

DOI: 10.11992/tis.201805011

网络出版地址: <http://kns.cnki.net/kcms/detail/23.1538.TP.20190610.1055.002.html>

# 一种恢复图像自然色彩的重构方法

陶霖密, 袁春, 王浩达

(清华大学 计算机科学与技术系, 北京, 100084)

**摘要:** 现代数码相机是通过颜色过滤矩阵在每个像素位置采集一个颜色分量, 重构出全彩色数字图像。压缩感知理论证明了该重构是误差有界的, 但在实际应用时却隐含着一个问题: 重构图像所需的稀疏编码字典是从图像数据库学习出来的, 而目前数字图像都是重构出来的, 因此存在着从重构的图像学习字典去重构图像的循环悖论。针对这个问题, 提出并构建了新的完全采样彩色图像的 Sandwich 图像数据集, 打破了压缩感知理论在应用于图像重构时的循环悖论, 使得压缩感知方法能够真正地重建自然彩色图像。Sandwich 图像数据集的构建及其训练得到的字典可以应用于如图像超分辨率重构、去噪、修复等领域。深入的图像重建实验表明, 使用 sandwich 图像集训练的字典不论是字典原子特性还是由其重构得到的图像质量均好于基于传统数据集的结果。  
**关键词:** 颜色过滤矩阵; 压缩感知; 稀疏编码; 字典学习; 图像重构; 自然色彩  
**中图分类号:** TP18   **文献标志码:** A   **文章编号:** 1673-4785(2019)05-0877-05

中文引用格式: 陶霖密, 袁春, 王浩达. 一种恢复图像自然色彩的重构方法 [J]. 智能系统学报, 2019, 14(5): 877-881.

英文引用格式: TAO Linmi, YUAN Chun, WANG Haoda. Reconstruction method for restoring the natural color of an image[J]. CAAI transactions on intelligent systems, 2019, 14(5): 877-881.

## Reconstruction method for restoring the natural color of an image

TAO Linmi, YUAN Chun, WANG Haoda

(Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China)

**Abstract:** Currently, the digital camera captures one color component at each pixel location through a color filter array and reconstructs a full-color digital image. The compressed sensing theory has proven that this reconstruction is based on the error bound with sparse coding dictionary. However, in practice, there is a hidden problem: the sparse coding dictionary needed for image reconstruction is learned from the image database while the current digital images are reconstructed. Thus, there is a cyclic paradox of reconstructing images from the reconstructed image learning dictionary. To solve this problem, this study proposes and constructs a fully sampled color image dataset called Sandwich to break the cyclic paradox of compressed sensing theory in the application of image reconstruction, enabling the compressed sensing method to reconstruct truly natural color images. The dictionary trained from the constructed Sandwich dataset can be applied to several domains, such as image super-resolution reconstruction, denoising, and restoration. The in-depth image reconstruction experiments show that the dictionary trained using the Sandwich image dataset is better than the dictionary trained using the traditional dataset in terms of both dictionary atomic characteristics and the reconstructed image quality.

**Keywords:** color filter array; compressed sensing; sparse coding; dictionary learning; image reconstruction; natural color

在单传感器相机成像过程中, 彩色图像传感器只采集到彩色图像的三分之一的信号, 彩色图

像是基于该原始信号重构出来的, 该重构过程通常称为去马赛克。原始图像信号去马赛克过程的彩色图像的质量至关重要, 算法性能表现直接决定重构图像的质量高低。重构过程中引入的噪声

收稿日期: 2018-05-09. 网络出版日期: 2019-06-11.

基金项目: 国家自然科学基金项目 (61672017).

通信作者: 陶霖密. E-mail: [linmi@tsinghua.edu.cn](mailto:linmi@tsinghua.edu.cn).

可能造成最终图像的人工伪色彩、拉链效应、灰度边缘等问题,图像丰富的特征细节和自然的色彩过渡往往无法兼得。传统的去马赛克算法往往只利用信号在空间和频谱上的区域相关性<sup>[1]</sup>,速度较快而适应性较差。为了克服这一问题,基于压缩感知<sup>[2]</sup>、深度学习<sup>[3]</sup>的图像重构算法被提出并得到广泛应用。压缩感知与深度学习相比于预先定义的先验模型如小波变换等,针对特定类型的信号学习得到的字典集会带来更好的重构效果<sup>[4-7]</sup>。但问题是这两者均依赖于从已有的数据中学习。一般来说,学习所需的训练集来源有两类:1)原始的不完全图像信号;2)独立的图像数据集切分成图像块(patch)集合。本质上原始信号本身已经缺失不足以提供更多信息,使用一个独立但相关的图像集训练能得到更好的重构结果。目前广泛使用的是以柯达无损图像数据集为代表的经由硬插值算法得到的图像组成的数据集。这种图像由raw格式原始信号经由插值算法计算得到,本身引入了重构误差,并非是无损数据。例如:常用的柯达图像数据集,其饱和度低下,在色彩过渡处出现大量由于插值所造成的灰度边缘,与人眼感知的自然色彩世界有较大出入。这种训练集中天生存在的误差将会被引入字典中,最终影响重构的成像质量,从而造成图像重构悖论。

