

指导教师： 杨涛

提交时间： 2015.03.29

The task of
Digital Image Processing

数字图像处理

School of Computer Science

No: _____

姓名： 张峻岑

学号： 2012303052

班 号： 10011207



收缩领域的图像复原处理

Uwe Schmidt Stefan Roth Department of Computer Science, TU Darmstadt

摘要

能很好的扩展较大的图像，比如说在消费者这一块百万像素的图像。以计算方式地昂贵的优化往往是罪魁祸首。然而，存在的高效的替代品，它们都没有达到原图像的质量水平。本文的目的是建立一个有效的图像复原方法，同时提供计算效率和高质量的图像恢复。为此我们建议建立收缩领域，以领域为基础随机的构架图像模型和优化算法相结合的单个但愿。基本收缩操作支持连接到微波的方法，但在这里使用的是随机领域背景 (in a random field context)。计算效率通过构造卷积和 DFT 为核心部件来实现；高恢复质量的实现是通过所有模型参数的损失为基础的培训和级联架构。不同于浩大的工程解决方案，我们的学习方法能够轻松的适应不同的效率和图像质量之间的取舍。我们展示了当前最高水平的恢复结果的计算效率，通过固有的并行对比发现其具有明显的加速潜力。

1. 介绍

去除图像伪影的图像复原方法如噪声、模糊、云纹等，在学术研究界均获得显著的关注，在数字成像的实际应用方面也是如此 [e.g.,6]。在学术研究中，重点主要是为了实现最大的图像质量，较大的忽略了计算恢复过程中的影响[18,21,31]。在数字成像的实际应用中，计算资源通常受到严重制约，然而这是因为相机上的硬件的处理能力比传统的个人电脑要低上许多倍。但是，即使是在台式个人电脑上哪怕是目前最先进的技术也往往需要几分钟来对一个 VGA 尺寸的图像（相当于 0.3 百万像素）去除干扰。现代数码相机能够拍摄 1600 万像素甚至更多的图像，而另一方面，总的来说现有技术并没有形成规模。值得注意的例外主要是 BM3D[5]，它具有高效率和图像质量，但是

是一个庞大的多年细化的工程。而且，它使用块匹配作为密钥计算部件的并行结构，比如说 GPU 和 DSP，是有挑战性的想法。人们希望在将来进步的恢复技术能够将嵌入式硬件直接在相机上使用，但由于图像分辨率的增加这也未必是不可能的。因此，为了弥补现有的图像复原技术的缺陷，同时实现高图像质量和计算效率，需要一种不同的图像复原方法。

在本文中，我们引入收缩领域，一个有原则的图像复原体系结构，从通常的随机场模型中获取的现有的优化算法。特别是，收缩字段欠其计算效率到从所谓 halfquadratic 优化添加剂形式衍生的特定种类的二次松弛技术的方法[11] - 以每像素级不会施加的操作只有圈和离散傅立叶变换 (的 DFT)。但是，与现有的添加剂半二次方法[15, 28]，我们充分利用学习通过损失为基础的培训与应用程序特定功能的丧失[比照。 13]，这使我们能够实现更高水平的恢复质量。此外，在其对比标准随机场，这是潜在的特定功能，收缩领域模型直接的潜在的相关“收缩功能”。这增加了添加剂的形式半二次方法的灵活性，因为我们可以证明，势函数总是导致单调的收缩功能。与此相反，我们可以做到学会非单调的收缩功能，类似于那些在小波图像去噪[12]的情况下被有区别的对待。更重要的是，使用收缩功能直接承认有效的学习，因为模型的预测和梯度 WRT 该模型的参数可以计算在封闭的形式。最后，我们的方法采用的预测级联[25]，使用多阶段模型的迭代优化。严格意义上来说，我们同时学习了随机场模型和迭代优化算法 [cf.2]。

所提出的方法有几个关键的好处：(1) 它的概念很简单和标准的推理程序随机场模型得出；(2) 实现了非常高的水平的图像

质量看齐，或超越，艺术的当前状态；(3) 在计算上是非常有效的与复杂性为 $O(D \log D)$ (其中 D 是像素数)；(4) 它提供了并行的高度使得它非常适合用于 GPU 和 DSP 的实现；(5) 不同于大量工程化技术，如 BM3D，所有参数都可以直接从使用简单的基于梯度的优化实施例的数据得知，使得它易于应用，适应新的设置，如效率和恢复质量之间的不同折衷。

1.1 其他相关工作

正规化或先验和收缩功能之间的联系已被广泛研究小波图像复原[e.g.,1,26]。半二次优化和收缩功能添加剂形式之间的连接已经由王等人指出。[28]基于对方法[28]，克里希南和宏泰[15]普及添加剂半二次优化非盲解的任务[e.g.,30]。我们从这个方面从这里开始，但相比之下不使用固定的潜在功能，但是采用更普遍的，易于学习的收缩功能。

连续条件随机域 (CRFs) 的图像复原的歧视性的训练已经由萨穆埃尔和塔彭[22]提出，最近被陈等人重新审视[3]。高斯模型及其相关的损失为基础的培训首先用塔彭等人介绍[27]。最近，jancsary 等人[13]改进这领域引入回归树 (RTFS)，更灵活的高斯 CRF 有了最小化损失的成果。相对于这些以前的方法，所提出的收缩领域承认更高效推理，并且可以通过标准的基于梯度的优化装置很容易训练。区别地学习随机场模型和其相关联的优化算法已经由 Barbu 提出并通过了[2]。本着同样的精神，Schmidt 等人,[25]最近训练有素 RTFS 的级联。而[2]是非常有效的，它产生给定的相同的模型复杂性较低的图像质量，并依赖于复杂的和耗时的学习过程。我们的工作是从概念上最相似的[25]，其也通常出于半二次推断。不过，在这里我们还得到了模型参数 (收缩功能滤波器响应) 专门从加半二次型。通过这样做，我们权衡造型灵活性 (相比[25]) 针对更高效推理和易于培训。

