PROCEEDINGS A

rspa.royalsocietypublishing.org

研究

已投稿至期刊

主题领域：机器学习，信号处理

关键词：高斯过程、反卷积、贝叶斯推断

通讯作者：Felipe Tobar电子邮箱：ftobar@uchile.cl

arXiv:2305.04871v2 [stat.ML] 9 May 2023

高斯过程反卷积

Felipe Tobar[1]、Arnaud Robert[2] 和 Jorge F. Silva[3]

[1]数据与人工智能倡议，智利大学[2]伦敦帝国理工学院计算系[3]智利大学电气与电子工程系

我们考虑反卷积问题，即从卷积过程 $y = x \star h + \eta$ 的观测值 $\mathbf{y} = [y_1, \ldots, y_N]$ 中恢复潜在源 $x(\cdot)$，其中 η 是加性噪声，$\mathbf{y}$ 中的观测值相对于 y 可能存在缺失部分，并且滤波器 h 可能未知。当 x 是连续时间信号时，我们提出一种解决该任务的新策略：对信号源 x 采用高斯过程（GP）先验，从而实现闭式贝叶斯非参数反卷积。我们首先分析直接模型，以确立模型定义良好的条件。随后转向逆问题，研究 i）贝叶斯反卷积可行所需的一些必要条件，以及 ii）在盲反卷积情况下，滤波器 h 在多大程度上可以从数据中学习或进行近似。所提出的方法称为高斯过程反卷积（GPDC），我们从概念上、通过说明性示例以及利用真实世界数据集，将其与其他反卷积方法进行了比较。

这是作者生成的作者定稿，即已被接收但尚未由期刊排版的稿件版本，用于在个人或公共知识库中分享。

© 作者。由英国皇家学会依据知识共享署名许可协议 http://creativecommons.org/licenses/by/4.0/ 发布，该协议允许不受限制地使用，但须注明原作者和信号源。

THE ROYAL SOCIETY PUBLISHING

1. 引言

在信号处理领域，连续时间信号源 x 与滤波器 h 之间的卷积记作[1]

$$f(t) = x \star h = \int_{\mathbb{R}} x(\tau)h(t-\tau)\mathrm{d}\tau, \tag{1.1}$$

可以理解为通过具有脉冲响应 h 的采集设备对 x 进行的广义（无噪声）观测。在这里， h 反映了观测设备的质量或精度，因为 h 越接近狄拉克 δ 函数，卷积后的量 f 就越接近信号源 x。

这一卷积模型，以及从一组含噪观测值中恢复信号源的需求 x ， f出现在许多场景中，包括：天文学 [1]、电信中的信道均衡 [2]、、去混响 [3]、、地震波重建 [4]，以及图像复原 [5]等。在这些场景中，实践者需要去除由非理想滤波器引入的不必要伪影 h，换言之，他们需要执行反卷积，以恢复 x 自 f。实践中，反卷积处理的是受噪声污染的观测值（由感测过程导致），这些观测值是 f的实现，在本文设定中记为 $\mathbf{y}$ 。此外，我们将事实上 假设存在缺失观测值，因为我们的表述将信号源 x 和卷积 f 视为连续时间对象，而观测值$\mathbf{y}$ 始终是有限的，这意味着观测值中存在“缺失部分”。图 1 展示了使用所提出方法处理一只鸟的图像的过程。

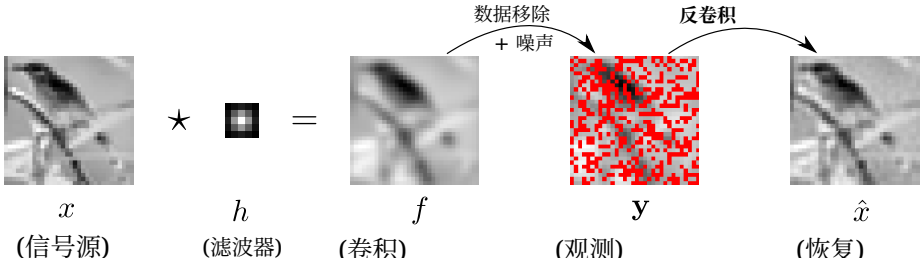


图 1：反卷积操作示意图：图像 x 与滤波器 h 进行卷积，生成模糊图像 f，随后图像受到破坏（红色部分表示加性噪声和数据移除），得到$\mathbf{y}$。反卷积旨在根据$\mathbf{y}$ 生成原始信号 x 的估计值，记为$\hat{x}$。此图像由所提出的方法生成。

作为逆问题的一种具体情形，受损信号的反卷积已在很大程度上从统计学角度得到研究 [6–8]。这种方法与线性滤波器理论密切相关，而由 [9,10] 等人奠定的基础至今仍是现代反卷积实现的核心。在浩瀚的反卷积文献中，一个普遍的结论是：在存在噪声或缺失数据时，反卷积前建立一个恰当的模型来重建 f （依据$\mathbf{y}$）至关重要。从贝叶斯角度来看，这归结为恰当选择信号源 x的先验分布；我们通过对信号源施加高斯过程先验来处理这一问题。

（a）贡献与组织结构

尽管反卷积问题无处不在，且备受关注， [11]，我们认为，从贝叶斯角度出发，根据有限数量的受噪（卷积后）样本恢复连续时间信号，**并为反卷积提供误差棒**，在很大程度上仍未得到充分探索。针对这一挑战，我们研究贝叶斯非参数[1]为简化记法，我们使用紧凑表达式 x， h 和 f 分别表示函数 $(x(t))_{t\in\mathbb{R}}$、$(h(t))_{t\in\mathbb{R}}$ 和 $(f(t))_{t\in\mathbb{R}}$ 。

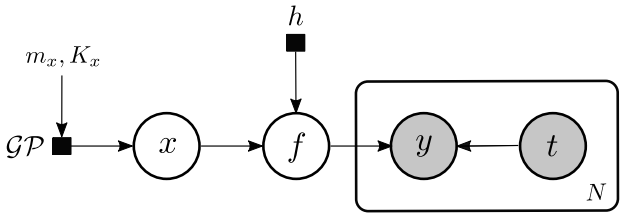


图2：层次高斯过程卷积的图模型。请注意，观测值和潜在随机变量分别以灰色圆圈和白色圆圈表示，而黑色方块表示固定参数。

我们的方法在潜在源 x 上采用高斯过程（GP）先验进行反卷积，因此称为高斯过程反卷积（GPDC）。本研究的主要贡献包括：i) 所提模型定义良好的条件及其样本生成方法；ii) 后验反卷积的闭式解，以及实现这种反卷积的条件；iii) 将其应用于盲反卷积及所需的近似；iv) 在一维和二维真实世界数据上对 GPDC 方法进行实验验证。

本文结构安排如下。第2节将卷积表示为高斯过程层次模型，并介绍相关文献。第3节研究直接模型（即 x 生成 f 和 y），并定义 h（t）以及 x 满足何种条件时，f（t）才能在逐点意义上定义良好。接着，第4节聚焦于逆问题（即从 y 估计 x），并从傅里叶表示的角度研究恢复。第5节讨论盲反卷积（即 h 未知），第6节和第7节分别介绍实验验证和结论。

2. 使用高斯过程的反卷积

（a）所提出的生成模型

考虑以下层次模型（图模型见图 2）：

$$\text{source process:} \quad x(t) \sim \mathcal{GP}(m_x(t), K_x(t)), \tag{2.1}$$

$$\text{convolution:} \quad f(t) = \int_{\mathbb{R}} x(\tau) h(\tau - t) \mathrm{d}\tau, \tag{2.2}$$

$$\text{observations:} \quad y_i \sim \mathcal{N}(f(t_i), \sigma_n^2), i = 1, \ldots, N, \tag{2.3}$$

其中，$\mathbf{t} = \{t_i\}_{i=1}^N \in \mathbb{R}^N$ 表示观测时间。首先，式（2.1）为信号源指定一个协方差为 K_x（t）的平稳高斯过程先验；我们假设对于所有 t，m_x（t）$= 0$。这一选择基于如下考虑：先验分布对于在缺失和噪声数据条件下实现反卷积至关重要，而且高斯过程具有众所周知的插值性质 [12]。其次，式（2.2）通过线性时不变滤波器 h定义连续时间卷积过程 f。第三，式（2.3）定义了高斯似然，其中，时刻 t 对 f 的含噪观测记为 $\mathbf{y} = [y_1, \ldots, y_N] \in \mathbb{R}^N$。请注意，观测值 y 只能看到 f（·）的“部分”，因此需要采用贝叶斯方法来量化与条件过程 x|y相关的不确定性。

（b）与经典方法和先前工作的关系

反卷积最简单的方法或许是逆傅里叶变换 方法，它通过对傅里叶变换逐点相除来执行反卷积 $\mathcal{F}^{-1}\{\mathcal{F}\{\mathbf{y}\}/\mathcal{F}\{\mathbf{h}\}\}$ ——其中$\mathbf{h}$ 是滤波器 h 的离散形式——以及维纳方法，该方法在均方误差意义上求出 x 的最优估计。这两种方法都假设滤波器已知，并在傅里叶域中进行除法运算，而这在实际应用中可能不稳定。最早将贝叶斯方法用于反卷积问题的研究可归功于 Bretthorst，他在 1992 年提出使用

在离散时间情形下采用多元正态（MVN）先验 [13]。一年后，Rhode 和 Whittenburg 研究了指数衰减正弦波的贝叶斯反卷积 [14],；随后，Sir David Mackay 在其著作中介绍了假设采用 MVN 先验时反卷积问题的贝叶斯推导，并将其与其他方法联系起来，重点讨论了图像复原 [15, 第 46]章。在统计学领域，反卷积指的是 x 和h 为概率密度而非时间序列的情形；这不在本文讨论范围之内。

在贝叶斯反卷积中使用不同的先验，是为了满足潜在时间序列 x的特定属性需求。例如，[16] 考虑了拉普拉斯先验混合，因此需要使用变分近似；而 [17] 则采用了全变分先验，旨在实现可靠的图像重建。然而，迄今为止，人们很少关注信号源 x 为连续时间信号的情况。现有的大多数方法仅针对离散时间的有限序列（或向量）设定先验；尽管通过时间量化，这些方法仍可应用于连续数据，但对于非均匀采样数据或密集时间网格，其计算复杂度会急剧增加。同样，克里金法相关文献通常会将其与高斯过程的相关文献进行比较，该领域也研究了反卷积问题（参见例如 [18,19]），但尚未从概率角度处理连续时间设置。此外，机器学习和计算机视觉领域近期的反卷积进展，主要集中于将神经网络模型与先验设计 [20] 和维纳滤波器 [21],等经典概念相结合，同时仍然考虑离散对象（而非连续时间），并且通常不允许存在缺失数据。

卷积模型主要被高斯过程领域采用，用于为多输出 [22–24]和非参数情形 [25,26]参数化协方差函数。随着深度学习的兴起，研究人员在高斯过程中复现了卷积神经网络（CNN）的卷积结构。例如，深度核学习（DKL） [27] 将卷积神经网络与高斯过程串联起来，使所得结构（同样是高斯过程）的核具备共享权重和偏置，这有助于检测图像不同区域中的共同特征，并增强其对物体平移的鲁棒性。另一个例子是卷积高斯过程 [28], ，它通过将核与图像块响应函数 串联来扩展深度核学习，使核具备卷积结构，而无需引入额外的超参数。这一概念已扩展至图和深度高斯过程——分别参见 [29,30] 。

