

指导教师： 杨涛

提交时间： 2016.03.20

CVPR2015 Paper Translation

No : 01

姓名 : 李武松

学号 : 2013302551

班号 : 10011304

快速，灵活的卷积稀疏编码

Felix Heide
Stanford University, UBC

Wolfgang
Heidrich KAUST, UBC

Gordon Wetzstein
Stanford University

概要

稀疏编码在机器学习和计算机视觉方面已成为越来越重要的工具。图像特征能够被（机器）学习，随后图像特征被用于图像的分类和图像的重建。与此相反的方法是块拼接，卷积稀疏编码是基于整张图片的操作，从而实现无缝捕捉邻域的相关关系。在这篇文章中，我们将会提出一个全新的方法来解决卷积稀疏编码问题，我们的方法令人瞩目的是十分快捷而且能够找出比发展水平报告中更加优秀的解。此外，这篇文章中提出的方法是第一个能够允许附加特定边界条件和支持不完备数据条件下的特征学习和解决一般的重建问题。

1. 介绍

越来越多的计算机视觉工作依赖于自然界照片的统计分析，低层次的工作得益于好的前期工作，前期工作包括修复、去噪、去模糊、和超分辨率，而识别、分类和其他更高级别的工作经常使用已经掌握的从自然图像提取特征作为先验知识。在本文中，我们回顾了一种图像特征的无监督学习策略：卷积稀疏编码，卷积稀疏编码是在人类视觉感知领域建模被引入[18]，但是在最近已经被证明它在诸如低/中级的特征学习和低层次的重建问题等诸多计算机视觉问题中有着重要的应用[21, 7]。作为更加复杂层次的层次结构或者高层次计算机视觉网络

挑战中的一部分[13, 22, 23]，除了这些应用外，CSC在许多其他的重建任务和基于特征的方法也有广泛应用，应用包括去模糊、去噪、修复、分类、定位和跟踪。

基于块的学习和重建方法的流行与CSC密切相关[5, 16, 24]。基于块拼接的特征学习常常包含相同信号的移位版本，同时当我们将信号分成小块的时候就可能会失去潜在基础信号结构。一个更简洁漂亮的来模拟这些问题方式是使用稀疏分布的卷积特征的总和。卷积稀疏编码不仅在合理的时间内寻找一个解是十分具有挑战的，而且寻找一个好的局部最小值也是十分不容易的。

一般而言，卷积CSC是一个非凸问题，现有的许多方法的全局收敛性得不到保证。最近发现快速卷积稀疏编码有了开创性进展——通过CSC在频率域的特征学习能够被有效解决。格罗斯等人。率先[9]提出了一个一维的音频信号的频域方法，随后论证了用于二维图像数据有效频域方法。

这是使CSC变得实用的第一步，这些频率的方法在（图像的学习和重建中）引入边界伪影，作为内在的全局方法，使它不完整的数据进行工作变得十分困难。

在优化的基础[6, 2, 19, 1, 20]，我们提出了一种新的基于分裂的卷积稀疏编码方法，我们不仅要阐述我们使我们能够轻松地结合适当的边界条件和稀疏的观测的公式，而且还要展

示我们所提出比目前技术发展水平更快，收敛到更好的解决方案。特别是，我们作出以下贡献：

- 我们提出了一个灵活的卷积稀疏编码的公式，有有效解的分解目的为一系列简单的，凸函数。这个公式适用于正倒向的优化框架。

- 我们证明所提出的方法允许适当的边界条件是不牺牲性能，收敛速度比其他方法更快更好的解决方案，我们验证了后者使用几个低级别的重建问题。

- 我们显示了没有与 CSC 求解证明的不完全观测。

- 我们证明提出的求解已知的特征也同样有效，如高斯混合项或物理激励的卷积基。

2. 数学框架

传统上，卷积稀疏编码问题的形式表示的形式：

$$\begin{aligned} \operatorname{argmin}_{\mathbf{d}, \mathbf{z}} \quad & \frac{1}{2} \|\mathbf{x} - \sum_{k=1}^K \mathbf{d}_k * \mathbf{z}_k\|_2^2 + \beta \sum_{k=1}^K \|\mathbf{z}_k\|_1 \\ \text{subject to} \quad & \|\mathbf{d}_k\|_2^2 \leq 1 \quad \forall k \in \{1, \dots, K\}, \end{aligned} \quad (1)$$

在 ZK 稀疏特征映射的近似数据项 X 卷积时相应的滤波器固定空间的支持。这里 $\mathbf{x} \in \mathbb{R}^D$, $\mathbf{z}_k \in \mathbb{R}^D$ 是矢量图像， $\mathbf{D}_k \in \mathbb{R}^M$ 是矢量的二维滤波器， $K = 1 \dots K$ ，和 * 是 **二维卷积算子的内积在矢量化的输入**。虽然上述方程是严格唯一有效的一个单一的目标图像，它可以很容易地推广到多个图像 X。

最近 Bristow 等人。[3, 4] 显示利用 Parseval 定理求解方程能够使效率显著提高 (1)，即一个信号能量等效为一个常数，它的傅里叶变换。我们将在以下忽略这个常数。式 1 可改写 [3, 4, 14] 为

$$\begin{aligned} \operatorname{argmin}_{\mathbf{d}, \mathbf{z}} \quad & \frac{1}{2} \|\hat{\mathbf{x}} - \sum_{k=1}^K \hat{\mathbf{d}}_k \odot \hat{\mathbf{z}}_k\|_2^2 + \beta \sum_{k=1}^K \|\mathbf{t}_k\|_1 \\ \text{subject to} \quad & \|\mathbf{s}_k\|_2^2 \leq 1 \quad \forall k \in \{1, \dots, K\} \\ & \mathbf{s}_k = \mathbf{S} \Phi^T \hat{\mathbf{d}}_k \quad \forall k \in \{1, \dots, K\} \\ & \mathbf{t}_k = \mathbf{z}_k \quad \forall k \in \{1, \dots, K\}, \end{aligned} \quad (2)$$