2. 背景：半二次优化

作为一个起点，我们考虑恢复从损坏的观察 y 一个图像 x 结合观察可能性和调用之前贝叶斯规则的图像：

$$p(x|y) \propto p(y|x) \cdot p(x) \quad (1)$$

$$\propto \mathcal{N}(y; Kx, I/\lambda) \cdot \prod_{i=1}^N \prod_{c \in \mathcal{C}} \exp(-\rho_i(f_i^T x_{(c)})) \quad (2)$$

这里的腐败过程建模与高斯可能性 (或数据项) ，其中 $Kx \equiv k \otimes x$ 表示 x 的卷积内核 (点扩散函数) k 和 λ 是关系到假定加性高斯噪声的强度。Regularization is provided through a Markov random field model (fields of experts [21]) with robust potential functions $\rightarrow i$ that model the responses $f_i^T x_{(c)}$ of filters f_i over all cliques $c \in \mathcal{C}$ of the image x .

后验分布 $p(x|y) \propto \exp(-E(x|y))$ 可

以通过其相关的吉布斯能量表示

$$E(x|y) = \frac{\lambda}{2} \|y - Kx\|^2 + \sum_{i=1}^N \sum_{c \in \mathcal{C}} \rho_i(f_i^T x_{(c)})$$

这使得通过寻找能量最小预测恢复图像 $\hat{x} = \arg \min_x E(x|y) = \arg \max_x p(x|y)$.

一种方式，以减少方程 (3) 是可直接使用梯度下降算法。另一种流行的方法

[10, 11, 28] ，一个我们分析和扩展在本文中，是先引入独立的辅助变量 Z 所有滤波器响应英尺 x (三) 获得增强能源函数 $E(X, Z | y)$ 的在于 。块坐标下降

的策略，最大限度地减少 WRT 之间交替 x 和 z ，然后用于最小化 $E(X, Z | y)$ 。这种方法通常比减少 $\tilde{E}(x|y)$ 较快的收敛，每下降一步往往是比较简单的开展。

具体地，辅助变量是以这样的方式到，电子导入 $E(X | Z, Y)$ 变为二次函数；最大限度地减少 $\tilde{E}(Z | X, Y)$ 只是相当于解决许多独立的单变量优化问题。在计算机视觉，首次由格曼提出了名义下“半二次”转正的同事[10, 11]这种做法。换句话说，该算法的每次迭代使用不同的二次松弛 $E(X | Z, Y)$ 的原始目标函数 $E(X | Y)$ ，通过辅助变量 z 确定。

半二次方法可以进一步分为添加剂[11]和乘法[10]的形式。没有进入细节，在实践中主要计算的区别在于

$\arg \min_x E(x | z, y) = \Omega(z, y)^{-1} \eta(y)$ 中的乘法形式，并且

$\arg \min_x E(x | z, y) = \Omega(z, y)^{-1} \eta(y)$ 的添加剂的形式。在这里， $RD \times D$ 是一个稀疏矩阵 D 是像素的数目，并 $\eta \in \mathbf{R}^D$ 是一个向量。

这意味着，所述二次函数可以通过求解线性方程系统被最小化，凡在乘法方式中， Z 仅影响方程系统矩阵 Ω ，并且在添加剂的形式仅右手的方程系统的侧 η 。因此，添加剂的形式是在一般的计算更加有吸引力，因为方程系统矩阵保持恒定期间迭代优化（例如，可重新使用 Ω 的分解，或 Ω 可能与基的变化对角化）。

然而，面临的一个挑战是，添加剂的形式是不能直接适用于多种重尾势函数实际意义 \rightarrow 。为了解决这个问题，并加快收敛，王等人 [28] 提出的延续方案，其中一个参数 β 被半二次优化（比照海藻酸钠。1）

中增加了。Concretely, the problem is

cast as

$$\hat{x} = \operatorname{argmin}_x E(x|y) = \operatorname{argmin}_{x, z}$$

!!1

$$E(x, z | y)$$

with

$$E_\beta(x, z | y) = \frac{\lambda}{2} \|y - Kx\|^2 + \sum_{i=1}^N \sum_{c \in C} \left(\frac{\beta}{2} (f_i^T x_{(c)} - z_{ic})^2 + \rho_i(z_{ic}) \right). \quad (4)$$

