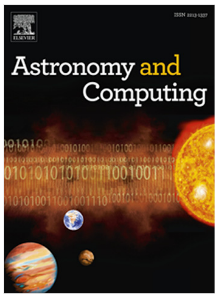
Accepted Manuscript



Vizic: A jupyter-based interactive visualization tool for astronomical catalogs

W. Yu, M.C. Kind, R.J. Brunner

PII: S2213-1337(16)30150-0
DOI: http://dx.doi.org/10.1016/j.ascom.2017.06.004
Reference: ASCOM 195

To appear in: *Astronomy and Computing*

Received date : 29 December 2016
Accepted date : 26 June 2017

Please cite this article as: Yu, W., Kind, M.C., Brunner, R.J., Vizic: A jupyter-based interactive visualization tool for astronomical catalogs. *Astronomy and Computing* (2017), http://dx.doi.org/10.1016/j.ascom.2017.06.004

Vizic: A Jupyter-based Interactive Visualization Tool for Astronomical Catalogs

Weixiang Yu[a,b,*], Matias Carrasco Kind[b,c], Robert J. Brunner[c,b]

[a] *Department of Physics, University of Illinois at Urbana-Champaign, Urbana IL, USA*
[b] *National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign, Urbana IL, USA*
[c] *Department of Astronomy, University of Illinois at Urbana-Champaign, Urbana IL, USA*

Abstract

The ever-growing datasets in observational astronomy have challenged scientists in many aspects, including an efficient and interactive data exploration and visualization. Many tools have been developed to confront this challenge. However, they usually focus on displaying the actual images or focus on visualizing patterns within catalogs in a predefined way. In this paper we introduce *Vizic*, a Python visualization library that builds the connection between images and catalogs through an interactive map of the sky region. *Vizic* visualizes catalog data over a custom background canvas using the shape, size and orientation of each object in the catalog. The displayed objects in the map are highly interactive and customizable comparing to those in the observation images. These objects can be filtered by or colored by their property values, such as redshift and magnitude. They also can be sub-selected using a lasso-like tool for further analysis using standard Python functions and everything is done from inside a Jupyter notebook. Furthermore, *Vizic* allows custom overlays to be appended dynamically on top of the sky map. We have initially implemented several overlays, namely, Voronoi, Delaunay, Minimum Spanning Tree and HEALPix grid layer, which are helpful for visualizing large-scale structure. All these overlays can be generated, added or removed interactively with just one line of code. The catalog data is stored in a non-relational database, and the interfaces have been developed in JavaScript and Python to work within Jupyter Notebook, which allows to create customizable widgets, user generated scripts to analyze and plot the data selected/displayed in the interactive map. This unique design makes *Vizic* a very powerful and flexible interactive analysis tool. *Vizic* can be adopted in variety of exercises, for example, data inspection, clustering analysis, galaxy alignment studies, outlier identification or just large scale visualizations.

Keywords: Jupyter, Python, Visualization, catalogs, large-scale structure of universe - methods: numerical

1 INTRODUCTION

For the past decades, conducting large-area sky surveys became one of the most powerful approaches to carry out observations within the astronomy community in order to understand the Universe. The Sloan Digital Sky Survey (SDSS) (York et al., 2000) set the foundations for such surveys followed by others, such as DES and Pan-STARRS (The Dark Energy Survey Collaboration, 2005; Kaiser et al., 2002). Modern telescopes and world-class supercomputing facilities produced exceptionally good images and exceptionally precise measurements of astrophysical sources, challenging astronomers to face unprecedentedly massive datasets. Moreover, future large sky, for example, LSST and Euclid (Tyson, 2002; LSST Science Collaboration et al., 2009; Laureijs et al., 2012), will obtain data of orders of magnitude larger than what we have today. These massive datasets place a new challenge on astronomers in the light of data exploration and visualization.

Various groups and individuals have tried to confront this challenge by developing new astronomical applications that are capable of displaying and visualizing large datasets. Notable examples include *Aladin Lite* (Boch and Fernique, 2014), *VisiOmatic* (Bertin et al., 2015), *ASCOT* and *Toyz* (Moolekamp and Mamajek, 2015), to mention just a few. The majority of these softwares are web-based applications that focus on the visualization of high-resolution images with additional functionalities of catalog over-plotting or image analysis. Meanwhile, *GLUE* (Beaumont et al., 2015), a desktop application that specializes in visualizing selection propagations between different plots of the same dataset, has became a popular tool for exploring hidden patterns within catalogs and images. Unfortunately all these mentioned tools, while being very powerful, have some drawbacks when it comes to the study of large-scale structure and flexible customization. Image-focused applications are limited by the fundamental characters of photometric observations, whereas images produced by sky surveys are fixed in terms of compositions, in other words, the sources are immoveable. In this regard, catalogs are more flexible considering astronomical sources in a catalog can be easily sorted and filtered. However, tools that are more friendly with tabular data are usually less good at interactive visualizations of geospatial data. They either render data points into simple scatter plots or summarize data distribution into pixels like with *VaeX* (Breddels, 2016), which lacks rich interactions with individual object while providing a very powerful way to visualize distributions of extremely large catalogs. The regarding limitations of existing softwares in addition to emerging new technologies for

*wyu16@illinois.edu