

CSST 科学数据处理系统月报

第 6 期

主办方：中国空间站巡天望远镜科学工作联合中心

时间：2021 年 04 月 10 日



科学数据处理系统研制人员参加空间及地面光学红外天文望远镜定标系统学术研讨会

3 月 26-28 日，由中国空间站工程巡天望远镜长三角地区科学中心与中国科学院上海天文台联合承办的“东方科技论坛-空间及地面光学红外天文望远镜定标系统学术研讨会”在上海天文台召开。

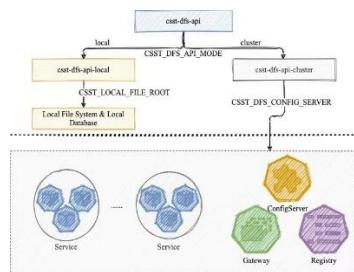
会议主题也是中国空间站工程巡天望远镜数据处理工作一直以来非常关心的技术问题，CSST 数据处理相关定标研究工作也已经开展了一段时间。因此，CSST 科学数据处理系统多名相关研制人员积极参加了本次会议。其中，科学数据处理系统在轨定标小组组长任娟娟以及邹虎、周志民、聂俊丹、苑海波、孟宪民、钟靖、陕欢源、刘凤山、沈世银、齐朝祥等人在会上针对相关议题进行了学术报告。

本次学术研讨会共同探讨了各类定标系统对科学产出的影响，并讨论了定标软硬件方面问题。本次会议瞄准空间和地面光学望远镜相关领域的国际前沿，积极推动我国在建的一批地面和空间望远镜的科学定标方案不断走向成熟，对提升我国在观测天文方面的关键核心技术具有重要意义。



数据流管理模块研制工作进展

数据流子系统(DFS)完成了第一个版本的设计与实现，API 原型 (CSST.DFS.API.LOCAL) 全部采用 Python 开发完成，提供在本地的文件数据提取与存储、数据库的存储与查询等相应服务，满足本地应用开发需求。DFS 基于微服务架构，后续系统将在第一版的基础上，进一步增加高性能实时数据库锥形检索、调用与访问接口扩充、容器封装等相关功能；完整的功能原型将于 2021 年 6 月 Cycle3 完成时实现与多色成像与 IFS 系统的对接。



多色成像模块研制工作进展

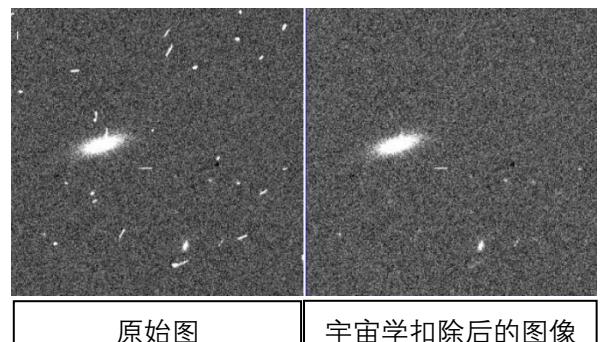
多色成像模块在 3 月份在多项研制工作上取得进展：

1. 完成 0 级数据说明文档：在 CSST 数据系统中首先完成了 0 级数据结构和接口设计说明，包括了对主巡天

数据格式、名称、存放方式、数据文件内容和文件头关键词的定义。（胡义，聂俊丹，张天萌，邵立）

2. 流水线原型研发进展：利用仿真图像数据，与数据流、数据库和 DQ 模块一起展开了 Cycle3 阶段流水线原型的研发。目前已完成探测器效应改正功能。宇宙线探测利用机器学习方法，在训练样本有限的情况下，获得了较好的结果（效果如下图所示）。（马斌，胡义，广州大学团队等）

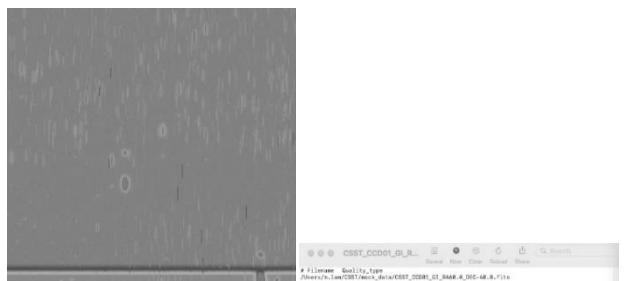
3. 流水线流程图绘制：梳理了整个多色成像数据软件的流水线流程，主要工作包括仪器效应改正，位置定标，流量定标到最后的星表提取，理清了各个工作包的工作流程（输入，输出，需要的外部数据，计算资源需求和实现的主要功能等）以及各个工作包之间的接口关系，为绘制完整流程图和完成系统设计方案文档奠定了基础。（罗常青等）