直观地说，当 $\beta \rightarrow \infty$ 如图 1 所示，辅助变量 Z 接近其相应的滤波器响应和公式。 (4) 收敛到原来的公式。

(3)。这种非盲图像去卷积法已经在近几年由 Krishnan 和 Fergus 普及。

要知道为什么这种做法是如此吸引人，它是由仔细观察交替优化过程得到的启发，这是总结了海藻酸钠。1。具体地讲，在两个更新步骤在计算上很便宜时 - 也就是我们现在假设 - 2D 卷积是用圆形（周期性）的边界条件。

$$f_{i,\beta}(v) = \operatorname{argmin}_z \left(\rho_i(z) + \frac{\beta}{2} (v - z)^2 \right) \quad (5)$$

$$g_\beta(z) = \left[\frac{\lambda}{\beta} K^T K + \sum_{i=1}^N F_i^T F_i \right]^{-1} \left[\frac{\lambda}{\beta} K^T y + \sum_{i=1}^N F_i^T z_i \right] = \mathcal{F}^{-1} \left[\frac{\mathcal{F} \left(\frac{\lambda}{\beta} K^T y + \sum_{i=1}^N F_i^T z_i \right)}{\frac{\lambda}{\beta} \tilde{K}^* \circ \tilde{K} + \sum_{i=1}^N \tilde{F}_i^* \circ \tilde{F}_i} \right]. \quad (6)$$

$Fx = [f^T x_{(c_1)}, \dots, f^T x_{(c_{|C|})}]^T \equiv f \otimes x$ 表示二维卷积与过滤器 F 。光学传递函数 F 是 $F(f)$ 的从滤波器（点扩散函数） f ，其中 F 表示离散傅立叶变换 (DFT) 而得。需要注意的是除法和乘法 () 应用于元素明智的公式。 (6)； F^* 表示 F 的复共轭。

Algorithm 1 Half-quadr. minimization with continuation

Require: β -schedule β_1, \dots, β_T with $\beta_{t+1} > \beta_t$

```

 $\hat{x}_0 \leftarrow y$ 
for  $t \leftarrow 1$  to  $T$  do
     $\hat{z}_{ic} \leftarrow \arg \min_{z_{ic}} E_{\beta_t}(z|\hat{x}_{t-1}, y) = f_{i, \beta_t}(f_i^T \hat{x}_{t-1(c)})$ 
     $\hat{x}_t \leftarrow \arg \min_x E_{\beta_t}(x|\hat{z}, y) = g_{\beta_t}(\hat{z})$ 
    
```

式。(5)是很便宜的计算,因为音响, !
 (v)是单变量函数,可以预先计算为v的所有可能的值,然后存储在一个查找表用于快速检索[15]。重要的是,只有加半二次型可通过更新式的图像x。(6)非常快速地在封闭的形式,因为所有的卷积矩阵(以及因此的整个方程组矩阵)可通过的DFT,这意味着,求解线性方程系统相当于逐元素师在变换域,随后通过逆对角化DFT保留在溶液中空间域[例如, 15, 28]。注意,这仅需要N+1圈和N+3的DFT具有一个总体复杂性为O(D log D)的算法,其中D是像素的数目。

2.1 收缩功能

$f_{i, \beta}$ 的作用是(公式5)被称为在小波图像复原文学收缩(或映射)函数[比照。12]。直观地说,其目的是紧缩小型过滤器/小波系数,也就是拉他们到零,因为他们被假定为噪声,而不是信号造成的。

就目前而言,收缩功能的形状由#和其相关联的潜在功能*i*单独确定(式(5)),参见图1(A-D),用于图示)。但是,我们观察到的所有 $f_{i, \beta}$ 根据公式。(5)单调增加的功能,而不管潜力pi。为了证明这个命题1,它有以下引理(证明在增刊材料)是有用的:

Lemma 1. For any function $f : \mathbb{R} \rightarrow \mathbb{R}$ and all $\epsilon \geq 0$, $\arg \min_z f(z) \leq \arg \min_z (f(z) - \epsilon z)$.

Proposition 1. For all $\epsilon, \beta \geq 0, v \in \mathbb{R}$ and any $\rho(z)$, the shrinkage function $f_\beta(v) = \arg \min_z (\rho(z) + \frac{\beta}{2}(v-z)^2)$ is monotonically increasing, i.e. $f_\beta(v) \leq f_\beta(v + \epsilon)$.

Proof.

$$f_\beta(v + \epsilon) = \arg \min_z \left(\rho(z) + \frac{\beta}{2}(v + \epsilon - z)^2 \right) \quad (7)$$

$$= \arg \min_z \left(\rho(z) + \frac{\beta}{2}(v - z)^2 - \epsilon \beta z \right) \quad (8)$$

It follows from Lemma 1 that $f_\beta(v) \leq f_\beta(v + \epsilon)$. \square

尽我们所知,这在之前还没有被观测到。更重要的是,这意味着可以通过直接建模的收缩功能,而不是潜在的功能获得的添加剂的半二次优化额外的灵活性。

3. 级联收缩场

正如下面我们刚刚提到的,将进一步证明,直接建模的收缩功能,是有吸引力。为达到此目的,我们取出势函数和式(5)中全部的相关联的优化问题,更换 $f_{i, \beta}$ 与模拟为高斯RBF核的线性组合灵活的收缩功能:

$$f_{\pi_i}(v) = \sum_{j=1}^M \pi_{ij} \exp\left(-\frac{\gamma}{2}(v - \mu_j)^2\right). \quad (9)$$

我们假设共享精度 γ 和放置在内核固定,等距离的位置 μ_j 。我们使用多达M=53高斯核,使有关的功能的形状,没有进一步的假设(两个例子示于图1(e-f))。

收缩功能被广泛研究了小波修复文献。而不是手动选择收缩功能但是,我们学习他们的数据通过设置权重

π_{ij} 方程的参数形式在式(9)中。这与之前的工作相比对比是鲜明的。企

图在区别地学习收缩功能小波恢复存在[例如, 12], 但并不常见。此外, 在这里, 小波图像恢复是完全不同的, 因为在恢复图像的像素不经由随机字段连接。

我们不知道使用了学习的一半二次优化这种特殊形式的背景下任何先前的工作。因此, 这种快速的优化方法的全部潜力还没有被解锁, 因为模型参数总是被手动选择的。

Algorithm 2 Inference with a cascade of shrinkage fields

```

 $\hat{x}_0 \leftarrow y$ 
for  $t \leftarrow 1$  to  $T$  do
     $\hat{x}_t \leftarrow g_{\Theta_t}(\hat{x}_{t-1})$ 
    
```

