SkICAT: A Cataloging and Analysis Tool for Wide Field Imaging Surveys

N. Weir (Caltech), U.M. Fayyad (JPL), S.G. Djorgovski (Caltech), J. Roden (JPL)

We describe an integrated system, SkICAT (Sky Image Cataloging and Analysis Tool), for the automated reduction and analysis of the Palomar Observatory-ST ScI Digitized Sky Survey. The Survey will consist of the complete digitization of the photographic Second Palomar Observatory Sky Survey (POSS-II) in three bands, comprising nearly three Terabytes of pixel data. SkICAT applies a combination of existing packages, including FOCAS for basic image detection and measurement and SAS for database management, as well as custom software, to the task of managing this wealth of data. One of the most novel aspects of the system is its method of object classification. Using state-of-the-art machine learning classification techniques (GID3* and O-BTree), we have developed a powerful method for automatically distinguishing point sources from non-point sources and artifacts, achieving comparably accurate discrimination a full magnitude fainter than in previous Schmidt plate surveys. The learning algorithms produce decision trees for classification by examining instances of objects classified by eye on both plate and higher quality CCD data. The same techniques will be applied to perform higher-level object classification (e.g., of galaxy morphology) in the near future. Another key feature of the system is the facility to integrate the catalogs from multiple plates (and portions thereof) to construct a single catalog of uniform calibration and quality down to the faintest limits of the survey. SkICAT also provides a variety of data analysis and exploration tools for the scientific utilization of the resulting catalogs. We include initial results of applying this system to measure the counts and distribution of galaxies in two bands down to $B_j \approx 21$ mag over an $\sim 70$ square degree multi-plate field from POSS-II. SkICAT is constructed in a modular and general fashion and should be readily adaptable to other large-scale imaging surveys.