Astronomical Data Analysis Software and Systems X
ASP Conference Series, Vol. 238, 2001
F. R. Harnden Jr., F. A. Primini, and H. E. Payne, eds.

A High Throughput Photometric Pipeline

Michael L. Reid[1], Denis J. Sullivan, Richard J. Dodd[1,2]

School of Chemical and Physical Sciences, Victoria University of Wellington, New Zealand

Abstract. The advent of large format CCD detectors in projects measuring time-critical astrophysical phenomena has resulted in an explosion in data generation rates. The current generation of gravitational microlensing and all-sky surveys have, on an ad-hoc basis, developed data processing software which currently meets their needs. However, the next generation of these projects will require faster, more advanced software to manage the data flow. As part of the Japanese/New Zealand microlensing collaboration, MOA (Microlensing Observations in Astrophysics), a software pipeline has been developed with the intention of being highly scalable, automated, portable, flexible, and robust. The core of the system is built around a high performance object database optimised for time series work and flexible reduction software that can use the PSF fitting packages DoPHOT and DAOphot II, as well as the ISIS Optimal Imaging Subtraction software. Evolution of the software has made it suitable for general purpose astronomical photometric reduction. This paper provides an overview of the software system.

1. Introduction

The Microlensing Observations in Astrophysics Project (MOA[3]) is a collaboration between Japanese and New Zealand scientists to search for gravitational microlensing events. Gravitational microlensing occurs when a massive object bends light from a luminous background object resulting in a time-dependent apparent brightness change of the background object. Such events are produced by rare alignments between the luminous source, massive body and the observer and to have a reasonable chance of observing some microlensing events the brightness of millions of stars must be measured over many nights. MOA performs nightly observations from Mount John University Observatory (MJUO) of ten million stars towards the Galactic Bulge and Magellanic Clouds.

The MOA Project has a custom microlensing detection system (Bond 2000) at MJUO but also exports data to member institutions. Lessons learnt from this

[1]Carter Observatory, Wellington, New Zealand

[2]Faculty of Science, University of Auckland, New Zealand

[3]`http://www.vuw.ac.nz/scps/moa/`

system were used to build software suitable for automated digital image (CCD) reduction which could be used with any telescope/detector combination and for any astronomical time-series research. The software is specifically designed for use in large astronomical projects with mosaic CCD detectors which produce too much data for the reductions to be guided by manual intervention.

2. Automated Reduction

A software package called Autophot has been developed to automate reduction of CCD images. High throughput is achieved by reducing the images in parallel using a task scheduler (most useful on multi-processor computers). The reduction can be performed using DAOphot II (Stetson 1987), or DoPHOT 3.1 (Schechter 1993) – ported by the authors to ANSI C. Autophot has been adapted to use the ISIS Optimal Image Subtraction software (Alard & Lupton 1998, Alard 2000) and modifications to ISIS to improve the consistency of the image subtraction are in progress. The system has been designed for robustness: failure to reduce one image does not halt the pipeline and prevent the reduction of other images.

A new task scheduler which uses Java Remote Method Invocation (RMI) is being developed to allow parallel reduction on a cluster of workstations. This will allow improvements in reduction throughput which scale, almost linearly, with the size of the cluster. Implementation of the software with portability in mind should allow heterogeneous computer types to be added to the cluster (including all the Windows PCs which are unused at night).

3. Data Storage

Knowledge of typical data access patterns required for fast time-series photometric analysis allows data storage to be optimised for fast retrieval. An object database implemented in C++ has been built ('StarBase') which has shown itself to provide fast access to approximately one hundred observations of several million-star star fields made by the MOA mosaic SITe CCD. The database files written by StarBase are platform-independent and can be read by the Java java.io.DataInputStream classes too. The database is customised by programming to a C++ API and a set of Java bindings have been tested.

The database implementation uses algorithms which were selected (or developed) to perform well on large data sets. As an example, built into the database is the ability to find non-variable stars and use them to estimate corrections for atmospheric transparency (called 'homogenisation').

4. Homogenisation

The least-squares atmospheric transparency estimation method suggested by Honeycutt (1992) takes about one hour and ∼256 MB RAM to perform on ∼2000 observations made by Sullivan et al. (2000) on the pulsating white dwarf GW Librae. An iterative method developed by the authors estimates the same transparency corrections to within one milli-magnitude of the least-squares estimates.