此外, 对于 ALG1 的迭代次数的 β 延续时间表, 一般都是手动选择。

在下文中, 我们显示了如何同时保持了这种方法的计算好处克服所有这些限制。为此目的, 我们得到的所有模型参数都从训练数据(比的大小和过滤器数目, 以及优化的迭代次数等)中得来的。

直接建模的收缩功能的最重要的好处是, 它可以使我们减少在优化过程, 以一个单一的二次最小化在每个迭代中, 我们表示为一个收缩区(SF)的预测:

$$g_{\Theta}(x) = \mathcal{F}^{-1} \left[\frac{\mathcal{F}(\lambda \mathbf{K}^T y + \sum_{i=1}^N \mathbf{F}_i^T f_{\pi_i}(\mathbf{F}_i x))}{\lambda \mathbf{K}^* \circ \mathbf{K} + \sum_{i=1}^N \mathbf{F}_i^* \circ \mathbf{F}_i} \right] \quad (10)$$

$$= \Omega^{-1} \eta. \quad (11)$$

收缩场 $\mathcal{N}(\Omega^{-1} \eta, \Omega^{-1})$ 是一个特定的高斯条件随机场, 其矩 η 和 Ω 通过学习模型参数确定 Θ , 所观察到的图像 y 和点扩散函数 K 。一个关键的好处是, 收缩场预测 $g_{\Theta}(x)$ 和它的梯度

$$\frac{\partial g_{\Theta}(x)}{\partial \Theta} \text{ w. r. t. } \Theta \text{ 可以以闭合形式计算}$$

模型参数, 这允许有效的参数学习 (Sec. 3.1). 这是相对于其他配方更复杂的学习过程, 这需要解决使用双级优化嵌套的最小化问题[3.22]。请注意, 我们完全消除的延续参数 β , 它被吸收入式的权重 π_i 在式(9)中和融合 λ (这将被学习) 在方程(10)。

由于半二次优化通常涉及方程的几个(或许多)迭代在式(5)和(6), 我们同样可以串联多个预测到收缩字段(CSF)的级联, 就像 ALG. 2 中总结的那样。一个 CSF 因此串联成高斯 CRF [cf. 25]。需要注意的是收缩功能的概念并不在以前的 CRF 级联中存在。RTF 级联[25], 例如, 使用回归树来指定一元和成对的因素; 因为得到的方程系统矩阵不能通过的 DFT 对角化, 它们不承认快速闭合形式推理如等式(10)。

3.1 学习

我们使用损失最小化来学习模型

参数 $\Theta_t = \{\lambda_t, \pi_{ti}, f_{ti}\}_{i=1}^N$ 对于在 Alg. 2 的每一个阶段(迭代) t 。

通过我们的级联的每一个阶段学习不同的模型参数, 我们基本上是量身定做随机场模型相关的优化算法的每一次迭代。对于非盲解卷积, 我们按照[25]和参数化预测与模糊核, 以使得在测试时间中实例特定模糊(K 的公式10)设置; 模型没有经过修改的具体模糊。贪婪地学习模型阶段逐阶段从 $t=1, \dots, T$, 在阶段 t 我们尽量减少成本函数:

$$J(\Theta_t) = \sum_{s=1}^S \ell(\hat{x}_t^{(s)}; x_{gt}^{(s)}) \quad (12)$$

与训练数 $\{x_{gt}^{(s)}, y^{(s)}, k^{(s)}\}_{s=1}^S$,

当 $\hat{x}_t^{(s)}$ 通过 Alg. 2 得到。我们可以，原则上，使用任何连续可微的损失函数，具体选择的（负）峰值信噪比（PSNR）：

$$\ell(\hat{x}; x_{gt}) = -20 \log_{10} \left(\frac{R\sqrt{D}}{\|\hat{x} - x_{gt}\|} \right), \quad (13)$$

其中 D 表示 x 的像素的数量和 R 的像素的最大亮度水平。（i. e., R=255）

我们最大限度地减少式（12）与基于梯度的 L-BFGS 方法（使用实施由[23]）。为此，我们，类似于[13]，区分预测恢复图像 \hat{x}_t 的损失（在阶段

t) 的 WRT 模型参数 Θ_t 作为

$$\frac{\partial \ell(\hat{x}_t; x_{gt})}{\partial \Theta_t} = \frac{\partial \ell(\hat{x}_t; x_{gt})}{\partial \hat{x}_t} \cdot \frac{\partial \Omega_t^{-1} \eta_t}{\partial \Theta_t} \quad (14)$$

$$= \hat{c}_t^T \left[\frac{\partial \eta_t}{\partial \Theta_t} - \frac{\partial \Omega_t}{\partial \Theta_t} \hat{x}_t \right] \quad (15)$$

$$\text{with } \hat{c}_t = \Omega_t^{-1} \left[\frac{\partial \ell(\hat{x}_t; x_{gt})}{\partial \hat{x}_t} \right]^T.$$

