

**BOES 관측데이터의 자동처리 프로그램 개발**  
**DEVELOPMENT OF AN AUTOMATIC PROCESSING PROGRAM FOR BOES DATA**

강동일<sup>1, 2</sup>, 박홍서<sup>1</sup>, 한인우<sup>2</sup>, G. Valyavin<sup>2</sup>, 이병철<sup>2</sup>, 김강민<sup>2</sup>  
<sup>1</sup>한국교원대학교  
<sup>2</sup>천문연구원

DONG-IL KANG<sup>1, 2</sup>, HONG-SUH PARK<sup>1</sup>, IN-WOO HAN<sup>2</sup>, G. VALYAVIN<sup>2</sup>,  
BYEONG-CHEOL LEE<sup>2</sup>, AND KANG-MIN KIM<sup>2</sup>

<sup>1</sup>Korea National University of Education, Kang-Nae, Chung-Buk, 363-791, Korea  
*E-mail: dikang@boao.re.kr*

<sup>2</sup>Korea Astronomy Observatory, Daejeon 305-348, Korea

(Received November 16, 2005; Accepted December 8, 2005)

**ABSTRACT**

We developed a data reduction program (RX) to process BOES data automatically. It processes a whole set of data taken during one night automatically - preprocessing, extraction to one-dimensional spectra and wavelength calibration. The execution is very fast and the performance looks pretty good. We described the performance of this program, comparing its procedure with that of IRAF. RX does not have functions for continuum normalization yet. We will develop those functions in the next works.

*Key words:* BOES — Echelle spectrograph — Automatic data reduction

**1. 서론**

BOES (BOao Echelle Spectrograph)는 한국천문연구원의 보현산천문대 1.8 m 망원경의 부대기기로 개발된 고분산 분광기이다. 2002년 가을에 개발이 완료되어 첫 관측을 시작하였고, 약 1년간의 기기 조정 기간을 거쳐 2003년 가을부터 본격적인 관측을 수행하고 있다. BOES의 개발로 국내에서 고분산 분광 관측 시대가 개막되었다고 할 수 있다. 그동안의 관측 결과 BOES는 2 m급 망원경에 설치된 동급의 분광기로서는 세계적으로 최고의 성능을 가진 것으로 평가된다.

이 논문은 BOES 관측 데이터를 자동으로 처리하는 프로그램 개발에 관한 것이다. 분광데이터 자동 처리 프로그램의 필요성은 몇 가지로 생각할 수 있다. 우선 기기 관리의 측면에서는 관측된 데이터를 신속히 - 거의 실시간으로 - 처리하고 데이터의 질을 조사하여 관측기기의 이상 유무 등을 점검하는데 있다. 다음으로는 관측된 데이터가 매우 많아 수동으로 처리하기에는 시간이 너무 많이 걸리는 경우를 생각할 수 있다. BOES의 경우에도 관측 프로그램에 따라 다량의 데이터가 생성되는 경우가 있는데, 분광선의 빠른 시간 변화를 모니터링 하는 경우가 한 예로서 수백 개의 프레임이 생기기도 한다. 또 다른 예로, 수십 혹은 수백 채널의 다천체 분광기의 경우에는 수동으로 관측데이터를 처리하는 것이 원천적으로 불가능

하고, 자동 처리 프로그램이 필수적으로 요구된다.

자동 데이터처리 프로그램 개발에서 문제가 되는 것은 수동 처리와 비교할 때 과연 얼마나 사용자의 연구 목적을 만족시킬 수 있는냐이다. 아주 초보적인 수준에서는 몇몇 이유, 이를테면 데이터의 S/N 비가 나빠서 자동 프로그램의 수행 자체가 불가능한 경우도 있을 것이다. 프로그램이 상당히 발전하여도 많은 사람들은 자동 프로그램을 신뢰하지 않고, 연구자 스스로 수작업으로 분석하는 것을 선호하는 경우가 많다. 그러나 향후 천문학의 변화 추세는 - 관측 프로그램의 대형화 등의 이유로 - 자동 데이터 처리 프로그램의 사용이 대세가 될 것이다. 이런 추세에 대비해서도 분광 관측 데이터 자동처리 프로그램 개발은 매우 필요하다고 생각된다.

분광 데이터의 자동 처리는 크게 다음의 두 단계로 나누어 생각할 수 있다.

- 1) 2차원 분광 데이터에서 1차원 스펙트럼을 추출
- 2) 1차원 스펙트럼의 정규화

1차원 스펙트럼의 정규화는 천체의 분광형 등에 따라 다른 알고리즘이 요구되는 등 상당히 어려운 문제이다. 특히 에셀 분광 데이터의 경우에는 정규화가 까다롭다. 본 논문에서는 우선 1차원 스펙트럼 추출 프로그램에 대하여 소개하기로 한다. 정규화 프로그램 개발은 향후 연구 과제로서 후속 논문으로 발표할 예정이다.

표 1. 차수별 스펙트럼의 파장 범위

구경 번호	차수 번호	시작 파장(Å)	중심 파장(Å)	끝 파장(Å)	화소당 파장 (Å / pixel)
1	122	3781.536	3825	3858.026	0.018652
2	121	3812.791	3857	3889.909	0.018805
3	120	3844.568	3889	3922.323	0.018960
10	113	4082.754	4130	4165.286	0.020125
20	103	4479.190	4531	4569.668	0.022063
30	93	4960.892	5018	5061.020	0.024415
40	83	5558.681	5623	5670.775	0.027333
50	73	6320.262	6393	6447.593	0.031049
60	63	7323.632	7408	7471.029	0.035942
73	50	9228.062	9334	9413.524	0.045224
74	49	9416.414	9524	9605.639	0.046141
75	48	9612.614	9722	9805.760	0.047097

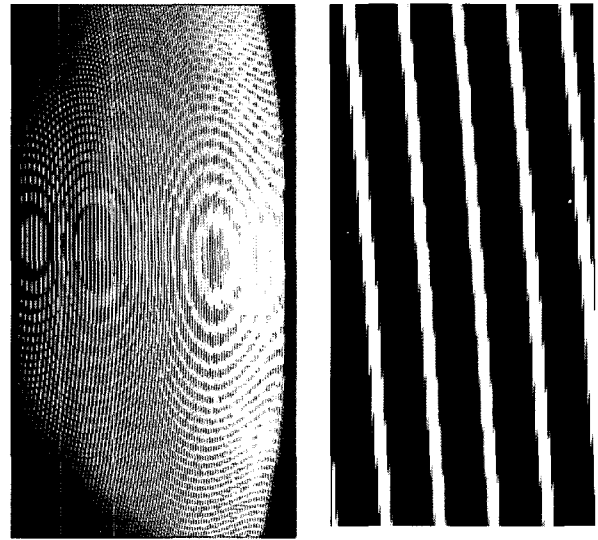


그림 1. BOES 데이터의 모습

그림 2. 그림 1을 확대한 모습

## 2. 에셀 데이터 처리 알고리즘

본 연구를 통해 개발한 프로그램의 이름은 Automatic Reducing Program for Echelle Spectroscopy의 첫 자를 따와서 RX[ar-ecs]로 간략히 하였다. 이번 절에서는 RX를 이용한 데이터 처리 알고리즘을 소개한다.

