

一种基于超完备字典学习的图像去噪方法

蔡泽民^{1,3}, 赖剑煌^{2,3}

(1. 中山大学数学与计算科学学院, 广东广州 510275; 2. 中山大学信息科学与技术学院, 广东广州 510275;
3. 广东省信息安全技术重点实验室, 广东广州 510275)

摘要: 基于超完备字典的图像稀疏表示因其具有稀疏性、特征保持性、可分性等特点而被广泛应用于图像处理. 本文提出一种超完备字典学习算法并应用于图像去噪. 将字典学习等价于一个二次规划问题, 并提出适合于大规模运算的投影梯度算法. 学习所得字典能有效描述图像特征. 基于此超完备学习字典, 获得图像的稀疏表示, 并恢复原始图像. 实验结果表明, 与小波类去噪方法相比, 本文的学习算法能更好地去除图像噪声, 保留图像细节信息, 获得更高的 PSNR 值.

关键词: 稀疏表示; 基追踪; 匹配追踪; 字典学习; 二次规划

中图分类号: TN911.73 **文献标识码:** A **文章编号:** 0372-2112 (2009) 02-0347-04

An Over-complete Learned Dictionary-Based Image De-noising Method

CAI Ze min^{1,3}, LAI Jian huang^{2,3}

(1. School of Mathematics and Computational Science, Sun Yat-sen University, Guangzhou, Guangdong 510275, China;
2. School of Information Science and Technology, Sun Yat-sen University, Guangzhou, Guangdong 510275, China;
3. Guangdong Province Key Laboratory of Information Security, Guangzhou, Guangdong 510275, China)

Abstract: Images' sparse representations over over-complete dictionaries have a wide application in image processing due to the properties of sparsity, integrity and separability. This paper proposes a dictionary learning algorithm which is applied to image denoising. The dictionary learning problem is expressed as a box constrained quadratic program and a fast projected gradient method is introduced to solve it. The learned dictionary describes the image content effectively. Experimental results show that: in comparison with the wavelet based denoising methods, our learning based algorithm has better denoising ability, keep more detail image information and improve the peak signal to noise ratio.

Key words: sparse representation; basis pursuit; matching pursuit; dictionary learning; quadratic program

1 引言

在数字图像获取和传输的过程中, 往往会使图像感染上噪声. 因此图像去噪成为图像处理应用的一个重要环节. 图像去噪问题可以模型化为:

$$I_{noisy} = I + v \quad (1)$$

其中 I_{noisy} 为含噪图像, v 为噪声项, I 为原始图像.

为了能更好地去除图像中的噪声, 并保持图像的原有质量. 我们认为, 对于图像数据的表示, 必须同时具备如下三个属性:

(1) 稀疏性: 仅采用少数的表示系数, 便能够获得图像数据的高质量恢复.

(2) 特征保持性: 图像的特征, 如边缘等, 在图像数据的表示中, 能够被保留下来.

(3) 可分性: 即具备将图像有用信息与噪声数据区分开的能力.

目前许多图像去噪方法根据图像频谱分布规律, 从频率上将图像中的有用信息与噪声分开, 例如小波方法去噪^[1-3]. 这种方法一般认为噪声能量集中于高频, 而图像有用信息的频谱则分布于一个有限区间. 然而, 在许多情况下, 图像有用信息中也有高频成分, 例如图像边缘, 所以采用这种方法在去除噪声的同时也损失了部分有用信息, 即缺乏特征保持性; 而噪声在低频部分也有一定的分量, 简单地滤除高频成分无法去除这部分噪声分量, 即没有有效的可分性.

基于超完备字典的图像稀疏表示方法能够同时满足这三个属性^[4]. 字典中的每个原子代表原始图像中的一个结构原型, 因此其在字典上具有稀疏表示, 而噪声

分量在该字典上不具备稀疏表示. 因此我们可以根据数据是否在字典上具有稀疏表示, 将原始图像与噪声区分开, 达到去噪的目的.

