,但是通过训练过程逐步产生字典.通过该字典对数据进行变换往往采用正交匹配追踪算法(Orthonormal Matching Pursuit, OMP)^[7]或基追踪算法(Basic Pursuit, BP)^[8].由于这些算法中需要进行奇异值分解

(Singular Value Decomposition, SVD)或主分量分析(Principle Component Analysis, PCA),由于其算法复杂度高,所以无法一次处理较大的图像.算法对于大图的处理是通过将大图分为小块图像进行处理的.

稀疏度和字典规模是衡量稀疏表示有效性的主要性能指标.稀疏度指表示系数中不为0的个数.字典规模指字典矩阵的列数.

本文提出了一种新的字典训练算法,该算法专门针对一类图像的稀疏表示问题.传统算法在字典训练之初就固定字典规模,这样就使字典训练过程中每次都要对所有原子进行更新.本文所提及的算法开始并不固定字典规模,在字典训练过程中逐步增加原子个数,根据样本的情况分层产生原子,并根据需要自适应调节原子个数,并且避免了对矩阵进行SVD或PCA,因此能够有效降低算法复杂度.实验证明,该算法在同样的稀疏度和字典规模下,拥有比传统算法更好的性能和更高的运算效率.

1 算法原理

取大小相等的图像作为训练样本,将二维图像变成一维信号进行处理,并且对该一维信号进行归一化.归一化是因为如果一个图像乘以某个系数与另一个图像相等,认为它们表示的是同一内容、同一

基金项目:中央高校基本科研业务费专项资金(No. CHD2009JC156)、西安邮电学院青年教师基金(No. ZL2010-21)和长安大学基础研究支持计划专项基金资助

第一作者:徐健(1981-),女,讲师,博士研究生,主要研究方向为图像信号处理与模式识别. Email: xujian_paper@126.com

收稿日期:2010-07-23;修回日期:2010-09-21

图像. 经过归一化处理, 训练样本的能量相等. 可以把训练样本看作起始点在原点的归一化的矢量, 那么不同的训练样本只是对应不同方向的矢量. 如果训练样本均为某一类图像, 那么这些图像信号具有较大的相似性. 因此这些矢量可以看作空间上占据单位超球的一小块球冠, 稀疏表示问题也可以看作利用较少的超平面来逼近这一小块球冠.

这种算法和美术上经常使用多边形切割来逼近圆形是同一个原理, 如图 1.

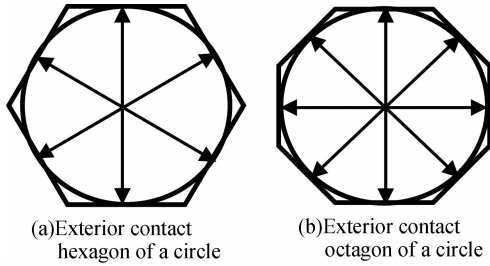


图 1 圆的外切多边形

Fig. 1 Exterior contact polygonal of a circle

在平面上的与原点距离为 1 的直线可以由与它相垂直的归一化矢量表示. 例如: 从原点出发与直线 $kx + y + b = 0$ 垂直的矢量为 $(k, 1)$, 该矢量归一化结果为 $l = \left(\frac{k}{\sqrt{k^2 + 1}}, \frac{1}{\sqrt{k^2 + 1}} \right)$, 当 b 变化时, 可以产生一组互相平行的直线. 可以确定 b 得到唯一的直线 l , 该直线与原点之间的距离为 1, 则 l 与 l 一一对应, 如图 2.

