

분산 저장된 천문 디지털자료의 통합관리시스템 구현에 관한 연구

A Study on the Construction of Integrated Management System of Distributed Astronomical Data

성현일* · 임인성**

Hyun-Il Sung · In-Sung Yim

차 례

- | | |
|------------------|-------------|
| 1. 서론 | 5. 지속적 관리방안 |
| 2. 관측자료의 특성 및 획득 | 6. 결 론 |
| 3. 분산된 자료의 현황 | • 참고문헌 |
| 4. 시스템 설계 및 구현방안 | • 부 록 |

초 록

한국천문연구원의 지역천문대인 보현산천문대, 소백산천문대, 레문산천문대의 관측자료 활용을 용이하게 해주는 통합관리 시스템을 구축하였다. 관측자료는 생산된 천문대 각각에 아카이브로 보관하며, 통합시스템에는 관측자료에 관한 MySQL 테이블만 저장한다. 지역천문대와 통합시스템에는 각각 관리자가 있어 자료의 보관 및 헤더정보 추출, 테이블 생성 및 웹 페이지 관리 등을 역할 분담하여 수행한다. 사용자는 통합시스템의 웹 검색페이지 한곳에서 모든 천문대의 관측자료를 검색할 수 있으며 다운로드 또한 검색결과 페이지에서 가능하다. 통합시스템에서 독립된 각각의 천문대 아카이브로 쿼리를 보내 검색하는 방식은 가상천문대의 기본개념과 동일하다. 이와 같이 본 연구는 국제가상천문대 사업에 참여하는 한국가상천문대 구축 작업의 일환이기도 하다.

키 워 드

관측자료, 아카이브, 통합관리 시스템

* 한국천문연구원 선임연구원

(Senior Researcher, Korea Astronomy and Space Science, hisung@kasi.re.kr)

** 한국천문연구원 책임연구원

(Principal Researcher, Korea Astronomy and Space Science, yim@kasi.re.kr)

• 논문접수일자 : 2008년 11월 10일

• 게재확정일자 : 2008년 12월 9일

ABSTRACT

We constructed the integrated management system of distributed astronomical data of Bohyunsan Optical Astronomy Observatory(BOAO), Sobaeksan Optical Astronomy Observatory(SOAO), and Mt. Lemmon Optical Astronomy Observatory(LOAO). The observed data was stored in the archive of the observatory and MySQL table is saved in the integrated management system. There are data archive managers in the observatories and they manage the integrated management system such as data storing, head information extraction, table generation, and management of web pages, etc. The users can search and download all the data in the integrated management system web page. The concept of this system - send queries from web page to the observatory archives - is the same of the basic concept of virtual observatory. This research is one of the main work in the Korean virtual observatory which is take part in the international virtual observatory alliance.

KEYWORDS

Astronomical Data, Archive, Integrated Management System

1. 서론

1.1 연구의 목적

컴퓨터 성능의 발달과 네트워크 속도의 향상에 힘입어 대용량 자료의 저장과 처리 및 전송이 용이해졌고, 이를 바탕으로 공동활용이 활발하게 되었다. 이제, 연구에 사용하는 자료들이 대부분 디지털화 된 형태로 제작보관되는 시대가 되었다. 천문학 분야도, 관측자료인 천체이미지와 스펙트럼을 건판에 저장하던 아날로그 시대를 거쳐 오늘날에는 CCD(Charge Coupled Device)와 같은 장비를 이용하여 자료를 디지털 형태로 얻고 있다. 뿐만 아니라 건판에 담긴 과

거의 자료들도 디지털화 작업을 통해 재생산되어 귀중한 자료로 이용되고 있다. 미국의 경우는 팔로마천문대(Palomar Observatory)에서 촬영한 천천 사진건판을 디지털화하여 Digitized Sky Survey(DSS) 만들기도 했다(Pennington et al, 1993).

오랜 역사를 바탕으로 하는 천문학은, 발달 정도가 그 나라의 과학기술 수준을 가늠하는 기준 즉, 국가의 위상을 나타내는 잣대가 되기도 한다. 천문학 연구를 위한 가장 기본적인 장비인 망원경의 경우 현재 대부분의 선진국들은 8~10m급 광학망원경을 보유하고 있다(<그림 1> 참조). 한편 우리나라가 보유한 최대 크기의 광학망원경은 직경이 1.8m에 불과하다.

선진국의 천문대들은 생산되는 관측자료들을 아카이브 하여 누구나 이용할 수 있도록 하고 있다(〈표 1〉 참조). 이에 따라 관측자료의 이용률이 높아지는 경향을 보이고 있다.

관측장비 측면에서 우리나라의 천문학 연구 환경은 외국과 비교하여 열악한 상태이기 때문에 기 확보된 관측자료를 많은 천문학자들이 이

용할 수 있도록 해 주는 인프라를 구축할 필요성이 더욱 절실하다. 이러한 맥락에서, 본 연구에서는 우리나라의 각 천문대에서 생산되는 천문관측자료들을 하나의 표준화된 시스템 아래에서 보관하고 시간과 공간의 제약없이 누구나 이용할 수 있는 디지털 아카이빙 체제를 구축하기 위한 기술적 방안을 제시하고자 하였다.



〈그림 1〉 세계의 대형망원경과 한국의 망원경

〈표 1〉 세계의 천문 아카이브 현황

아카이브명	설명	관측 파장
AAT Archive	Anglo-Australian Telescope Archive	가시광
DSS Archive	Digitized Sky Survey Archive	가시광, IR
HST Archive	Hubble Space Telescope Archive	가시광, IR
ING Archive	Isaac Newton Group of telescopes Archive	가시광, IR
ISO Archive	Infra-red Space Observatory Archive	IR
IUE Archive	International Ultraviolet Explorer Archive	UV
JCMT Archive	James Clerk Maxwell Telescope Archive	sub-mm
ROSAT Archive	Röntgen Satellite Archive	X-ray
UKIRT Archive	UK Schmidt Telescope Archive	IR

1.2 연구의 방향

천문대에서 얻어진 관측자료들은 관측자들이 자신의 자료를 저장매체에 직접 담아가서 분석을 하며, 천문대는 각 천문대의 자료보관 원칙에 따라 얻어진 자료를 개별적으로 보관하여 왔다. 관측자료의 사용권한은 우선 관측자에게 주어지고, 관측시설을 제공하는 천문대는 일정기간이 지난 관측자료에 대한 사용권한을 갖는 것이 일반적이다. 천문대는 가치 있는 관측자료들이 더 많이 이용될 수 있도록 데이터베이스화하여 공개하고 있다(Moran 2007; Arviset et al. 2007).