设计合适的字典是图像稀疏表示的一个关键问题. 本文提出一种超完备字典学习算法, 将字典训练问题转化为一个带边界约束二次规划问题, 并研究适合于大规模运算的解法, 使其适用于具有高维特性的图像处理问题. 训练得到的字典具有非负性的特征, 与自然图像数据相适应. 实验结果表明, 基于本文算法训练得到的字典具有很好的去噪能力, 能够更好的保持图像的特征, 去噪后图像具有更高的峰值信噪比.

2 超完备字典学习新算法

首先从含噪图像出发, 从中随机提取 K 张子图像, 构成训练样本集 $\{s_i\}_{i=1}^K$. 我们的目标是学习出一个包含样本结构, 且每个样本 s_i 在其上具有稀疏表示的字典 D , 并得到它们的表示系数 α_i , 即

$$\min_{D, \alpha} \frac{1}{2} \|S - D\alpha\|_F^2 + \sum_{i=1}^K \lambda_i \|\alpha_i\|_0 \quad (2)$$

其中矩阵 $S = (s_1, s_2, \dots, s_K)$. 采用迭代的方法. 首先给定字典 D , 求解所有训练样本 s_i 在 D 上的稀疏表示 α_i , 得到稀疏表示矩阵 $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_K)$; 然后根据 α 更新字典 D . 即每次迭代包含如下两个步骤:

2.1 固定字典 D , 求 α

在这一阶段, 我们假设字典 D 是固定的, 求解各训练样本在 D 下的表示系数, 即:

$$\min_{\alpha_i} \|s_i - D\alpha_i\|_2^2 + \lambda_i \|\alpha_i\|_0, \quad \text{其中 } i = 1, 2, \dots, K \quad (3)$$

利用匹配追踪算法^[5], 可求解问题(3).

当求得所有训练样本 s_i 的稀疏表示向量 α_i , 则获得稀疏表示矩阵 $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_K)$.

2.2 固定 α , 更新字典 D

在这一阶段, 稀疏表示矩阵 α 是已知的, 问题(2)可忽略后面各项, 得到:

$$\min_D \frac{1}{2} \|S - D\alpha\|_F^2 \quad (4)$$

式(4)等价于一个带边界约束的二次规划问题. 二次规划的标准形式为^[6]:

$$\begin{aligned} \min_x f(x) &= \frac{1}{2} x^T G x + d^T x \\ \text{s. t. } a_i^T x &= b_i, \quad i \in \xi \\ a_i^T x &\geq b_i, \quad i \in \vartheta \end{aligned} \quad (5)$$

其中 G 为对称矩阵, ξ 和 ϑ 为下标集, d 、 x 和 $\{a_i\}$ 、 $\{b_i\}$, $i \in \xi \cup \vartheta$ 是列向量. 下面我们分析字典更新问题(4)与二次规划问题(5)之间的等价性. 首先给出相伴向量的定义 1 以及命题 1.

定义 1(相伴向量) 将矩阵 $A = [a_{ij}] \in R^{m \times n}$ 的元

素按列的顺序排列成列向量 $\text{vec}(A) = (a_{11}, a_{21}, a_{m1}, a_{12}, \dots, a_{m2}, \dots, a_{1n}, \dots, a_{mn})^T$, 则 $\text{vec}(A)$ 称为矩阵 A 的相伴向量.

命题 1 若乘积矩阵 AB 是一个方阵, 则有:

$$\text{tr}(AB) = (\text{vec}(A^T))^T (\text{vec}(B))$$

$$\text{vec}(AB) = (I \otimes A) (\text{vec}(B))$$

其中 I 为单位阵, \otimes 为矩阵的 Kronecker 乘积.

