ST-ECF NEWSLETTER

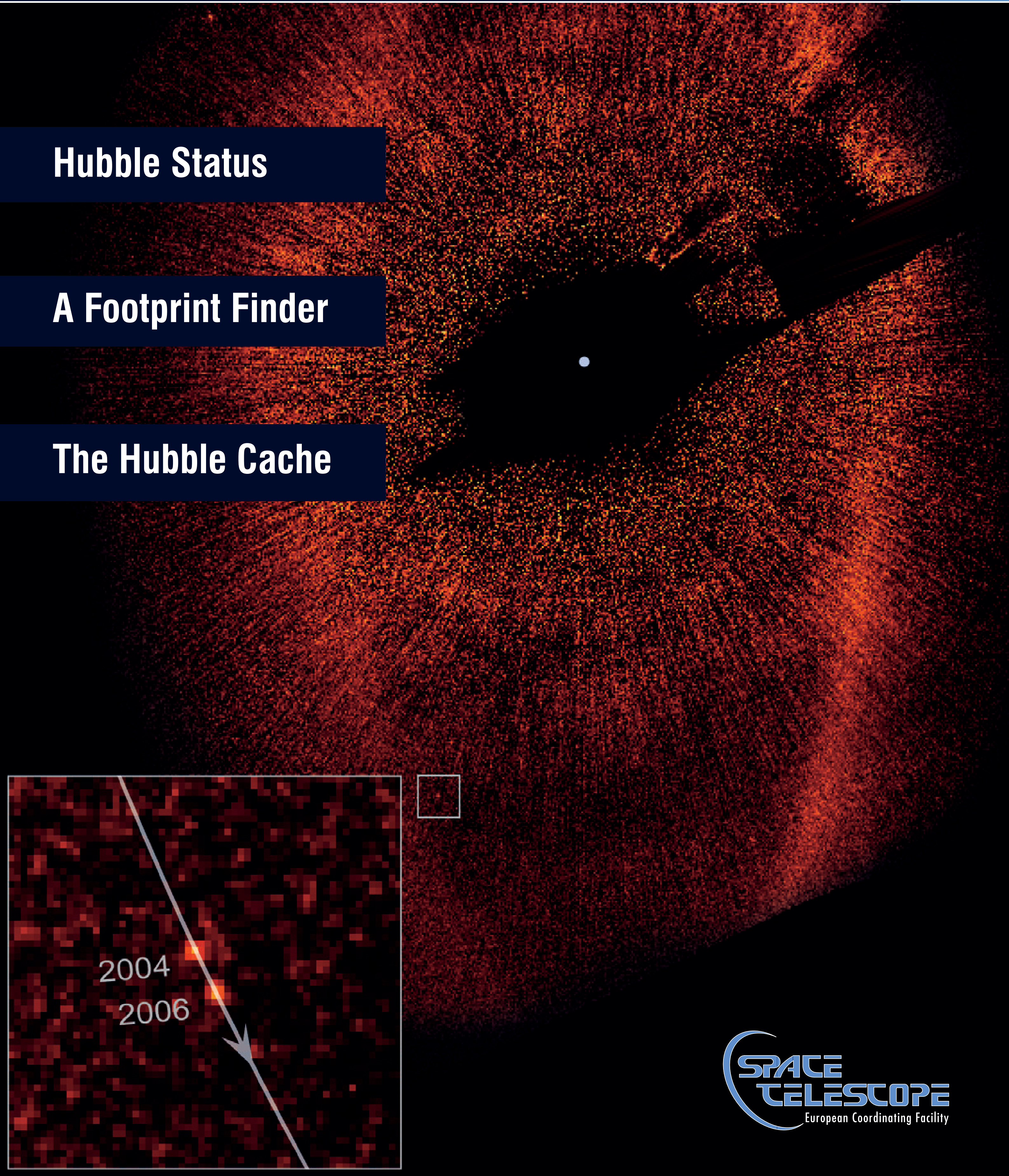
December 2008

45

Hubble Status

A Footprint Finder

The Hubble Cache



Space Telescope European Coordinating Facility

HUBBLE STATUS

Jeremy Walsh

The past few months have been eventful times for Hubble and as a result the new scheduled launch date for the next servicing mission (SM4) is now May 2009. At the time of the last Newsletter, the launch looked set for late August. However, the necessity of having another shuttle ready on a second launch pad, in case a rescue mission to return astronauts from a damaged *Atlantis* was required, resulted in a slip to 14 October. Following a flight software update, which necessitated safing of all the science instruments, WFPC2 again safed during its recovery on 10 September with a discrepancy in telemetry followed by an unrelated safing of the NICMOS cryo-cooler (NCC) on the following day. A high-speed rotor in the NCC appeared to be spinning too fast and the current on one of the rotors also showed a high out-of-limits value causing the instrument to also safe. In addition, two pressure sensors on the NCC showed disparate values when they should have been similar, leading to the indication that there may have been a blockage. After extensive investigation it was decided to leave the NCC off until after SM4.

Fig 1: *Rare sight of two space shuttles at Kennedy Space Center readied for launch in September 2008. Space shuttle Atlantis is in the foreground on Launch Pad A with Endeavour in the background on Launch Pad B.*

When both shuttles were together on the pad at Kennedy Space Center (see Figure 1) and the final round of pre-launch tests were in progress, the telescope safed again on 27 September with a failure of the Control Unit / Science Data Formatter (CU/SDF) in the Science Instrument Command and Data Handler (SI C&DH) subsystem. The CU/SDF is the interface between the telescope computer and the instruments and therefore critical for operation of all of the main science instruments. The exceptions are the Fine Guidance Sensors (FGS) that do not run through the CU/SDF and so could continue to operate. A review board was convened to investigate the failure and found that a serious failure had occurred in the Side A electronics of the SI C&DH. This unit is crossed-strapped with a spare (Side B) so it was decided to attempt to switch on this backup. This unit had not been tested since before launch over 18 years ago, and the change involved halting many systems aboard the telescope, so was not a procedure that could be undertaken lightly. A decision was taken very quickly after the failure of the Side A SI C&DH to postpone SM4 until a new unit could be flown on the telescope. Without a new unit, Side B would represent a single point failure, which could jeopardise the successful operation of the scientific instruments, and thus a cause for concern for continued lifetime of science operations. Since the spare unit on the ground is not in a flight-ready configuration, extensive testing needs to be done before it can be ready for swapping with the in-flight unit (this unit is designed to be changed out by an astronaut). Together with the requirement of availability of two shuttles for the HST Servicing Mission, this pushed the launch to May 2009.