虽然卷积在很大程度上推动了高斯过程核函数的设计，但沿着“相反方向”开展的研究，即利用高斯过程工具箱解决一般反卷积问题的研究，却寥寥无几。据我们所知，使用高斯过程先验进行贝叶斯反卷积的尝试仅有以下几类：专门研究如何从光谱偏振观测中检测磁信号 [31]；仅考虑离散时间脉冲响应[32]；或者，近期针对非平稳 Matérn 协方差这一特殊情况，使用多元正态先验 [33], 。这种参数化方法由 [34] 提出，尤其已用于散射雷达数据 [35]。

基于这些采用高斯过程进行反卷积的研究所获得的实验结果，我们的工作重点是分析和研究核函数与滤波器，考察它们恢复 $x(t)$ 的能力，其中 $\mathbf{y}$ 为观测信号。所提出的策略称为 GPDC，预计与以往方法相比具有更优越的建模能力，同时能够获得易于计算的闭式后验反卷积分布密度。然而，在实施 GPDC 之前，还有一些方面需要解决，具体包括：i）根据 的分布律，方程（2.2）中的积分何时有限；ii）何时能够从 x中恢复 x；iii）如何处理盲场景。以下各节将重点讨论这些问题。

3. 直接模型的分析

(a) 关于 x 和 h的假设及其对 f的影响

我们假设 x 是平稳高斯过程，且其协方差核 K_x 可积（即 $K_x \in L_1$）；这是确保 x 具有定义良好的傅里叶功率谱密度（功率谱密度）所必需的。此外，

尽管样本路径 x 通常不可积，但它们局部可积，这是因为 $K_x \in L_1$，因此，我们假设滤波器 h 衰减得足够快，使得式（2.2）中的积分有限。当 h 具有紧支撑，或其被拉普拉斯函数或平方指数函数所支配时，这一点总能得到保证。如果这些性质（可积性和稳态性）对 x成立，那么根据以下结果，它们也适用于 f。

注 1。 若 $f(t)$ 在公式 (2.2) 中对任意 $t \in \mathbb{R}$逐点有限，则 $f(t) \sim \mathcal{GP}(0, K_f(t))$ 具有平稳核

$$K_f(t) = \int_{\mathbb{R}^2} h(\tau')h(\tau)K_x(\tau - (\tau' - t))d\tau d\tau'. \tag{3.1}$$

引理 1。 若卷积滤波器 h 和协方差 K_x 均可积，则 K_f 公式 *(3.1)* 中的函数可积。

证明。 对 $K_f(t)$ 在公式 (3.1) 中的可积性，直接对 $\int_{\mathbb{R}} |K_f(t)|\mathrm{d}t$应用 Fubini 定理和三角不等式（两次）即可得到 $\int_{\mathbb{R}} |K_f(t)|\mathrm{d}t < \infty$ （完整证明见附录）。

□

注 2。 引理 2 给出了 关于 可积性的一个充分条件 K_f，因此从理论上证明了：*i*）可对 f进行基于傅里叶的分析；*ii*）可借助高斯过程工具，从贝叶斯推断的角度处理反卷积。不过，引理 2 中的条件 并非必要条件；例如，若滤波器 h 是（不可积的）*Sinc* 函数，则公式 (2.1)-(2.3) 中的生成模型仍然定义良好（见例2）。

(b) 从卷积模型中采样

我们提出两种从 f （见式（2.2））中采样的方法：我们可以 i）从 $x \sim \mathcal{GP}(0, K_x(t))$ 中抽取一条轨迹，并将其与 h进行卷积；或者 ii）直接从 $f \sim \mathcal{GP}(0, K_f(t))$中采样。然而，这两种方法都存在严重局限。第一种方法无法通过数值方式实现，因为卷积结构意味着 f 中的每个元素都依赖于 x 的无限多个值（对于一般的 h）。第二种方法通过直接对 f的有限版本采样，避开了这一困难；然而，这样一来， x 就被积分消去了，这意味着我们不知道 x 的哪一条样本轨迹对应于 f 这些（有限个）样本。

联合采样信号源过程和卷积过程的有限版本（亦称边缘化版本）的关键，在于这样一个事实：尽管式（2.2）确立的 x 与 f 之间的关系是确定性的，但它们有限版本之间的关系却变成了随机的。我们将 x 和 f 的有限版本 [2] 分别记为 $\mathbf{x} = x(\mathbf{t}_x)$ 和 $\mathbf{f} = f(\mathbf{t}f)$，其中 $\mathbf{t}_x \in \mathbb{R}^{N_x}$ 和 $\mathbf{t}_f \in \mathbb{R}^{N_f}$。因此，我们可以按照 $p(\mathbf{x}, \mathbf{f}) = p(\mathbf{f}|\mathbf{x})p(\mathbf{x})$ 分两个阶段进行分层采样：首先采样 $\mathbf{x} \sim \mathrm{MVN}(0, K_x(\mathbf{t}_x))$，然后在给定 $\mathbf{x}$ 的条件下采样 $\mathbf{f}|\mathbf{x} \sim p(\mathbf{f}|\mathbf{x})$，即 [3]

$$\begin{aligned} p(\mathbf{f}|\mathbf{x}) &= \mathrm{MVN}\left(\mu_{\mathbf{f}|\mathbf{x}}, \sigma^2_{\mathbf{f}|\mathbf{x}}\right) \\ \mu_{\mathbf{f}|\mathbf{x}} &= K_{fx}(\mathbf{t}_f, \mathbf{t}_x)K_x^{-1}(\mathbf{t}_x)\mathbf{x} \\ \sigma^2_{\mathbf{f}|\mathbf{x}} &= K_f(\mathbf{t}_f) - K_{fx}(\mathbf{t}_f, \mathbf{t}_x)K_x^{-1}(\mathbf{t}_x)K_{xf}(\mathbf{t}_x, \mathbf{t}_f), \end{aligned} \tag{3.2}$$

其中， $K_{fx}(t_1, t_2) = K_{fx}(t_1 - t_2)$ 表示 f 与 x 之间的平稳协方差，其逐元素定义为

$$K_{fx}(t_1, t_2) = K_{fx}(t_1 - t_2) = \int_{\mathbb{R}} h(t - (t_1 - t_2))K_x(t)\mathrm{d}t. \tag{3.3}$$

式 K_{fx} 中的（3.3）的可积性与引理 2中 K_f 的可积性类似地得到，假设 h， $K_x \in L_1$。

[2]在此，我们采用紧凑记法 $x(\mathbf{t}) = [x(t_1) \ldots, x(t_n)]$, 其中 $\mathbf{t} = [t_1, \ldots, t_n]$。 [3]我们采用反向参数记法 $K_{xf}(\mathbf{t}_x, \mathbf{t}f) = K_{fx}^{\top}(\mathbf{t}f, \mathbf{t}_x)$，以避免使用转置。

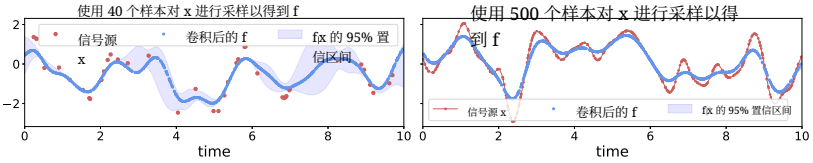


图 3：分层采样**x** 和**f**，对于区间 [0,10]中的**x**，分别考虑 40 个（左）和 500 个（右）样本。 K_x 和 h （以及因此的 K_{xf}和 K_f）均为长度尺度为 $l=\sqrt{0.05}$ （速率为 $\gamma=10$）的平方指数核。

此外，我们回顾一下， $f(t)|\mathbf{x}$ 的条件均值由下式给出

$$\mathbb{E}\left[f(t)|\mathbf{x}\right]=\int_{\mathbb{R}} h(t+\tau)\mathbb{E}\left[x(\tau)|\mathbf{x}\right]\mathrm{d}\tau. \tag{3.4}$$

该表达式表明， $f(t)$ 在给定 **x** 时的期望值，可以通过以下步骤计算：首先求出 x 在给定 **x**时的平均插值，记作 $\mathbb{E}\left[x(\tau)|\mathbf{x}\right]$,，然后按照式（2.2）对该插值进行卷积。

此外，我们还回顾一下， $f(t)|\mathbf{x}$ 的条件方差由下式给出

$$\mathbb{V}\left[f(t)|\mathbf{x}\right]=\int_{\mathbb{R}^2} h(t+\tau)\mathbb{V}\left[x(\tau),x(\tau')|\mathbf{x}\right]h(t+\tau')\mathrm{d}\tau\mathrm{d}\tau'.$$

注意，由于高斯过程的后验方差会随着观测值数量的增加而减小，因此当 $\mathbf{t}_x$ 变得越来越密集时，$\mathbb{V}\left[x(\tau)\ x(\tau')|\mathbf{x}\right]$ 趋近于零，相应地，$\mathbb{V}\left[f(t)|\mathbf{x}\right]$也会趋近于零。因此，**x** 中的元素越多，方差 $\mathbb{V}\left[f(t)|\mathbf{x}\right]$ 就越小， $f|\mathbf{x}$ 的轨迹也就越集中于式（3.4）中的后验均值附近。这与缺失源数据情况下的离散卷积有关：此时会“先插值，再卷积”，这凸显了对 x进行插值（即先验）的重要性。定义 1 给出了平方指数（SE）核，随后例 1 展示了采样过程，以及 $f|\mathbf{x}$ 如何集中于其均值附近。

定义 1。 定义为平稳核

$$K_{SE}(t)=\sigma^2\exp\left(-\frac{1}{2l^2}t^2\right) \tag{3.5}$$

称为平方指数（*SE*），其参数为幅度 σ 和长度尺度 l。此外，平方指数核的长度尺度也可以用其逆长度尺度（或速率） $\gamma=\frac{1}{2l^2}$来定义。平方指数核的傅里叶变换同样是平方指数核，因此，使用平方指数核的高斯过程的功率谱密度的支撑集为整个实数轴。

例 *1*。设 K_x 和 h 为长度尺度为 $l=\sqrt{0.05}$ （速率为 $\gamma=10$）的平方指数核，则 K_f和 K_{xf} 也都是平方指数核。图 *3* 展示了在区间 $t\in[0,10]$上采样得到的 *x* 和 *f*：在左图（相应地，右图）中，我们对一个 *40* 维（相应地，*500* 维）向量 *x* 进行采样，并以红色显示，用其生成一个 *1000* 维向量 *f*，以蓝色显示；两幅图中还同时给出了 $p(f(t)|x)$ 的 *95%* 误差棒。请注意，在固定区间内，*x* 的维度越大，条件密度 $p(f(t)|x)$ 就越集中。

注释 3。卷积过程的高斯过程表示使我们能够分别生成无限维过程 x 和 f 的有限样本**x**、**f** （在任意输入上）。这代表了经典离散时间卷积的连续概率对应形式：*i)* 可通过计算实现；*ii)* 适用于数据可能经过非均匀采样的现实场景。

4. 逆问题：贝叶斯非参数反卷积

现在，我们转而研究从有限向量 x 中恢复 $\mathbf{y}=[y_1,\ldots,y_N]$（在时间点 $\mathbf{t}=[t_1,\ldots,t_N]$处观测得到）的逆问题。在式（2.1）至（2.3）所定义的设定中，反卷积 $x|\mathbf{y}$是一个高斯过程，定义如下：

$$x|\mathbf{y}\sim\mathcal{GP}(\mu,\sigma^2) \tag{4.1}$$

$$\mu(t)=K_{xy}(t,\mathbf{t})K_{\mathbf{y}}^{-1}\mathbf{y}, \tag{4.2}$$

$$\sigma^2(t_1,t_2)=K_x(t_1,t_2)-K_{xy}(t_1,\mathbf{t})K_{\mathbf{y}}^{-1}K_{yx}(\mathbf{t},t_2), \tag{4.3}$$

其中，$K_{\mathbf{y}}=K_f(\mathbf{t},\mathbf{t})+\mathbf{I}_N\sigma_n^2$是观测值向量$\mathbf{y}$的边际协方差矩阵，$\mathbf{I}_N$是 N阶单位矩阵，而 $K_{xy}(t_1,t_2)$表示 $x(t_1)$与 $y(t_2)$之间的交叉协方差，它又等于 $K_{xf}(t_1,t_2)=K_{xf}(t_1-t_2)$。

（a）何时可以进行反卷积?