보현산천문대와 소백산천문대, 레몬산천문대도 각 천문대에서 보관하고 있는 관측자료들을 각각의 천문대 내에 데이터베이스화하고 있다(성현일 등 2003). 그러나 현재는 각 천문대의 자료들을 독립적인 데이터베이스로 제작 중이므로 이를 이용하는 사용자 입장에서는 자료의 검색과 추출이 불편할 수밖에 없다. 이렇게 개별 시스템으로 구성 중인 데이터베이스를 하나의 관리체계 아래에 통합하면 자료의 관리와 이용이 쉬어질 뿐만 아니라 각 천문대의 관측자료 형태를 표준화 시킬 수도 있다. 또한, 동일한 천체에 대한 관측자료일지라도 각기 다른 천문대에서 관측이 수행될 경우 각각 따로 보관되어지는데, 통합된 관리체계 아래에서는 자료의 물리적 위치와 무관하게 검색을 통해 이들 관측자료들을 모두 찾아 동시에 이용할 수 있어 연구의 효율도 높아진다.

2장에서는 대상자료의 분석에 대해 설명하고, 3장에서는 분산되어 있는 관측자료의 현황을, 4장에서는 시스템 설계 및 구현방안을, 5장에서는 지속적 관리방안을 설명하고 6장에서는 결론을 다루고자 한다.

2. 대상자료 분석

2.1 관측자료 획득

관측을 희망하는 연구자는 관측계획서를 천문대에 제출하고 심사를 통해 선정이 되면 주어진 기간에 천문대에서 직접 관측을 하게 된다. 관측은 자신이 원하는 관측기기를 부착하여 원하는 필터 또는 분해능으로 적절한 노출을 주어 얻게 된다. 보현산천문대와 소백산천문대 그리고 레몬산천문대는 망원경 및 각종 장비의 구동을 도와주는 관측보조자를 두고 있어 이들의 도움을 받으며 관측을 수행 한다. 얻어진 관측자료는 우선 관측장비와 연결된 컴퓨터에 저장되고 관측자는 이 자료를 자신이 준비한 저장매체에 담아서 가져가고 있다.

관측자료는 자료의 특성에 따라 분석하는 도구가 달라지게 되는데 천문학분야의 측광관측과 분광관측 자료의 분석에 가장 많이 사용되는 프로그램은 IRAF(Image Reduction and Analysis Facility, <http://iraf.noao.edu/>)이다. 모든 자료는 CCD를 이용하여 디지털화된 형태로 얻게 되는데, CCD의 특성에 의한

효과들을 제거하는 전처리 과정을 먼저 거치게 된다. 전기적 신호의 불균질성에 의한 영점(bias), 열잡음에 의한 암잡음(dark), 각 픽셀의 반응도 차이에 따른 바닥(flat) 등을 보정해 주면 전처리가 끝나게 되고 정량화가 가능한 수준에 이르게 된다.

2.2 관측자료의 형식

일반적으로, 관측자료는 fits(flexible image transport system) 파일의 형태로 얻어진다

(Wells 1981). fits 파일은 천문학 분야의 자료를 위해 개발된 파일의 형태로써 NASA와 국제천문연맹에서 승인하였으며 대부분의 천문관측자료는 이 형태로 이루어져 있다. fits 파일은 자료의 정보를 담고 있는 헤더 부분과 자료의 내용을 담고 있는 부분으로 나뉜다. 헤더 부분은 <그림 2>와 같이 관측자, 망원경, 관측기기에 대한 정보뿐만 아니라 대상천체의 이름, 좌표, 파일형태, 관측날짜, 관측시간, 천체고도 등을 표시하고 있으며 CCD의 각종 특성에 관한 정보도 담고 있다.

```

SIMPLE = T / simple fits format
BITPIX = -32 / unsigned 16-bit data
NAXIS = 2 / 2-dimensional image
NAXIS1 = 2200 / number of columns
NAXIS2 = 2048 / number of rows
BZERO = 32768.00 / real = tape*b scale + bzero
BSCALE = 1.000000 / real = tape*b scale + bzero
OBSERVER = 'H.I.Sung' / observer
OBSERVAT = 'BOAO' / observatory
TELESCOP = 'BOAO 1.8M F/8' / telescope
DETECTOR = 'BOAO CCD 2K' / detector
CCDPICNO = 35237 / CCD sequence
OBJECT = 'HH223' / object name
IMAGETYP = 'object' / image type: Bias, Dark, Flat, Object
COMMENTS = ' ' / comments
DATE-OBS = '2006-04-29' / date (dd/mm/yy) of obs.
UT = '17:04:11.00' / universal time (hh:mm:ss.ss) exposure start
ST = '16:10:08.80' / sidereal time (hh:mm:ss.ss)
HA = '20:51:26.97' / hour angle (hh:mm:ss.ss)
RA = '19:18:41.832' / right ascension (hh:mm:ss.sss)
DEC = '19:10:22.15' / declination (dd:mm:ss.ss)
EPOCH = 2006.3250136576 / epoch of ra and dec
AIRMASS = 1.4022257983 / airmass
DROTATOR = 302.5018700000 / derotator angle (decimal degree)
POSANGL = 333.1234567890 / slit position angle from E-W (decimal degree)
FOCUS = -2591 / secondary mirror position (encoder)
FILTER = 'HaW(150A)' / filter name
EXPTIME = 1800.000 / exposure time (sec)
GAIN = 1.80000 / gain, electrons per adu
RDNOISE = 7.00000 / readout noise, electron
CHIPTEMP = -100.45420 / CCD chip temperature (celsius)
COLBIN = 1 / binning factor, column
ROWBIN = 1 / binning factor, row
COLORIG = 0 / subregion origin, column
ROWORIG = 0 / subregion origin, row
COLSECT = '[ 20:2048: 132]' / image sector [prescan:active:overscan]
ROWSECT = '[ 0:2048: 0]' / image sector [prescan:active:overscan]
COLCENT = 1043.5000000000 / ccd center, column
ROWCENT = 1023.5000000000 / ccd center, row
COLPIXEL = 0.3437746770 / pixel size in arcseconds, column
ROWPIXEL = 0.3437746770 / pixel size in arcseconds, row
COLSEG = '[0:1000, 1:1024]' / frame format
COMMENT1 = ' ' / comments
COMMENT2 = ' ' / comments
COMMENT3 = ' ' / comments