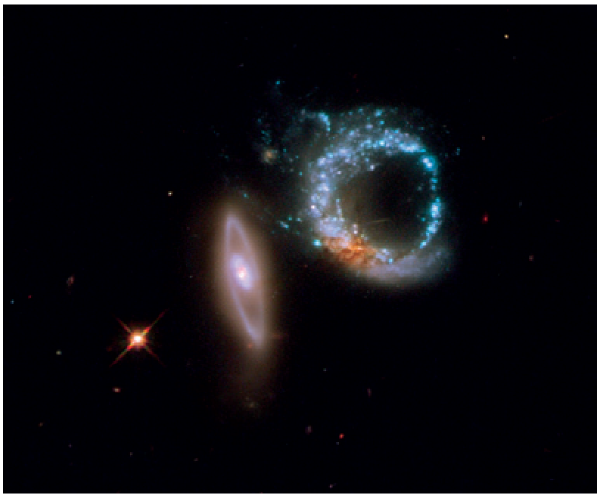


Fig 2: *WFPC2 image of the peculiar interacting galaxy pair Arp 147. This image was acquired two days after Hubble was brought back online following the events described in the text and demonstrated that the camera is again working exactly as it was before going offline.*

After developing the turn-on procedure for the SI C&DH on the ground, the switch was made to Side B on 23 October. At first all appeared to proceed well, but then there was a glitch on the SI C&DH and the Advanced Camera for Surveys (ACS) Solar Blind Channel (SBC) safed during turn-on of the low voltage power supply. Subsequent investigation showed that the two events were unrelated and that a momentary short that resolved itself probably caused the interruption on the Side B SI C&DH. It was decided to repeat the turn-on procedure and this time all went well. WFPC2 was restarted and took its first image on 23 October (Figure 2) and the SBC was successfully powered up on 29 October. After further investigation of the NICMOS cryo-cooler problem it was determined that the differing pressure readings may not have indicated a fault and turn-on was attempted on 18 November. Since the neon coolant in the NCC has warmed up, the complete cool-down cycle must now be run, taking about one month. Provided this cool down proceeds successfully, NICMOS should be able to resume operations in late December.

The slip of the launch from October 2008 to May 2009 implies that there is an insufficient number of high quality proposals from Cycle 16 remaining to keep the telescope observing efficiently for the four months to May 2009. Thus a supplemental Call for Proposals was announced on 17 November, with the very short deadline of 8 December. The Call is restricted to large (>75 orbits) and innovative or high risk shorter proposals for WFPC2, ACS SBC, NICMOS (assuming it returns to its cooled state) and FGS. The successful proposals will be judged by a subset of the Cycle 16 Time Allocation Committee (TAC) and panellists, but with all assessments by e-mail, with the aim of having the successful proposals beginning observation in late January 2009. Meanwhile the new instruments, the Cosmic Origins Spectrograph (COS) and the Wide Field Camera 3 (WFC3), remain in their storage containers at the Kennedy Space Center in Florida.

THE HUBBLE CACHE

Felix Stoehr, Jonas Haase, Daniel Durand (CADC) & Alberto Micol

ABSTRACT

How do scientists actually obtain data from the Hubble Space Telescope? There are two ways: the first is to submit an observing proposal and win observing time in a highly competitive selection process. The second option is much easier: after a proprietary period of (usually) one year, all data from Hubble become public and can be obtained from the Hubble archive by simple download over the internet. The ST-ECF and the Canadian Astronomy Data Centre (CADC) have now joined together to move the Hubble archive up to the next level by making retrieval of data much faster and more convenient. This also opens up the possibility of data-mining (as an example see Schade et al., 2002) as well as the option of serving Hubble data through the Virtual Observatory (VO).

INTRODUCTION

The Hubble archive is hosted around the world at three sites: at the Space Telescope Science Institute (STScI) in Baltimore, USA, at CADC in Victoria, Canada and at the ST-ECF in Garching, Germany. About 35 TB of data are delivered to users each year, whether they be professional astronomers or amateur astronomy enthusiasts, with the STScI delivering the largest fraction.

The datasets offered are calibrated to remove as much of the instrument signature as is feasible. In an huge ongoing effort, instrument teams at STScI continuously provide new data reduction software and calibration reference files to try to take out as many instrument effects as possible. For each user request, the original telemetry data from the telescope is reprocessed and recalibrated on-the-fly using a request handler system. Originally CADC and ST-ECF introduced the On-The-Fly-Calibration pipeline (Pirenne et al., 1998) to apply the calibration procedures to the raw FITS data and this was then later augmented by the On-The-Fly-Recalibration pipeline delivered by the STScI (Swam et al., 2002), which made it possible to recreate the raw files of all active instruments (ACS, STIS, NICMOS and WFPC2) from telescope telemetry data. This was clearly superior to the previous system, which essentially froze the raw data in time. Another advantage of the system was that it conserved storage space as only the Hubble telemetry files and a few smaller auxiliary files needed to be stored. This was an important resource aspect when data were stored on expensive optical disks in jukeboxes. Indeed, since its beginnings in the early 1990s the Hubble archive has gone through several generations of storage media, always trying to minimise the cost of storage. Hugely expensive laser disks were supplanted by CDs then DVDs and finally by spinning disks.

The high quality of the data facilitates scientific (re)-use and, at present, more purely archival scientific papers (41%) are published than papers where at least one of the authors was the principal investigator (PI) of the corresponding programme (38%) (Rick White, private communication). The remaining 21% of the papers have partial archival content: using PI data but also using additional data out of the Hubble archive.

The on-the-fly reprocessing and recalibration procedure is rather unusual for today's archives. Most of them deliver either entirely static data products or update their holdings for distinct data releases (e.g., SDSS). It does, however, guarantee the best possible quality of all datasets at all times.

