



巡天空间望远镜 科学数据处理系统介绍

刘超 国家天文台

上海天文台、紫金山天文台、南京天光所；

北京大学、北京师范大学、长春光机所、广州大学、暨南大学、空间技术
研究院、天津大学、网络信息中心、云南大学等

2020年11月27日@厦门ChinaV0年会





概述

□ 载人空间站巡天空间望远镜（CSST）

- 2米口径空间望远镜
- 在轨运行十年间将对17500平方度天空进行高空间分辨率、高灵敏度巡天观测，获取百亿天体的多色图像及无缝光谱数据。
- 有望在宇宙学、星系演化、银河系等多个天体物理领域取得突破，将使我国光学天文研究在未来10-20年取得国际领先地位





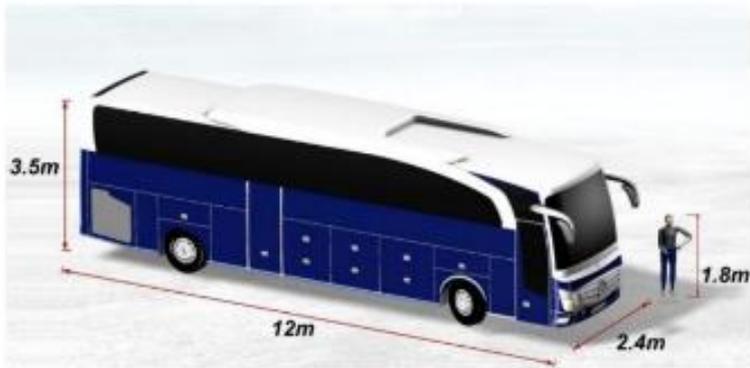
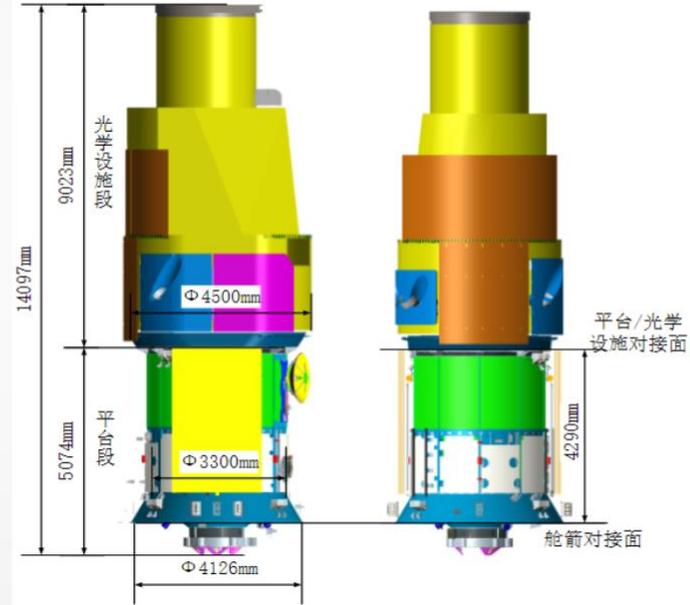
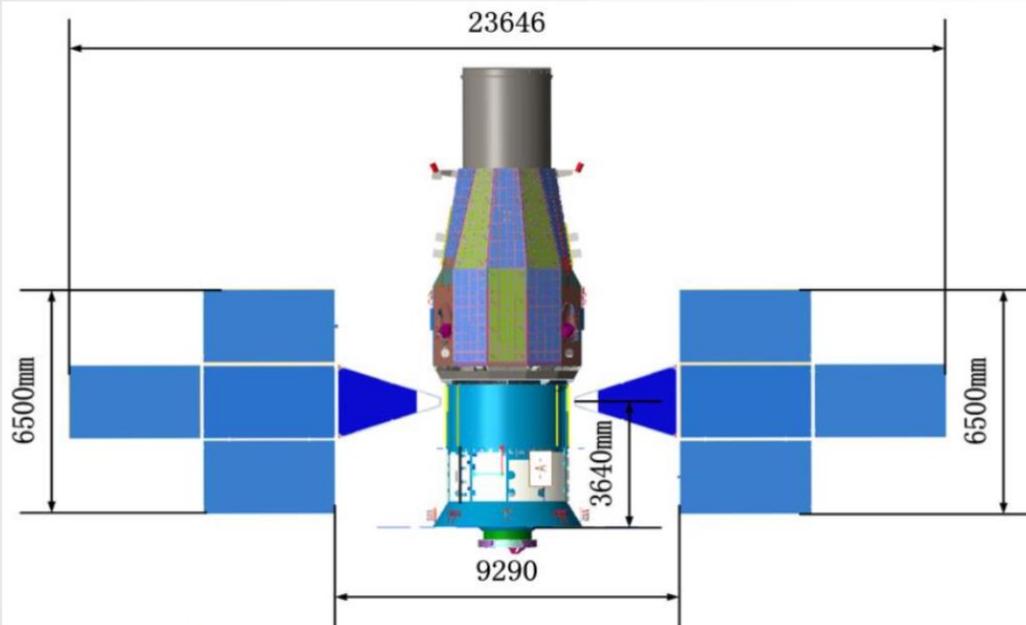
视场是HST-ACS的350倍



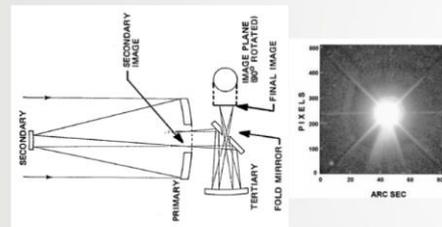
- CSST ($1.1^\circ \times 1^\circ, 1 \text{deg}^2$)
- EUCLID ($0.79^\circ \times 0.71^\circ$)
- WFIRST ($0.79^\circ \times 0.43^\circ$)
- HST-ACS ($202'' \times 202''$)
- HST-WFC3 ($160'' \times 160''$)



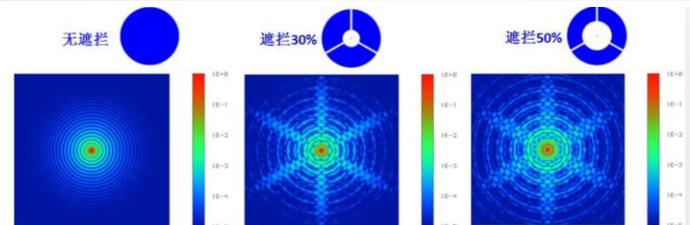
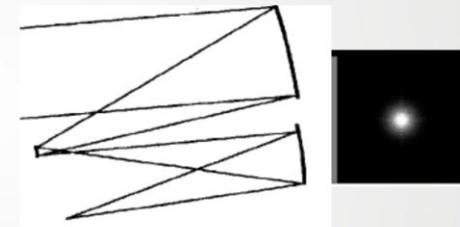
巡天空间望远镜结构



Coaxial system (Korsch TMA)



Off-axis system (COOK TMA)





后端模块

光学
设施

主光机系统 (长春光机所、上海技物所)

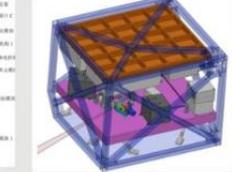
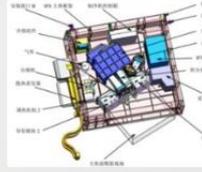
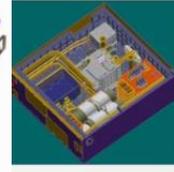
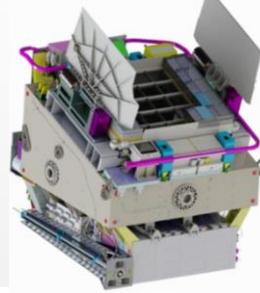
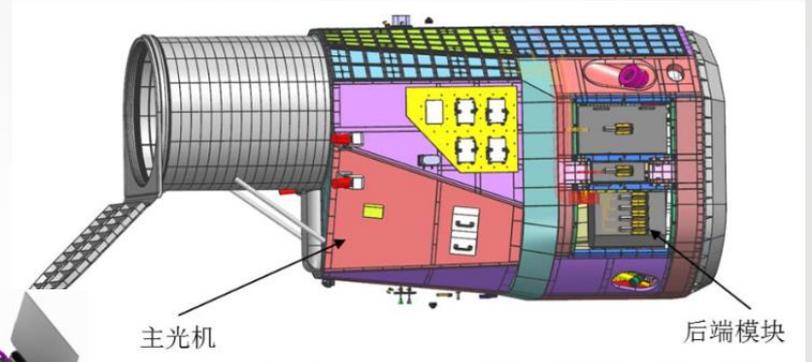
巡天模块 (光电所、国家天文台)

太赫兹模块THz (紫金山天文台)

多通道成像仪MCI (上海技物所、上海天文台)

积分视场光谱仪IFS (上海技物所、上海天文台)

星冕仪模块CPIC (南京天光所)



巡天模块

太赫兹模块

多通道成像仪

积分视场光谱仪

星冕仪模块

91%数据
70%时间

9%数据
20%时间



后端模块指标

巡天模块	波段	天区面积	分辨率	点源极限星等/等效宽波段成像观测星等
深度多色成像观测	NUV/u/g/r/i/z 短波红外	约17500平方度	平均0.15", 最优0.12"	$g \leq 26$, 其他平均 ≤ 25.5
极深度多色成像观测		约400平方度(不连续天区)		深度+1
深度无缝光谱观测	GU, GV, GI	约17500平方度	空间分辨率: 平均0.3", 最大0.4" 光谱分辨率: $R \leq 200$	$GU \leq 22$, $GV \leq 23$, $GI \leq 23$
极深度无缝光谱观测		约400平方度(不连续天区)		深度+1

MCI模块: 高精度测光&紫外可见光超深场

- 观测视场面积: $\leq 7.5' \times 7.5'$
- 观测深度: CBV: ≤ 30 等, CBU与CBI平均星等: ≤ 29 等
- 测光精度: $u \geq 2\%$, $r \geq 0.7\%$, $z \geq 1\%$

