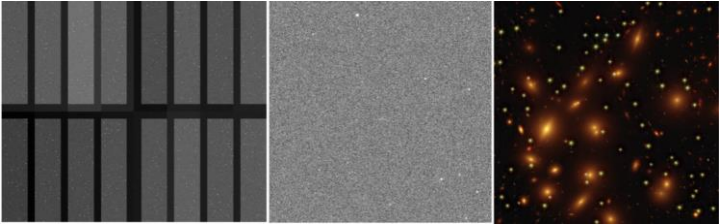


中期研制进展

多通道成像仪数据处理模块中期进展

多通道成像仪 (MCI)科学数据处理模块在 CSST 科学数据处理系统项目中中期评估中，已按计划完成从 0 级数据 (原始图像)到 1 级科学数据和星表数据，再到 2 级科学图像和星表数据的数据处理流水线开发。相关流水线已经用于 MCI 仿真数据处理和分析，完成以下关键技术：(1) 单次曝光图像处理相关工作包的流水线封装和算法开发和优化，包括仪器效应改正，位置定标，流量定标和星表提取工作包；(2) 图像合并工作包的开发和研制，针对超深场图像叠加进行优化；(3) 图像相减和暂现源搜寻工作包的开发和研制，实现快速准确进行暂现源搜寻功能。下图为多通道成像仪科学数据处理结果示意图。左图为输入 0 级图像，中图为 1 级图像，即单次曝光图像，右图为 2 级图像，即深场叠加图像。



积分视场光谱仪数据处理模块中期进展

积分视场光谱仪 (IFS) 数据处理模块在 CSST 科学数据处理系统项目中中期评估过程中，设计了包括宇宙线扣除、光谱追迹、文件格式、图像流量、光谱谱形、信噪比和波长定标等功能项和技术指标在内的 11 项评估项目，并通过仿真数据输入、数据定义文档和数据处理软件结果的背靠背比较，对 IFS 单次曝光处理、IFS 多次叠加及 IFS 科学数据处理三条流水线的主要功能项及两条主要技术指标进行了评估。

评估结果表明 IFS 数据处理模块已经完成了三条流水线的主要功能设计和软件实现。此外，通过仿真数据的验证表明，现有软件处理结果在现有硬件设计指标的基础上可以达到信噪比和波长定标两条主要技术指标的要求（见下表）。通过本次中期评估发现 IFS 数据处理软件模块在宇宙线扣除、流量定标和硬件参考文件等技术环节中存在需要进一步提高的地方，将在后续工作中改进。

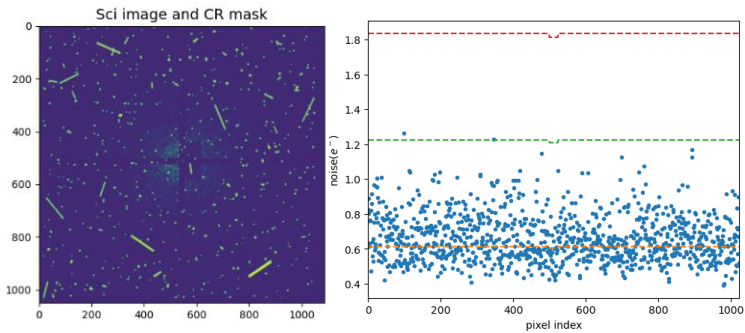
技术指标	评估细则	评估结果
在曝光时间满足单次300s并20次曝光累计情况下，数据处理系统应测量得到的天体的信号指标：0.2角秒的空间分辨单元内对B波段面亮度 ≤ 17 等(AB)/平方角秒的目标源，所得单个光谱单元在B波段有效波长处($\sim 440\text{nm}$)的信噪比 $S/N \geq 10$ 。	利用20次300s曝光叠加数据，检查B波段有效波长处的信噪比 $S/N > 10$ 的区域的表面亮度。	$S/N > 10$ 的空间区域轮廓与17.8星等/平方角秒轮廓大致重合。
在B波段定标的波长不确定度(RMS)优于0.035nm	对于给定波长处附近的灯谱线，采用灯谱数据中的N-1条灯谱进行波长定标，并和标准灯谱波长进行对比。	蓝端图像中达到的波长定标精度可以优于0.2Å；而红端图像由于可用定标谱线较多，波长定标精度可以优于0.1Å。

星冕仪数据处理模块中期进展

系外行星成像星冕仪 (CPI-C) 数据处理模块在 CSST 科学数据处理系统项目中中期评估中，已完成星冕仪数据处理流水线的初步开发及测试工作，顺利通过了中期评估。