为了打破了深度学习、压缩感知等方法在应用时的循环悖论,使得这些方法能够真正地重建自然彩色图像,本文提出了一种新的恢复自然彩色图像的重构方法,通过挖掘图像信号在不同字典下的稀疏表示的差异来提升图像色彩表现。这种基于压缩感知的图像重构方法的核心贡献是Sandwich图像数据集,以及构建这种数据集的方式。通过这些图像数据集训练的字典,重构得到的图像有更好的边缘色彩表现,而且不出现伪色彩。实验证明,在峰值信噪比(PSNR)和视觉质量方面,该方法优于大多数传统的去马赛克算法,包括不限于使用基于Kodak数据集字典重构的图像和使用传统基于压缩感知方法重构得到的图像。

## 1 压缩感知和稀疏编码

压缩感知理论的基础是存在一种变换方法<sup>[2]</sup>,使得信号变换到某空间后是稀疏的,即很少系数不为零。现有的大量研究结果表明,基于和具体信号无关的固定的变换,比如小波变换,通常不

能使变换后的信号足够稀疏。针对特定类型的信号学习得到的稀疏变换会带来更好的稀疏性。

假设 $D \in \mathbb{R}_n \times K$ 为 $K$ 原子的超完备字典( $K > n$ ), $x \in \mathbb{R}_n$ 为能在字典 $D$ 下形成稀疏表示的信号。那么在 $\alpha_0 \in \mathbb{R}_K$ 的前提下,信号 $x$ 能被表示为 $x = D\alpha_0$ 。

基于压缩感知的图像去马赛克(Demosaicing)可被表示成如下问题:

$$y = Mx \quad (1)$$

式中: $x$ 表示未知的全色彩的原始图像全部信息; $M$ 表示捕获图像相机使用的已知的CFA(color filter array)模版; $y$ 为经过CFA获取到的图像信息。

图像重构的核心目标就是从测量得到的信号 $y$ 中恢复出原始的图像信息 $x$ 。压缩感知理论证明了原始信号 $x$ 是可求解的,只要满足以下条件:

1)在特定的字典 $D$ 中, $x$ 有一个足够稀疏的表示:

$$\gamma : x = D\gamma \quad (2)$$

2)其中投影矩阵 $MD$ 必须满足一些特定条件,比如有限等距性质(restricted isometry property, RIP性质)。

RIP性质保证了观测矩阵不会把两个不同的稀疏信号映射到同一个集合中(保证原空间到稀疏空间的一一映射关系),要求从观测矩阵中抽取的每 $M$ 个列向量构成的矩阵是非奇异的。在具体情况中一般采用固定大小的稀疏矩阵,往往通过构造恰当的测量矩阵使得满足RIP等约束条件。

稀疏编码(sparse coding)可被描述为

$$\hat{\gamma} = \arg \min_{\gamma} \|\gamma\|_p \text{ s.t. } \|x - D\gamma\|^2 < \varepsilon \quad (3)$$

式中: $x$ 是原始信号; $\hat{\gamma}$ 是需要得到的稀疏信号。

本文需要解决的问题:

- 1)  $\gamma$ 是否存在唯一解,如果有,如何得到;
- 2) 如何判断求得的 $\gamma$ 是最稀疏的。

**唯一性定理** 结合式(1)和式(3),可以得到:

$$\hat{\gamma} = \arg \min_{\gamma} \|\gamma\|_p \text{ s.t. } \|y - MD\gamma\|^2 < \varepsilon \quad (4)$$

式(4)也可以被正则化描述成:

$$\hat{\gamma} = \arg \min_{\gamma} \|\gamma - MD\gamma\|^2 + \lambda \|\gamma\|_p \quad (5)$$

式中 $y$ 为经观测得到的信号。一般地 $p$ 取值区间为 $[0, 1]$ 。当 $p=0$ 时,由于 $\ell_0$ 函数是不连续的,直接求解式(5)是一个NP难问题。因此采用一些近似算法,譬如OMP和其他典型的基于 $\ell_1$ 范式的近似算法,如LASSO和 $\ell_1$  Magic来求解。

