tomoCAM：通过 GPU加速和非均匀快速傅里叶变换实现快速基于模型的迭代重建

Dinesh Kumar, [a, c]* Dilworth Y. Parkinson[b, c] and Jeffrey J. Donatelli[a, c][a]数学系，劳伦斯伯克利国家实验室，伯克利，加利福尼亚州，[b]先进光源，劳伦斯伯克利国家实验室，伯克利，加利福尼亚州，以及[c]能源研究应用高级数学中心，劳伦斯伯克利国家实验室，伯克利，加利福尼亚州。电子邮箱：*dkumar@lbl.gov*

arXiv:2304.12934v1 [physics.med-ph] 18 Apr 2023

摘要

基于 X 射线的计算机断层扫描（CT）是一种成熟的技术，用于根据物体的二维投影确定其三维结构。在过去几十年中，同步辐射源上的断层成像仪器在光源亮度和探测器技术方面取得了重大进展。这些进展带来了新的观测和发现，并提升了诸如更快的帧率、更大的视场、更高的分辨率和更高的维度等能力。这些进展使材料科学界能够将断层测量的范围扩展到越来越多的原位 和操作中 测量。在这些新型实验中，样品可能快速演化、具有复杂几何形状，并受到视场限制，从而限制了可采集的投影数量。在此类情况下，用于重建图像的标准滤波反投影（FBP）通常会产生质量较差的重建图像。迭代重建算法，例如基于模型的迭代重建（MBIR），已证明能够在这些限制条件下成功生成高质量的重建图像，但通常需要拥有数百个计算节点的高性能计算资源，才能在合理的时间内解决该问题。

本文介绍了**tomoCAM**，这是一种新的基于 GPU 加速的基于模型的迭代重建实现。该实现利用非均匀快速傅里叶变换（NUFFT）高效计算 Radon 算子和反投影算子，并通过异步内存传输最大限度地提高数据传输至 GPU 内存的吞吐量。与传统的 MBIR 代码相比，生成的代码速度显著更快，并以可负担的计算时间和资源实现了 MBIR 所带来的重建质量提升。**tomoCAM**具有 Python 前端，可从基于 Jupyter 的框架访问，从而能够直接集成到同步辐射设施的现有工作流程中。关键词— X 射线断层成像、微型计算机断层扫描、同步辐射、断层重建、GPU

1. 引言

利用同步辐射技术进行微米和纳米断层成像，对于揭示现代材料的内部结构至关重要，尤其是在动态环境中。其应用领域广泛，包括研究陶瓷基复合材料的裂纹和劣化；这类新型轻质材料用于在高温高压下运行的喷气发动机（Forna-Kreutzer 等，2021）；研究石油、盐水和二氧化碳在岩石中的流动（Walsh 等，2014）；以及分析电池中的枝晶形成，这会导致容量降低并最终失效（Dienemann等，2023）。如今，许多同步辐射微型计算机断层扫描设施配备了相机，可每秒采集数千帧、每帧包含数百万像素的图像（MacDowell 等，2012；Nikitin 等，2022；Ge 等，2018；Thuering等，2011；Mokso等，2017）。这些仪器技术的进步推动用户不断拓展同步辐射光束线可成像对象的边界。越来越多的研究人员正在开展原位 （Larson & Zok, 2018；French等，2022）和操作中 （Kulkarni 等，2020；Dienemann 等，2023）测量。对于微型计算机断层扫描重建，通常首先尝试的技术是滤波反投影 （FBP）方法（Herman，2009），该方法可在 *tomopy* （G¨ursoy 等，2014）和*tomocupy* （Nikitin，2023）中使用。然而，在许多动态实验中，观测样品变化迅速，通常无法采集足够的投影以满足角向香农采样条件（Crowther等，1970）并克服噪声。在这种情况下，FBP 并不适用。采用该方法得到的重建图像往往噪声过高，并出现条纹伪影，使得后续分析难以开展，甚至无法进行。

作为替代方案，同步迭代重建技术（SIRT）（Tarantola & Valette，1982）和基于模型的迭代重建（MBIR）（Venkatakrishnan 等，2013年；Mohan等，2014年）等迭代方法旨在缓解这些不足。它们将重建表述为一个优化问题。通过迭代过程获得解，该过程旨在最小化测量数据与样品数字表示的前向模型（拉东变换）之间的不匹配。如多项研究所示（Trampert & Leveque，1990年；Zhang等，2014年；Venkatakrishnan 等，2013年；Mohan 等，2014年），这种迭代方法能够将先验知识（例如全变分约束）纳入优化过程。然而，目前基于 CPU 的 MBIR 实现通常需要大型计算集群，才能实现与数据采集时间相当的周转时间。这不仅增加了实验到分析的闭环所需的额外时间，也给材料科学家带来了额外负担，因为他们必须掌握使用计算集群的新技能。本文介绍了**tomoCAM**，这是一种基于非均匀快速傅里叶变换方法的 GPU 加速 MBIR 实现。借助现代 GPU 设备提供的计算能力以及相对低廉的计算机内存成本，如今可以在单台机器上于合理时间内完成这些重建图像。

为设计我们的 GPU 加速算法及其实现，我们基于非均匀快速傅里叶变换（NUFFT）构建了 Radon 和反投影算子（Greengard & Lee，2004年；Fessler & Sutton，2003年），在保持高精度的同时显著降低了计算成本。我们还利用了 CUDA 软件开发套件（NVIDIA

等，2020年）中原生提供的高度优化 cuFFT 库。我们遵循（Venkatakrishnan 等，2013年；Mohan 等，2014年）提出的数学框架，加入了全变分约束，有助于在保留清晰边缘的同时减少噪声。我们实现的一个重要特性是能够灵活地引入不同的约束。约束的选择不受算法设计的限制。

我们通过对已知模体和 Tomobank 公开提供的实验数据开展一系列数值实验，测试了我们的计算框架（Carlo 等，2018 年）。我们将重建图像与通过滤波反投影（G¨ursoy等，2014 年）和 **SVMBIR** （ Team，2020 年）获得的结果进行了比较，后者是一种公开的基于 CPU 的软件包。

在数值实验中，我们发现：

- 与 FBP 相比，**tomoCAM** 生成的重建质量更高、噪声更少，并且所需的投影数量更少，
- 此外，我们观察到，在单台机器上，**tomoCAM** 的运行速度约为**SVMBIR**的 15 倍。

.