这表明相对于计算繁杂的卷积，更为有效的方法是在频域中的算法。

在这里， $\hat{\cdot}$ 是一个信号的频率表示， \odot 是信号的分量形式， Φ 是离散傅里叶变换 (DFT) 矩阵， \mathbf{S} 是一个滤波器上的 (小) 空间支持

注意方程式是十分重要的。2 和 1 实际上只相当于圆形边界条件下的假设 [22]。kavukcuoglu 等人。

[13] 指出，边界条件是影响优化甚至圆卷积模型边界条件的一个重要障碍，因为像素接近边界，在一般情况下，被滤波器覆盖的像素比中心像素少。而启发式算法 [3] 也可以学习滤波器具有非常小的空间的支持，这种假设不一定成立的大滤波器或一般重建问题。我们提出以下，卷积稀疏编码的一般公式：

$$\begin{aligned} \operatorname{argmin}_{\mathbf{d}, \mathbf{z}} \quad & \frac{1}{2} \|\mathbf{x} - \mathbf{M} \sum_{k=1}^K \mathbf{d}_k * \mathbf{z}_k\|_2^2 + \beta \sum_{k=1}^K \|\mathbf{z}_k\|_1 \\ \text{subject to} \quad & \|\mathbf{d}_k\|_2^2 \leq 1 \quad \forall k \in \{1, \dots, K\}. \end{aligned} \quad (3)$$

在这里， \mathbf{M} 是一个对角矩阵或块，这样它将线性系统的形式 $(\mathbf{M} + \mathbf{I}) \mathbf{x} = \mathbf{b}$ 分成许多小的、独立的被有效解决的系统。例如，边界处理 \mathbf{M} 可以是二进制的对角矩阵，掩码的边界估计 $\mathbf{P}_K \mathbf{K} = \mathbf{1} \mathbf{D}_K * \mathbf{Z}_K$ 。这使得我们在边界地区无法使用滤波器，从而保持问题的卷积性质不需要循环边界或其他条件。此外，我们发现，我可以充分学习和基于不完整数据重建。

不幸的是，公式 3 不能直接解决“傅里叶变换”在文献中的讨论 (公

式 2)。在下面，我们得出一个公式，让我们不仅能够灵活有效求解方程 3，但我们也证明了我们的公式，解决传统的卷积稀疏编码问题（公式 1）的速度比以前的方法更快和收敛性更好。

2.1. 有效的目标分割

为了有效的求解公式 3。我们重新约束包括通过指标函数的很大的目的（ \bullet ）等，这凸集合的约束是 $C = \{V \mid \text{ksvk2 } 2 \leq 1\}$ 。这将产生以下无约束的目标：

$$\underset{\mathbf{d}, \mathbf{z}}{\operatorname{argmin}} \frac{1}{2} \|\mathbf{x} - \mathbf{M} \sum_{k=1}^K \mathbf{d}_k * \mathbf{z}_k\|_2^2 + \beta \sum_{k=1}^K \|\mathbf{z}_k\|_1 + \sum_{k=1}^K \operatorname{ind}_C(\mathbf{d}_k), \quad (4)$$

可表示为下列函数的总和

$$\underset{\mathbf{d}, \mathbf{z}}{\operatorname{argmin}} f_1(\mathbf{Dz}) + \sum_{k=1}^K (f_2(\mathbf{z}_k) + f_3(\mathbf{d}_k)), \quad \text{with} \quad (5)$$

$$f_1(\mathbf{v}) = \frac{1}{2} \|\mathbf{x} - \mathbf{Mv}\|_2^2, f_2(\mathbf{v}) = \beta \|\mathbf{v}\|_1, f_3(\mathbf{v}) = \operatorname{ind}_C(\mathbf{v}).$$

在这里， $\mathbf{Z} = [\mathbf{ZT1} \ \mathbf{ZT} \cdots \mathbf{K}]^T$ 和 $\mathbf{D} = [\mathbf{D1} \cdots \mathbf{DK}]$ 是一个级联的 Toeplitz 矩阵，每一个代表一个与各自 DK 卷积滤波器。方程（5）的总和组成的功能网络，这些都是简单的优化个别，而它们的和是具有挑战性的。下面的 [1]，我们 F1 与 M 包括因为将数据项分为两个不同的子问题的 M 和 D 分别但不共同。

2.2. 目标的概括

为了获得这一结果更直观地，我们认为一般目标服从（5）

$$\underset{\mathbf{z}}{\operatorname{argmin}} \sum_{i=1}^I f_i(\mathbf{K}_i \mathbf{z}), \quad (6)$$

\mathbf{K}_i : $\mathbb{R}^{b_i \times a_i}$ 是任意的矩阵，
 f_i : $\mathbb{R}^{b_i} \rightarrow \mathbb{R}$ 是封闭的，适当的，凸函数， $i \in \{1, \dots, I\}$ ，使得 $f_i(\mathbf{K}_i \mathbf{z})$:
 $\mathbb{R}^{a_i} \rightarrow \mathbb{R}$; I 是在求和函数个数。