星冕仪数据处理流水线可以从星冕仪 0 级数据（原始数据）出发，进行相机本底、平场、暗电流、坏点等仪器效应以及宇宙线和天空背景等观测效应的改正，生成星冕仪 1 级数据产品。目前软件功能需求分析、流水线与数据库和调度系统的接口设计、各类型数据产品定义等工作均已完成。星冕仪流水线已初具功能，完成了仪器效应改正、宇宙线识别和标记、1 级数据生成，0 级和 1 级数据质量检测等多个数据处理功能的开发和封装集成，并与数据库与调度系统等支撑系统完成集成。团队基于仿真数据对流水线开展了单元测试、功能测试、性能测试等测试，各项指标均满足项目要求。其中宇宙线识别模块基于多幅图像，识别率 $>90\%$ ；数据处理后残余噪声不大于 3 倍理论噪声，满足性能指标要求。下图左图为宇宙线标记结果；右图为数据处理后的残余噪声以及和理论噪声。

除数据流水线外，团队还开展了星冕仪观测图像仿真、仪器定标算法、相机性能测试和数据分析等工作。星冕仪数据处理团队后续将结合仪器测试数据，深入开展仪器定标和流水线性能提升等方面的工作。



太赫兹数据处理模块中期进展

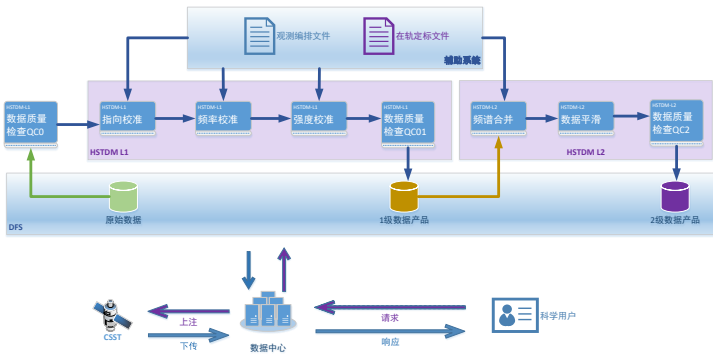
太赫兹（HSTDM）数据处理模块在 CSST 科学数据处理系统项目中期评估完成的研制工作如下：

1、1 级数据处理流水线：完成了各功能模块的代码编写（包含包括了指向校准、频率校准、强度校准、数据质量检查 QC0 和 QC1）及单元测试，分别完成了 0 级数据和 1 级数据的格式定义及其 toml 文件编写，1 级数据格式定义及其 toml 文件编写。其中，1 级数据流水线代码遵循 PEP8 规范，单元测试包含了 8 项测试用例，覆盖了 QC0、QC1 及流水线工作包测试。流水线工作包的测试是以仿真数据作为输入，输出是经过 1 级流水线处理生成的 1 级数据，数据通过程序判读与人工复查确保符合太赫兹 1 级数据定义。0 级仿真数据生成考虑了接收机的中频响应特性、宇宙微波背景 CMB 以及人为可调的随机高斯噪声。

2、2 级数据处理流水线：完成了频谱合并功能模块代码编写，可以实现 1）依据 obsid 检索 1 级流水线的相关数据产品；2）对通过 QC1 检查的 1 级流水线产品中的同一天区的频谱数据进行合并；3）将合并后的频谱数据生成初步定义的 2 级数据产品。

3、性能测试：通过模拟生成 500 组观测数据（包括 450 组 ON 点观测数据，25 组 BLK 类型数据，25 组 OFF 数据），模拟 500s 有效观测时间内产生的数据，在现有运行平台上 1 级流水线总耗时约为 14min。1 天内 HSTDM 模块的有效观测时长不到 12h（依据 HSTDM 观测编排效率不足 50%估算），则处理所产生的 0 级数据时间小于 21h，符合小于 24h 的性能要求。

