

天文学的新革命 -

虚拟天文台

崔辰州

北京天文台 (100012)

cb@yac.bao.ac.cn

2001年3月25日至27日是中国国家天文台兴隆观测站分配给张艳霞博士使用我国口径最大的2米光学望远镜进行观测的时间。为了这次观测，张博士已经准备了很长时间。早在上一年她就向兴隆观测站提出了观测申请。由于国内外的天文学家都可以申请使用这架2米望远镜，观测站要从众多的申请中筛选出计划合理、目标明确且有较好科学价值的方案，然后才分配给这些申请者一般不超过3个夜晚的观测时间。为能申请成功，在向观测站提出观测申请之前，她必须花不少时间根据观测季节、观测站的地理位置选取观测目标，安排观测顺序；在观测时还要完成各种辅助工作，如望远镜的调试、校准等。观测后把观测数据刻成光盘或保存到磁带上带回天文台，最后才能进行数据分析处理工作。为了这次观测，她至少得准备几个月的时间。但最倒霉的事情终于发生了：由于那几天华北地区沙尘暴强烈，望远镜无法观测，几个月的准备不得不付诸东流。

申请、准备、观测、数据处理，这沿袭了上百年的观测方式将随着虚拟天文台的诞生而发生彻底改变。

虚拟天文台 (Virtual Observatory) 的基本工作原理

巡天，就是对整个天区进行观测、普查。如果利用可视化工具和望远镜巡天得到的数据把所观测天体再现出来，就是一个数字虚拟天空。如果综合利用 射线巡天、X射线巡天、紫外巡天、光学巡天、红外巡天和射电巡天所得到的观测数据，便可以构成一个全波段的数字虚拟天空。通过可视化软件按用户要求显示这些巡天望远镜所观测到的任何部分的天空，这就仿佛一架虚拟的大望远镜。如果再开发出功能强大的计算工具、检索工具、统计分析工具，这就是一个独一无二的虚拟天文台。到时天文学家要取得数据便只需要打开电脑、登录网站、输入几个参数、点几下儿鼠标即可。

虚拟天文台概念的产生

二十世纪九十年代，也就是刚刚过去的这十年，天文望远镜的设计工艺和制造技术、大接收面积探测器（如2048×2048像素CCD）的制造工艺有了革命性的发展；计算机的计算能力成倍提高，互联网络的通信能力日益增强、覆盖范围不断扩大。科学技术的突飞猛进为古老天文学的又一次革命奠定了坚实的基础。

望远镜设计与制造能力的提高已使得越来越多的大型空间观测设备投入使用，为天文学开启了光学波段以外的如 射线、X射线、远红外等观测窗口。技术的进步同样造就了新一代大口径地面光学、近红外望远镜以及单天线或多天线厘米波、毫米波射电观测设备。这些新近投入使用的大型观测设备或进行巡天观测或对特定目标进行观测，但它们都有一个共同的特点：大数据量产出。例如我国将在2004年建成的LAMOST望远镜每夜观测得到的数据

约为 2GB¹，使我们大开眼界的哈勃太空望远镜每天的数据约为 5GB，设想中的大口径综合巡天望远镜 (Large-Aperture Synoptic Survey Telescope) 的日产数据量更高达 10TB²，相当于 500 个 20GB 硬盘的容量。

如此巨大的数据产出，在天文学历史上第一次使天文学家得到的数据多得用不了，那个辛辛苦苦观测许久但数据还是不充足的年代已经一去不复返。数亿个天体目标多个波段的数据资料的产生对数据处理、分析工具提出了更高甚至全新的要求，也为数据挖掘、高级模式识别、大规模交叉相关统计分析等这些新兴科学方法提供了用武之地。此外也在天文学历史上第一次使得高度复杂数值模拟结果可以与观测数据进行有科学意义的相互验证，结束了数值模拟结果因缺乏足够的观测数据而很难验证的历史。巨大而复杂的观测数据不断产生，迫切需要得力的工具、方法来发现蕴含在这些数据海洋中的宝藏。就是在这样的情况下，美国在名为“新千年中的天文学与天体物理学”科学展望中以首要位置提出了建立“国家虚拟天文台 (National Virtual Observatory)” (以下简称“虚拟天文台”) 的设想。这个虚拟天文台将把空间与地面观测设备得到的多波段巡天的海量数据有机地联合起来，同时提供利用这些数据资源进行科学研究所必需的各种计算机、网络方面的软硬件资源。未来的虚拟天文台将成为新天文发现的催化剂，为天文学带来一场新的革命。