The drawbacks associated with this procedure are that it often takes minutes, hours or even days before the data have been processed and made available to the user. Having a quick look at the actual data in order to decide whether or not the data will fit the astronomer's needs is not possible. It is also not possible to offer programmatic access (such as a static URL) to the files and it follows that direct access through Virtual Observatory (VO) protocols is also not possible. The same holds true for scientific projects that would like to data mine large parts or the entire Hubble archive. Again, such projects cannot be done because of prohibitively long calibration times. Finally, changes in the metadata associated with the datasets that are induced by software or reference file changes cannot be easily captured and transferred into the database system that is used for the queries to the archive. In order to address these issues, CADC and ST-ECF started a common effort to build the Hubble Cache.

THE CACHE

With falling prices for CPU and disk space, processing the entire Hubble archive and storing the calibrated files on spinning disks came within reach. As a first step, all Hubble datasets are reprocessed and recalibrated with the latest software and reference files and put into a fast distributed storage system. Then a system of software agents, running continuously, detects newly arriving datasets or datasets affected by recent changes and submits them to a grid processing system in a fully automatic way. The agents also monitor the processing and the system state (see Figure 1), deal with exceptions, extract metadata from the processed files and update the database system. The software was written in Python, making heavy use of the object-oriented paradigm and runs in identical versions at CADC and ST-ECF. A sketch of the system is given in Figure 2 and the layout of the database is shown in Figure 3.

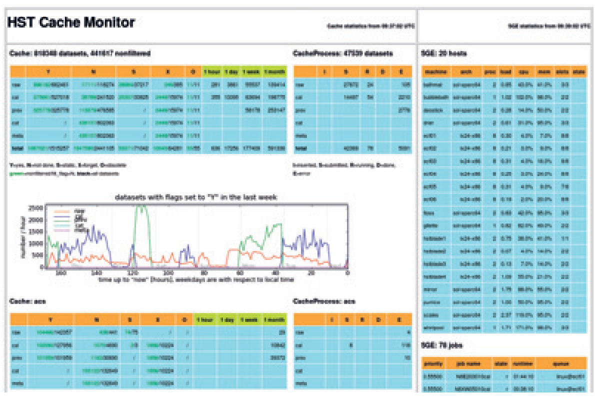


Fig 1: *HST Cache monitor. A web interface to continuously updating statistics on the progress of Cache processing.*

We use the Sun Grid Engine to drive the grid of about 75 processing nodes (Solaris and Linux systems) that are available at each site. The Sun Grid Engine turned out to be very easy to install and maintain, to be extremely reliable and flexible and at the same time provide a lot of functionality. With the current processing power at hand, reprocessing the whole Cache holdings (about 38 TB of data) takes about two months at each site. A number

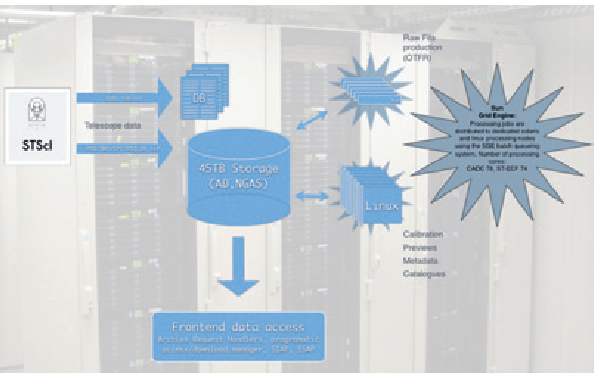


Fig 2: *Schematic overview of the data flow within the Cache system.*

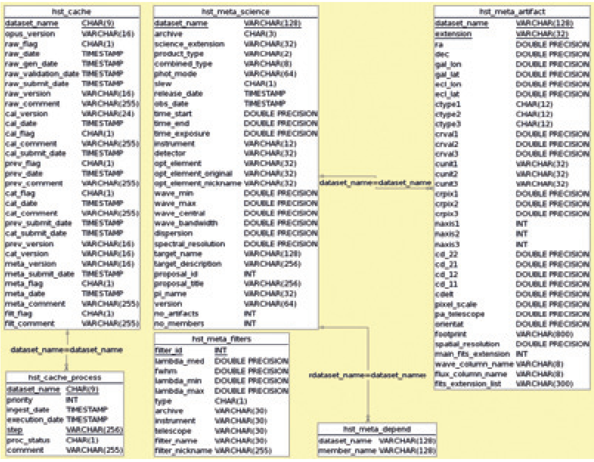


Fig 3: *Structure of the database used by the Cache system.*

of data exchange mechanisms ensure a high degree of interoperability and redundancy between the archives. For historical reasons we use different distributed storage systems: AD was developed and is used at CADC, NGAS is used by the ST-ECF and was developed by ESO.

PROCESSING STEPS

Since different hardware is needed for the raw FITS file production and in order to keep the processing of a dataset manageable, several independent steps have been introduced.

• **Production of raw FITS files (using Solaris-based hardware):** Hubble telemetry data and metadata runs through the OTFR pipeline, provided by the STScI, to generate raw FITS files.

• **Calibration (on Linux hardware):** Raw FITS files are calibrated in the OTFC pipeline, using the newest available calibration software from the STSDAS IRAF package. Whenever possible a MultiDrizzled output product is provided as well.

• **Preview (Linux):** Uniform preview products in both FITS and PNG formats are created from calibrated data to be used for fast web and VO access.

• **Metadata (Linux):** See next section.

• **Catalogues/object lists (Linux):** These are planned as a further improvement to the harvested metadata and possibly to produce a general searchable source database.

METADATA

The quality of the query interface (see Figure 4) that can be presented to the users strongly depends on the quality of the metadata that is available. In order to maximise metadata quality we extract the values directly from the

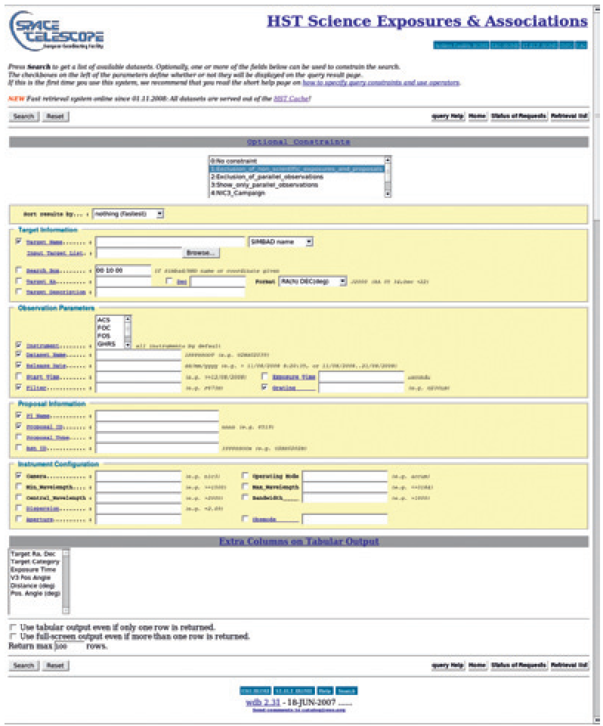


Fig 4: *Hubble archive query interface.*

recomputed FITS headers where possible. This information is then complemented with entries from the original STScI databases as well as with data we compute directly from the pixels. We decided to follow the CADC internal Common Archive Observation Model (CAOM, Dowler et al., 2008) as well as VO characterisation standards wherever it made sense. In order to be able to display footprints in VO tools, or to be able to compute whether or not a celestial object really is located on an image, we use the footprint finder algorithm described on page seven of this Newsletter.