最后，我们提供了一个 Python 前端，将 **tomoCAM** 的功能开放给广泛使用的 NumPy 软件包（Harris 等，2020 年），并使用 pybind11（Jakob 等，2017 年）实现。这使得将 tomoCAM 集成到现有工作流程中变得十分容易。代码可从 https://github.com/lbl-camera/tomocam 免费获取。

2. 拉东变换与非均匀快速傅里叶变换

本节简要概述与断层成像相关的基本概念，包括拉东变换（及其伴随算子）以及它与傅里叶变换的联系。为进行断层测量，通过旋转相机或所研究的样品，在不同角度采集一系列图像，这些图像称为投影。

2.1. 拉东变换

拉东变换是任何断层重建的基础。它通过线积分（2）将函数 $f(\mathbf{x}\ z)$，$\mathbf{x}\ \in\mathbb{R}^2$，$z \in \mathbb{R}$ 变换为 $R\,f(t,\ \hat{}n,\ z)$ $t \in \mathbb{R}$，$\hat{}n \in \mathbb{S}^1$，见图 1。给定一组有向直线$\ell_{t,\ \hat{n}}$，其定义为

$$\ell_{t,\hat{n}} = \{\mathbf{x} : \langle \mathbf{x}, \hat{n} \rangle = t\} = \{t\hat{n} + s\hat{n}_\perp : s \in \mathbb{R}\} \tag{1}$$

其中，$\hat{}n_\perp$ 是 X 射线束的方向，$\hat{}n$ 是在与 ℓ 相同平面内且垂直于射线束的方向，t 是从原点到 ℓ 的距离。函数 f 的拉东变换 Rf 定义为:

$$Rf(t, \hat{n}, z) = \int_{\ell_{t,\hat{n}}} f(\mathbf{x}, z) = \int_{-\infty}^{\infty} f(t\hat{n} + s\hat{n}_\perp, z)\, ds \tag{2}$$

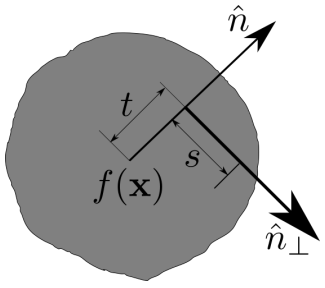

图 1. f 的拉东变换是其沿着每条垂直于 $\hat{}n$ 的直线的线积分

断层测量可以准确地建模为由 f 表示的样品密度的 Radon变换。通过对该方程 (2) 求逆，可以从数据中重建 f，这对断层重建至关重要。根据中心切片定理，f 沿方向 $\hat{}n$ 的 Radon变换的傅里叶变换等价于 f 沿 $\hat{}n$ 的傅里叶变换，即:

$$\begin{aligned}\mathbb{F}_1[R\,f](k,\hat{n},z) &:= \int_{-\infty}^{\infty} e^{-2\pi ikt} R\,f(t,\hat{n},z)dt \\ &= [\mathbb{F}_2\,f](k\hat{n},z),\end{aligned} \tag{3}$$

其中，z 是沿旋转轴的维度，$\mathbb{F}_d$ 表示 d 维傅里叶变换。Radon变换及其伴随算子是应用于三维数据、逐切片处理的二维算子。为简化起见，在后续记号中我们将省略对 z 的依赖。假设 f 和 $\mathbb{F}_1 f$ 在处处均可积，则 Radon变换 (2) 的逆变换为：

$$f(\mathbf{x}) = \int_0^{\pi}\int_{-\infty}^{\infty} e^{2\pi ik\langle \mathbf{x},\hat{n}(\theta)\rangle}\int_{-\infty}^{\infty} e^{-2\pi ikt} y(t,\hat{n}(\theta))dt\,|k|\,dk\,d\theta, \tag{4}$$

其中，我们将 θ 依赖关系表示为 $\hat{n}(\theta) = (\cos\theta,\ \sin\theta)$。

精确计算公式 (4) 的计算成本非常高。在实践中，公式 (4) 通常通过非均匀快速傅里叶变换（G¨ursoy 等，2014年）进行高效近似，或者直接在实空间中使用 Shepp-Logan（Shepp&Logan，1974年）、Ram-Lak（G.N.& A.V.，1971年）和 Butterworth（Butterworth 等，1930年）等各种滤波器进行估计。这些滤波器根据傅里叶采样密度进行近似和加权，因此得名滤波反投影。此外，这些滤波器还经过专门设计，用于抑制较高的傅里叶频率。这是重建层析数据时最常用的方法，部分原因在于其能够以极快的速度完成反演。然而，当视野受到部分遮挡或样品处于演化状态时，可能无法采集足够的投影来充分采样傅里叶空间。在这种情况下，滤波反投影会导致较差的图像质量。

2.2. 非均匀快速傅里叶变换（*NUFFT*）

在离散均匀网格上，拉东变换的逆变换需要使用一维快速傅里叶变换（FFT）沿径向线计算傅里叶系数，然后将非均匀极坐标网格 $\{\mathbf{k}_j = k_j(\cos\theta_j, \sin\theta_j)\}$ 上的二维傅里叶逆变换映射到笛卡尔网格 $\{(x_n, y_n)\}$，这可以表示为求和

$$f_n = \sum_{j=1}^{M} c_j e^{2\pi ik_j(x_n\cos\theta_j + y_n\sin\theta_j)}, \quad n \in [1,N], \tag{5}$$

其中，c_j 是 $\mathbf{k}_j$处的傅里叶系数，N 是表示均匀笛卡尔网格上样品密度的离散点数量，M 是表示投影数据的极坐标网格点数量。然而，直接计算公式（5）的计算成本很高，其复杂度为 $\mathcal{O}(MN)$。同步辐射光源获取的数据通常可达到 $M = \mathcal{O}(10^{10})$ 像素，而最终重建图像的尺寸 N 与最终重建图像的数量级相近。