RX가 수행하는 관측 데이터의 전처리(pre-processing)에 대한 알고리즘은 Massey(1997)를, 에셀 관측 데이터에 대한 알고리즘은 Willmarth & Barnes(1994)와 Churchill(1995) 등을 참고하였으며 필요에 따라 데이터 처리 과정을 적절히 수정 및 재구성하였다. 따라서 이번 절에서 제시하는 데이터 처리 알고리즘은 에셀 분광 관측 데이터의 전형적인 처리 과정은 아니고 본 연구의 개발 과정에서 최적화된 알고리즘이다.

### 2.1. 에셀 분광 데이터

그림 1은 BOES로 관측한 전형적인 CCD 프레임이다. 그림에서 세로 방향으로 활처럼 구부러진 선(혹은 띠)이 여러 개 배열되어 있는 것을 알 수 있다. 그림 2는 이 그림 일부를 확대한 것이다. 에셀 분광 데이터에서 각각의 구부러진 선은 보통 분광기에서 나오는 하나의 스펙트럼으로 생각할 수 있는데, 이런 관점에서 에셀 분광기는 여러 개의 스펙트럼을 동시에 기록하는 분광기라고 생각할 수 있다. 에셀 데이터에서 각각의 구부러진 스펙트럼을 분광 차수(spectrum order)라 부르는데 이것은 분광기가 기록하는 고유의 스펙트럼 번호이다. BOES의 경우에는 131개의 분광 차수를 갖는다(김강민 2003). 그림 1

에서 스펙트럼의 방향을 분산방향이라고 한다. 가로방향은 차수 방향(혹은 구경 방향)으로 오른쪽으로 갈수록 차수가 감소하고 파장은 증가한다. 하지만 실제로 CCD에 기록되어 우리가 사용하는 차수는 약 85개 정도이다. 그리고 단파장 영역의 9개 차수는 끝 부분이, 장파장 영역의 1개 차수는 중간 부분이 잘리기 때문에 대개는 이들을 제외한 75개 정도의 차수만을 이용한다. 우리가 사용하게 되는 75개 차수에 대해 파장이 증가하는 방향으로 일련 번호를 매기고 이것을 구경(aperture) 번호라 부른다.

에셀 분광기에서 각 구경에 해당하는 파장 범위는 구경마다 다르다. 구경번호가 증가할수록 장파장이 기록된다. 표 1에 BOES의 각 구경에 해당하는 파장의 범위를 일부 나타내었다. 표 1에서 볼 수 있듯이 인접한 구경의 파장 범위는 끝 부분에서 상당 부분 겹치는 것이 에셀 분광기의 특성이다. BOES의 경우에는 구경 번호가 작은 쪽에서는 절반 이상이 겹치지만 구경 번호가 커지면서 겹치는 부분이 점점 줄어들다가 마지막 두 구경 간에는 대략 6Å 정도의 간격이 생긴다. 그림 3에서 각 구경별 파장의 범위를 시각적으로 확인할 수 있다.

에셀 분광 데이터의 특징은 한 프레임에 많은 스펙트럼이 동시에 기록되고, 스펙트럼의 형태가 구부러져 있다는 점이다. 따라서 같은 스펙트럼 하나만 기록되는 보통 분광 데이터에 비하여 데이터 처리 과정이 훨씬 복잡하고 어렵다.

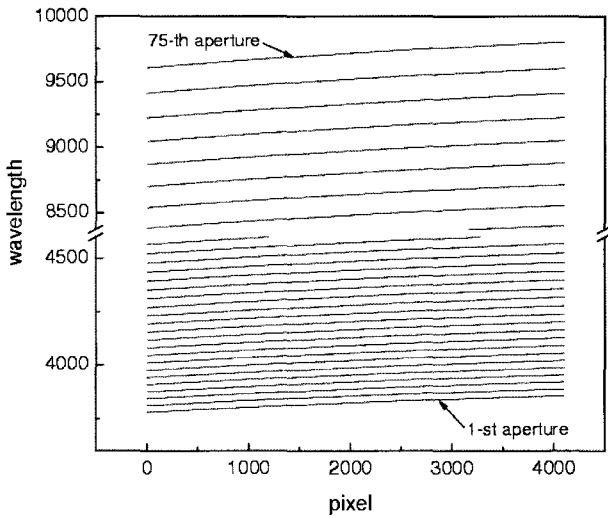


그림 3. BOES 구경별 파장 범위

2.2. 관측 데이터의 종류

BOES 관측 데이터는 다음의 4 종류가 있다.

2.2.1. 바이어스(bias) 프레임

CCD 칩의 고유한 전기적 특성을 보정하는데 필요하다. 즉 셔터를 닫고 노출시간을 0으로 하면, CCD 칩에서 나오는 신호는 그 값이 0이어야 하지만 고유한 전기적 특성에 의하여 일정한 신호가 나온다. 이를 바이어스라고 한다. 바이어스 데이터는 몇 개의 프레임을 얻어 중간 값(median)을 계산하여 보정에 이용한다.

2.2.2. 플랫(flat) 프레임

CCD 화소의 감도 차이와 간섭 패턴 등을 보정하기 위하여 필요하다. 플랫 프레임은 흡수선이나 방출선이 없는 Tungsten-Halogen 램프를 광원으로 하여 얻는다. 플랫 데이터는 충분한 S/N 비를 얻을 수 있도록 노출 시간을 적절히 주어야 하고, 통계적으로 상호 비교할 수 있도록 충분한 수의 프레임을 얻어야 한다. BOES의 경우에는 대개 하루 밤 관측에 10 개 정도의 플랫 데이터를 얻는다.

2.2.3. ARC 프레임

관측된 스펙트럼의 파장 동정에 이용된다. 관측 파장 범위 내에서 충분히 많은 방출선을 내는 광원을 이용하는 데, BOES의 경우에는 Thorium-Argon 광원을 이용한다.

2.2.4. 천체 프레임

연구 대상이 되는 천체로부터 얻는 프레임이다.

