Version 1 of the Hubble Source Catalog

Brad Whitmore, whitmore@stsci.edu

Hubble has been in orbit for nearly 25 years, and has compiled an impressive legacy of observations. The statistics alone make the point with a dozen different instruments, roughly a hundred observing modes, several hundred different filters and gratings, tens of thousands of targets, and over a million observations. While this great volume and diversity is one of the great strengths of Hubble, it also makes it difficult to effectively use the archives in some cases, which is the motivation for the Hubble Source Catalog (HSC; https://archive.stsci.edu/hst/hsc/).

For example, imagine that you are interested in obtaining a color–magnitude diagram for all Hubble observations of the Small Magellanic Cloud (SMC). A search of the Hubble Legacy Archive (HLA; hla.stsci.edu) would show 7289 observations in this region (see Figure 1)!

After retrieving the data you would need to: (1) combine the various images; (2) perform photometry on roughly 10 million separate detections of 1 million individual objects; and (3) figure out how to match all those objects, in order to sort out duplicates and sum the data for repeat visits with the same instrumental configuration. Imagine the amount of time and effort this would take. With the HSC, you can now perform this task in minutes!

In many ways, the HSC provides entry into the world of database astronomy, by collecting useful information into a database that can be easily searched. The Sloan Digital Sky Survey (SDSS; see upper right panel in Figure 2) is leading the way in this rapidly growing field of astronomy. Taking a page from the SDSS book, the goal of the HSC is to combine the data from all the sources observed by Hubble over the years into a single master catalog.

HSC is still in its infancy, including at present only observations from Wide Field Planetary Camera 2 (WFPC2), the Wide Field Camera of the Advanced Camera for Surveys (ACS/WFC), and Wide Field Camera 3 (WFC3). Other instruments will be added in the future, including the ACS High Resolution Camera (ACS/HRC) and the Near Infrared Camera and Multi-Object Spectrometer (NICMOS).

Version 1 of the HSC, released in February 2015, is already a very powerful new tool. The purpose of this article is to introduce HSC to the community and demonstrate how it can be used for research. In particular, we highlight the expanded opportunity for archival research proposals in Cycle 23.

As demonstrated in Figure 2, the great depth and increased spatial resolution of the Hubble images provide a dramatic increase in the number of sources in a given region as compared with SDSS.

Detailed examples are provided to guide potential users in common ways to use the HSC and avoid common pitfalls (http://archive.stsci.edu/hst/hsc/help/HSC_faq.html#use_case). Results from one case study are shown in Figure 3. It shows how variable stars can be found by searching for stars with large photometric scatter in repeat observations, in this case for the dwarf galaxy IC 1613. This figure also

Figure 1: HLA footprints for a search of the SMC using a radius of 2 degrees. A color-magnitude diagram containing 385,675 data points, created by the HSC in less than 2 minutes, is shown in the upper right.

Figure 2: HLA footprint, SDSS image and catalog, and HSC sources for a portion of M101. The pink circles show the HSC detections.
demonstrates that the photometry in the HSC is typically accurate to a few hundreds of a magnitude based on repeat observations of non-variable stars.

While variability is a good example of how the HSC might be used (and in fact has already sparked an ESA-funded project to build a Hubble Catalog of Variables), several other potential examples are also listed in section 7.4 of the Hubble Space Telescope Primer for Cycle 23 http://www.stsci.edu/hst/proposing/documents/primer/primer.pdf.

Examples include: (1) extractions from extremely large data sets (e.g., Figure 1); (2) cross-matching with other catalogs (e.g., SDSS, Two Micron All Sky Survey, spectral catalogs, personal catalogs, etc.); (3) identification and determination of object properties (e.g., star clusters and associations, colors, elongations); (4) measurements of astrometric properties (e.g., proper motions, cluster kinematics, identification of Kuiper Belt objects); and (5) measurement of photometric redshifts.

Now, with the reader’s interest hopefully piqued, here are some technical details.

• The HSC is derived from visit-based source lists from the HLA, built using the SOURCE EXTRACTOR software (Bertin & Arnouts 1996).

• The matching algorithms used by the HSC are described in Budavari & Lubow (2012).

• The HSC can be accessed in three ways. (1) For most cases, the Discovery Portal of the Mikulski Archive for Space Telescopes is the best choice http://mast.stsci.edu/, (2) For large queries, a CASJobs (CAS: Catalog Archive Server) capability is available, similar to the one developed for SDSS. (3) For certain detailed queries, the HSC homepage (http://archive.stsci.edu/hst/hsc/) is appropriate.

• A draft journal-level publication describing the HSC, the quality of the data, and the potential for doing science with the catalog is available. This can be found, along with other useful information, in a FAQ—available at http://archive.stsci.edu/hst/hsc/help/HSC_faq.html.

• Figure 4 shows the current photometric accuracy available for the different instruments. Note that ACS artificially appears to be much better than WFC3, primarily because these HLA source lists do not currently attempt to go as deep as the WFC3. These comparisons will be made more uniform in the fall of 2015.

• Figure 5 shows the current astrometric accuracy for the various instruments.

While the HSC represents a tremendous new resource for astronomers, it must be used with care. Unlike SDSS, with a uniform set of filters and all-sky coverage over a substantial part of the sphere, the Hubble database consists of tiny pieces of sky using several different cameras, hundreds of filters, and exposure times that range from a fraction of a second to thousands of seconds. Hence the HSC is a very different type of catalog. It will require caution when making use of it.

Potential users should pay special attention to the “five things you should know about the HSC” (http://archive.stsci.edu/hst/hsc/help/HSC_faq.html#five_things). This website illustrates common artifacts and limitations that can occur. In Figure 6
we highlight one of these “five things,” namely that coverage can be non-uniform. The figure shows a particularly dramatic example, caused by overlapping images and a poor choice of the number of images to require for a match. The lesson is that users should keep the unique characteristics of the HSC in mind. In particular, they should always look at the relevant images, rather than just blindly pulling things out of the HSC database.

Catalogs have been a mainstay since the beginning of astronomy. Historical examples include the Messier and Hershel catalogs, and the New General Catalog. More recent examples include 2MASS, Hipparcos, and SDSS. Examples of those coming in the future will include the catalogs of the Panoramic Survey Telescope & Rapid Response System (Pan-STARRS), Gaia, and the Large Synoptic Survey Telescope (LSST).

In many ways the Hubble Source Catalog will be unique, first and foremost because of the depth and spatial resolution of Hubble Space Telescope. It will also be unique because of the inherent non-uniformity and patchwork nature of Hubble observations. This irregularity will require care when developing search criteria. Nevertheless, it is clear that the HSC will be a powerful new tool for research with Hubble data, even with its limitations, and will be an important reference for future telescopes, such as James Webb Space Telescope, and survey programs, such as LSST.