

CSST 科学数据处理系统月报

第 16 期

主办方：中国空间站巡天望远镜科学工作联合中心

时间：2022 年 3 月 10 日

简讯

CSST 科学数据处理系统 2021 年度总结研讨会召开

2022 年 1 月 13-15 日，载人空间站工程巡天空间望远镜（CSST）科学数据处理系统 2021 年度总结研讨会以线上线下相结合的方式召开。此次会议分为北京、上海和南京三个分会场。来自中国科学院国家天文台、上海天文台、紫金山天文台、南京天光所、长春光机所、计算机网络中心、中国空间技术研究院钱学森实验室、北京大学、广州大学、昆明理工大学、云南大学、中山大学、北京师范大学、河北师范大学、上海交通大学等 16 个单位 100 多名研制团队成员参加了本次会议。国家天文台副台长刘继峰出席会议并致辞。应用总体也通过线上全程参加了会议并给出重要指导。

经过全体研制人员 2021 年的不懈努力和积极协调，科学数据处理系统各个软件配置项和模块的研制，取得了显著进展，基本完成了既定计划，多项重要技术研究取得突破。共约 40 多位系统研制骨干人员汇报了 2021 年度研制工作进展、阶段成果、软件工程化规范等内容；最后，会议对下一研制周期的工作计划进行了部署和安排。整个会议过程中来自各模块的研制团队成员进行了充分的交流和讨论，为项目下一年的顺利开展奠定了良好基础。



CSST 科学数据处理系统软件系统设计说明内部评审会召开

2022 年 1 月 20 日，中国科学院国家天文台在北京组织召开了载人空间站工程巡天空间望远镜科学数据处理系统软件系统设计说明的内部评审会。评审组听取了项目组做的《载人空间站工程巡天空间望远镜科学数据处理系统软件系统设计说明》报告，备查了三个配置项任务书，形成了评审意见。项目组目前正在对问题进行闭环。

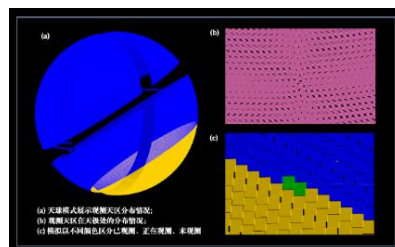
CSST 科学数据处理系统质量管理小组启动会召开

2022 年 1 月 19 日，国家天文台组织召开了软件工程化管理培训和暨质量小组启动会。由总体部专家给大家做了关于软件工程与 KJZ 软件研制管理要求的讲座，然后针对软件工程管理的内容和项目过程中质量管理的内容进行了讨论。

突出进展

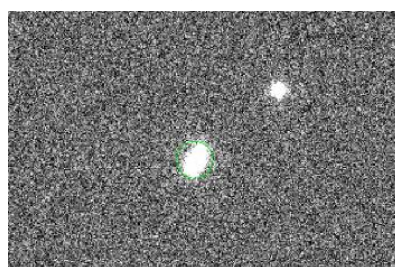
观测编排软件模块研制工作进展

根据 CSST 观测使用的 CCD 设计和目前观测天区划分，实现在万维望远镜平台对观测天区的动态可视化展示，可分为天球模式和观天模式两种可视化视角。直观呈现 CSST 对全天观测的覆盖情况。考虑到观测人员的实际需求，天区可视化呈现结果能够根据观测计划和项目状态、项目进度和当前时间进行动态显示，以不同颜色区分已观测、正在观测、未观测不同类别，或手动选择显示观测天区范围，使可视化信息更直观。未来将进一步与其他模块进行对接，考虑对观测天区进行标签化管理，使其更具有参考价值。



天体测量模块研制工作进展

天体测量团队近期开展了太阳系移动天体图像的模拟仿真工作，基于 galsim 软件采用亚像素移动法实现了 CSST 对太阳系移动天体观测图像的模拟，已交付第一批测试数据给 CSST 数据处理团组测试使用，下图展示了太阳系小行星 Tomatic 在 CSST 凝视恒星的观测模式下的模拟图像。团队还利用 Gaia 超高精度天体测量数据，筛选得到了 199 个极有可能为双星类星体或透镜型类星体的源，为 CSST 天球参考架的建立提供重要基准源。针对银盘上超密集星场图像所面临的目标检测难题，正在开展集中攻关。团队内部组建了 3 个并行小组，已对他人和自研的多种目标检测方法进行了研究和试验，不断提高目标检测率。