我们的目标是确定在什么情况下，$\mathbf{y}$ 能够提供足够的信息来确定x，并尽可能降低不确定性。为此，我们考虑 $x|\mathbf{y}$（式（4.1）中的高斯过程）的窗口化（傅里叶）谱表示：

$$\hat{x}_w(\xi)=\mathcal{F}\{w(t)x(t)\}(\xi)=\int_{\mathbb{R}}w(t)x(t)e^{-j2\pi\xi t}\mathrm{d}t, \tag{4.4}$$

其中，$\mathcal{F}\{\cdot\}$ 表示连续时间傅里叶变换（FT）算子，$\xi\in\mathbb{R}$ 是频率变量，而窗口 $w:\ \mathbb{R}\to\mathbb{R}$ 具有紧支撑（因此上面的积分有限）。选择这种窗口化谱表示，是因为高斯过程 $x|\mathbf{y}$ 是非平稳的（其功率谱密度未定义），且遵循分布 $x|\mathbf{y}$ 的样本轨迹$\mathcal{GP}(\mu,\sigma^2)$通常并非勒贝格可积（无法计算其傅里叶变换）。基于式（4.4）中引入的表示，我们定义成功恢复的概念，并将其与 K_x 和 h的谱表示联系起来。

定义 2. 我们称 $x(\cdot)$ 可以从观测值 中成功恢复 $\mathbf{y}=[y_1,\ldots,y_N]$，如果 $\forall\xi\in\mathbb{R}$，则式 $\hat{x}_w(\xi)$ 中的谱表示(4.4) 可以通过增加观测值的数量，以任意精度恢复，其中窗口 N为任意紧支撑窗口 $w(\cdot)$。

定理 1。 对于方程 (2.1)-(2.3)所定义的情形，要成功地从 x 中恢复$\mathbf{y}$，一个必要条件是 h 的谱支撑包含 K_x的谱支撑，即 $\operatorname{supp}\mathcal{F}\{K_x\}\subseteq\operatorname{supp}\mathcal{F}\{h\}$。

证明。 由方程（2.1）-（2.3）以及傅里叶变换的线性性质可知，窗口化频谱 $\hat{x}_w(\xi)$ 是一个复值的[4] 高斯过程，其均值和边缘方差为（计算见附录）：

$$\mu_{\hat{x}_w}(\xi)=\left(\hat{m}_x(\xi)+\hat{K}_x(\xi)\hat{h}(\xi)e^{-j2\pi\mathbf{t}\xi}K_{\mathbf{y}}^{-1}\mathbf{y}\right)\star\hat{w}(\xi), \tag{4.5}$$

$$\sigma^2_{\hat{x}_w}(\xi)=\hat{K}_x(\xi)\star|\hat{w}(\xi)|^2 - \|\hat{w}(\xi)\star\hat{K}_x(\xi)\hat{h}(\xi)e^{-j2\pi\xi\mathbf{t}}\|_{K_{\mathbf{y}}^{-1}}, \tag{4.6}$$

其中，$\|\mathbf{v}\|_A=\mathbf{v}^\top A\mathbf{v}$ 表示 $\mathbf{v}$ 关于矩阵 A 的马氏距离，而 $\hat{}\,h=\mathcal{F}\{h\}$ 被用作 h 的傅里叶变换的紧凑记法。根据定义 2，当$\hat{x}_w(\xi)$ 的后验方差趋于零时，即式 (4.6) 右侧的两项相互抵消时，便实现了成功恢复。要使这一情况发生，这两项必须具有相同的支撑集。然而，由于这对任意窗口 $\hat{}\,w(\xi)$ 都应成立，我们要求 $\operatorname{supp}\hat{}\,K_x(\xi)=\cup_{i=1}^N\operatorname{supp}\hat{}\,K_x(\xi)\hat{}\,h(\xi)e^{-j2\pi\xi t_i}$。由于 $\cup_{i=1}^N\operatorname{supp}e^{-j2\pi\xi t_i}=\mathbb{R}$，真正需要的是 $\operatorname{supp}\hat{}\,K_x(\xi)=\operatorname{supp}\hat{}\,K_x(\xi)\hat{}\,h(\xi)$，而当 $\operatorname{supp}\hat{}\,K_x(\xi)\subseteq\operatorname{supp}\hat{}\,h(\xi)$ 时，这一条件便成立。 □

[4]关于复值高斯过程，请参见 [36–38]。

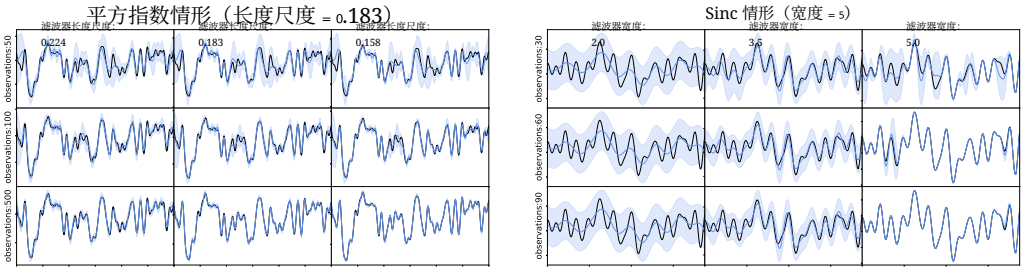


图 4：当 h 和 K_x 均为平方指数（左）或 Sinc（右）时，GPDC 的实现。每种情形下，我们都考虑了逐渐增加观测值数量 $\mathbf{y}$ （从上到下），以及h 的不同参数（从左到右）——所有宽度和长度尺度参数的取值均标注在图标题中。图中显示了真实源信号 x （黑色）、GPDC 后验反卷积均值（蓝色）以及 GPDC 的 95% 误差棒（浅蓝色）。与定理 1 和备注 4一致，请注意，在平方指数情形（左）中，增加观测值能改善恢复；而在 Sinc 情形（右）中，当滤波器过窄时（第一列和第二列），即使观测值数量不断增加，也无法实现恢复。

注 4。 定理 1 给出了从 x 中根据 $\mathbf{y}$ 的ˆ频谱内容中有多少 x（由 K_x表示）在与 h 进行卷积时未被抑制，从而可以从 $\mathbf{y}$中提取出来的必要 条件。成功恢复的更强充分条件显然取决于 $\mathbf{y}$ 的测量位置。直观地说，恢复取决于是否有足够的观测值，以及向量$e^{-j2\pi\xi\mathbf{t}}$ 是否与 $K_y^{-1}(\mathbf{t},\mathbf{t})$ 的特征向量对齐；该矩阵定义了式 (4.6)中的范数，因此类似于 *[39,40]*中的采样定理。

定义 3 介绍了 Sinc 核 [41]。随后，示例 2 阐释了定理1 和备注 4中的论断，其中 GPDC 恢复是根据频谱支撑和观测值数量来评估的。

定义 3. 定义如下的平稳核

$$K_{Sinc}(t) = \frac{\sigma^2 \sin(\Delta\pi t)}{\Delta\pi t} \tag{4.7}$$

称为（以零为中心的）*Sinc* 核，其参数为幅度 σ 带宽 Δ。虽然 *Sinc* 函数不可积（见引理2），但它存在傅里叶变换，形式为一个以零为中心、宽度为 Δ 的矩形。因此，使用 *Sinc* 核的高斯过程不会生成频率高于 $\Delta/2$的路径。

例 2。 我们来考虑 *GPDC* 的两种情形：第一种情形中，两个 K_x 和 h 都是平方指数（*SE*）核，因此 支撑集 $\mathcal{F}\{K_x\} =$ 支撑集 $\mathcal{F}\{h\} = \mathbb{R}$。第二种情形中，两个 K_x 和h 都是 *Sinc* 核 *[41]*，，其功率谱密度的重叠部分取决于它们的参数。我们针对不同数量的观测值和参数，在这两种情形下实现了 *GPDC*；具体而言，对于 *Sinc* 情形，我们考虑了 支撑集ˆ$K_x \not\subset$ 支撑集ˆh 的情形。图 4（左）显示，在 *SE* 情形中，无论 h 和 K_x的参数如何，增加观测值数量总能改善恢复效果。相反，请注意，对于 *Sinc* 情形（图 4，右），即使观测值数量很多，潜在源 x 仍无法恢复。这是因为卷积滤波器的窄带特性 会移除 x 的某些谱分量 h。所有幅度参数均设为 *1*；宽度（*Sinc*）和长度尺度参数（*SE*）见图 4 。

5. 盲反卷积情形

（a）训练与可观测性

实现 GPDC 需要选择超参数，即 K_x 和 h的参数，以及噪声方差 σ^2。根据具体应用，滤波器可能是已知的。例如，在去混响中， [42] h 由房间的几何形状决定；而在天文学中， h 由大气和/或望远镜决定 [43,44]。当 h 未知时，这种情况称为 盲反卷积，其参数必须与其他超参数一起从数据中学习。

我们通过最大化对数似然来拟合所提出的模型，见公式（2.1）至（2.3）

$$l(K_x, \sigma^2, h) = -\frac{n}{2}\log 2\pi - \frac{1}{2}\log\det K_{\mathbf{y}} - \frac{1}{2}\mathbf{y}^\top K_{\mathbf{y}}^{-1}\mathbf{y}, \tag{5.1}$$

其中 $K_{\mathbf{y}} = K_f(\mathbf{t},\mathbf{t}) + \mathbf{I}_N$，其中 $\mathbf{I}_N$是大小为 N的单位矩阵，而 K_f与 K_x之间的关系可由公式（3.1）得出。因此， K_x， h和 σ^2通过 l出现在公式（5.1）中。 $K_{\mathbf{y}}$。