非均匀快速傅里叶变换（NUFFT）为计算方程（5）提供了一种精确且高效的方法。该方法首先使用沿径向线进行的一系列一维快速傅里叶变换，在极坐标网格上计算傅里叶系数。然后，将计算得到的系数与紧支撑扩散核 φ进行卷积，并在均匀网格上计算该卷积。随后，对均匀网格上的卷积值执行傅里叶逆变换，再除以该核的傅里叶变换 $\hat{\varphi}$，即

$$c_j = \sum_t \rho(t,\theta_j)e^{-2\pi itk_j} \tag{6}$$

$$F_r = \sum_{\|\mathbf{k}_r^c - \mathbf{k}_j\| < W} c_j\varphi(\mathbf{k}_r^c - \mathbf{k}_j), \tag{7}$$

$$f_n \approx \hat{\varphi}^{-1}{\mathbb{F}_2}^{-1}(F) \tag{8}$$

其中 $\{\mathbf{k}_r^c\}$ 为笛卡尔网格，$\mathbb{F}_2$ 为二维傅里叶变换，公式（6）和（8）通过快速傅里叶变换计算，W 为公式（7）中卷积的扩散宽度。对于适当选择的核函数，NUFFT 的误差为 ϵ，前提是将 W 选择为跨越约 $w = \log_{10}(1/\epsilon)$ 个网格

点（每个维度）。公式（6）至（8）的计算复杂度为 $\mathcal{O}(MlogM_t + w^2N + NlogN)$，其中 M_t 为极坐标网格径向方向上的点数。由于 $w^2 \ll M$，与直接计算公式（5）相比，这将带来巨大的加速。拉东变换同样可以通过按相反顺序执行上述步骤来计算。在本研究中，我们使用 cuFINUFFT（Shih等，2021）库来计算非均匀快速傅里叶变换。关于这一主题的详细讨论，参见（Dutt 和 Rokhlin，1993；Fessler 和 Sutton，2003；Greengard 和 Lee，2004；Barnett等，2019；Barnett，2021）。

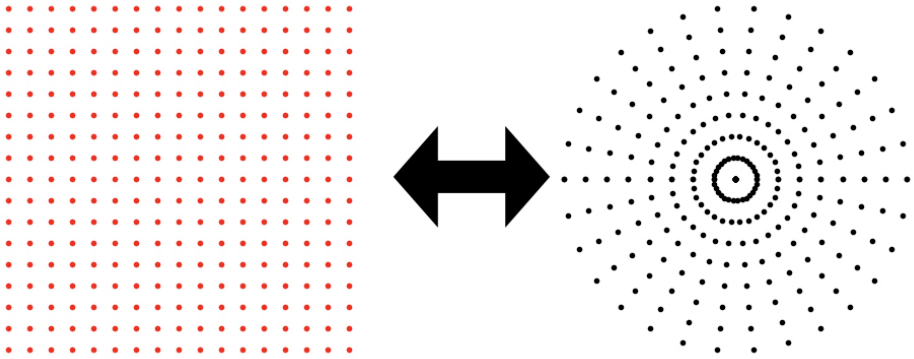

图2。NUFFT 用于将均匀网格上的强度转换为极坐标网格上的傅里叶变换，反之亦然。

3. 基于模型的迭代重建

FBP 方法的另一种方法是采用迭代方法，例如 MBIR。尽管这些方法的处理时间更长，但与 FBP 方法相比，它们能够生成质量更高的重建图像。当可用的投影数量较少时，这一点尤为明显。这是因为迭代方法能够将先验信息作为优化过程中的约束纳入其中。关于更详细的讨论，请参阅 (ASTM-E1441-19, 2019)。迭代方法通过最小化其拉东变换与投影数据 b 之间的差异来寻求一个解f，即，

$$f = \arg\min_f \underbrace{\|Rf - b\|^2}_{\mathcal{G}} + \underbrace{g(f)}_{\mathcal{H}}. \tag{9}$$

这里我们将目标函数设为一个最小二乘问题。目标是迭代搜索 f，使其最小化 Rf 与 b 之间差异的 ℓ^2-范数，同时通过 g 对违反约束的情况进行惩罚。现在，我们对方程 (9) 关于 f 求导，并令结果等于 0。第一项的梯度为

$$\nabla\mathcal{G} = R^*\left(Rf - b\right), \tag{10}$$

其中， R^* 是公式（2）的伴随算子，

$$R^*\rho(\mathbf{x}) = \int_0^{\pi}\int_{-\infty}^{\infty} e^{2\pi ik\langle \mathbf{x},\hat{n}(\theta)\rangle}\int_{-\infty}^{\infty} e^{-2\pi itk}\rho(t,\hat{n}(\theta))dt\ dk\,d\theta, \tag{11}$$

即公式（4）去掉缩放因子 $|k|$ 后的形式。利用 NUFFT，可以高效地计算算子 R 和 R^* 使用非均匀快速傅里叶变换。

在此展示的结果中，我们在方程 (9) 中将 $\mathcal{H}$ 设为全变差惩罚项。我们遵循 (Venkatakrishnan 等，2013年；Mohan等，2014年) 中提出的数学方法，将 $\mathcal{H}$ 建模为 q-广义高斯随机场（qGGRMF），

$$g_m = \sum_n w_{mn} h_{mn}, \quad \forall n \in \{n \mid \|m-n\|_\infty \leq 1\} \tag{12}$$

$$h_{mn} = \frac{\left(\frac{|f_m - f_n|}{\sigma}\right)^2}{c + \left(\frac{|f_m - f_n|}{\sigma}\right)^{2-p}} \tag{13}$$

其中， h 定义在 m 的一跳邻域上， m 和 n 是三维均匀网格上的整数坐标。权重 w_{mn} 是将单位分割的高斯权重，并且与 m 和 n 之间的距离成反比。超参数 c、 p 和 σ 用于控制惩罚项的强度。项 $\mathcal{H}$ 是一个代数表达式，可以轻松求导。

在本研究中，我们采用了 (Giselsson & Boyd, 2014) 中详细介绍的带重启单调加速梯度方法，但也可以使用其他优化器。

4. 实现

在实现软件解决方案时，性能是一个关键因素。本节将讨论一些对**tomoCAM**性能有重大影响的重要实现细节。这些因素包括内存管理，以及通过高效的 GPU 缓存来隐藏 PCIe 延迟。为了兼顾高性能和易用性，我们综合使用了 C++、CUDA 和 Python。**tomoCAM** 的数据结构使用 C++实现，而大多数数学函数则使用 CUDA 编写。为了高效处理大型数据集，我们采用了两级分区方案，以便将数据无缝流入和流出 GPU 内存。针对典型同步辐射微米或纳米 CT 正弦图中数量庞大的像素（可能超过 $\mathcal{O}(10^{10})$），我们对**tomoCAM**的内存使用进行了仔细优化。例如，为了最大限度地减少内存占用，我们将包含经常访问数据的大型数组（如最新的解、投影数据和梯度）作为引用而非副本传递，这是 C++中的默认行为。我们实施了多种策略来最大限度地减少内存占用，包括：

