

KMTNet 실시간 자료처리 파이프라인 개발 DEVELOPMENT OF REAL-TIME DATA REDUCTION PIPELINE FOR KMTNet

김동진, 이충욱, 김승리, 박병곤
한국천문연구원

D. -J. KIM, C. -U. LEE, S. -L. KIM, AND B. -G. PARK
Korea Astronomy and Space Science Institute, Daejeon 305-348, Korea

E-mail: keaton03@kasi.re.kr

(Received November 25, 2012; Accepted January 07, 2013)

ABSTRACT

Real-time data reduction pipeline for the Korea Microlensing Telescope Network (KMTNet) was developed by Korea Astronomy and Space Science Institute (KASI). The main goal of the data reduction pipeline is to find variable objects and to record their light variation from the large amount of observation data of about 200 GB per night per site. To achieve the goal we adopt three strategic implementations: precision pointing of telescope using the cross correlation correction for target fields, realtime data transferring using kernel-level file handling and high speed network, and segment data processing architecture using the Sun-Grid engine. We tested performance of the pipeline using simulated data which represent the similar circumstance to CTIO (Cerro Tololo Inter-American Observatory), and we have found that it takes about eight hours for whole processing of one-night data. Therefore we conclude that the pipeline works without problem in real-time if the network speed is high enough, e.g., as high as in CTIO.

Key words: KMTNet; pipeline; data reduction; database

1. 서론

한국천문연구원에서는 미시중력렌즈 방법을 이용하여 다수의 외계행성을 발견하는 외계행성 탐색시스템 개발사업(Korea Microlensing Telescope Network; KMTNet)을 2009년부터 진행하고 있다. 이 시스템은 동일한 성능을 가진 3대의 망원경을 칠레, 호주, 남아프리카 공화국에 설치하여 우리은하 중심 방향에 대한 24시간 연속 관측을 수행한다(Kim et al., 2010, 2011; Park et al., 2012). 각각의 망원경은 미국 임시관측소에서 조립과 성능시험을 거친 후 현지에서 설치되며, 현재 시험 중인 망원경은 2013년 초에 칠레에 설치된다(Kappler et al., 2012; Poteet et al., 2012). 망원경이 설치되는 초기에는 아직 연구용 카메라가 제작되지 않아 4K × 4K 카메라를 이용하여 시험관측을 수행하고, 2013년 9월경에 본격적인 관측연구를 위해 18K × 18K 연구용 카메라를 설치하여 운영할 계획이다.

한국천문연구원에서는 대용량자료의 원활한 처리를 위하여 자료처리계획을 수립하고 각각의 처리모듈을 개발하고 있으며, 여러 번에 걸친 반복시험을 통해 각각

의 기능과 루틴들의 안정성을 확인한 바 있다(김동진 등, 2009, 2011; Lee et al., 2010, 2012). 특히 외계행성 탐색시스템의 본격적인 가동을 앞두고 있는 현 시점에서 관측 자료의 생성, 전송, 측광 자료처리, 데이터베이스화, 백업 등의 전 과정을 관리하고 운영할 수 있는 파이프라인의 개발과 성능시험은 매우 중요하다. 따라서 여러 번의 디버깅 과정을 통한 파이프라인의 안정화 과정과 성능예측이 절실히 요구된다. 이에 이 논문에서는 우리가 개발한 파이프라인의 안정성과 성능을 제시하고자 한다. 2장에서는 영상차감법의 중요한 과정인 관측영상과 기준영상의 정렬 단계에서 영상 정렬 오차를 최소화하기 위한 정밀 포인팅 관측 방법에 대해 기술하고, 3장에서는 시험운영 전산기와 측광 파이프라인 구성, 영상차감법에 대해 기술한다. 그리고 4장에서는 MOA 그룹의 실제 관측영상을 기반으로 KMTNet 모의 관측 영상을 생성하여 시험운영 전산기와 측광 파이프라인의 성능시험 결과를 정리하였다.