```

<그림 2> 관측자료의 정보를 담고 있는 fits 파일의 헤더 부분

2.3 천문대 설치 조건

천문대는 날씨가 맑고 대기가 안정적이며, 광공해가 적고 접근성이 좋은 곳을 찾아 건설한다. 대기의 영향을 적게 받기 위해서는 고도가 높은 산으로 올라가야 되는데, 전 세계에서 천문대 설치 최적지로 현재 알려진 곳은 북반구에서는 하와이, 남반구에서는 칠레 등이다. 하와이의 마우나케아산에는 미국뿐만 아니라 여러 나라의 대형망원경 10여기가 설치되어 운영중인데, 단일경으로는 세계 최대 크기인 일본의 8.2m 수바루망원경과 쌍둥이망원경인 10.5m 켄망원경 등이 있다. 칠레의 안데스산맥에는 6.5m 마젤란망원경 2기와 8.2m VLT 망원경 4기 등이 있으며, 8.3m 제미니망원경은 하와이와 칠레에 하나씩 설치되어 있다.

우리나라의 경우, 최적의 관측지는 보현산과 소백산으로 알려져 있으며 현재 이들 장소에 각각 천문대가 건설되어 관측연구가 이루어지고 있다.

3. 분산된 자료의 현황

한국천문연구원에서 운영 중인 연구용 망원경으로는 보현산천문대(BOAO)의 1.8m 망원경과 소백산천문대(SOAO)의 61cm 망원경, 그리고 레몬산천문대(LOAO)의 1m 망원경이 있다. 이들 망원경을 이용하여 획득한 관측자료는 각각의 천문대에서 보관하고 있다.

3.1 국내 관측자료

3.1.1 보현산천문대

보현산천문대의 1.8m 망원경은 1996년에 건립되어 관측에 들어갔다. 관측장비는 과학기술의 발달에 힘입어 계속 발전해 왔는데 현재, 측광 관측장비로는 2KCCD 카메라와 적외선 이미징 카메라 KASINICS(KASI Near Infrared Camera System)가 있고, 분광관측장비로는 고분산 Echelle 분광기가 있다.

보현산천문대는 지금까지 관측한 자료를 모두 DAT 테이프와 CD 및 하드디스크에 백업하여 보관하고 있다. 1996년부터 2002년까지 관측한 약 500GB 정도의 자료가 50개의 DAT 테이프와 800장의 CD에 저장되어 있으며, 이들 자료 중 1997년 4월부터 2002년 12월까지 관측된 약 400GB의 자료를 아카이브로 구축하였다. 현재는 관리와 이용이 용이하도록 모든 자료를 하드디스크에 저장하고 있다.

3.1.2 소백산천문대

1978년 9월 소백산에 건설된 소백산천문대는 61cm 반사망원경의 가동과 함께 현대적인 광학천문학 연구의 막을 열었다. 주로 밝은 천체에 대한 광전측광을 수행해 오던 소백산천문대는 1992년부터 CCD 카메라를 이용하여 어두운 천체에 대한 관측연구를 수행하고 있다. 그 후 2001년부터는 광시야 2KCCD 카메라를

이용하여 정밀측광 관측연구를 주도하고 있다.

소백산천문대는 내부관측자료만 보관하고 있으며 1년에 30GB 이상의 자료가 축적되고 있다.

3.2 국외 관측자료

레몬산천문대의 1m 광학망원경은 2001년 말에 설치되어 약 1년 6개월 동안 시험관측후 2003년 9월부터 본격적으로 천체 관측 연구에 활용되고 있다. 해발 2,776m에 위치한 레몬산천문대는 측광 가능한 날이 약 200일/년 이상으로 국내에 비해 관측 여건이 매우 좋은 곳이다.

모든 관측은 대전의 한국천문연구원에서 원격으로 이루어지고 있다. 본격적인 관측이 이루어진 2004년부터 매년 약 300~400GB의 자료가 축적되고 있다.

3.3 기타자료

한국천문연구원의 지역천문대에서 연구용으로 생산된 관측자료 이외에 연구용망원경과 소형망원경으로 수년 동안 쌓아온 천문현상 관측자료들이 있다.

대표적인 천문현상에는 태양과 달에 관련된 일식과 월식이 있으며, 행성들과 관련된 목성식, 토성식, 금성의 태양면 통과, 행성직렬 등이 있다. 또한 태양계 내의 작은 천체들과 관련된 유성우, 소행성, 혜성 등의 출현이 있다.

천문현상 자료 아카이브는 1991년부터 2005년까지 얻어진 자료를 축적하고 있으며, 이미지의 개수는 4,000여장이다(성현일 등 2006).

4. 시스템 설계 및 구현방안

한국천문연구원의 지역천문대에서 생성된 관측자료는 각 천문대에서 보관하고 있으며, 데이터베이스화 작업도 개별적으로 진행 중이다. 현재, 보관 중인 관측자료에서 자신이 원하는 자료를 찾기 위해서는 각각의 데이터베이스에 접속하여 자료를 검색하고 검색결과에서 필요한 자료를 다운로드 받게 되어 있다.

이러한 불편함을 줄이고 연구의 효율을 높이기 위해서는 각기 다른 지역에 분산 저장된 자료들을 한 곳에 있는 자료처럼 한꺼번에 검색하고 다운로드 할 수 있는 시스템이 만들어져야 한다. 또한, 자료의 축적과 관리도 동일한 관리시스템 아래에서 보다 체계적으로 이루어져야 한다.

산재된 자료를 동시에 검색 할 수 있는 데이터베이스 시스템을 설계하고 구현하는 방법은 가상천문대(Virtual Observatory) 구축을 향한 첫걸음이기도 하다. 현재 세계 각국은 가상천문대 구축작업을 활발히 진행하고 있다 (<표 2> 참조).

4.1 아카이빙 대상

각 천문대에 보관중인 관측자료들을 분석하여 사용 가능한 자료를 골라내는 작업을 먼저 하여야 한다. 날씨와 기기의 상태에 따라 자료의 품질이 나쁜 경우가 있는데, 이런 자료는 이용할 수 없으므로 미리 선별하여 데이터베이스에 들어가지 않도록 한다.