- 主机内存中从不以复数形式存储数据。这也有助于减少与 GPU 内存之间传输的数据量。
- 分区不会复制数据，而是包含指向父数组中内存位置的指针。
- 计算全变分约束时，会就地更新梯度。
- 投影数据会重新排列为正弦图形式，以便快速连续读取。

4.1. GPU 优化

GPU 在执行复杂计算方面效率极高，但 PCIe 总线上的延迟仍然是 GPU 加速软件实现的一大瓶颈。为了尽量缩短运行时间并最大限度地提高从 CPU 到 GPU 内存的数据吞吐量，我们采用了多种技术，包括异步传输、OpenMP 线程和两级数据分区方案。分区沿旋转轴进行，首先根据可用 GPU 设备的数量将数据划分为相应数量的分区。随后，每个分区会进一步划分为更小的数据块，其最佳大小取决于 GPU 设备的可用内存。子分区会以流式方式传输到 GPU 内存中；为尽量减少内存占用，不会为数据创建深拷贝。图 3 概述了这一流程。通过采用这些技术，我们能够显著降低 PCIe 瓶颈的影响，并提升 GPU 加速软件实现的性能。tomoCAM 的其他优化和功能包括：

- 由于旋转轴可能未与图像中心对齐，我们利用傅里叶移位性质，高效地将旋转轴移至图像中心。
- 我们使用 OpenMP 线程在所有可用的 GPU 上并行启动一级分区，并将数据流式传输至 GPU 内存。
- 为提高缓存效率，我们利用 GPU 的共享内存存储会被多次访问的数据，例如在计算全变分 约束时。

- 通过 pybind11 项目（Jakob 等，2017年）提供了 Python 前端和 NumPy 兼容性。

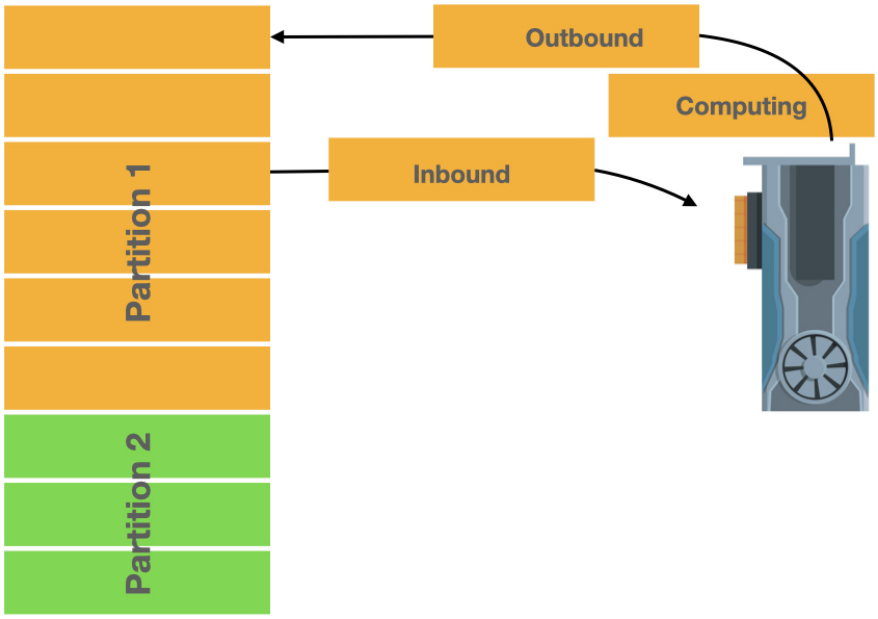

图 3. 大型数组采用两级分区方案沿旋转轴进行分区。

tomoCAM 已发布 可通过 https://github.com/lbl-camera/tomoc 作为开源项目公开获取 am.

5. 数值实验

我们使用公开可用的仿体和实测数据集对**tomoCAM** 进行了测试。这里，我们展示了使用**tomoCAM**、**SVMBIR** （Team，2020年）以及 Tomopy 软件包中提供的 gridrec 进行滤波反投影所得到的重建结果比较（G¨ursoy 等，2014年）。在绘图之前，每个重建结果和线剖面（B）均使用 s 进行缩放，并使用 Δ 进行平移，其中 $s,\ \Delta = \arg\min_{s,\ \Delta} \|A - sB + \Delta\|$ 以真实值（A）为基准。在实验数据的情况下，我们使用**tomoCAM**得到的结果对**SVMBIR** 和 gridrec 的重建结果进行重新缩放。**SVMBIR** 中使用的全变分约束与**tomoCAM**中使用的约束略有不同，详见（Team，2020年）中的理论部分。**SVMBIR** 使用 10 个最近邻，而**tomoCAM** 使用其中的 26 个来计算（12）。我们认为，可以对**tomoCAM**和 **SVMBIR** 的参数进行微调，以产生等效结果。与**SVMBIR** 的主要比较旨在展示性能提升，而不是比较两种不同的约束或图像质量。所有测试均在单台机器上完成，

- $2\times$ Intel(R) Xeon(R) CPU E5-2620 v4 @ 2.10GHz
- $4\times$ Tesla P100 GPU
- 128GB 随机存取存储器

在第一个实验中，我们比较了使用三种代码对泡沫仿体进行的重建。我们使用 **foam ct phantom** 软件包（Pelt 等，2022）生成了一个泡沫仿体及其大小为（$128\times16\times2048$）的投影数据。图 4(a) 展示了该模体的一个完整切片，并将其与使用各代码（**SVMBIR**、**tomoCAM** 和 gridrec）获得的重建结果（4(b-d)）进行了比较。随后，在 4(e-h) 中展示了每幅图像的 放大 区域。然后，在 4(i) 中比较了各个放大区域的线剖面。结果显然表明，**tomoCAM** 和 **SVMBIR** 均能有效抑制噪声。**tomoCAM** 的一项主要优势是，它能够在快一个数量级的时间内获得相当的结果。

接下来，我们评估通过 Tomobank（Carlo 等，2018）获取的两个不同同步辐射光源实验数据集的重建图像。对于每个数据集，可用的投影数量明显低于通常预期的数量，这遵循了一般经验法则，即投影数量应与相机传感器中的像素列数相同。