파장동정이 필요한 연구에서는 위에서 제시한 프레임

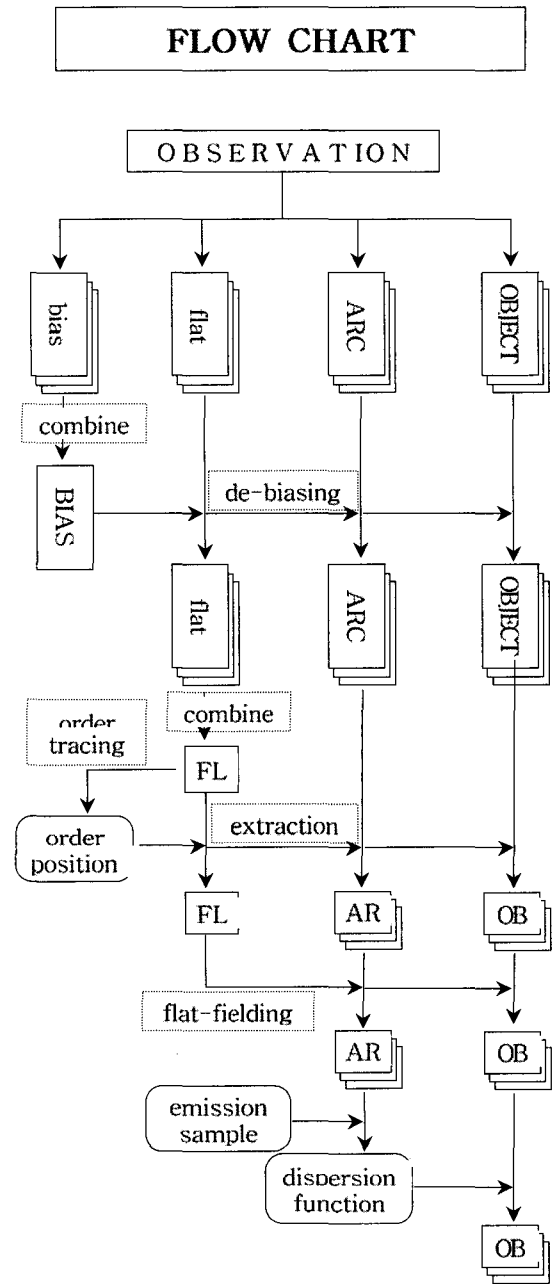


그림 4. 자료처리의 순서도

외에도 암잡음 (dark) 이나 비교성(comparison star)에 대한 프레임을 얻기도 한다. 이렇게 얻은 4 종류의 분광 관측 데이터는 대개 그림 4와 같은 순서도를 따라 처리되어 1차원 스펙트럼을 뽑아내고 파장을 동정하게 된다. 다음에 그림 4의 각 단계에 대하여 자세히 설명하기로 한다.

2.3. 마름질(trimming)

BOES는 50 화소 폭의 prescan 영역과 102 화소 폭의 overscan 영역을 두고 있는데 이 영역에 대해서 마름질을 한다. 경우에 따라서는 데이터의 질이 좋지 않은 분산 방향의 양끝 부분에 대해서도 어느 정도 마름질을 하기도 하지만 이는 사용자가 선택해야 할 문제이다. 마름질은 획득한 모든 프레임에 대해 실시한다.

보통 마름질 이전에 나쁜 화소 보정(bad pixel correction)이나 overscan 보정 (Petr 2002)을 하지만 본 연구에서는 생략하였다. 이런 기능은 차후 보강할 예정이다.

2.4. 플랫, 바이어스 프레임 합치기

(마스터 플랫, 마스터 바이어스 프레임 만들기)  
앞에서도 언급하였듯이 통계적 신뢰성을 높이기 위해서 가능한 한 많은 수의 바이어스와 플랫 프레임이 필요한데, BOES의 경우 보통 10장 정도를 얻는다. 이 프레임들은 다음 순서 (Gazinur 2005)를 통해 하나의 마스터 프레임으로 합쳐진다.

- 1) 전체 프레임을 3 개씩 묶어 중앙치(median)로 평균한다. 3개씩 묶은 후 나머지가 있을 경우에는 마지막 묶음에 포함시켜 평균한다.
- 2) 중앙치로 평균된 프레임들을 산술평균 (average)하여 하나의 마스터 프레임을 만든다.

이 과정은 바이어스, 플랫 프레임에 동일하게 적용되되 플랫 프레임의 경우에는 바이어스 보정 후에 실시한다.

2.5. 바이어스 보정 (de-biasing)

앞에서 얻은 마스터 바이어스를 이용한 바이어스 보정은 플랫, ARC, 천체프레임 등 모든 데이터에 대해 실시한다. 이것은 단순히 각 프레임에서 마스터 바이어스 프레임을 빼주면 된다.

2.6. 구경의 위치 확인 (aperture tracing)

2차원으로 기록된 CCD 프레임에서 각 구경의 위치를 결정하는 과정이다. BOES의 경우, 분광기가 매우 안정되어 있으므로 하루에 관측된 모든 프레임에서 각 구경의 위치는 거의 동일하다. 따라서 구경의 위치 결정은 S/N 비가 좋은 마스터 플랫 프레임을 이용하여 얻은 위치 정보를 모든 프레임에 적용한다.

구경의 위치는 CCD 프레임의 중앙에서 구경이 증가하는 방향으로 프레임을 자른 단면 프로파일을 구해서 확인

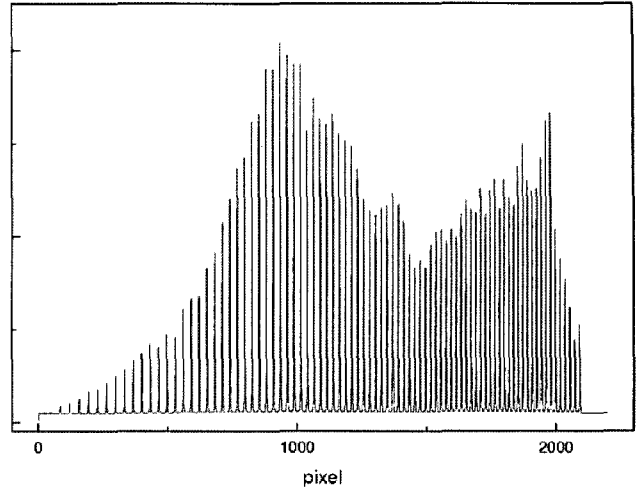


그림 5. 플랫 데이터 중심의 단면 프로파일

할 수 있다. 그림 5 에서 보이는 85개의 봉우리들이 단면 프로파일에서의 구경에 해당하고 그 정점들이 각 구경의 중심점이라고 생각할 수 있다. 2차원 프레임에서 구경은 활처럼 휘어져 나타나므로 정점에서 분산 방향으로 이동하면서 인접한 정점을 찾아 이어나가면 구경의 중심선을 구할 수 있다. 이 과정은 다음과 같이 정리할 수 있다.

- 1) 마스터 플랫의 중심에서 구경 방향의 단면 프로파일을 구한다 (그림 5).
- 2) 이 프로파일에서 화소  $x$ 의 값  $I(x)$ 을 다음 식 (1)로 이동 평균 (moving average)하여 전체적으로 부드럽게 만든다.

$$I(x_n)' = \frac{1}{m} \sum_{k=0}^{m-1} I(x_{n-k}) \quad (1)$$

- 3) 입력 파라미터의 차수 검색 범위 (표 2 참조) 내에서 각 차수별 중심 좌표를 결정한다. 어떤 화소  $x$ 에서 단면 프로파일의 세기를  $I(x)$  라고 하면, 다음의 두 조건을 동시에 만족하는 화소가 중심 좌표가 된다.

- ①  $I(x-3) < I(x-2) < I(x-1) < I(x) > I(x+1) > I(x+2) > I(x+3)$
- ②  $I(x) > \text{문턱값}$

여기서 문턱값은 단면 프로파일에서 최대값의 1% 정도로 하였다.