在这个求解问题中,字典集与最终信号的稀疏编码算法同样重要,其中字典 $D$ 可以是典型的数学基,譬如小波、Garbor、傅里叶基等。然而这些数学基往往不能很好表达复杂的自然图像。

## 2 Sandwich 图像数据集

目前广泛的用于测试图像重建、压缩、字典学习等的图像数据集是 Kodak 无损图像数据集。Kodak 图像数据集中图像为 24 位全彩色 png 格式, 由 Eastman Kodak 公司开放无限制使用。这些图像采集来源已不可考, 由于其广泛被用于字典学习的数据来源和不同图像重构算法性能的评价, 有必要谨慎考虑。

如图 1 所示, Kodak 数据集中图像整体饱和度偏低<sup>[8]</sup>, 亮度较现代数字相片也较暗, 与人视觉感知到的图像色彩有较大的出入, 作为 benchmark 其不太贴合自然世界中图片的情况。除此之外, 良好的数据集图像在色彩转换的边缘处应当流畅自然, 然而 Kodak 图像数据中明显可见在色彩边缘处出现了大量的灰黑边。这种情况隐藏在字典学习的原始信息中, 会引发一系列的副作用, 从而导致最终重构的图像也加上了不自然的黑边, 称为拉链或锯齿效应。作为基准线测试数据集, Kodak 数据中仅有 24 幅图片信息, 不能满足图像重构多字典学习的测试要求。

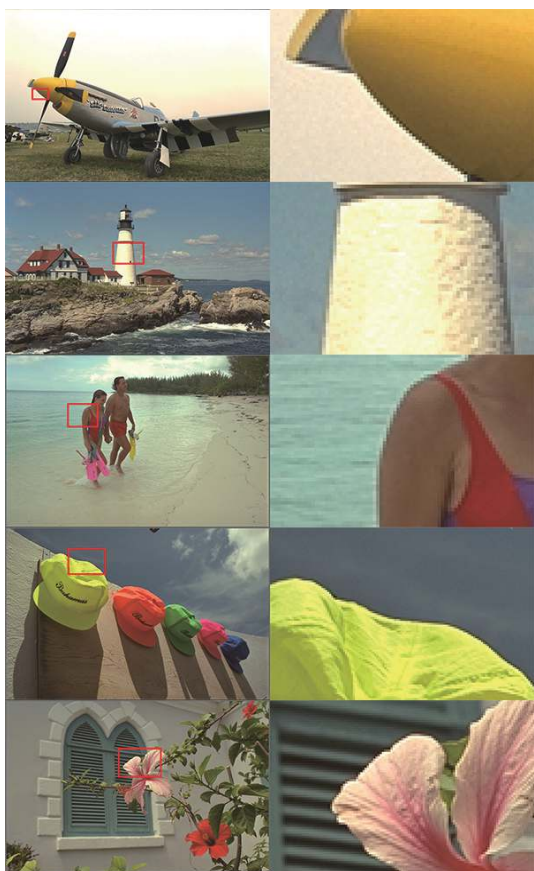


图 1 柯达图像数据集中存在的灰度边缘  
Fig. 1 Gray edges in the Kodak image database

传统使用的数字图像数据集基本都是由单感光器经过 CFA 模版采样<sup>[9-11]</sup>, 然后经过计算插值

得到的, 同样存在和 Kodak 数据集类似的问题, 为彩色图像的重建引入了多种噪声<sup>[12]</sup>。

为了在源头上避免多余的噪声和误差, 使用 Foveon X3 彩色图像传感器(图 2)采集了几百幅图像作为数据集(下文称 Sandwich 图像集)。Foveon X3 彩色图像传感器具有 3 层结构, 没有任何 CFA 模版, 每个像素都直接采样得到 RGB 三色的信息, 无需插值, 避免了插值、重构等计算带来的误差<sup>[12]</sup>。

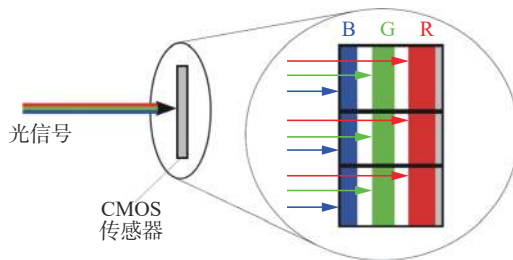


图 2 Foveon X3 图像传感器  
Fig. 2 Foveon X3 imaging sensor

Sandwich 图像数据集中的图像是 2 600×1 800 大小的高清彩色图像, 实际应用中可裁剪出任意大小的子图使用。在字典学习过程中有足够充分的图像信息可供使用, 从而使得描述图像时能得到更好的稀疏性。

