中国科学院大学
University of Chinese Academy of Sciences

海量天文数据在线可视化研究

答 辩 人：张磊

导　　师：崔辰州 研究员

答辩日期：2019年5月30日

答辩提纲

1.研究背景简介

2.系统技术路线

3.瓦片数据生成

4.China-VO SkyView

5.研究总结与成果

研究背景简介

- 研究背景
- 国内外研究现状
- 研究意义
- 研究内容

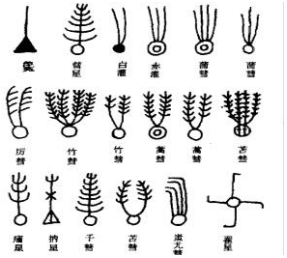
背景简介 - - 1.1 研究背景

·天文学是一门以观测为基础的学科

位于山西的5000年前陶寺天文台遗址

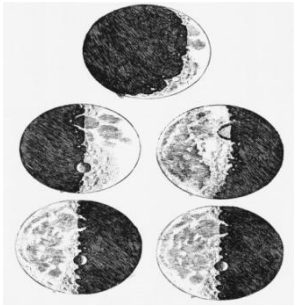
伽利略将望远镜指向星空

胶卷、天文CCD相机

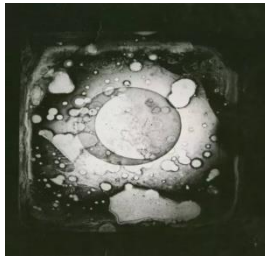


西汉马王堆
帛书彗星图

宋代苏州
石刻天文图



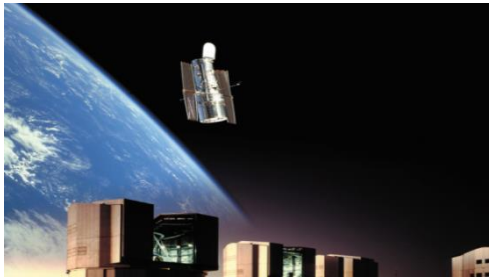
1610年伽利略手绘月球表面



1839年人类第一张
照相天文图片

作者拍摄的
昴星团

·现代天文学从不同角度获取天体数据进行更加全面的研究



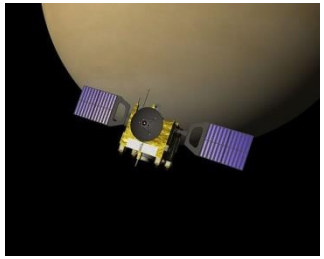
光学/红外望远镜

天马射电望远镜

激光干涉引力波天文台（LIGO）

多信使天文学

·对大量太阳系天体进行了遥感探测或登陆研究工作



金星快车探测器

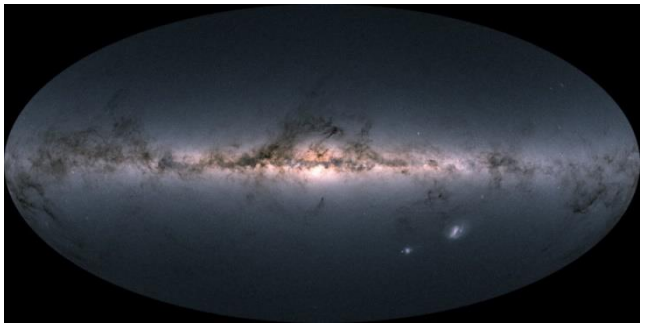
阿波罗11号探测器

好奇号号火星探测器

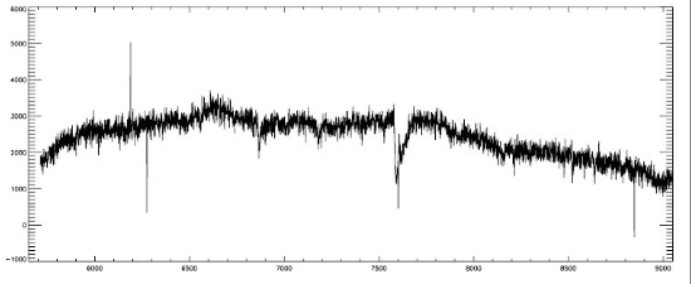
伽利略号木星探测器

·天文学家获取了图像、星表、光谱等各种类型的数据

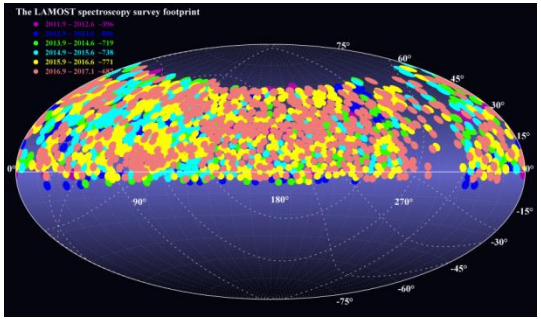
GAIA DR2发布的数据



GAIA第一张全天彩色图像



LAMOST望远镜获取的首条天体光谱



LAMOST望远镜巡天观测足迹

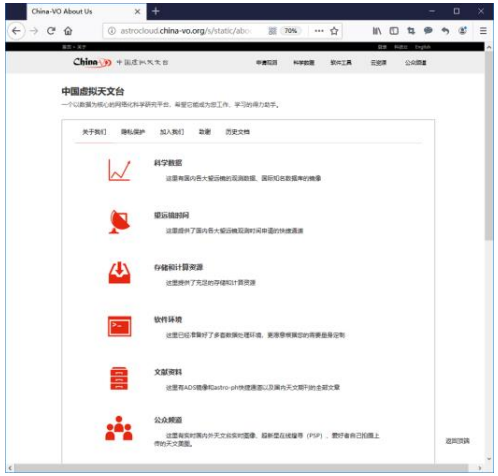
中国虚拟天文台（China-VO）

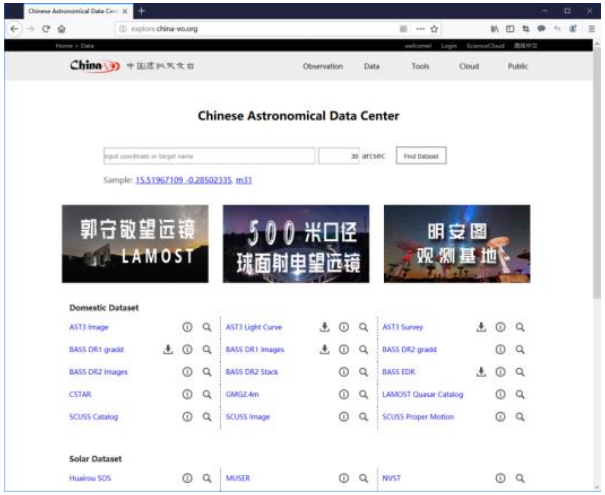
·虚拟天文台借助先进的计算机和网络技术，力图实现分布在世界各地的各种大型天文数据的互联、融合和发现