IMPLEMENTATION

The hardest aspect of developing a software framework for a telescope such as Hubble is to be able to deal with the heterogeneous nature of the data. Between different instruments and even within data from any given Hubble instrument there exists a host of exceptions and failure modes that one has to accommodate. Below are a number of design choices that were made to help deal with this.

• **The Predictor:** The backbone of the Cache is a method that returns the expected output files at each processing step. Hubble has a bewildering number of instrument modes and associated output products, so ensuring that the same files are produced every time, no matter how much the software or reference files change, is paramount.

• **Object-oriented Python Code:** Another way of dealing with the plentiful exceptions among Hubble instruments is an object-oriented code base written in Python. Methods and exceptions only needed for one instrument can easily be overloaded on a case-to-case basis and a flexible scripting language makes it easy to extend and adapt the Cache code as more necessary functionality is (re)-discovered.

• **Single Source:** The number of input tables to the Cache is kept to an absolute minimum to avoid hitting the inconsistencies and special naming problems that are found in the rather large set of derived Hubble database tables. This means that a number of historical special cases had to be rediscovered and folded into the code, but it also presented the opportunity to mandate absolute consistency across instruments.

• **Associations:** Handling of associations of data differs wildly among the Hubble instruments. Sometimes members are visible as datasets in their own right, sometimes they are not. That is why it is important to treat associations block-wise at all times, that is, processing steps requested for a single member will always be run on the entire association to ensure that it stays consistent internally.

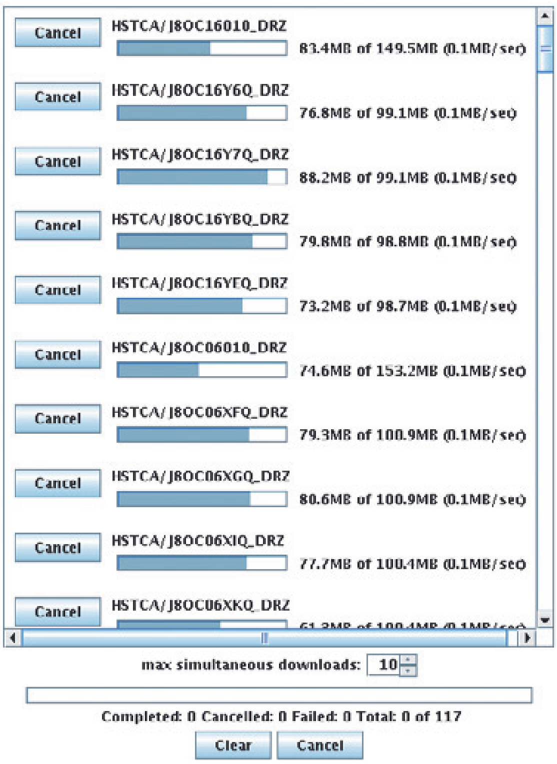


Fig 5: *Download Manager in action.*

For the impatient astronomer: programmatic access to Hubble data

Professional astronomers, who often need to analyse large amounts of data in an automated fashion, and curious amateur enthusiasts no longer have to suffer long waiting times and byzantine data request procedures. All Hubble datasets are now available pre-processed and can be downloaded directly, either via the ST-ECF or the CADC query interfaces mentioned above, or directly via file proxies using a web browser, command line tool or similar.

The URL to download any given file is:
(ST-ECF)
archive.eso.org/archive/hst/proxy/ecfproxy?file_id=<FILE_ID>
(CADC)
cadc.hia.nrc.gc.ca/getData?archive=HSTCA&file_id=<FILE_ID>

The <FILE_ID> to insert is the dataset name followed by an extension without '.fits'.

As an example, to get o6d701030_x1d.fits the URL would be:

(ST-ECF)
archive.eso.org/archive/hst/proxy/ecfproxy?file_id=o6d701030_x1d
(CADC)
cadc.hia.nrc.gc.ca/getData?archive=HSTCA&file_id=o6d701030_x1d

Please note that the separator between dataset name and extension can be either "." or "_". All older instruments, up to and including WFPC2, use a dot, STIS, NICMOS and ACS have underscores. For an overview of the possible extensions and examples of file ids please consult the instrument-specific help pages describing the filenames at:

(ST-ECF)
archive.eso.org/cms/hubble-space-telescope-data/filenames

(CADC)
cadc.hia.nrc.gc.ca/hst/hst_filenames.html

If you use a command line tool such as **curl** or **wget** to download a file you might have to specify an output file name as they do not always get the name from the HTTP header.

For example:
```
(ST-ECF) curl -o o6d701030_x1d.fits 'archive.eso.org/
archive/hst/proxy/ecfproxy?file_id=o6d701030_x1d'
(CADC) curl -o o6d701030_x1d.fits 'cadc.hia.nrc.gc.ca/
getData?archive=HSTCA&file_id=o6d701030_x1d'
```

The Download Manager can generate an expanded list of such commands for any given dataset.

Documentation of the respective proxies can be found at:
(ST-ECF) archive.eso.org/archive/hst/proxy/ecfproxy
(CADC) cadc.hia.nrc.gc.ca/getData/doc

• **Common software and command line interface at CADC and ST-ECF:** Since the work force is limited on both sites, the same database structure, software library and user interface is used to minimise maintenance. All site dependencies have been encapsulated or emulated (e.g., NGAS *v.* AD file storage system). A central CVS repository keeps the Cache software in synchronisation and external software is tested and installed in cooperation.