IFS模块: 延展目标源精细成像光谱观测

- 视场面积: $\leq 6'' \times 6''$
- 切割单元: $\geq 0.2''$
- $R \geq 1000$

CPIC模块: 系外行星直接成像探测和多波段测光

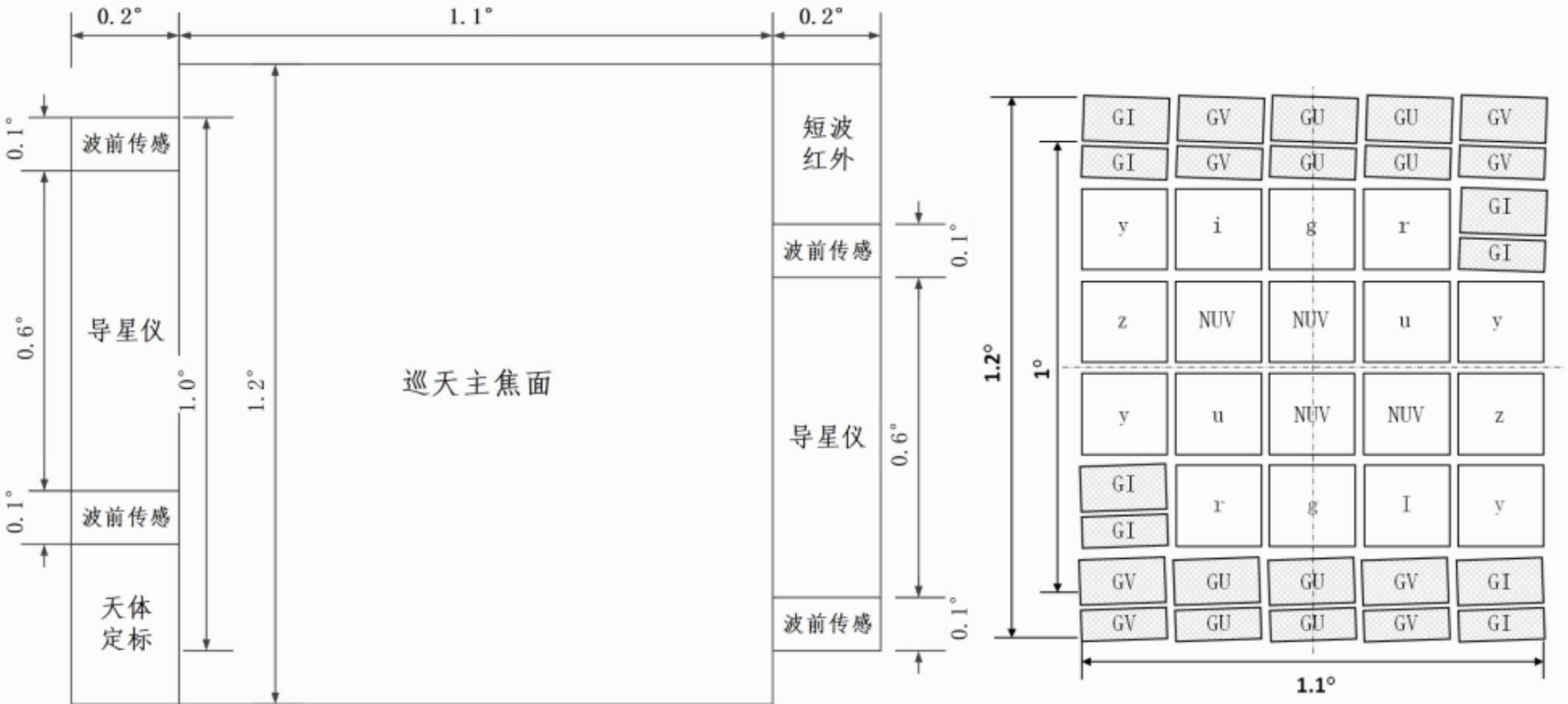
- 成像对比度: $\geq 10^{-8}$ (0.60-0.90 μm)
- 内工作角: $\geq 0.55''$ @ 0.633 μm

THz模块: 高速度分辨率和高灵敏度的谱线观测

- 频段范围: 0.41~0.51THz
- 频率分辨率: $\geq 100\text{kHz}$
- 观测灵敏度 $\geq 150\text{mK}$



巡天模块



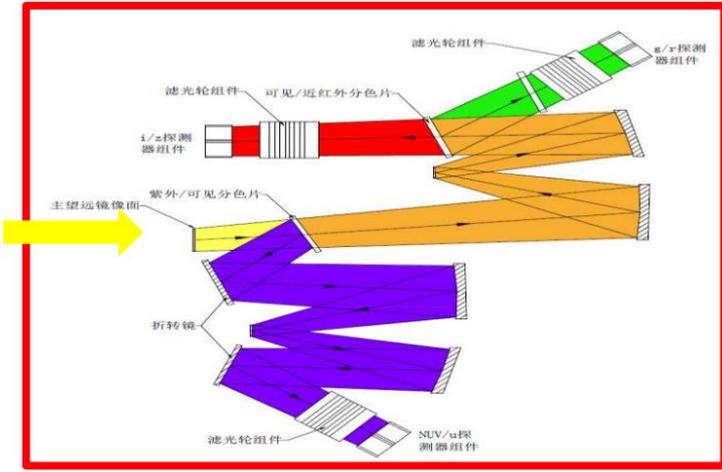


巡天模块

指标项目		指标要求	备注
主焦面	物理面积	$\geq 262435\text{mm}^2$	焦面尺寸应与主光学系统视场匹配，同时为波前传感器、导星仪留有充足可用的视场布局空间
	感光面积	$\geq 234000\text{mm}^2$	
	平面度	$\leq 60\ \mu\text{m (PV)}$	包含探测器感光面自身平面度和探测器拼接的平面度
观测波段	多色成像	覆盖 $0.255\sim 1.0\mu\text{m}$ ，至少6个波段	
	无缝光谱	覆盖 $0.255\sim 1.0\mu\text{m}$ ，至少3个波段	
光电性能	读出噪声	$\leq 5e^-/\text{pix}$	读出频率150kHz
	暗电流	$\leq 0.02e^-/\text{pix}/\text{s}$	像元尺寸按 $10\mu\text{m}$ 计；制冷需求满足模块功耗与散热条件约束
	读出时间	$\leq 40\text{s}$	
	满阱电荷数	最低值 $\geq 70\ \text{ke}^-$ ，平均值 $\geq 90\text{ke}^-$	像元尺寸按照 $10\mu\text{m}$ 计
图像性能	图像量化位数	16bit	
	图像动态范围	$\geq 82.9\text{dB}$	对应满阱电荷数最小值 70ke^- ，计算公式： $20\log(\text{满阱}/\text{读出噪声})$
模块效率	多色成像	近紫外波段 ≥ 0.26 ，u波段 ≥ 0.32 ，g波段 ≥ 0.58 ，r波段 ≥ 0.63 ，i波段 ≥ 0.62 ，z波段 ≥ 0.25	满足 $2\times 150\text{s}$ 曝光条件下的 5σ 星等：g波段 $\geq \text{AB } 26$ 等，NUV、u、r、i和z波段平均 $\geq \text{AB } 25.5$ 等
	无缝光谱	$255\sim 400\text{nm}$ 波段 ≥ 0.24 ， $400\sim 600\text{nm}$ 波段 ≥ 0.44 ， $600\sim 900\text{nm}$ 波段 ≥ 0.43 ， $900\sim 1000\text{nm}$ 波段 ≥ 0.18	满足 $4\times 150\text{s}$ 曝光条件下的等效宽波段 5σ 星等：GU波段 $\geq \text{AB } 22$ 等，GV波段 $\geq \text{AB } 23$ 等，GI波段 $\geq \text{AB } 23$ 等
无缝光谱	光谱分辨率 $R=\lambda/\Delta\lambda$	$R\geq 200$	各光栅组件视场范围内按不同区域及波长平均
成像通道像质	在主光学提供的像质条件下（80%能量集中度半径）	近紫外/u/g/r波段 $\leq 0.135''$ ，i波段 $\leq 0.145''$ ，z波段 $\leq 0.165''$	按照每个波段中心波长考察，包含模块内部原因导致的与主系统焦面失配；主光机静态像质：NUV、u、g和r波段 $\leq 0.13''$ ，i波段 $\leq 0.14''$ ，z波段 $\leq 0.16''$
无缝光谱通道像质		三个波段的+1或-1级光谱以及零级像REE80均值 $\geq 0.3''$ ，最大值 $\geq 0.4''$	像面范围内的平均值，按照每个波段的中心波长考察



MCI 和 IFS



多通道成像仪MCI:

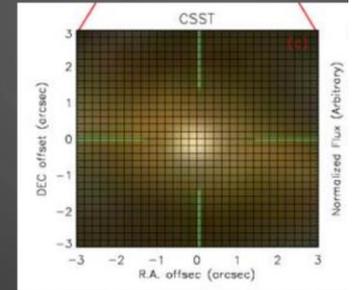
- 中继光学扩焦 (28m→41m) + 光路分色设计, 实现紫外(0.255-0.43)、光学1(0.43-0.7)、光学2(0.7-1.0微米)的**三通道**同时成像观测
- 同时搭载**30**宽带、窄带等滤光片
- 每通道一台9K CCD/sCMOS (0.05"/pix, FOV = 7.68' x 7.68')
- 发展媲美HST的深场观测能力
- 实现多通道同时高精度测光观测

基本能力:

- 空间分辨率0.2角秒
- 视场6"X6"
- 视场内填充率>90%
- 光谱覆盖范围0.35-1微米
- 光谱分辨率: $\Delta\lambda=3.5$ 埃

要注意的:

- 不适合看特别暗的目标源
 - 2米镜, 且同时在空间和光谱维度上切割
 - 单次曝光时长短: 300S
- 可能需要多次观测才能达到想要的灵敏度

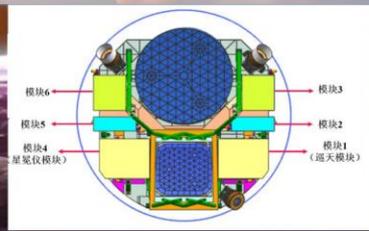




CPI C ❖ 项目总体任务目标

项目任务总体情况介绍

- ◆ 研制超高对比度成像星冕仪，用于空间站巡天光学设施
- ◆ 技术指标：对比度 10^{-8} （超过现有水平两个数量级）
- ◆ 科学目标：成像观测类太阳光谱恒星周围的成熟系外行星，对行星形成和演化理论提供重要观测证据，为后续搜寻存在生命的行星奠定重要基础。



参量	指标
射频频率范围	0.41~0.51 THz
瞬时中频带宽	≥ 2 GHz (两路1GHz)
频谱分辨率	优于100 kHz
系统灵敏度	< 150 mK $3\sigma/100\text{kHz带宽}/200\text{s}$
波束宽度	≤ 100 "
频谱动态	优于30 dB
输入信号亮温	5~350 K
在轨寿命	≥ 10 年
在轨工作时间	~ 4 年

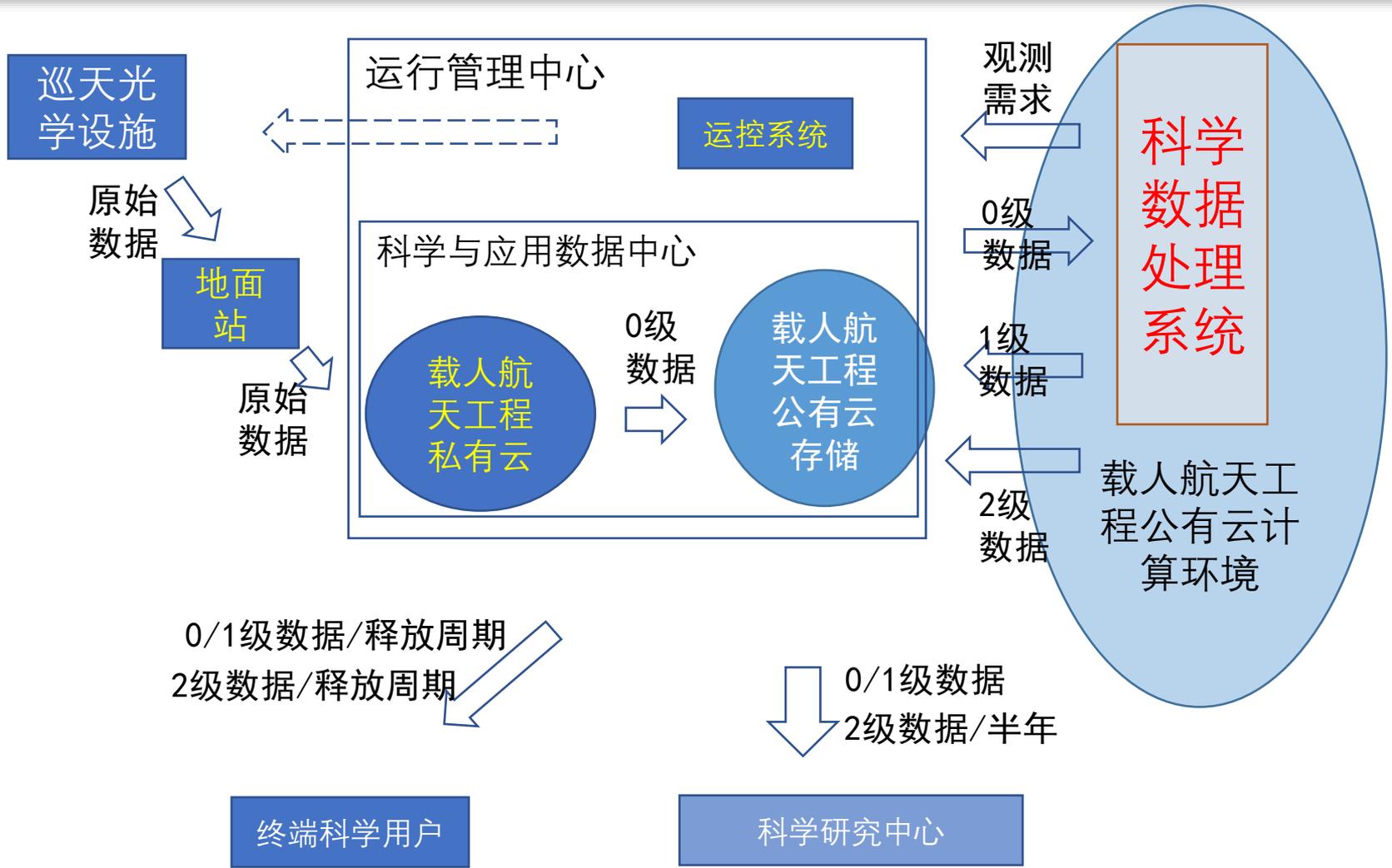


数据处理系统任务目标

- 在建造期间（发射之前，-2023.12）完成以下任务：
 - ▣ 研制科学数据处理系统软件
 - ▣ 实现CSST科学观测数据处理、为开展相关科学研究提供必需的观测基础数据
 - 五个后端模块、六类不同数据，涉及单色、多色成像，无缝光谱，积分视场光谱、星冕仪数据、太赫兹信号等
 - 生成1、2级数据产品，以供科学研究团队进一步分析并生成高级数据产品
 - ▣ 科学观测需求编排
 - 满足观测期间对望远镜每日观测天区的科学规划
- 在运行期间，数据系统完成数据处理并生成1/2级数据产品



研制方案——主要功能





研制方案——主要功能

□ 数据处理系统的主要功能

▣ 科学观测需求编排

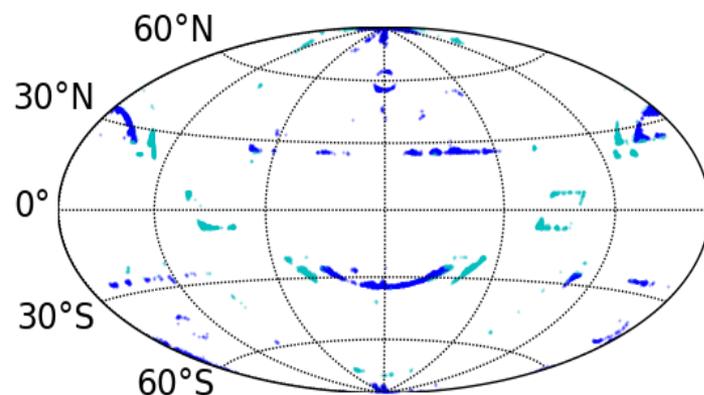
- 编制各个后端模块一段固定周期的科学观测需求，包括每次曝光使用的仪器、观测天区位置、观测时间、曝光时间、观测类型标识等信息
- 科学观测需求包括既定的巡天观测需求、各个后端模块的核心科学需求、一般申请观测需求和机遇观测需求；提出在在轨定标观测和科学测试观测需求

▣ 科学数据处理

- 输入0级数据，形成1级数据，再生成2级数据产品
- 满足最大多数科学用户的需求

▣ 观测数据仿真

- 实现观测图像的仿真，**为数据处理提供测试验证数据**





研制方案——主要功能

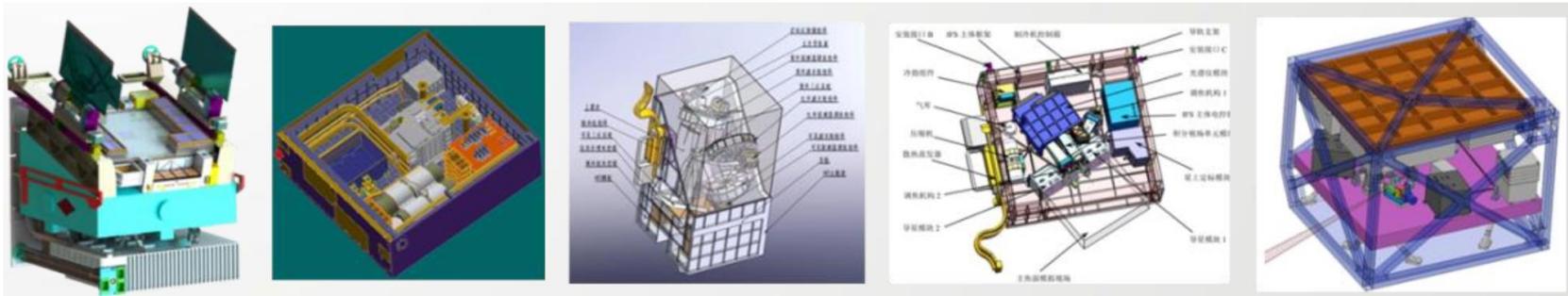
后端
模块



0级
数据
(输入)



1/2级
数据
(输出)