然而，最大似然可能无法以唯一的方式恢复所有超参数，因为 h 和 K_x 在 $K_{\mathbf{y}}$ 中相互纠缠，因此可能无法辨识。例如，如果 h 和K_x 都是平方指数核（如例 1 所示），其长度尺度分别为 l_h 和 l_x，那么 $K_{\mathbf{y}}$ 也属于平方指数核，长度尺度为 $l_y = l_x + l_h$；因此，无法从学得的l_y中恢复原始长度尺度。更直观地说，卷积的不可观测性可以理解如下：给定的信号 f可能由以下任一方式生成：i）快速的信号源 x 与宽滤波器 h，或 ii）缓慢的信号源与窄滤波器。此外，仅凭 y 无法确定 h 的时间位置，因为似然函数对 x 和 h的时间平移不敏感。由于反卷积问题存在对称性，学习超参数时应尽可能利用关于 h 和 K_x的信息，尤其是在盲场景下。

(b) h 的易处理近似

GPDC 的闭式实现仅取决于能否成功计算公式 (3.1) 和 (3.3) 中的积分。对于某些 K_x 和 h的选择，例如定义 1 和 3 中的平方指数核或 Sinc 核，这些积分可以解析计算，但通常难以处理。虽然有人可能会认为计算这些积分会妨碍 GPDC 的广泛适用性，但请注意，当滤波器 h 表示为位于 $\{l_i\}_{i=1}^M$处的 Dirac δ 函数之和 $\{w_i\}_{i=1}^M$ 时，即:

$$h(t) = \sum_{i=1}^{M} w_i \delta_{l_i}(t), \tag{5.2}$$

协方差 K_f 和交叉协方差 K_{fx} 就会变为核函数计算值的求和:

$$K_f(t) = \sum_{i=1}^{M}\sum_{j=1}^{M} w_i w_j K_x(l_i - (l_j - t)) \tag{5.3}$$

$$K_{xf}(t_1, t_2) = K_{xf}(t_1 - t_2) = \sum_{i=1}^{M} w_i K_x(l_i - (t_1 - t_2)). \tag{5.4}$$

离散时间滤波器 h 本身就值得关注，尤其是在数字信号处理领域，但它也有助于使 GPDC 具备更广泛的适用性。这是因为，当公式 K_f、 K_{xf}中的积分形式（公式 (3.1) 和 (3.3）无法通过闭式形式计算时，可以采用不同的积分技术进行近似，例如求积方法、蒙特卡罗方法，甚至重要性采样（详细的近似方法见附录）。

在训练方面（盲反卷积所必需），离散时间滤波器的另一个优点是，其超参数（权重 $\{w_i\}_{i=1}^M$）不会与 $K_{\mathbf{y}}$纠缠在一起，而是各自以双线性形式出现在其中——见式（5.3）。这并不意味着参数是可辨识的，但双线性结构简化了优化。最后，请注意，由于 K_y 的样本近似（例如，使用求积或蒙特卡罗）维数较低

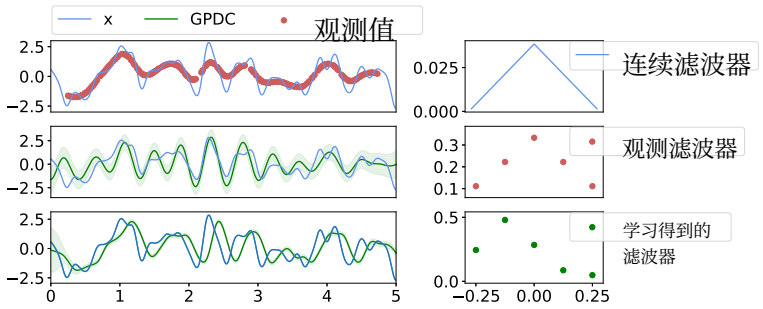


图5：采用离散近似滤波器的 GPCD h。上图：真实源信号 x、观测值 y 和滤波器 h。真实信号 x 由具有谱混合核 K_x（见定义4）的高斯过程生成，其参数为 $\sigma=1$、 $\gamma=50$、 $\nu=1$、 $\sigma_n=0.01$。中图：GPDC（绿色）使用仅根据真实 h 计算的离散滤波器，仅学习了 K_x；我们的方法恢复了 $\gamma=11.8$、 $\nu=1.9$ 和 σ_n =0.03，并且由于 h 与 K_x相乘所导致的不可观测性，幅度 σ 被固定为1。下图：盲情形下的后验反卷积：所有超参数、 h、 K_x 和σ^2均通过最大似然进行学习。学习得到的滤波器绘制在右下角（绿色圆点），恢复的参数为 $\gamma=38.8$、 $\nu=1e^{-6}$ 和 $\sigma_n=0.01$（σ 仍固定为1）。请注意，尽管参数值的恢复看似不理想，但盲情形下的反卷积（左下图，绿色）仍被正确检测，只是由于上述不可观测性问题，存在一个无法辨识的时间偏移。

与数据相比，虽然它们的阶数较高，但不会带来显著的计算开销，因为计算式（5.1）仍主要受通常的 GP 计算成本 $\mathcal{O}(N^3)$所主导。定义 4和 5分别介绍谱混合核和三角滤波器，随后示例 3 将说明如何按照上述描述，在盲情形和非盲情形下对训练 GPDC 进行离散化。

定义 4. 定义如下的平稳核

$$K_{SM}(t)=\sigma^2\exp\left(-\frac{1}{2l^2}t^2\right)\cos\left(2\pi\nu t\right) \tag{5.5}$$

称为（单分量）谱混合（*SM*），其参数为幅度 σ、长度尺度 l 和频率 ν。与定义 1中的平方指数核类似，*SM* 核的长度尺度可通过其速率 $\gamma=\frac{1}{2l^2}$来定义。*SM* 核的傅里叶变换是一个以频率 ν 为中心的平方指数核。

定义 5. 定义如下的函数

$$h_{Tri}(t)=\sigma^2\max\left(1-\frac{2|t|}{\Delta},0\right) \tag{5.6}$$

称为三角滤波器，也称为巴特利特窗，其参数为幅度σ 和宽度 Δ。

例3。 我们考虑以下情形： K_x 是一个平方指数核（见定义 1），而 h 是一个三角形（见定义 5）。我们不计算 K_f和 K_{xf}的积分，而是考虑两种情况：*i*）如第 5(a)节所述，通过最大似然为 h 和 K_x 学习离散近似；*ii*）根据真实的 h固定一个离散滤波器，并如第 5(b)节所述，仅训练 K_x。图 5 展示了这些实现。请注意，离散滤波器（仅有*5*个点）能够从 x 中合理地恢复潜在源$\mathbf{y}$。此外，根据 h的不可观测性，在盲情形下，真实信号源与反卷积结果之间存在一个不可辨识时滞。

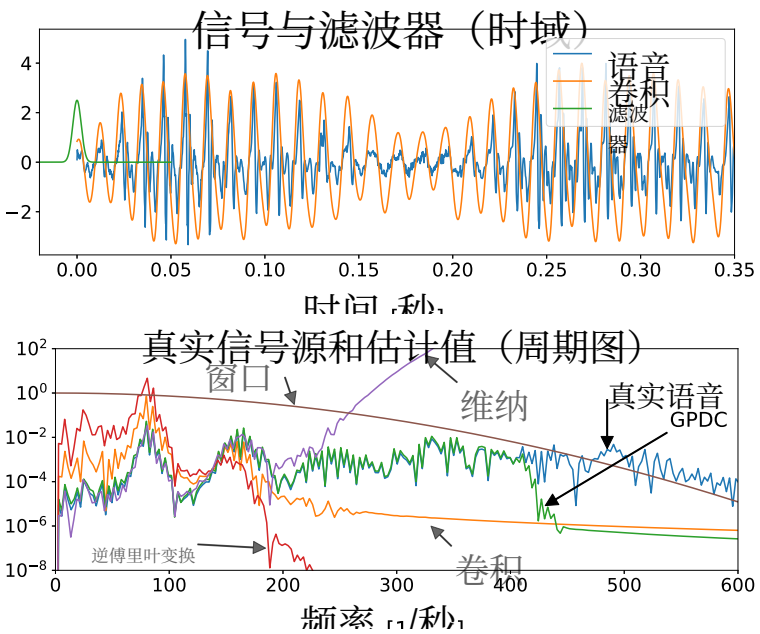


图6：使用GPDC、维纳方法和逆傅里叶变换方法恢复语音信号。信号显示在时域（顶部），估计结果显示在频域（底部）。

6. 实验

我们在两种实验设置下测试了 GPCD。第一种模拟了声学去混响场景，其目的是将 GPDC 在无噪声、无数据缺失情况下（即 $\sigma_n^2 = 0$）的点估计与标准反卷积方法进行比较，以验证其准确性。第二种实验研究了从低分辨率、含噪声且不完整的观测值中恢复潜在图像的问题，分别考虑滤波器 h 已知和未知的情况（即盲超分辨率）。以下实验以及论文中示例所使用的代码均可在 https://github.com/GAMES-UChile/Gaussian-Process-Deconvolution中找到。

（a）声学反卷积。

我们考虑一个 350 [毫秒] 语音信号（来自 http://www.mcsquared.com/reverb.htm），以 5512.5 赫兹重新采样（即 2000 个采样点），对其进行标准化，然后与长度尺度为 $l = 2.2$ [毫秒] 的平方指数滤波器进行卷积。图 6（上）展示了语音信号、滤波器及卷积结果。由于已知语音具有分段平稳特性，我们使用平方指数核实现 GPDC，然后按照第 5 节（a）中的说明，通过最大似然学习超参数。学习得到的超参数为 $\sigma = 0.95$，$l = 5 \cdot 10^{-4}$，$\sigma_{\text{noise}} = 0.12$。请注意，这些数值与数据预处理相符（由于信号经过标准化，方差之和约为 1），且长度尺度也与图 6（上）所示的情况一致。

将所提出的 GPDC 卷积方法（假设滤波器已知 h，该滤波器由上文介绍的平方指数核给出）与维纳反卷积方法和逆傅里叶变换方法进行了比较，相关方法见第2节。图6（下）展示了数据的周期图，以及维纳、逆傅里叶变换和 GPDC 的结果，还有滤波器 h。可以看到，在所考察的频谱范围内的大部分区域，GPCD 都忠实地跟踪了真实的潜在源；相比之下，维纳方法出现了偏离，而逆傅里叶变换的结果则迅速衰减，这可能是由数值不稳定性导致的。值得注意的是，GPDC估计恰好在滤波器开始衰减的频率处不再精确跟踪真实信号源，这与定理1一致。表1 展示了所考察方法在时域和频域中的性能，具体而言，通过均方误差（MSE）、KL散度