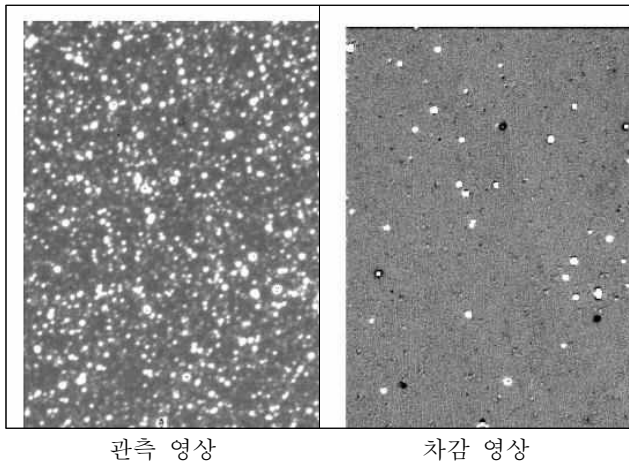


그림 1. 관측 영상과 차감 영상. 차감 영상의 하얀 점은 기준 영상보다 밝아진 별이고, 검은 점은 어두워진 별이다. Saturation 등으로 masking 작업을 하였을 경우 masking 값에 따라 변광성처럼 보이기도 한다. 기준 영상과 관측 영상의 정렬이 정밀하지 않으면 차감 영상에는 검은 점과 하얀 점이 인접하여 나타난다.

2. 고정밀 포인팅관측

우리는 외계행성 탐색시스템 파이프라인을 구성하는 주측광엔진으로 Alard & Lupton(1998)이 개발한 영상차감법을 사용한다. 그림 1은 관측영상과 영상차감법으로 차감한 영상의 일부를 확대한 영상으로, 우측 차감영상의 상단에 보이는 검은색 줄은 기준영상과 관측영상의 정확히 일치하지 않아 정렬과정에서 생긴 오프셋을 나타낸다. 이런 오프셋은 관측이 이루어질 때 얼마나 정확히 포인팅이 되었는지에 따라 영상의 가장자리에 있는 별들이 측광과정에 포함되거나 포함되지 않을 수 있다. 이 때문에 망원경의 포인팅 정밀도를 높여 최대한 많은 별들이 측광에 포함되도록 하는 것이 매우 중요하다. 은하중심부 관측은 미리 정의된 4개의 영역만을 관측하기 때문에 교차상관함수(cross correlation function)를 이용하여 타깃의 포인팅 오프셋을 정확히 알아낼 수 있다. 이 방법은 프로젝트 초기에 관측영역으로 망원경을 보낸 후 readout 속도가 빠른 오토가이더 칩을 이용하여 영상을 얻고 이를 포인팅 기준영상으로 저장한다. 이후, 동일한 영역으로 망원경을 포인팅할 때에는 오토가이더 영상과 포인팅 기준영상을 교차상관시켜 오프셋을 구해 이 값만큼 망원경을 이동시킨다. 이 방법을 적용할 경우 오프셋의 정확한 측정이 가능하여 망원경의 포인팅 정밀도는 1" 이내가 되고, 이를 통해 영상 정렬오차를 최소화 할 수 있다.

3. 자료처리 파이프라인의 구성

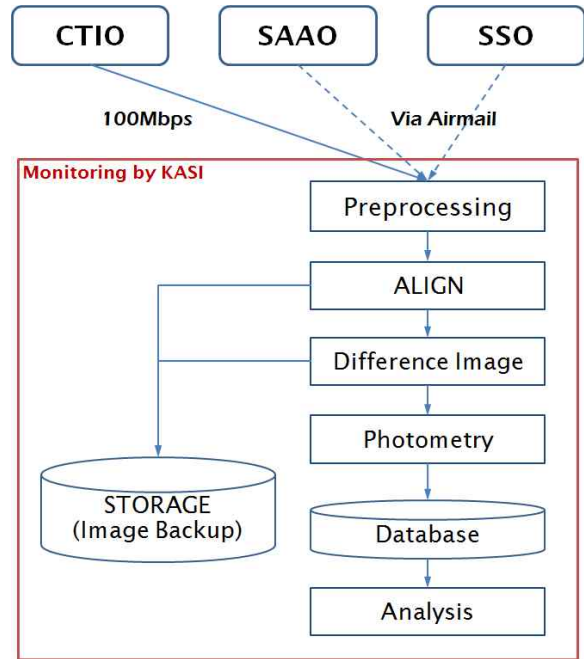


그림 2. KMTNet 자료 흐름도. 칠레는 네트워크로 관측 자료를 가져오고 남아공과 호주는 항공우편으로 관측 자료를 가져온다.