下图为太赫兹数据处理模块分解示意图。



数据系统研制进展

序号	模块名称	月度进展
1	多色成像	在仿真图像中加入变化的卫星轨迹，有形状和亮度的周期性变化，进行了卫星轨迹算法更新；开展了测光红移测试。包括输入星表、输出合并星表的测光红移拟合；畸变和位置定标算法正在更新，将畸变固定，尝试不同密度星场在此基础上进行位置定标；协助 QC2 进行合并星表检验；与在轨定标，MCI 共同开展 CTE 改正测试，包括 HST 方法复现，Eculid 方法测试等。
2	无缝光谱	布署 Cycle8 研制任务并梳理现阶段的困难；完善无缝光谱模块相关功能、接口和测试内容，配合进行规范和优化；对 Mosaic 成像至伪直接成像生成的问题数据进行溯源。
3	数据流管理	进行腾讯云的测试环境部署；进行 2 级科学数据文件的服务实现；测试 Apache Doris 作为 2 级科学数据的数据仓库。
4	流水线运行管理	按 Cycle8 计划，在腾讯云测试环境做 Cycle7 流水线部署并开展压力测试，测试负载加大，会导致非逻辑错增多，通过容错设置可解决；流水线逻辑主控部分，对 UI 功能调整做相关设计。
5	多通道成像仪	组织 Cycle8 的各条流水线相关工作；完成流水线功能-接口-测试汇总表；合作组织仪器效应公共模块讨论，开展 CTE 效应改正。
6	积分视场光谱仪	对单次曝光开始添加背景光扣除功能；多次曝光合成算法中，采用仿真方法检验 LSF 和 PSF 的变化；开展流水线 API 接口检查和修改。

序号	模块名称	月度进展
7	星冕仪	开展星冕仪可见相机测试、数据分析、定标方案优化；进行 CTE 效应仿真算法优化；进行数据处理任务分配和细化；开展 API 接口优化工作。
8	太赫兹	对流水线规范化进行梳理；进行太赫兹在轨定标工作的部署。
9	天体测量	完成流水线接入准备，包括流水线梳理、gitlab 使用、公共模块下载安装；完成算法模块的 Python 格式梳理；对接口进行讨论。
10	观测数据仿真软件	基本实现新版的仿真流程，优化内存消耗，仿真效率提高一倍，并基于新版仿真软件完成了 2 平方度的仿真测试。升级了无缝光谱多级像的 PSF 模块，添加了实验室 pre/over-scan 和 16 通道新格式输出模块，开发测试 CTI 模块，持续开展仿真软件的运行服务。
11	数据可用性标记	QC 对 Cycle8 研制任务进行梳理，进行人员整合；QC2 对合并星表进行了初步检验，对发现的问题进行梳理及讨论。
12	在轨定标	定标星场的可观测时间的估算和统计分析（初选的 9 个星场）；计算适合定标观测的星场坐标范围，整理定标星场的量化选取需求，为下一步迭代选取定标星场提供参考；在轨测试阶段的在轨定标测试的讨论和规划；CTE 建模方法的调研整理；在轨定标数据处理流水线代码优化；MCI 和 IFS 模块正在进行实验室测试数据分析，以此分析结果来优化在轨数据处理流水线，同时获取初步的定标参数和参考文件。
13	观测需求编排	正在基于国际虚拟天文台联盟 UWS 协议优化后台运行的编排任务的监控与控制；优化一般观测申请的展示逻辑及细节；优化网站前端显示布局；维护警报监测系统，增加对 GCN 的新版警报发布方式的支持。
14	流水线封装	对接腾讯云计算资源，正在开展腾讯云性能测试，对其进行各项性能指标进行测试，并对各个数据软件进行资源消耗监控；持续推进流水线优化设计，重新对流水线进行规范化设计是 Cycle8 的重要任务之一，目前已经对主巡天多色成像、无缝光谱、天体测量、星冕仪和太赫兹模块进行了函数级重构，特别强调了模块接口代码需要进行类型标注。

近期节点和计划

2023 年下半年 Cycle8 任务节点：

时间节点	计划任务安排	时间节点	计划任务安排
23 年 8 月	数据处理系统完成中期评审	23 年 11 月	在轨定标方案评审
23 年 8 月-23 年 12 月底	仿真软件更新版本	23 年 11 月	各模块完成单元测试
23 年 8 月-24 年 1 月	数据处理软件已有算法完善和优化	23 年 12 月	完成观测需求编排科学专家咨询
23 年 8 月-24 年 1 月	数据处理软件 2 级流水线开发	23 年 12 月	流水线完成集成测试
23 年 8 月-24 年 1 月	配合开展地面测试	23 年 12 月	根据鉴定件标定参数完成仿真软件的修改和小天区仿真数据生成
23 年 8 月-24 年 1 月	定标产品生成算法迭代和更新	23 年 12 月	完成第二次科学效能评估
23 年 8 月-24 年 1 月	观测需求编排软件的更新迭代	23 年 12 月	完成配置项软件设计评审
23 年 8 月-24 年 1 月	配合完成地面测试	23 年 12 月	完成 2 级数据定义专家咨询
23 年 10 月	流水线设计优化	24 年 1 月	完成 C8 阶段结束，完成总结和任务验收
23 年 11 月初	单元测试平台完善和优化		