

Radio Star Candidates from FIRST and 2MASS Databases

Yanxia Zhang and Yongheng Zhao

¹National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100012, China
email: zyx@bao.ac.cn; yzhao@lamost.org

Abstract. We positionally cross-identified FIRST (the Faint Images of the Radio Sky at Twenty centimeters) catalogue with 2MASS (the Two Micron All Sky Survey) pointed-source database and collected the data from radio band and near infrared band. Then the data were cross-matched with the Véron-Cetty & Véron 2006 catalog and the Tycho-2 catalog, respectively. Therefore the known samples of quasars and stars are obtained. We applied principal component analysis (PCA) on the known sample. The overall sample may be projected in the principal component space. From the space, we can easily locate the area that radio stars occupy, and select out radio star candidates. With the follow-up observation of these candidates, the properties of radio stars may be studied.

Keywords. techniques: miscellaneous; methods: statistical; methods: data analysis; astronomical data bases: miscellaneous

1. Data Sample and Chosen Attributes

We cross-matched the Two Micron All Sky Survey (2MASS) catalogue with the Faint Images of the Radio Sky at Twenty centimeters (FIRST) catalogue within 5 arcsec radius, and obtained 153135 entries with one to one matching between the FIRST and 2MASS catalogues. The entries were then cross-identified with the Véron-Cetty & Véron 2006 catalog and the Tycho-2 catalog within 5 arcsec radius, respectively. Similarly, we obtained 2389 quasars and 1353 stars from the 2MASS and FIRST catalogues. The chosen attributes from different bands are $\log F_{\text{peak}}$ (F_{peak} : peak flux density at 1.4 GHz), $\log F_{\text{int}}$ (F_{int} : integrated flux density at 1.4 GHz), f_{maj} (fitted major axis before deconvolution), f_{min} (fitted minor axis before deconvolution), f_{pa} (fitted position angle before deconvolution), $j - h$ (near infrared index), $h - k$ (near infrared index), $k + 2.5 \log F_{\text{int}}$, $k + 2.5 \log F_{\text{peak}}$, $j + 2.5 \log F_{\text{peak}}$, $j + 2.5 \log F_{\text{int}}$.

2. Method

Principal component analysis (PCA) is a statistical method that permits the determination of the minimum number of independent or uncorrelated variables underlying a larger number of observed variables. Thus, PCA is used as a technique for both data compression and analysis (Zhang & Zhao, 2003 and reference therein).

3. Result and Discussion

We applied PCA on the known sample, and plotted them in the principal component space shown in Fig. 1. In order to clearly to see the difference between stars and quasars, we gave the distribution of the first principal component (PC1) and the second principal

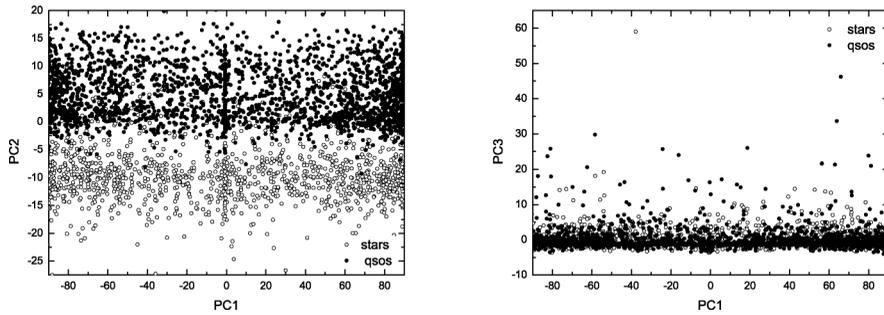


Figure 1. The distribution of the known stars (open circles) and quasars (filled circles) in the principal component space.

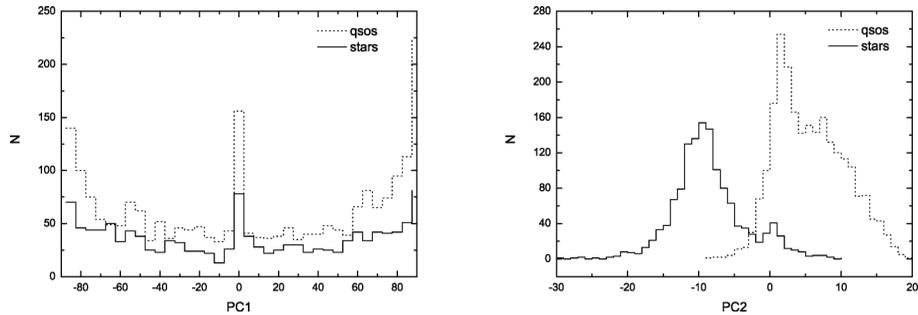


Figure 2. The distribution of the known stars and quasars as a function of PC1 (left panel) and PC2 (right panel), respectively.

component (PC2), as shown in Fig. 2. It is obvious from Fig. 1 and Fig. 2 that PC2 is the important factor to separate quasars from stars and has a good cutoff value, i.e. $PC2 = -2.6$. As a result, the objects are identified as stars when $PC2 < -2.6$, and as quasars when $PC2 > -2.6$. The classification result is indicated in Table 1. The number of misclassified stars and quasars is 162 and 58, separately. The accuracy of stars and quasars adds up to 88.0% and 97.6%, respectively. The overall accuracy amounts to 94.1%. So PCA as unsupervised method is an effective approach used for the classification problem in astronomy. If we want to obtain radio star candidates from 2MASS and FIRST catalogues, we can project the overall sample into the principal component space, and select the sources whose $PC2 < -2.6$ as radio star candidates, similarly choose those whose $PC2 > -2.6$ as quasar candidates.

Acknowledgements

This paper is funded by National Natural Science Foundation of China under grants No.10473013, No.10778724 and No.90412016.

References

Zhang, Yanxia & Zhao, Yongheng 2003, *PASP* 115, 1006