· China-VO致力于研究和开发符合国内实际情况和需求的，符合国际虚拟天文台联盟规范的虚拟天文台技术和服务

·重点研发领域：

China-VO系统平台建设
国内外天文数据的统一访问
现有天文工具的虚拟天文台集成
天文设备的虚拟天文台集成
虚拟天文台基础上的科普教育





中国天文数据中心

·China-VO维护和运行，整合国内大部分数据资源

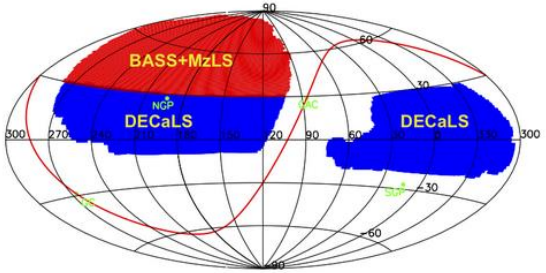
·现有数据量超过4PB，2030年超过50PB。

BASS巡天数据

·以中国为首开展的大视场多波段深度巡天项目

·大量精美的图像

·海量的星表数据



BASS巡天天区范围

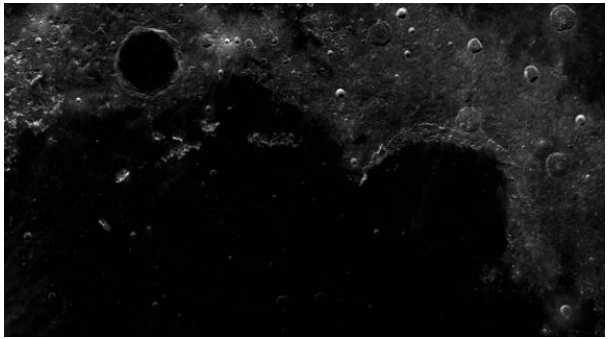
BASS DR2 图像

月球数据

·我国嫦娥二号获取

·全月高分辨率影像数据

嫦娥二号卫星

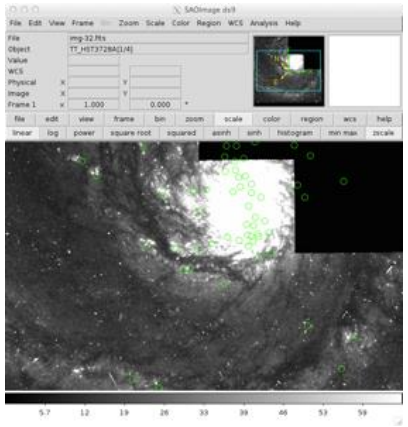


嫦娥二号月球数据

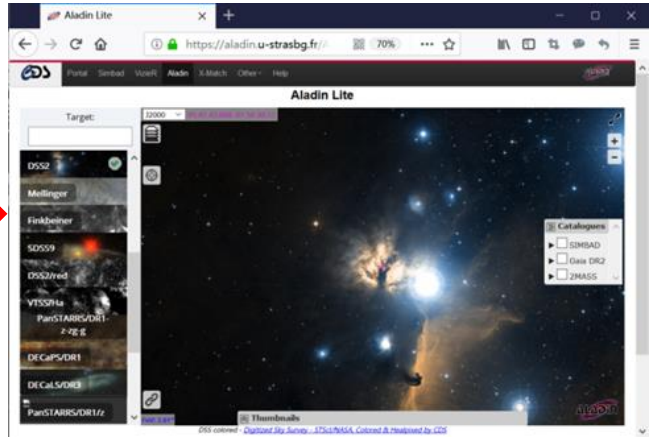
背景简介 - - 1.2 国内外研究现状

·分层可视化、在线可视化

全部加载
极度耗时

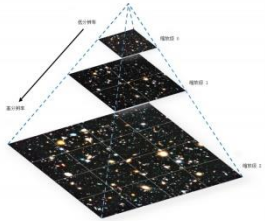


传统的数据传输方式（ds9）

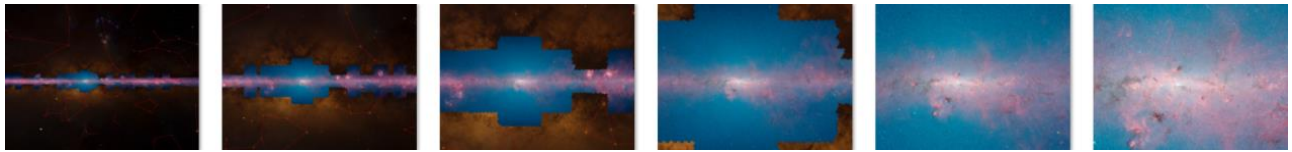


瓦片数据传输方式（Aladin Lite）

按需加载
节约时间



分层瓦片天图金字塔



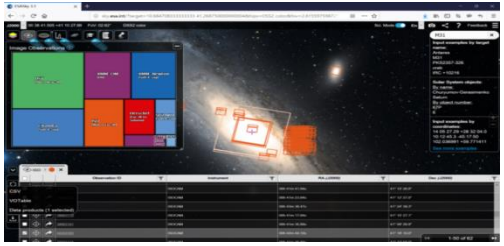
逐层显示目标天体的更多细节

背景简介 - - 1.2 国内外研究现状

科学研究领域：

·在数据获取方面，系统作为各天文数据中心档案的可视化门户，为天文学家获取数据提供了一种方便的方式。

·在传播方面，系统已被用于学术出版物的视频摘要，帮助科研人员快速可视化推广其科研成果。



可视化的方式获取数据（以ESASky为例）