외계행성 탐색시스템 자료처리 파이프라인의 최종 목표는 각 관측소에서 생산되는 하루 최대 200 GB의 자료를 효율적으로 처리하는 것이며, 이를 위한 가장 좋은 방법은 완전 자동화된 파이프라인을 현지 관측소에 설치하는 방법이다. 하지만 이 방법은 전산자원의 관리에 있어 많은 어려움을 가지고 있기 때문에 자료처리를 위한 주전산기는 천문연구원에 설치하기로 결정하였다. 따라서 자료처리 파이프라인은 자료의 전송을 시작으로 측광자료처리, 데이터베이스화, 백업 등의 전체 단계가 천문연구원의 주전산기에서 사람의 개입 없이 또는 최소화하여 이루어지도록 구성하여야 한다. 특히 칠레관측소처럼 초고속 네트워크를 사용할 수 있는 환경에서는 전산망을 이용하여 전 과정에 대한 통합화 및 자동화가 가능하지만, 인터넷 속도가 느린 남아프리카공화국 관측소와 호주 관측소의 경우에는 항공우편을 통해 데이터를 전송하고 이를 주전산기에 다시 저장하기 때문에 이에 따른 사람의 개입이 요구된다. 그림 2는 KMTNet 측광파이프라인의 관측 자료의 흐름도이다. 관측소에서 관측한 영상을 한국천문연구원으로 전송하면 전처리 서버가 자료를 받아 전처리 작업 후 영상차감법을 수행한다. 그리고 측광 작업을 거쳐 데이터베이스로 자료를 업로드하는 자동화된 일련의 과정을 거치게 된다.

우리는 사람의 개입 없이 완전 자동화가 가능한 칠

표 1. 자료처리 파이프라인 모듈 구성

	과정	모듈 이름
STEP 1	관측자료 전송	KMTNet_MOD_STEP1_SENDER
STEP 2	전처리	KMTNet_MOD_STEP2_PREPROC
STEP 3	영상 정렬	KMTNet_MOD_STEP3_ALIGN
STEP 4	영상 차감	KMTNet_MOD_STEP4_SUBTRACT
STEP 5a	모든 별에 대한 측광	KMTNet_MOD_STEP5a_ALLPHOT
STEP 5b	차감 영상에서의 측광	KMTNet_MOD_STEP5b_VARPHOT
STEP 6	데이터베이스 업로드	KMTNet_MOD_STEP6_DBUPD
STEP 7	영상 백업	KMTNet_MOD_STEP7_ALLBAK

레 파이프라인의 구성과 성능시험을 위해 이미 구축한 시험운영 전산기(김동진 등, 2011)의 설정을 수정하여 재구성하였다. 기존의 시험운영전산기는 여러 대의 컴퓨터가 하나의 저장공간을 NFS(Network File System)로 공유하였으나, 네트워크 병목현상 등의 문제가 발생하여 이번 구성에는 각 서버에 개별 저장공간을 사용하도록 구성하였다.

특히 이번 파이프라인에서는 효율적인 자료처리와 측광자료의 관리를 위하여 관측영상을 일정한 크기로 자르고 하루치로 묶어서 처리하는 관측세트 개념을 도입하여 자료처리에 활용하였다. 즉, 관측영상을 1,152 × 1,152 크기의 단위영상으로 나누고 각각의 부분영상의 하루치를 시간 순서로 정렬하여 하나의 관측세트로 구성하고 주전산기의 코어 하나가 하나의 관측세트를 처리하는 방식이다. 관측영상을 단위영상으로 자르지 않고 처리 할 경우 하나의 CPU가 영상정렬, 차감, 측광 등의 모든 작업을 진행하기 때문에 다중코어를 가진 노드의 모든 자원을 제대로 활용할 수 없다. 뿐만 아니라 전체 영상을 주메모리에 모두 올리기 때문에 여러 작업을 동시에 진행한다면 주메모리 부족으로 자료처리 시스템이 멈추거나 리부팅 되는 경우도 발생한다. 단위영상으로 처리를 할 경우 잘라진 단위영상을 동시에 여러 CPU로 처리할 수 있고 주 메모리 사용량도 많지 않기 때문에 전산 자원을 더욱 효율적으로 사용할 수 있고 빠르게 측광 정보를 얻을 수 있다는 장점이 있다. 각각의 관측세트에 대한 처리가 끝나고 측광 정보를 분석하여 변광성이 발견되면 변광성 목록을 업데이트한다. 각 단계별 자료처리 과정을 표 1에 정리하였다. 모듈은 C-shell, IRAF(Image Reduction and Analysis Facilities), DIAPL(Difference Image Analysis Package; Wozniak & Pych, 2006) 등을 이용하여 구성되었고, 전체 전산자원의 관리는 Sun Grid Engine job scheduler를 사용한다.