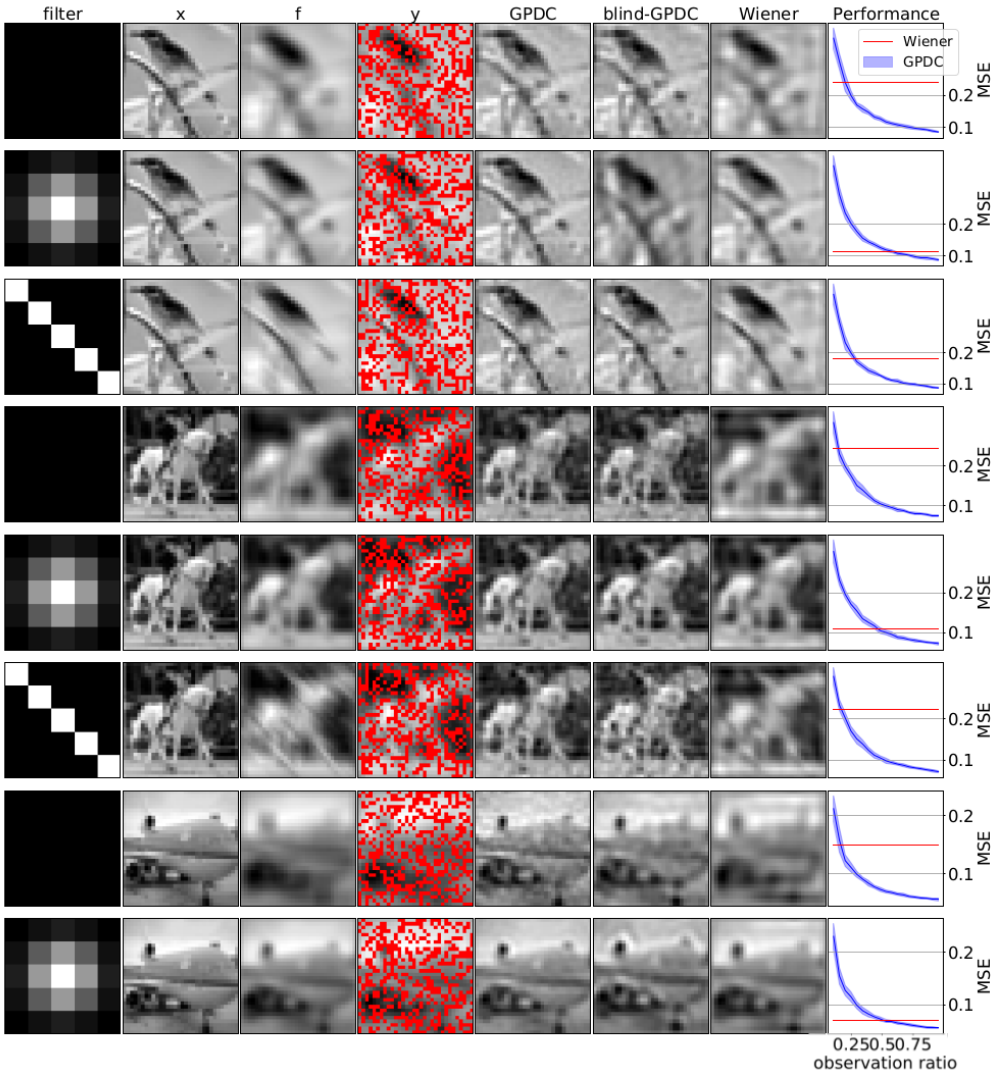


图 7：根据低分辨率、含噪声且不完整的观测值对 8 幅图像进行反卷积。这 8 幅图像由 3 幅源图像生成，并对其应用了所述 5 种滤波器之一。从左到右：使用的滤波器 h、真实图像 x、卷积 f、观测值 $\mathbf{y}$ （红色表示缺失像素占 25%）、GPDC估计、盲GPDC估计和维纳估计。最右侧的图展示了 GPDC（蓝色）和维纳滤波器（红色）的性能随观测像素数量的变化（GPDC 的误差棒为基于 20 次试验计算的 2 倍标准差）。

和 Wasserstein距离来衡量真实功率谱密度（PSD）与其估计值之间的差异 [45]（数值越低越好）。请注意，GPDC 在所有指标上均优于维纳方法和逆傅里叶变换方法。时域估计结果见附录中的其他图表。

(b) 盲图像超分辨率。

对于 32×32 图像 x，我们使用 5×5 滤波器 h创建了卷积图像 f，并获得了一个受噪声污染且存在缺失数据（保留了 60% 的像素）的观测结果 $\mathbf{y}$。我们实现了 GPDC，以便在 h **已知** 和 **未知**时，均能从 $\mathbf{y}$恢复 x。因此，对于每幅图像，我们考虑了 i）离散滤波器 h 已知并由 GPDC 使用的情况，以及 ii）从 h 中学习滤波器的情况。

度量	GPDC	维纳	逆傅里叶变换
均方误差（时间域）	**19.0**	35.0	41.9
均方误差（功率谱密度）	**0.015**	0.058	0.153
库尔贝克–莱布勒散度（功率谱密度）	**0.05**	0.20	0.45
瓦瑟斯坦距离（功率谱密度）	**2124.3**	3643.8	4662.6

表1：音频去混响实验在时间和频率域中的定量评估

通过最大似然估计数据。对于文稿主体（图 7），我们展示了使用三种不同滤波器（常量、单峰和对角）的结果，并在附录中展示了其他形状的结果。所有情形下，我们都假设信号源（真实图像）采用平方指数核，并针对每种情形学习长度尺度 l、幅度 σ 和噪声参数 σ_n。对于盲情形，离散滤波器 h 也通过学习得到。

将 GPDC 与（来自 `scikit-image` 的）维纳滤波器进行比较，该滤波器应用于**完整的、无噪声的**图像 f。之所以将所提出的 GPDC 与应用于真实图像 f 而非生成的观测值 $\mathbf{y}$ 上的标准维纳滤波器进行比较，原因有二：i）它构成了一个利用全部图像信息的基准；ii）维纳滤波器的标准实现无法处理缺失数据（如 $\mathbf{y}$ 中的缺失数据）。图7 展示了来自 CIFAR-10 [46] 的 3 幅不同图像的组合结果，以及针对每幅图像最多使用三种不同滤波器的结果。请注意，非盲和盲 GPCD 都能够提供比维纳滤波器更能代表真实图像（x）的重建结果。相对于维纳基准，GPDC 的优越性能通过相对于真实图像 x 的均方误差来衡量，该误差是观测数据量的函数（图 7中最右侧的图）。附录还提供了使用不同图像和滤波器的其他示例。

7. 讨论与结论

我们提出了高斯过程反卷积（GPDC），这是一种用于连续时间信号贝叶斯非参数反卷积的方法，其基础是对潜在信号源 x 采用高斯过程先验，因此我们可以为反卷积估计提供误差棒。我们研究了直接生成模型（其定义良好的条件以及如何从中采样）、逆问题（给出了成功恢复信号源的必要条件）、盲反卷积情形，并展示了若干示例，以及在去混响和超分辨率场景中的实验验证。通过在傅里叶域分析式（4.2），可以发现本方法与经典理论之间的一个关键联系：我们可以将 GPDC 解释为维纳反卷积的一种非参数、可处理缺失数据的扩展。从这个意义上说，所提出的方法为经典反卷积提供了高斯过程视角，从而以高斯过程的丰富资源扩充了反卷积工具箱，例如现成的协方差函数、稀疏近似以及大量专用软件。

当受限于离散且无缺失数据情形时，我们的方法会简化为标准的维纳反卷积，文献中称之为 [9]；因此，我们重点比较 GPDC 与维纳反卷积方法。然而，尽管可以从经典反卷积的角度理解 GPDC，我们所考虑的设置与数字通信中的设置截然不同。我们关注的是利用一组有限观测 x（信号源 x）来估计连续时间对象$\mathbf{y}$，这些观测值应理解为卷积过程 $f = x \star h$的受噪声污染且部分缺失的实现，并量化该估计的不确定性。由于我们对信号源的信息有限（x 只能通过$\mathbf{y}$ 根据公式（2.3）获知），我们采用贝叶斯方法，并在 x上施加先验，该先验是一个高斯过程。据我们所知，尚无先前工作从概念层面探讨这一设置

从理论角度探讨这一术语，但少数研究将其应用于具体问题（参见 [31–35]）。

就新颖的反卷积方法而言，文献中的趋势是将神经网络与经典概念（如先验设计）相结合 [20] 以及维纳滤波器 [21]，，并且主要应用于图像，这或许是因为深度学习在计算机视觉领域取得了强劲且持续的进展。然而，尽管近期已有一些关于高斯过程在特定设置下应用于反卷积问题的研究，据我们所知，目前尚无一种方法能够针对盲情形和非盲情形，从有限数量的含噪观测中恢复连续时间对象（对于图像而言，则为连续空间对象），同时具备我们所提出方法的两项主要特性：i）采用贝叶斯立场，从而能够确定误差棒；ii）提供第4(a)节中受傅里叶方法启发的保证。就此而言，我们声称，这是首个与具体应用无关的方法研究，确立了层次模型正确定义的条件、反卷积过程的能力，以及对不确定性进行有原则量化的方法。

我们希望研究结果能为其在不同科学领域的应用铺平道路，也能激励高斯过程领域的研究者进一步拓展本研究。具体而言，我们设想未来的研究工作可聚焦于以下方向：

- 提出反卷积的充分条件（而不仅仅是必要条件），以补充定理 1，从而衔接谱估计与高斯过程的交叉领域 [47]。
- 通过可训练的多分辨率卷积来控制时间相关性，从而构建稀疏高斯过程。这将提供计算效率高的稀疏高斯过程，并确保重建结果清晰。
- 开发所述方法的稀疏高斯过程版本，以便将 GPDC 用于大型数据集的反卷积

数据可获取性。本文使用的所有数据均为合成数据或公开数据。实验 1 使用的音频数据来自 `http://www.mcsquared.com/reverb.htm`，而实验 2 使用 CIFAR-10 [46]。

资助。我们感谢 Google 以及以下 ANID-Chile 资助项目提供的经费支持：Fondecyt- Regular 1210606、电气与电子工程高级中心（Basal FB0008）和数学建模中心（Basal FB210005）。

参考文献

1. Wang H, Sreejith S, Lin Y, Ramachandra N, Slosar A, Yoo S. 2022 基于神经网络的天文应用点扩散函数反卷积。*arXiv* 预印本 *arXiv:2210.01666*. 2. Clapp T, Godsill S. 1997 移动通信中的贝叶斯盲反卷积。载于*IEE*移动通信系统自适应信号处理专题讨论会（参考编号：*1997/383*）第 9/1–9/6 页。3. Willardson ML, Anderson BE, Young SM, Denison MH, Patchett BD. 2018 混响室中高振幅声音的时间反转聚焦。美国声学学会杂志**143**，696–705。4. Arya VK, Holden H. 1978 地震数据反卷积——综述。*IEEE*地球科学电子学汇刊**16**，95–98。5. Hong H, Shi Y. 2017 沿模糊路径进行运动模糊快速反卷积。加拿大电气与计算机工程杂志**40**，266–274。6. Tarantola A. 2005模型参数估计的逆问题理论与方法。SIAM。7. Stuart AM. 2010 逆问题：贝叶斯视角。《数值学报》**19**，451–559。8. Aster RC, Borchers B, Thurber CH. 2018参数估计与逆问题。爱思唯尔。9. Wiener N. 1964平稳时间序列的外推、插值与平滑。麻省理工学院出版社。10. Kalman RE. 1960 线性滤波与预测问题的新方法。基础工程学杂志**82**，35–45。11. Riad S. 1986 反卷积问题：综述。*IEEE*论文集**74**，82–85。12. Rasmussen C, Williams C. 2006面向机器学习的高斯过程。麻省理工学院出版社。