奇异小天体-2018RR2视频摘要截图

科普教育领域：

广泛用于几乎所有级别的天文教育与科普活动，如正式的学校课堂以及非正式的科普场馆等，产生了引人注目的参与式教育、科普体验。

背景简介 - - 1.3 研究意义

(1)首次尝试将国内获取的天文数据以分层可视化方式进行展示，对国内其它天文数据的可视化提供参考借鉴。

(2)将天文数据快速高效地分享给使用者，有利于数据的开放共享。

(3)促进可视化技术在中国天文数据的科研和科普的应用。

(4)本研究的最终成果是首个基于我国多个巡天数据集的渐进分层多分辨率在线可视化系统——China-VO SkyView，将提升China-VO的数据服务能力。

背景简介 - - 1.4 研究内容

系统实现
技术选择
性能对比
HiPS技术

系统技术路线

瓦片数据处理

BASS巡天数据
嫦娥二号全月7米DOM数据

总体设计
开发技术
数据展示
信息叠加层
后续开发

China-VO SkyView

系统技术路线

- 系统实现
- 技术选择
- 性能对比
- HiPS技术

系统技术路线 - - 2.1 系统实现

China-VO SkyView在线可视化系统实现

实现China-VO SkyView涉及几个主要问题：

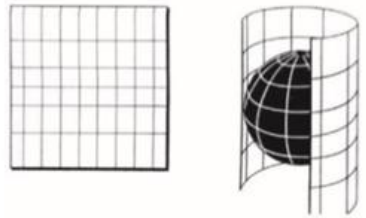
·分层可视化技术 ·瓦片生成工具 ·在线展示与后期开发

本工作计划使用现有工具进行，以缩短系统开发时间

系统技术路线 - - 2.2 技术选择

Leaflet

·基于正方形逐层划分
·Google Sky、WikiSky使用
·天文科研领域使用较少

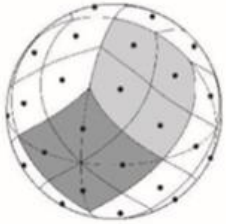


TOAST

·基于三角形逐层划分（HTM）
·主要在WWT软件中使用
·成熟的瓦片生成工具
·成熟的在线展示系统

HiPS

·基于菱形逐层划分（HEALPix）
·IVOA标准，20多个数据中心使用
·550多个HiPS数据集
·成熟的瓦片生成工具
·简单的在线展示工具
·强大的二次开发API