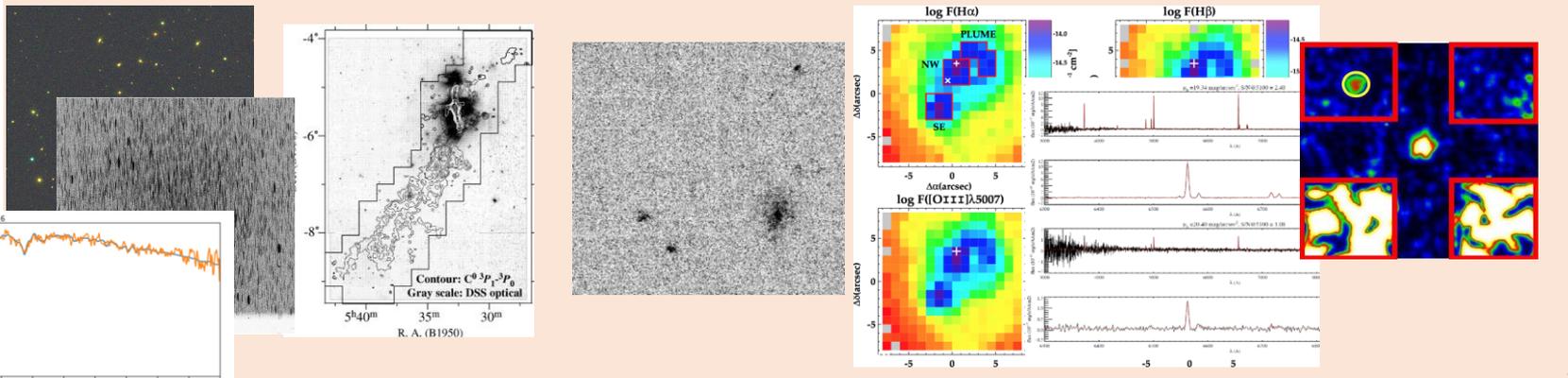
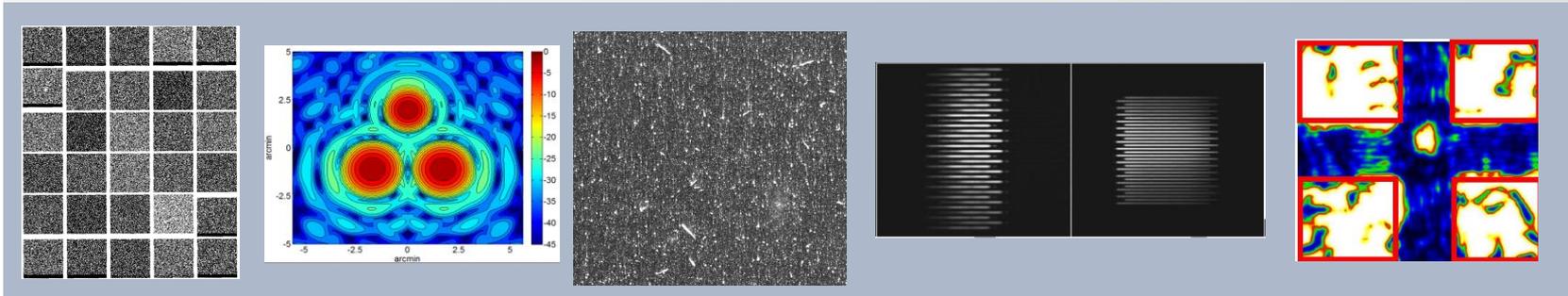
巡天模块

太赫兹模块

多通道成像仪

积分视场光谱仪

星冕仪模块





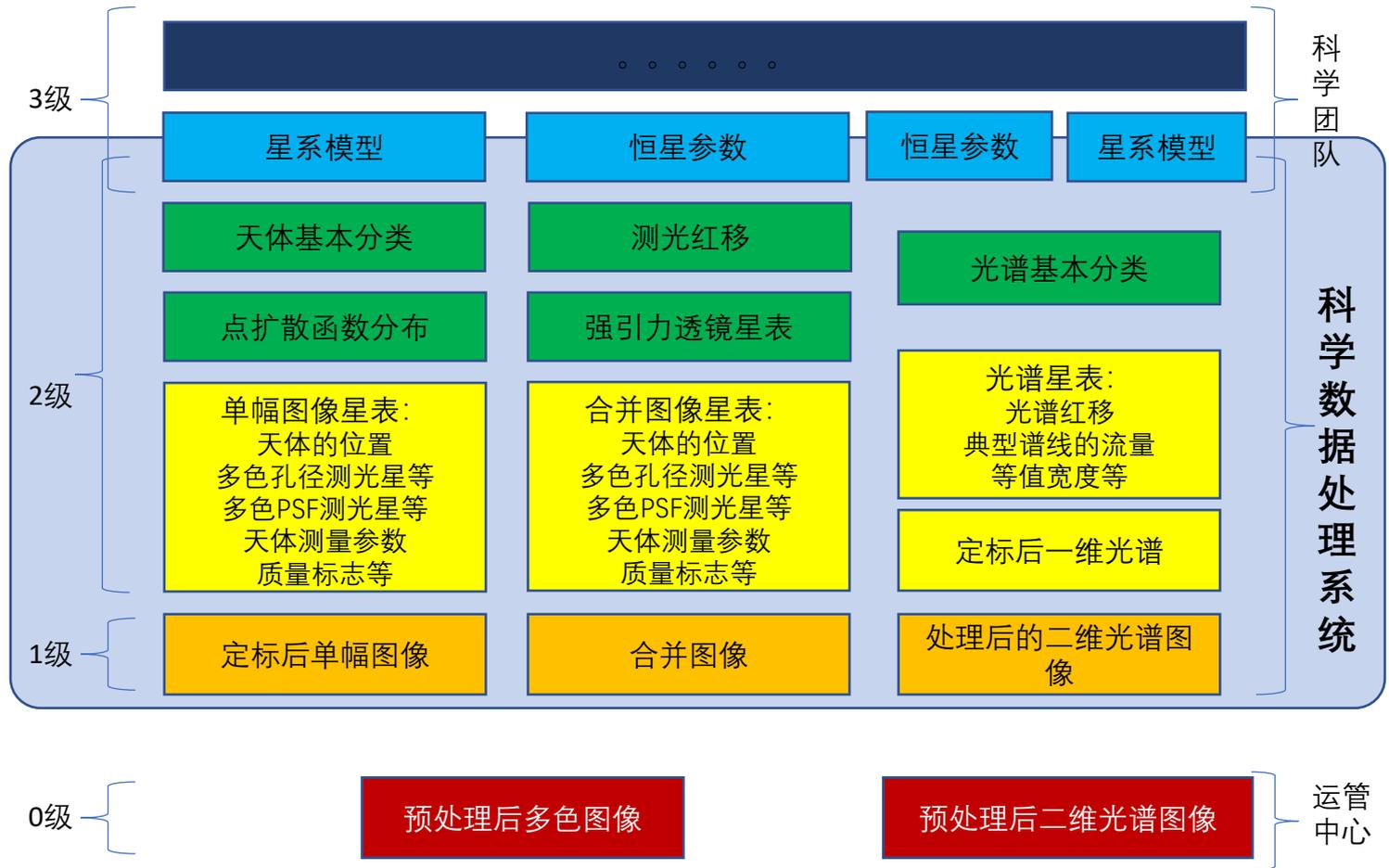
研制方案——数据产品

很多.....

