

The Kohonen Map: Chaos to Order in 17000 Easy Steps

D. Brett, R. West, P. Wheatley
University of Leicester, UK

Abstract

Successful use of the Kohonen Self-Organising Map scheme in the area of lightcurve classification is presented along with preliminary results of the method applied to analysis of the XMM-Subaru x-ray and optical-counterpart deep-field data.

1. INTRODUCTION

With the advent of tera-scale datasets the use of automated analysis systems, for handling tasks such as the classification of thousands or even millions of lightcurves, is becoming a tool we can no-longer do without. It is with this aim that we have applied the Kohonen Self-Organising Map (from hereon referred to as the KSOM) scheme [2] to the task of grouping, or clustering, lightcurves of similar shape on a 2-D grid (which will be referred to as map-space). From this representation of the dataset it is then possible to find cluster boundaries using the U-Matrix method [3], and thus extract information about similar lightcurve shapes from the map.

2. METHOD AND RESULTS

The KSOM works through a process of competitive learning, allowing variations in the dataset to express themselves across an n-dimensional space (in this case 2-D for ease of visualisation). The greater the distance in map-space between two data the greater the difference between their underlying shapes, thus areas on the map in close proximity to each other can represent slight variations within a class of lightcurve. In the case of lightcurves, such as the ROTSE dataset [1], the parameters which cause variations across the map are the lightcurve shape (a set of values at X phase-offsets which describe a shape), but in the case of data like the XMM-Subaru dataset these parameters are distinct and non-related, such as morphological classification or a colour ratio. In this way the KSOM is a very adaptable algorithm.

Using this method both the ROTSE and a subset of the XMM-Subaru data have been analysed. In the case of the ROTSE singly periodic lightcurves the network performs admirably, distinguishing clearly with little overlap between the various clusters (see figure 1). It is only in those classifications of star where the underlying physical process, and hence morphology of the lightcurve, are the same that there is clear overlap. Indeed this shows further description of the data via other parameters, such as oscillation period of the lightcurve, may be required to extract the complete classifications from the data.

The XMM-Subaru data, though quite a small dataset, has had the KSOM algorithm applied to it. It has been found that there are groupings apparent in the data (in this instance the grouping parameters were such things as x-ray to optical brightness and colour ratios as opposed to normalised-magnitudes at different phase-offsets, as in the case of the ROTSE data).

3. CONCLUSIONS

It has been shown that the Kohonen Self-Organising Map can be a very useful tool in the challenging task of large data-set classification, such as with stellar lightcurves and that the algorithm is applicable to other astronomical data clustering tasks, such as with the XMM-Subaru data subset.

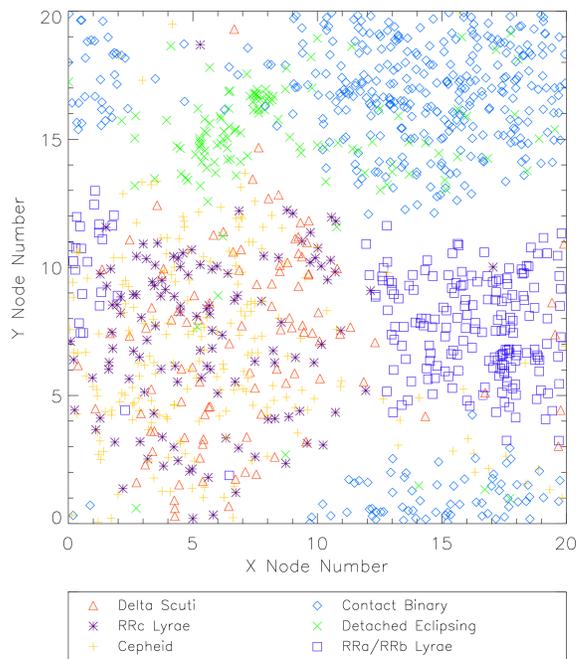


FIGURE 1: The final structure of a network trained using 1206 pre-classified ROTSE lightcurves.

REFERENCES

- [1] Akerlof, C. et al. *AJ*, 119, 190-1913
- [2] Kohonen, T. *Proc. IEEE* 78, 1464-1480
- [3] Ultsche, A. & Siemon, H.P. *Proc. Int. Neural Networks Conf. (Dordrecht: Kluwer)*, 305.