数据系统研制进展

序号	模块名称	月度进展
1	多色成像	在阿里云服务器上调试优化了单次曝光数据处理流水线, 在单个计算节点下 3 小时可以获得一次曝光 18 个探测器的图像和星表; 在轨定标流水线完成了部分测试天区的数据处理; 利用 HST 数据产生了 CSST 视场大小的近邻星系 M81 的近似 g 波段合成图像, 后续开展数据处理尝试; 开展宇宙线探测公共模块的搭建和说明文档的撰写。
2	无缝光谱	制定 Cycle5 阶段任务计划及具体实施方案; 完成 Cycle5 阶段研制任务具体部署工作。
3	数据流管理	完成数据流管理 Cycle4 总结和 Cycle5 计划; 开始推进测试环境硬件升级的采购工作; 在 4 节点的测试环境中部署 YugabyteDB 数据库, 用于数据处理业务库的测试; 正在开展 IFS 模块的数据导入工作。
4	流水线运行管理	在前期一级流水线的 Cycle3 原型 (多色成像)、Cycle4 原型 (无缝光谱) 基础上, 完善流水线引擎功能, 支持云环境下的多节点调度, 在 Cycle5、Cycle6 阶段支持二级流水线; 参与集成讨论, 构建二级流水线集成框架, 简化流水线业务逻辑, 重点关注云上的数据级并行; 讨论 Cycle5 阶段计划。
5	多通道成像仪	完成 Cycle5 软件开发计划; 开展宇宙线相关公共模块工作; 生成图像相减/暂现源所需预处理图像; 开展多色成像-MCI 公共模块讨论。
6	积分视场光谱仪	开发环境迁移至技改专用服务器; 各模块以周为单位进行细节功能更新; 启动 IFS 科学数据处理中宽线成分分解模块。
7	星冕仪	完成 Cycle4 阶段任务总结报告及 Cycle5 阶段任务开发计划。
8	太赫兹	完成太赫兹模块数据处理 Cycle5 阶段任务书; 完成《数据处理软件配置项 Cycle5 软件开发计划》中太赫兹模块在轨定标部分的相关内容。
9	天体测量	太阳系移动天体仿真图像的模拟实现了基于 galsim 软件太阳系移动天体仿真图像的模拟, 已交付第一批测试数据给数据处理团队测试; 密集星场目标检测工作由 3 个团队并行对多种目标检测方法进行了研究, 目标检提取率不断提高, 正在开展研讨和新方法研发; 基于天测参数进行类星体分类识别方法方面取得进展。
10	观测数据仿真软件	参与年度总结研讨会汇报并讨论了观测仿真软件的年度工作进展及下一研制周期各功能模块的升级和优化, 展示了观测仿真软件的发布原型及使用演示; 结合会议讨论进一步梳理并优化包含自定义星表、仪器参数和点扩散函数数据库等数据模块接口, 以确保 CSST 主巡天仿真程序的顺利发布; 调试仿真程序在线发布的用户接口和编写用户使用手册, 为仿真程序的发布和 CSST 科学预研究的顺利开展提供保证。
11	数据可用性标记	利用 Cycle4 仿真数据, 完成基于 0 级数据的统计工作, 图像背景的不均匀性与散点可通过阈值判断对坏点与偏离较大的宇宙线进行筛选。
12	在轨定标	进行 Cycle4 总结和 Cycle5 规划, 按计划逐步开展各项工作; 完成三个不同数密度星场不同曝光时间的仿真, 并使用集成的多色成像流水线完成了仿真数据的处理, 目前正在基于数据处理结果开展分析; 根据内审专家意见在系统分析方案里补充了星冕仪和太赫兹模块的在轨定标相关内容; 继续完善在轨定标方案; 为三月底的在轨定标内容专家咨询会做准备。
13	观测需求编排	完成 Cycle4 总结, 明确 Cycle5 阶段的相关任务; 调研了 CSST 在轨工作模式; 根据现有天区分划及仪器视场计算观测区域, 在万维望远镜平台中进行可视化呈现; 修改提案用户、专家权限设置方式, 修正部分权限及角色 bug。

近期节点和计划

2022 年上半年 Cycle5 节点:

时间节点	计划任务安排	时间节点	计划任务安排
1-6 月	实现统一程序框架并完成主要功能封装	3-6 月	设计和实现 2 级数据流水线标准 API 接口
3 月	完成在轨定标策略方案书咨询	3-6 月	开展探测器地面测试
3 月	完成仿真软件 V1.0 版本的发布	3-6 月	完成需求规格说明书编写
3 月	向数据中心提交数据库需求	5 月	完成主要公共模块的定义和开发
3 月	向总体部提交观测需求编排接口需求	6 月	实现完成完整 1 级数据处理流水线
3-6 月	开展系统软件设计	6 月	实现 2 级流水线原型
3-6 月	开展算法科学验证	6 月	实现编排生成功能、观测可视化平台功能、观测申请系统功能和警报系统功能的
3-6 月	软件外部接口定义	6 月	发布第二版仿真软件