公式 6 最近的工作是出于图像反

卷积 [10, 1]，其中有一个类似的目标，包括简单的凸函数的和。方程 6 中的问题可以改写为：

$$\underset{\mathbf{z}}{\operatorname{argmin}} \sum_{i=1}^I f_i(\mathbf{K}_i \mathbf{z}) = f(\mathbf{Kz}), \quad \text{with} \quad (7)$$

$$\mathbf{K} = \begin{bmatrix} \mathbf{K}_1 \\ \vdots \\ \mathbf{K}_I \end{bmatrix} \quad \text{and} \quad f(\mathbf{v}) = \sum_{i=1}^I f_i(\mathbf{v}_i),$$

在 \mathbf{v}_i 选择 i 支持 V 。使用的公式与堆叠矩阵 \mathbf{K} 对现有的优化框架允许我们映射方程（6），如图所示的 Chambolle 和 [10] Pock 的方法，对交替方向乘子法在 [1]，并在 [20] 的随机优化方法。

为清楚起见，我们描述了解决方程只有一个可能的方法（7）：交替方向乘子法，用来表示解决问题的形式：

$$\underset{\mathbf{z}}{\operatorname{argmin}} h(\mathbf{y}) + g(\mathbf{z}) \quad \text{subject to} \quad \mathbf{Ay} = \mathbf{z}. \quad (8)$$

有关算法的详细讨论可以在 [2] 中找到。使用标准的缩放的形式 [2] 和设置 $\mathbf{H}=\mathbf{0}$, $\mathbf{K}=\mathbf{K}$ 和 $\mathbf{G}=\mathbf{F}$ ，这看起来直观推导算法之前，收益率算法。1、解决方程（7）。我们观察到，由此产生的最小化成为可分离的所有的网络连接。

Algorithm 1 ADMM for a sum of functions in Eq. (7)

- 1: for $k = 1$ to V do
 - 2: $\mathbf{y}^{k+1} = \underset{\mathbf{y}}{\operatorname{argmin}} \|\mathbf{Ky} - \mathbf{z} + \lambda^k\|_2^2$
 - 3: $\mathbf{z}_i^{k+1} = \operatorname{prox}_{f_i}(\mathbf{K}_i \mathbf{y}_i^{k+1} + \lambda_i^k) \quad \forall i \in \{1, \dots, I\}$
 - 4: $\lambda^{k+1} = \lambda^k + (\mathbf{Ky}^{k+1} - \mathbf{z}^{k+1})$
 - 5: end for
-

因为在这个 CSC 公式， \mathbf{M} 包含于 $\mathbf{F1}$ ，我们可以解决在 ALG 2 线的更新。1 在傅里叶域。注意，分割 \mathbf{M} 与滤波器相结合，通过 $\mathbf{F1}$ 代入 \mathbf{DZ} (\mathbf{DZ}) 导致交替方向乘子法扩大标准

的应用；标准方法[2]

作为一个单独的操作。符号快捷 $\text{PROX} \cdot$ 是邻近算子在[19]。许多不同的近端运营商在文献中已知的[19]，制定我们的方法，使用这些运营商让我们能够轻松地应用已知的导子。虽然我们得到了一个解方程 7 使用拆解，这对于 Chambolle 和 Pock 的方法或随机优化[20]同样是可能的。

2.3. 针对特异 CSC 的子问题的交替方向乘子法

我们继续讨论一般 ADMM 算法在上一小节介绍了卷积稀疏编码的应用。从算法解决二次子。1 (2 号线) 给出

$$y_{\text{opt}} = \underset{y}{\text{argmin}} \|Ky - \tau\|_2^2 = (K^T K)^{-1} (K^T \tau) \quad (9)$$

在这里，我们 $\tau = Z - \lambda K$ 作为一个符号的快捷方式。这取决于我们解决的滤波器（即 $Y = D$ ）或为特征映射（即 $Y = Z$ ），我们得到

$$\begin{aligned} d_{\text{opt}} &= (Z^T Z + 2\mathbb{I})^{-1} (Z^T \tau_1 + \tau_2 + \tau_3) \quad \text{for } y = d \\ z_{\text{opt}} &= (D^T D + \mathbb{I})^{-1} (D^T \tau_1 + \tau_2) \quad \text{for } y = z \end{aligned} \quad (10)$$

在这里， Z 是一个级联的 Toeplitz 矩阵为各自的稀疏编码 τ 我选择了 Z_k 和第 i 个支持 τ 内德为式 (7)。操作符的共轭转置，用符号[14]。在这两种情况下，可以变量重排在方程的方程组 (10)，使 $(Z^T Z + 2\mathbb{I})$ 和 $(D^T D + \mathbb{I})$ 的块对角[3, 14]，使得反演的有效的在 $J \in \{1\} \dots D$ 不同的块并行化。逆可以有效每个块 J 使用 Woodbury 公式进行计算，公式如下：

$$\begin{aligned} (Z_j^T Z_j + 2\mathbb{I})^{-1} &= \frac{1}{2} \mathbb{I} - \frac{1}{2} Z_j^T (2\mathbb{I} + Z_j Z_j^T)^{-1} Z_j \\ (D_j^T D_j + \mathbb{I})^{-1} &= \mathbb{I} - \frac{D_j^T D_j}{1 + D_j D_j^T}, \end{aligned} \quad (11)$$

其中的第二个等式，因为在 D_j 是一个行向量。我们通过 Cholesky 分解计算 $(2\mathbb{I} + Z_j Z_j^T)^{-1}$ 。相反，在[4]直接反演（由于代码更新，这是在每一次迭代的方法做的），这导致了一个重要的缓存分解显著降低运行时间如下。