13. Bretthorst GL. 1992 贝叶斯插值与反卷积。技术报告，圣路易斯华盛顿大学化学系。14. Rhode C, Whittenburg S. 1993 贝叶斯反卷积。《光谱学快报》**26**，1085–1102。15. MacKay DJ. 2003 信息论、推断与学习算法。剑桥大学出版社。16. Adami KZ. 2003 贝叶斯反卷积中的变分方法。粒子物理、天体物理与宇宙学中的统计问题第 143 页。17. Babacan SD, Molina R, Katsaggelos AK. 2008 使用全变分先验的变分贝叶斯盲反卷积。*IEEE*图像处理汇刊 **18**，12–26。18. Jeulin D, Renard D. 1992 使用克里金法进行图像反卷积的实际局限。《显微镜学、微量分析与微结构》**3**，333–361。19. Goovaerts P. 2008 不规则地理单元条件下的克里金法与半变异函数反卷积。《数学地球科学》**40**，101–128。20. Ren D, Zhang K, Wang Q, Hu Q, Zuo W. 2020 使用深度先验的神经盲反卷积。载于*IEEE/CVF*计算机视觉与模式识别会议论文集第 3341–3350 页。21. Dong J, Roth S, Schiele B. 2020 深度维纳反卷积：维纳方法与深度学习相结合，实现图像去模糊。载于 Larochelle H, Ranzato M, Hadsell R, Balcan M, Lin H 编，《神经信息处理系统进展》第 33 卷，第 1048–1059 页。Curran Associates, Inc. 22. Boyle P, Frean M. 2005 相关高斯过程。载于 Saul LK, Weiss Y, Bottou L 编，《神经信息处理系统进展》第 *17* 卷，第 217–224 页。麻省理工学院出版社。23. Álvarez M, Lawrence ND. 2009 用于多输出回归的稀疏卷积高斯过程。载于 Koller D, Schuurmans D, Bengio Y, Bottou L 编，《神经信息处理系统进展》第 *21* 卷，第 57–64 页。Curran Associates, Inc. 24. Parra G, Tobar F. 2017 用于多输出高斯过程的谱混合核。载于 Guyon I, Luxburg UV, Bengio S, Wallach H, Fergus R, Vishwanathan S, Garnett R 编，《神经信息处理系统进展》第 30 卷。Curran Associates, Inc. 25. Tobar F, Bui T, Turner R. 2015 使用非参数核的高斯过程学习平稳时间序列。载于《神经信息处理系统进展》第 *28* 卷，第 3501–3509 页。Curran Associates, Inc. 26. Bruinsma W. 2016 广义高斯卷积过程模型。硕士论文，剑桥大学工程系。27. Wilson AG, Hu Z, Salakhutdinov R, Xing EP. 2016 深度核学习。载于 Gretton A, Robert CC 编，第十九届国际人工智能与统计学会议论文集第 51 卷《机器学习研究论文集》第 370–378 页，西班牙加的斯。PMLR。28. van der Wilk M, Rasmussen CE, Hensman J. 2017 卷积高斯过程。载于 Guyon I, Luxburg UV, Bengio S, Wallach H, Fergus R, Vishwanathan S, Garnett R 编，《神经信息处理系统进展》第 *30* 卷，第 2849–2858 页。Curran Associates, Inc. 29. Walker I, Glocker B. 2019 图卷积高斯过程。载于 Chaudhuri K, Salakhutdinov R 编，第三十六届国际机器学习会议论文集第 97 卷，第 6495–6504 页，美国加利福尼亚州长滩。PMLR。30. Blomqvist K, Kaski S, Heinonen M. 2019 深度卷积高斯过程。载于欧洲机器学习与数据库知识发现联合会议第 582–597 页。施普林格。31. Asensio Ramos A, Petit P. 2015 贝叶斯最小二乘反卷积。《天文学与天体物理学》**583**，A51。32. Tobar F, Rios G, Valdivia T, Guerrero P. 2017 使用高斯过程先验从测量混合中恢复潜在信号。*IEEE*信号处理快报**24**，231–235。33. Arjas A, Roininen L, Sillanpää MJ, Hauptmann A. 2020 盲层次反卷积。载于*IEEE*信号处理机器学习国际研讨会第 1–6 页。34. Paciorek CJ, Schervish MJ. 2006 使用一类新的非平稳协方差函数进行空间建模。《环境计量学：国际环境计量学会官方期刊》**17**，483–506。35. Ross S, Arjas A, Virtanen II, Sillanpää MJ, Roininen L, Hauptmann A. 2022 面向非相干散射雷达数据的层次反卷积。《大气测量技术》**15**，3843–3857。36. Boloix-Tortosa R, Murillo-Fuentes JJ, Payán-Somet FJ, Pérez-Cruz F. 2018 用于回归的复高斯过程。*IEEE*神经网络与学习系统汇刊**29**，5499–5511。

37. Ambrogioni L, Maris E. 2019 用于时间序列分析的复值高斯过程回归。信号处理**160**, 215 – 228. 38. Tobar F, Turner R. 2015 使用高斯过程对复值信号进行建模。载于*IEEE* 国际声学、语音与信号处理会议论文集，第 2209–2213 页。39. Shannon CE. 1949 噪声环境下的通信。无线电工程师学会会刊**37**, 10–21. 40. Nyquist H. 1928 电报传输理论中的若干专题。美国电气工程师学会汇刊**47**, 617–644. 41. Tobar F. 2019 带限高斯过程：Sinc 核。载于《神经信息处理系统进展》第 *32* 卷，第 12749–12759 页。Curran Associates, Inc. 42. Naylor PA, Gaubitch ND. 2010语音去混响。施普林格科学与商业媒体。43. Starck JL, Pantin E, Murtagh F. 2002 天文学中的反卷积：综述。太平洋天文学会刊**114**, 1051. 44. Tobar F, Araya-Hernández L, Huijse P, Djuri´c PM. 2021 傅里叶对的贝叶斯重建。*IEEE* 信号处理汇刊**69**, 73–87. 45. Cazelles E, Robert A, Tobar F. 2021 平稳时间序列的 Wasserstein-傅里叶距离。*IEEE* 信号处理汇刊**69**, 709–721. 46. Krizhevsky A. 2009 从微小图像中学习多层特征。多伦多大学技术报告。47. Tobar F. 2018 贝叶斯非参数谱估计。载于 Bengio S, Wallach H, Larochelle H, Grauman K, Cesa-Bianchi N, Garnett R 编，《style id='32'>神经信息处理系统进展》第 31 卷。Curran Associates, Inc.

附录

本节包含定理 1 所需的详细计算、音频反卷积实验的详细图表，以及图像反卷积实验的更多示例。

（a）引理 1 的扩展证明

引理 2. 如果卷积滤波器 h 和协方差 K_x 均可积，那么 $K_f(t)$ 可积。

证明。 这与两个函数卷积可积性的标准证明思路相同，只是稍作修改，因为 $K_f(t)$ 的定义——文章中的式（5）——包含两个卷积的复合，而不只是一个。因此，利用 Fubini 定理和三角不等式（两次），可得

$$
\begin{aligned}
\int_{\mathbb{R}} |K_f(t)| \mathrm{d}t &= \int_{\mathbb{R}} \left| \int_{\mathbb{R}} h(\tau') \int_{\mathbb{R}} h(\tau) K_x(\tau - (\tau' - (t)) \mathrm{d}\tau \mathrm{d}\tau' \right| \mathrm{d}t && \text{[Fubini on eq. (5))]} \\
&\leq \int_{\mathbb{R}} \int_{\mathbb{R}} |h(\tau')| \int_{\mathbb{R}} |h(\tau)| \left| K_x(\tau - (\tau' - (t)) \right| \mathrm{d}\tau \mathrm{d}\tau' \mathrm{d}t && \text{[triangle ineq. twice]} \\
&= \int_{\mathbb{R}} |h(\tau')| \int_{\mathbb{R}} |h(\tau)| \int_{\mathbb{R}} \left| K_x(\tau - (\tau' - (t)) \right| \mathrm{d}t \mathrm{d}\tau \mathrm{d}\tau' && \text{[Fubini]} \\
&= ||h||_1 ||h||_1 ||K_x||_1 \\
&< \infty. && [h, K_x \in L_1]
\end{aligned}
$$

□

（b）定理 1 的计算

$$
\begin{aligned}
\mathbb{E}\left[\hat{x}_w(\xi)|\mathbf{y}\right] &= \mathbb{E}\left[\mathcal{F}\left\{x(t)w(t)\right\}|\mathbf{y}\right] \\
&= \mathcal{F}\left\{\mathbb{E}\left[x(t)|\mathbf{y}\right]w(t)\right\} \\
&= \mathcal{F}\left\{\left(m_x(t) + K_{xy}(t,\mathbf{t})K_y^{-1}(\mathbf{t})\mathbf{y}\right)w(t)\right\} \\
&= \left(m_x(t) + \mathcal{F}\left\{K_{xy}(t,\mathbf{t})\right\}K_y^{-1}(\mathbf{t})\mathbf{y}\right) \star \hat{w}(\xi) \\
&= \left(\hat{m}_x(\xi) + \hat{K}_{xy}(\xi)e^{-j2\pi\mathbf{t}\xi}K_y^{-1}(\mathbf{t})\mathbf{y}\right) \star \hat{w}(\xi) \\
&= \left(\hat{m}_x(\xi) + \hat{K}_x(\xi)\hat{h}(\xi)e^{-j2\pi\mathbf{t}\xi}K_y^{-1}(\mathbf{t})\mathbf{y}\right) \star \hat{w}(\xi)
\end{aligned} \tag{7.1}
$$