共18类

共16类

共14类，另有
辅助数据9类



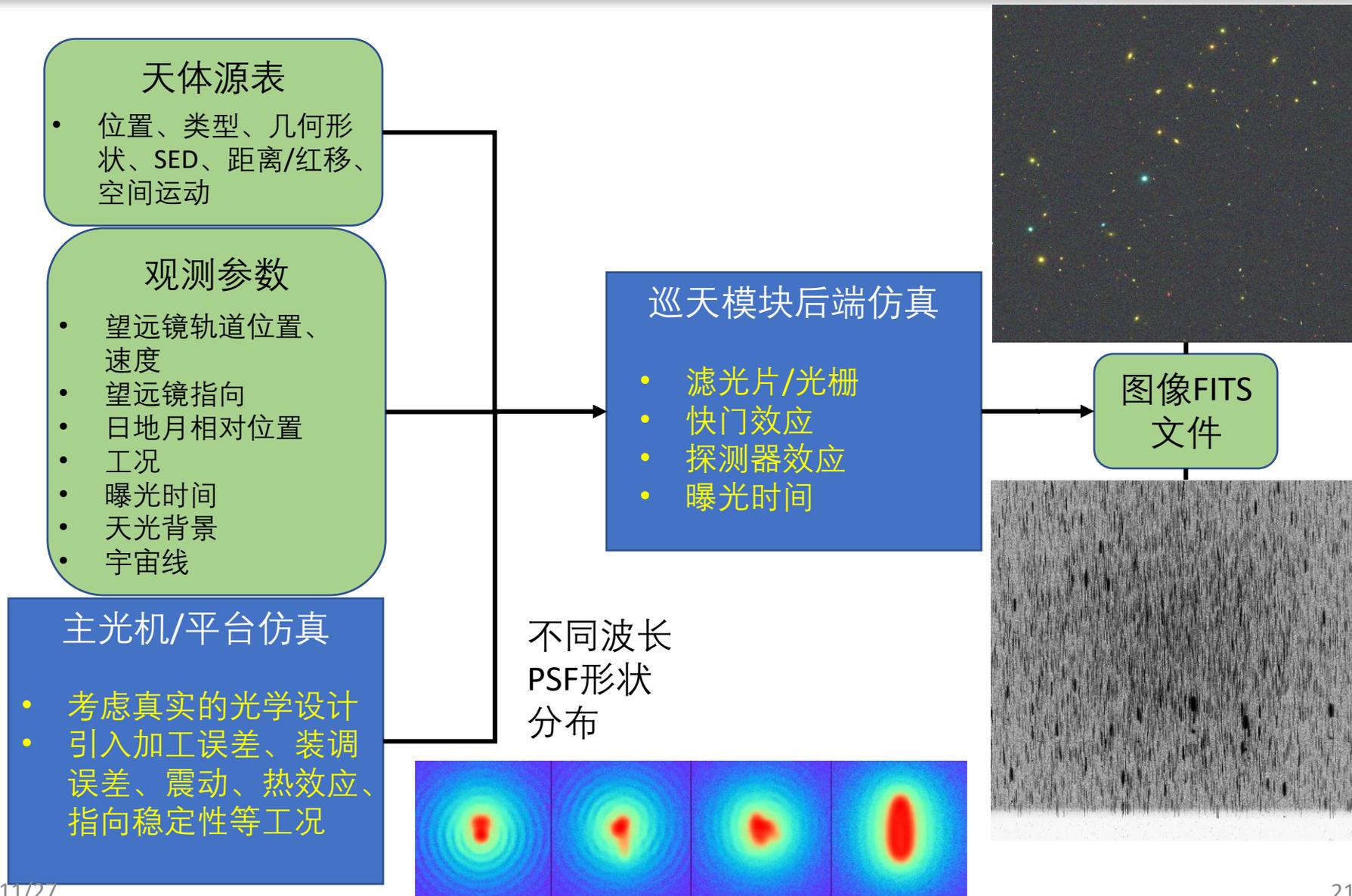
科学团队

科学数据处理系统

运管中心



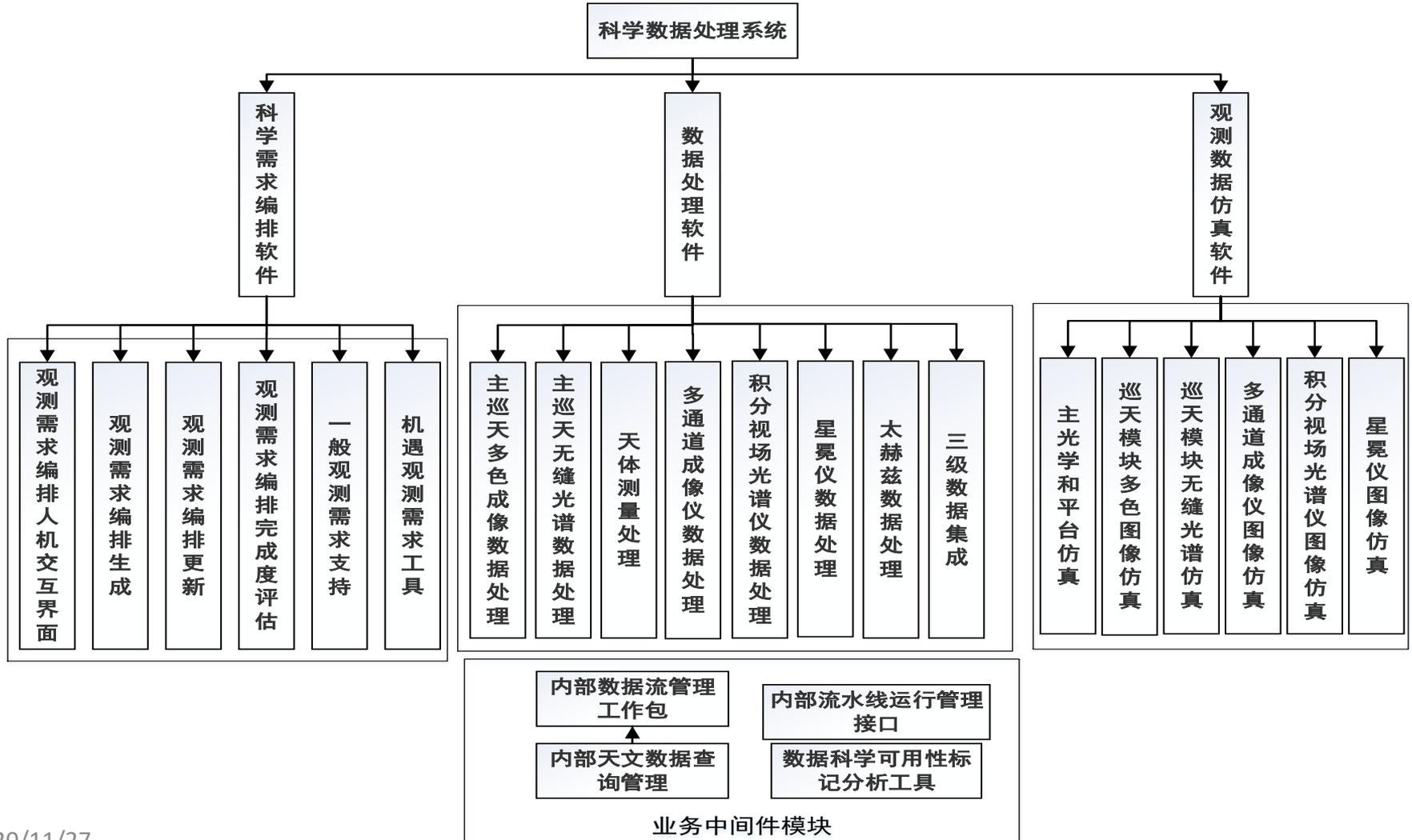
研制方案——主要功能





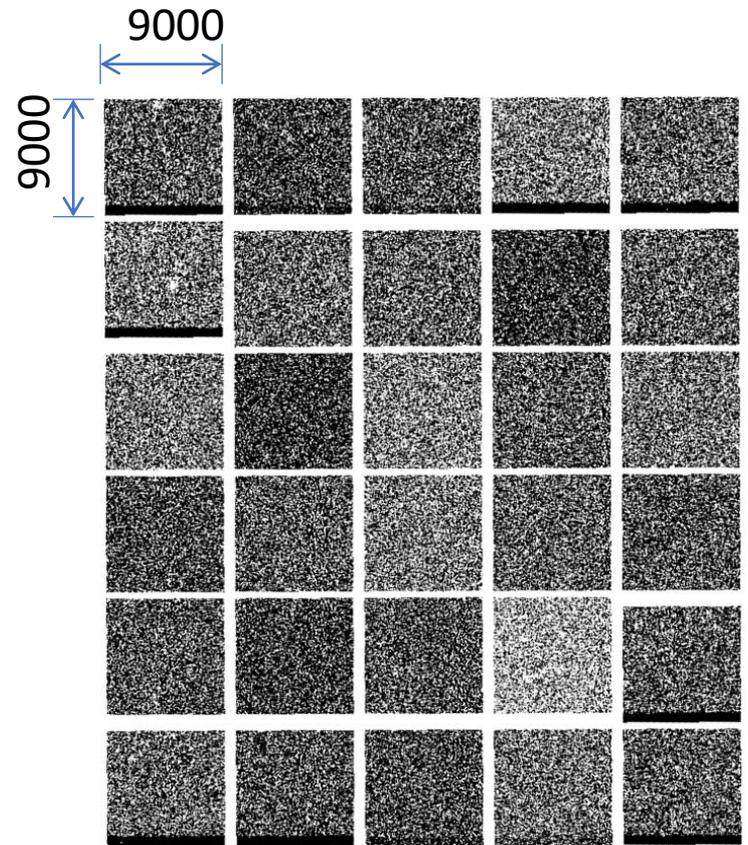
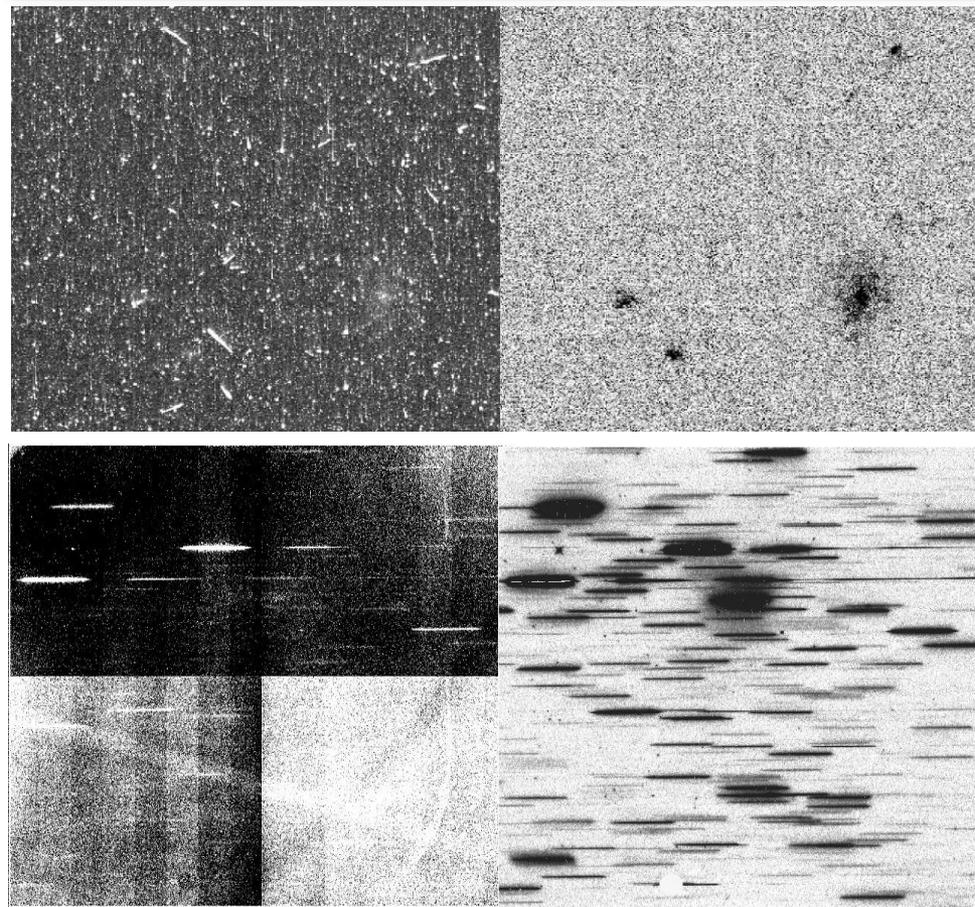
研制方案——系统组成

三个软件配置项（D类软件）





任务特点



X 60万幅图像

巡天模块1幅图像~25亿像素!



