

Automatic photometric processing methods for star variability identification

Sergey Bratarchuk^{1*}, Zlata Potilicina²

¹RTU MTAF AERTI, Lomonosova Str, 1v, Riga, LV-1003, Latvia

²Longgenesis, Hong-Kong

*Corresponding author's e-mail: sergey.bratarchuk@gmail.com

Received 30 March 2019, www.cmmt.lv

Abstract

In the task of variable star detection exists a problem of missing data. By using shared telescope networks like LCO, users often face the concurrence for the observation time. This concurrence does not let to make a lot of photos of the same part of the sky. The author of the research proposes a new method for the solution of the missing data or unevenly based data problem in the task of variable stars' detection. Method is based on the addition of your own variable star data by using the data of other researchers. Author suggests an algorithm that identifies the star of interest on the series of photos. Algorithm automatically identifies the stars on the different images independently from the shift or rotation of the stars on the image. Then the algorithm extracts the data about the flux and magnitude of the stars on the image. In this way, by getting data about the magnitude and flux of the star from different sources, it is possible to fill the gaps in data that will increase the probability that a star will be identified as a variable one.

Key words

Computer vision, Variable stars, Cepheids, image processing

1 Introduction

A variable star is a star whose apparent magnitude as seen from Earth fluctuates. The variability of a star was first detected in the 16th

century. For next centuries, it was only possible to conduct observational studies the brightest stars in the sky. Since the invention of the computer, astronomical research has become less and less related to the direct contact of

a researcher with a telescope and more to the acquisition and analysis data, which also enabled deeper exploration of variable stars. There are many developed algorithms for determining the period, however, the problem of variable star detection still exists [1]. Nowadays astronomers use telescope networks where photography of a part of the sky can be requested via the Internet. As a result, there is competition for telescope operating time, which sometimes does not allow multiple photographs of the same star to be taken over a long period. The lack of photographs can lead to a periodic detection algorithm not detecting a variable star, especially if the period of this star is relatively short. The development of methods that can supplement data with the data of other scientists is topical. Such methods are not yet fully developed, but the methods that have been developed are fragmentary. There are, for example, VaST software that allows you to identify variable stars in a series of images when the overlay area is over 40%. This points to how the research problem is topical to computational methods of astronomy and not fully resolved. The aim of this work is to develop an instrument for automatic identification of star variability in a series of celestial images.

2 Detection of variable stars using sky image photometry

2.1 DETECTION PRINCIPLE

There are 3 phases for obtaining research data from a shared subscription resource: data collection (software for telescopes or observatories), data processing and detection

of various types of objects or events. In most cases, in the data acquisition process, attention is paid not to existing algorithms that control telescopes, but rather to study how to use methods for processing astronomical images more effectively and efficiently. However, special attention is paid to data processing and object detection [2].

Astronomical images are usually captured with a CCD photodetection matrix. These detectors are similar to those used in mobile phones and other devices. Typically, images are stored in astronomy distributed FITS file format [3]. FITS is the data format most commonly used in astronomy for the transportation, analysis and archiving of scientific data files. FITS has much more than just one image format (such as JPG or GIF) and is mainly designed to store scientific datasets consisting of multidimensional arrays (images) and two-dimensional tables containing image information (date of capture, subject name and coordinates, etc.) [4] These images are almost always grayscale. If color pictures are required, color information is obtained by sequentially using multiple color filters.

The main principle of the variable star detection method is based on the identification of changes between images in a series of images taken at intervals of a few minutes to several days, for example like in the Figure 1. Changes can be detected by comparing the images by pixels or elemental image fragments. This technique is the analysis of image differences [5] Some scientists, based on the principle discussed above, have proposed their own methods and algorithms with their own improvements, such as [6 - 10]. These methods are also used in research on the detection of stars in new variables such as [11, 12].