칠레 관측소의 경우에는 관측한 영상을 네트워크 전송을 통해 실시간으로 전송받는다. 이를 위해 우리는

칠레관측소와 100 Mbps 속도의 네트워크 사용계약을 체결하고 실시간 파일전송 프로그램을 개발하였다. 이 프로그램(STEP1)은 칠레 관측소의 관측 컴퓨터에 관측 영상 파일이 저장될 때 커널에서 발생하는 파일 시스템의 CLOSE_WRITE 코드를 식별하면 해당 관측영상을 Aspera 고속 전송 프로토콜을 사용하여 전송하도록 구성되어있다. 파일 시스템 변경 코드는 파일이 생성되거나 삭제, 이동될 때마다 커널에서 발생하는 코드로 별도의 검색프로그램을 사용하는 것 보다 더욱 빠르게 파일이 입력된 것을 확인할 수 있고, 디스크의 파일 시스템을 직접 검색하지 않기 때문에 관측 컴퓨터의 자원 활용에도 유리하다. 따라서 칠레관측소에서 관측한 영상 1장이 하드디스크에 저장되면 약 1분 뒤에 한국천문연구원 주전산기로 동일한 파일이 전송된다. 칠레와 한국은 최소 622 Mbps 대역폭의 전산망으로 연결되어 있어 100 Mbps를 사용하는데 큰 어려움이 없다.

관측영상의 전송이 완료되고 일정시간이 지나 관측이 종료된 것으로 판단되면 천문연구원의 주전산기에서 파이프라인이 작동된다. 우리는 이 시점을 정하기 위해 현지 관측소를 기준으로 일출(태양고도 0도) 후 1시간이 적당하다고 판단하였다. 따라서 2013년 6월 15일을 기준할 때 칠레에서 아침 7시 40분에 해가 뜨면 한 시간 뒤인 아침 8시 40분, 즉 한국에서는 오후 8시 40분에 한국천문연구원의 주전산기에서 파이프라인이 자동으로 작동된다.

관측소에서는 관측영상의 전처리를 위해 관측 전과 후에 바이어스, 다크, 플랫영상을 합성하여 매일 매일의 보정영상들을 생성한다. 특히 플랫영상을 얻기 위해서는 하늘의 밝기가 빠르게 변하는 상황에서 각 필터별로 적당한 노출시간을 조절하여 관측해야 하는데, 이를 위해 태양 고도에 따른 플랫영상의 노출시간을 계산하는 프로그램을 사용한다(이충욱 등, 2007). 따라서 매일 정해진 시간 안에 전처리영상의 획득(STEP2)이 가능하다. 전처리를 마친 영상들은 1,152 × 1,152 크기로 자른 후 하루를 기준으로 관측세트를 구성한다. 영상차감에 사용하는 기준영상이 아직 존재하지 않는 관측 초기에는 관측된 영상 중 FWHM이 가장 좋은 영상 5장을 골라 정렬 및 합성하여 기준영상을 생성한다. 기준영상이 만들어진 이후에는 영상정렬 모듈(STEP3)을 이용하여 관측세트를 기준영상의 좌표와 동일하게 정렬한다.

영상정렬이 이루어진 후에는 관측영상과 기준영상을 선형 회귀분석하여 콘볼루션 커널을 구하고 이를 기준 영상에 적용하여 합성영상을 만들어 관측영상에서 빼면 차감영상을 얻게 되는데(STEP4) 차감영상에서 주변보다 빛이 밝아지거나 어두워진 영역을 찾아 중심 좌표를 구한다. 이 좌표들은 변광성으로 분류되어 우선 측광대상으로 분류되고, 변광성목록을 업데이트한 후 측광과정

표 2. 칠레 자료처리를 위한 전산기 사양. 호주와 남아프리카 공화국 자료처리를 위해 유사한 성능의 전산기 2세트를 추가하여 연동할 계획이다.