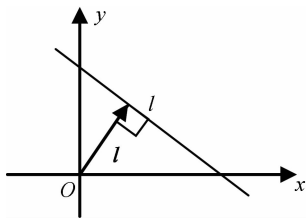


图 2 在平面上与原点距离为 1 的直线与归一化矢量一一对应

Fig. 2 A straight line with distance one to the origin uniquely correspond to a normalized vector on a 2-Dimensional Plane

同理在三维空间中, 一组平行平面的方程可以使用与这组平行平面相垂直的矢量来表示. 例如平面方程 $ax + by + cz + d = 0$, 矢量 (a, b, c) 必然与该组平面相垂直. (a, b, c) 矢量归一化可以得到

$$P = \left(\frac{a}{\sqrt{a^2 + b^2 + c^2}}, \frac{b}{\sqrt{a^2 + b^2 + c^2}}, \frac{c}{\sqrt{a^2 + b^2 + c^2}} \right)$$

可以确定唯一的 d 得到与原点的距离为 1 平面 P , 则平面 P 与矢量 P 一一对应.

以此类推, 一个 n 维的矢量必然能够确定 n 维空间中的一组超平面, 如果该矢量进行归一化, 可以确定唯一的超平面与之对应.

如果一类信号被看作是 n 维矢量, 这类信号必然只占据超球面的一小块球冠. 这样可以使用一组超平面来逼近球冠, 达到稀疏表示的目的.

假设初始的图像矢量 y_1 和 y_2 (图 3), 采用 K-均值聚类的方法可以得到聚类中心 a . 这些聚类中心构成了第一层的原子.

用每个矢量减去它们各自在聚类中心的投影得到差向量 e_1, e_2 , 将 e_1, e_2 归一化可以得到第二层待聚类的矢量 e'_1 和 e'_2 . 由于每个矢量与其聚类中心上的投影相减所得到的矢量都与聚类中心矢量垂直, 那么这些聚类都处于与聚类中心相垂直的超平面上.

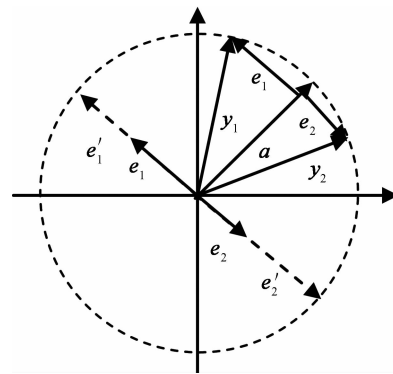


图 3 算法的几何解释

Fig. 3 Geometrical interpretation of algorithm

可以证明该算法的收敛性(图 4): 假设样本为 y . y 在聚类中心 c_1 上的投影为 y' , 与聚类中心之间的夹角为 φ_1 , 则 y' 与 y 之间的差值为 e_1 . 由图 4 可以看出 $e_1 = y \cdot \sin \varphi_1$. 在 y 与 \hat{a}_1 都归一化的情况下, y 与 \hat{a}_1 夹角越小, 则 e_1 越小. 如果 \hat{a}_2 是第二层聚类中心, e_1 在 \hat{a}_2 上的投影是 y_2 , y_2 与 e_1 间的夹角为 φ_2 , y_2 与 e_1 之间的差为 e_2 . $e_2 = e_1 \cdot \sin \varphi_2 = y \cdot \sin \varphi_1 \cdot \sin \varphi_2$.

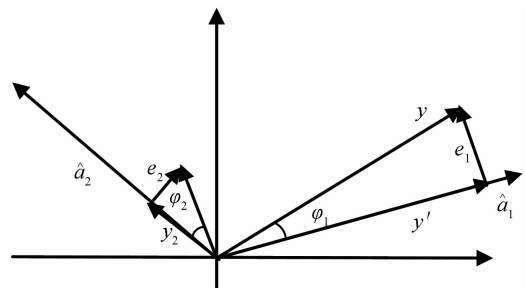


图 4 算法收敛性证明

Fig. 4 The convergence of the algorithm

以此类推, 经过 n 次迭代, $e_n = y \cdot \sin \varphi_1 \cdot \sin \varphi_2 \cdots \sin \varphi_n$, 只要 $\varphi_1, \varphi_2, \dots, \varphi_n$ 都小于 90° , 则 $e_1 > e_2 > \dots > e_n$, 算法收敛. 又由于 $\varphi_1, \varphi_2, \dots, \varphi_n$ 不可能等于 90° , 如果等于 90° , x 在该聚类中心上的投影长度为 0. 因此该算法一阶收敛.