〈표 2〉 세계의 가상천문대 현황

가상천문대	가상천문대 명칭	국가
KVO	Korean Virtual Observatory	한국
EURO VO	European Virtual Observatory	유럽연합
China-VO	Chinese Virtual Observatory	중국
VO-i	Virtual Observatory India	인도
CVO	Canadian Virtual Observatory	캐나다
SVO	Spanish Virtual Observatory	스페인
DRACO	Italian Virtual Observatory	이탈리아
ArVO	Armenian Virtual Observatory	아르메니아
VO France	Observatoire Virtuel France	프랑스
GAVO	German Astrophysical Virtual	독일
HVO	Hungarian Virtual Observatory	헝가리
JVO	Japanese Virtual Observatory	일본
NVO	National Virtual Observatory, United States	미국
RVO	Russian Virtual Observatory	러시아
AstroGrid	Virtual Observatory United Kingdom	영국
Aus-VO	Australian Virtual Observatory	호주

관측시 작성하는 로그북에는 관측 당시의 여러 가지 상황들이 기록되어 있다. 일차적으로, 로그북에 기록된 사항을 기반으로 품질이 떨어지는 관측자료를 버린다. 그 다음에 fits 파일을 열어 볼 수 있는 다양한 도구들을 이용하여 각 파일들을 열어 이미지에 문제가 없는지 검사하여 잡음이 심하거나, 추적이 제대로 되지 않은 것, 시상이 많이 나쁜 파일들을 버린다.

관측자료는 전처리 과정을 거쳐야 하는데, 이를 위해서는 bias, flat, 그리고 dark 파일이 필요하다. 천체 관측시, 천체뿐만 아니라 전처리에 필요한 자료들도 동시에 얻게 되는데 각 천체의 관측자료와 이를 위한 전처리 파일들을 짝지어 주는 작업이 있어야 한다. 대부분의 경우 동일한 날짜에 얻어진 전처리용 파일을 이용하기 때문에 파일을 짝지어 주는 작업을 자동화 시킬 수 있다. 한 발 더 나아가, bias, flat, 그리고 dark 를 자동으로 일괄 보정해주는 전처리 시스템인

파이프라인을 구축한다면, 자료를 이용하는 연구자의 연구효율을 훨씬 높일 수 있다.

4.2 시스템 설계

데이터베이스 관리 시스템(Database Management System, DBMS)은 공개용으로 널리 사용되고 있는 MySQL을 채택하였다. MySQL은 관계형 데이터베이스 관리 시스템인 SQL (Structured Query Language)의 하나로, 사용자는 유틸리티를 이용하여 테이블 구조로 저장되어 있는 MySQL 데이터베이스의 데이터들에 접근할 수 있다(Kimula & Takahashi 2002). MySQL은 상용 데이터베이스와 비교하여 뒤지지 않는 뛰어난 관계형 데이터베이스로, 안정성 측면에서도 우수하고 보안이나 각종 함수도 많아 프로그래밍에 용이하다. 또한, MySQL은 UNIX

와 Solaris, AIX, FreeBSD, Linux, Windows 등 다양한 플랫폼에서 사용 가능하다(조준의 2000).

4.2.1 단일 천문대 아카이빙 방법

천문대 한 곳의 관측자료를 아카이빙하고 검색가능한 데이터베이스화하는 작업은 MySQL을 이용한 전형적인 데이터베이스 제작방법과 비슷하게 진행된다. 다만, 관측자료는 일반 사진자료와 달리 영상을 전처리하는 과정이 필요하므로 어떤 상태의 관측자료를 아카이브 할 것인가를 고려하여야 한다.

따라서 아카이브에서 관측자료를 제공하는 방법은 두 가지가 있다. 첫번째는, 대상천체의 자료와 함께 bias 영상, dark 영상, flat 영상 등 전처리에 필요한 자료를 원시자료의 형태 그대로 각각 제공하는 방식이다. 두번째는 bias 보정, dark 제거, flat fielding 등의 전처리를 수행해서 얻어진 대상천체의 자료만을 제공하는 방식이다. 이 경우에는 전처리에 필요한 자료들은 아카이빙 할 필요가 없다. 그러나 전처리는 작업자에 따라 결과가 약간씩 다르게 나올 수 있는 것이므로 전처리 되지 않은 자료를 직접 전처리를 하기를 원하는 사용자들도 있다.

전처리 하지 않은 원시자료의 형태로 제공하게 될 경우는 아카이브에 많은 관측자료를 저장하여야 하며 사용자가 필요로 하는 파일의 숫자와 전체 용량이 커져 자료 전송에 어려움이 생기게 된다. 한편, 전처리를 수행하는 경우에는 기상조건과 기기상태의 변화 등 관측조건에 따른 특성들을 고려하기가 어렵기 때문에 사용자

가 요구하는 정확도를 달성하기가 어려운 단점이 있다. 따라서 보현산천문대와 소백산천문대, 레몬산천문대의 관측자료 아카이브는 모든 자료를 원시자료의 형태로 보관하고 있다.

아카이빙은 MySQL 시스템에서 자료의 저장과 검색, 추출이 가능한 형태로 색인하여 저장한다(〈표 3〉 참조). 관측후 생성되는 모든 자료는 데이터베이스 내에 추가하지만, 관측자료는 관측 후 자료처리와 분석에 많은 시간이 필요하다. 이런 특성 때문에 관측 후 1년 동안은 관측자만이 그 자료에 접근하여 사용할 수 있도록 하고 있다.

4.2.2 통합 아카이빙 방법

각 지역천문대의 관측자료 아카이브는 통합된 시스템 아래에서 운영해야 관리와 사용 모두 효율적이 된다. 통합하는 첫 번째 방법은

〈표 3〉 MySQL 내의 테이블 내용

컬럼명	내용	데이터형
Name_File	파일명	varchar(15)
Type	자료종류	varchar(15)
Name_Obj	관측대상 이름	varchar(15)
Coord_RA	RA(관측날짜)	varchar(15)
Coord_Dec	Dec(관측날짜)	varchar(15)
Date	관측날짜	varchar(15)
UT	관측시간	varchar(15)
Filter	사용필터	varchar(15)
Time_Exp	노출시간	varchar(15)
Observer	관측자명	varchar(15)
Observatory	천문대명	varchar(15)
Instrument	관측기기	varchar(15)
Name_Fitsfile	fits 파일명	varchar(50)
Coord_RA_2000	RA(2000)	decimal(20,17)
Coord_Dec_2000	Dec(2000)	decimal(20,17)
Coord_RA_1950	RA(1950)	decimal(20,17)
Coord Dec 19500	Dec(1950)	decimal(20,17)

자료를 물리적으로 한 곳에 모으는 것이고, 두 번째 방법은 지역천문대에 아카이브 되어 있는 자료를 네트워크로 연결하여 한 곳에 있는 것처럼 가상시스템을 만드는 것이다.