항목	구성사양
CPU	Intel Xeon X5650 2.66 GHz(6 core) × 2
노드	3
RAM	24 GB/node
HDD	300 GB(SAS 6 Gbps) 120 GB(SSD 3 Gbps)
OS	CentOS 6.3

을 수행한다. 차감영상의 측광과정은 두 가지 옵션을 가지고 있다. 하나는 검출할 수 있는 모든 좌표를 이용하여 측광을 수행하는 방법(STEP5a)이고, 다른 하나는 변광성에 대해서만 측광하는 방법(STEP5b)이다. 이렇게 2가지 옵션을 구분한 이유는 모든 별에 대한 측광을 실시할 경우 많은 연산시간이 소요되어 실시간 처리가 어렵기 때문이다. 따라서 실시간으로 변광성을 발견하고 분석하기 위해서는 변광성에 대해 우선 측광을 실시한다. 그리고 우리은하 중심부의 관측 시준이 종료되거나 관측할 수 없는 날 등 주전산기의 여유 시간이 있을 때 모든 별에 대해 재 측광을 실시한다.

변광성에 대한 측광이 이루어지면 자료의 검색과 정렬을 위해 측광데이터베이스를 업데이트한다(STEP6). 이미 김동진 등(2011)은 외계행성 탐색시스템 측광데이터베이스를 구축한 바 있으나, 이번 연구에서는 대용량 데이터베이스의 관리 향상 및 검색 속도 증가를 위하여 테이블에 최소의 자료만 입력하도록 테이블의 구조를 수정하였다. 즉 하나의 테이블에 많은 자료를 입력하는 것보다 저장공간을 조금 더 차지하더라도 여러 개의 테이블로 데이터베이스를 분할하여 입력 자료의 양을 줄이는 방법이다. 이는 검색 시간을 줄이는데 도움이 될 뿐 아니라 데이터테이블에 문제가 발생하였을 경우 백업 및 복구 작업에 있어서도 매우 유리하다. 데이터베이스로 측광 자료 업데이트가 종료되면 STEP3과 STEP4 과정에서 생성된 정렬영상과 차감영상을 저장장치로 백업한다(STEP7). 측광자료처리와 데이터베이스, 저장장치의 효율적인 운영을 위해 모든 서버는 독립적으로 운영된다.

4. 모의 영상을 이용한 성능시험

남반구에서 우리은하 중심영역을 가장 오래 관측할 수 있는 시기는 6월 중순이다. 이때는 칠레관측소를 기준으로 오후 7시에 관측을 시작하여 다음날 6시까지 관측이 가능하기 때문에 약 11시간을 관측할 수 있다. 은하 중심부 관측전략에 의하면 2분 30초마다 4개의 관측영역을 관측하므로 이 때, 총 264장을 관측할 수 있다. 여

기에 전처리를 위해 관측한 바이어스, 다크 및 플랫폼 영상을 더하면 하루 최대 약 300장의 관측영상을 얻게 된다. 18K × 18K CCD 영상의 한 장 크기가 약 650 MB 이므로 하루 최대 관측되는 파일의 크기는 약 200 GB 가 된다.

연구용 카메라는 9K × 9K의 CCD 칩 4장을 모자이크 배열하여 구성한다. 각각의 칩에는 영상정보를 읽기 위해 8개의 앰프가 설치되어 있고, 각 앰프는 1,152 × 9,216 화소영역을 읽는다. 4개의 칩에서 읽은 자료는 총 32개의 1,152 × 9,216 영상이 적층되는 FITS MEF(Multi Extension Format) 형태로 저장된다(Atwood et al., 2012). 현재 연구용 카메라는 아직 개발 단계에 있기 때문에 파이프라인의 성능시험을 위해서는 동일한 크기를 갖는 모의 영상이 필요하다. 이를 위해 우리는 연구용 카메라와 비슷한 특성의 칩을 사용하여 우리은하 중심부를 관측한 MOA(Microlensing Observations in Astrophysics; Bond et al., 2001) 그룹의 2011년 MOA-2011-BLG-034 영역 자료를 사용하였다. 우리는 각각의 2K × 4K 관측 영상을 1,152 × 1,152 단위영역을 자른 후 8개를 일렬로 합성하여 마치 1개의 앰프에서 읽은 1,152 × 9,216 크기의 데이터영역처럼 만들고, 이를 32개 적층형태로 저장하여 연구용 카메라의 포맷과 동일한 형태를 가진 모의영상을 생산하였다.