2 算法流程

初始化 $Y=[y(1), y(2), \dots, y(N)] \in \mathfrak{R}^{n \times N}$, 每一列为一个训练样本. 给出各个分层的信号的原子数 $k_{(1)}, k_{(2)}, \dots, k_{(l)}$. l 为分层的数目.

迭代过程:

1) 将 N 个样本采用 K-均值聚类的方法将 $Y=[y(1), y(2), \dots, y(N)] \in \mathfrak{R}^{n \times N}$ 划分成 $k_{(1)}$ 个不同的类别 $\theta_{(1)}(\hat{a}_i), i=1, \dots, k_{(1)}$, 其中 $\theta_{(1)}(\hat{a}_i)$ 是一个矢量集. 当观察矢量满足

$$d(y(t), \hat{a}_i) = \min \{d(y(t), \hat{a}_j), j=1, \dots, k_{(1)}\} \quad (1)$$

认为 $y(t) \in \theta_{(1)}(\hat{a}_i)$, 其中 $d(y(t), \hat{a}_i)$ 是 $y(t)$ 和聚类中心 \hat{a}_i 之间的距离. 将聚类中心 $[\hat{a}_1, \hat{a}_2, \dots, \hat{a}_{k_{(1)}}]$ 作为第一层原子.

2) 假设 $[p_1, p_2, \dots, p_N]$ 为 $[y_1, y_2, \dots, y_N]$ 在各所属聚类中心上的投影, 则 $e_i = y_i - p_i, i=1, 2, \dots, N$.

3) 对任意的 j 和 $s, j \neq s$. 当 $\|e_j - e_s\|_2^2 < \epsilon$ 或 $\|e_j + e_s\|_2^2 < \epsilon$, 则去掉 e_s , 只保留 e_j , 其中 $\|\cdot\|_2^2$ 表示向量 2 范数的平方. 该步的意义是: $\|e_j - e_s\|_2^2 < \epsilon$ 则表示 e_j 和 e_s 距离非常近, 可以用同一个原子表示. 如果 $\|e_j + e_s\|_2^2 < \epsilon$ 表示 e_j 和 $-e_s$ 距离非常近, 也可以用同一个原子表示, 表示系数只相差一个符号. 对于这类样本合并处理可以降低算法复杂度. 合并后, 将 e_i 归一化得到 e'_i .

4) 对 e'_i 进行 K-均值聚类, 得到 $\theta_{(j)}(\hat{a}_i), i=1, \dots, k_{(j)}$. 字典 $D=[\theta_{(1)}(\hat{a}_1) \quad \theta_{(2)}(\hat{a}_1) \dots \theta_{(j)}(\hat{a}_1)]$

如果 $l=J$, 停止迭代. 否则, 转 5).

5) 使用 OMP 算法计算 Y 在字典 D 下的稀疏表示矩阵 X , 令 X 满足:

$$\min_{D, X} \{\|Y - DX\|_F^2, \|x_i\|_0 \leq T_0\} \quad (2)$$

式中, x_i 为 X 的第 i 列, T_0 为稀疏度.

6) 计算残差 $E=Y-DX, e_i$ 为 E 的第 i 列, 转 3).

3 实验结果

本文选用 MIT 的人脸库, 该人脸图库包括 2 428 张人脸, 每张图像均为 19×19 . 实验中随机选择 50 张人脸构成测试集, 剩余的 2 378 张人脸构成训练集. 每张图像都变成 361×1 的列向量.

