Astronomical Data Analysis Software and Systems XIX
ASP Conference Series, Vol. 434, © 2010
Y. Mizumoto, K.-I. Morita, and M. Ohishi, eds.

The Spitzer Source List

Harry I. Teplitz, Peter Capak, Timothy Brooke, Sachin Shenoy, Carolyn Brinkworth, Vandana Desai, Iffat Khan, and Russ Laher

Spitzer Science Center, MS 100-22, Caltech, Pasadena, CA 91125, USA

Abstract. The Spitzer Science Center will produce a source list (SL) of photometry for a large subset of imaging data in the Spitzer Heritage Archive (SHA). The list will enable a large range of science projects. The primary requirement on the SL is very high reliability, with areal coverage, completeness and limiting depth being secondary considerations. The SHA at the NASA Infrared Science Archive (IRSA) will serve the SL as an enhanced data product. The SL will include data from the four channels of IRAC (3-8 microns) and the 24 micron channel of MIPS. The Source List will include image products (mosaics) and photometric data for Spitzer observations of about 1500 square degrees and include around 30 million sources. We describe the plans and timeline for development of the Spitzer Source List. We demonstrate the verification of the Source List pipeline using Spitzer Legacy catalogs at "truth tables". Finally, we discuss the range of use cases which will be supported.

1 Introduction

The Spitzer Science Center (SSC) will produce a source list (SL) of photometry for a large subset of imaging data obtained by the IRAC and MIPS instruments onboard the *Spitzer Space Telescope* (Werner et al. 2004) during its recently-completed cryogenic mission. The SL will enable a large range of science projects, but is not intended to meet the standards of a mission wide catalog (e.g. 2MASS). The primary requirement on the SL is very high reliability, even at the cost of completeness. The Spitzer Heritage Archive (SHA) at the NASA/IPAC Infrared Science Archive (IRSA) will serve the SL as an enhanced data product and will ensure that appropriate caveats and warnings are prominently placed. The SL products are planned for public release in approximately April 2011.

This SL will ensure high reliability by including only a subset of Spitzer data that is well behaved and can be processed and verified autonomously. This means observing programs must meet a minimum set of requirements for processing by the SL pipeline. These minimum requirements include data obtained with IRAC channels 1-4 in high-dynamic range or mapping mode and MIPS channel 1 (24 microns) in scan or photometry mode; it will not include any other observing modes such as IRAC sub-array, Moving Object mosaics, MIPS SED, MIPS total power, IRS peak-up imaging, nor MIPS 70 or 160 micron imaging. The SL pipeline will mask the most difficult regions within otherwise tractable mosaics.

The pipeline will create new mosaics that span observing programs and employ a limited set of namelists specific to the most common observing modes and strategies. These new mosaics will improve upon the standard mosaics supplied in the lights out SSC pipeline.

Sources will be extracted from these new mosaics using a combination of MOPEX/APEX (Makovoz & Marleau 2005) and SExtractor (Bertin & Arnouts) software and each channel being separately processed. The SL will provide photometry for isolated, high significance ($> 10\sigma$) point like sources with good (> 4) imaging coverage. Confused sources, low significance sources ($< 10\sigma$), extended sources, and moving targets will be excluded. The cleaned list will be served by the SHA as a searchable list of objects with detections in multiple bands merged into a single source by closest positional match, and multiple matches flagged. An object entry will include either measured photometry or upper-limits measured in all bands. The quality of the SL will be ensured, in part, by comparison with the Spitzer "legacy" surveys, for which PI teams delivered catalogs of well-validated photometry.

2 Requirements

Final SL products will be required to be easily searchable and highly documented. An important goal of the project is to minimize misuse of the list. No requirement will be placed upon completeness, but validation tests show that we can expect 80-90% complete compared to Legacy surveys at 10σ. Typically legacy products go deeper, and we expect only 50-70% compared to full Legacy-survey catalogs (e.g. c2d class A+B sources down to 5σ).

The driving requirement on the SL pipeline shall be the need to ensure high reliability of reported sources. Requirements will be set separately for "extragalactic" and "galactic" fields. The primary difference between the fields is the complexity of the background (together with source crowding). Galactic fields will be defined as those with E(B-V)> 0.2, or with a specific applicable description in the program abstract. The SL shall meet reliability level achieved by Legacy teams:

- For extragalactic fields, a maximum of 0.01% spurious sources will be included
- For galactic fields without structured background, a maximum of 0.05% spurious sources when detected at 10 sigma in two bands
- For galactic fields with structure in the background, a maximum of 0.2% spurious sources when detected at 10 sigma in two bands

In the case of MIPS-only fields (about half of MIPS data), these requirements will be relaxed to be: Extragalactic: 0.05%; Galactic, low BG: 0.1% (appx, still under study); Galactic, structured BG: 0.25% (appx, still under study). Finally, the case of extremely crowded galactic fields is still under study, and some of these fields may have to be excluded.