数据系统主要特征

□ 数据量大：

- 需要在自动化、并行的云环境下处理，总文件和数据库存储量超过**30PB**
- 数据库包含约数千亿条记录
 - 数千万原始观测图像（探测器级），~百亿星系，~百亿恒星，数千万探测到的信号源（可能是天体，也可能是噪声）
 - 采用空间球面坐标检索模式
- 数据处理加工难度高：数据种类多，算法复杂，没有现成软件可以沿用，10年数据量总算力需求**4.75PFLOPSx年**（仅使用CPU的情况下）
 - 部分算法可转移到GPU计算环境中
- 基本数据处理本身就是科学研究的重要内容，技术挑战性都前所未有



数据系统特征

□ 数据计算特点

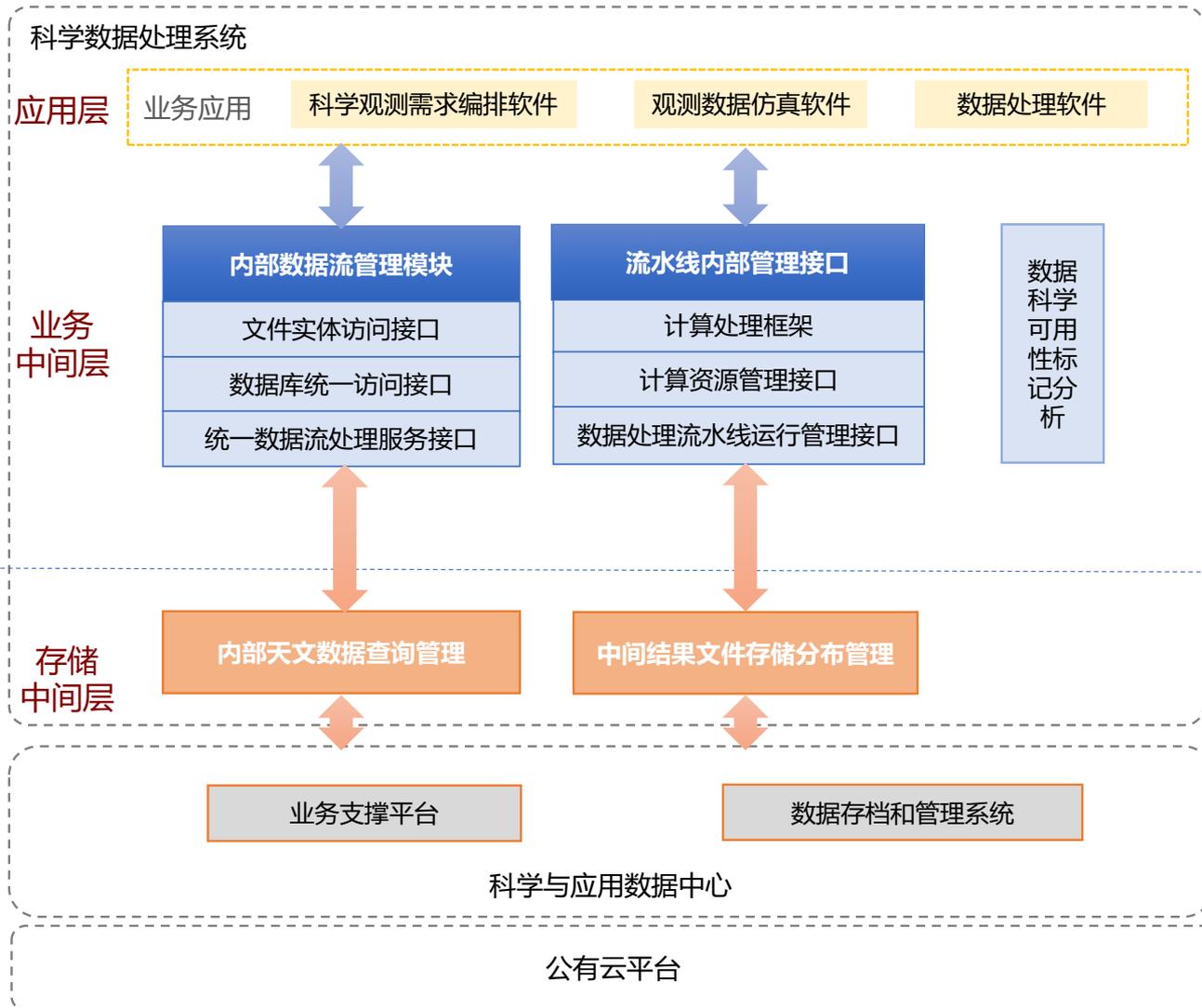
- 每次拍摄产生30x150MB图像文件
- 每幅图像文件初期处理几乎是相互独立的
- 同一天区不同时间、不同颜色观测的结果需要融合，但此时天区与天区之间是独立处理的
- 对天体的光学流量进行校准（定标）时，需要相邻天区之间的数据比较，但不需要全天天区的关联操作
- 对天体位置的确定也是在局部天区进行的
- 数据处理具有局部独立的特点，便于开展逻辑简单的多任务并行运算

□ 数据库使用特点

- 空间球面坐标查询（类似地理信息系统的查询模式）
- 每次图像处理都需要访问数据库，检索图像文件位置，保存处理结果



软件架构





详细设计方案——数据处理软件

□ 科学数据日常处理流程

- ▣ 科学数据处理系统从内部数据库中读取最新观测数据记录，启动日常处理流程，将0级数据经过逐级处理，完成仪器效应修正、宇宙线和天光背景去除等操作；
- ▣ 提取图像中天体的位置、流量、光谱、形态等重要物理信息；
- ▣ 不同后端仪器根据其仪器特征、数据特性需要不同的处理流程。
- ▣ 数据处理的结果为科学研究所使用的1级科学数据。

□ 数据产品生成

- ▣ 在每次数据产品释放之前，应对已处理的1级数据进行再处理，合并图像、合并星表、估算更多天体参数，如红移、星系参数、恒星参数等，形成2级数据。
- ▣ 将0—2级数据重新组织形成可发布的数据库和相应文件结构。