图 3 展示出了 Sandwich 图像数据集中的一些样例。从图 3 中可明显看出, 图像边缘处色彩过渡自然流畅, 不会出现黑/灰色边缘, 即所谓的拉链效应。在不出现生硬的过渡黑边的情况下, 其图像结构信息依然保留完好, 视觉效果符合人眼视网膜感知自然世界特性。另一方面, 在不同的光源情况下体现了不同的亮度与饱和度。



图 3 Sandwich 图像数据集  
Fig. 3 Sandwich image database

图像数据集的评价: 图像数据集不能产生明显的黑边, 支持表达足够丰富的应用场景, 视觉效果评测是检查图像数据集质量的一个重要方面。但是, 视觉效果评测并不是一个客观评价标准, 人眼视觉的差异性导致了这个评测的结构难以精确地表达。从数据集学习得到的字典, 针对不同的图像处理应用领域, 譬如灰度图去噪、彩



色图像去噪、非均匀图像去噪修复等进行仿真实验,比较最终重构的图像 PSNR 等客观数据。

### 3 图像重构实验

基于传统、广泛使用的 Kodak 图像数据集,经由 K-SVD 算法学习到的字典<sup>[13-14]</sup>(下文统称 Kodak 字典)是完备的,每个字典原子 patch 大小为  $8 \times 8 \times 3$ , 总共有 192 个原子。这些原子中具有明显的黑灰边信息(见图 4),且彩色字典的边缘几乎全部都呈现灰色,可见其 RGB 三通道具有强烈的相关性。同时,字典的原子饱和度低下,不足以表达丰富的色彩信息。

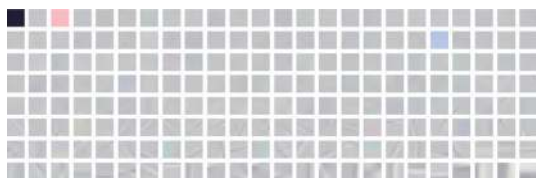


图 4 基于柯达图像学习的字典

Fig. 4 Dictionary learned from the Kodak image database

基于 Sandwich 数据集,经由 K-SVD 算法学习到的字典(下文统称 Sandwich 字典)与 Kodak 字典相同, Sandwich 字典原子 patch 大小为  $8 \times 8 \times 3$ , 总共有 192 个原子,也是完备的。Sandwich 字典中,其原子呈现出完全不同的特性(见图 5)。Sandwich 字典的色彩过渡自然,未出现不自然、生硬的黑边现象。由于足够丰富的场景和光源使得字典饱和度和亮度得到充分的体现。字典原子三通道间的差异较大,没有强制的正相关性,符合视觉感知的图像色彩特性。另外在这种字典中还具有良好的局部不变特征(如角点、斑点等)。

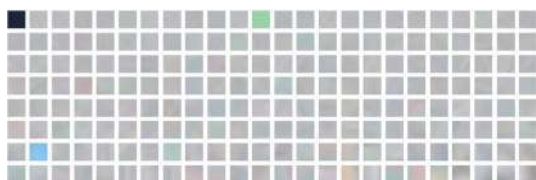


图 5 基于 Sandwich 图像学习的字典

Fig. 5 Dictionary learned from the Sandwich image database

本文采用相同的 Sandwich 图片,首先下采样得到彩色图像传感器所获取的三分之一图像信息,然后分别基于 Kodak 字典和 Sandwich 字典进行恢复重构。图 6 分别为从 Kodak 字典恢复的重构图像,和 Sandwich 恢复图像。图 6(a) 平白增加了原本不存在的黑边,该黑边信息从 Kodak 图像中引入。



(a) 基于 Sandwich 字典所重构的图像



(b) 基于 Kodak 字典所重构的图像

图 6 水果图图像重构实验对比

Fig. 6 Comparisons of the original and reconstructed fruit image

图 7 分别显示了从 Kodak 数据中学习到的字典造成了重构图像中原本不存在的杂色或伪色彩(图 7(a)),以及 Sandwich 字典重构的图像就不存在伪色彩(图 7(b)),其原因是 Kodak 数据集中的图像数据饱和度低、三通道间相关性强、缺失自然色彩信息。广泛使用的图像基准集 Kodak 图像数据集中在边缘处存在大量黑边,在压缩感知的字典学习过程中这种三通道高相关性天然存在的缺陷将会被引入到最终被重构的字典中。这种存在于字典之中的缺陷最终被引入到重构的图像中。最后针对应用不同图像数据集字典重构的图像比较 PSNR(peak signal to noise ratio, 峰值信噪比),表明基于新的 Sandwich 数据集所重构的图像具有更好的 PSNR(见表 1)。