생산된 모의 영상을 칠레관측소에서 한국천문연구원으로 보내는 것과 동일한 환경을 시험하기 위하여 두 대의 컴퓨터를 사용하였다. 한 대는 칠레 현지의 관측 컴퓨터로, 다른 한 대는 한국천문연구원의 전처리컴퓨터로 가정하였다. 앞서 만든 모의 관측세트를 관측컴퓨터의 특정 디렉터리에 한 장씩 복사함으로써 칠레에서 관측 후 영상이 저장되는 효과를 연출하였다. 파일의 복사가 이루어지면 바로 전처리컴퓨터로 자료전송이 이루어졌다. 파일의 평균전송속도는 Aspera 45 Mbps 프로토콜을 사용할 때 42 Mbps였으며, 한 장의 영상 전송에 평균 2분 3초가 소요되었다. 따라서 Aspera 프로토콜의 용량을 100 Mbps로 업그레이드하면 전송에 1분이 소요될 것으로 예상된다.

천문연구원 자료처리 파이프라인의 첫 단계는 전처리 작업이다. 성능시험에서 측정된 시간은 관측 영상 1장 당 약 30초의 시간이 소요되었으며 264장의 처리에 약 130분이 걸렸다. 전처리와 같은 파일단위의 작업은 한 대의 컴퓨터에서 분산처리를 하면 디스크의 입출력 병목현상과 메모리 사용률 급증 때문에 분산처리를 하지 않았을 때의 전처리 시간과 큰 차이가 없다. 따라서 전처리 과정은 1개의 연산기만 사용하는 단일 작업으로 수행하도록 설정하였다. 표 2는 성능시험에 사용한 칠레 자료처리 전산기의 사양이고 그림 3은 전산기의 구성도이다. 전산기의 각 노드 간 연결을 10 Gbps로 구성

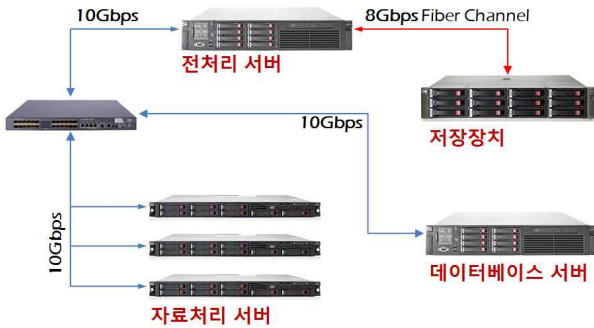


그림 3. 성능평가 시험에 사용한 전산기의 구성도.
 모든 서버는 10 Gbps 네트워크로 연결되어있고, 전처리 서버와 저장공간은 8 Gbps 광케이블로 연결하였다.

한 것은 대용량의 관측영상의 이동 시 발생할 수 있는 네트워크 병목현상을 줄이기 위해서이다.

전처리 작업이 끝난 관측 영상은 1,152 × 1,152 단위영역으로 잘라 주전산기로 전송된다. 단위영역의 영상정렬, 영상차감, 측광은 Sun Grid Engine 잡스케줄러로 제어하여 작업에 부하가 발생하지 않는 범위에서 안정적으로 전산자원을 사용하도록 구성하였다. 초기 구축한 주전산기는 10 Gbps 네트워크 환경에서 NFS로 구성된 저장공간을 여러 대의 컴퓨터가 공유하는 방식이었으나 저장공간의 입출력 접속이 많아져 네트워크 대역폭이 포화되는 현상이 발생되었다. 또한 대역폭 포화상태가 장시간 지속되면 NFS 연결이 끊어지는 문제도 발생하였다. 이에 본 연구에서는 저장공간을 공유하는 대신 각 서버에 개개의 입출력 장치를 사용할 수 있게 시스템을 재구성하였다. 최적의 입출력장치를 선정하기 위하여 하나의 노드에서 1,152 × 1,152 단위영역 1개를 측광 처리를 하였을 때 HDD, SSD, RAM 드라이브 장치에 따른 연산 시간을 1차 측정하고, 단위 영역 12개(노드에 장착된 코어의 최대 개수)를 동시에 측광하였을 때 드라이브 장치별로 연산 시간이 증가하는 것을 2차 측정하여 표 3에 나타내었다. KMTNet 측광 파이프라인은 다중 코어를 사용하여 분산처리를 하는 방식이기 때문에 대량으로 발생하는 자료의 입출력을 드라이브 장치가 제때 소화해 내지 못하면 전체 연산 시간에도 영향을 미칠 수 있기 때문에 이 실험을 통해 최적의 드라이브 장치를 선택할 필요가 있다. 표에서 알 수 있듯이 RAM 드라이브를 사용할 때의 성능이 가장 좋았으나, 시스템 구성에 있어 많은 비용과 어려움이 수반되어 SSD를 사용하여 시스템을 구성하였다. 재구성한 시스템으로 영상정렬(STEP3)에서 측광(STEP5)과정까지 소요된 시간은 표 4에 나타내었다. 개별 저장공간을 사용하도록 재구성한 시스템으로 약 5시간 30분의 연산

