

# CSST 科学数据处理系统月报

## 第 8 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 06 月 10 日

### 新闻

#### 中国空间站工程巡天望远镜科学数据处理系统 召开第 2 次调度会

5 月 17 日，中国空间站工程巡天望远镜(简称 CSST)科学数据处理系统第 2 次调度会以线上线下相结合的方式在中国科学院国家天文台召开。来自国家天文台、上海天文台、紫金山天文台和南京天光所的指挥线领导和技术骨干成员参加了本次会议。会议由国家天文台副台长、科学数据系统指挥刘继峰主持。

会议首先由四家单位的负责领导分别介绍了各单位的科学数据处理系统工作进展情况，并提出工作中遇到的技术难题和需要协调的问题；然后大家针对提出的各种问题一一讨论、共同协商，最后提出解决方案。会后形成会议纪要。本次会议按计划进行，达到调度会的目的。CSST 科学数据处理系统调度会按季度举行，各单位表示将积极合作，全力以赴，共同推进 CSST 科学数据系统项目顺利进行。

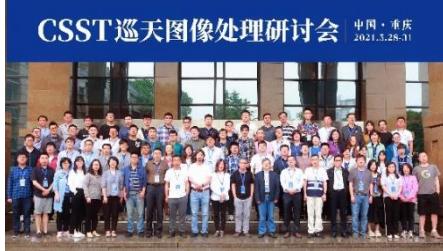


#### CSST 巡天图像处理研讨会在重庆举办

5 月 29 日-30 日，CSST 科学数据处理系统第一次“科学数据处理-科学研究”联合研讨会在重庆召开。本次会议由国家天文台联合中心主办，重庆邮电大学和重庆物理学会协办，会议主题聚焦主巡天图像处理技术方法。来自 CSST 科学数据系统和各个科学课题(包括宇宙学、星系和 AGN、恒星和银河系结构、星团、星族等)的骨干和研究人员约 100 人参加了本次会议，线上参与研讨的研究人员 200 人次。

本次会议的目的是加强 CSST 科学数据处理系统与科学课题之间的交流合作。会上，CSST 科学数据处理系统和各科学课题就图像处理各层面问题开展研讨，提出了互相的需求，最后就存在的问题和需要加强合作的方面展开了深入的讨论。本次会议促进了科学数据处理系统研制和科学课题研究的顺利开展。会议报告录像已整理发布在联合中心网站。

### CSST 巡天图像处理研讨会



### 突出进展

#### 无缝光谱模块研制工作进展

为了研发出一套适用于 CSST 大视场无缝光谱巡天的自动化光谱处理软件，无缝光谱处理团队核心成员汪霞、陈建军等面向 CSST 无缝光谱仿真团队提供的模拟数据，通过对哈勃空间望远镜 aXe 无缝光谱处理模块进行改进，成功实现 CCD 图像预处理及多目标自动化光谱提取。

新研发的无缝光谱处理模块能实现对 CSST 无缝光谱 0 级图像进行多通道合并、本底、暗流、平场等仪器效应改正(如图 1)，并能对多个目标源的位置、形态准确进行自动化识别及扣除天光背景，进而成功得到所有探测目标源(恒星和星系)的一维光谱(如图 2)。

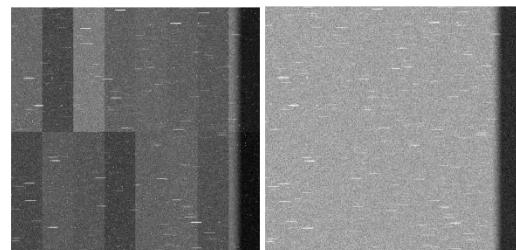


图 1：左图为 CCD 预处理前，右图为 CCD 预处理后。

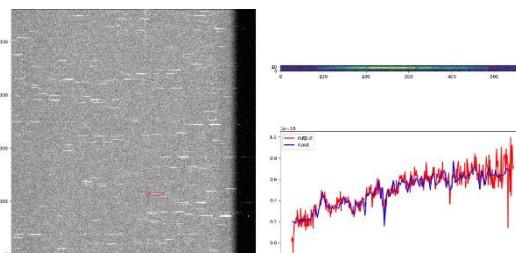
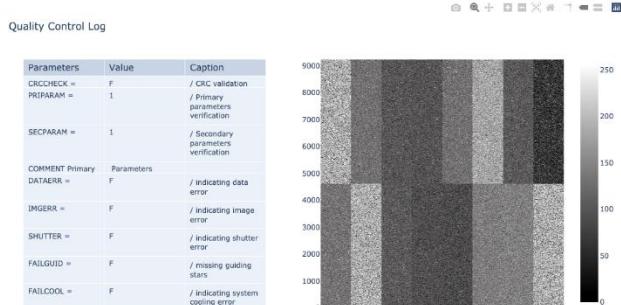