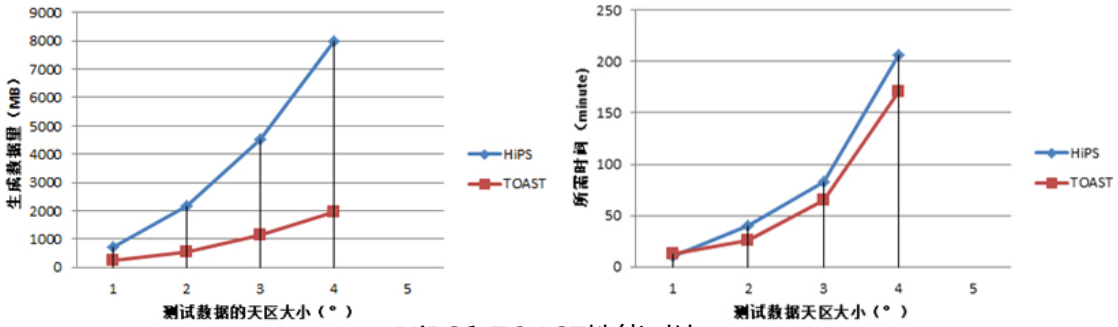
系统技术路线 - - 2.3 性能对比

·测试环境:CentOS，2GB内存，120GB硬盘大小。

·测试数据：2MASSJ波段1、2、3、4平方度原始图像（分别为45、144、312、583张图像）。

·HiPS瓦片切分工具：Hipsgen.jar
TOAST瓦片切分工具：Montage

·结果显示HiPS的生成数据量远大于TOAST，这是由于在同一切分层级下HiPS生成瓦片数量较多。
在耗时方面，可以通过调整线程和RAM来缩短所需时间。



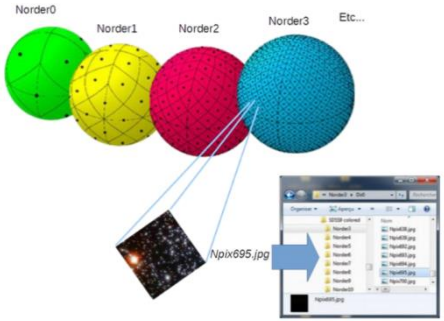
HiPS和TOAST性能对比

系统技术路线 - - 2.4 HiPS

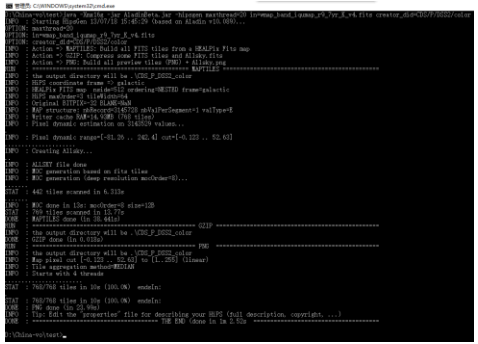
·HiPS本质上是将各种空间分辨率的巡天数据映射到存储在基本文件系统内的HiPS瓦片集合。通过HEALPix层级及其单元索引可以完全标识HiPS瓦片。

·HiPS采用文件夹的方法管理瓦片数据，将其存储在对应层的文件夹中。

·HiPS文件夹的前缀为“Norder”，代表层（Order），文件夹内为瓦片，瓦片索引的前缀为“Npix”。



基于目录和文件的HiPS体系结构



HiPS瓦片生成工具

```
<!-- include Aladin Lite CSS file in the head section of your page -->
<link rel="stylesheet" href="//aladin.u-strasbg.fr/AladinLite/api/v2/latest/aladin.min.css" /
>

<!-- you can skip the following line if your page already integrates the jQuery library -->
<script type="text/javascript" src="//code.jquery.com/jquery-1.12.1.min.js" charset="utf-8">
</script>

<!-- insert this snippet where you want Aladin Lite viewer to appear and after the loading of
 jQuery -->
<div id="aladin-lite-div" style="width:400px;height:400px;"></div>
<script type="text/javascript" src="//aladin.u-strasbg.fr/AladinLite/api/v2/latest/aladin.mi
n.js" charset="utf-8"></script>
<script type="text/javascript">
    var aladin = A.aladin('#aladin-lite-div', {survey: "P/DSS2/color", fov:60});
</script>
```

