Optimal Bandmerging versus Simple General Source Association

Russ R. Laher and John W. Fowler

Spitzer Science Center (SSC), California Institute of Technology, Pasadena, CA 91125

Contact: laher@ipac.caltech.edu

ADASS XVII, London, UK, September 23-26, 2007

ABSTRACT

For collating point-source flux measurements derived from multiple infrared passbands of Spitzer Space Telescope data -- e.g., the four wavelength passbands of the Infrared Array Camera (IRAC) and the three wavelength passbands of the Multiband Imaging Photometer for Spitzer (MIPS) -- it is best to use the Spitzer Science Center's "bandmerge" software, rather than the relatively simple method of general source association (GSA). This type of processing is useful for making catalogs detailing measurements at multiple wavelengths for a specific object. The former method uses both source positions and positional uncertainties to form a chi-squared statistic that can be thresholded for optimal matching, while the latter method finds nearest neighbors across bands that fall within a user-specified radius of the primary source. Our assertion is supported by our study of completeness (C) versus reliability (R) for the two methods, which involved matching MIPS 24 μm and IRAC 3.6 μm point sources in the SWIRE-survey Chandra Deep Field South. In this study, completeness is defined as the number of true matches divided by the total number of sources with detections in both passbands, and reliability is defined as one minus the number of false matches divided by the total number of sources (regardless of whether they have detections in both passbands). Both methods can achieve C = 98%, but with R = 92.7% for GSA versus R = 97.4% for bandmerge. With almost a factor of three lower in unreliability, bandmerge is the clear winner of this comparison. High reliability is important because astronomers study objects with peculiar properties, which are likely to have false matches.

Keywords: image processing, point sources, bandmerging, general source association, SWIRE Chandra Deep Field South, Spitzer Space Telescope, infrared, astronomy

MERGING METHODS STUDIED IN THIS SHOOTOUT

The "bandmerge" software program, which was developed at the Spitzer Science Center (SSC), uses decision theory to perform optimal merging of point sources from up to seven different Spitzer passbands (four IRAC wavelengths and three MIPS wavelengths). Input source positions and positional uncertainties are required in separate files for each band. This program is the centerpiece of the SSC bandmerge-GUI package.

The SSC "gsa" and "mgsa" programs perform general source association, i.e., point-source matching based only on positional proximity. The gsa program merges two point-source lists. The mgsa program iteratively executes the gsa program to merge three or more point-source lists. Only input azimuthal/elevation position coordinates are required (e.g., Right Ascension and Declination). This simple method is useful when positional uncertainties are unavailable. GSA is general and, therefore, not limited to Spitzer data.

DATA ANALYZED: CHANDRA DEEP FIELD SOUTH

In this study, we examined the case of merging point sources from two different infrared passbands of data taken by the Spitzer Space Telescope from observations of the Chandra Deep Field South. Specifically, we used 3.6-μm data from the Infrared Array Camera (IRAC) and 24-μm data from the Multiband Imaging Photometer for Spitzer (MIPS). The data were acquired as part of the Spitzer Wide-area InfraRed Extragalactic survey (SWIRE; Lonsdale *et al.*, PASP 115:897, 2003).

The Spitzer false-color mosaic image at the right, generated by the SWIRE team, was made from IRAC images of a small portion of the Chandra Deep Field South. The image is composed of three single-infrared-wavelength mosaics. Image data from IRAC passbands corresponding to 3.6 μm, 4.5 μm, and 8.0 μm are shown in false-color blue, green, and red planes, respectively.

The scene contains some bright stars, which have the six-point spider-diffraction pattern and appear blue/white in color, but is dominated by thousands of barely resolved elliptical-like galaxies, which appear blueish, and PAH-band-emitting spiral-type galaxies, which appear reddish. There are also a few asteroids, which appear a deep red color due to their cold temperatures.

Our Monte Carlo simulation employed 9,912 MIPS-24 detections that had real matches in IRAC-1 and 991 that did not (the real data appeared to have about this rate of detection in IRAC-1 of sources detected in MIPS-24) and 27,313 IRAC-1 detections with no MIPS-24 counterparts. These detection counts match the real SWIRE data.

SWIRE Chandra Deep Field South (http://swire.ipac.caltech.edu/swire/public/imageGallery/imageGallery.html)

PERFORMANCE COMPARISONS

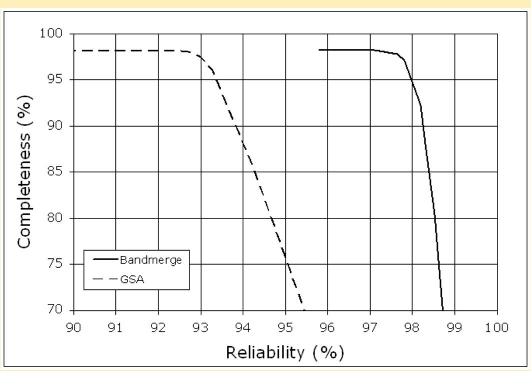

Figure 1. Completeness versus reliability for bandmerge and GSA. Three independent Monte Carlo simulations were run, resulting in negligible changes in bandmerge's superiority over GSA but fluctuations of about 0.1% in both C and R for both methods (1-σ).

NT = number of true matches	NS = number of possibly correctly matched sources , ~ 0.9 N
NF1 = number of type-1 false matches	C = NT / NS
NF2 = number of type-2 false matches	R = 1 - (NF1 + NF2) / N
N = total number of sources	

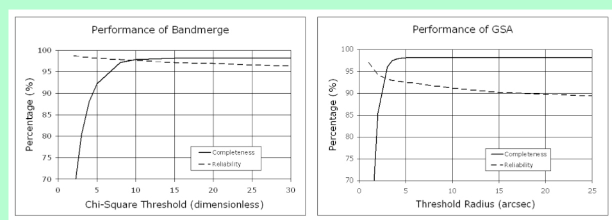

Figure 2. Variation of completeness and reliability with user-specifiable threshold, which is the dimensionless chi-square threshold in the case of bandmerge (left panel) and the search radius, in arcseconds, in the case of GSA (right panel).

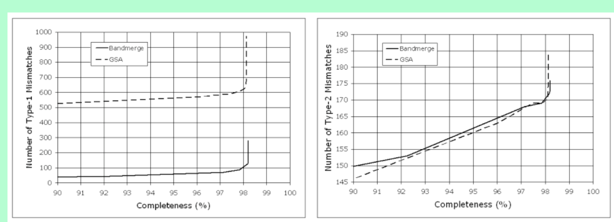

Figure 3. Defining "unreliable" here as a MIPS-24 detection that had no real IRAC-1 partner, but matched an IRAC-1 detection that had no real MIPS-24 counterpart anyway (type 1), we find that the lower rate of unreliability for bandmerge relative to GSA is the biggest difference between the two methods (left panel). The other kind of unreliable match is the MIPS-24 detection that has an IRAC-1 partner, but gets matched to the wrong IRAC-1 detection (type 2); this occurs less frequently and is a weaker function of threshold (right panel). Both methods suffer about the same from type-2 mismatches.