数据可用性标记流水线 Cycle3 原型工作进展

数据可用性标记模块作为数据系统接收和处理数据的前哨站，是数据处理过程和数据流中的第一步。截至目前，数据可用性标记完成了导星系统 0 级数据定义，QCO 数据流水线所处位置及其原型代码，以及 0 级数据质量标记列表原型等。

最近工作的一个亮点：确立了在数据处理流水线中所处的位置并构建了一个原型代码。在 Cycle3 阶段利用此代码我们可以完成仿真数据读取→判断数据完整性→成图→生成数据质量报告的全过程。通过判断和质量报告，未来将有助于监测还原观测时刻数据的生成状况和及时反馈存在的问题。目前本模块的质量报告内容还正在完善，争取在下一个阶段添加导星系统的更多判断。



仿真数据快速成图

仿真数据的数据质量报告

数据系统研制进展

序号	模块名称	月度进展
1	多色成像	完成 0 级数据说明文档；流水线原型研发工作完成探测器效应改正功能。
2	无缝光谱	完成兴隆基地 80cm 望远镜无缝光谱终端安装调试及初步观测；初步完成无缝光谱处理流水线各模块流程图设计工作；启动无缝光谱 1 级和 2 级数据定义工作；启动面向无缝光谱仿真数据处理流水线实际搭建工作。
3	数据流管理	完成数据流子系统第一个版本的设计与实现；初步完成与 IFS 团队开展的基于数据流子系统的第一 次联调工作；完成面向未来数据处理应用需求的集群数据流子系统 CSST. DFS. API. CLUSTER 的初步设计。
4	流水线运行管理	结合 C3 流水线原型的应用需求，设计了 C3 流水线原型的集成框架；对流水线引擎执行模块的迭代改进，完善时间戳计量及日志输出；在流水线管理部分设计了容错机制，提高流水线引擎的可靠性；流水线前端显示的迭代改进及功能测试，增加堆叠图显示；基于已有功能改进，继续完善容器化流水线规范。
5	多通道成像仪	制定了 MCI 1 级/2 级数据接口设计说明的工作计划。
6	积分视场光谱仪	梳理了 Cycle3 阶段的任务要求并制定任务计划；启动 Git 项目版本管理；完成了标准星模拟的科学图像仿真。
7	星冕仪模块	完成并提交了星冕仪 0 级数据的初步定义；完成了星冕仪波前定标图像仿真。
8	太赫兹模块	继续完善 0 级数据的定义和数据处理流程架构的优化。
9	天体测量	初步完成模拟观测图像所需的天测计算软件模块的联调，获得了 CSST 近 10 年的卫星平台在轨道上的地心天球参考系下的三维位置和速度仿真数据，实现了三维位置和速度与卫星轨道 6 参数的相互转换；实现了 fits 格式模拟星表数据的提取接口程序。
10	观测数据仿真软件	完成主巡天观测仿真，包括无缝光谱和主巡天成像共 500 个连续指向曝光，约 10 平方度；仿真数据已经交付数据处理团队进行分析，检验无误后将及时面向中国天文界公开发布。
11	数据可用性标记模块	完成导星系统 0 级数据定义，完成 0 级数据标记基本定义；正在进行导星系统对巡天编排仿真设计工作。
12	在轨定标	初步完成 HST 主要在用终端设备 ACS/WFC3/COS 的在轨定标资料调研整理，正在列出 CSST 主巡天在轨定标内容。
13	观测需求编排	模块任务划分为通用功能、一般观测申请支持、观测需求编排系统、ToO 观测决策支持系统四大部分，并进行人员细化分工；初步理清了工作方法，开始进行各项任务的需求分析工作。

近期节点和计划

2021 年上半年 Cycle3 节点：

时间节点	计划任务安排	时间节点	计划任务安排
2 月下旬	启动科学观测需求编排软件部分的研制	5 月中	0 级数据定义通过科学专家评审
3 月初	多色成像、科学可用性评估、数据流管理、流水线运行管理等模块接口定义	5 月底	开展各个模块 1-2 级数据的初步定义
3 月	提供约 7 平方度满足 0 级数据定义格式的成像仿真数据	6 月中	完成 1-2 级数据初步定义的科学专家咨询
3 月底	所有模块完成 0 级数据的定义	7 月初	完成基于多色成像数据处理的完整流水线原型
4 月中	各个模块集中制定数据处理流水线的详细定义	7 月	完成系统设计文档的评审