根据 比尔-朗伯定律 （Swinehart，1962），b 在公式（9）中定义为 $-\log(I/I_0)$，其中 I 是测得的强度，I_0 是样品未遮挡视野时的光束强度。**tomoCAM** 所使用的超参数为 $p = 1.2$、$\sigma = \frac{1}{0.7}$ 和 $c = 0.0001$。对于 gridrec，我们选择了二阶 Butterworth滤波器，截止频率设为 0.25，这对于同步辐射断层重建而言是典型设置。对于模体，我们分别为 Tomobank 数据集 ID 25（TB-25）和 Tomobank 数据集 ID 86（TB-86）的 **SVMBIR** 运行选择了 $T = 1$ 和 $\sigma_x = \{0.98, 2.1, 1.1\} \times 10^{-4}$，以生成与**tomoCAM**质量相近的重建图像。

我们在绘制图像和线剖面时，采用了与图 4 类似的模式。图像的第一行展示完整切片，随后是放大的区域，接着从每个放大区域的中间获取线剖面。鉴于 **tomoCAM** 和 **SVMBIR**在数学上具有相似性，我们预计，在进行彻底的超参数搜索后，两者将取得相近的结果。表 1 展示了各代码所用时间的比较。

		重建时间（秒）		
		tomoCAM	**SVMBIR**	gridrec
数据集	size			
模体	(128, 16, 2048)	93	810	0.21
TB-25	(400，128，2048)	862	12730	2.53
TB-86	(202，128，2448)	1210	14273	3.73

表1。 重建各种数据集所需时间的比较。对$\boldsymbol{tomoCAM}$ 和$\boldsymbol{SVMBIR}$ 均进行了*100*次迭代的计时。

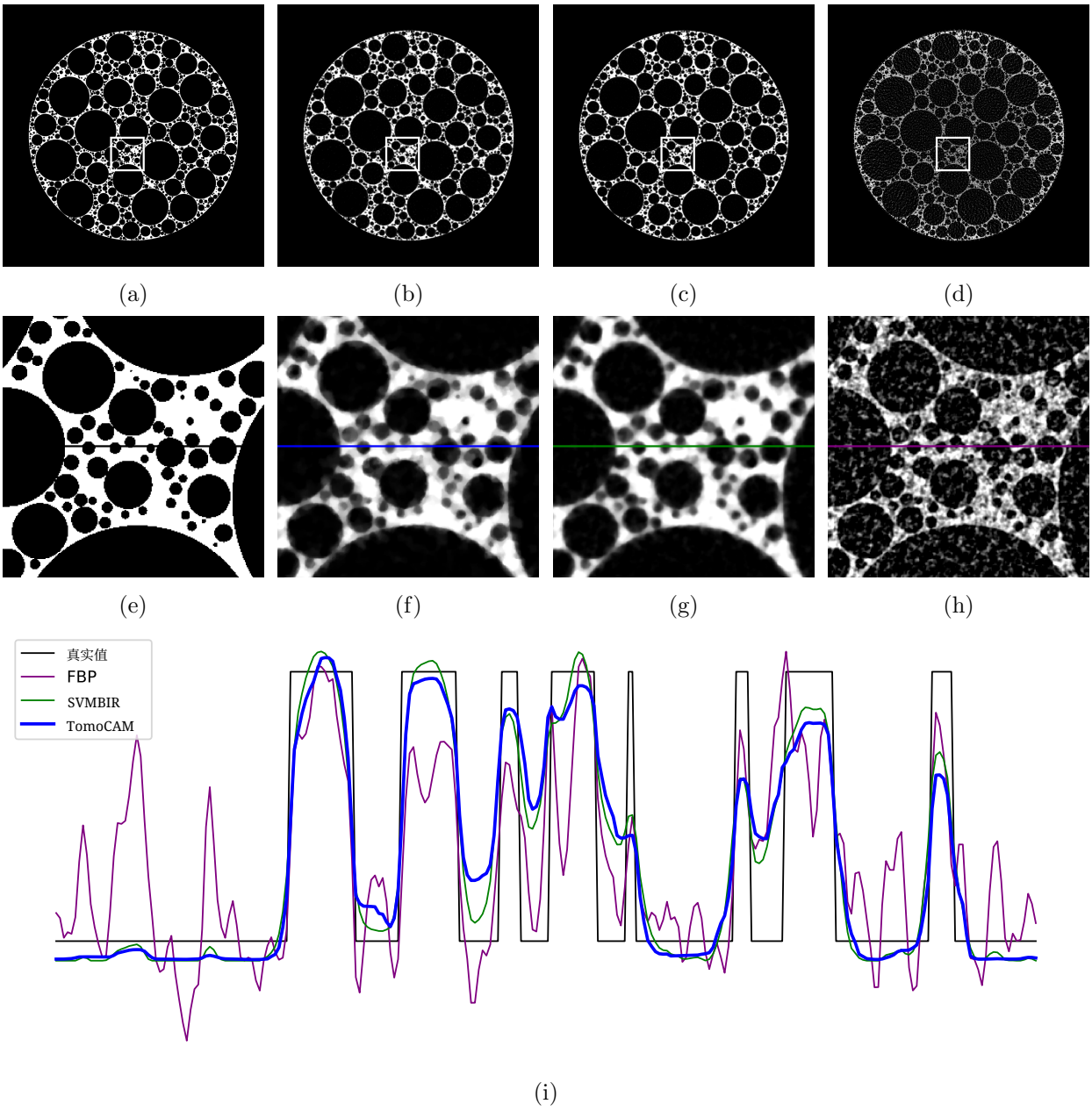

图4. 使用128个投影对泡沫仿体进行重建方法的比较：(a) 真实值，(b) **tomoCAM**，(c) **SVMBIR**，以及(d) gridrec。(e)、(f)、(g)和(h)分别是(a)、(b)、(d)和(e)中方框所示感兴趣区域的放大图，(i)显示了(e)、(f)、(g)和(h)中的线剖面。与gridrec相比，使用**tomoCAM** 和**SVMBIR** 进行重建所得图像的噪声较低。在此，**tomoCAM** 比**SVMBIR**快约 $9\times$ 。