对算法中的邻近算子。1 (3 行) 是简单的派生和众所周知的文献[19]：

$$\begin{aligned} \text{prox}_{\theta f_1}(v) &= (\mathbb{I} + \theta M^T M)^{-1} (v + \theta M^T x) \quad \text{Quadratic} \\ \text{prox}_{\theta f_2}(v) &= \max \left(1 - \frac{\theta \beta}{|v|}, 0 \right) \odot v \quad \text{Shrinkage} \\ \text{prox}_{\theta f_3}(v) &= \begin{cases} \frac{Sv}{\|Sv\|_2} & : \|Sv\|_2 \geq 1 \\ Sv & : \text{else} \end{cases} \quad \text{Projection} \end{aligned}$$

$\text{prox}_{\theta f_1}$ 的称作逆 M 通常是一块对角矩阵（通常是对角线）

2.4. 交变曲面的问题

以上我们所描述的算法. 1，它可以解决双凸问题 (3) Z 或 D 时，各自变量 \mathbf{n} 固定。共同解决，我们遵循交流他们之间的标准方法，产生算法. 2。

Algorithm 2 CSC learning using coordinate descent

- 1: Algorithm penalty parameters: $\rho_d \in \mathbb{R}^+$, $\rho_z \in \mathbb{R}^+$
 - 2: Initialize variables: $d^0, z^0, \lambda_d^0, \lambda_z^0$
 - 3: **repeat** {Outer iterations}
 - 4: **Kernel update:** Solve Eq. (5) w.r.t. d :
 $d^i, \lambda_d^i \leftarrow \underset{d}{\text{argmin}} f_1(Zd) + \sum_{k=1}^K f_3(d_k)$ using Alg. 1 with $\rho = \rho_d, \lambda = \lambda_d^{i-1}$
 - 5: **Code update:** Solve Eq. (5) w.r.t. z :
 $z^i, \lambda_z^i \leftarrow \underset{z}{\text{argmin}} f_1(Dz) + \sum_{k=1}^K f_2(z_k)$ using Alg. 1 with $\rho = \rho_z, \lambda = \lambda_z^{i-1}$
 - 6: **until** No more progress in both directions.
-

有了这种交替的方法，我们已经构建了一个坐标下降的个人拉格朗日乘子与从先前的迭代中被初始化。在实践中，我们运行一个子程序 \mathbf{fi} 直到足够有效取得进展。的坐标下降步长内德的进步使得每个局部优化。使用一个

常数 P 次数为每一步给了我们足够高效性能好。我们停止该算法时，两者的优化，可以进一步降低的目标，一个局部最小值被发现。它也遵循我们的算法单调递减的目标，其迭代序列目的。 d_i, z_i

2.5. 实施细则

对于式 (3)，我们发现，参数 $\beta = 1$ 提供了计算稀疏性和数据本文测试结果之间的良好折衷。我们有非常惊讶， β 其他设置导致相似的结果。算法 2，我们选择了启发式值 $\rho_D = 1 / (100 \cdot \max(x))$ 和 $\rho_z = 1 / (10 \cdot \max(x))$ 。

3. 分析

3.1. 复杂性分析

本节分析了所提出的优化方法的复杂性，并与其他方法相比，理论运行时。 D 是在 X 一个图像的像素数， N 为图像的数目， k 是学习内核的数量， M 的滤波器支持的大小，和 P 内迭代算法中的步骤 (.2)，计算成本见表 1。我们观察到，布里斯托的方法具有显著的立即明显优于蔡勒等人。当钾。其主导成本是二维线性系统的反演。注意，在对比的空间方法，布里斯托的方法，也是我们的，是该方法的 n 滤波器尺寸 M 独立，我们考虑两种情况： $K > N$ 和 $K \leq N$ 。在 n 第一案 ($K > N$)，我们利用反演技巧解释在式 (11)。在这里，每个 D 矩阵 ZJ 是一个 $N \times K$ 矩阵。因此，利用公式 (11)，减少逆成本从 K^3 到 KN^2 。因为我们的高速缓存的分解，这个成本仅仅是一个局部迭代。对于其他的 ($P-1$) 迭代的 backsolves 成本只有 KN (而不是 K^2 的天真的反转)。在第二种情况下，当 $K \leq N$ ，我们每一次的迭代有 P D 矩阵的 Cholesky 分解为 K^3 的全部成本，但是所有其他 ($P-1$) 迭代，只有回来解决成本 K^2 。因此，通过缓存的分解，我们能够通过 $P-1$ 实现的加速线性系统 + $(P-1) / K$

在这种情况下。我们对 $P = 10$ 的设置，甚至中等数量的 $K = 100$ 滤波器导致加速的 $9 \times$ 。在下面，我们表明，不仅每一次迭代的复杂性降低，但收敛性提高。

Method	Cost (in flops)
Zeiler et al. [22]	$PN \cdot (\underbrace{KD}_{\text{Conjugate gradient}} + \underbrace{KDM}_{\text{Spatial convolutions}} + \underbrace{KD}_{\text{Shrinkage}})$
Bristow et al. [3, 4]	$PN \cdot (\underbrace{K^3D}_{\text{Linear systems}} + \underbrace{KD \log(D)}_{\text{FFTs}} + \underbrace{KD}_{\text{Shrinkage}})$
Ours ($K > N$)	$\underbrace{KN^2D + (P-1)KN}_{\text{Linear systems}} + \underbrace{PN \cdot (KD \log(D) + KD)}_{\text{FFTs}} + \underbrace{KD}_{\text{Shrinkage}}$
Ours ($K \leq N$)	$\underbrace{K^3D + (P-1)K^2D}_{\text{Linear systems}} + \underbrace{PN \cdot (KD \log(D) + KD)}_{\text{FFTs}} + \underbrace{KD}_{\text{Shrinkage}}$