- 4) BOES 프레임은 장시간에 걸쳐서 아주 서서히 그 위치가 변해서 우리가 원하는 차수가 검색 범위를 넘어서거나 혹은 차수간의 간격이 달라져 사용하려는 차수의

개수보다 많거나 적은 차수가 범위 내에서 확인되는 경우가 있다. 이 경우에는 우리가 원하는 차수만을 선택하기가 어려워진다. 이 문제를 해결하기 위해 다음과 같은 방법으로 차수의 절대위치를 확인한다.

- ① 2.8 절에서 설명할 1차원 추출 알고리즘을 이용하여 ARC 프레임의 1차원으로 추출한다.
  - ② 보다 넓게 정한 차수 검색 범위에서 75개 이상의 구경을 확인하고 이 중에서 비교적 방출선이 많은 46번째 구경을 길잡이 구경 (guide aperture)으로 정한다. 아직은 이것이 46번째 구경인지 알 수 없다.
  - ③ 2.10절에서 설명할 방출선의 위치 정보가 들어 있는 파일을 읽어 온다. 이 파일에는 우리가 사용하려는 75개 구경에 있는 방출선의 위치와 파장의 정보가 들어 있다.
  - ④ 45, 46, 47번째 구경에서의 방출선 위치 정보를 길잡이 구경에 각각 중첩하여 방출선이 서로 일치하는 비율을 구한다. 서로 일치하는 경우에는 일치율이 대개 0.9를 초과하고 일치하지 않는 경우에는 0.2를 넘지 않기 때문에 일치 여부를 쉽게 확인할 수 있다. 이렇게 하여 길잡이 구경의 실제 구경번호를 확인하게 된다.
- 5) 프레임의 중심 단면에서 각 구경별 중심 좌표들이 결정되었으므로 이것을 이용하여 전 영역의 중심 좌표를 다음과 같이 구한다.

- ① 중심으로부터 분산방향으로 한 화소만큼 이동한다.
- ② 이동한 화소와 이 화소의 구경 방향 (혹은 한 구경 내에서는 공간 방향이라고도 한다.) 좌우 화소를 포함한 3화소 중 큰 값을 중심 좌표로 정한다.
- ③ 위 과정을 구경의 끝까지 진행하면 전체 구경에서의 중심 위치를 결정할 수 있다. 그림 6의 밝은 부분이 구경이 기록된 부분이고 중심이 검은 선으로 표시되어 있다.

플랫 프레임으로부터 구한 스펙트럼의 위치 정보는 전체 프레임과 ARC 프레임의 1차원 추출 시 그대로 이용된다.

### 2.7. 배경 값 제거

스펙트럼이 CCD에 기록될 때 스펙트럼 값뿐만 아니라 배경 (background) 값도 같이 기록 된다 (그림 7). 구경을 1차원으로 추출하기 이전에 이 배경 값을 먼저 제거해야 한다. 배경 값의 결정 및 제거는 다음 순서를 따른다.

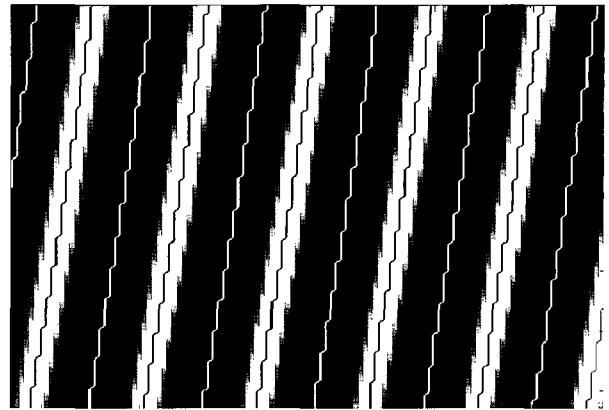


그림 6. 밝은 부분이 스펙트럼이 기록된 부분이고, 어두운 부분이 배경이다. 스펙트럼의 중심을 검은 선으로 표시하였고, 배경 선을 흰 선으로 표시하였다.

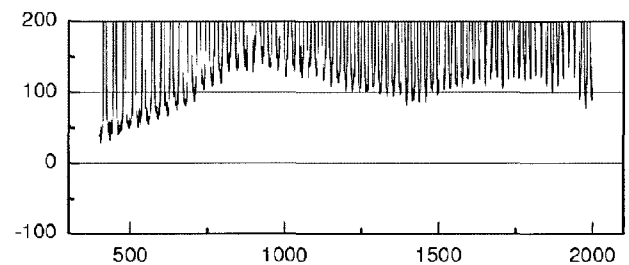


그림 7. 배경 값을 가지고 있는 CCD의 단면

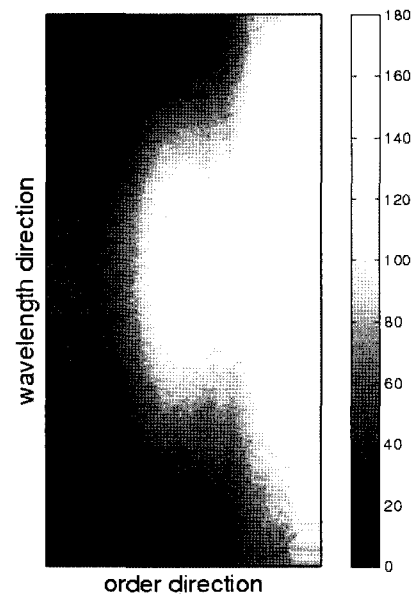


그림 8. 2차원으로 보이는 배경 값

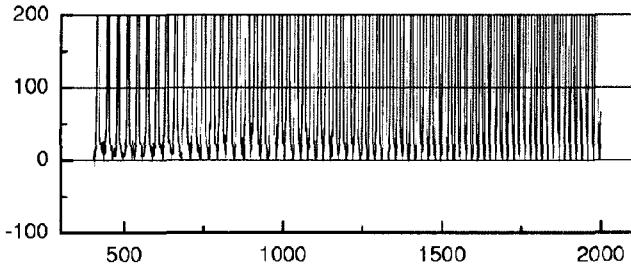


그림 9. 배경값이 제거된 CCD의 단면

1) 인접한 두 구경의 중심을 따라 배경선을 그린다. 첫 번째 구경과 마지막 구경의 바깥쪽에도 배경선을 그린다 (그림 6).

2) 배경선을 이용해서 배경이 그리는 함수를 구한다. 배경선은 분산 방향을 따라 위치하지만 함수는 구경 방향으로 결정한다. 75개 구경에 대해 76개 배경선이 있으므로 이 배경선의 값을 이용하여 구경 방향 배경 함수를 구한다. 함수는 스플라인 (spline)이나 선형 내삽 (linear interpolation)으로 할 수 있는데, 두 방법의 차이는 거의 없었다. 이러한 배경 함수를 구경방향으로 모두 결정하면 전체적으로 2차원 배경 프레임의 모습으로 보인다 (그림 8).

3) 배경 프레임이 구해지면 이를 원래 프레임에서 빼기 연산하여 CCD에 기록된 배경값을 모두 제거 한다 (그림 9).