HiPS在线展示工具代码

瓦片数据生成

- BASS巡天数据
- 嫦娥二号全月7米DOM数据

瓦片数据生成 - - 3.1 BASS巡天数据

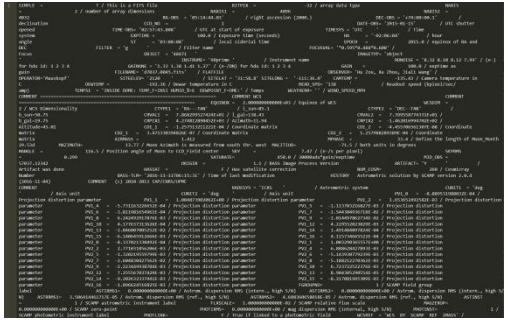
数据介绍

·BASS DR1图像和星表数据

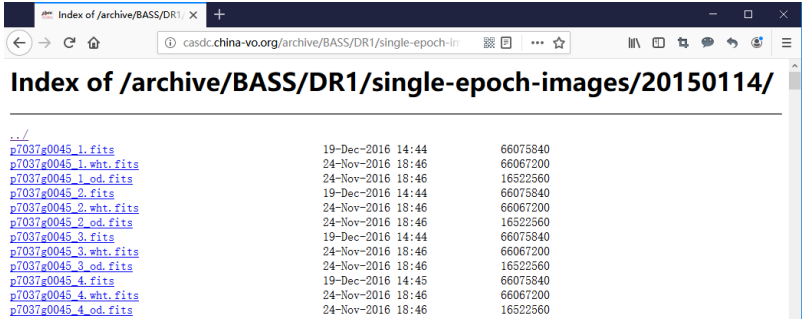
·BASS DR2图像和星表数据

·数据格式：FITS，天文界常用的数据格式，专门为在不同平台之间交换数据而设计

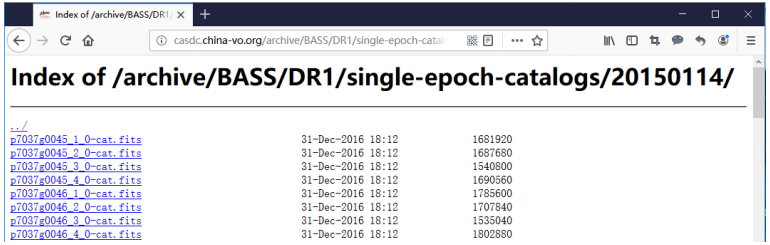
·FITS由文件头和数据组成，在文件头中存储有对该文件的描述，如观测时间、观测对象、拍摄温度、曝光时间等信息，便于后期数据分析使用



FITS文件头信息



China-VO数据档案中的BASS DR1图像数据



China-VO数据档案中的BASS DR1星表数据

瓦片数据生成 - - 3.1 BASS巡天数据

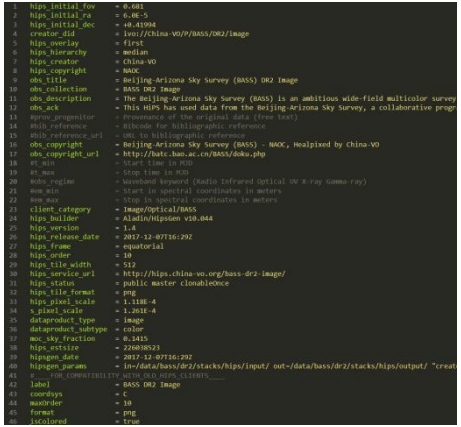
瓦片生成与在线展示

·图像HiPS瓦片生成:Hipsgen.jar

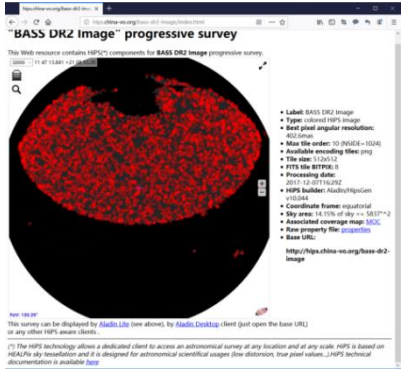
·星表HiPS瓦片生成:Hipsgen-cat.jar

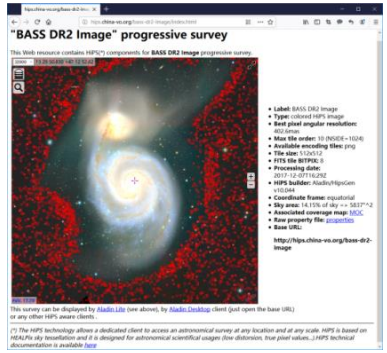
·在线展示：Alidin Lite

HiPS瓦片文件结构



包含生成HiPS相关信息的属性文件





在线展示BASS DR2图像和星表数据

瓦片数据生成 - - 3.2 月球数据

数据介绍

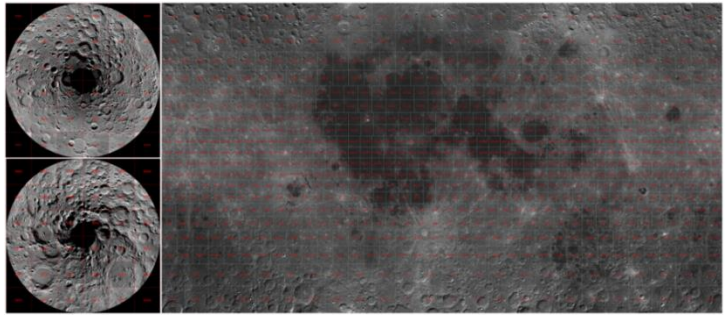
·嫦娥二号全月7米分辨率数字正射影像（DOM）数据，同时具有地图几何精度和影像特征。