表 1: 我们的方法和其他近似的方法的代价。

3.2. 收敛

对于 2 个不同大小的数据集，我们绘制的经验收敛的算法，并将其与图 1 中的数据状态进行比较。在这两种情况下，我们学习 $K = 100$ 滤波器。数据集是水果集 [22] 与 $N = 10$ 的图像。在这种情况下，我们有 13 个迭代算法收敛，而 [4, 3] 有一个缓慢下降的目标，并充分融合后约 300 次迭代。为了能够比较客观的价值观，所有比较的方法在这里实施循环卷积，边缘逐渐变细的应用。

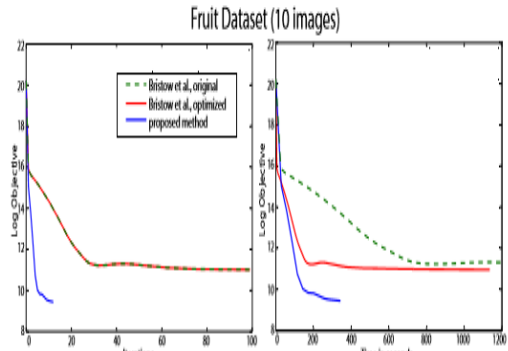


图 1: 2 个数据集的收敛性 (左 10 个图像，右 100 个) 该算法收敛到一个更好的解决方案，在更短的时间比竞争的方法。

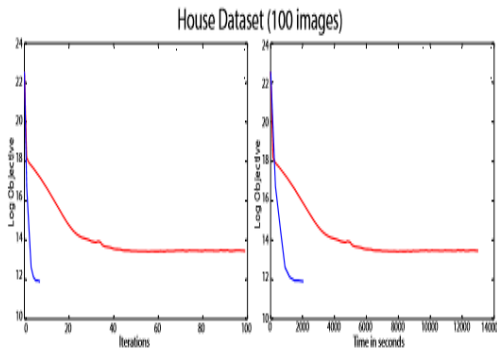


图 2: 可视化滤波器从水果集 1、7、8 和 13 次迭代后。进化是从随机分布的高斯团块样结构和最终收敛到 \mathbf{f}_i , 承担相似的伽柏滤波器。

我们也情节的收敛在一个绝对的时间尺度表明所取得的加速。我们进一步比较我们的方法 [3, 4] 我们的分解策略从一线在情商 (11)。而这一策略提高他们的方法, 因为在每一次迭代的速度逆做更有效的高效, 性能达不到我们的方法。对于二次数据集, 我们随机抽样的 100 个图像的大小为 100 从一个城市的场景 (数据集和原始图像的补充) 100。在这种情况下, $K \leq N$ 和 N, K 大。因此, 求解线性系统成为主要的成本 (见表 1) 和效益 \mathbf{f}_i 缓存 (不能在 [3, 4] 做的) 变得更加明显。因此, 特别是对于大型数据集, 我们的方法更快的收敛速度, 通常更好的优化。注意, 我们已经与 [3, 4] 和我们改进的分解策略从 \mathbf{f}_i 一线在情商 (11)。我们也比较卷积编码, 基于补丁的稀疏编码。其中一个主要的挑战是大型数据集, 其中大部分补丁为基础的方法变得不可行。从 10 个图像中考虑学习 1000 个 1000 像素的像素。把所有补丁纳入学习需要 1000 万个培训补丁。K-SVD, 例如, 不能处理的, 在我们的电脑多的数据, 所以我们跑了 10 100×100 像素的图像比较。在训练集和 100 次迭代使用的所有补丁, K-SVD 耗时 13.1 小时收敛而我们的方法只需要大约 4.5 分钟 (包括英特尔至

强 E5 / 酷睿 i7 机器与 132 GB 的 RAM)。除了以上的收敛性分析, 给出了 \mathbf{f}_i 滤波器演变的整个学习过程在图 2 的水果集。最初的 \mathbf{f}_i filters 似乎随机然后变成高斯斑点状结构后的几次迭代。8 次迭代之后, 观察到的结构是那些经常发生在基于块的字典学习方法非常相似, 而滤波器最终收敛于 Gabor 形状。