첫 번째 방법의 경우, 일년 내내 지속적으로 생산되는 관측자료를 한 곳에 모으려면 수시로 자료를 옮겨야 하는 단점이 있는 반면, 데이터베이스 시스템을 보다 간단하게 구성할 수 있다는 장점이 있다.

두 번째 방법의 경우, 각각의 지역천문대에서 생산된 관측자료는 그 천문대 내에 아카이브하여 보관하기 때문에 자료를 이동하지 않아도 되는 장점이 있지만, 중앙의 데이터베이스 시스템과 각 지역천문대를 연결하는 고속의 네트워크를 설치하여야 하고 이를 다룰 수 있는 데이터베이스 프로그래밍 기술이 수반되어야 한다는 단점이 있다.

장기적인 관점에서 볼 때, 각 지역천문대에서 생산되는 관측자료는 점점 늘어나고, 네트워크 기술의 발달로 자료의 전송속도가 더욱 빨라질 것이므로 자료를 생산지인 지역천문대에 저장하고 가상시스템으로 묶는 두 번째 방법을 채택하는 것이 바람직하다. 한국천문연구원 대전 본원의 전산시스템을 중심으로 지역천문대를 네트워크로 연결하여, 각 지역의 관측자료를 중앙에서 관리하는 시스템을 구축하면 된다(〈그림 3〉 참조).

4.2.3 시스템 구현시 쟁점과 해결방안

시스템을 개발과정의 쟁점사항으로는 기술



〈그림 3〉 대전 본원과 지역천문대의 네트워크 연결

적인 부분과 기술외적인 부분이 있다. 세부항목으로는 1)자료의 표준화, 2)자료의 통합 및 연계, 3)자료의 품질 개선, 4)보안성 및 안정성 확보, 5)소요예산 확보, 6)인력 확보, 7)규정 및 제도의 정비, 8)유관부서간 협조 및 참여 등이 있다.

각 천문대에서 생산되는 파일들은 모두 fits 파일의 형식이지만, fits 파일의 헤더 부분은 천문대마다 약간씩 다른 내용을 포함하고 있다. 그러므로 자료의 표준화는 기본적으로 수행되어야 할 작업이다. 관측자료 자체의 변형 없이 이 문제를 해결하기 위해 MySQL 테이블 작업시 fits 파일들의 헤더에서 공통되는 주요 키워드를 골라 표준으로 삼는 방법을 선택했다. 헤더의 다른 정보와는 무관하게 이들 키워드만으로 자료 검색을 하게 된다. 자료의 통합 및 연계는 본 연구의 시작점이기도 하다. 분산

된 자료를 물리적으로 합치지 않고, 자료의 정보만을 한 곳에서 관리하는 방법을 선택하여 자료의 축적과 관리가 용이한 형태로 시스템을 구축하였다. 관측자료는 관측할 때의 기상상태와 기기 작동상태에 따라 품질이 달라진다. 사용가능한 관측자료는 일정 수준 이상의 품질이어야 하지만, 저품질의 자료일지라도 전처리 등의 과정을 거치면 품질이 향상되기도 한다. 따라서 전처리에 필요한 기본적인 자료들이 함께 얻어진 관측자료들은 저품질일지라도 데이터베이스에 포함시키고 있다. 보안성의 관점에서 자료의 공개시기가 문제가 된다. 관측자료는 일정기간 동안은 관측자가 자료에 대한 모든 권한을 갖게되어 있다. 자료의 분석과 연구 결과의 생산이 가능한 일년정도의 기간이 지나면 모든 자료는 외부에 공개하는 것을 원칙으로 하고 있다. 관측 즉시 모든 자료는 데이터베이스화 되지만, 일년 이내의 자료에 대한 접근 권한은 관측자에게만 부여하여 이 문제를 해결한다. 시스템 전체의 구성을 위한 소요예산과 인력의 확보는 한국기상천문대 구축사업에서 지원이 가능하다. 한편, 관측자료의 데이터베이스화 작업을 시작하면서 자료공개에 대한 규정을 확립하였다. 외국의 천문대와 마찬가지로 국내 천문대에서도 천문대별로 내부규정을 마련하여 일년이 지난 자료는 자동적으로 공개가 가능하도록 하고 있다. 관측자는 이 규정을 준수하는 것을 전제로 관측신청을 하게 된다. 유관부서의 협조와 참여문제는 관측 자료를 생성하는 지역천문대와 전체를 통합관리하는 대전 본원과의 유기적인 협조체제를

구축하여 해결한다. 지역천문대는 생성되는 자료를 저장하고 관리할 수 있는 틀을 제공 받는다는 잇점이 있으며, 대전 본원에서는 외부의 사용자들에게 자료를 제공하여 자료이용의 효율을 극대화 한다는 잇점이 있으므로 부서들 간의 협조체제 구축은 자연스럽게 이루어진다.

4.3 메타데이터 요소설계

4.3.1 메타데이터 요소의 도출과정

메타데이터의 정의는 데이터에 대한 데이터이지만, 오늘날에는 디지털 형태의 정보 자원에 대한 소재나 내용을 기술하고 이를 검색하고 이용하기 위한 요소들의 집합이라는 의미로 통용된다. 관측자료의 메타데이터 요소는 이용대상인 천문전공자의 검색 효율성과 자료 제공자인 지역천문대 관리자의 메타데이터 작성의 편의성 및 관리의 효율성을 고려한 메타데이터 추출을 모색하였다.