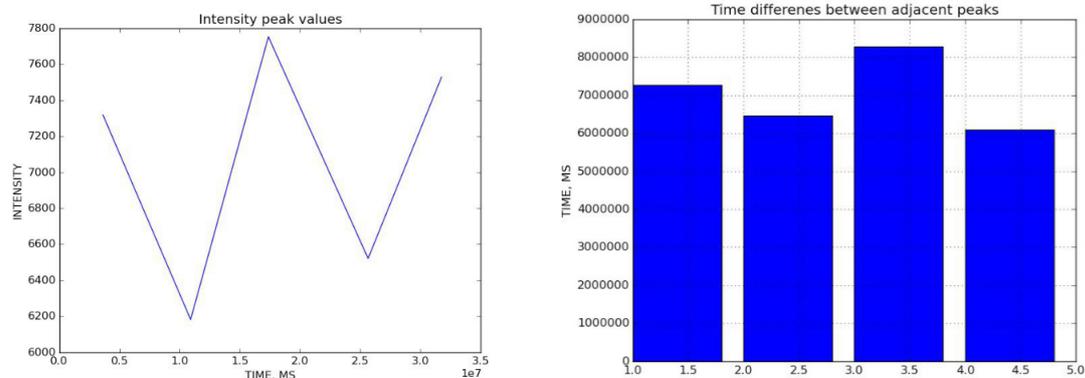


FIGURE 1 Intensity peak values and time difference between peaks of a variable star

In the literature, one of the sub-methods of subtraction, or subtraction, is often found in these methods. This data processing technique is based on the principle of subtraction of matrices - $B = A - Ar$, where B is the resultant matrix, photometric matrices of A at some point of time, and Ar is the reference image taken under particularly favorable weather conditions in a moderate atmosphere with low humidity. Most areas in the resulting image are dark, representing areas where two images are close to the same brightness. However, a number of bright point sources appear in the image of the residue. These points indicate areas where the brightness of two images differs. This technique is used by large projects such as OGLE [13] and ROTSE [11].

Based on the techniques and techniques discussed above, programmers have also offered solutions to the problem of detection and identification of variable stars. One such solution is VaST. This computer program is suitable for identifying variable stars in astronomical image series (images from FITS files). The software allows you to process images with redirected and rotated images of the same scale (the image has the same pixel height and width and the image's spatial angle). The coating should be at least 40%. This computer program is not fully automatic and requires user intervention [14]. There is also a library of algorithms AsPyLib that is suitable for Python programming language. Using this library, you can detect variable stars, as well as analyze their light curves. The detection process is not fully automated, and the support and updating of this module has ended in 2015 [15].

2.2 PROBLEM OF UNEVENLY DISTRIBUTED DATA

Literature shows that studies related to variable star data processing can highlight 2 unsolved problems - lack of automation and insufficient data. For example, [12] The authors of the study admit that visual inspection is required. There is also a problem of lack of data, [12] authors admit that some stars were not clearly defined. The authors comment that this might be due to the irregular nature of the light variability of this star, or whether the amount of data was not large enough to be clearly classified as long-term variable stars. The problem of the lack of large amounts of data is also not fully resolved.

The number of images of these stars

is important for the resultant detection of variable stars, for calculating this star period and for classification. This is especially true for changing stars with a small period, which can change brightness within hours or even faster. In circumstances where the time of use of telescopes is strictly limited, research on variable stars may be difficult.

There have been attempts to solve the problem of unevenly selected data over time. There are algorithms and techniques that are well suited for processing uneven data over time. The most commonly used methods for determining the frequencies of variable stars are the following: Lomb-Scargle method [16], Fast Chi-Squared technique [17], SigSpec [18], Cubic spline method [19], Welch-Stetson algorithm [20].

Literature shows that automatic identification of variable stars is a topical task in astronomy and this task has not been resolved to the end. Existing data acquisition and processing approaches may not always guarantee variable star detection. In the case of unsuccessful coincidences, the time intervals at which the starry sky photographs were taken may coincide with the maximum or minimum of the star's brightness. This leads to the need to get more images for each part of the sky. This would exclude the possibility that the algorithm will not identify the variability of the star's brightness. There is also a problem with the competition for the telescope's lifetime, which prevents the reception of unlimited number of sky images. Another approach to automatic detection of variable stars is needed.

In this study, only pulsating star images were processed, more precisely cefeids, because these stars, unlike the double star of the eclipse, the sinusoidal light curve with the threshold, unlike the irregularly pulsating stars, have a regular period, unlike Mira's pulsating stars, the pulsation period is relatively short.

2.3 METHOD

2.3.1 Development of cefeid identification program

Automatic identification of cefeids in images is a task that needs to be solved in order to successfully study the same stars, whose images were taken in different studies or researches. The main problem in this task is that sometimes the star of interest may be in the pictures taken during different studies. This leads to a situation that the same star

may have different coordinates on the images (because it was consciously or unconsciously captured, maybe with different telescopes and perhaps from different angles of view).