本文实验全部使用信噪比 (Signal to Noise Ratio, SNR) (单位: dB) 为衡量图像质量的指标

$$\text{SNB} = 10 \times \log_{10}(\|Y\|^2 / \|Y - Y'\|^2) \quad (3)$$

式中 Y 为原始图像, Y' 为经过处理的图像.

为了测试本文的新算法的稀疏表示效果, 使用

了正交匹配追踪算法对一组图像进行稀疏表示. 该算法能够有效地解决 0 范数的优化问题. 算法得到的稀疏图像的信噪比与 K-SVD 的稀疏表示的信噪比如表 1, 利用稀疏表示系数恢复出来的图像实例如图 5, 其中 SP(Sparseness) 为稀疏度.

本算法训练得到的字典不仅可以用于稀疏表示, 还可以用于图像去噪, 当图像加入随机高斯噪声, 信噪比为 23.809 7 dB 时, 去噪效果如图 6.

该稀疏表示用于压缩采样恢复也可以取得较好的效果. 对 19×19 的原始图像随机采 100 个点, 即取 Φ 为随机采样阵, 采样率 $\text{rate} = 100/361 = 27.7\%$. 压缩采样结果如图 7.

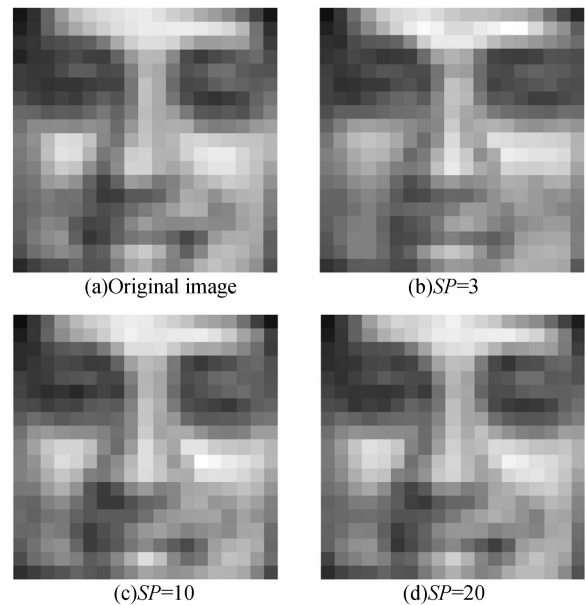


图 5 新算法的恢复效果图

Fig. 5 The reconstruction result of the novel algorithm

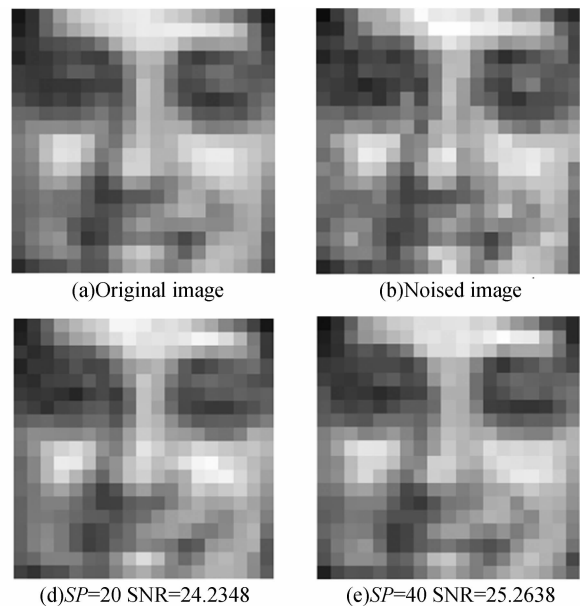


图 6 稀疏表示的去噪效果

Fig. 6 Experimental results of the denoising

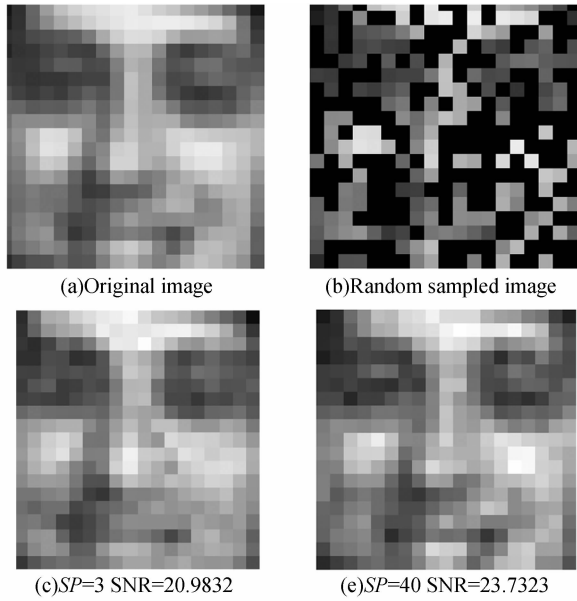


图 7 压缩采样的恢复效果

Fig. 7 Experimental results of the compressive sampling

表 1 新算法与 K-SVD 的对比实验结果

Table 1 The result comparison between the novel algorithm and K-SVD

| SP | Dictionary size | Average SNR/dB | |
|----|-----------------|----------------|-----------------|
| | | K-SVD | Novel algorithm |