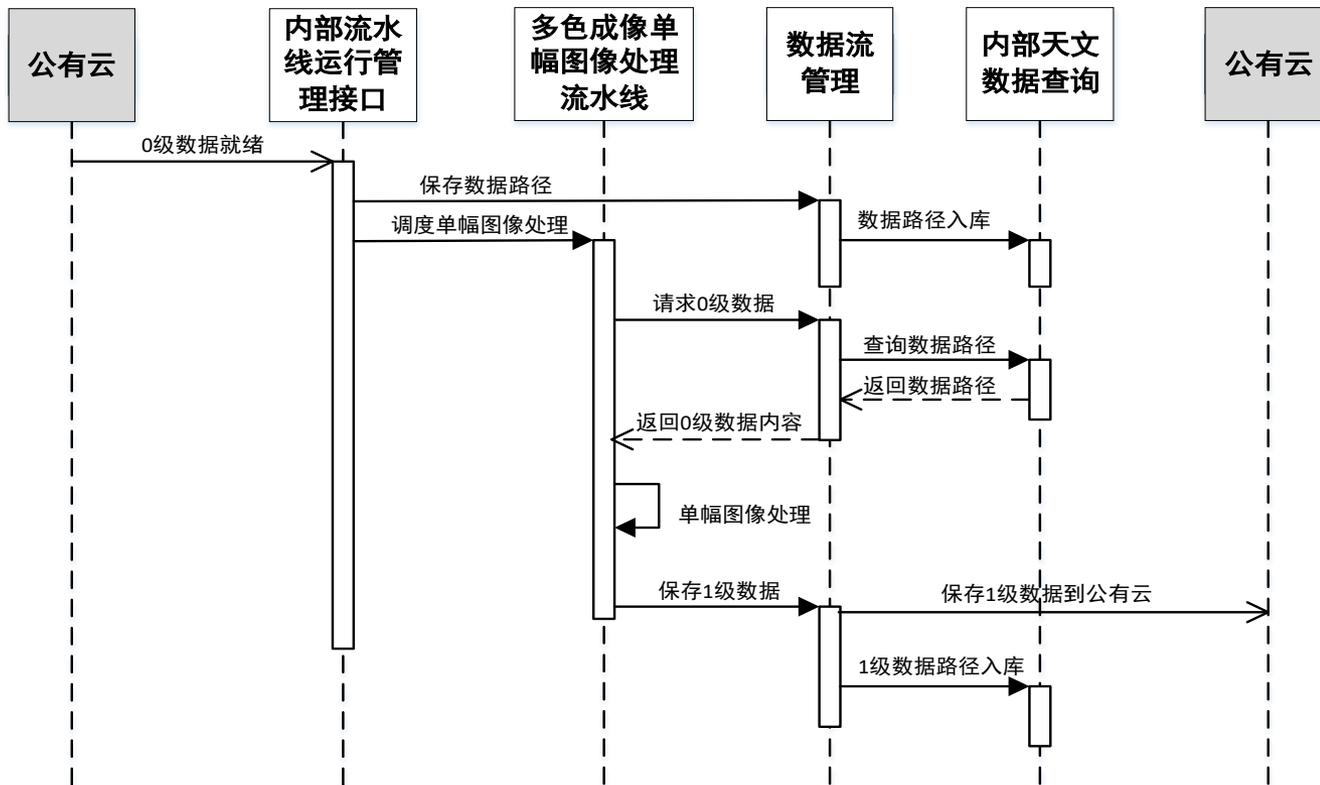
□ 历史数据再处理流程

- ▣ 当科学数据处理系统软件进行了升级和更新、相关定标数据得到改进之后，所有之前的历史数据（0级数据）需要重新进行处理并形成新版本的1级数据。



详细设计方案——数据处理软件

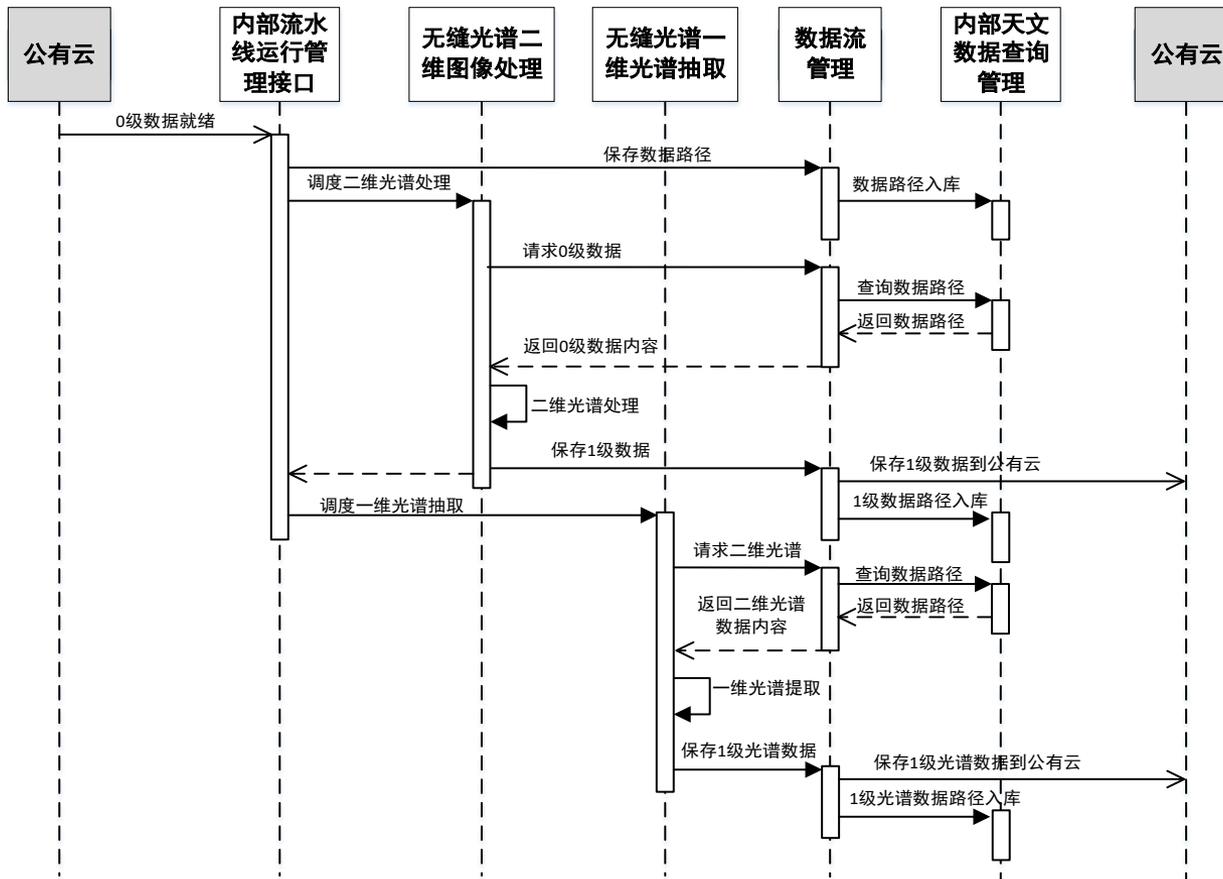
- 实现主要功能的流水线
- 多色成像日常处理流水线：





详细设计方案——数据处理软件

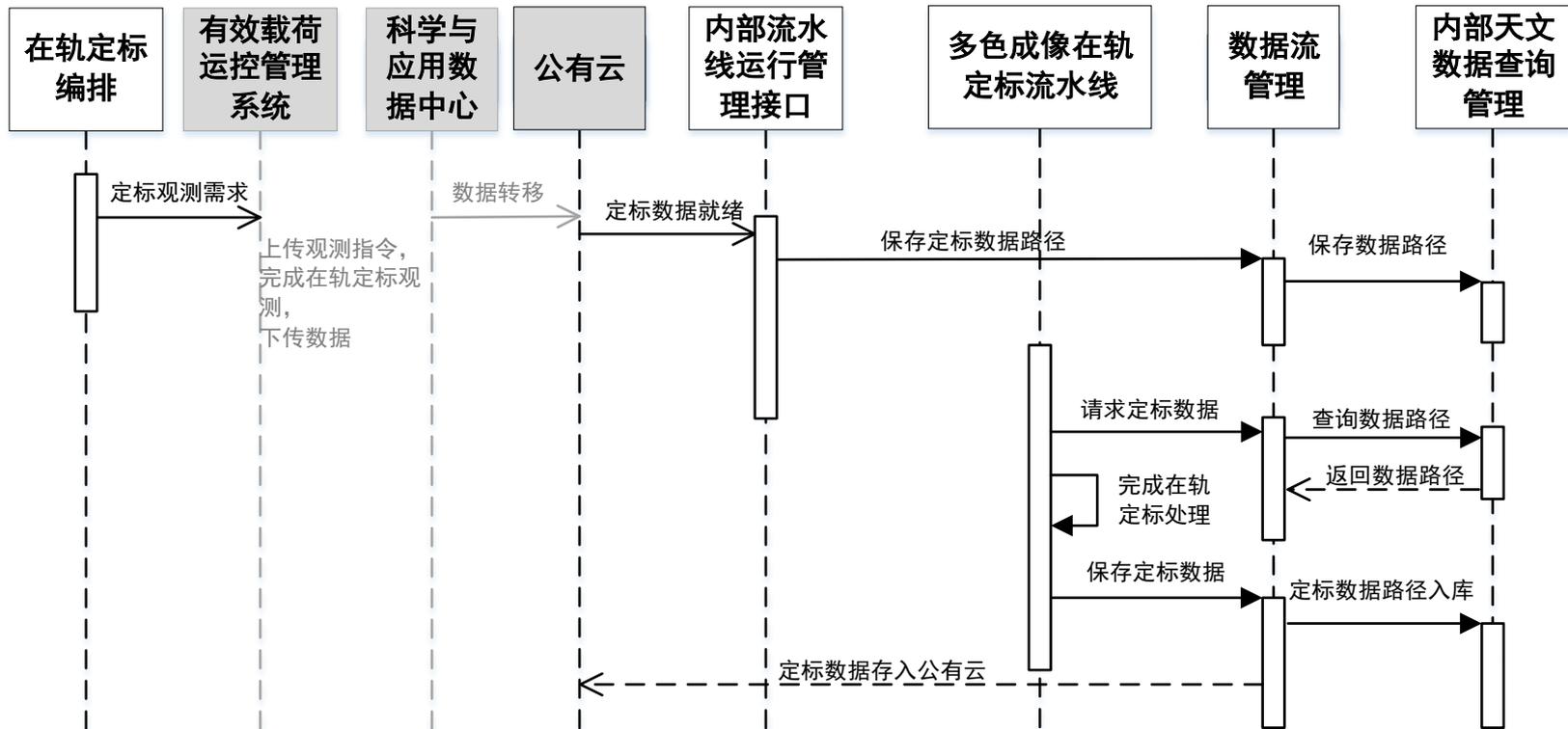
无缝光谱数据日常处理流水线：





详细设计方案——数据处理软件

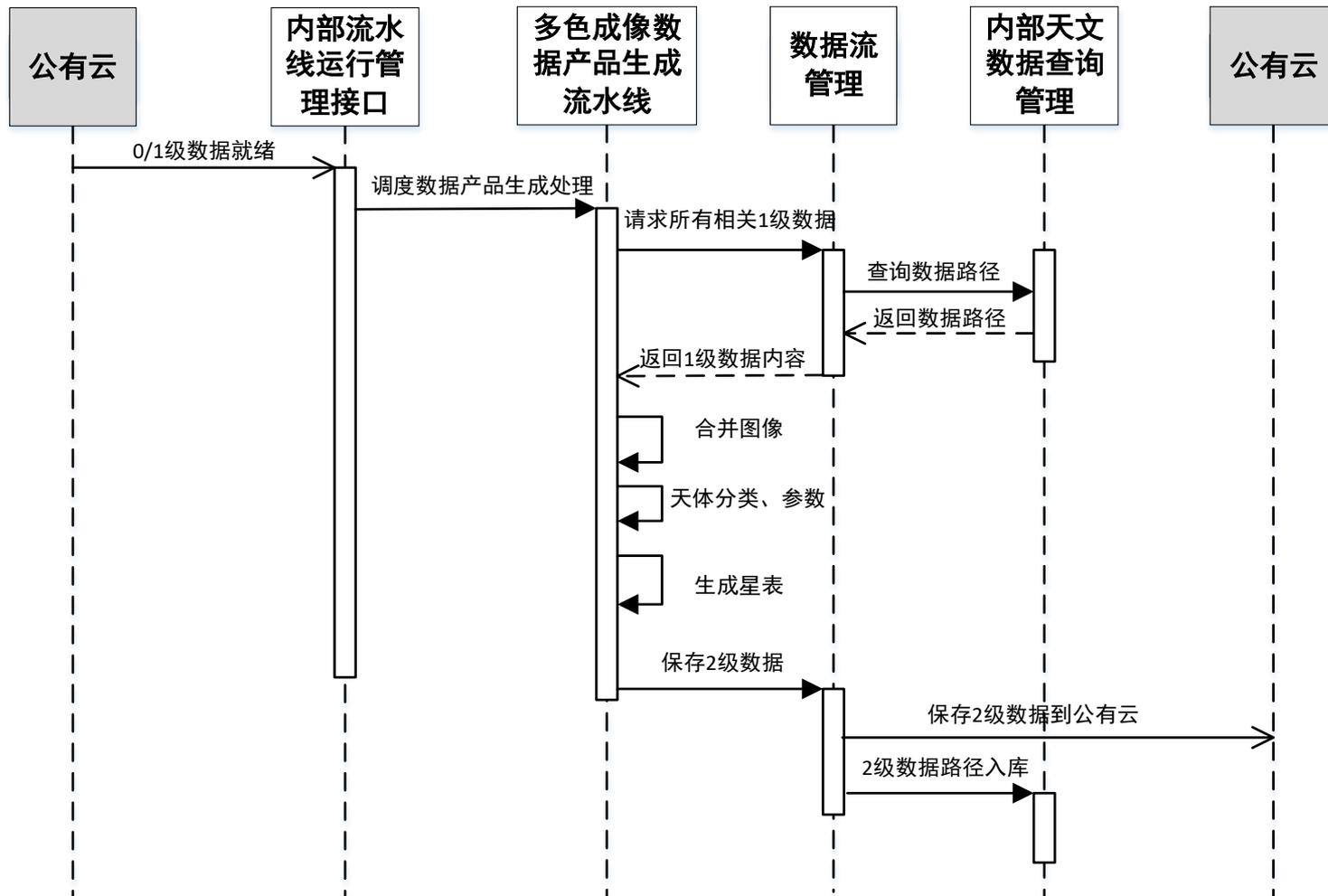
□ 多色图像定标处理流水线：





详细设计方案——数据处理软件

多色图像数据产品生成流水线：



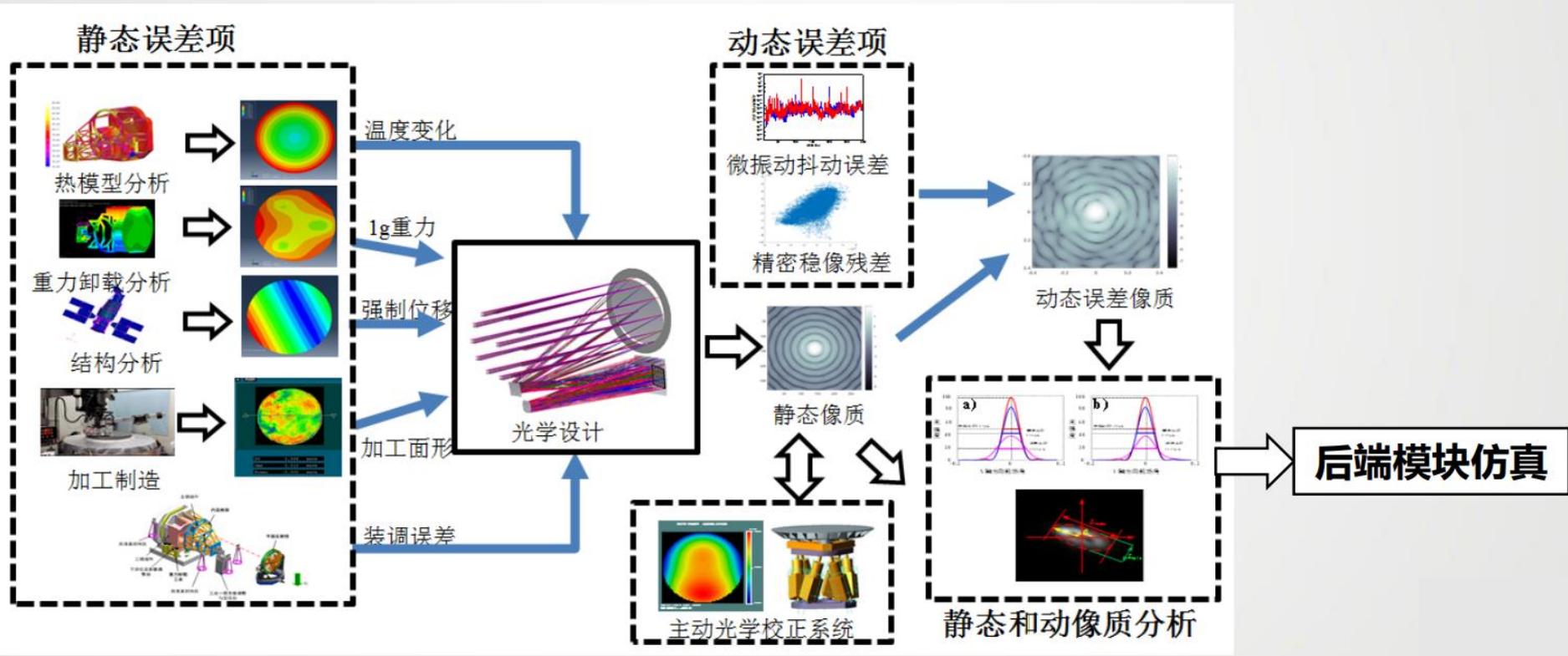


详细设计方案——观测数据仿真软件

天体源表

位置 米型 几何形

光学处理软件



调误差、震动、热效应、指向稳定性等工况

PSF形状分布



组织管理和任务分工

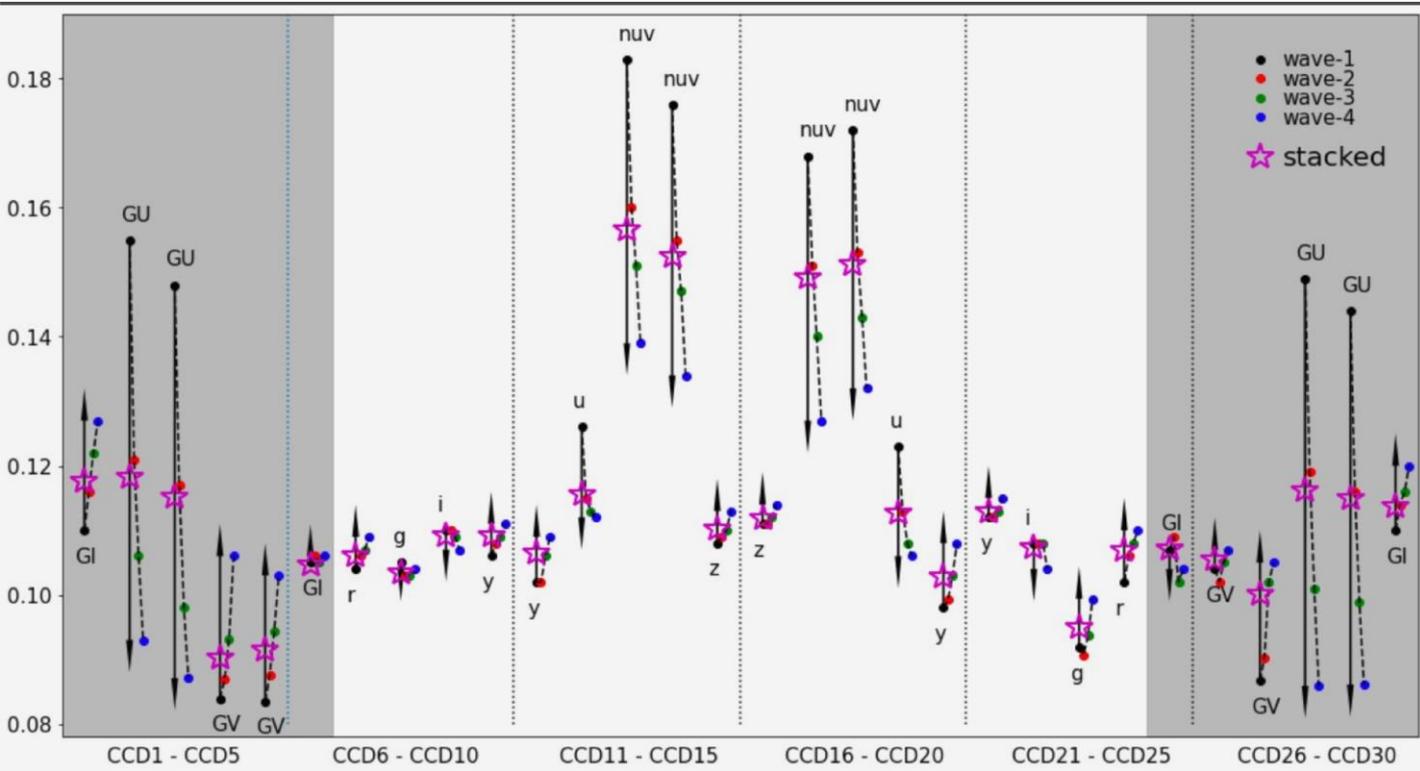
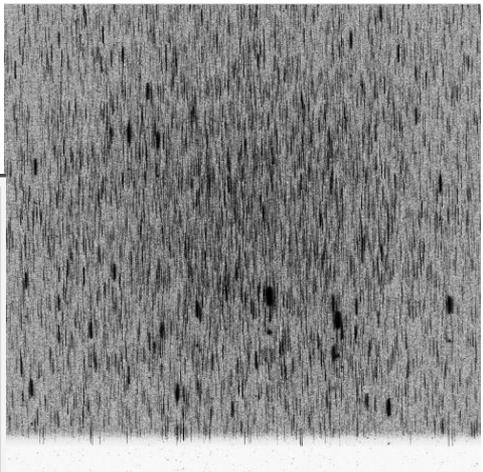
软件/模块	负责人姓名	负责人单位	其他参与单位	
观测需求编排软件	崔辰州	国家天文台	上海天文台	