·共有844个分幅，每个分幅包括一个数据文件（tif）、一个同名坐标信息文件（tfw）和一个同名投影文件（prj），数据总量为747.9GB。

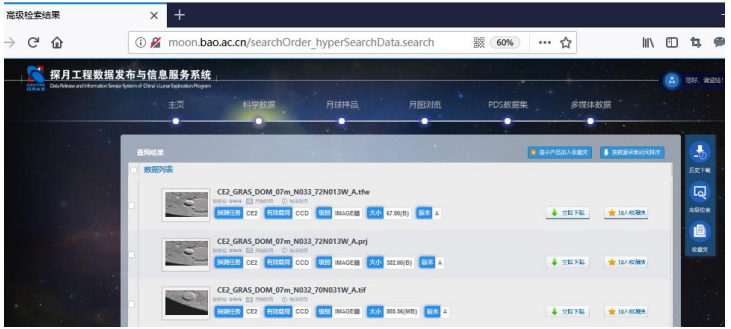
·北极（70°N~90°N）和南极（70°S~70°S）为极区方位投影，70°S~70°N之间为正轴等角墨卡托投影。

字段	含义
CE2	表示探测任务，即嫦娥二号任务。
GRAS	表示数据制作单位，即地面应用系统。
xxx	表示产品数据类型，即 DEM 或 DOM。
aaa	表示分辨率，分别为 07m，20m 和 50m。
bbbb	表示地图图幅号，第一位为字符，后三位为数字，如 C001。
yyMzzzF	表示分幅数据的中心地理位置，yy 为 2 位纬度值，M 为北纬（N）或南纬（S），zzz 为 3 位经度值，F 为东经（E）或西经（W）。
v	表示产品版本，取值 A~Z，即第一版为‘A’，最后一版位‘Z’。
ext	为产品格式后缀，为 tif（数据文件）和 tfw（坐标信息文件）。

嫦娥二号分幅数据产品命名规范
CE2_GRAS_xxx_aaa_bbbb_yyMzzzF_v.ext



嫦娥二号全月7米分辨率DOM分幅格网图

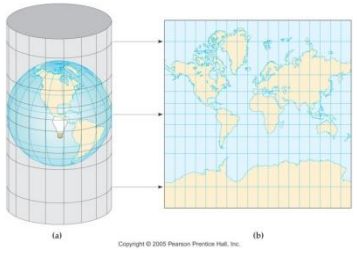


嫦娥二号全月7米分辨率DOM数据组（tif、tfw、prj）

瓦片数据生成 - - 3.2 月球数据

瓦片生成

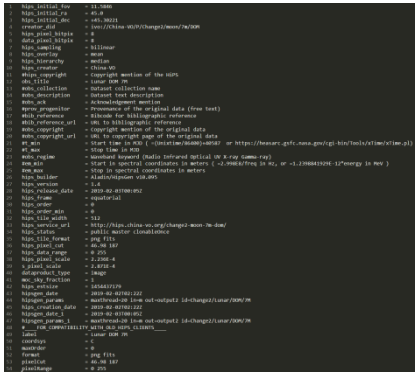
相关投影和数据格式：
·墨卡托投影：按等角条件，将经纬网投影到圆柱面上，将圆柱面展为平面。
·TIFF：标签图像文件格式，具有文件头、文件目录和文件数据的3级体系结构。
·GeoTIFF：TIFF格式的一种扩展，是基于TIFF的地理空间栅格数据存储格式，并且已成为TIFF规范的一部分。



墨卡托投影示意图

相关处理工具：
·QGIS：开源的地理信息系统软件，可以将TIFF格式的数据转换为GeoTIFF。
·GDAL：开源的栅格空间数据转换库，提供对GeoTIFF等地理数据格式的支持。