2.8. 1차원 추출

2차원 프레임으로부터 1차원 스펙트럼 추출은 2.6절에서 구한 구경의 중심에서 공간방향을 따라 적분 구간 내 화소 값을 더하면 된다. 적분구간은 스펙트럼이 퍼져나가서 거의 바닥 값에 이르는 범위까지로 충분히 설정한다. 본 연구에서는 적분 구간 내의 화소 값들을 단순히 더했지만 통계적으로 보다 의미 있는 결과를 얻기 위해서는 가중치 적분 (weighted integration)이나 적절한 적분 구간을 결정하는 수학적 방법에 대한 추후 연구가 필요하다.

2차원 프레임을 1차원 스펙트럼으로 추출하면 4102×2048 이던 크기가 4102×75로 줄어든다 (그림 10).

2.9. 바닥 고르기(flat-fielding)

바닥고르기는 CCD가 갖는 화소간의 반응율을 보정하기 위한 것인데, 측광 데이터의 경우처럼 화소 단위로 보정하지는 않고, 1차원으로 스펙트럼을 추출할 때 적분한 구간에 위치한 화소의 평균 반응율에 대한 보정을 한다. 다음 절차를 따른다.

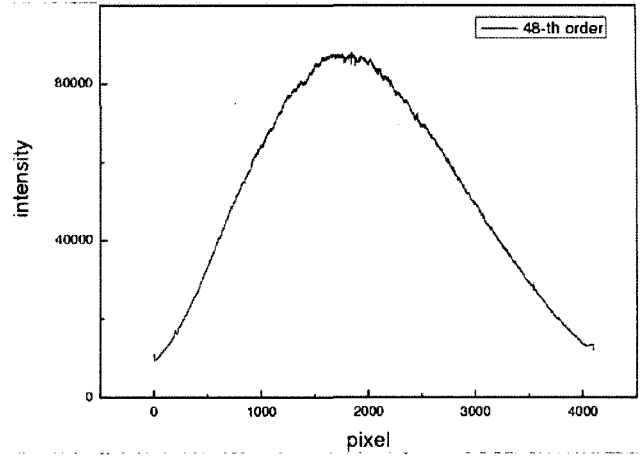


그림 10. 1차원으로 추출된 스펙트럼의 모습

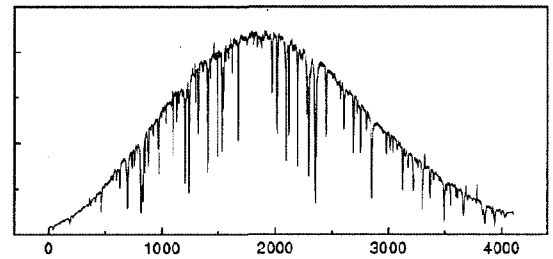


그림 11. 바닥고르기 이전의 스펙트럼

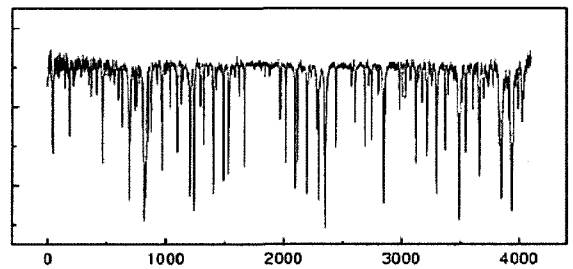


그림 12. 바닥고르기로 평평해진 스펙트럼

- 1) 플랫 프레임 각 구경에서의 최대값으로 해당 구경 전체를 나누어 정규화(normalize) 한다.
- 2) 정규화 된 플랫 프레임으로 ARC 프레임과 천체 프레임을 나누기 연산하면 바닥고르기가 완료된다.

그림 11은 바닥 고르기 이전의 스펙트럼이다. 그림에서 보는 바와 같이 스펙트럼은 블레이징 함수 (blazing

function)를 보이고 있지만 바닥 고르기 과정을 마친 그림 12에서 스펙트럼은 화소간 반응율과 블레이징 함수의 효과가 보정되어 평탄해진(flatten) 모습을 보여 준다.

분광데이터 처리의 개념은 각 구경을 1차원으로 추출하여 파장을 동정하는 것이다. 여기까지의 과정으로 2차원 데이터를 1차원으로 추출하여 바닥고르기를 마쳤으므로 이것으로 다음 단계에서 파장을 동정하기 위한 준비가 모두 끝난 셈이다.

2.10. 파장 동정 및 분산 함수 결정

분광 데이터를 수동으로 파장을 동정하기 위해서는 기존에 제작된 방출선 그림 (emission line map)을 보고 육안으로 방출선을 하나씩 확인하여 전체 파장을 동정한다. 이것은 데이터 처리에서 가장 많은 시간을 차지하며, 상당한 경험을 필요로 하는 과정이기도 하다. 이 과정을 자동으로 처리하기 위해서는 다음과 같은 새로운 접근 방법이 필요하다.

분광기는 기기 자체의 큰 변동이 없으면 시간이 지나도 CCD의 거의 같은 위치에 스펙트럼을 기록한다. 따라서 ARC 프레임에서 방출선의 위치를 미리 알고 있다면 이후로 획득한 ARC 프레임에서의 방출선도 거의 같은 위치에서 찾을 수 있다. 방출선의 위치 정보는 수동으로 위치와 파장을 찾아 기록한 파일을 이용한다. 우리가 파장 동정하려는 ARC 프레임에 이 방출선 위치 정보를 겹치면 거의 비슷한 위치에서 대부분의 방출선을 찾을 수 있다. 파장 동정의 세부 과정은 다음과 같다.

- 1) 방출선 좌표가 기록된 파일 (EmissionModel.mat)을 읽어 온다.
- 2) 읽어 온 방출선의 좌표  $x$ 로부터 적정 범위 (7 화소) 이내에 있는 어떤 화소  $x'$ 에서의 값을  $I(x')$ 이라 할 때, 다음의 조건을 만족하는 새로운 방출선의 좌표  $x'$ 을 찾는다.

$$I(x'-3) < I(x'-2) < I(x'-1) < I(x') > I(x'+1) > I(x'+2) > I(x'+3) \quad (2)$$

위 조건을 만족하는  $x'$ 을 찾으면 우리는 이것이 파장  $\lambda$ 에 해당하는 방출선이라고 생각한다.

- 3) 앞에서 찾은 방출선의 중심을 구하기 위하여 다음과 같은 방법으로 방출선의 경계  $x_1, x_2$ 를 결정한다.