命题 1 的证明是简单的. 根据这一命题以及 F 范数的性质, 我们有:

$$\begin{aligned} \|S - D\alpha\|_F^2 &= \text{tr}[(S - D\alpha)^T (S - D\alpha)] \\ &= \text{tr}(D\alpha\alpha^T D^T) - 2\text{tr}(S\alpha^T D^T) + \text{tr}(SS^T) \\ &= (\text{vec}(D^T))^T (I \otimes \alpha\alpha^T) (\text{vec}(D^T)) \\ &\quad - 2(\text{vec}(\alpha S^T))^T (\text{vec}(D^T)) + \text{tr}(SS^T) \end{aligned} \quad (6)$$

因为训练样本集是已知的, 可忽略最后一项 $\text{tr}(SS^T)$, 再根据自然图像数据的特点, 可对字典中的元素加入非负性约束, 从而字典更新问题(4)转化为:

$$\begin{aligned} \min_{\text{vec}(D^T)} \frac{1}{2} (\text{vec}(D^T))^T (I \otimes \alpha\alpha^T) \text{vec}(D^T) \\ - (\text{vec}(\alpha S^T))^T \text{vec}(D^T) \\ \text{s. t. } \text{vec}(D^T) \geq 0 \end{aligned} \quad (7)$$

通过比较最优化问题(7)和(5), 可知问题(4)是二次规划的一个特例.

本文采用投影梯度类方法求解问题(7). 首先, 定义可行集为 $\Omega = \{x \in R^{nV} | x \geq 0\}$. 假设当前的可行解为 $\text{vec}(D^T)_k$, 通过迭代式 $\text{vec}(D^T)_{k+1} = \text{vec}(D^T)_k + e_k g_k$ 计算 $\text{vec}(D^T)_{k+1}$, 并将其投影到非负平面 $(\text{vec}(D^T)_{k+1})_i = \max(0, (\text{vec}(D^T)_{k+1})_i)$, ($i = 1, 2, \dots, nV$), 记这一操作为 $P_\Omega(\text{vec}(D^T))$. 定义 g_k 为 $\text{vec}(\alpha S^T) - (I \otimes \alpha\alpha^T) \text{vec}(D^T)$, $e_k > 0$ 为搜索步长. 根据 Barzilai 和 Borwein 等人的研究^[7], 我们初始化步长为

$$e_k^{(0)} = \frac{r_{k-1}^T y_{k-1}}{y_{k-1}^T y_{k-1}} \quad (8)$$

其中 $r_{k-1} = \text{vec}(D^T)_k - \text{vec}(D^T)_{k-1}$, $y_{k-1} = g_k - g_{k-1}$. 为了能更快收敛到最优解, 步长 e_k 必须满足 Armijo 条件^[6]:

$$f(x_k + e_k g_k) \leq f(x_k) + \alpha e_k \nabla f_k^T g_k \quad (9)$$

采用回溯线性搜索方法^[6], 我们能获得满足 Armijo 条件的步长 e_k . 基于回溯线性搜索策略的投影梯度算法收敛速度快. 算法只须经过有限次迭代, 便可获得问题(7)的解, 从而完成字典的更新, 本文所给出的实验均未超过 30 次. 通过 L 次更新之后, 便可得到训练字典 D .

3 基于稀疏表示的图像去噪框架

这一节我们讨论如何将训练得到的字典 D 应用于图像去噪^[4]. 首先从输入图像 I_{noisy} 中提取出 K' 张子图

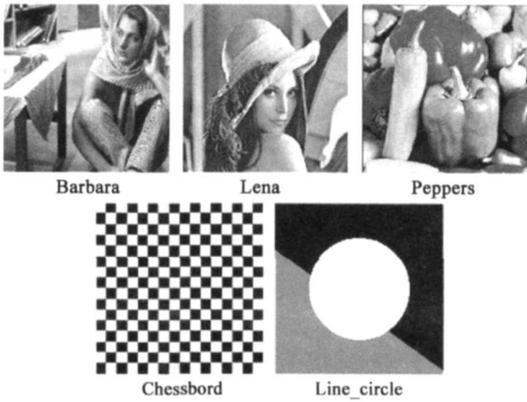


图1 原始图像(大小为512×512)