메타데이터 요소를 도출하기 위한 과정은 다음과 같다. 1)자료가 생성되는 지역천문대의 fits 파일 헤더정보를 조사한다. 2)메타데이터 기술요소를 도출하기 위해 각 지역천문대 자료의 헤더정보를 나열한다. 3)헤더정보 요소와 속성들을 나열하고 상호 매핑하여 공통요소를 추출한다. 4)도출된 공통 요소 가운데 필수적인 요소를 선별하고 불필요한 요소를 배제한다. 5)공통 요소 이외의 요소 가운데 필요하다고 판단되는 요소를 선별하고 추가한다.

4.3.2 도출된 메타데이터 요소

도출된 메타데이터 세트는 필수요소와 부가 요소로 크게 나눌 수 있다. 필수요소인 관측대상 이름 (Object name), 좌표 (Right Ascension, Declination), 기준년도 (Epoch), 추출영역범위 (Search box size)는 모든 지역천문대 메타데이터에서 공통적으로 사용되는 요소이며, 자료를 식별하기 위한 가장 기본이 되는 요소이므로 필수요소로 정하였다. 요소들 중 이름과 좌표는 두가지 모두 관측대상을 지정하기 위한 것이므로 둘 중 하나를 선택하여 사용하면 된다.

관측날짜 (Observation data), 관측자 (Observer), 관측천문대 (Observatory), 자료형태 (Data type), 관측기기 (Instrument)와 같은 부가요소는 대부분의 메타데이터에서 사용하고 있는 요소이지만 특정한 기준에 부합하는 자료의 추출을 위해 적용하는 기준에 해당한다.

메타데이터의 구성을 위한 테이블스키마 관련 프로그램은 부록에 포함하였다.

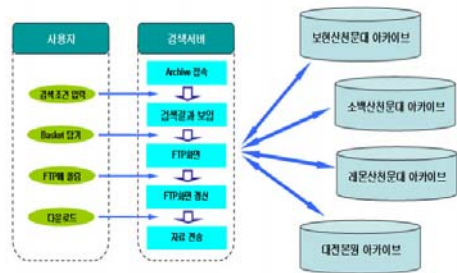


〈그림 4〉 관측자료 검색 웹 화면

4.4 통합 시스템 및 주요모듈 설계

관측자료를 검색하는 웹 페이지는 보현산천문대 관측자료의 단일 아카이빙 시스템을 기반으로 〈그림 4〉와 같이 구성했다. 검색화면에서는 여러 가지 검색조건을 입력하여 원하는 자료를 찾게 되는데, 반드시 입력하여야 하는 필수검색조건과 추가검색조건으로 나눌 수 있다. 필수검색조건은 관측대상의 이름 또는 좌표 등이고, 추가검색조건은 관측날짜, 자료형태, 관측자, 관측기기, 관측천문대 등이다.

사용자가 웹 서버에 접속하여 검색화면에서 관측자료를 검색하여 결과를 보고 자료를 추출해 내는 과정은 〈그림 5〉와 같다. 사용자는 검색화면에서 검색조건을 입력하면 추가검색조건인 관측천문대가 설정되어 있는 경우 해당 아카이브에 접속하여 조건에 맞는 자료가 있는지 검색을 한다. 만약 추가검색조건인 관측천문대를 지정하지 않았다면 검색서버에 연결되어 있는 모든 천문대의 아카이브에 쿼리를 보내 검색결과를 받아오게 된다.



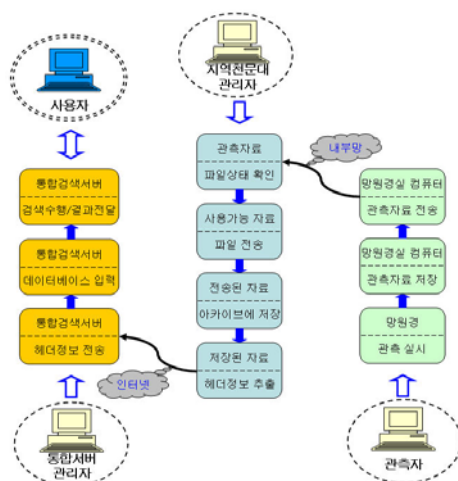
〈그림 5〉 관측자료 검색후 지역천문대에서 자료를 추출해 내는 과정

사용자는 검색결과를 보고 원하는 자료를 체크하여 Basket에 담게 된다. 이때 자료의 품질을 검사할 수 있도록 미리보기 기능을 부여하였다. 관측자료는 기본적으로 fits 형태로 저장되어 있지만 미리보기를 원하는 자료는 jpeg 형태로 변환되어 사용자의 화면에 보여진다. 파일을 Basket에 담으면 해당 파일들의 목록이 저장된다.

Basket에 담은 파일을 가져오기 위해서는 FTP를 이용할 수도 있고, CD에 담아서 우편으로 받을 수도 있다. 자료의 양이 많지 않는 경우에는 FTP를 주로 이용하게 된다. FTP에 올리기 명령을 보내면 각각의 아카이브에서 검색서버의 FTP 디렉터리에 파일들이 복사되고 사용자는 FTP 디렉터리에 담겨진 파일을 다운로드하면 된다. 외국의 천문대들은 네트워크 속도가 느린 경우가 많아 한국에서 외국 아카이브 자료를 이용하고자 할 때는 다운로드 방식보다 CD에 담아서 우편으로 받는게 더 효율적인 경우가 있지만 국내의 지역천문대들은 고속의 네트워크가 연결되어 있어 자료를 다운로드하는 방식이 더 편하다.

통합시스템의 전체적인 구성은 <그림 6>과 같다. 이 시스템과 관련된 사람은 관측자와 지역천문대 관리자, 통합서버 관리자, 그리고 사용자로 나누어진다. 먼저 관측자가 망원경을 이용하여 관측을 실시하면 망원경실의 컴퓨터에 자료가 저장된다. 일정량의 자료가 저장되면 이 자료들을 그 천문대 내부의 네트워크

를 이용하여 천문대 내 아카이브 컴퓨터의 임시저장 디스크로 전송한다. 전송된 관측자료는 파일의 상태를 확인하여 사용 불가능한 자료들을 제거하는 선별과정을 거친뒤 아카이브의 디스크 내에 저장된다. 그리고 데이터베이스화를 위해, 저장된 파일의 헤더 정보를 추출한다. 이 일련의 과정은 지역천문대의 관리자가 담당한다. 저장된 헤더 정보는 천문대 외부망을 이용하여 통합서버로 전송된다. 통합서버 관리자는 전송된 헤더 정보를 이용하여 데이터베이스의 테이블 목록을 제작하여 통합검색서버의 데이터베이스에 입력한다. 각 지역천문대에서 관측된 자료의 정보가 통합서버에 모이면 사용자는 통합서버의 웹에 접속하여 모든 천문대의 관측자료를 검색할 수 있게 된다. 검색결과 다운로드를 원하는 관측자료는 외부망을 이용하여 지역천문대로부터 전송받게 된다.