虚拟天文台的科学价值

虚拟天文台将在未来十年中成为一个把天文学带入发现新时代的关键角色。虚拟天文台将是无可比拟的：拥有成 TB 上 PB³ 的数据资源、数百万天体从 射线直到射电波段的图像信息、功能强大的数据挖掘与分析工具、PB 量级的超级计算与数据处理能力、高速连接各大数据中心的宽带网络资源。由于虚拟天文台是以高速发展的通信技术、计算机科学为支撑，与传统概念上的天文台不同，故有其自己鲜明特点。首先，虚拟天文台是快速发展的。随着数据量的迅速增大、计算机网络的快速发展，虚拟天文台的各种软硬件设施都要及时更新以保持其先进性。第二，虚拟天文台是分布式的。无论其数据还是各种计算机软硬件资源都将分布在全球不同的国家和地区。第三，虽然物理构成上具有分布性，但功能上是有机的统一体，对用户提供整体的服务。第四，虚拟天文台是面向全球的。它的资源将为全球天文学家所共享。此外，拥有独一无二的各种资源的虚拟天文台必须承担起与它相称的社会责任，为公众了解天文，学习、利用天文提供服务。

得益于计算机计算能力、存储能力以及探测器技术的突破性进展，大范围多波段巡天计划不断出笼。利用多波段的巡天数据，我们可以产生出天体的全方位图像甚至宇宙的数字电影，从而使全新的天文发现成为可能。在历史上，光学与射电观测的结合导致了类星体的发现；红外天文学的产生导致了活动星系核、恒星形成区的发现。多波段巡天的结合更将带来更多全新的成果。虚拟天文台的建成将使天文学的内涵更加丰富。下面就以数字银河系为例来了解一下虚拟天文台的科学威力。

数字银河系

我们知道银河系是由银晕、银盘、核球、银心、旋臂等多种结构组成的复杂的天体系统。每个组成部分又是由具有一定年龄、质量、化学组成、动力学特征、分布特征的恒星以及非恒星物质的气体尘埃来体现。这些成分的特征为我们了解银河系形成提供了线索。但到目前为止，我们还没有能力同时研究如此众多的变化因素从而彻底了解我们的银河系。虚拟天文台将利用能提供银河系各种成分物理性质的大量观测数据建立一个数字化的银河系模

¹ 1GB 相当于 10^9 字节

² 1TB 相当于 10^{12} 字节

³ 1PB 相当于 10^{15} 字节

型，研究它的几何学、动力学特性，并以此模型与目前天文学家设想的各种模型进行比较，还可以利用模型的动力学特征寻找可能带来并合、潮汐等现象的运动成分。

为了建立这样的模型，虚拟天文台要结合光学与红外巡天数据，利用天体的位置、亮度、颜色数据产生它们的三维立体分布；利用远红外、21 厘米氢及一氧化碳分子光谱波段巡天图像量化尘埃的分布及其消光函数，并和恒星的三维分布模型反复拟合最终达成一致；利用红外、射电资料识别出巨流和恒星形成区；利用自行数据拟合恒星的运动；利用不同年代的数据研究变星。要实现这样的模型，虚拟天文台必须具备综合利用多波段、多历元巡天数据，使用用户自编程序进行图像处理、海量多维数据可视化以及相关统计分析的能力。

虚拟天文台的体系结构

虚拟天文台是高技术发展的产物，同时又对科技的发展提出了新的挑战。对成 TB 上 PB 数据的存贮、管理、检索、统计分析、可视化等等，都是期待解决的问题。更难的挑战可能还是这些数据的复杂性。要知道对几千万甚至更多天体中的每一个目标都有几十或几百项数据，处理起来谈何容易。