嫦娥二号全月7米分辨率DOM数据分层HiPS瓦片生成



HiPS属性文件

China-VO SkyView

- 总体设计
- 开发技术
- 数据展示
- 信息叠加层
- 后续开发

China-VO SkyView - - 4.1 总体设计

·系统定位：将作为中国虚拟天文台可视化门户，方便使用者访问我国及镜像的国外天文数据

·访问地址：http://skyview.china-vo.org

·主要功能：将China-VO的数据资源进行统一在线可视化展示，为使用者提供数据服务

China-VO SkyView在线可视化系统所处的位置

China-VO SkyView - - 4.2 开发技术

HiPS

一种天文科研领域使用广泛的分层可视化技术，有较成熟的瓦片生成与在线展示工具。

Nginx

一款应用广泛的web服务器软件，具有跨平台、用户群体广泛、安全性高等优点。

Web技术

HTML5、CSS3、JavaScript、jQuery、Alidin Lite API、Sublime、SecureCRT等。

开发环境

Windows10操作系统，Inter(R) Core(TM) i5-4300U CPU @1.90GHz，4G内存。

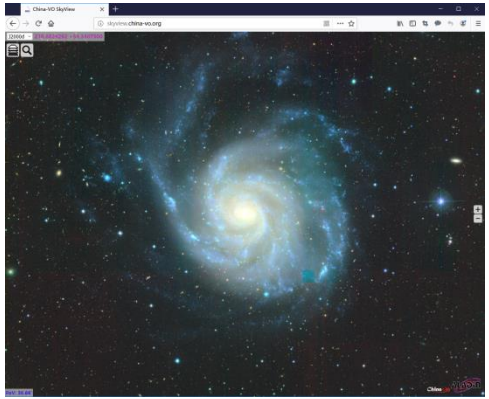
China-VO SkyView部分代码

China-VO SkyView - - 4.3 数据展示

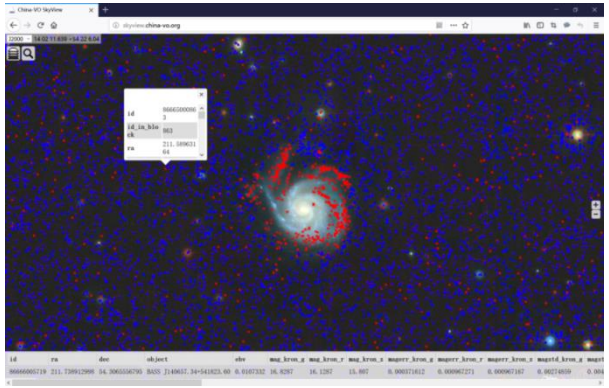
1）国内巡天数据和行星数据

·BASS 图像数据
·BASS 星表数据
·嫦娥二号全月7米分辨率DOM

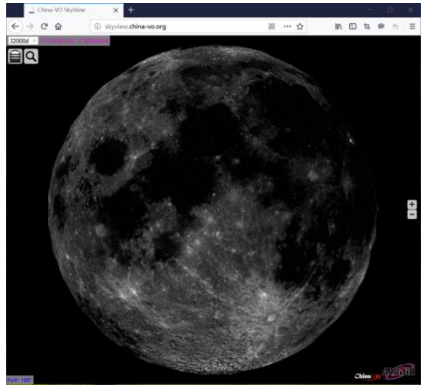
China-VO的国内HiPS巡天数据



BASS DR2图像数据



BASS DR1和DR2星表数据



嫦娥二号全月7m分辨率DOM

China-VO SkyView - - 4.3 数据展示

2）镜像国外数据

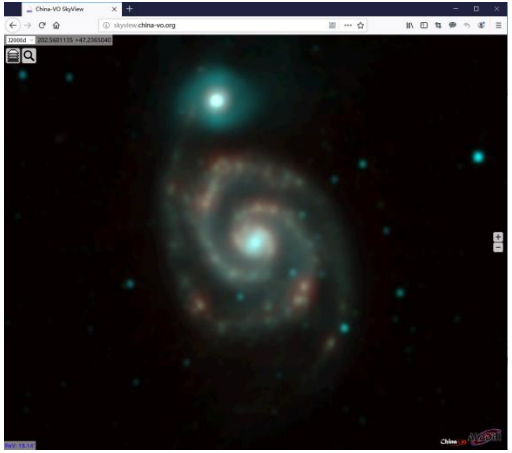
·SHS巡天
·GALEX-GR6-Color巡天
·AllWISE-color巡天等

Mirrors

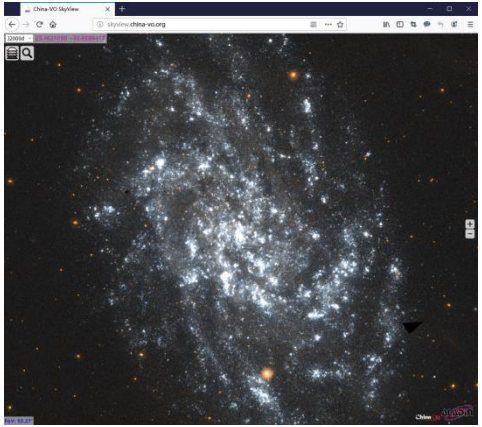
#	Title	Tile	Last modif	URL
1	2MASS color J (1.23um), H (1.66um), K (2.16um)	jpeg	2016-04-22	http://hips.china-vo.org /m/CDS_P_2MASS_color/
2	DSS colored	jpeg	2015-05-11	http://hips.china-vo.org/m/CDS_P_DSS2_color/
3	DSS2 Red (F+R)	jpeg,fits	2014-10-24	http://hips.china-vo.org/m/DSS2Merged/
4	SDSS9 color	jpeg	2014-10-24	http://hips.china-vo.org /m/CDS_P_SDSS9_color/
5	AllWISE color Red (W4) , Green (W2) , Blue (W1) from raw Atlas Images	jpeg	2016-07-01	http://hips.china-vo.org /m/CDS_P_allWISE_color/
6	GALEX GR6 AIS Color composition	png,jpeg	2016-12-01	http://hips.china-vo.org /m/CDS_P_GALEXGR6_AIS_color/
7	SHS - digitized H-alpha photographic survey	png,jpeg,fits	2014-04-22	http://hips.china-vo.org/m/SHS/
8	PLANCK R2 LFI color composition 30-44-70 GHz	jpeg	2015-02-06	http://hips.china-vo.org /m/CDS_P_PLANCK_R2_LFI_color/
9	Herschel PACS (color composition)	jpeg	2015-10-16	http://hips.china-vo.org /m/ESAVO_P_HERSCHEL_PACS-color/