像 \hat{x}_t 一样，我们可以有效地计算 \hat{c}_t 通过求解线性方程组明确地区分变换域中的要素。衍生物对于特定的模型参数以及进一步的细节，诸如边界处理由于周期性卷积和参数约束，这里省略了简洁和使方程更易读；然而，所有的细节可以在补充材料中找到。

在方程（12），每个阶段被训练得所述损耗在无论 t 为多少阶段实际上

应用在级联中使每个阶段尽可能的小；这也适用于[25]的级联模式。然而，与此相反的[25]的级联，它使用非可微回归树，以确定高斯 CRF 的参数和要求的定制的训练，我们的收缩功能是光滑和微调的。因此，我们就不需要基于梯度的和的组合优化之间交替（生长回归树）。此外，我们可以使

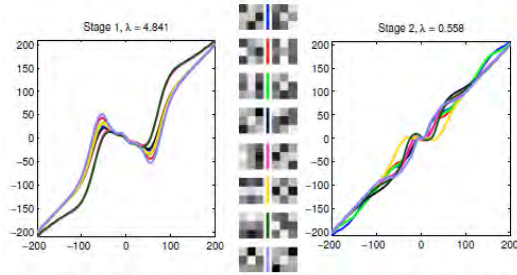


Figure 2. First two stages of learned $CSF_{3 \times 3}$ model. The shrinkage functions are color-matched with their corresponding filters.

用基于标准的梯度方法，以通过最小化共同训练模型的所有 T 阶段：

$$J(\Theta_{1, \dots, T}) = \sum_{s=1}^S \ell(\hat{x}_T^{(s)}; x_{gt}^{(s)}), \quad (16)$$

其中只有最终预测 \hat{x}_t 的损失是相关

的。衍生物 w. r. t. 各个阶段的模型参数是有效地计算，并采取相同的基本形式如式（15），其允许容易实现。请注意，所有阶段甚至可以同时应用边界操作，如填充和截断共同学习。所有的细节和推导都在补充材料。

4. 实验

σ	KSVD [7]	FoE [9]	BM3D [5]	LSSC [18]	EPLL [31]	opt-MRF [3]	CSF_{pw}^5	$CSF_{3 \times 3}^4$	$CSF_{5 \times 5}^4$	$CSF_{7 \times 7}^5$	ARF-4 [2]	RTF ₅ [24]
15	30.87	30.99	31.08	31.27	31.19	31.18	29.99	30.78	31.12	31.24	30.70	—
25	28.28	28.40	28.56	28.70	28.68	28.66	27.47	28.29	28.58	28.72	28.20	28.75

Table 1. Average PSNR (dB) on 68 images from [21] for image denoising with $\sigma = 15, 25$; left part quoted from Chen *et al.* [3]. Training of our CSF models and denoising carried out *without* 8-bit quantization of noisy images to allow comparison with [2, 3].

训练。虽然类高斯的CRF的，可以在我们的方法中的一个阶段而得知被限制（相对于[13]），这限制来自于快速预测和学习的实质益处。这意味着我们可以训练我们的模型在比较大的数据集 - 甚至是一个简单的matlab实现。产生用于我们的去噪实验训练数据，我们裁剪一个 256×256 像素区域从每个 400 个图像的伯克利分割数据集[19]，即我们的训练这样设置大致包含 2500 万像素。

我们贪婪地培养了我们的模型增加容量的四款不同配置的 5 个阶段：

- CSF_{pw}^5 Pairwise model with fixed $f = \{[1, -1], [1, -1]^T\}$.
- $CSF_{3 \times 3}^5$ Fully trained model with 8 filters of size 3×3 .
- $CSF_{5 \times 5}^5$ Fully trained model with 24 filters of size 5×5 .
- $CSF_{7 \times 7}^5$ Fully trained model with 48 filters of size 7×7 .

因此， $CSF_{m \times m}^T$ 表示与 $m^2 - 1$ 的 T 分期级联在大小 $m \times m$ 过滤器（如果 $T < 5$ ，在测试时间只有 T 分期进行了评估在测试时间；预测可以停止在任何阶段）。需要注意的是有更多的配置也是可能的，并会导致不同的性能和速度的折衷，它可以被选择以适应特定的应用。图。2 和 1（E-F）示出了学习型号分别为 $CSF_{3 \times 3}$ 和 CSF_{pw} 的前两个阶段，这是很好的例子，通过我们的观察，几乎所有学到的收缩功能没有单调递增，这意味着它们不可能被学习的潜在功能（cf. Sec. 2）获得。

去噪。我们首先计算图像去噪（即 $K = 1$ ）的任务，为此，我们训练我们的模型以消除高斯噪声标准差 $\sigma = 25$ 。嘈杂的训练图像通过加入模拟的高斯噪声的干净图像获得。我们随后量化的噪声图像的强度值，以 8 位，使训练数据更加逼真。在实践中，噪声图像是总是整数值和范围有限，如强度值 $\{0 \dots, 255\}$ 。

训练后的模型，我们评估他们在 68（8 位量化噪音）测试图像最初是由 [21] 提出，其中有后来成为一个参考的图像去噪设置；图 4 示出一个去噪例子。我们比较当前最先进的技术的各种不同选择。结果在表格 2 表明，回归树字段（RTFS）[24] 的一个（5 级）级联达到最佳性能（训练了与相同的数据作为我们的模型）。这并不奇怪，因为更灵活的 RTFS 不作任何假设，噪声（相对于所有其他方法在表格 2），并能有效地处理额外的量化噪声。至于其他的方法，我们通过我们最强大的 $CSF_{7 \times 7}^5$ 模型以 0.22 分贝跑赢大盘的最强竞争对手 BM3D。它也可以被观察到，我们的模型的结果后只有 3-4 阶段饱和，因此，“收敛”非常快。