4. 学习特征

4.1. 滤波学习

我们训练我们的滤波器对果实和城市集 [22] 与局部对比度归一化应用。图 3 显示结果滤波器收敛后 (我们经过 13 次代, 布里斯托经过 300 次迭代)。

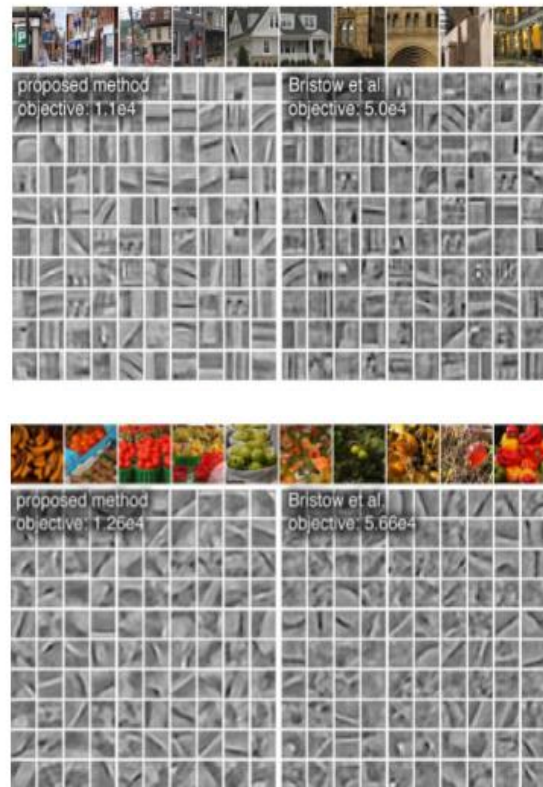


图 3: 城市 and 水果数据集上的过滤波器: 22。我们的数据显示缩略图随着 \mathbf{f}_i filters 学到的方法 (左), 在 [3, 4]。在这两种情况下, 我们的方法 \mathbf{f}_i finds 局部最优的目的, 比方法低 3-4。

虽然滤波器在 \tilde{f} 第一眼看上去相似，我们的结果包含较少的数据规范的特征，这使得它们更一般的为我们展示了对证券交易委员会的其他类型的图像重建。5.1

4.2. 边界条件

公式 3 是一个优雅的配方，允许一般边界条件必须融入学习并重建步骤。通常，我们把边界掩模，这样它就不会有助于目标函数。如图 4 可见，边界尚进行外推数据 \tilde{f} 切割线和其他高频结构继续在图像边界，但很快脱落之后。

4.3. 从稀疏数据中学习

混合矩阵方程 3 中不仅允许一般边界约束，但对于任何类型的线性算子的应用。在图 5 中，我们证明可以使用不完整的数据的学习 \tilde{f} 滤波器。我们随后使用 \tilde{f} 滤波器从图像不完整的测量学，图像修复缺失的部分和评估所获得的峰值信噪比为不同程度的不完全性。正如预期的那样，质量的重建下降，减少量的观测。然而，不完全测量学习 \tilde{f} 滤波器可能对许多应用是有趣的（如：）

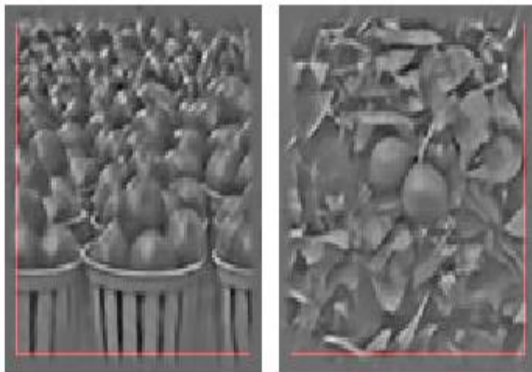
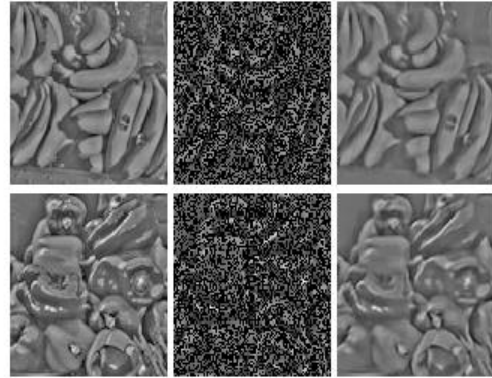


图 4：建议制定允许我们使用圆形的边界条件，证明两个例子。在实践中，图像边界（红色）以外的区域外推，但不影响目标。



Subsampling	90 %	70 %	50 %	30 %	10 %
PSNR	23.2 dB	21.3 dB	19.2 dB	16.5 dB	14.8 dB

图 5：从稀疏观察中学习。最上面的行显示 50% 个采样的例子。原始图像被显示在左栏，随机抽样观测中心，并重建整个图像使用 \tilde{f} 滤波器从这些稀疏的观察正确。下表滤波器：不同的数据采样率的学习评价 \tilde{f} 滤波器。我们发现 16 的总 100 学习 \tilde{f} 滤波器各采样率高于表。请看 \tilde{f} 滤波器的全套的补充。你可以看到，不到 50% 的抽样，重建质量有明显下降 \tilde{f} 由于 \tilde{f} 滤波器降低质量。

滤波器自适应学习去马赛克)，但目前没有任何现有的支持（ $E\tilde{f}$ 系数）的卷积稀疏编码方法。图 5 显示，即使采样率高达 50%，学习 \tilde{f} 滤波器和修复重建的质量是比较好的。



图 6: 修复的例子表明: 原始图像 (左), 随机抽样的不完全观测 (中间偏左), 重建了 \mathbf{f}_i 滤波器算法 (中右), 和 \mathbf{f}_i 滤波器从 [3, 4] (右)。除了这个例子, 我们评估了一个更大的数据集在图 7 的重建质量。

Image	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
PSNR ours	23.46	25.60	24.64	24.66	30.13	28.10	24.42	24.97	22.07	26.15	25.90	20.76	24.20	24.06	23.60	24.28	22.26	26.03	21.36	28.36	22.89	21.52
PSNR [3,4]	23.06	24.58	24.28	24.71	29.40	27.27	23.99	24.62	21.79	25.13	25.22	20.50	23.92	23.57	23.37	23.91	21.77	25.74	21.10	27.80	22.74	21.42