The authors offer an approach for identifying cepheid stars in a series of celestial images. The algorithm is implemented in a program written in Python 2.7 programming language using algorithm libraries photutils (one of the Astropy modules that provide photometric detection of astronomical sources and tools) and astropy (a module for processing astronomical images and data). The program's input data is a series of images with a specific sky fragment taken at different times, possibly within the framework of various studies. The software processes data in FITS format.

The steps of the algorithm:

1. The indication of the path to the directory where the star image series is located; The 1st image in the directory is considered a reference image (reference image - is an image that will be compared with other images in the series for cepheid identification);
2. The coordinates of the variable star in the reference image are entered;
3. The image captured by the telescope includes both the sky, the stars and other cosmic objects, so the next step was to recognize the stars in the pictures. For this purpose, the DAOFind algorithm from the photutils library was used. This algorithm was developed and patented by Peter Stetson [21]. DAOFIND searches images for local density maxima that have a peak amplitude greater than threshold (approximately; threshold is applied to a convolved image) and have a size and shape similar to the defined 2D Gaussian kernel. The output of this algorithm contains the coordinates of all the detected stars, the sharpness of the star in the image, the flux, the instrumental magnitude of the star (magnitude) (Figure 2).
4. Sometimes the shape and characteristics of different stars are similar, so that the identification of a certain star must consider not only the shape and characteristics of an identifiable star, but also the shape and characteristics of its adjacent star. So, the identification should be done using a pattern that includes both an identifiable star and adjacent

stars. This was also done by calculating a template from the reference image. The template has a square shape whose size was calculated dynamically (depending on the image size). Then all the distances from the identifiable star to the other "adjacent stars" were calculated. These distances also will be used to identify the star in other images in the series.

5. Processing of other images in the series (algorithm for single-star processing):
 - 1) All existing stars are detected by the image using DAOFind.
 - 2) The same stars (even in different images) will have similar sharpness. Therefore, the sharpness of all the stars identified in the image is compared with the same properties in the reference image. Only a small group of stars will be left.
 - 3) Small squares are created (the area is an automatically calculated dependence on the image size), in which the center is a candidate who can be an identifiable star, this square also includes other stars. The distances from the candidate star to the other star square are calculated. These distances were compared to the distances the reference image and the desired variable star are identified in the image.

2.3.2 Testing of algorithm using generated images

To test whether the algorithm can identify the desired star in the image series, data was generated for testing. Image series with 3 complexity levels were generated:

1. The pictures contain 10 to 50 stars with randomly generated coordinates and other properties; In a series of images, each image is guaranteed to have an identifiable star and all the images have the same dimensions;
2. Figures contain 50 to 500 stars with randomly generated coordinates and other properties; In a series of images, each image is guaranteed to have an identifiable star and the images have different dimensions;
3. Figures range from 500 to 1000 stars with randomly generated coordinates and other properties; Images in the series may not have identifiable stars and images have different dimensions;

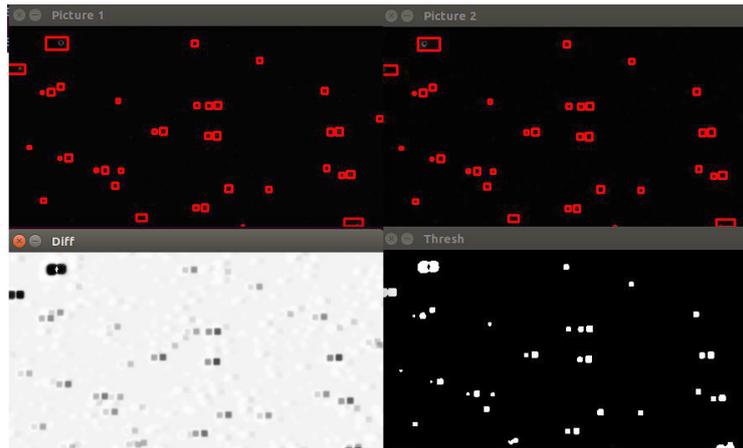


FIGURE 2 Steps of the identification of the stars using DAOFind

2.3.3 Testing of using real astronomical images

After processing the generated images, you must process the real images of the sky. The images were taken from both the LCO archive and some images were taken by the authors using LCO telescopes.