Update: 2019-02-03

China-VO镜像的国外HiPS巡天数据



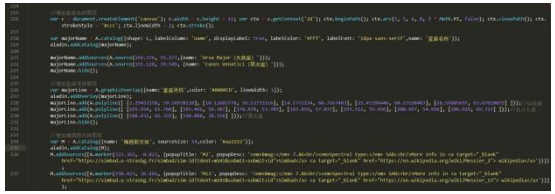
AllWISE-color巡天数据



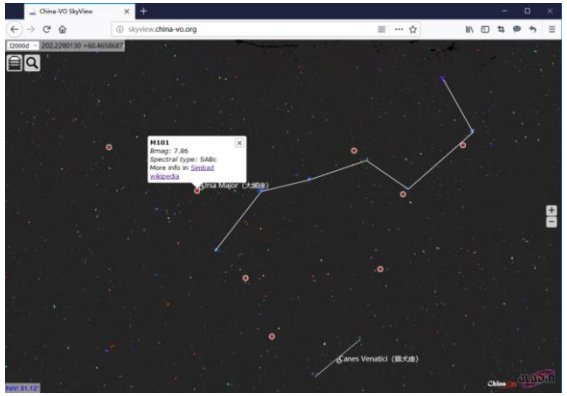
GALEX-GR6-Color巡天数据

China-VO SkyView - - 4.4 信息叠加层

使用Aladin Lite API编写相关代码添加叠加图层



部分叠加层代码示例

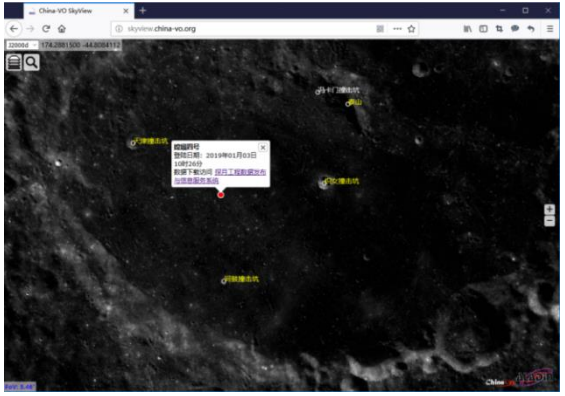


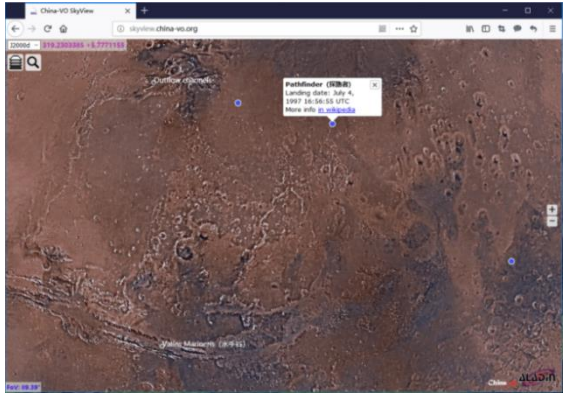
巡天图像

·星座名称
·星座连线
·梅西耶天体

月球图像

·主要地形特征标记
·月球上的中国名
·嫦娥和阿波罗探测器着陆点位置





火星图像

·主要地形特征标记
·探测器着陆点位置

China-VO SkyView - - 4.5 后续开发

(1) 本系统和国外成熟的系统相比有较大差距，主要体现为数据集较少和功能应用单一。

(2) 在数据集方面，随着国内各类型天文项目的设计和推进，越来越多的国内第一手天文数据能够在系统中予以集成和展示。同时随着国外地面和空间项目数据的公开，会有更多国外镜像数据源可以接入系统。

(3) 在功能方面，将增加光谱检索、与其他外部服务建立联系、论文可视化、时域探索、任务计划工具等功能。

(4) 本系统将成为数据资源众多、功能强大的China-VO可视化门户，为使用者提供专业的天文数据服务。

研究总结与成果

研究总结与成果 - - 5.1 研究总结

(1) 对主流的分层可视化技术、瓦片生成与在线展示工具、在线可视化系统进行了深入研究，确定了我国在线可视化系统的技术路线。

(2) 基于HiPS分层可视化技术及Aladin Lite，将我国的BASS巡天数据和嫦娥二号月球数据进行了分层可视化展示。

(3) 基于Web技术，实现了China-VO SkyView在线可视化系统，将China-VO现有的14个国内和镜像国外HiPS数据集进行了在线展示。此外还增加了星座连线、月球上的中国名、探测器着陆位置等信息叠加层。

(4) 通过该系统可以有效地将China-VO的天文数据分享给相关科研工作者和普通公众，实现科研数据的开放共享，促进科普教育工作的开展。

研究总结与成果 - - 5.2 研究成果

1. 张磊, 樊东卫, 崔辰州, et al. 海量巡天数据在线可视化研究[J]. 天文学进展.（已接收）

2. 杜康玉、李珊珊、张磊. 嫦娥奔月. 中国科学院十大优秀科普微视频，2017.

THANKS

中国科学院大学
University of Chinese Academy of Sciences