像,并按列重排为列向量.考虑如下的最优化问题:

$$\begin{aligned} \{\hat{\alpha}_j, \hat{I}\} = \arg \min_{\alpha_j, I} \lambda \|I - I_{noisy}\|_2^2 \\ + \sum_{i,j} \mu_j \|\alpha_j\|_0 + \sum_{i,j} \|D\alpha_j - R_j I\|_2^2 \end{aligned} \quad (10)$$

其中第一项衡量的是含噪图像 I_{noisy} 与原始图像 I 之间的总体相似程度;第二项是稀疏性约束;第三项 $R_j I$ 中表示第 j 张子图像, R_j 是用于提取子图的矩阵, $D\alpha_j$ 是重建得到的子图,我们希望它们之间的误差小.

给定训练所得字典 D ,先求解每张子图在其上的稀疏表示 $\hat{\alpha}_j$,即对于每一张子图,求解最优化问题:

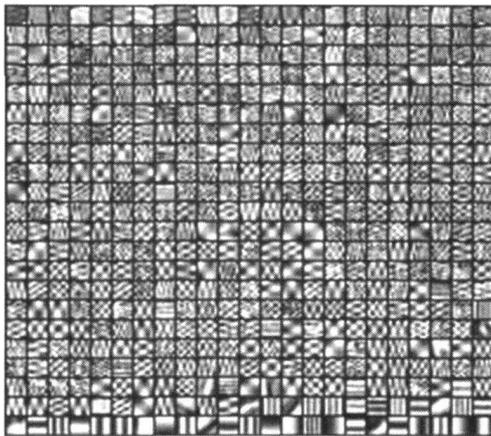


图2 Lena图字典

$$\hat{\alpha}_j = \arg \min_{\alpha} \mu_j \|\alpha\|_0 + \|D\alpha - R_j I\|_2^2 \quad (11)$$

匹配追踪算法^[5]可用于求解这一问题,得到每张子图的表示系数 $\hat{\alpha}_j$.

当求得所有子图的稀疏表示,则问题(10)转化为:

$$\hat{I} = \arg \min_I \lambda \|I - I_{noisy}\|_2^2 + \sum_{i,j} \|D\hat{\alpha}_j - R_j I\|_2^2 \quad (12)$$

这是简单的二次项,其解为:

$$\hat{I} = (E + \sum_{i,j} R_j^T R_j)^{-1} (Y_{noisy} + \sum_{i,j} R_j^T D \hat{\alpha}_j) \quad (13)$$

其中 E 为单位阵, \hat{I} 即为所求的去噪后图像.

4 实验结果与分析

实验的目的是检验算法对于加性高斯白噪声的去

噪能力,并与小波类去噪方法进行比较.实验数据包含三张自然图像和两张合成图像,如图1所示.

我们从含噪图像中提取出6000张子图像用于训练字典,每张子图的大小为 8×8 ,所训练出的字典包含256个原子,即字典 D 的大小为 64×256 冗余度为4.图2是利用本文算法训练得到的Lena图字典.将训练所得字典应用于去噪框架,首先从 I_{noisy} 图像中随机提取出子图像,计算它们在字典 D 上的稀疏表示.进而利用重构式(13)得到估计图像 \hat{I} (λ 取 $30/\sigma$).为了能与小波类方法进行比较,我们采用目前在小波域上得到较好效果的BLS-GSM方法^[2],同时也与传统的软阈值方法进行对比,比较结果见表1.