<그림 6> 통합시스템의 전체 구성 및 흐름도

5. 자료 관리방식

자료의 생성에서 제공까지의 전 과정을 살펴보면 <그림 6>과 같다. 이 과정은 관리의 관점에서 크게 3부분으로 나눌 수 있으며, 각 부분에 관련되는 사람은 관측자, 지역천문대 관리자, 통합서버 관리자 등이다. 관측자가 수행하는 관측실시와 자료저장, 전송 등의 과정은 자료의 축적에 해당한다. 축적된 자료는 각 지역천문대 관리자의 파일 품질 확인 및 선별, 유용한 파일 아카이브 저장, 헤더정보 추출 등의 자료관리 과정을 거친다. 이렇게 각 천문대별로 준비된 자료는 통합서버로의 헤더정보 전송과 데이터베이스 입력 등의 자료 공개를 위한 단계를 거친다. 그러므로 관리의 전 과정은 자료축적, 자료관리, 자료공개의 세부분으로 크게 나눌 수 있다. 자료의 획득, 처리 및 공개와 관련되지 않은 인력 및 자원의 활용 등 일반적 관리는 본 장의 관리방안에서 제외하였다.

5.1 자료 축적

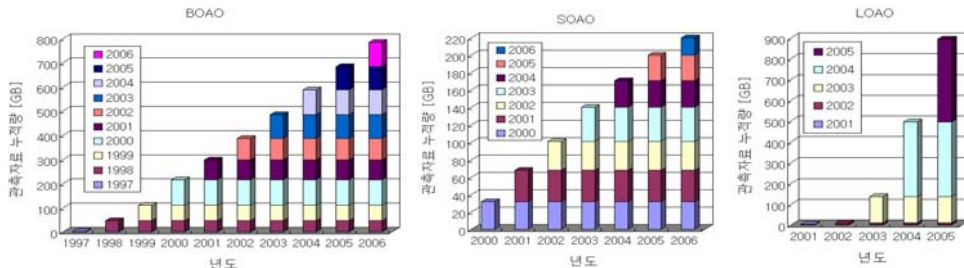
각 지역천문대는 망원경과 관측기기가 건설

후 안정화되는 시점까지 관측자료의 생산량이 매년 조금씩 증가하였고, 관측시스템이 안정화된 이후에는 해마다 생산되는 관측자료의 양이 일정한 수준에 이르러 큰 변동이 없었다. 매년 축적되는 관측자료의 양은 보현산천문대의 경우 약 100GB이며, 소백산천문대의 경우는 30GB, 레몬산천문대의 경우는 약 300~400GB 정도이다(<그림 7> 참조).

각 천문대는 늘어나는 관측자료를 보관하기 위하여 대용량 자료 저장 시스템을 구축 중에 있다. 관측자료는 관측이 끝나는 즉시 아카이브에 저장되고 검색서버와 연결되어 검색 결과에 표시되지만, 관측 후 1년이 지난 자료만 다운로드가 가능하다.

5.2 자료 관리

관측자료의 아카이브를 위해서 각 천문대에 자료의 자동저장 시스템을 구축 중에 있다. 관측자가 이용하는 컴퓨터는 2개로, 관측기기를 제어하는 것과 자료를 저장하는 것이 있는데, 관측자료가 저장되는 컴퓨터는 그 천문대의 아카이



<그림 7> 지역천문대별 관측자료 축적 현황

브 서버와 연결되어 관측이 끝난 뒤 오전 시간에 자동으로 자료가 아카이브 서버에 옮겨지도록 프로그래밍 할 예정이다.

각 지역천문대는 아카이브 서버의 안정적인 운영을 위해 관리자를 두며, 대전 본원의 통합서버 관리자는 검색서버의 프로그램뿐만 아니라 검색서버와 지역천문대의 연결, 그리고 사용자 관리도 하여야 한다. 한꺼번에 많은 사용자가 몰릴 경우 네트워크 속도의 저하와 사용자 FTP 디렉터리의 용량 초과 여부를 체크하여 서버가 항상 최상의 성능을 유지할 수 있도록 관리 한다.

5.3 자료 공개

관측자료는 1년 동안 관측자가 사용권한을 가지며 관측 후 1년이 지나면 외부에 공개하여 누구나 사용할 수 있게 하고 있다. 따라서 모든 관측자는 자료를 빨리 분석하여 연구결과를 얻고자 노력하게 된다. 그러나 관측 후 1년 이내에 논문화되지 못하는 관측자료들이 많으며 이 자료들의 활용을 극대화시키기 위해 많은 천문대에서 데이터베이스의 검색 성능을 높여려는 노력을 하고 있다. 본 연구의 통합검색시스템도 흠어져 있는 관측자료를 쉽게 검색하고 잘 활용할 수 있는 기반을 제공해 주는데 있다.

6. 결론

선진국과 비교하여 상대적으로 열악한 관측

시설을 보유한 한국 천문학계에서는 기존 관측자료의 활용을 최대화 할 필요성이 더욱 커지고 있었다. 따라서 기 확보된 관측자료를 많은 천문학자들이 이용할 수 있도록 해 주는 인프라의 구축이 절실했다. 이에 한국천문연구원 지역의 지역천문대인 보현산천문대, 소백산천문대, 레몬산천문대의 관측자료 활용을 용이하게 해주는 통합관리 시스템을 구축하게 되었다.

관측자료는 생산된 천문대 각각에 아카이브로 보관하며, 통합시스템에는 관측자료에 관한 정보 즉, MySQL 테이블만 저장한다. 테이블은 관측이 끝난 후 곧 새로운 정보를 추가한다. 따라서 지역천문대와 통합시스템에는 각각 관리자가 있어 역할을 분담하여 업무를 수행한다. 지역천문대의 관리자는 자료의 품질 확인 및 선별, 유용한 파일의 아카이브 저장, 헤더정보의 추출 등을 하고, 통합시스템 관리자는 추출된 헤더정보의 전송과 테이블 생성 및 데이터베이스 입력, 웹 페이지 관리 등의 작업을 한다 추출된 헤더정보를 주고 받는 전송과정은 지역천문대와 통합시스템의 관리자가 공동으로 작업을 수행하게 된다. 이 시스템의 경우 관측자료의 물리적 이동이 없기 때문에 관리가 용이한 장점이 있다.