DOWNLOAD MANAGER

With all the Hubble files available online, the limiting factor in getting the data to the user is now the network bandwidth. Typically the files are downloaded in serial order and the maximum speed obtained is that of a single stream. This speed is in most cases significantly lower than the total bandwidth available, which is often due to the TCP/IP configuration of all the nodes along the path from the user to the data centre, but there are also limitations due to the inner workings of the TCP/IP protocol itself.

In order to speed up the downloading process, we use the Download Manager (Figure 5) that was developed at CADC. This program allows for parallel downloads and can thus use the full bandwidth available.

OUTLOOK

The Hubble Cache has been online at CADC and ST-ECF since 1 November 2008. The step to extract the metadata from the FITS files is in the works and is expected to be complete in early 2009. The main entry points to the Cache are:

cadc.hia.nrc.gc.ca/hst/science.html and
archive.eso.org/hst/science

at CADC and ST-ECF, respectively.

REFERENCES

Dowler P. et al. 2007, ADASS XVI, ASP Conference Series, 376, 347

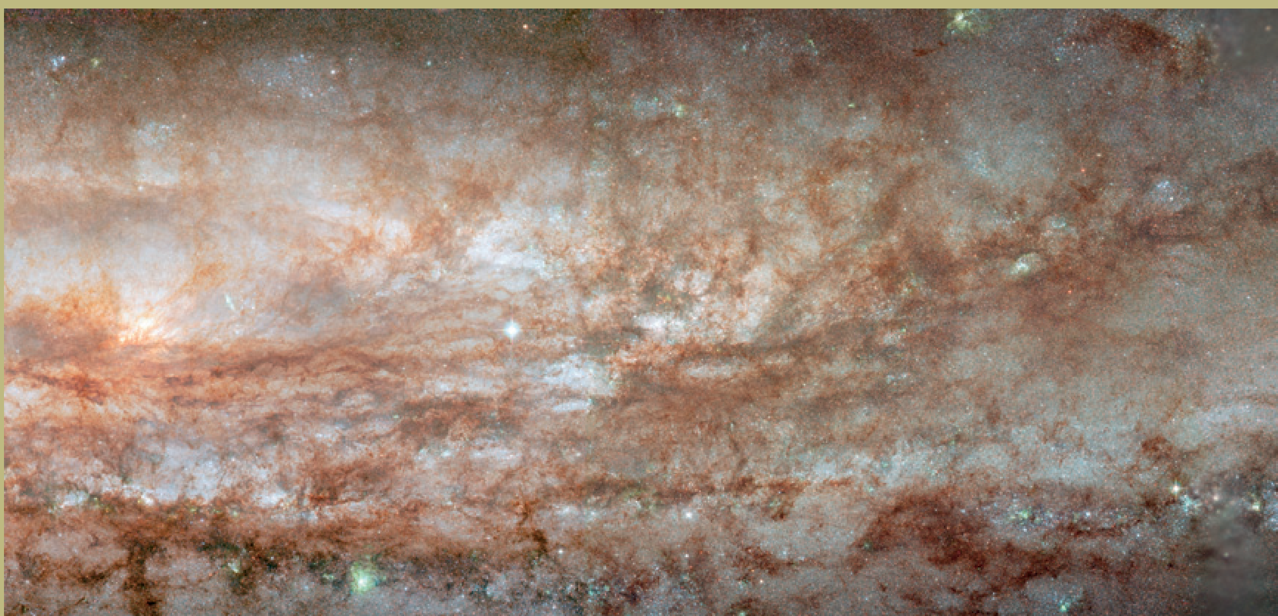
Pirenne B., Micol A., Durand D. & Gaudet S. 1998, ADASS VII, ASP Conference Series, 145, 341

Schade D. et al. 2002, ADASS XI, ASP Conference Series, 281, 36

Stoehr F. 2008, ST-ECF Newsletter 45, 7 (this issue)

Swam M. S., Hopkins E. & Swade D. A. 2001, ADASS X, ASP Conference Series, 238, 291

NGC 253 ABLAZE WITH THE LIGHT FROM THOUSANDS OF YOUNG, BLUE STARS [heic 0819]



Credit: NASA, ESA, J. Dalcanton and B. Williams (University of Washington), T.A. Rector/University of Alaska Anchorage, T. Abbott and NOAO/AURA/NSF

NGC 253 is one of brightest spiral galaxies in the night sky, easily visible with small telescopes, and it is composed of thousands of young, blue stars. It is undergoing intense star formation. The image demonstrates the sharp "eye" of Hubble's Advanced Camera for Surveys, which is able to show individual stars. The dark filaments are clouds of dust and gas. NGC 253 is the dominant galaxy in the Sculptor Group of galaxies and it resides about 13 million light-years from Earth. The observations come from a detailed study, called the ACS Nearby Galaxy Survey Treasury (ANGST) programme.

FOOTPRINT FINDER

Felix Stoehr

ABSTRACT

Plotting tools such as VirGO or Aladin show professional astronomers or astronomy enthusiasts the exact extent of the region of the sky that has been observed by a telescope. But how can one actually obtain the polygon outline of an existing astronomical image? We present a way of doing just that.

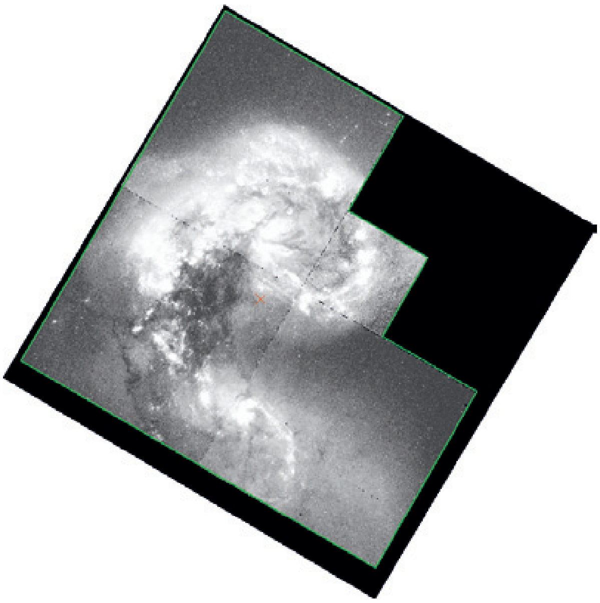


Fig 1: *Combined image of the four chips of a single WFPC2 observation. The green line shows the computed footprint and the red cross marks the barycentre.*

INTRODUCTION

The footprint, which we define as the accurate polygon outlines of the illuminated parts of an astronomical image, is very valuable information. It can be used in Virtual Observatory (VO) tools in order to show where the observed regions are without having to download and display all the very large images themselves. Footprints can also be used to compute whether or not a given object indeed was observed on that image.