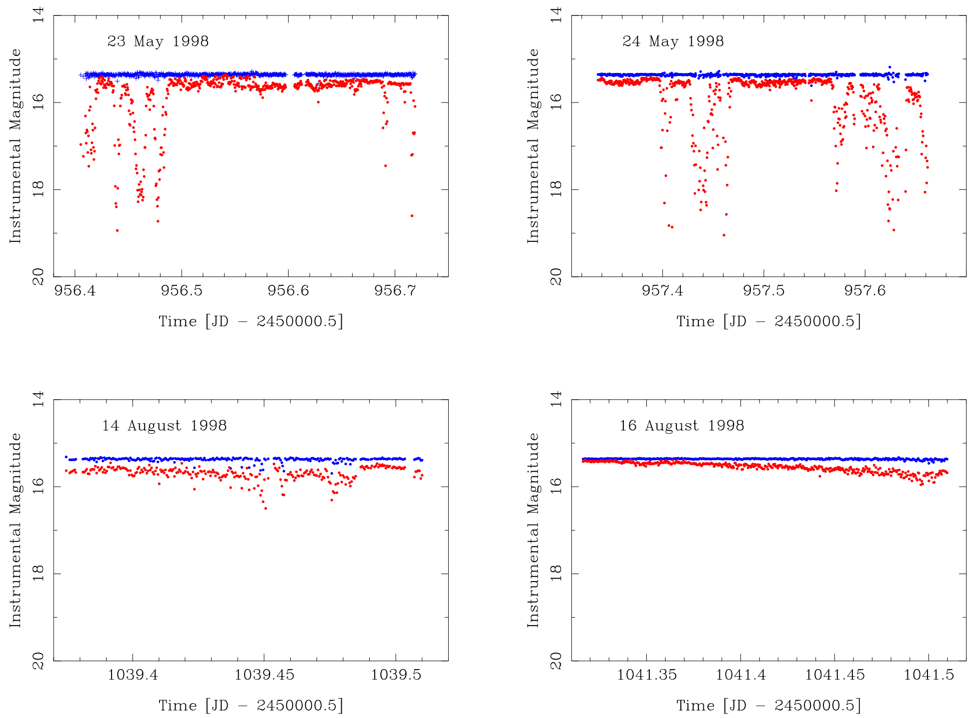

Figure 1. Observations of a reference star near GW Librae. The lighter (red) points are the observed instrumental magnitudes and the darker (blue) points are the instrumental magnitudes corrected for atmospheric transparency variations (presumably due to transient clouds).

The iterative method requires 30 seconds and 10 MB RAM and is able to deal with an ensemble of stars with missing observations. Using an algorithm which scales well with increasing data had an enormous effect in this case, and has guided development of the data pipeline. Reduction of these white dwarf data (which were not obtained by the MOA collaboration) demonstrates that the pipeline can be used in other astronomical projects.

5. Performance - Case Study

This reduction pipeline software has been tested on a number of MOA observation targets including the possible planetary microlensing event MACHO-98-BLG-35 (Rhie et al. 2000) and the finite-source microlensing event MACHO-95-BLG-30 (Alcock et al. 1997), successfully reducing observations made in both poor and good seeing conditions. Observations made on MACHO-95-BLG-30 by the MOA Project were originally reduced manually (by the authors) using IRAF/DAOphot over six man-months. The same images were reduced using Autophot/DoPHOT with the same hardware (a 4–CPU SGI Indigo) but removal of human interaction allowed a reduction time of four hours using parallel processing. Estimation of atmospheric transparency corrections by manual plotting

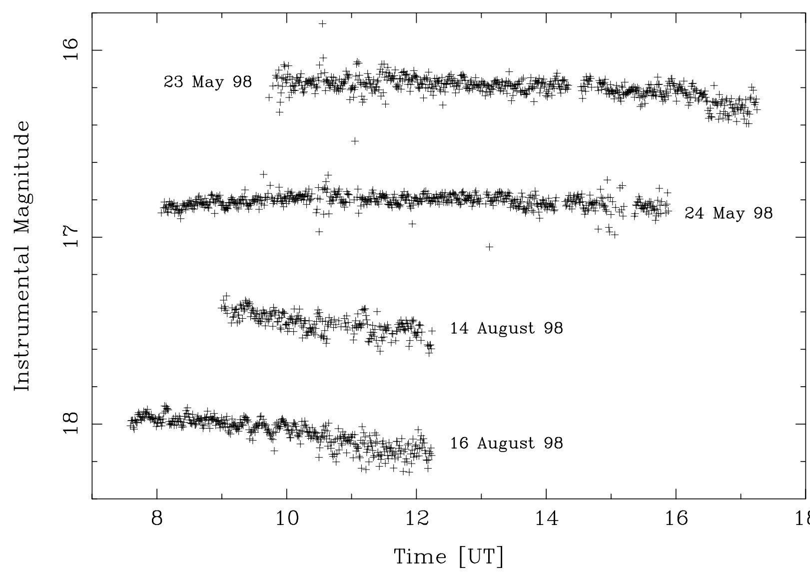

Figure 2. Observations of GW Librae with most atmospheric effects removed (nights earlier than 16 August have been offset for clarity). The brightness pulsations of the white dwarf are now visible.

and examination of star lightcurves took ∼weeks, a process now performed by StarBase in 30 seconds.

6. Acknowledgments

The authors would like to thank all members of the MOA Project and acknowledge financial support from Carter Observatory New Zealand, Victoria University Science Faculty Leave and Grants Committee, and the Marsden Fund of New Zealand. Additional thanks go to the Astronomical Data Analysis Software and Systems X (ADASS 2000) Committee for conference financial assistance.

References

Alard, C. 1998, A&AS, 144, 363

Alard, C.& Lupton, R. 1998, ApJ, 503, 325

Alcock, C., et al. 1997, ApJ, 491, 436

Bond, I. 2000, "Microlensing 2000", ASP Conf. Series (in press)

Honeycutt, R. 1992, PASP, 104, 435

Rhie, S., et al. 2000 ApJ, 533, 378

Schechter, P. 1993, PASP, 105, 1342

Stetson, P. 1987, PASP, 99, 191

Sullivan, D., et al. 2000, Baltic Astronomy, 9, 223