(a) 基于 Sandwich 字典所重构的图像



(b) 基于 Kodak 字典所重构的图像

图 7 灰砖图图像重构实验对比

Fig. 7 Comparisons of the original and reconstructed brick image

表1 图像重构实验结果 PSNR 对比

Table 1 Comparison of the original and reconstructed of PSNR

图像	Kodak数据集	Sandwich数据集
水果图像	33.58	<b>37.39</b>
灰砖图像	42.12	<b>46.40</b>

## 4 结束语

具有正常色彩感知视觉的成年人,其视网膜中红、绿、蓝三色传感器在视网膜中随机分布。人脑感知自然世界的过程本身就是一个不完全信号恢复过程。但是在基于压缩感知的图像重构过程中存在的图像-重构-图像循环悖论表明,通过新的图像数据库可以打破该循环悖论,从而显著提高重构图像的质量。

本文从图像数据库这一全新的角度研究了图像重构问题,开发的 Sandwich 数据库是一种普适的图像数据库,可以广泛地用于各种图像重构、图像质量评估等研究。

## 参考文献:

- [1] KIMMEL R. Demosaicing: image reconstruction from color CCD samples[J]. *IEEE transactions on image processing*, 1999, 8(9): 1221–1228.
- [2] DONOHO D L. Compressed sensing[J]. *IEEE transactions on information theory*, 2006, 52(4): 1289–1306.
- [3] GHARBI M, CHAURASIA G, PARIS S, et al. Deep joint demosaicking and denoising[J]. *ACM transactions on graphics*, 2016, 35(6): 191.
- [4] ZHANG Meiqing, TAO Linmi. A patch aware multiple dictionary framework for demosaicing[C]//Proceedings of the 12th Asian Conference on Computer Vision. Singapore, Singapore, 2014: 236–251.
- [5] ZHANG Lei, WU Xiaolin, BUADES A, et al. Color demosaicking by local directional interpolation and nonlocal adaptive thresholding[J]. *Journal of electronic imaging*, 2011, 20(2): 023016.
- [6] MAIRAL J, ELAD M, SAPIRO G. Sparse representation for color image restoration[J]. *IEEE transactions on image processing*, 2008, 17(1): 53–69.
- [7] MOGHADAM A A, AGHAGOLZADEH M, KUMAR M, et al. Compressive framework for demosaicing of natural images[J]. *IEEE transactions on image processing*, 2013,

22(6): 2356–2371.

- [8] FRANZEN R. Kodak lossless true color image suite[EB/OL]. (1999-11-15)[2018-03-22]. <http://r0k.us/graphics/kodak>.
- [9] GUNTURK B K, GLOTZBACH J, ALTUNBASAK Y, et al. Demosaicking: color filter array interpolation[J]. *IEEE signal processing magazine*, 2005, 22(1): 44–54.
- [10] GUNTURK B K, ALTUNBASAK Y, MERSEREAU R M. Color plane interpolation using alternating projections [J]. *IEEE transactions on image processing*, 2002, 11(9): 997–1013.
- [11] LI Xin, GUNTURK B, ZHANG Lei. Image demosaicing: a systematic survey[C]//Proceedings of Visual Communications and Image Processing 2008. San Jose, United States, 2008: 68221J.
- [12] NAKAMURA J. Image sensors and signal processing for digital still cameras[M]. Boca Raton: CRC Press, 2005.
- [13] LOSSON O, MACAIRE L, YANG Y. Comparison of color demosaicing methods[J]. *Advances in imaging and electron physics*, 2010, 162: 173–265.
- [14] CANDÈS E J, ROMBERG J, TAO T. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information[J]. *IEEE transactions on information theory*, 2006, 52(2): 489–509.

## 作者简介:



陶霖密,男,1962年生,副教授,主要研究方向为人机交互、计算机视觉与模式识别。承担国家重点基金情感计算项目,以及与IBM、INTEL、SIEMENS 的国际合作基金项目等。发表学术论文多篇。



袁春,男,1969年生,副研究员,博士生导师,主要研究方向为机器学习、计算机视觉、视频分析与处理。先后负责和参与微软、欧盟、国家自然科学基金、863 和 973 等重要研究课题。取得专利 15 项,发表学术论文 70 余篇。



王浩达,男,1992年生,硕士研究生,主要研究方向为机器学习、计算机视觉。