- ①  $x'$ 으로부터 멀어지면서  $I(x')$ 가 감소하다가 다시 증가하기 시작하거나 0보다 작아지기 시작하는 지점
- ② ①을 만족하는 경우가 없을 때에는  $x' \pm 5$ 인 지점

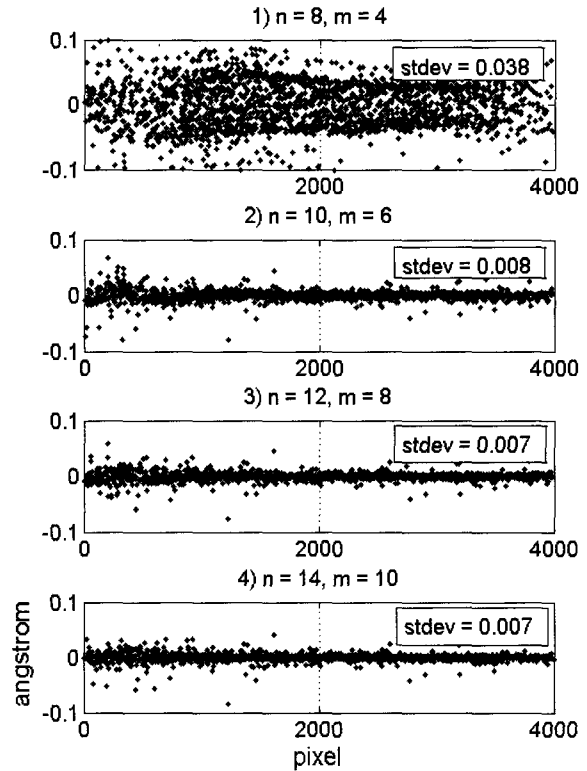


그림 13. 식 2 로 계산한 방출선의 좌표와 이것을 식 3을 이용해 결정된 분산함수로 다시 계산한 방출선 좌표와의 잔차를 나타낸 그래프

- 4) 방출선의 경계  $x_1, x_2$ 로부터 방출선의 중심  $x_c$ 를 다음 공식으로 결정한다.

$$x_c = \frac{\sum_{i=x_1}^{x_2} (I(i) \times i)}{\sum_{i=x_1}^{x_2} I(i)} \quad (3)$$

- 5) 화소와 파장과의 관계를 나타내는 분산 함수 (dispersion function)를 결정한다. 화소와 파장의 관계가 2차원이라면 분산 함수도 2차원으로 표현할 수 있다. 대개 다음과 같은 다항식을 이용한다.

$$W(j, x) = \sum_{j=1}^m \sum_{i=1}^n a_{ij} C_i(x) C_j(j) \quad (4)$$

- $n$  : 분산 방향 다항식의 차수
- $m$  : 스펙트럼 차수 방향 다항식의 차수

위에서 함수  $C$ 로는 Chebyshev 혹은 Legendre 함수를 많이 사용한다. 분산 함수를 결정한다는 것은 식 3에서  $a_{ij}$ 를 구하는 것인데, 이미 구한 방출선의 중심 위치와 파장을 식 3에 대입하여 최소자승법으로 결정한다.

그림 13은 2차원 함수 맞추기에서 다항식의 차수에 따른 결과를 비교한 그림이다. 그림의 가로축은 화소이고, 세로축은 방출선의 파장과 2차원 Chebyshev 다항식으로 다시 계산된 파장과의 차이 값으로 단위는 Å이다. 스펙트럼의 분산방향 다항식의 차수  $n$ 과, 스펙트럼의 구경방향 다항식의 차수  $m$ 을 바꾸어 가면서 결과를 비교하였다. 일반적으로 다항식의 차수가 증가할수록 오차는 감소하지만 지나치게 큰 값은 통계적으로 의미가 없고 인위적인 패턴을 만들 수 있으므로 피해야 한다. 여기에서는 분산 방향으로 12차 정도, 구경 방향으로 8차 정도가 적당하다.

### 2.11. 우주선(cosmic-ray) 제거

우주선은 그 형태와 세기가 매우 다양하기 때문에 제거하기가 상당히 까다롭다. 여러 가지 방법을 사용해 본 결과 현재까지는 2차원 중앙치 필터(2-dimensional median filtering)가 가장 나은 결과를 보여주었다. 우주선 제거와 관련하여서는 보다 세련되고 수준 높은 알고리즘의 개발이 필요하다. 다음 순서로 우주선을 제거하였다.

- 1) 1차원으로 추출된 스펙트럼이지만 스펙트럼이 다수(75개)의 구경으로 이루어져 있으므로 데이터는  $4102 \times 75$  크기의 2차원 행렬로 취급할 수 있다. 이 행렬을  $A$ 라 하고 이것에 2차원 중앙치 필터링(하영호 외 2003)을 하여 새로운 행렬  $B$ 를 얻는다.
- 2)  $A$ 와  $B$ 의 잔차  $R$ 을 구하고  $R$ 의 각 구경별 표준편차  $\sigma_j$ 를 구한다. ( $j$ : 구경 번호)
- 3)  $R_{ij} > a\sigma_j$ 를 만족하면 우주선으로 판단한다. 경험적으로  $a$  값은 5 정도가 적절하였다. ( $j$ : 분산방향 화소)
- 4) 우주선의 경계는 방출선의 경계 결정(2.10절)과 같은 방식으로 구한다.
- 5) 우주선의 경계 내에 있는 모든 값은 경계점으로 선형 내삽(linear interpolation) 하여 새로 결정한다.

### 2.12. 천체 스펙트럼의 처리와 저장

2.6절에서 2차원 프레임을 1차원으로 추출하기 위한 스펙트럼의 위치 정보를 구하였고, 2.10절에서 구경별 스펙트럼의 파장 정보를 구하였다. 이것으로 천체 프레임을 처리하기 위한 모든 준비가 끝난 셈이다.

다음의 순서대로 천체 프레임을 처리한다.

- 1) 스펙트럼의 위치 정보를 이용하여 배경 값을 제거하고 1차원 스펙트럼을 추출한다.
- 2) 분산함수의 정보를 규격화된 형식으로 바꾸어서 파일 헤더를 작성한다.
- 3) 파일 헤더와 1차원으로 추출된 데이터를 FITS 파일 형식에 맞추어서 저장한다. 이를테면 천체프레임 'es16202.fits'에 대해 다음의 두 파일이 생성된다.
  - eces16202.fits : 1차원 추출과 파장 동정을 마친 파일
  - ceces16202.fits : eces16202.fits 를 연속선에 대해 정규화 한 파일

## 3. 프로그램 설명

### 3.1. 개발도구

본 연구의 개발도구로는 『MathWorks』사의 MatLab ver 7.0.4 Service Pack 2 를 사용하였다. MatLab은 매우 강력한 대화형 프로그래밍 언어(조한근 2004)로서 다음과 같은 몇 가지 장점 때문에 개발 도구로 채택하였다.

첫째, 언어를 배우기 쉽다. 기본 데이터 형으로 행렬을 사용하므로 데이터 표현과 문법이 간단하고 직관적이어서 다른 언어를 사용하던 사용자는 물론, 초보자도 쉽게 익힐 수 있다.

둘째, 고속의 데이터처리 속도이다. MatLab은 다차원의 행렬 연산을 반복 연산 없이 직접 수행할 수 있으며, 반복 연산도 데이터의 벡터화(vectorization)를 통해 벡터 간 연산을 하면 매우 고속으로 처리한다. 이러한 MatLab의 데이터처리 속도는 C나 Fortran에 뒤지지 않는다.

셋째, 풍부한 라이브러리이다. 알고리즘 개발, 수치모델링, 과학과 공학적인 그래픽, GUI(graphical user interface) 개발 등과 관련한 매우 풍부한 함수와 라이브러리를 제공하여 사용자의 아이디어를 쉽게 구현하도록 돕는다.