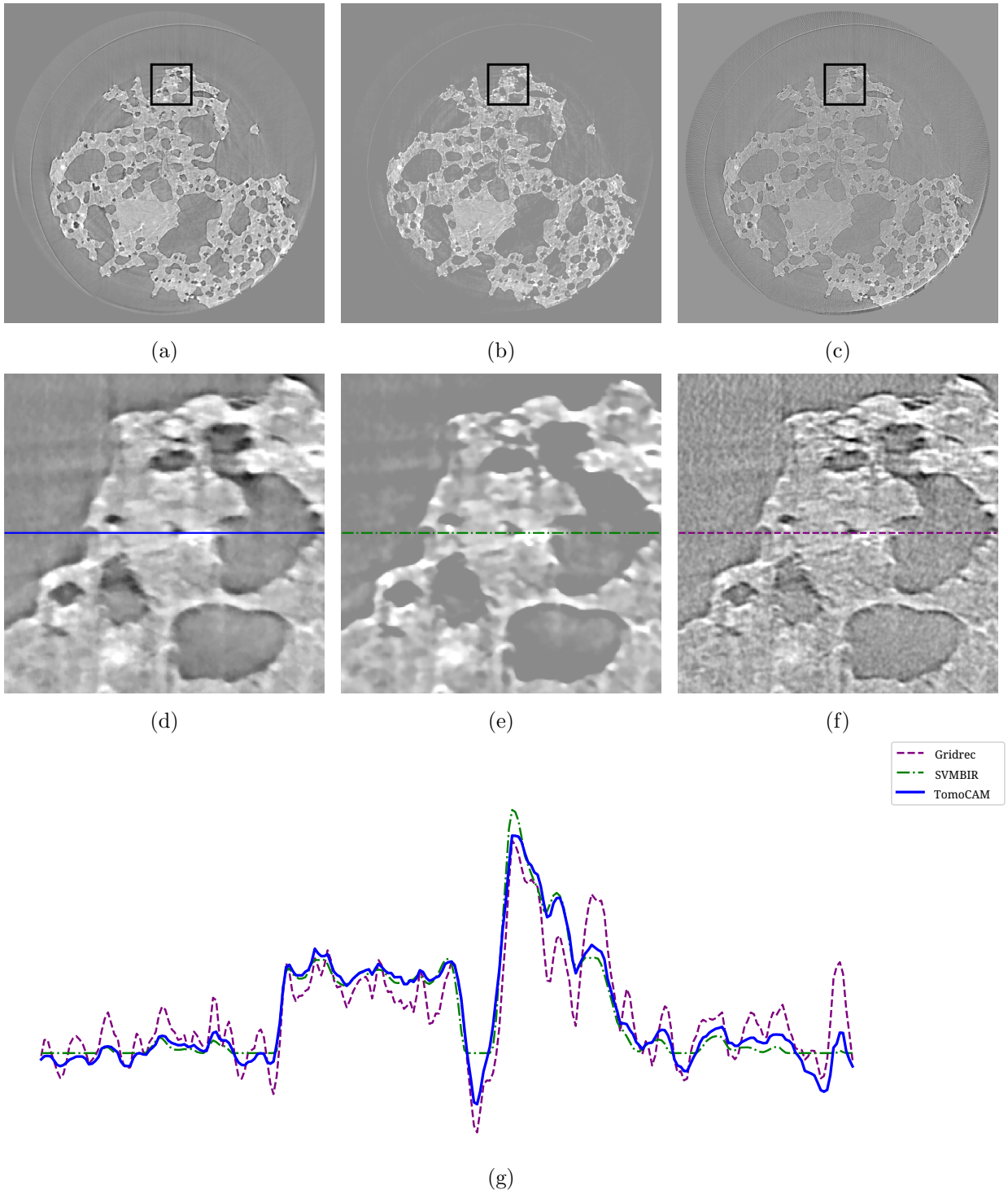

图 5. Tomobank 数据集编号：25 的重建图像：使用 400 个投影对岩石渗透率进行原位 研究，其中 (a) 为 **TomoCAM**，(b) 为**SVMBIR**，(c) 为 gridrec。(d)、(e) 和 (f) 分别为 (a)、(b) 和 (c) 中方框所示的放大感兴趣区域。(e) 显示了 (d)、(e) 和 (f) 的线剖面。所有重建图像均应用了圆形掩膜。与 gridrec 相比，虽然 **tomoCAM** 和 **SVMBIR** 都能有效抑制噪声，但**tomoCAM** 的速度约为 **SVMBIR** 的 $15\times$ 倍。