표 3. 입출력장치에 따른 단위영역 당 연산 시간

	HDD	SSD	RAM 드라이브
1개 단위영역 측광 처리	588 sec	587 sec	583 sec
12개 단위영역 동시 측광 처리	752 sec	688 sec	630 sec

표 4. NFS와 개별 저장공간(SSD) 사용에 따른 61,440개의 단위영역 연산 시간 비교(3개 노드 36코어 사용)

	연산 시간
NFS(10 Gbps) 공유 저장공간 사용	630 min
각 노드에 SSD 저장공간 사용	330 min

시간이 소요되었고 이는 NFS 공유 저장공간 방식을 사용한 기존 시스템에 비해 약 2배 개선된 결과를 얻었다. 이때 1,152 × 1,152 단위영역에서 발견한 변광성의 수는 약 5,500개이고, 각 관측세트의 측광을 마치고 데이터베이스의 업데이트에 소요된 시간은 1초미만으로 전체 파이프라인의 처리시간에 큰 영향을 미치지 않았다. 데이터베이스 입력과정이 종료되면 측광과정에서 얻은 정렬영상과 차감영상을 별도의 저장공간으로 옮겨 정기적인 테이프 백업과정을 수행한다.

5. 결론 및 토의

우리는 이 연구를 통해 칠레관측소의 실시간 자료처리 파이프라인을 구성하였다. 파이프라인의 안정성 시험결과 파일의 전송부터 데이터베이스 업데이트까지 전 과정에 걸쳐 안정적으로 작동되는 것을 확인하였다. 또한 칠레관측소와 비슷한 관측조건을 구현하기 위하여 MOA 그룹에서 획득한 우리은하중심 관측자료를 이용하여 모의 관측자료를 만들어 시험하였다. 그 결과, 자료전송에는 1분, 전처리과정에 2시간 10분, 측광처리 작업에 5시간 30분이 각각 소요되어 실시간 처리에 충분한 성능을 가지고 있음을 확인하였다. 이를 토대로 칠레에서 관측시간이 제일 긴 6월 15일을 기준으로 관측 후 자료처리 과정을 예측할 수 있었다. 칠레에서 관측한 영상 1장은 1분 뒤에 한국으로 전송된다. 칠레에서 관측이 종료되고 1시간 후인 한국시간 오후 8시 40분에 전처리 작업을 시작으로 측광처리, 데이터베이스 업데이트 및 영상 백업까지 모두 7시간 40여분이 소요되어 다음날 새벽 4시 20분경이면 모든 처리가 완료된다. 이때 데이터베이스에 기록되는 변광성의 총 개수는 560만 개를 가정하였다. 만약 변광성의 개수가 2배 더 많아진 천만개를 가정하면 자료처리는 오전 10시에 완료될 것

으로 예상된다.

이번 개발과정에서 3가지 부분의 성능향상을 이루었다. 첫째, 측광엔진으로 사용하는 영상차감법이 가진 고정밀 영상정렬과정을 구현하기위해 교차상관 보정 프로그램을 개발하고 이를 통해 망원경의 정확한 포인팅이 이루어질 수 있도록 관측전략을 세웠다. 둘째, 자료의 실시간 전송을 위해 운영체제 커널 단계에서 파일 핸들링 이벤트를 검출하고 Aspera 프로토콜을 이용하여 파일을 전송하는 프로그램을 개발하였다. 셋째, 파이프라인의 안정성과 효율성의 극대화를 위해 하루치 관측영상을 $1,152 \times 1,152$ 크기의 단위영역으로 나누어 자료처리 세트를 구성한 후 Sun Grid Engine job scheduler로 각 세트를 처리하는 파이프라인을 구성하여 파이프라인의 안정성과 효율성을 극대화하였다.