### 3.2. RX를 이용한 데이터 처리 과정

#### 3.2.1. 데이터 준비

분광 관측 데이터를 처리하기 위하여 관측자는 천체 프레임 외에 여러 개의 바이어스 프레임, 플랫 프레임, ARC 프레임 등을 획득한다. 데이터 처리는 보통 하루에 얻은 데이터를 기본 단위로 한다. 물론 며칠에 걸쳐 더욱 많은 프레임을 획득하여 한꺼번에 처리할 수도 있다. 어떤 방식이든 RX를 이용하여 데이터를 처리하려면 모든 데이터



파일을 한 폴더에 저장하여야 한다.

3.2.2. 입력 파라미터 결정

이 프로그램은 GUI(graphical user interface)를 통한 대화형 프로그램이 아니라 모든 입력 파라미터(input parameters)를 미리 입력해 놓고 데이터가 저장된 폴더를 지정하면 데이터 처리의 전 과정을 일괄 처리 방식으로 수행한다.

파라미터는 프로그램이 BOES 데이터를 처리하기 위해 필요한 BOES 프레임의 특성 정보를 가지고 있다. 이러한 특성 값들은 대부분 그 변화가 미약하여서 시스템의 큰 변화가 없으면 개발도중에 최적화된 값들을 사용하여도 무방하다고 판단된다.

파라미터는 SetPara 함수 안에서 미리 정해진 값으로 결정된다. 사용자가 이 값을 바꾸기 위해서는 SetPara.m 파일을 열어서 수정할 수 있다. 표 2에 입력 파라미터들에 대해 설명해 놓았다.

3.2.3. 프로그램의 실행

프로그램의 실행은 다음의 순서를 따른다.

- ① 실행 명령어 'RX'를 명령창에 입력하면 프로그램이 진행된다.
- ② 데이터가 저장된 폴더를 묻는 폴더 선택 창에서 해당 폴더를 지정해 준다. 프로그램은 지정된 폴더 내의 확장자가 .FITS 인 모든 파일의 헤더를 읽어서 관측 데이터를 천체, 바이어스, 플랫, ARC 등으로 분류하고, 프레임의 크기와 종류, 광섬유의 번호 등을 실행 창에 보여주고 이상 여부를 묻는다.
- ③ 이상이 없음을 확인한 후 'Y' 키 혹은 '엔터' 키를 누르면 데이터 처리를 실시한다.
- ④ 프로그램은 이후 모든 데이터 처리 과정을 스스로 진행하여 결과물인 1차원 추출 후 파장동정이 종료된 천체 프레임의 데이터를 FITS 파일 형식으로 해당 폴더에 저장하고 끝난다.

4. 프로그램의 성능

이 프로그램의 성능을 첫째 프로그램 수행상의 안정성, 둘째 1차원으로 추출된 스펙트럼의 질, 셋째 파장 동정의 정밀성, 마지막으로 자동화 프로그램으로서 갖는 처리 속도 면에서 IRAF를 이용한 수동처리 결과와 비교하여 보았다.

표 2. 입력 파라미터

파라미터	값	의미
Path	C:\	데이터가 저장된 폴더
nFiber	1	광섬유 번호
naxis1	2200	CCD 가로방향 화소수
naxis2	4102	CCD 세로방향 화소수
Xprescan	50	CCD 가로방향의 prescan 수
Xoverscan	102	CCD 가로방향의 overscan 수
Yprescan	0	CCD 세로방향의 prescan 수
Yoverscan	0	CCD 세로방향의 overscan 수
nX	2048	사용할 CCD의 가로방향 수
nY	4102	사용할 CCD의 세로방향 수
nOrder	75	사용할 스펙트럼 차수의 개수
PeakGapMin	15	스펙트럼간의 최소 간격
PeakGapMax	40	스펙트럼간의 최대 간격
CheckPrimePeaks	0	prime peaks의 결정 여부
Peak.Start	350	스펙트럼의 검색의 시작점
Peak.Threshold	200	스펙트럼 확인을 위한 문턱값
ExtAperture	5	1차원 추출을 위한 적분 범위
EmissionRange	3	방출선을 검색하는 범위
EmissionHeight	0.005	방출선을 확인하는 문턱값
Polytype	1	분산함수에 사용할 함수
FittingOrder.x	12	분산함수의 화소방향 차수
FittingOrder.y	8	분산함수의 스펙트럼방향 차수

4.1. 안정성

관측으로 얻은 원시 데이터를 다루다 보면 프로그램의 개발 과정에서 미처 인지하지 못한 문제로 인해 오류가 발생하여 프로그램이 멈추는 일이 많다. 프로그램의 진행에 영향을 미치는 대표적인 요인으로는 다음과 같은 것을 예로 들 수 있다.

4.1.1. 데이터의 크기 차이

관측을 하다보면 크기가 다른 프레임을 획득하게 되는 경우가 있다. 서로 다른 크기의 프레임은 행렬 연산 시 오류를 일으킨다. 프로그램을 시작할 때 이 부분을 확인하고 진행해야 한다.

4.1.2. 구경의 위치 변화

광섬유가 달라지거나 시간이 지남에 따라 구경의 위치가 변할 수 있다. 이럴 경우 구경의 위치를 잘못 결정할 수 있는데, 2.6절에서의 길잡이 구경을 이용한 방법으로 현재는 BOES가 보이는 이러한 구경의 위치 변화에 적절히 대처하도록 하였다.

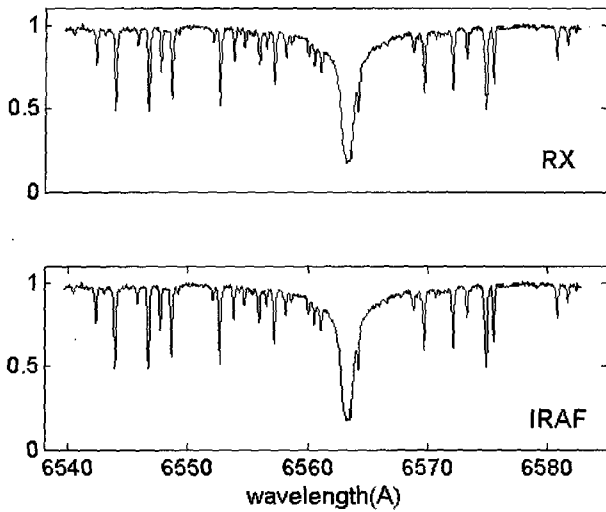


그림. 14 RX와 IRAF으로 처리한 스펙트럼 비교

#### 4.1.3. 0값의 삽입

가끔 CCD를 읽어 올 때, 원래는 없던 0의 값을 가진 화소가 끼어 들어서 그 이후로 읽어 오는 화소의 위치가 1화소씩 밀려나는 현상이 발견된다. 이런 현상이 항상 같은 위치에서 발생하는 것이 아니므로 자료처리 과정에서 이런 현상에 대비해야 한다.

그 외에도 다량의 BOES 관측 데이터를 통한 수차례 시험으로 간혹 발생하는 오류에 대처하도록 프로그램이 최적화 되었으며 현재는 안정성이 상당히 확보되었다.