Image	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
PSNR ours	23.21	25.35	24.69	24.70	29.88	27.82	24.24	25.44	21.88	26.29	26.05	20.55	24.26	23.80	23.46	24.58	21.90	25.86	21.16	28.11	22.96	21.43
PSNR [3,4]	22.91	24.68	24.44	24.72	29.33	27.28	23.81	25.21	21.60	25.72	25.58	20.33	23.92	23.44	23.03	24.32	21.55	25.70	21.00	27.72	22.62	21.35

图 7: 为 \mathbf{f}_i 重建质量滤波器学习算法 (表, 中心行) 和 \mathbf{f}_i 滤波器在 [3] 提出, 4 (表, 底排)。所有的重建都进行 50% 的抽样。上表显示了重建结果与 \mathbf{f}_i 滤波器从 [22] 果数据了解到, 下一个节目从城市数据集的 \mathbf{f}_i 滤波器重建。该数据集包括 22 个图像, 没有一个是集学习 \mathbf{f}_i 训练部分滤波器。随着图像 4 例外, 我们的算法的结果在更高质量的重建都 \mathbf{f}_i 滤波器组。

5. 重建

在这一部分中, 我们评估的算法重构信号时, \mathbf{f}_i 滤波器是已经学过的或已知的先验。重建问题是所有 \mathbf{f}_i 滤波器从算法的代码更新步求解。2。

5.1. 重构验证

图 6 显示了一个不完整的测量图像重建的例子。这是一个修复的问

题, 我们测试 \mathbf{f}_i 滤波器从水果数据库学习。我们比较了重建质量滤波器使用学习方法和学习 \mathbf{f}_i 滤波器的方法 [3, 4]。不仅是我们的 \mathbf{f}_i 滤波器产生更好的重建的边缘和尖锐特征, 但我们的框架允许我们在不牺牲性能 \mathbf{f}_i 碑文在 CSC 求解解决这个修复问题, 这在以往的工作中由于“傅里叶的把戏是不可能的”。图 7 中的实验评估包

含 22 个图像重建质量，只有一个例外，表明改进的两滤波器组学在城市和水果集相比以前的工作质量。

5.2. 非标准化数据

在大多数的结果中，我们显示对比度归一化的例子。然而，非标准化的数据，可以直观地处理的方式。无论是 f_i 过滤器是从非规范化培训数据或为直流偏移低频词学是学习后直接添加到 f_i 滤波器组。对于前一种方法观察是典型的 Gabor 像 f_i 滤波器具有不同的直流偏移。后者可以被解释为增加平滑之前在低频词的形式而非严格稀疏。重建必须共同解决的 f_i 滤波器系数 f_i 系数以及低频词，图 8 所示。我们还比较了这种方法与一个国家的最先进的压缩感知方法，从东等。[8]。从图 7 使用相同的数据集（50% 随机抽样与非规范化的数据）的方法达到 23.5 dB 的平均峰值信噪比达到 29 dB 时，我们。这一初步结果表明，稀疏卷积编码的进一步调查可能会导致许多卓有成效的应用程序，即使在低级别的功能学习以外。



图 8: 修复非规范化的数据: randomly subsampled, 不完全观测 (左, 右, 中心) 所提出的 f_i 滤波器 (中心重建左、右)。

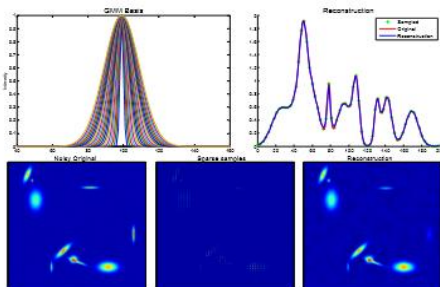


图 9: 已知卷积基础的重建。在这个例子中 f_i 过滤器是从一维高斯采样 (左上) 和用于 f_i 卷积模型稀疏的样本集 (右)。同样的实验在 2D, 在目标信号受到噪声 (左下), 采样 (底部中心), 然后重建从只有 6.25% 的噪声测量 (右下)。

5.3. 已知基础重建

我们也评估所提出的方法 fitting 卷积模型稀疏和噪声测量时, f_i 滤波器是已知的。最近这些类型的重建的例子已经出现在计算成像领域的基础是出于物理模型 (例如, [12, 11])。总的来说, physically motivated 卷积稀疏编码有宽

在雷达、声纳、超声和地震成像等领域的应用, 但我们留下了详细的评价这些应用到未来的工作。图 9 显示了稀疏采样数在一维、二维重建。过滤器是由一个高斯分布和二维的例子, 测量采样进一步损坏的独立同分布的高斯噪声的标准偏差 $\sigma = 0.01$ 。这个实验说明 CSC 的效用非特征学习型应用, 如一般的高斯混合模型。所提出的方法是能够从只有 6.25% 的样品中的高质量潜在信号恢复

6. 讨论

总之, 我们提出了一种新的方法, 使用卷积稀疏编码的学习和重建问题。我们的配方允许适当的边界条件是 f_i 灵活的, 它允许特征不完全观察学习, 或任何类型的线性算子的估计。我们表明, 我们的框架是更快的速度比的艺术和收敛到更好的解决方案。

未来的工作, 虽然速度比现有的方法已经制定, 我们是天生的并行性和运行时可以明显 f_i 明显改善由 EF f_i 高效的 GPU 实现。它是评估学习和高维问题特征重建有趣, 如三维高光谱图