| 3 | 361×900 | 16.798 9 | 17.367 9 |
| 10 | 361×900 | 19.035 6 | 20.308 3 |
| 20 | 361×900 | 20.745 9 | 22.433 6 |
| 3 | 361×1 200 | 17.077 7 | 17.397 8 |
| 10 | 361×1 200 | 19.384 9 | 20.327 3 |
| 20 | 361×1 200 | 21.164 6 | 22.589 2 |

4 算法复杂度分析

假设第一层原子的个数为 $k_{(1)}$, 第二层原子的个数为 $k_{(2)}$. 假设字典的大小为 1 200, 那么 $k_{(1)} + k_{(2)} = 1 200$.

为了加快运算速度, 流程中的 OMP 普遍使用 Batch-OMP 快速算法^[9]. 假设样本数为 N , 样本维数为 n , 稀疏度为 T_0 .

第 J 次迭代, 字典大小为 $k_{(J)}$, 假设 K-均值聚类时聚类中心平均更新 iternew 次稳定, 其中 iternew 为更新次数. 迭代一次的乘法运算量为 $Nnk_{(J)}$. 对任意的 j 和 $s, j \neq s$. 计算 $\|e_j - e_s\|_2^2$, 计算乘法量为 $k_{(J)}(k_{(J)} - 1)n/2$

对任意的 N 个样本全部使用 OMP 求出稀疏度为 T_0 的稀疏表示, 运算量为^[9]

$$N(2nk_{(J)} + T_0^2 k_{(J)} + 3T_0 k_{(J)} + k_{(J)}^3)$$

计算残差 $E = Y - DX$ 的运算量为 $Nnk_{(J)}$.

因此第 J 次迭代的总计算量为

$$T_{(J)} = (\text{iternew} + 1) * Nnk_{(J)} + k_{(J)}(k_{(J)} - 1)n/2 + N(2nk_{(J)} + T_0^2 k_{(J)} + 3T_0 k_{(J)} + T_0^3) \quad (4)$$

因此当原子层数为 2 时, 总运算复杂度为 T_{Allnew}

$T_{\text{Allnew}} = T_{(1)} + T_{(2)}$ 实验采用的 CPU 为 1.8 GHz, 内存为 512 M. 新算法与 K-SVD 的运行时间比较见表 2.