图3 噪声 $\sigma=20$ 的Barbara图像去噪后结果比较

表1 基于学习的方法与小波类方法的比较结果

图像	噪声标准差 σ	PSNR 值 (dB)			
		含噪图像	小波软阈值方法	BLS-GSM	本文学习算法
Barbara	10	28.14	29.82	33.28	34.52
	20	22.11	26.24	29.22	30.95
	30	18.57	24.38	26.94	28.66
	50	14.16	22.27	24.53	25.56
Lena	10	28.11	31.89	35.10	35.29
	20	22.12	28.59	32.12	32.27
	30	18.61	26.65	30.25	30.22
Peppers	10	28.14	24.09	28.12	27.81
	20	22.08	29.66	34.72	34.84
	30	18.58	27.35	32.58	32.65
Chessboard	50	14.14	24.57	29.87	29.43
	10	28.14	30.33	39.93	41.74
	20	22.09	25.81	34.74	38.15
	30	18.58	23.50	31.85	35.13
Line_circle	50	14.14	20.89	28.14	31.30
	10	28.12	35.88	43.27	44.03
	20	22.12	31.53	38.24	39.54
	30	18.59	29.00	35.73	36.75
Line_circle	50	14.13	25.89	32.78	32.78

从表1可以看到,本文基于学习的方法具有更强的去噪能力,特别是对纹理信息比较多的图像处理效果更好(如 Barbara 图和 Chessboard 图),且能保留图像的细节信息,具有更高的峰值信噪比.在噪声标准差小于 $\sigma = 50$ 的情况下,基于学习的方法降噪之后 PSNR 平均值为 34.00dB,比基于小波软阈值方法和 BLS-GSM 方法分别高出 6.40dB 和 1.02dB.需要指出的是,采用基于学习的方法包含了字典设计(训练)和稀疏表示求解两个过程,而小波类方法采用的是固定的小波字典,因此基于学习的方法计算量上比小波类方法要大.图3给出了图像 Barbara 在 $\sigma = 20$ 时,采用各种方法的不同去噪效果.

5 结束语

本文提出一种基于超完备字典学习的图像去噪方法.通过学习获得一个包含图像信息的超完备字典,并求解各子图像在此字典上的稀疏表示,通过重建恢复图像.实验结果表明,与传统的小波类去噪方法相比,基于学习的方法能更好的去除图像中的高斯白噪声,保留图像的细节信息,提高去噪图像的 PSNR 值,且具有更好的视觉效果.

参考文献:

- [1] D L Donoho. De noising by soft thresholding[J]. IEEE Trans on Information Theory, 1995, 41(3): 613- 627.
- [2] J Portilla, V Strela, et al. Image de noising using scale mixtures of Gaussians in the wavelet domain[J]. IEEE Trans on Image Processing, 2003, 12(11): 1338- 1351.
- [3] 杨晓慧,焦李成,等.基于第二代 bandelets 的图像去噪[J]. 电子学报,2006, 34(11): 2063- 2067.
Yang Xiaohui, Jiao Licheng, et al. Second generation bandelets

based image denoising[J]. Acta Electronica Sinica, 2006, 34 (11): 2063- 2067. (in Chinese)

- [4] M Elad, M Aharon. Image denoising via sparse and redundant representation over learned dictionaries[J]. IEEE Trans on Image Processing, 2006, 15(12): 3736- 3745.
- [5] S G Mallat, Z Zhang. Matching pursuit with time frequency dictionaries[J]. IEEE Trans on Signal Processing, 1993, 41 (12): 3397- 3415.
- [6] J Nocedal, S J Wright. Numerical Optimization[M]. New York: Springer Verlag, 2006.
- [7] J Barzilai, J Borwein. Two point step size gradient methods [J]. IMA Journal of Numerical Analysis, 1988, 8(1): 141- 148.

作者简介:



蔡泽民 男,1980 年生于广东汕头.博士研究生.2000 年到 2004 年在广州中山大学攻读信息与计算科学专业.研究兴趣包括数字图像处理,模式识别,机器学习与计算机视觉.

Email: caizemin@mail2.sysu.edu.cn



赖剑煌 男,1964 年生于海南东方县,中山大学信息科学与技术学院教授、博士生导师.现任广东省图像图形学会副理事长、秘书长,中国图象图形学会常务理事.1989 年获中山大学应用数学硕士学位.1999 年获中山大学基础数学博士学位.主要研究方向为数字图像处理、模式识别、机器学习、小波分析及其应用.

通信作者: Email: sljh@mail.sysu.edu.cn