Three objects for capturing were selected. These objects are classic Cefeids - AH Velorum (period 4.2 days), BG Crucis (period 3.3 days) and Beta Doradus (period 9.8 days). By accessing the LCO image archive (which contains all the pictures taken with this telescope network), the authors have noticed that other scientists had not photographed these stars and these sky fragments using this telescope network, so the authors of the work were the first to do it using LCO telescopes. These photos are now available to all other users of the LCO telescope network in the scientific archive.

After that, all the taken photos were analyzed using the developed software. Then the authors used the Lomb-Scargle method to detect the period.

After that the algorithm was once again tested using the images of 3 Cepheids CU Mon, AK Cep and W Sgr. The images were retrieved from the LCO scientific archive.

3 Conclusions

1. The developed algorithm when processing artificially generated images successfully processed and identified the star in images with coverage less than 25%;
2. When processing data from the LCO archive, it was determined that using data only from one study or research it

is not possible to accurately determine the period of cefeid because of the lack of data;

3. When processing the images that was taken by the authors, the correct period was not calculated for any star. Continuity of observations and number of images are important for the correct determination of the period. The breaks were due to the windows expired error (the observation time has expired but the photograph was not taken). When contacting the support service, the response was that this error occurred when a lot of people also wanted to observe the same telescope and at the same time. The support service confirmed that there was competition and offered to attract other users of the telescope, to ask or to help them to take photos (collaborative project). So, to determine the cefeid period, the number of images must be increasing, especially if the period is a few days, and several photos must be taken in one day;
4. When processing data that was retrieved from LCO scientific archive (AK Cep, CU MON and W Sgr), the periods of the stars were calculated correctly. Overall over 1000 images were processed and thus it was possible to correctly determine the periods;
5. VaST software requires at least 40% image overlay for successful star identification. The method and the program based on it, developed within this research, can successfully identify a star even if the image coverage is less than 40%. For example, in 25% of the

generated images, the stars of interest were successfully identified;

In this study, an algorithm and software for automatic identification of cefeids in a series of star images were developed. The developed method also includes the necessary steps to identify cefeids. The identification process

utilized cefeid's sharpness in the image, as well as the cefeid's distance to the other stars (in pixels). Developed software, unlike the VaST software, can identify a star in images with a total area of less than 40%, and unlike the methods offered by the AsPyLib algorithm library, this software is automated.