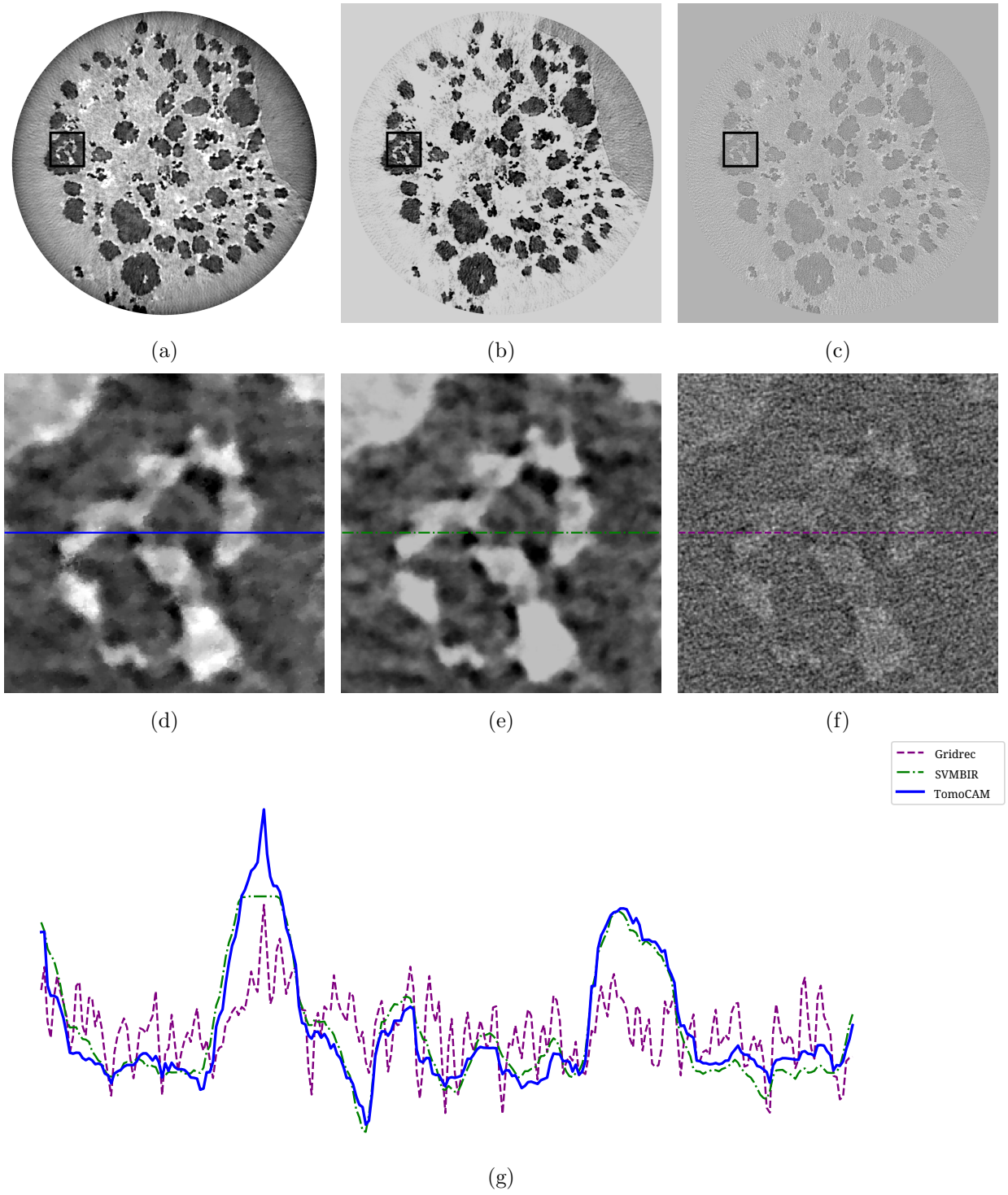

图 6。TomoBank 数据集 ID：86，使用 202 个投影进行稀疏投影角度纳米计算机断层成像数据的重建结果，其中 (a) 为 **tomoCAM**，(b) 为 **SVMBIR**，(c) 为 gridrec。(d)、(e) 和 (f) 分别是 (a)、(b) 和 (c) 中方框所示感兴趣区域的放大图。(e) 显示了 (d)、(e) 和 (f) 的线剖面。所有重建图像均应用了圆形掩模。与 gridrec 相比，**tomoCAM** 和 **SVMBIR** 在抑制噪声方面都表现出色，而 **tomoCAM** 的速度约快 15× 倍。

6. 结论

本文介绍了 tomoCAM，这是一款新型 GPU 加速软件，可用于重建高质量的断层成像图像。**tomoCAM** 能够在硬件要求相对适中的情况下，于合理时间内对大型数据集进行基于模型的迭代重建。与常用的滤波反投影方法相比，其重建图像噪声更低，速度则比仅 CPU 的 MBIR 实现快一个数量级。

我们为**tomoCAM**开发了一个基于 Python 的前端，专门用于接收 NumPy 数组作为重建的输入和输出。这使**tomoCAM** 能够无缝集成到束线科学家的现有工作流程中。尽管在可用投影数据稀缺时，使用 MBIR 尤为有利，但目前的 MBIR 实现非常耗时。因此，即使 MBIR 显然具有优势，这也是束线科学家不使用它的主要原因。**tomoCAM** 解决了这一问题，使 MBIR 重建更具实用性，具体方式是

- 提高效率：运行时间缩短了一个数量级，使其速度快于之前的 MBIR 版本，
- 降低硬件要求：可在配备 GPU 的普通台式机等小型设备上运行，从而提高了可及性，
- 简化超参数搜索：**tomoCAM** 的速度使超参数搜索更加便捷，从而能够更快速、更高效地开展实验，
- 增强兼容性：该实现提供了 Python 接口，因此易于集成到现有的 FBP 工作流中。

致谢 本工作得到了能源研究应用高级数学中心（CAMERA）的支持。该中心由能源部（DOE）科学办公室的高级科学计算研究和基础能源科学项目资助（项目编号：DE-AC02-05CH11231）。我们感谢橡树岭国家实验室的 S.V. Venkatakrishnan 提供了许多宝贵的讨论。我们还要感谢同事 J.A. Sethian、Z. Hu 和 K. Pande 审阅本稿件。

参考文献

ASTM-E1441-19,（2019年）。计算机断层扫描（*CT*）标准指南。标准。美国材料与试验协会，西康舍霍肯，宾夕法尼亚州。Barnett, A.（2021年）。应用与计算调和分析，**51**，1–16。Barnett, A.、Magland, J. 与 afKlinteberg, L.（2019年）。*SIAM* 科学计算杂志，**41**，C479–C504。Butterworth, S. 等。（1930年）。无线工程师，**7**（6），536–541。Carlo, F. D.、G¨ursoy, D.、Ching, D. J.、Batenburg, K. J.、Ludwig, W.、Mancini, L.、Marone, F.、Mokso, R.、Pelt, D. M.、Sijbers, J. 与 Rivers, M.（2018年）。测量科学与技术，**29**（3），034004。**网址：** *https://doi.org/10.1088/1361-6501/aa9c19*Crowther, R. A.、DeRosier, D. J. 与 Klug, A.（1970年）。伦敦皇家学会会刊。*A*辑：数学与物理科学，**317**（1530），319–340。**网址：** *https://royalsocietypublishing.org/doi/abs/10.1098/rspa.1970.0119*Dienemann, L. L.、Geller, L. C.、Huang, Y.、Zenyuk, I. V. 与 Panzer, M. J.（2023年）。*ACS* 应用材料与界面，**15**（6），8492–8501。PMID：36719130。**网址：** *https://doi.org/10.1021/acsami.2c20787*Dutt, A. 与 Rokhlin, V.（1993年）。*SIAM J. Sci. Comput.* **14**，1368–1393。Fessler, J. 与 Sutton, B.（2003年）。*IEEE* 信号处理汇刊，**51**，560–574。Forna-Kreutzer, J. P.、Ell, J.、Barnard, H.、Pirzada, T. J.、Ritchie, R. O. 与 Liu, D.（2021年）。材料与设计，**208**，109899。**网址：** *https://www.sciencedirect.com/science/article/pii/S0264127521004524*French, J.、Dahlkamp, C.、Befus, E. 与 Czabaj, M. W.（2022**224**，109453）。复合材料科学与技术，。**网址：** *https://www.sciencedirect.com/science/article/pii/S0266353822001956*