表 2 新算法和 K-SVD 运行时间比较

Table 2 Run time comparison between the novel algorithm and K-SVD

| Dictionary size | Run time/min | |
|-----------------|--------------|-----------------|
| | K-SVD | Novel algorithm |
| 361×900 | 270.32 | 10.54 |
| 361×1 200 | 480.45 | 18.75 |

5 结论

本文发明了一种新的过完备字典训练方法, 证明了该方法的收敛性, 并分析了该算法的计算复杂度. 将该方法用于压缩感知, 去噪等方面取得了比传统的 K-SVD 更好的效果.

基于训练的过完备字典以其自适应形和灵活性等各方面优势较以往基于分析的字典有更加广阔的应用前景. 但是有两个问题有待解决: 第一, 本文所提及的算法的稀疏表示结果信噪比仍然有限; 第二, 训练样本数、字典大小和图像大小三者之间的关系还没有明确证明.

参考文献

- [1] CANDES E J. Compressive sampling [C]//Proceedings of International Congress of Mathematicians. Madrid, Spain. European Mathematical Society Publishing House, 2006: 1433-1452.
- [2] BARANIUK R. Compressive sensing [J]. *IEEE Signal Processing Magazine*, 2007, 24(4): 118-120.
- [3] MAIRAL J, ELAD M, SAPIRO G. Sparse representation for color image restoration [J]. *IEEE Transactions on Image Processing*, 2008, 17(1): 53-69.
- [4] AHARON M, ELAD M, BRUCKSTEIN A. K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation [J]. *IEEE Transactions on Signal Processing*, 2006, 54(11): 4311-4322.
- [5] HE Z, CICHOCKI A. K-EVD clustering and its applications to sparse component analysis [C]. Independent Component Analysis and Blind Signal Separation, Charleston, SC, USA. LNCS, 2006, 3889: 90-97.
- [6] XIE Z, FENG J. KFCE: a dictionary generation algorithm for sparse representation [J]. *Signal Processing*, 2009, 89(10): 2072-2077.
- [7] TROPP J A, GIBERT A C. Signal recovery from random measurements via orthogonal matching pursuit [J]. *IEEE Transactions on Information Theory*, 2007, 53(12): 4655-4666.
- [8] KUNIS S, RAUHUT H. Random sampling of sparse trigonometric polynomials, II. Orthogonal matching pursuit versus basis pursuit [J]. *Foundations of Computational Mathematics*, 2008, 8(6): 737-763.
- [9] RUBINSTEIN R, ZIBULEVSKY M, ELAD M. Efficient implementation of the K-SVD algorithm using batch orthogonal matching pursuit [J/OL]. Technical Report-CS Technion, 2008 [2008-03-15]. http://www.cs.technion.ac.il/users/wwwwb/cgi-bin/tr-get.cgi/2008/CS/CS-2008-08_revised.pdf.

Self-adaptive Image Sparse Representation Algorithm Based on Clustering and Its Application

XU Jian^{1,2}, CHANG Zhi-guo^{3,4}

(1 *School of Communication and Information Engineering, Xi'an University of Posts & Telecommunications, Xi'an 710121, China*)

(2 *School of Electronic and information Engineering, Xi'an Jiaotong University, Xi'an 710049, China*)

(3 *College of Information and Engineering, Chang'an University, Xi'an 710064, China*)

(4 *Shaanxi Engineering and Technique Research Center for Road and Traffic Detection, Xi'an 710064, China*)

Abstract: A dictionary training algorithm was proposed for sparse representation of images and its convergence was proved. The geometrical explanation of the algorithm is to approximate the hyperspherical cap with least hyperplanes. The algorithm clustered the error vectors of each step, and signed the cluster center as new atoms which made the dictionary more suitable for sparse representation of samples. Compared with the traditional algorithm, the new one has higher adaptability, lower requirement of sample number and dictionary size, higher convergence rate, and lower complexity. Finally, the experiment of compressive sensing and denoising demonstrates that dictionary training by this algorithm has good effect.

Key words: Sparse representation; Clustering; compressive sensing; Dictionary; Atom; Sparseness