$$
\begin{aligned}
\mathbb{V}\left[\hat{x}_w(\xi)|\mathbf{y}\right] &= \mathbb{V}\left[\mathcal{F}\left\{x(t)w(t)\right\}|\mathbf{y}\right] \\
&= \mathbb{E}\left[\overline{\mathcal{F}\left\{x(t)w(t) - \mathbb{E}\left[x(t)w(t)|\mathbf{y}\right]\right\}}\mathcal{F}\left\{x(t)w(t) - \mathbb{E}\left[x(t)w(t)|\mathbf{y}\right]\right\}|\mathbf{y}\right] \\
\text{[lin. exp.]} &= \mathbb{E}\left[\overline{\mathcal{F}\left\{(x(t) - \mathbb{E}\left[x(t)|\mathbf{y}\right])w(t)\right\}}\mathcal{F}\left\{(x(t) - \mathbb{E}\left[x(t)|\mathbf{y}\right])w(t)\right\}|\mathbf{y}\right] \\
\text{[conv. thm]} &= \mathbb{E}\left[\mathcal{F}\left\{(x(-t) - \mathbb{E}\left[x(-t)|\mathbf{y}\right])w(-t) \star (x(t) - \mathbb{E}\left[x(t)|\mathbf{y}\right])w(t)\right\}|\mathbf{y}\right] \\
\text{[def. conv.]} &= \mathbb{E}\left[\mathcal{F}\left\{\int (x(-\tau) - \mathbb{E}\left[x(-\tau)|\mathbf{y}\right])w(-\tau)(x(t-\tau) - \mathbb{E}\left[x(t-\tau)|\mathbf{y}\right])w(t-\tau)\right\}|\mathbf{y}\right] \\
\text{[lin. conv.]} &= \mathcal{F}\left\{\int w(-\tau)\mathbb{E}\left[(x(-\tau) - \mathbb{E}\left[x(-\tau)|\mathbf{y}\right])(x(t-\tau) - \mathbb{E}\left[x(t-\tau)|\mathbf{y}\right])|\mathbf{y}\right]w(t-\tau)\right\} \\
\text{[def. cov.]} &= \mathcal{F}\left\{\int w(-\tau)\mathbb{V}\left[x(-\tau), x(t-\tau)|\mathbf{y}\right]w(t-\tau)\right\} \\
\text{[def. cov]} &= \mathcal{F}\left\{\int \left(K_x(t) - K_{xy}(-\tau,\mathbf{t})K_y^{-1}K_{yx}(\mathbf{t},t-\tau)\right)w(-\tau)w(t-\tau)\mathrm{d}\tau\right\} \\
&= \hat{K}_x(\xi) \star |\hat{w}(\xi)|^2 - \mathcal{F}\left\{\int w(-\tau)K_{xy}(-\tau,\mathbf{t})K_y^{-1}K_{yx}(\mathbf{t},t-\tau)w(t-\tau)\mathrm{d}\tau\right\} \\
&= \hat{K}_x(\xi) \star |\hat{w}(\xi)|^2 - \overline{\mathcal{F}\left\{w(t)K_{xy}(t,\mathbf{t})\right\}}\,K_y^{-1}\mathcal{F}\left\{K_{yx}(\mathbf{t},t)w(t)\right\} \\
&= \hat{K}_x(\xi) \star |\hat{w}(\xi)|^2 - \overline{\hat{w}(\xi) \star \hat{K}_x(\xi)\hat{h}(\xi)e^{-j2\pi\xi\mathbf{t}}}\,K_y^{-1}\hat{w}(\xi) \star \hat{K}_x(\xi)\hat{h}(\xi)e^{-j2\pi\xi\mathbf{t}} \\
&= \hat{K}_x(\xi) \star |\hat{w}(\xi)|^2 - \|\hat{w}(\xi) \star \hat{K}_x(\xi)\hat{h}(\xi)e^{-j2\pi\xi\mathbf{t}}\|_{K_y^{-1}}
\end{aligned} \tag{7.2}
$$

(c) 扩展音频实验

对于音频去混响实验（参见文章正文中的图 6 ），我们还展示了所有详细考察方法的估计结果。首先，图 8 展示了真实信号（及滤波器），而图 9 展示了所提出的 GPDC 的估计结果（均值和 95% 误差棒）。图 10 和11 分别展示了维纳方法和逆傅里叶变换方法的估计结果；在本实验中，我们将这两种方法作为 GPDC 的基准方法。还需说明的是，由于这两种方法在频域中直接进行反卷积，因此不会保留信号的相位，故需要将其与真实信号对齐——为便于可视化，我们已对它们进行了对齐。

基于信号源的估计结果 x（由逆傅里叶变换、维纳滤波器和所提出的 GPDC 提供，分别见图 11、10 和 9 ），我们可以看出，逆傅里叶变换所提供的重建效果最差（因为它没有考虑信号源具有随机性这一事实）。随后，维纳滤波器所提供的重建效果优于逆傅里叶变换，因为维纳滤波器考虑了滤波器和信号的动态特性，但它缺少一个似然模型，这意味着它无法区分（含噪）观测值 $\mathbf{y}$ 与卷积 f。最后，请注意，

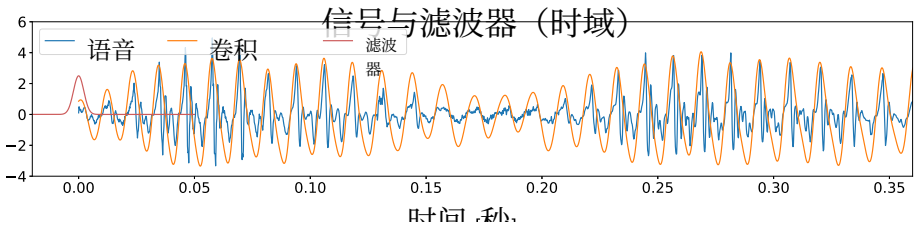


图 8：真实音频信号（信号源和卷积结果）以及卷积滤波器。

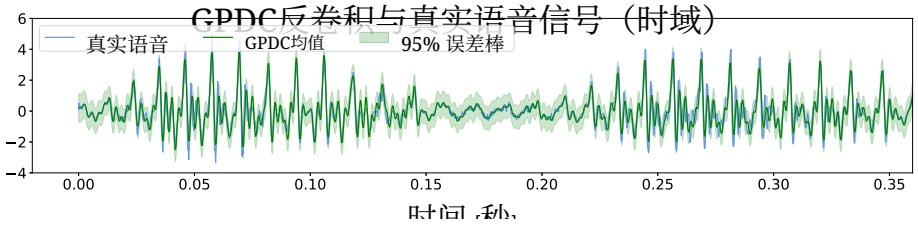


图 9：语音信号的反卷积：GPDC估计、真实信号源及 95% 误差棒。

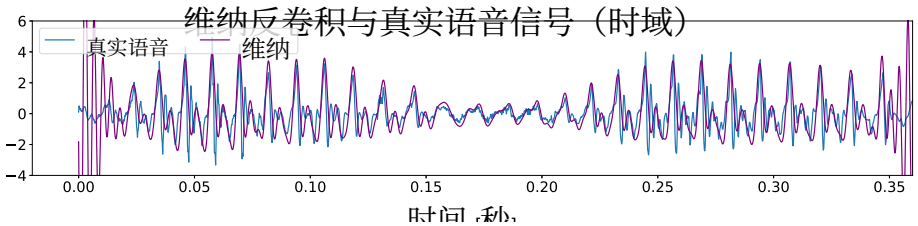


图 10：语音信号的反卷积：维纳估计及真实信号源。

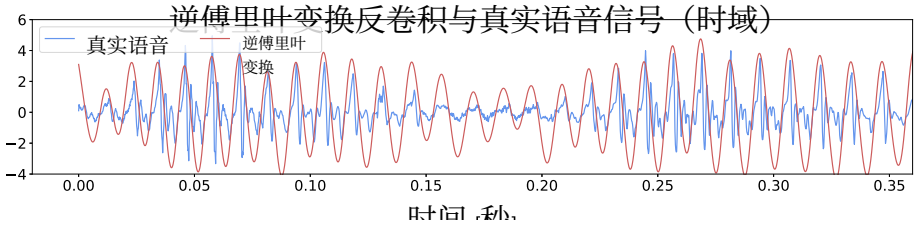


图 11：语音信号的反卷积：逆傅里叶变换估计及真实信号源。

所提出的 GPDC 优于基准方法，因为它同时考虑了信号源的动态行为（通过 K_x）以及观测值$\mathbf{y}$中的噪声。

（d）扩展图像实验

本节针对图像反卷积或 超分辨率实验，使用其他测试图像和滤波器进行了进一步的模拟。该实验旨在将所提出的GPDC和盲GPDC（其中滤波器 h 未知，因此从图像中学习）与维纳反卷积方法进行比较验证。我们考虑了CIFAR-10数据集中的以下 (32×32) 图像：一架飞机（图 12）、一只鹌鹑（图 13）、一只鸟（图14）、一匹马（图15）

和第二架飞机（图16）。在所有实验中，我们考虑了 5 个不同的滤波器 h ，尺寸为(5×5) 像素。这些滤波器在每幅图中均按从上到下的顺序显示，并以 0 到 1 的灰度进行颜色编码，具体如下：

(i) h_0：恒定滤波器；(ii) h_1：在原点处具有较大值（即 1）、而在原点附近接近于零的滤波器；(iii) h_2：类似高斯 RBF 衰减的径向滤波器；(iv) h_3：由在 0 和 1 之间均匀分布的随机值组成的滤波器；(v) h_4：对角滤波器。

最后，对于所有图像和滤波器，在观测值数量从 10% 到 90% 的情况下，均实现了 GPDC 和盲 GPDC（用于性能曲线）；图中展示的是缺失数据比例为 30% 的情况。

为了了解所提出的 GPDC 及其盲变体相对于维纳方法的性能，请注意，**维纳基准是在完整且无噪声的图像上实现的**，而所提出的模型只能使用存在缺失数据和噪声的图像版本。我们注意到，对于滤波器 h_1 和 h_2，尽管维纳方法优于 GPDC 和盲GPDC，但所提出的模型会随着观测值数量的增加而持续改进，甚至盲 GPDC 在视觉上也能够恢复真实图像。对于更复杂的滤波器（h_0、h_2、h_4），我们可以看到，只要图像中有超过 25% 的部分可用，两种 GPDC 方法的性能就会大幅优于维纳基准。对于所有图像都能一致地观察到这一行为，从而验证了所提出的 GPDC 方法在处理不同图像时的稳健性。

在以下所有图中，结果从左到右依次为：滤波器（取值介于 0 和 1 之间，以灰度编码）、真实图像 x、卷积 f、观测$\mathbf{y}$ （红色表示 30% 的缺失像素）、GPDC估计、盲GPDC估计、维纳估计，以及 GPDC 随观测像素数量变化的性能（20 次试验的误差棒）。每幅图中的每一行都是同一实验的一次实现，但使用了不同的滤波器。

