RESUMEN DE LA TESIS PARA OPTAR AL
GRADO DE: Doctor en Ingeniería Eléctrica
POR: Pablo Andrés Huijse Heise
FECHA: 31/05/2014
PROFESOR GUÍA: Pablo Antonio Estévez Valencia.

FINDING PERIODICITIES IN ASTRONOMICAL LIGHT CURVES USING INFORMATION THEORETIC LEARNING

The analysis of time-variable astronomical phenomena is of great interest as it helps to improve our understanding of the structure and topology of our Universe, the mechanisms of galaxy and stellar evolution, etc. The basic tool to study variability in the sky is the light curve. Light curves are time series of stellar brightness and their analysis reveals key information about the physics behind the variable phenomena. Periodic variable stars are particularly interesting. Periodic variable stars are used to estimate the size and distance-scales of our Universe, and the period itself is key in the estimation of other stellar parameters (mass, radius), stellar classification and exoplanet detection. The precise estimation of the period is critical in order to accomplish these scientific tasks. Astronomy is experiencing a paradigm change due to the extent volumes of data generated by current astronomical surveys. In less than 10 years, hundreds of Petabytes of astronomical images and time series catalogs will be produced. Conventional astronomy does not possess the tools required for this massive data mining operation. Nowadays there is a growing need for methods with solid statistical background to do automatic astronomical time series analysis. These methods need to be robust, fully-automated and computationally efficient.

In this doctoral research I have developed methods for periodicity detection and period estimation in light curves that are based on information theoretic concept of correntropy and advanced signal processing techniques. These methods are intended for automatic and efficient periodic light curve discrimination in large astronomical databases. Correntropy is a generalization of the conventional correlation to higher order statistics. In this thesis I propose the slotted correntropy estimator, the correntropy kernelized periodogram (CKP) and the correntropy non-negative matrix factorization spectrum (CNMFS). The slotted correntropy extends correntropy to unevenly sampled time series such as light curves. The CKP is a generalized periodogram that can be computed directly from the samples without regards on their sampling. The CNMFS is a high resolution spectrum that is localized on the fundamental frequency of the process.

The results presented in this thesis show that information theoretic based criteria perform better than conventional methods used in astronomy such as the LS periodogram, analysis of variance, string length and the slotted autocorrelation function (second-order methods). Including the higher-order moments of the time series into the estimation makes the proposed information-theoretic methods more robust against noise and outliers, giving them the upper hand in term of the precision of the detected periods. The proposed methods are also general as they do not pose any assumption on the underlying periodic signal (e.g. sum of sine-waves), and can be adapted heuristically (CKP) or automatically (CNMFS) to different periodic light curve shapes. The proposed methods are less prone to return a harmonic, sub-harmonic or an alias of the underlying period, a usual problem with conventional methods. The results also show that the proposed methods are more robust and less dependent on the number of samples and the time span of the light curve, i.e. the period can be recovered even if few samples or only a short piece of the light curve is available. This suggests that these methods may outperform conventional methods for early or online periodicity discrimination on surveys that are currently operating (VVV, DECam).