사용자는 천문대 아카이브 각각을 검색하지 않아도 통합시스템의 웹 검색페이지 한곳에서 모든 천문대의 관측자료를 검색할 수 있다. 또한 자료의 다운로드도 검색결과 페이지에서 가능하다. 다운로드 시에는 통합시스템에서 외부를 통해 지역천문대의 아카이브에 접속해

야 하므로 고속의 네트워크가 요구된다.

통합시스템에서 독립된 각각의 천문대 아카이브로 쿼리를 보내 검색하는 방식은 가상 천문대의 기본개념과 동일하다. 이와 같이 본 연구는 국제가상천문대 구축에 참여하는 한국의 가상천문대 기술수준을 향상시키는 작업의 일환이기도 하다. 또한 통합관리 시스템을 이용함으로써 얻어지는 연구의 효율성 증대는 한국 천문학 발전의 밑거름이 될 것이다.

본 연구에서 획득한 기술은 천문 관측자료 뿐만 아니라 분산 저장된 모든 자료들의 효율적 이용을 위한 통합관리의 일반적인 모형으로 자리 매김 가능할 것이다.

참고문헌

Arviset, C., Dowson, J., Ortiz, I., Parrilla, E., Salgado, J., Zender, J. "ESA Planetary Science Archive" *Astronomical Data Analysis Software and Systems*, XVI (376): 163–171.

Kimula, H. & Takahashi, M. 2002. 오준혁 옮김. 『입문 SQL』. 서울: 영진.COM.

Moran, K. 2007. "Astronomical Archives: Their Historical and Scientific Value and How to Preserve them for the Future." *ASP Conference Series*, 377: 329–340.

Pennington, R. L., Humphreys, R. M.,

Odehahn, S. C., Zumach, W., & Thurmes, P. M. 1993. *Publications of the Astronomical Society of the Pacific*, 105: 521–526.

Wells, D. C., Greisen, E. W., Harten, R. H. 1981. *A&AS*, 44: 363–370.

성현일, 김순욱, 배영호, 최준영. 2006. 천문 이미지 디지털 아카이빙 시스템 개발. 『천문학논총』, 21(1): 1–9.

성현일, 김상철, 남현웅, 김봉규, 임인성, 윤요나. 2003. 보현산천문대 관측자료 Archive 시스템 설계 및 구축. 『천문학논총』, 18(1): 43–49.

조준익. 2000. 『UNIX 서버와 DB구축』. 서울: 미래컴.

부록.

테이블스키마 : 헤더정보 추출 프로그램

```

<%@ include file="../db_conn.jsp"%>
<%
String data = request.getParameter("data") + "\n";
String sql="";
sql = "insert into expos_catalogue (catalog_code, data_type, object,
    ra, decdec, epoch, date_obs_ut, start, filter, exp_time,
    observer, instrument) values ( ?, ?, ?, ?, ?, ?, ?, ? ,
    ?, ?, ?)";
Connection con = DB_Connection();
PreparedStatement pstmt = null;

String catalog_code = "", data_type = "", object = "", ra = "",
    dec = "", date_obs_ut = "", startStr = "", filter = "", exp_time
    = "", oberver = "", instrument = "", epoch = "";

try{
    BufferedReader inBuf = new
        BufferedReader(new FileReader(data File));
    String line;
    int start = 0;
    int end = 0;
}
    
```



```

while((line = inBuf.readLine()) !=null){
try{
start = 0;
end = 13;
catalog_code = Util.replaceStr(line.substring(start, end),
".fits", "").trim();
start = 14;
end = 20;
data_type = line.substring(start, end).trim();
start = 21;
end = 39;
object = line.substring(start, end).trim();
start = 40;
end = 52;
ra = line.substring(start, end).trim();
start = 53;
end = 65;
dec = line.substring(start, end).trim();
start = 66;
end = 81;
epoch = line.substring(start, end).trim();
start = 82;
end = 92;
date_obs_ut = line.substring(start, end).trim();
start = 93;
end = 104;
startStr = line.substring(start, end).trim();
if(startStr.indexOf(".")!=-1){
startStr = startStr.substring(0, startStr.indexOf("."));
}
start = 105;
end = 114;

filter = line.substring(start, end).trim();
start = 115;
end = 125;
exp_time = line.substring(start, end).trim();
if(exp_time.indexOf(".")!=-1){
exp_time = exp_time.substring(0, exp_time.indexOf("."));
}
start = 126;

end = 143;
observer = line.substring(start, end).trim();
if(catalog_code.substring(0,3).equals("ccd")){
instrument = "1KCCD";
}else if(catalog_code.substring(0,2).equals("2k")){
instrument = "2KCCD";
}else if(catalog_code.substring(0,2).equals("sp")){
instrument = "SP";
}

pstmt = con.prepareStatement(sql);
pstmt.setString(1, catalog_code);
pstmt.setString(2, data_type);
pstmt.setString(3, object);
pstmt.setString(4, ra);
pstmt.setString(5, dec);
pstmt.setString(6, epoch);
pstmt.setString(7, date_obs_ut);
pstmt.setString(8, startStr);
pstmt.setString(9, filter);
pstmt.setString(10, exp_time);
pstmt.setString(11, observer);
pstmt.setString(12, instrument);
int cnt = 0;
try{
cnt = pstmt.executeUpdate();
pstmt.close();
}catch(Exception e){
out.println("<B><FONT COLOR=W#FF8080W#>" +
catalog_code + " 중복데이터</FONT></B><br>");
}finally{
if(pstmt!=null) pstmt.close();
}
}catch(Exception e){
out.println("<B><FONT COLOR=W#FF8080W#>" +
catalog_code + " </FONT></B>" + e.toString() + "<br>");
}finally{
if(pstmt!=null) pstmt.close();
}
}
}catch(Exception e) {
out.println("<B><FONT COLOR=W#FF8080W#>" + catalog_code
+ " Insert 실패</FONT></B>" + e.toString() + "<br>");
} finally {
if(pstmt!=null) pstmt.close();
if(con!=null) DB_Disconnection(con);
}
}%>

```