#### 4.2. 스펙트럼의 질

2005년 6월 24일 Ceres의 관측 데이터를 RX와 IRAF로 각각 처리하였다 (그림 14).  $H_{\alpha}$ 선 근처 연속선 레벨에서의 표준편차  $C\sigma = 0.037$ 에 대해 두 스펙트럼의 차이 값이 보이는 표준편차는  $R\sigma = 0.017$ 을 보였다. 이것은 RX를 이용한 1차원 추출이 S/N 비 내에서 충분히 신뢰할 수 있음을 의미한다.

#### 4.3. 파장 동정의 정밀성

파장 동정의 결과는 IRAF와 RX에서 공통적으로 확인된 방출선의 중심 좌표를 비교하였다. 그림 15의 <RX - IRAF>는 같은 방출선을 RX와 IRAF로 계산한 좌표의 간차를 나타낸 것이다. 이것은 서로 다른 1차원 추출 알고리즘과 파장동정 알고리즘의 결과이다. 그림에서 보듯이 CCD의 화소간 경향성이 나타나고 어느 정도 분산을 보인다.

<RX-IRX>는 RX와 IRX (IRAF에서 1차원 추출을 하고

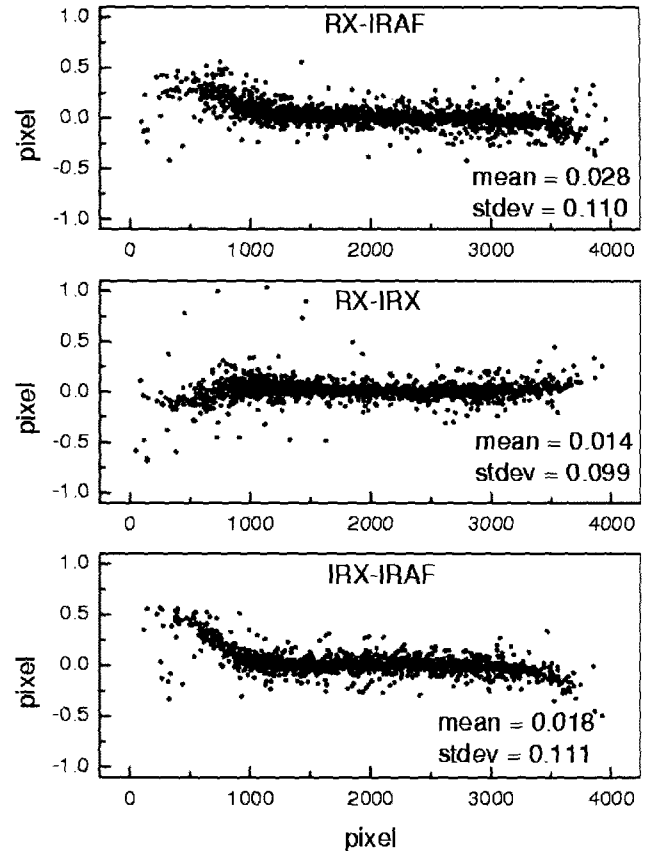


그림 15. RX와 IRAF의 파장동정 결과 비교. 가로축은 분산방향이고 세로축은 두 결과의 잔차이다.

RX로 파장 동정)와의 간차를 나타낸 것이다. 즉, 서로 다른 추출 알고리즘, 같은 파장 동정 알고리즘의 결과이다. 그림에서 보듯이 양 끝에서 나타났던 경향성은 보이지 않는다.

<IRX - IRAF>는 IRX와 IRAF와의 간차를 나타낸 것이다. 즉, 같은 추출 알고리즘, 다른 파장동정 알고리즘의 결과이다. <RX-IRAF>에서 보이던 CCD 화소 간의 경향성이 다시 나타났다.

이것으로 보아 파장 동정에 미치는 1차원 추출 알고리즘 (extraction algorithm)의 영향 보다는 파장 동정을 위한 방출선의 중심을 결정하는 알고리즘 (centroid algorithm)이 중요한 역할을 하는 것을 알 수 있다.

#### 4.4. 처리 속도

PC (Pentium 4, 3.0 GHz CPU)로 2005년 1월 21일부터 24일까지 관측된 바이어스 프레임 28개, 플랫폼 프레임 12개, ARC 프레임 15개와 19개의 천체 프레임을 처리하

는 데에 대략 15분 정도가 소요되었다. 구경의 위치와 각 구경에서의 방출선 위치 정보를 획득한 상태에서 천체 프레임 하나를 처리하는 데에는 십 수 초 정도가 걸리므로 데이터 처리는 거의 실시간으로 수행되어 BOES 관측 데이터를 처리하는데 있어 속도는 전혀 문제가 되지 않는다고 할 수 있다.

## 5. 결론 및 논의

RX는 BOES 관측 데이터를 자동 처리하는 프로그램으로 처음 개발된 사례이다. 아직 개발의 초기 단계이지만 프로그램의 안정성과 속도 그리고 추출된 스펙트럼의 질이 대체로 만족할 만한 것으로 판단된다. 궁극적으로 RX를 이용하여 추출된 스펙트럼을 천문학적 연구에 응용하여 그 신뢰성을 점검해야 할 것이다. 이를 위해 해결하여야 할 RX의 몇 가지 문제점과 개선 사항들은 다음과 같다.

첫째, 아직 RX는 GUI 환경을 충분히 갖추지 못한 상태이다. 한편으로는 간단히 프로그램을 실행할 수 있다는 장점이 있지만 사용자가 중간 진행 상황을 잘 확인할 수 없다는 단점이 되기도 한다. 프로그램 수행을 위한 파라미터 등을 쉽게 지정할 수 있고, 프로그램 진행 상황을 보면서 중간 중간에 사용자가 처리 과정에 개입할 수 있는 GUI환경을 개발하면 직관적으로 프로그램의 진행 과정을 이해할 수 있을 뿐 아니라 처리 과정에서 발생하는 문제점도 신속히 확인 할 수 있을 것이다.

둘째, 연속선의 정규화는 별의 분광형에 따라 연속선 결정의 알고리즘이 달리 적용되어야 하고 자동화를 구현하기가 매우 어려울 것으로 예상된다. 이를 구현하기 위해서는 더욱 수준 높은 프로그램의 수정 및 개발이 이어져야 하겠다.

셋째, 본 프로그램의 개발 과정에서 가장 어려우면서 중요한 부분이 사용자가 원하는 구경의 정확한 위치 결정과 파장을 동정하는 부분이었다. 그림 15에서 보듯이 아직 이 프로그램의 알고리즘은 완벽하게 검증되지 못하였다. 이 부분에 대한 정밀한 검증과 새로운 알고리즘의 개발이 계속되어야 하겠다.

## 참고문헌

- 김강민, 2003, 보현산천문대 분광기 시스템의 개발과 성능 분석, 경북대 박사학위논문
- 조한근, 2004, MATLAB 프로그래밍, 한티미디어
- 하영호, 남재열, 이응주, 이철희, 2003, 디지털영상처리, 244
- Churchill, C. W., 1995, Introduction to Echelle Data

Reduction Using the Image Reduction Analysis Facility

Gazinur, G., 2005, 개인서신

Messey, P., 1997, A User's Guide to CCD Reductions with IRAF

Petr, S., 2002, Publ. Astron. Inst 90, 5-8

Willmarth, D. & Barnes, J., 1994, A user's guide to reducing echelle spectra with IRAF