Method	PSNR	St.	CSF_{pw}	$CSF_{3 \times 3}$	$CSF_{5 \times 5}$	$CSF_{7 \times 7}$
BLS-GSM [20]	27.98	1	26.60	27.54	27.46	27.70
5×5 FoE [9]	28.22	2	27.26	27.93	28.26	28.38
LSSC [18]	28.23	3	27.31	28.02	28.34	28.45
BM3D [5]	28.31	4	27.36	28.05	28.37	28.52
RTF ₅ [24]	28.74	5	27.36	28.08	28.39	28.53

Table 2. Average PSNR (dB) on 68 images from [21] for image denoising with $\sigma = 25$. On the right, each row shows the results from the respective stage of our models.

我们还比较反对近期推出的 OPT-MRF 陈等人。[3]有两个原因：首先，它是目前图像复原表现最好的 CRF 之一，通过使用更好的优化技术与最初提出的[22]模型架构来实现。其次，它采用了模型配置非常类似我们的，那就是 48 滤镜尺寸 7×7 的，这是从数据充分吸取（包括相关的潜在功能）。此外，我们比较的快速主动随机场（ARF）模型巴尔布[2]，它采用大小 5×5 24 个过滤器。由于他们两人都没有受过训练，有 8 位量化噪声图像进行评估，我们用自己的设置给不给我们的模型不公平的优势。因此，我们还训练并没有量化评估我们的模型。结果在表格 1 表明，我们超越[2, 3]，并且还可以与基于 RTF 级联模型竞争[24]（受训与非量化的图像），其额外的灵活性似乎没有在这里还清由于图像噪声是真正的高斯。结果进一步表明，我们还可以争夺的噪音水平 $\sigma = 15$ ，为此，我们培养更多的模型。

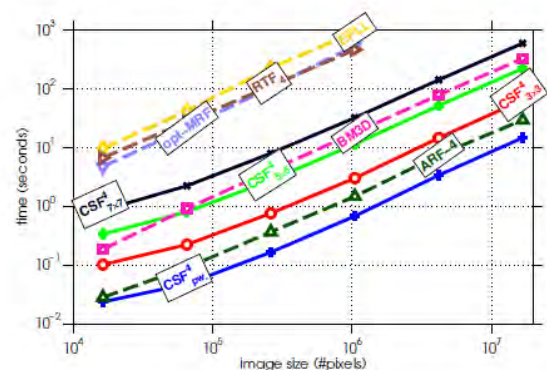
运行。运行时的比较图图像去噪图 3 显示，我们的模型扩展到超过 16 百万像素的合理运行时图像尺寸（最大 10 分钟，我们用一个简单的单线程 matlab 实现最佳模式，只有 23 秒一个 GPU）。

虽然 RTFS 的级联[24]是非常灵活的，并产生恢复结果，其相对复杂和高度优化的 C++ 实现取决于多线程来提高运行时性能先进设备，最先进的。比较单线程性能（图 3），它是约一个数量级慢相比我们 $CSF_{7 \times 7}$ （其表现出竞争力的性能，比照表 1）。我们超

越 BM3D 以更快的运行与我们的 $CSF_{5 \times 5}^4$ 模型（cf. Tab. 2）。

此外，我们的模型的推理过程是非常适合的 GPU 或 DSP 的并行化。为了评估潜在的加速比，我们使用相同的代码与 Matlab 的内置 GPU 的能力，并能获得显著改善运行时间（表 3）。但是，我们应该期待更多的加速通过使用更强大的 GPU 最近使用 CUDA 或 OpenCL 的优化实现（[3]举一个 $40 \times$ GPU 加速了大概多线程 CPU 代码）。

而 ARF 模型[2]（设计用于实时去噪）是更有效的（仅 CPU）的比我们 CSF 阶段，过滤器和过滤器大小相同数量的，它显示出较差的结果：它执行 0.38 分贝逊于 $CSF_{5 \times 5}^4$ （表 1），甚至我们 $CSF_{3 \times 3}^4$ 模型，只有 8 3×3 过滤器优于 ARF 的修复质量方面。而 ARF 是尽可能快地 $CSF_{3 \times 3}^4$ 两次，我们可以通过重新使用过滤器（cf. Sec. 3.1）加快 CSF 起来。



Method	128 ²	256 ²	512 ²	1024 ²	2048 ²	4096 ²
CSF _{pw} ⁴	0.02	0.05	0.17 (0.03)	0.7 (0.05)	3.4 (0.18)	15 (0.8)
CSF _{3x3} ⁴	0.10	0.22	0.76 (0.15)	3.0 (0.27)	14.6 (0.78)	65 (3.7)
CSF _{5x5} ⁴	0.34	0.80	2.78 (0.44)	11.5 (0.80)	52.0 (2.42)	223 (10.8)
CSF _{7x7} ⁴	0.86	2.23	8.00 (0.92)	32.3 (1.72)	143 (5.27)	603 (23.2)
ARF-4 [2]	0.03	0.09	0.37	1.5	7.5	29
BM3D [5]	0.18	0.92	4.09	18.0	78.9	330
opt-MRF [3]	4.73	21.7	108	538	-	-
RTF ₄ [24]	6.71	27.7	113	469	-	-
EPLL [31]	9.76	41.9	229	930	-	-

Figure 3. Runtime comparison for image denoising. Single threaded runtimes (in seconds) with an Intel Core i7-3930K CPU at 3.20GHz; small numbers in parentheses from simple Matlab based GPU execution on a NVIDIA GeForce GTX 480. Runtime of our models shown after 4 stages where performance saturates; using fewer stages takes proportionally less time, e.g. 2 stages take half the time. Note the logarithmic scales on both axes (top). The table columns show runtimes for image sizes up to 4096×4096 pixels (about 16.78 megapixels).