In principle every astronomical image stored in a FITS container comes with a header that describes the position of the image on the sky. With the so called World Coordinate System (WCS) information of this header, it is quite easy to compute the positions on the sky (usually as right ascension and declination coordinates) of all the four corners of the image.

However, these four corners of the image are only a very rough description of the actual footprint in many cases. For space observations, for example, it is now often common practice to store images "North up", independently of the orientation of the telescope. The surrounding parts of the new image are then filled with zeros or some flag indicating missing data. In addition, multiple images are often combined to create high-level data products. One example of such a high-level data product is shown in Figure 1. The four different chips of the WFPC2 camera on Hubble have been combined into one image leaving a corner region empty. In addition, when looking carefully at the image, it turns out that there are actually empty regions everywhere around the illuminated pixels. This can get more and more complex when images from completely different observations are combined as is shown in Figure 2 where 62 ACS direct images from the ST-ECF HLA ACS Grism Project (see page 12 of this Newsletter) have been combined into one using MultiDrizzle.

Here we propose a flexible algorithm to compute the actual footprint polygons of the illuminated parts of any image.

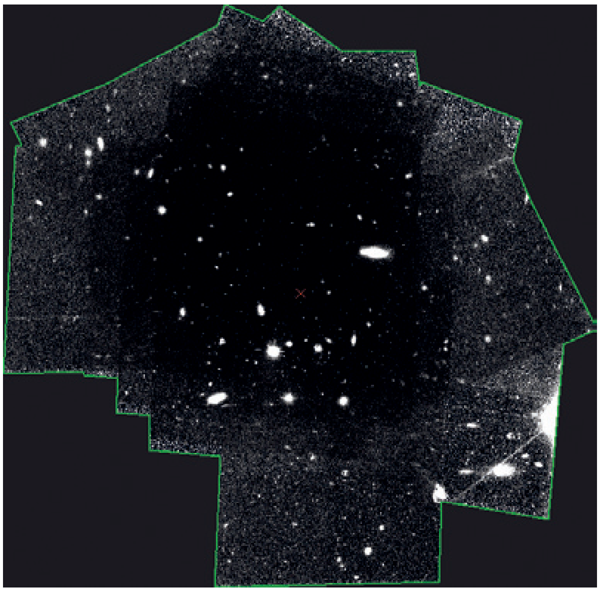


Fig 2: *Image of 62 ACS exposures combined together using MultiDrizzle.*

ALGORITHM

The full algorithm is a combination of eight rather well known sub-algorithms.

1) As a first step, the illuminated pixels that are sitting on the border with regions with "empty" pixels are identified. For this we use the first FITS header extension that does contain data (as default) and select all those non-zero values (zero or NaN are the defaults) pixels that have at least one of the eight neighbours that has a value equal to zero. In practice "equal to zero" means that the deviation from 0 should be smaller than a given epsilon value. This border pixel selection is called "8-connect border" in the literature.

2) The border pixels that are identified in the first step are then sorted into a tree data structure that allows us to very quickly identify the state of the pixels around a given pixel. This is crucial to allow for acceptable performance of the next steps. We use the Python dictionary type that is a ready-to-use tree-like data structure.

3) Next, we compute the distinct "chips" of the image, these are groups of pixels in which each pixel touches at least one of the other border pixels of the group. This computation can be done easily using a standard friends-of-friends group finder algorithm (Huchra & Geller, 1982) with a linking length of $\sqrt{2}$.

4) Only groups that are larger than a given threshold are kept for further processing. The default is that each group must be made of more than 5% of the total number of border pixels.

5) For each of the remaining groups, the pixels have to be ordered. This is done by identifying the pixel in the lower left corner and then walking clockwise starting from that pixel, in each step identifying the following clockwise pixel that is a border pixel too. This algorithm is similar to the method of escaping from a labyrinth (works for simply-connected mazes only) by always keeping the right hand on the wall. This step also nicely circumvents pathological cases of, for example, a 1-pixel-wide line in the image.

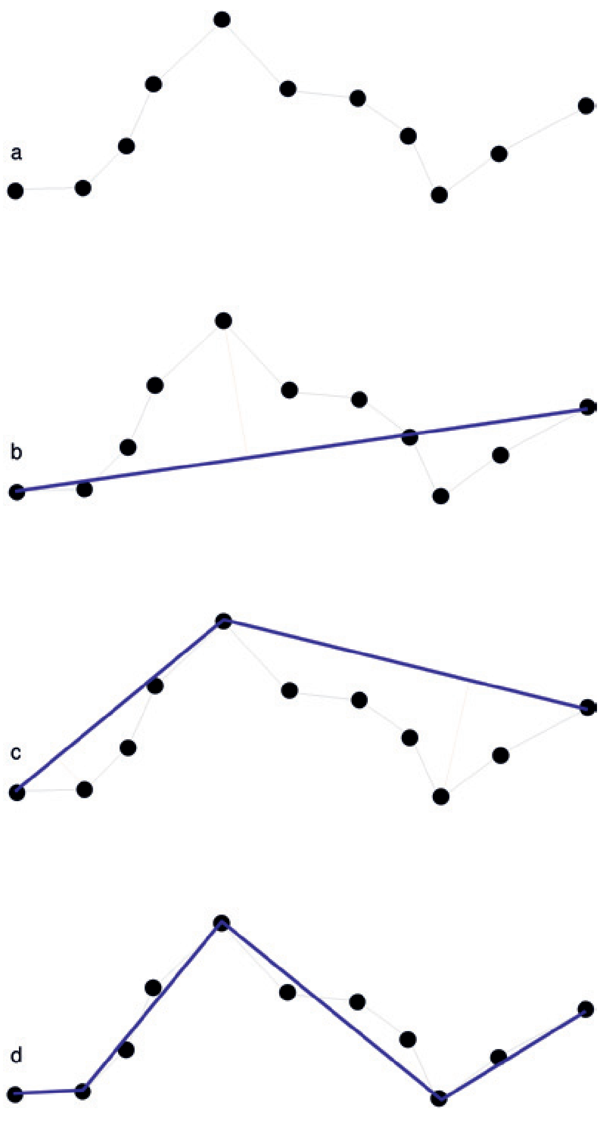


Fig 3: *Douglas-Peucker-algorithm: In the first step the start and end points of the polygon that should be simplified (a) are connected with a straight line (b). If there is a point further away from that line than a given threshold then the original straight line is split in two (c) and the process is continued recursively (d) until no point is more distant from the simplified polygon line than the threshold distance.*