针对如此海量且复杂的数据，天文学家们设想虚拟天文台可由这样五个功能模块构成：天文数据库、数据描述协议、数据访问接口、检索与计算服务、高端应用。天文数据库按照逻辑联系及虚拟天文台数据标准存贮星表、图片、光谱等各种数据，同时提供访问服务（例如用 WWW 浏览器或 FTP 工具）。数据描述协议就是一套结构化的语言，用以对虚拟天文台的一些组成元素进行描述，比如对数据库的描述、对数据库提供的服务的描述、对数据库中数据表的描述、对基本数据的描述等。按统一的标准对这些元素进行描述是进行数据分析处理和数据间的相互操作所必需的。数据访问接口类似数据库软件中的 ODBC，它将为虚拟天文台中的所有数据、服务提供统一的应用接口。虚拟天文台中的所有部分都将共享这一接口，从而使虚拟天文台为用户提供的服务具有地理分布的透明性和数据调用的透明性。用户只管去享用这些服务而不用关心所使用的数据、计算机到底在地球的哪个角落。检索和计算服务是用户要进行研究工作不可缺少的。高端应用是直接为用户打交道的各种应用软件，像大数据量多变量的可视化、统计相关分析工具等。这些软件及其算法尚处于最初探索阶段，是未来相当一段时间重要的研究领域。

虚拟天文台除了这些组成部分，科普教育服务也是它不可或缺的。虚拟天文台能使科普教育工作者，例如教师，天文馆、太空馆、科学馆工作人员，科普作家，利用这些丰富的资源和服务建立科普网站、编写科普教材、开发高级教学软件等。

虚拟天文台的实施构想

上面介绍了这么多虚拟天文台的功能、组成，但到目前为止虚拟天文台还基本上仅是一个设想，与具体实现还相差甚远。当前最重要的任务是建立起庞大的天文数据基础设施和为访问这些数据所必须的良好规范的成文的规范协议。在这基础上才可能谈论如何开发各种工具、建立完善各项服务功能。美国的天文学家们为虚拟天文台构想了一个四步走的建设方案。第一步，在虚拟天文台计划正式启动之前进行概念设计，制定出相关文档，成立一些工作组和开发中心，为虚拟天文台的正式启动做好准备工作。第二步，正式启动后，用一年半左右的时间建立起数据检索、数据发布和初级数据证认服务。第三步，再用一年半的时间使虚拟天文台初步具备大规模相关证认能力，各项工作进入全面实施阶段。第四步，在此基础上用两年左右时间基本建成全功能的虚拟天文台，具备所有规模的相关证认能力并可提供优化的计算与网络服务。

SkyView，虚拟天文台初体验

SkyView 是由 NASA 开发维护的一个网上虚拟天文台，可以产生从 射线到射电波段任何部分天区的图像。针对不同层次的用户，SkyView 提供了非专业用户、基本、高级、Java、X - window 五种用户界面。对于普通天文爱好者这样的非专业用户，只要给出要显示的天区位置或天体名称，所要图像的观测波段，SkyView 就可以让你看到所要天区或天体的图像。SkyView 在进行图像提取时所进行的一系列复杂处理我们根本不用关心，坐享其成就可以了。

图一 SkyView 非专业用户界面

下面图中这组图像是 SkyView 显示的蟹状星云脉冲星，其中 射线图像数据来自 EGRET X 射线数据来自 PSPC 光学数据来自 DSS 红外数据来自 IRAS 射电数据来自 VLA NVSS。

图二 SkyView 提取的蟹状星云脉冲星不同波段图像

从 SkyView 我们可以粗略体验到虚拟天文台所具有的巨大威力。当然，这还不足冰山之一角。未来功能完备的虚拟天文台将具有远超过 SkyView 的数据资源，更强大的计算、分析和显示功能。虚拟天文台将为天文学美好的明天打下坚实的基础，为古老的天文学掀开崭新的一页。我国天文学家对虚拟天文台也有很高的热情并积极参加到建设中去。相信随着 LAMOST 望远镜、太阳空间望远镜（SST）、五百米口径射电望远镜（FAST）等一批大型观测设备的建成，我国天文学将在国际天文学界拥有更高发言权，在未来的虚拟天文台中发挥重要的作用。

参考网站：

- | | | |
|-----------------|---|---|
| 美国国家虚拟天文台 | ： | http://www.srl.caltech.edu/nvo |
| 首次虚拟天文台国际讨论会 | ： | http://www.astro.caltech.edu/nvoconf/ |
| SkyView 网上虚拟天文台 | ： | http://skyview.gsfc.nasa.gov |
| 虚拟天空 | ： | http://virtualsky.org |
| 中国 LAMOST 大科学工程 | ： | http://www.lamost.org |

致谢 中国国家天文台的赵永恒研究员、张彦霞博士、郑宪忠博士在本文完成过程中给予了热心帮助，在此向他们表示衷心感谢！