数据处理软件	多色成像处理	张天萌	北京师范大学、天津大学、广州大学等	
	无缝光谱处理	刘凤山	北京大学、北京师范大学等、广州大学等	
	天体测量处理	齐朝祥	国家天文台、南京大学、暨南大学等	
	MCI数据处理	陕欢源	上海天文台	
	IFS数据处理	沈世银	上海天文台	
	CPIC数据处理	赵刚	南京天光所	
	THz数据处理	敖宜平	紫金山天文台	
	内部数据流管理	王锋	广州大学	国家天文台、湖南文理学院等
	内部流水线运行管理	吴开超	网络信息中心	国家天文台等
观测数据仿真软件	李国亮	紫金山天文台	国家天文台、长春光机所、云南大学等	



当前进展情况

科学数据仿真

- C1完成了7平方度主巡天仿真 (ver0.0)，包含宇宙线、均匀背景、仪器效率、理想化PSF (PSF卷积、光线追迹)
- C2修改了技术路线，以PSF分布连接主光机和后端仿真，基本满足需求
- C2加入像场畸变等效应
- PSF随波长的剧烈变化对无缝光谱的影响需要研究 (?)





未来计划

□ Cycle3计划（—2021.6）

- 需求分析文档（需要科学输入）
- 系统分析和设计
 - 细化各流水线的具体过程，数据模型和各个接口
- 开展在轨定标研究（同硬件、科学团队协作）
- 配合科学研究
 - 明确和细化与科学研究的接口
 - 提供满足科学团队需要的仿真数据
 - 需要统筹安排并与数据系统研制进度进行协调和同步
- 全面开始研制建设
 - 强化工程化规范
 - 细化工程管理制度
 - 继续广招人才



敬请指正!

谢谢!