Ge, M.、Coburn, D. S.、Nazaretski, E.、Xu, W.、Gofron, K.、Xu, H.、Yin, Z. 与 Lee, W.-K.（2018年）。《应用物理快报》，113（8），083109。网址：https://doi.org/10.1063/1.5048378 Giselsson, P. 与 Boyd, S. P.（2014年）。载于第53届IEEE决策与控制会议，CDC 2014，美国加利福尼亚州洛杉矶，2014年12月15–17日，第5058–5063页。IEEE。网址：https://doi.org/10.1109/CDC.2014.7040179 G.N., R. 与 A.V., L.（1971年）。《美国国家科学院院刊》。Greengard, L. 与 Lee, J.-Y.（2004年）。《SIAM评论》，46（3），443–454。G¨ursoy, D.、De Carlo, F.、Xiao, X. 与 Jacobsen, C.（2014年）。《同步辐射学报》，21（5），1188–1193。网址：https://doi.org/10.1107/S1600577514013939 Harris, C. R.、Millman, K. J.、van der Walt, S. J.、Gommers, R.、Virtanen, P.、Cournapeau, D.、Wieser, E.、Taylor, J.、Berg, S.、Smith, N. J.、Kern, R.、Picus, M.、Hoyer, S.、van Kerkwijk, M. H.、Brett, M.、Haldane, A.、del R´ıo, J. F.、Wiebe, M.、Peterson, P.、G´erard-Marchant, P.、Sheppard, K.、Reddy, T.、Weckesser, W.、Abbasi, H.、Gohlke, C. 与 Oliphant, T. E.（2020年）。《自然》，585（7825），357–362。网址：https://doi.org/10.1038/s41586-020-2649-2 Herman, G. T.（2009年）。平行束滤波反投影，第135–158页。伦敦：Springer London。网址：https://doi.org/10.1007/978-1-84628-723-7 8 Jakob, W.、Rhinelander, J. 与 Moldovan, D.（2017年）。pybind11——c++11与 Python 之间的无缝互操作。Https://github.com/pybind/pybind11。Kulkarni, D.、Normile, S. J.、Connolly, L. G. 与 Zenyuk, I. V.（2020年）。《物理学杂志：能源》，2（4），044005。网址：https://dx.doi.org/10.1088/2515-7655/abb783 Larson, N. M. 与 Zok, F. W.（2018年）。《复合材料A辑：应用科学与制造》，107，124–134。网址：https://www.sciencedirect.com/science/article/pii/S1359835X17304645 MacDowell, A. A.、Parkinson, D. Y.、Haboub, A.、Schaible, E.、Nasiatka, J. R.、Yee, C. A.、Jameson, J. R.、Ajo-Franklin, J. B.、Brodersen, C. R. 与 McElrone, A. J.（2012年）。载于《X射线断层摄影进展VIII》，S. R. Stock 编，卷8506，第850618页。国际光学与光子学学会，SPIE。网址：https://doi.org/10.1117/12.930243 Mohan, K. A.、Venkatakrishnan, S. V.、Drummy, L. F.、Simmons, J.、Parkinson, D. Y. 与 Bouman, C. A.（2014年）。载于2014年IEEE声学、语音与信号处理国际会议（ICASSP），第6909–6913页。Mokso, R.、Schlep¨utz, C. M.、Theidel, G.、Billich, H.、Schmid, E.、Celcer, T.、Mikuljan, G.、Sala, L.、Marone, F.、Schlumpf, N. 与 Stampanoni, M.（2017年）。《同步辐射学报》，24（6），1250–1259。网址：https://doi.org/10.1107/S1600577517013522 Nikitin, V.（2023年）。《同步辐射学报》，30（1），179–191。网址：https://doi.org/10.1107/s1600577522010311 Nikitin, V.、Tekawade, A.、Duchkov, A.、Shevchenko, P. 与 De Carlo, F.（2022年）。《同步辐射学报》，29（3），816–828。网址：https://doi.org/10.1107/S1600577522003095 NVIDIA、Vingelmann, P. 与 Fitzek, F. H.（2020年）。CUDA，版本：10.2.89。网址：https://developer.nvidia.com/cuda-toolkit Pelt, D. M.、Hendriksen, A. A. 与 Batenburg, K. J.（2022年）。《同步辐射学报》，29（1），254–265。网址：https://doi.org/10.1107/S1600577521011322 Shepp, L. A. 与 Logan, B. F.（1974年）。《IEEE核科学汇刊》，21（3），21–43。Shih, Y.、Wright, G.、And´en, J.、Blaschke, J. 与 Barnett, A. H.（2021年）。CoRR，abs/2102.08463。网址：https://arxiv.org/abs/2102.08463 Swinehart, D. F.（1962年）。《化学教育杂志》，39（7），333。网址：https://doi.org/10.1021/ed039p333 Tarantola, A. 与 Valette, B.（1982年）。《地球物理学评论》，20（2），219–232。网址：https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/RG020i002p00219 Team, S. D.（2020年）。基于超体素模型的迭代重建（SVMBIR）。可从 https://github.com/cabouman/svmbir 获取的软件库。Thuering, T.、Modregger, P.、Grund, T.、Kenntner, J.、David, C. 与 Stampanoni, M.（2011年）。《应用物理快报》，99（4），041111。网址：https://doi.org/10.1063/1.3618672

Trampert, J. & Leveque, J.-J.（1990年）。《地球物理研究杂志：固体地球》，**95**(B8)，12553–12559。**网址：** *https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/JB095iB08p12553*Venkatakrishnan, S. V.、Drummy, L. F.、Graef, M. D.、Simmons, J. P. & Bouman, C. A.（2013年）。载于由 C. A. Bouman、I. Pollak & P. J. Wolfe 编辑的计算成像 *XI*，第 8657 卷，第 75–86 页。国际光学与光子学学会，SPIE。**网址：** *https://doi.org/10.1117/12.2013228*Walsh, S. D. 、 Mason, H. E. 、 Du Frane, W. L. & Carroll, S. A. （2014年）。《国际温室气体控制杂志》，**22**，176–188。**网址：** *https://www.sciencedirect.com/science/article/pii/S175058361400005X*Zhang, R.、Thibault, J.-B.、Bouman, C. A.、Sauer, K. D. & Hsieh, J.（2014年）。《*IEEE*医学影像汇刊》，**33**(1)，117–134。