6) For each group, the result of the previous step is already a polygon that has as many corners as the group has border pixels. Although this is the most accurate representation of the footprint, in most, if not all, cases users are interested in a simplified version of the polygon with far fewer corners. We use the Douglas & Peucker (1973) algorithm to reduce the number of polygon corners by keeping only the most important ones. This is done the following way (Figure 3): the first and last points of the original polygon line are connected with a straight line. Then, for each point in between, the distance of that point to the straight line is computed. If one or more points are farther away than a user-specified threshold value, the point with the largest distance is kept and the procedure is repeated recursively for both new lines: the one from the start to the new point and the second from the new point to the end point.

7) For each of the simplified polygons we can easily compute the area and the barycentre (Meister, 1769).

8) In order to decide whether one footprint is contained in another one and whether or not it is contained an odd number of times (then this footprint describes a hole) or an even number (then it describes a chip or an “island” in a hole) in others, we use a standard point-in-polygon algorithm (Sutherland et al., 1974). Starting with the footprint that is contained in the largest number of other footprints, it is then possible to compute for each footprint its parent footprint and hence the full hierarchy tree.

FEATURES

The proposed algorithm can deal with arbitrarily shaped pixel regions including concave shapes (Figure 4). It is constructed to automatically detect multiple chips and, because of step 4), it is also very robust with respect to image defects and bad pixels as those are groups with small numbers of border pixels. Also, by design, it computes the full footprint hierarchy. It is relatively fast: the footprints of a typical ACS image with 4k x 4k pixels can be computed in about 15 seconds on a standard desktop computer. Given the way the border pixels are computed, it is enough to keep only three rows of the full image in memory at any given time. Pyfits, which is used to read the FITS files, allows access to the image row-by-row and so the memory used by the implemented code can be kept very low. In particular, at no time does the whole FITS image have to be loaded into memory. The algorithm should work out of the box for most FITS images and the only important free parameter is the threshold that determines the accuracy of the final simplified polygon. The resulting footprints are written to a text file. In addition, the user can request a plot of the polygon or have a region file written that can be imported into ds9. No input other than the FITS image file itself is necessary.

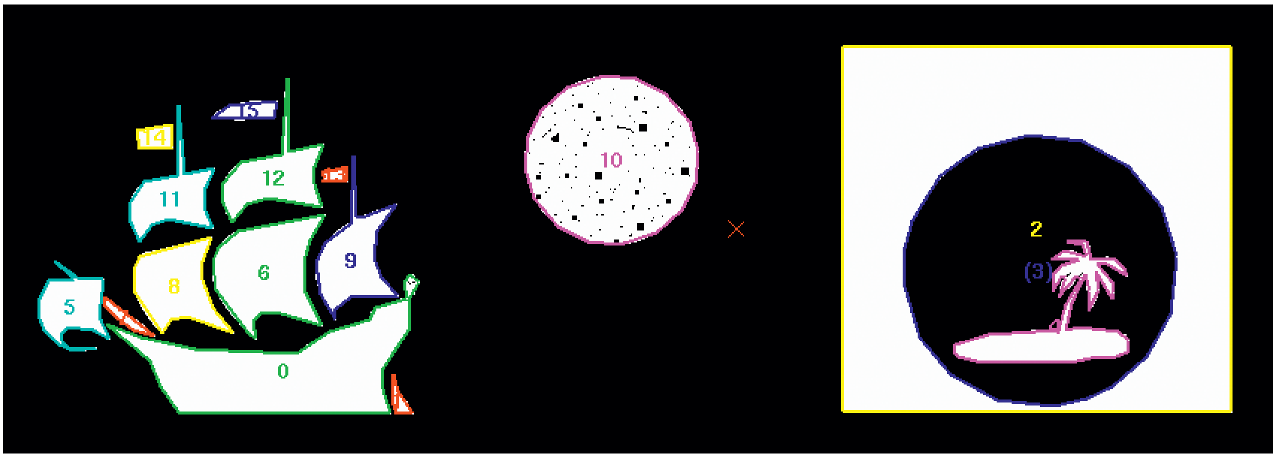


Fig 4: *Image showing the different features of the footprint finder algorithm including multiple chip detection, dealing with concave and convex shapes, dealing with bad pixels, computing the barycentres and computing the hierarchy of holes (denoted with parentheses) and islands.*

DOWNLOAD

The Python implementation of this algorithm may be obtained from the ST-ECF website:

www.stecf.org/software/ASTROsoft/Footprintfinder/

It requires the matplotlib module, which contains pylab, for computation and plotting and the pyfits package for the reading of FITS files. The footprint finder code has been tested successfully with matplotlib version 0.98.0 and pyfits version 1.3. A typical usage would be:

footprintfinder.py -p myimage.fits

STATUS AND OUTLOOK

The footprint finder is routinely used on all images of the Hubble Cache (see this Newsletter on page three) at CADC and ST-ECF. CADC will also use it on images in their other archives and STScI are planning to use the algorithm to compute the footprints of their HLA products.

REFERENCES

Huchra J. P. & Geller M. J. 1982, ApJ, 257, 423

Douglas D. & Peucker P. 1973, The Canadian Cartographer, 10(2), 112

Meister A. L. F. 1769, Novi Comm. Soc. Reg. Scient. Gotting., 1, 144 and associated plates

Sutherland I. et al. 1974, ACM Computing Surveys, vol. 6 no. 1

HUBBLE SEES MAGNETIC MONSTER IN ERUPTING GALAXY [heic 0817]

The Hubble Space Telescope has found the answer to a long-standing puzzle by resolving giant but delicate filaments shaped by a strong magnetic field around the active galaxy NGC 1275. It is the most striking example of the influence of these immense tentacles of extragalactic magnetic fields, say researchers.

