

The Hubble Source Catalog

Stephen H. Lubow

Space Telescope Science Institute,
3700 San Martin Drive, Baltimore, MD, U.S.A.
email: lubow@stsci.edu

Abstract. The Hubble Source Catalog (HSC) is designed to enhance the science obtained from the Hubble Space Telescope by combining the tens of thousands of visit-based source lists in the Hubble Legacy Archive (HLA) across filters and detectors into a single master catalog. The catalog contains data from the major Hubble imaging instruments: Wide Field Planetary Camera 2 (WFPC2), Advanced Camera for Surveys (ACS), and Wide Field Camera 3 (WFC3). It is based on cross matching and astrometry algorithms developed by Budavari & Lubow (2012). We recently released Version 2 that is three times the size of Version 1 and includes some new features. The catalog can be accessed through a variety of interfaces (see <http://archive.stsci.edu/hst/hsc/>). The HSC provides descriptions of astronomical objects involving multiple wavelengths and epochs. High relative positional accuracy of objects is achieved across the Hubble images, often with sub-pixel precision of a few milliarcseconds.

Keywords. astrometry, catalogs

1. Introduction

The Hubble Space Telescope (HST) has been collecting data for over 25 years and has provided a remarkably detailed view of the universe. The observations are highly diverse. HST is not generally used as a survey instrument. Instead, over 100 guest observer proposals are selected each year that cover a wide range of science objectives. A given astronomical object may have been observed with several different detectors, a range of filters, exposure times, and orientations for a variety of science objectives. In some cases, objects are covered by thousands of overlapping exposures. The primary objective of the Hubble Source Catalog (HSC) project is organizing this information into a coherent catalog. Due to the large amount of data and its complexity, the catalog production is a computationally challenging task. The effort began in 2009 with a NASA Advanced Information Systems Research Program grant to Johns Hopkins University and the Space Telescope Science Institute (STScI) and is now being supported by the STScI. After three beta versions (Budavari & Lubow 2012), Version 1 of the HSC was released in 2015 (Whitmore *et al.* 2016) and Version 2 was released in 2016. Version 2 has three times as many sources as Version 1 and includes spectroscopic cross matches. The science usage of the HSC is just getting started. The HSC is being used by the Hubble Catalog of Variables project based in Athens (see paper by Gavras, this volume).

2. Implementation

The HSC is built through a processing pipeline described in Budavari & Lubow (2012) and Whitmore *et al.* (2016). Briefly, HST images are run through Source Extractor software (Bertin & Arnouts 1996) to produce source lists that contain the properties of sources. These source lists are cross matched by position to the standard catalogs PAN-STARRS (Panoramic Survey Telescope and Rapid Response System, Stubbs *et al.* 2010),

SDSS (Sloan Digital Sky Survey, Ahn *et al.* 2014), and 2MASS (Two Micron All-Sky Survey, Skrutskie *et al.* 2006) which improves the absolute astrometry of the source lists to within about 100 milliarcseconds (mas). The cross matching is done through a histogram method that finds the most common offset. Improvements of relative astrometry are made by cross matching overlapping source lists with each other that typically results in accuracies of about 10 mas, using the algorithm developed by Budavari & Lubow (2012). The HSC provides substantially improved astrometry over what is currently available in the HST data headers. Once the astrometric corrections are complete, sources in different source lists are cross matched to determine matches that describe astronomical objects.

The HSC software is entirely based on a relational database system, Microsoft SQL Server. The catalog consists of database tables together with various functions that provide rapid information retrieval. Many of the geometric calculations are done using the JHU spherical library (Budavari *et al.* 2010). The database has a size of about 1.2 TB.

The HSC can be accessed through several user interfaces: a graphical user interface called the Discovery Portal, a MAST forms interface, the Hubble Legacy Archive image viewer that provides HSC source positions overlaid on the HST images, and an SQL query interface through CASJobs (see <https://archive.stsci.edu/hst/hsc/>).

3. Version 2 Results

Version 2 of the HSC was released in September 2016. It contains more recent data than Version 1, including four more years of ACS data and one more year of WFC3 data. The ACS source lists were reprocessed with more accurate parameters settings and go about one magnitude fainter than Version 1. As a result there are about three times as many sources in Version 2. Version 2 includes for the first time HST spectroscopic data that is cross matched with HSC sources. Users can then get spectroscopic information about HSC sources.

The information about the individual source detections in a match includes temporal and brightness data in various filters. Fig. 1 shows the distribution of the match size. It plots the number of objects (matches) than have at least the number of sources on the horizontal axis. Of the 92 millions objects in the HSC, more than 60 million that have more than a single source detection. Over 1 million objects have more than 25 source detections. Fig. 2 plots the distribution of the time spans covered by the matches. Over 10 million objects have time coverage of more than 1 year. Some matches cover more than 20 years.

The HSC provides information about the source fluxes and magnitudes through aperture photometry. We determine the accuracy of the photometry by comparing the differences in the fluxes from pairs of sources detected in the same detector and filter that reside in the same match. Ideally, there should not be any flux difference because we are comparing flux values for the same astronomical object that are in the same wavelength band and with the same detector. We define the flux difference as $f_{i,j} = |f_i - f_j| / \max(f_i, f_j)$ for sources i and j with fluxes f_i and f_j . This quantity is always less than or equal to unity. Fig. 3 plots the distribution of flux differences normalized such that the distribution function is unity at a flux difference of zero. We see that the flux errors are typically about 5 percent. The errors are somewhat larger for WFPC2.