References

- [1] Potiljicina Z 2017 .The identification of variable stars using computer vision approaches and methods *Research and technology – step into the future* **12**(2) 14
- [2] Djorgovski S G, Mahabal A A, Drake A J, Graham M J, Donalek C 2013 *Sky Surveys* In: Oswalt T.D., Bond H.E. (eds) Planets, Stars and Stellar Systems. Springer, Dordrecht
- [3] EESA *A short introduction to astronomical image processing* **E-source:** www.spacetelescope.org/projects/fits_liberator/improc (Accessed 15 Feb. 2018)
- [4] NASA 2014 *A Primer on the FITS Data Format* **E-source:** https://fits.gsfc.nasa.gov/fits_primer.html (Accessed 19. Feb. 2018)
- [5] Sokolovsky K V, Gavras P, Karamelas A, Antipin S V, Bellas-Velidis I, Benni P, Bonanos A Z, Burdanov A Y, Derlopa S, Hatzidimitriou D, Khokhryakova A D, Kolesnikova D M, Korotkiy S A, Lapukhin E G, Moretti M I, Popov A A, Pouliaxis E, Samus N N, Spetsieri Z, Veselkov S A, Volkov K V, Yang M, Zubareva A M 2016 Comparative performance of selected variability detection techniques in photometric time series data *Monthly Notices of the Royal Astronomical Society* **464**(1) 274–92
- [6] Tomaney A B, Crotts A P S 1996 Expanding the realm of microlensing surveys with difference image photometry *Astronomical Journal* **112** 2872
- [7] Alard C, Lupton R H 1998 A method for optimal image subtraction *Astronomical Journal* **503**(1) 325-31
- [8] Becker A C, Homrighausen D, Connolly A J, Genovese C R, Owen R, Bickerton S J, Lupton R H 2012 Regularization techniques for PSF-matching kernels - I. Choice of kernel basis *Monthly Notices of the Royal Astronomical Society* **425**(2) 1341–9 **E-source:** <https://academic.oup.com/mnras/article/425/2/1341/1193862> (Accessed 19. Feb. 2018)
- [9] Zackay B, Ofek E O 2016 Proper image subtraction—optimal transient detection, photometry, and hypothesis testing *The American Astronomical Society* **830**(1) 27 **E-source:** <http://iopscience.iop.org/article/10.3847/0004-637X/830/1/27/meta> (Accessed 19 Feb. 2018)
- [10] Bramich D M, Horne K D, Alsubai K A, Bachelet E, Mislis D, Parley N 2016 Difference image analysis: automatic kernel design using information criteria *Monthly Notices of the Royal Astronomical Society* **457**(1) 542-74
- [11] Whallon N 2011 *The automated detection and classification of variable stars* PhD thesis. The University of Michigan **E-source:** <https://lsa.umich.edu/content/dam/physics-assets/physics-documents/senior-theses/Nikola%20Whallon.pdf> (Accessed 6 Feb. 2018)
- [12] Zhang T M, Wang X F, Chen J C, Zhang J J, Zhou L, Li W X, Liu Q, Mo J, Zhang K C, Yao X Y, Zhao X L, Zhou X, Nie J D, Huang F, Jiang Z J, Ma J, Wang L Z, Wu C, Zhou Z M, Zou H, Wang L F 2015 The THU-NAOC transient survey: the performance and results from the first year *National Astronomical Observatories of Chinese Academy of Sciences and IOP Publishing Ltd* **15**(2) 215 **E-source:** <http://iopscience.iop.org/article/10.1088/1674-4527/15/2/006> (Accessed 19 Feb. 2018)
- [13] Szymanski M K 2005 The optical gravitational lensing experiment. Internet access to the OGLE photometry data set: OGLE-II BVI maps and I-band data *Acta Astronomica* **55** 43-57
- [14] VaST 2017 *VaST (Variability Search Toolkit): A software for variable star detection on a series of astronomical images* **E-source:** <http://scan.sai.msu.ru/vast/> (Accessed 16 Feb. 2018)
- [15] AsPyLib 2015 *AsPyLib: a python library for FITS image processing* **E-source:** <http://www.aspylib.com/> (Accessed 19. Feb. 2018)
- [16] Lomb N R 1976 *Ap&SS* **39** 447
- [17] Palmer D M 2009 A fast Chi-Squared technique for period search of irregularly sampled data *Astrophysical Journal* **695**(1) 496-502 **E-source:** <http://adsabs.harvard.edu/abs/2009ApJ...695..496P> (Accessed 15. Feb. 2018)
- [18] Reegen P 2007 SigSpec. I. Frequency- and phase-resolved significance in Fourier space *Astronomy and Astrophysics* **467** 1353-71
- [19] Akerlof C, Alcock C, Allsman R, Axelrod T, Bennett D P, Cook K H, Freeman K, Griest K, Marshall S, Park H-S, Perlmutter S, Peterson B, Quinn P, Reimann J, Rodgers A, Stubbs C W, Sutherland W 1994 Application of cubic splines to the spectral analysis of unequally spaced data *Astrophysical Journal* **436**(2) 787-94 **E-source:** http://articles.adsabs.harvard.edu/cgi-bin/nph-iarticle_query?db_key=AST&bibcode=1994ApJ...436..787A&letter=&classic=YES&defaultprint=YES&whole_paper=YES&page=787&epage=787&send=Send+PDF&filetype=.pdf (Accessed 6. Feb. 2018)
- [20] Welch D L, Stetson P B 1993 Robust variable star detection techniques suitable for automated searches - New results for NGC 1866 *Astrophysical Journal* **105**(5) 1813-21 **E-source:** <http://adsabs.harvard.edu/full/1993AJ....105.1813W> (Accessed 15 Feb. 2018)
- [21] Stetson P B 1987 DAOPHOT - A computer program for crowded-field stellar photometry *Astronomical Society of the Pacific* **99** 191-222 **E-source:** http://articles.adsabs.harvard.edu/cgi-bin/nph-iarticle_query?db_key=AST&bibcode=1987PASP...99..191S&letter=&classic=YES&defaultprint=YES&whole_paper=YES&page=191&epage=191&send=Send+PDF&filetype=.pdf (Accessed 19. Feb. 2018)

AUTHORS

Sergey Bratarchuk, 7 September 1976, Riga, Latvia



Current position, grades: RTU MTAF AERTI Doc, Dr.Sc.Ing
University studies: Riga Technical University
Scientific interest: Computer vision, Signal processing, Non-destructive testing

Zlata Potilicina, 14 February 2000, Riga, Latvia



Current position, grades: Longenesis, developer
Scientific interest: data mining, computer vision, machine learning