이 연구결과를 바탕으로 우리는 동일한 시스템을 남아프리카공화국과 호주의 자료처리에도 적용할 계획이다. 다만 남아프리카공화국과 호주의 경우에는 항공우편을 통한 자료전송을 계획하고 있기 때문에, 이를 제외한 나머지 부분에 대해서 동일한 자료처리 시스템과 파이프라인을 한국천문연구원에 구성할 예정이다. 남아프리카 공화국의 경우, 현재 추진 중인 SKA(Square Kilometre Array) 프로젝트(Hall et al., 2009)가 원활히 진행되면 현지 사이트의 전산망속도가 크게 향상될 것으로 예측되어, 항공우편 방식에서 네트워크를 이용한 실시간 자료전송 시스템으로의 전환이 가능하리라 기대한다.

외계행성 탐색 시스템은 3개 관측소에서 하루 최대 600 GB의 관측 자료를 생성하며 이에 따른 방대한 저장공간을 필요로 한다. 특히 하드디스크의 안정적인 운영시간이 5년인 것을 고려하여 저장장치의 고장이나 용량 부족 등의 문제가 발생하지 않도록 증설계획을 세워 저장공간으로 인해 문제가 발생하지 않도록 확장하는 것이 매우 중요하다고 판단된다.

참고 문헌

- 김동진, 이충욱, 김승리, 박병곤, 이재우, 2009, KMTNet 자료처리 시스템 설계와 측광데이터베이스 구축, PKAS, 24, 83
- 김동진, 이충욱, 김승리, 박병곤, 2011, 중력렌즈 사건 측광 데이터베이스 및 프레임워크 개발, PKAS, 26, 41
- 이충욱, 김승리, 2007, 최적화 된 바닥고르기 영상 관측, 천문기기워크숍, 1, 109
- Alard, C. & Lupton, R. H., 1998, A Method for Optimal Image Subtraction, ApJ, 503, 325
- Atwood, B., O'Brien, T. P., Colarosa, C., et al., 2012, Design of the KMTNet Large Format CCD Camera, SPIE, 8446, 6
- Bond, I. A., Abe, F., Dodd, R. J., et al., 2001, Real-time Difference Imaging Analysis of MOA Galactic Bulge Observations During 2000, MNRAS, 327, 868
- Hall, P. J., Schilizzi, R. T., & Lazio, T. J. L. W., 2009, The Square Kilometre Array, IEEE, 97, 1482
- Kim, S. -L., Park, B. -G., Lee, C. -U., et al., 2010, Technical Specifications of the KMTNet Observation System, SPIE, 7733, 107
- Kim, S. -L., Park, B. -G., Lee, C. -U., et al., 2011, Wide-field Telescope Design for the KMTNet Project, SPIE, 8151, 43
- Kappler, N., Kappler, L., Poteet, W. M., et al., 2012, Prototype Enclosure Design for the Korea Microlensing Telescope Network (KMTNet), SPIE, 8444, 43
- Lee, C. -U., Koo, J. -R., Kim, S. -L., Lee, J., Park, B. -G., & Han, C., 2010, Detection of Variable Stars in the Open Cluster M11 Using Difference Image Analysis Pipeline, JASS, 27, 289
- Lee, C. -U., Kim, D. -J., Kim, S. -L., Park, B. -G., & Cha, S. -M., 2012, Korea Microlensing Telescope Network: Data Management Plan, SPIE, 8451, 1
- Park, B. -G., Kim, S. -L., Lee, J. W., Lee, B. -C., Lee, C. -U., Han, C., Kim, M., Moon, D. -S., Moon, H. -K., Rey, S. -C., Sung, E. -C., & Sung, H., 2012, Korea Microlensing Telescope Network: Science Cases, SPIE, 8444, 47
- Poteet, W. M., Cauthen, H., et al., 2012, Design and Fabrication of Three 1.6-meter Telescopes for the KMTNet Project, SPIE, 8444, 5
- Wozniak, P. R. & Pych, W., 2006, Difference Image Analysis Package, <http://users.camk.edu.pl/pych/DIAPL>