

# Hepsptycho 多 GPU 及 AI 并行软件开发进展

Saturday, 29 June 2024 14:00 (15 minutes)

Ptychography 是一种相干衍射成像技术, 能够提供在 X 光衍射波长极限内样品的相位分辨率重建图像。在光学实验中, 使用小孔或其它光学仪器将 X 光探针聚焦到待扫描样品上。样品每个预定位置的衍射图样由放置在样品 1m 距离 (或更远) 之后的探测器捕获。相邻两个扫描位置需要一定的重叠比例, 以此来保证被记录的实验数据能够多倍冗余, 为 ePIE 算法或者 DM 算法的重建提供足够多的信息。

针对 ePIE 算法和 DM 算法, 现有的开源软件并未提供 QT 界面及提供完备的工作流, 无法为用户提供友好的实验元数据参数的快速导入 GUI 及基于 CUDA C++ 的多 GPU 并行重建迭代过程。Hepsptycho 软件由高能物理研究所计算中心与多学科中心 B2 线站联合开发, 程序核心均由 CUDA C++ 实现, 以实现在 NVIDIA GPU 上的并行计算, 同时使用 MPICH 作为多 GPU 卡协同并行重建的框架, 实现了迭代收敛的计算时间随 GPU 增加而线性递减。其元数据由 HDF5 文件提供导入, 可以大大降低用户的软件学习成本。

在 ePIE 算法的基础上, 计算中心与 B2 线站初步改写了算法核心模块, 将其升级为快速稳定收敛的 mPIE 算法。

针对相干衍射实验数据量巨大而导致从硬盘读入缓慢的问题, Hepsptycho 借鉴开源软件的单进程内多线程读入经验而非多 MPI 进程并行读入经验, 大幅提高了 IO 的速度。

同时, 由计算中心及 B2 线站联合开发的 AI 快速重建模块 (名称暂定为 W1-Net 神经网络模块), 正在逐步集成到 Hepsptycho 软件中, 这为用户进行实际光学实验前的参数快速调试提供了强大的模拟支撑。

在国产化加速卡及软件框架应用方面, hepsptycho 正在联合中科曙光及百度的科学家, 尝试使用国产 DCU+ 百度 paddlepaddle 推理框架实现 AI 的训练及推理过程, 而非 NVIDIA GPU+ TensorFlow 的组合, 以实现“国产替代”的目标。

## Summary

**Primary author:** Dr 磊, 王 (IHEP)

**Co-authors:** 洋洋, 穆 (IHEP); 邢, 成焯 (IHEP); 广才, 常 (IHEP); 家瑞, 胡 (IHEP); 京燕, 石 (IHEP); 法制, 齐 (IHEP); 誉, 胡 (IHEP); 建利, 刘 (IHEP); 浩凯, 孙 (IHEP); 世园, 符 (IHEP); 锐, 刘 (IHEP); 阳, 杨 (IHEP); 爽, 王 (IHEP)

**Presenter:** Dr 磊, 王 (IHEP)

**Session Classification:** 先进光源数据与软件