像数据[15]和[17]的4D光场。最后，这将是有趣的应用建议的框架，更复杂的层次结构和网络[13, 22]，可能是特别有用的高级计算机视觉应用，如识别。

结论卷积稀疏编码是一个强大的框架，有可能取代或补充流行的补丁为基础的学习和重建方法。这是适用于范围广泛的计算机视觉问题，如特征学习，去噪，修复，和去马赛克。所提出的方法，我们希望有一个实用的方法来解决一般问题的高效和紧凑最灵活的方式。

致谢

费利克斯海德是一四年从英属哥伦比亚大学奖学金的支持。戈登 Wetzstein 是压缩感知的英特尔战略研究联盟的支持。沃尔夫冈希的阿卜杜拉国王科技大学基线资助和 NSERC 发现给予支持。

参考文献

- [1] M. S. Almeida and M. A. Figueiredo. Frame-based image deblurring with unknown boundary conditions using the alternating direction method of multipliers. In Proc. ICIP, pages 582 - 585, 2013. 1, 2, 3
- [2] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. Foundations and Trends in Machine Learning, 3(1):1 - 122, 2011. 1, 3
- [3] H. Bristow, A. Eriksson, and S. Lucey. Fast convolutional sparse coding. In Proc. CVPR, pages 391 - 398, 2013. 1, 2, 3, 4, 5, 6, 7
- [4] H. Bristow and S. Lucey. Optimization methods for convolutional sparse coding. arXiv:1406.2407, 2014. 1, 2, 3, 4, 5, 6, 7
- [5] A. M. Bruckstein, D. L. Donoho, and M. Elad. From sparse solutions of systems of equations to sparse modeling of signals and images. SIAM review, 51(1):34 - 81, 2009. 1
- [6] V. Cevher, S. Becker, and M. Schmidt. Convex optimization for big data: Scalable, randomized, and parallel algorithms for big data analytics. IEEE Signal Processing Magazine, 31(5):32 - 43, 2014. 1
- [7] B. Chen, G. Polatkan, G. Sapiro, D. Blei, D. Dunson, and L. Carin. Deep learning with hierarchical convolutional factor analysis. IEEE Trans. PAMI, 35(8):1887 - 1901, 2013. 1
- [8] W. Dong, G. Shi, X. Li, Y. Ma, and F. Huang. Compressive sensing via nonlocal low-rank regularization. IEEE Transactions on Image Processing, 23(8):3618 - 3632, 2014. 8
- [9] R. B. Grosse, R. Raina, H. Kwong, and A. Y. Ng. Shiftinvariance sparse coding for audio classification. In Proc. UAI, pages 149 - 158, 2007. 1
- [10] F. Heide, M. Rouf, M. B. Hullin, B. Labitzke, W. Heidrich, and A. Kolb. High-quality computational imaging through simple lenses. ACM Trans. Graph., 32(5):149, 2013. 3
- [11] F. Heide, L. Xiao, A. Kolb, M. B. Hullin, and W. Heidrich. Imaging in scattering media using correlation image sensors and sparse convolutional coding. OSA Opt. Exp., 22(21):26338 - 26350, Oct 2014. 1, 8
- [12] X. Hu, Y. Deng, X. Lin, J. Suo, Q. Dai, C. Barsi, and R. Raskar. Robust and accurate transient light transport decomposition via convolutional sparse coding. OSA Opt. Lett.,

- 39(11):3177 - 3180, 2014. 1, 8
- [13] K. Kavukcuoglu, P. Sermanet, Y. Boureau, K. Gregor, M. Mathieu, and Y. LeCun. Learning convolutional feature hierarchies for visual recognition. In Proc. NIPS, 2010. 1, 2, 8
- [14] B. Kong and C. C. Fowlkes. Fast Convolutional Sparse Coding (FCSC). Technical report, UCI, May 2014. 1, 2, 3
- [15] X. Lin, G. Wetzstein, Y. Liu, and Q. Dai. Dual-coded compressive hyperspectral imaging. OSA Opt. Lett., 39(7):2044 - 2047, 2014. 8
- [16] J. Mairal, F. Bach, J. Ponce, and G. Sapiro. Online dictionary learning for sparse coding. In Proc. ICML, pages 689 - 696. ACM, 2009. 1
- [17] K. Marwah, G. Wetzstein, Y. Bando, and R. Raskar. Compressive light field photography using overcomplete dictionaries and optimized projections. ACM Trans. Graph. (SIGGRAPH), 32(4):46:1 - 46:12, 2013. 8
- [18] B. A. Olshausen and D. J. Field. Sparse coding with an overcomplete basis set: A strategy employed by v1? Vision Research, 37(23): 3311 - 3325, 1997. 1
- [19] N. Parikh and S. Boyd. Proximal algorithms. Foundations and Trends in Optimization, 1(3):123 - 231, 2013. 1, 3
- [20] M. Schmidt, N. L. Roux, and F. Bach. Minimizing finite sums with the stochastic average gradient. arXiv:1309.2388, 2013. 1, 3
- [21] A. Szlam, K. Kavukcuoglu, and Y. LeCun. Convolutional matching pursuit and dictionary training. arXiv:1010.0422, 2010. 1
- [22] M. D. Zeiler, D. Krishnan, G. W. Taylor, and R. Fergus. Deconvolutional networks. In Proc. CVPR, pages 2528 - 2535, 2010. 1, 2, 4, 5, 6, 7, 8
- [23] M. D. Zeiler, G. W. Taylor, and R. Fergus. Adaptive deconvolutional networks for mid and high level feature learning. In Proc. ICCV, pages 2018 - 2025, 2011. 1
- [24] D. Zoran and Y. Weiss. From learning models of natural image patches to whole image restoration. In Proc. ICCV, pages 479 - 486, 2011. 1