Blur kernel	[16]	[25]	CSF _{pw} ¹	CSF _{pw} ²	CSF _{pw} ³
Ground truth	32.73	33.97	32.48	33.50	33.48
Levin et al. [17]	30.05	30.40	29.63	30.34	30.42
Cho and Lee [4]	29.71	29.73	29.10	29.86	29.99
Fergus et al. [8]	28.38	29.10	28.36	29.02	29.01

Table 3. Average PSNR (dB) on 32 images from [17] for image deconvolution. Rows correspond to different blur kernel (estimates) provided by [17], while columns correspond to non-blind deconvolution methods. Left part of table quoted from [25], showing results from Levin et al. [16] and Schmidt et al. [25].

此外，我们的标准基于梯度的学习过程更容易和更快，而且能使学习更强大的模型，如CSF_{7x7}⁵。

计算学习目标函数 $j(\Theta)$ (公式 12)

和它的梯度 $\partial j(\Theta) / \partial \Theta$ 为 $S = 400$ 图像 256×256 像素需要总共只有大约 7 秒 (CSF_{pw})，24S (CSF_{3x3}⁴)，73S

(CSF_{5x5}⁴)，或 161S (CSF_{7x7}⁵) 我们

简单 matlab 实现 (英特尔酷睿 i7-3930K 六核的 3.20GHz，六并行线程)。这使我们能够通过使用 200L-BFGS 迭代彻底训练我们的模型。我们的方法的另一个重要特性是它的可预测的运行时间，这是相对于需要迭代推理的收敛方法 (如选择性加入 MRF 和 RTF) 取决于输入的数据。根据我们的经验，

运行时间变化更对于去卷积，这主要是由于该模糊核。

联合训练。而联合训练模型的各个阶段都有产生更好的结果的潜力，我们只部分证实了这一点在我们的去噪实验。由于我们学习的目标函数是不是凸，优化常常被卡在更糟糕的局部最优利用贪婪训练的时候比。因此，我们尝试了第一次训练的每个阶段贪婪 (预培训)，然后在“调整”的模式，方法是从训练前获得的参数，联合训练。虽然这是保证不会降低 (训练集) 的性能，它不总是改善效果很大，尤其是与增加的模型容量。联合调整

CSF_{pw}⁵ 的所有 5 个阶段。并以 27.67 分贝还清，约 0.31 分贝从 27.36 分贝提高 PSNR 性能 (cf. Tab. 2)。然而，调整我们的其他车型全部 5 个阶段很难有差别。连服 3 级调整，我们观察到只有轻微的改进，如从 28.02 分贝到 28.09 分贝为 CSF_{3x3}⁴，并从 28.34 分贝到 28.36 分贝为 CSF_{5x5}³。

非盲解卷积。其结果于表 3 显示，我们的方法也可以成功地应用到图像解卷积的盲去模糊，其中内核的估计被用于去模糊的图像的情况下。对于反褶积的任务，我们培养了 CSF_{pw}。模型尺寸 288 综合图像模糊 320×320 像素。一半的模糊训练图像中，我们使用的估计，而不是正确的模糊核 k 以应付使用错误的内核的估计在测试时间 (所建议的 [25])。我们 CSF_{pw}⁵ 模型优于由莱文等人的非盲解卷积方法

[16], 并可以与来自 Schmidt 等人的结果竞争[25]对于所有估计内核(表 3)。我们还应用了同样的教训 CSF_{pw}^5 模型近期基准 Kohler 等的相机抖动 [14], 在那里我们能够改善后表现最好的方法的结果徐贾[30]关于 0.56 分贝平均, 是以 0.15 分贝比[25]最好的结果更好。还原的每个基准的 800×800 大小的彩色图像只需要大约与我们的第二模型。

参考

- [1] A. Antoniadis and J. Fan. Regularization of wavelet approximations. *Journal of the American Statistical Assoc.*, 96(455):939–967, 2001.
- [2] A. Barbu. Training an active random field for real-time image denoising. *IEEE TIP*, 18(11):2451–2462, Nov. 2009.
- [3] Y. Chen, T. Pock, R. Ranftl, and H. Bischof. Revisiting loss-specific training of filter-based MRFs for image restoration. In *GCPR 2013*.
- [4] S. Cho and S. Lee. Fast motion deblurring.