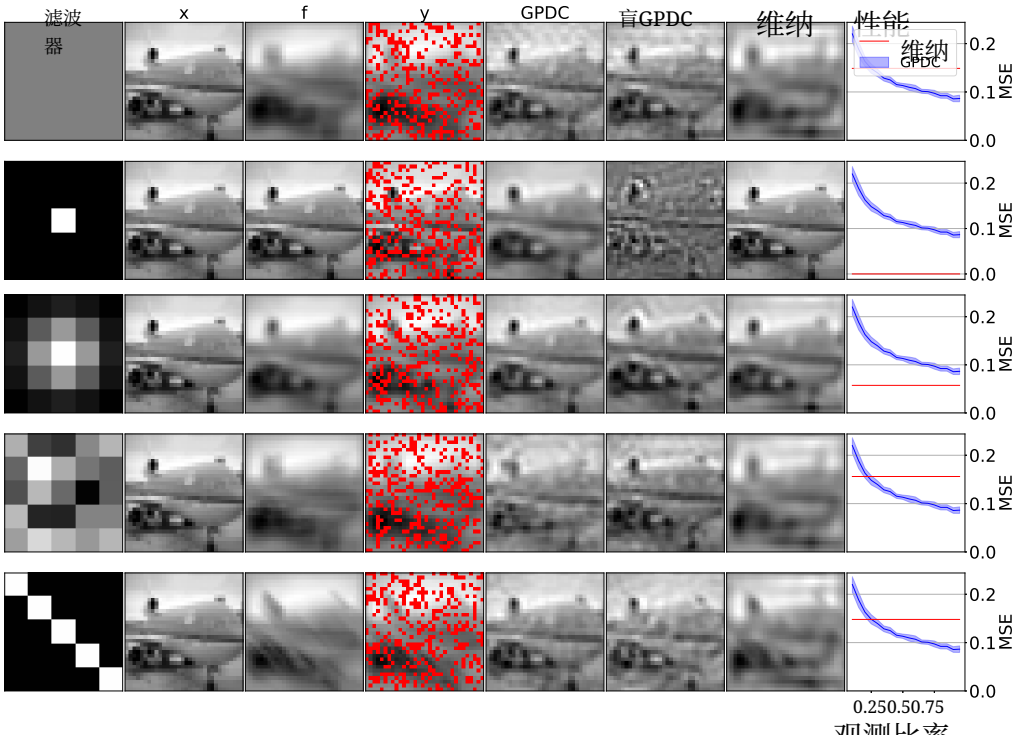


图12：平面的图片。

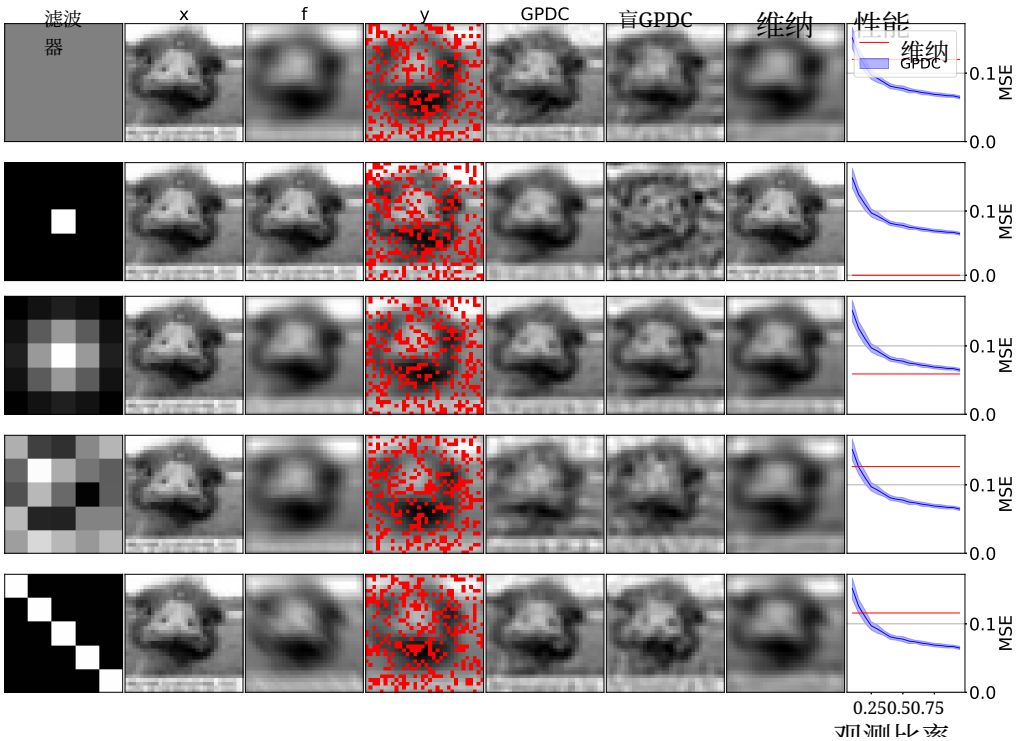


图13：一只鹌鹑的图片。

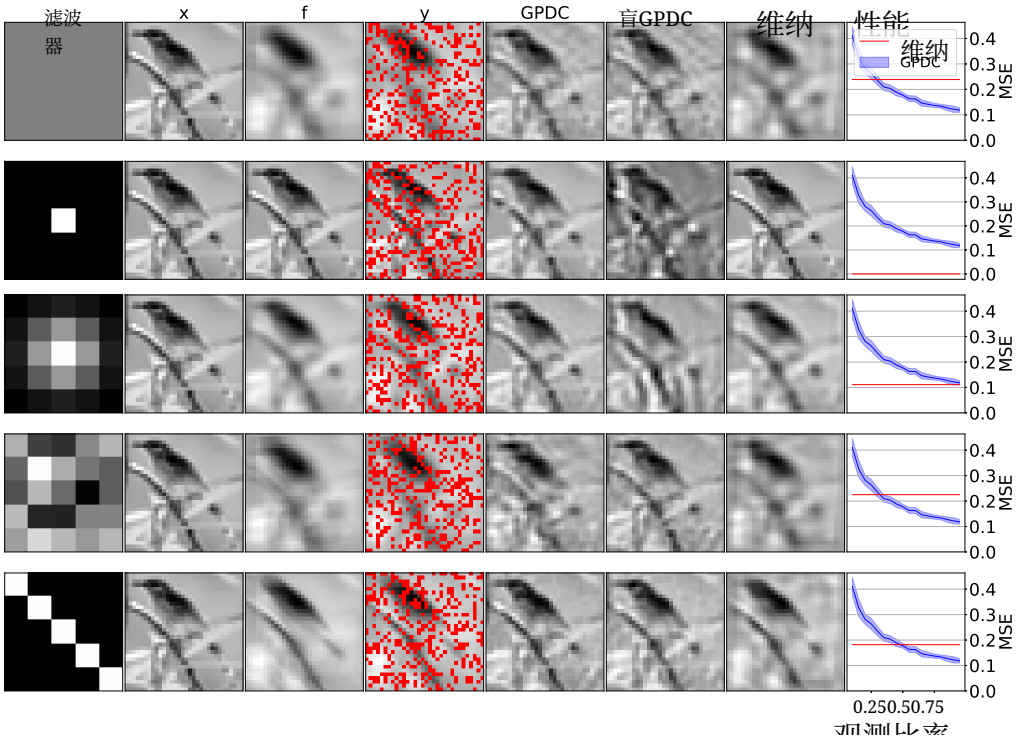


图14：鸟的图片。

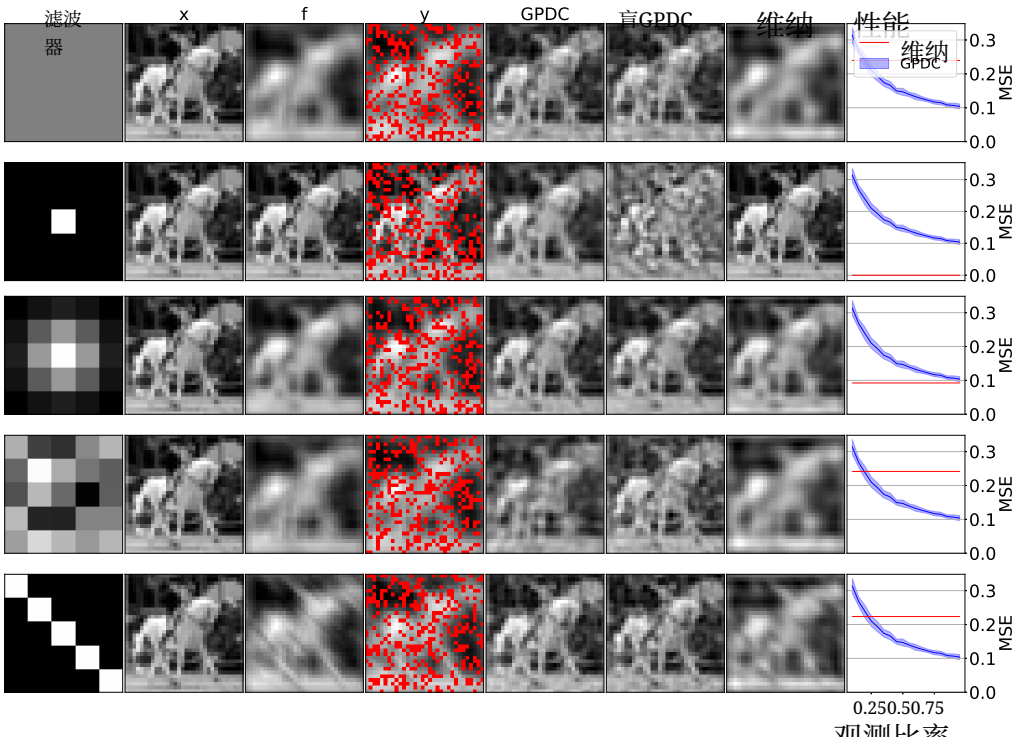


图15：马的图片。

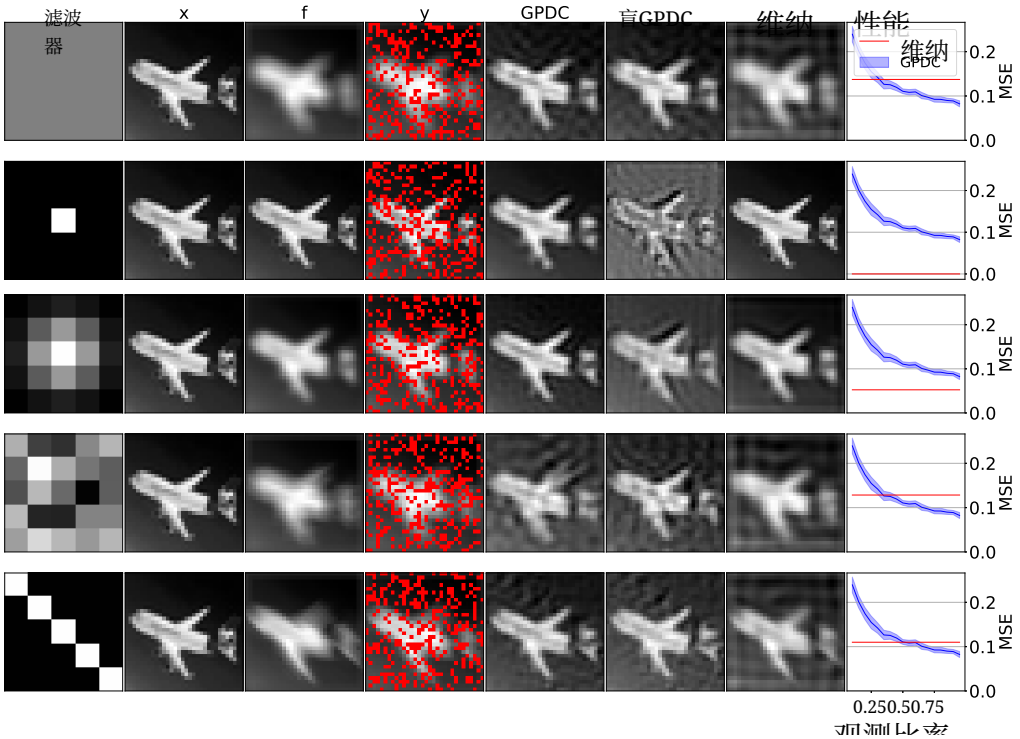


图16：平面的第二张图片。