图 2：左图是经过预处理的光谱图像，其中红框区域是一个星系的 1 级像；右图-上是目标光谱抽图，右图-下是输入和抽出光谱对比。

### 数据可用性标记模块研制工作进展

数据可用性标记模块完善了可视化界面的快捷功能，包括简单的存档、放大缩小展示、区域选取、拖拽操作、reset 坐标系等功能，同时还可展示可用性标记列表参数。




**数据系统研制进展**

序号	模块名称	月度进展
1	多色成像	利用仿真数据完成了探测器效应改正、位置定标和流量定标的模块运行检验；在流水线检验过程中发现并反馈了仿真数据的问题；完成 2 级数据产品定义初稿。
2	无缝光谱	完成无缝光谱 1、2 级数据定义的第二个版本；完成无缝光谱处理流水线的定义及系统分析方案；面向 CSST 仿真数据成功实现图像预处理及多目标自动化光谱提取。
3	数据流管理	完成数据库结构的详细设计，在分布式环境下部署数据库；面向 Cycle3 流水线的需求，模拟了数据生成，导入数据库；完成面向 Cycle3 流水线的观测记录和 0 级数据数据访问服务，及相应的 API 访问的设计和实现；仪器效应改正的编码实现。
4	流水线运行管理	结合 Cycle3 原型需求继续软件原型的迭代改进。调研容器化的 webssh、jupyter，为后期的在线应用调试做技术储备；完善流水线引擎的网络定义，调整启动模板，支持参数化配置；定义周期性任务 cron 的通用容器镜像，支持后期的各类系统级、应用级任务集成；设计了跨节点数据传输的流控实验，基于 postgresql 存储过程实现，验证技术可行性；完善流水线管理 UI 界面，增加手动布局功能。
5	多通道成像仪	完成 MCI 仿真处理软件关于亚像素效应的功能更新；完成 MCI 一级数据产品及入库标签的初步定义。
6	积分视场光谱仪	完成旋涡星系 NGC6217 的科学仿真、图像仿真、单次曝光处理；完成一级数据输出接口的初步说明。
7	星冕仪	进行 HST 数据处理并改进行程；进行数据处理软件设计和模块定义；仿真按照 0 级数据 fits 格式做出接口，仿真过程中做一些日志存储、信息显示、输入文件等工作。
8	太赫兹	开始对去年完成的程序代码标注；通过改善观测仿真程序算法，使得计算效率提高了将近 100 倍，大大加快观测源的观测仿真进程。
9	天体测量	完成了 CSST 天体测量数据处理接口详细设计，给出了主巡天天体测量数据头文件信息详细说明，形成了技术说明文档。
10	观测数据仿真软件	完成针对 Cycle3 仿真数据分析测试出现的问题于近期完成了对相关模块的更新升级，在紫台超算平台开始运行本轮仿真的升级数据；进一步讨论了下一阶段主巡天仿真需要考虑的模块升级内容。
11	数据可用性标记	流水线方面新增了可视化原型功能；QC 模块对数据库提出输出表格需求，定义为与 obsid 绑定。
12	在轨定标	讨论初步列出的 CSST 正式巡天阶段的主巡天主要是探测器相关方面和多色成像方面的在轨定标内容，根据建议进行修改并进一步完善；调研 Roman 在轨定标资料；将进一步讨论正式巡天阶段的主巡天无缝光谱方面的在轨定标内容及试运行阶段的在轨定标内容；进行 Euclid 和 JWST 在轨定标调研。
13	观测需求编排	基于 Python-Flask 搭建了系统框架，见 <a href="https://nadc.china-vo.org/csst-bp/">https://nadc.china-vo.org/csst-bp/</a> ，目前已支持浏览器操作、oauth2 单点登录、用户-角色-操作 三级权限控制等机制等功能。开发组将基于该框架进行原型开发，逐渐迭代明确系统功能及内外接口。项目组目前已开展系统分析设计文档撰写工作，根据理解的深入不断修改丰富。


**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

编辑：于清 3月

审核：刘超、王慎

签发：刘继峰

刘继峰