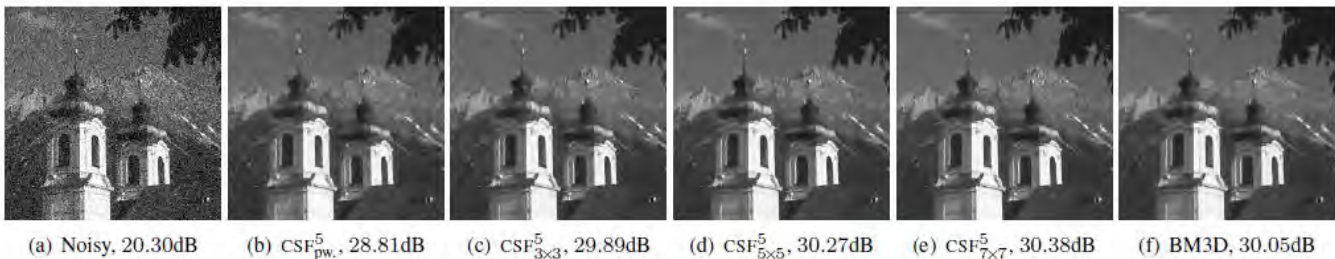


Figure 4. Denoising example ($\sigma = 25$, cropped): Comparison of our trained models and BM3D [5]. Best viewed magnified on screen.

5. 结论和未来工作

我们提出的收缩领域, 适用于恢复的高清晰度图像, 这是基于半二次优化的添加剂形式的扩展的新型随机场模型。通过与收缩功能的替代潜力, 我们增加了模型的灵活性和使所有模型参数的高效学习。图像去噪及解卷积级联收缩领域的实验表明, 运行速度快, 高品质的恢复可以手动完成。

今后的工作。下一步是一种更有效的 GPU 实现, 以进一步提高运行时为大的图像尺寸。另一个方向是培养我们的模型对其他图像质量指标[如 29]。最后, 可以进一步研究非单调收缩功能或探索更强大的多元收缩运营商。

- ACM T. Graphics, 28(5), 2009.
- [5] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE TIP*, 16(8):2080–2095, Aug. 2007.
- [6] DxO Image Science. Optics Pro 9: PRIME. <http://download-center.dxo.com/Press/opticspro/v9/pr/PR-DxO-Optics-Pro-9-EN.pdf>, Oct. 2013.
- [7] M. Elad, B. Matalon, and M. Zibulevsky. Image denoising with shrinkage and redundant representations. In *CVPR 2006*.
- [8] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman. Removing camera shake from a single photograph. *ACM T. Graphics*, 3(25), July 2006.

- [9] Q. Gao and S. Roth. How well do filter-based MRFs model natural images? In *Pattern Recognition (DAGM) 2012*.
- [10] D. Geman and G. Reynolds. Constrained restoration and the recovery of discontinuities. *TPAMI*, 14(3):367–383, 1992.
- [11] D. Geman and C. Yang. Nonlinear image recovery with halfquadratic regularization. *IEEE TIP*, 4(7):932–946, 1995.
- [12] Y. Hel-Or and D. Shaked. A discriminative approach for wavelet denoising. *IEEE TIP*, 17(4):443–457, 2008.
- [13] J. Jancsary, S. Nowozin, and C. Rother. Loss-specific training of non-parametric image restoration models: A new state of the art. In *ECCV 2012*.
- [14] R. Köhler, M. Hirsch, B. Mohler, B. Schölkopf, and S. Harmeling. Recording and playback of camera shake: Benchmarking blind deconvolution with a real-world database. In *ECCV 2012*.
- [15] D. Krishnan and R. Fergus. Fast image deconvolution using hyper-Laplacian priors. In *NIPS*2009*.
- [16] A. Levin, R. Fergus, F. Durand, and W. T. Freeman. Image and depth from a conventional camera with a coded aperture. *ACM T. Graphics*, 26(3):70:1–70:9, July 2007.
- [17] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman. Efficient marginal likelihood optimization in blind deconvolution. In *CVPR 2011*.
- [18] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman. Non-local sparse models for image restoration. In *ICCV 2009*.
- [19] D. Martin, C. Fowlkes, D. Tal, and J. Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *ICCV 2001*.
- [20] J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli. Image denoising using scale mixtures of Gaussians in the wavelet domain. *IEEE TIP*, 12(11):1338–1351, Nov. 2003.
- [21] S. Roth and M. J. Black. Fields of experts. *IJCV*, 82(2):205–229, Apr. 2009.
- [22] K. G. G. Samuel and M. F. Tappen. Learning optimized MAP estimates in continuously-valued MRF models. In *CVPR 2009*.
- [23] M. Schmidt. minFunc. <http://www.di.ens.fr/~mschmidt/Software/minFunc.html>, 2012.
- [24] U. Schmidt, J. Jancsary, S. Nowozin, S. Roth, and C. Rother. Cascades of regression tree fields for image restoration. *arXiv:1404.2086*, 2014.
- [25] U. Schmidt, C. Rother, S. Nowozin, J. Jancsary, and S. Roth. Discriminative non-blind deblurring. In *CVPR 2013*.
- [26] E. P. Simoncelli. Bayesian denoising of visual images in the wavelet domain. In *Bayesian Inference in Wavelet-Based Models*, volume 141 of *Lecture Notes in Statistics*, pages 291–308. Springer, 1999.
- [27] M. Tappen, C. Liu, E. H. Adelson, and W. T. Freeman. Learning Gaussian conditional random fields for low-level vision. In *CVPR 2007*.
- [28] Y. Wang, J. Yang, W. Yin, and Y. Zhang. A new alternating minimization algorithm for Total Variation image reconstruction. *SIAM J. Imaging Sciences*, 1(3):248–272, 2008.
- [29] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: From error visibility to structural similarity.

IEEE TIP, 13(4):600–612, Apr. 2004.

[30] L. Xu and J. Jia. Two-phase kernel estimation for robust motion deblurring. In *ECCV 2010*.

[31] D. Zoran and Y. Weiss. From learning models of natural image patches to whole image restoration. In *ICCV 2011*.

