Education and Heritage in the era of Big Data in Astronomy
Proceedings IAU Symposium No. 367, 2020
R. M. Ros, B. Garcia, S. R. Gullberg, J. Moldon & P. Rojo, eds.
doi:10.1017/S1743921321000028

Processing methods and approaches for the analysis of images of the eclipsed solar corona taken during campaigns with the participation of amateur astronomers

A. Stoev[1], P. Stoeva[1], S. Kuzin[2], M. Kostov[1] and A. Pertsov[2]

[1]Space Research and Technology Institute – Bulgarian Academy of Sciences, Stara Zagora Department, Stara Zagora, Bulgaria

[2]Lebedev Physical Institute, Russian Academy of Sciences, Russia, Moscow

Abstract. The increase in the amount of scientific information in heliophysics is related to both quantitative – increasing the number of high-power telescopes and the size of light receivers coupled to them, and qualitative reasons – new modes of observation, large-scale and multiple studies of the solar corona in different ranges, large-scale numerical experiments to simulate the evolution of various processes and formations, etc. The paper discusses the role and importance of methods for processing images of the solar corona, the store of obtained "raw" data and the need to access high-performance computing systems in order to obtain scientific results from the observational experiments, the need of international collaboration and access to the data in the era of increase in the amount of scientific information in heliophysics.

Keywords. Sun: corona, eclipses, techniques: image processing

Introduction: The brightness of the corona is a million times less than the brightness of the photosphere and is approximately equal to the surface brightness of the Moon. Because the Moon is not visible near the Sun, the solar corona can only be observed with the naked eye during the full phase of solar eclipses. Outside the eclipses, the corona is observed from the Earth's surface with the help of special telescopes – coronagraphs. The data obtained from observations of total solar eclipses (TSE) are usually "raw" observational materials, from various ground observations and experiments in the path of totality. Large amounts of data are also accumulating in the field of climatology, which studies the response of the Earth's atmosphere to a sudden deficit of solar radiation during the full phase. The size of astronomical radiation receivers – mainly CCDs – is growing significantly faster than the size and even the number of telescopes themselves. And the increase in the number of pixels in a CCD receiver is a proportional increase in the amount of information received from it. In heliophysics, scientific problems have arisen that require obtaining a large number of consecutive images of the same area of the corona with different exposures in search of rapid variability of its dynamic and kinematic characteristics. It is also important that all information obtained from such observations be processed immediately and added to the heliophysical database.

Big data in heliophysics – problems and approaches: The increase in the amount of scientific information in heliophysics is related to both quantitative – increasing the

number of high-power telescopes and the size of light receivers coupled to them, and qualitative reasons – new modes of observation, large-scale and multiple studies of the solar corona in different ranges, large-scale numerical experiments to simulate the evolution of various processes and formations, etc. (Zhang & Zhao 2015). Typically, methods for processing images of the solar corona include the selection of characteristics and boundaries, as well as various types of filtering and conversion of histograms. In the case of images of the corona taken during the TSE, the task is to extract the characteristics and various structures of the corona, such as arcs, helmets, arches, loops and especially their anomalous development and connections with the underlying active photospheric and chromospheric formations. The task of determining their characteristics, as a rule, comes down to identifying the boundaries of these structures and their interaction. Another feature of modern heliophysical experiments is the need to store the obtained "raw" data and the need to access high-performance computing systems in order to obtain scientific results from the observational experiment and to analyze them. They require hundreds of gigabytes to preserve a "snapshot" of the solar corona. Such calculations are performed on distributed clusters with thousands of processors and the ability to work with such data in the database allows us to trace the history of the evolution of individual objects in the solar corona (Farivar *et al.* 2013). Heliophysical data obtained during expeditions for observations of total solar eclipses are completely heterogeneous – they are obtained with different tools using different methods, different observers and teams choose a completely different approach to different objects and processes in the solar corona.

Working with heterogeneous and distributed data is becoming an important problem of heliophysical science. The solution to such situations will require a global initiative and cooperation to develop universal standards for annotation, search and access to heliophysical data obtained during the TSE. Within this framework, centralized repositories are likely to be set up to store information for heliophysical data archives. With the help of standard protocols, it will then be possible to select from them those that contain information about the required coronal area, heliophysical event, observational experiment, exposures, filters, observation coordinates, etc. A characteristic feature of the current situation related to observations of the Sun during TSE is that the data become so diverse and so complex that they themselves cease to be valuable. The only important thing is to know how to analyze this data to get a scientifically significant result.

Stages of developmen of universal standards for annotation, search and access to heliophysical data obtained during TSE: Primary image processing, Astrometric image calibration, Photometric calibration of images, Match images, Creating a catalog. The ultimate goal of this international collaboration is to explore and develop a horizontally scalable communication network for corona imaging during total solar eclipses, which will be available to researchers around the world to process large amounts of raw data from current and future observations of TSE. The architecture of the communication network should provide ample opportunities for customization and modification of algorithms used at all stages of research. In this regard, the need to strengthen international cooperation of professionals and amateur astronomers is evident as well as the standardization of the observational experiments in the field of photometry and spectrometry of the solar corona during TSE, the creation of a database of observations for the purpose of further processing and interpretation. Training of amateur astronomers, organization of international schools and conferences, publication of specialized literature, Web-based teaching aids, instructions and consultations are also needed for the extensive research of this wonderful magnificent natural phenomenon – a total solar eclipse.

https://doi.org/10.1017/S1743921321000028 Published online by Cambridge University Press

References

Zhang, Y. & Zhao, Y. 2015, Data Science Journal, 14, 11
Farivar, R., Brunner, R. J., Santucci, R., et al. 2013, Astronomical Data Analysis Software and Systems XXII, ASP Conf. 475, 91

https://doi.org/10.1017/S1743921321000028 Published online by Cambridge University Press