3 Pipeline

The SL pipeline uses the standard MOPEX/APEX and SExtractor software. Namelists have been created for a number of specific use cases (extragalactic shallow, medium deep; galactic structured background; galactic simple background; extreme crowding; etc). Initial processing extracts many sources and then a quality assurance (QA) module rejects spurious objects. Final lists for each channel are combined (bandmerged) using nearest-neighbor matching; the existence of multiple possible matches will be reported.

The most difficult part of the SL pipeline processing is ensuring high reliability. This task falls largely into three categories – artifact mitigation, masking of difficult regions, and rejection of spurious sources during the QA module. Common data artifacts are well characterized in the documentation available at the SSC website[1]. The pipeline includes the masking of significant area (typically up to 5%) within otherwise well behaved mosaics. Specifically, regions are bright stars, extended objects, or near large data artifacts (such as mux-bleed) are rejected. Finally, the QA module excludes sources that are too compact or too extended. Cosmic rays are too compact; galaxies and false detections within extended emission are too extended. Specifically, for the concentration is defined by: IRAC – the peak pixel flux divided by aperture flux; and MIPS – the central flux (1 PSF) divided by extracted flux. The QA module rejects sources with FWHM $> 2\times$ that of the PSF. It also rejects a source if peak pixel is not within positional uncertainty (0.9 arcsec) of the centroid.

4 Validation

Test cases were selected from the SWIRE (extragalactic; Londsdale et al. 2003) and c2d (galactic; Evans et al. 2003) legacy surveys.

The SWIRE case has low coverage and sets a good baseline for what can be achieved in fields without structured background or extreme crowding. In a 0.2 square degree test mosaic, the SL pipeline achieves results comparable to the data products supplied by the Legacy team. The spurious source rate is $< 0.01\%$. In fact, measurement the rate is limited by the small number of total sources – that is, there are almost no spurious sources in the extracted list. The completeness is better than 80% relative to the Legacy catalog. The SL excludes regions that the Legacy catalog does not; in addition, the Legacy catalog goes down to about 5σ, though it imposes a two-band detection requirement that the SL does not. Some differences remain, however, due to photometric scatter and differing background subtraction algorithms.

The c2d Legacy survey provides typical fields for comparison in the galactic case. The SL pipeline was tested against both structured and unstructured backgrounds. Again, the requirements were met: $< 0.1\%$ spurious sources in 0.4 square degrees, with $> 80\%$ completeness with respect to c2d "class A" sources.

4.1 Science Validation

The best validation of the SL will come from its use for science analysis. As a test, several published results were reproduced using SL photometry. As an example here, we show the selection of AGN based upon their mid-IR color. Dust heated by an AGN to several hundred K (or more) will dominate the 3-6 micron flux. Color-color selection can efficiently identify such sources from broad-band photometry of relatively bright galaxies. In the figure, we show the IRAC color-color plot for galaxies measured in 0.2 square degrees of the COSMOS field, and

[1] http://ssc.spitzer.caltech.edu/

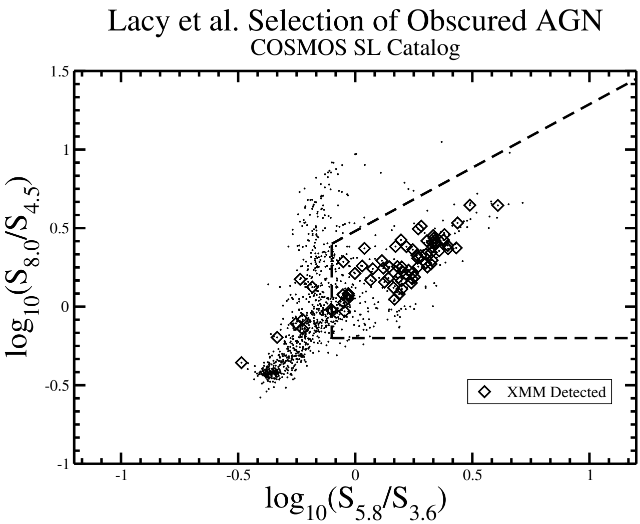

Figure 1. We show the Lacy et al. (2004) AGN selection box applied to objects in the COSMOS field extracted by the Source List pipeline. Diamond symbols indicate objects detected by XMM.

the AGN-selection criteria defined by Lacy et al. (2004). The AGN, as defined by XMM detection, predominantly fall within the selection box.

Further validation will be performed at SSC after the SL pipeline runs through the archive to create a beta version of the products. Final processing will occur after the validation phase.

Acknowledgments. This work is based in part on observations made with the *Spitzer Space Telescope*, which is operated by the Jet Propulsion Laboratory, California Institute of Technology under NASA contract 1407. Support for this work was provided by NASA through an award issued by JPL/Caltech.

References

Bertin, E., & Arnouts, S. 1996,A&AS, 117, 393
Evans, N. J., II, et al. 2003,PASP, 115, 965
Lacy, M., et al. 2004,ApJS, 154, 166
Lonsdale, C. J., et al. 2003,PASP, 115, 897
Makovoz, D., & Marleau, F. R. 2005, PASP, 117, 1113
Werner, M. W., et al. 2004, ApJS, 154, 1