An important byproduct of the HSC is improved astrometry. For each match, we determine a match position as the average of the sources positions that make up the match. We determine relative astrometric errors by the residual shifts in positions of sources from their match positions. The distributions of these errors in cases where there is more than one source in a match is shown in Fig. 4. The figure shows the distributions

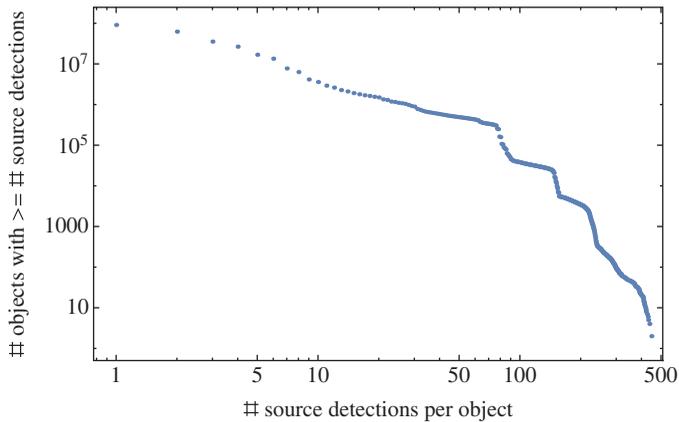


Figure 1. Plot of the number of objects (matches) than have at least the number of sources on the horizontal axis.

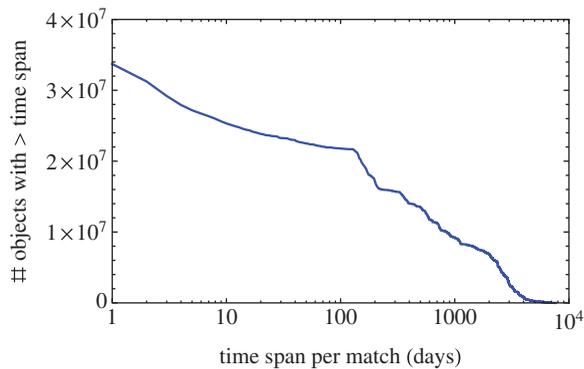


Figure 2. Plot of the number of objects (matches) than cover a time span of least the value on the horizontal axis.

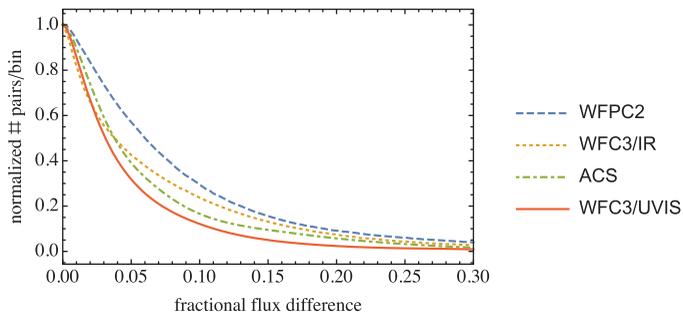


Figure 3. Distribution of flux errors by detector. WFC2 has the largest number of pairs at a fractional flux difference of 0.1, followed respectively by WFC3/IR, ACS/WFC, and WFC3/UVIS.

both before and after the HSC corrections are applied. The areas under the two curves plotted in the left panel are equal when extended to larger errors. The figure shows the striking improvement in astrometry with the HSC corrections. The peak value (mode) for the HSC astrometry is only 2.3 mas and the median value is about 10 mas. The lower plot shows the same data, but with a logarithmic vertical scale and a horizontal scale

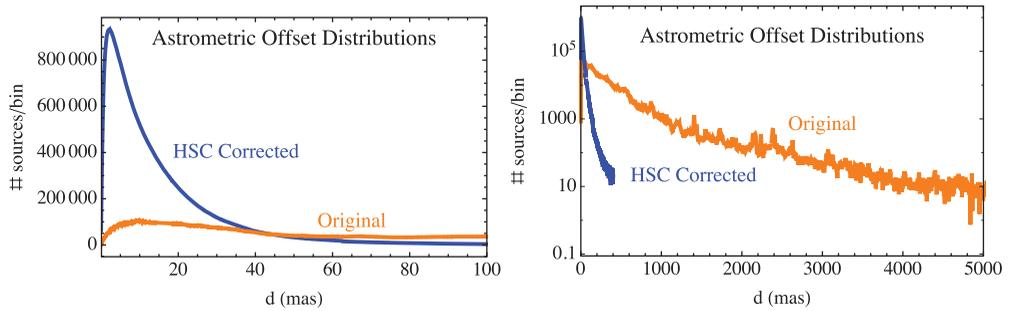


Figure 4. The distribution of relative astrometric errors for HSC sources in milliarcseconds.

that goes out to larger errors. We see that the original HSC astrometry has an extended tail that goes out to large errors that is not present in the HSC corrected astrometry.

4. Future Plans

The HSC is continuing to improve the quality of the catalog and the range of data. In the near term we plan to take advantage of the availability of the Gaia catalog for improving HSC astrometry (Gaia Collaboration *et al.* 2016).

References

- Ahn, C. P., *et al.* 2014, *ApJS*, 211, 17
 Bertin, E. & Arnouts, S. 1996, *A&AS*, 117, 393
 Budavari, T., Szalay, A. S., & Fekete, G. 2010, *PASP*, 122, 1375
 Budavari, T. & Lubow, S. H. 2012, *ApJ*, 761, 188
 Gaia Collaboration, Brown, A. G. A., Vallenari, A., *et al.* 2016, arXiv:1609.04172
 Skrutskie, M. F., *et al.* 2006, *AJ*, 131, 1163
 Stubbs, C. W., Doherty, P., Cramer, C., Narayan, G., Brown, Y. J., Lykke, K. R., Woodward, J. T., & Tonry, J. L. 2010, *ApJS*, 191, 376
 Whitmore, B. C., Allam, S. S., Budavari, T., *et al.* 2016, *AJ*, 151, 134