NGC 1275 is one of the closest giant elliptical galaxies and lies at the centre of the Perseus Cluster of galaxies. It is an active galaxy, hosting a supermassive black hole at its core, which blows bubbles of radio-wave emitting material into the surrounding cluster gas. Its most spectacular feature is the lacy filigree of gaseous filaments reaching out beyond the galaxy into the multi-million degree X-ray emitting gas that fills the cluster.

These filaments are the only visible-light manifestation of the intricate relationship between the central black hole and the surrounding cluster gas. They provide important clues about how giant black holes affect their surrounding environment.

A team of astronomers using the NASA/ESA Hubble Space Telescope Advanced Camera for Surveys have for the first time resolved individual threads of gas which make up the filaments. The amount of gas contained in a typical thread is around one million times the mass of our own Sun. They are only 200 light-years wide, are often surprisingly straight, and extend for up to 20 000 light-years. The filaments are formed when cold gas from the galaxy's core is dragged out in the wake of rising bubbles blown by the black hole.

This stunning image of NGC 1275 was taken using the NASA/ESA Hubble Space Telescope's Advanced Camera for Surveys in July and August 2006. It provides amazing detail and resolution of the fragile filamentary structures, which show up as a reddish lacy structure surrounding the central bright galaxy NGC 1275. Also seen in the image are impressive lanes of dust from a separate spiral galaxy. It lies partly in front of the giant elliptical central cluster galaxy and has been completed disrupted by the tidal gravitational forces within the galaxy cluster. Several striking filaments of blue newborn stars are seen crossing the image.

NASA, ESA and Andy Fabian (University of Cambridge, UK)

PROCESSING ACS GRISM DATA FOR THE HUBBLE LEGACY ARCHIVE

Martin Kümmel, Harald Kuntschner, Jeremy Walsh, Felix Stoehr, Richard Hook & Wolfram Freudling

ABSTRACT

The ST-ECF's contribution to the Hubble Legacy Archive has mostly been focused on the creation of science-ready extracted spectra from slitless spectroscopy data. The group started with NICMOS grism spectral extractions that were described in earlier articles. The next step is described here: the extraction of the larger set of grism spectra from the Advanced Camera for Surveys.

INTRODUCTION

In a coordinated effort the Canadian Astronomy Data Centre (CADC), the Space Telescope Science Institute (STScI) and the ST-ECF are currently building the Hubble Legacy Archive (HLA) [1]: a collection of high level Hubble data products and access tools to ease scientific analysis in the age of the Virtual Observatory. The ST-ECF contribution to the HLA concentrates on slitless spectroscopy, while the CADC and the STScI contributions centre mostly on imaging products (ACS, WFPC2 and STIS). The NICMOS G141 data [2,3] formed part of the first data release (DR1) in February 2008, and the next release will contain ACS/WFC G800L slitless spectroscopic data. This article briefly reports on the status of the HLA ACS slitless project.

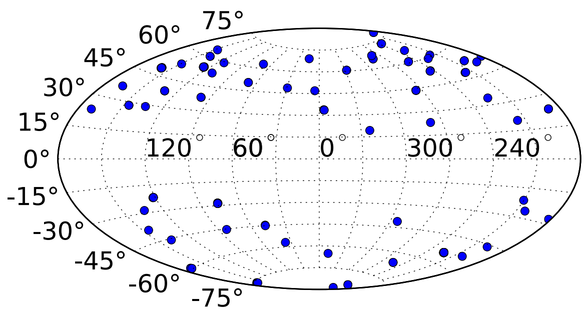


Fig 1: *ACS/WFC G800L pointings in galactic coordinates.*

THE ACS GRISM DATA

The slitless mode of the Wide Field Channel (WFC) of the Hubble Advanced Camera for Surveys (ACS) delivers spectra with a resolution of R∼100 in the wavelength range 0.6—1.0 μm. There are around 150 ACS/WFC G800L datasets, and Figure 1 shows their location in galactic coordinates. Many of these datasets are close to each other and thus not individually recognisable. The total sky coverage of the ACS/WFC G800L data is ∼600 arcmin^2. Figure 2 illustrates some properties of a typical dataset. The lower and the upper panels show a basic data set consisting of a slitless image and its corresponding direct image taken in the F850LP filter, respectively. A single image typically contains from many hundreds to over a thousand spectra to a limiting magnitude of $m_{F775W} > 26$. Since there is no selection of objects by slits, and the strongest (first order) spectra are about 120 pixels in length, there is almost always some degree of contamination between spectra, both in the spatial and spectral directions. For brighter objects, the zeroth order and the higher order spectra, which only contain a few percent of the received flux in the first order, also contribute to the contamination over distances of many hundreds of pixels. The blue circles in Figure 2 mark two object pairs with overlapping spectra in the slitless image.

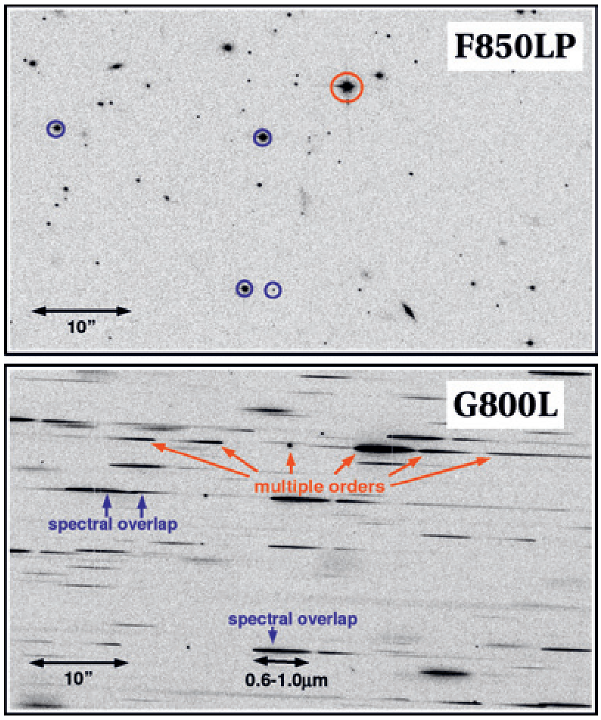


Fig 2: *Example of the basic dataset of one slitless spectroscopic image (bottom) and the corresponding direct image (top) with exposure times of 800 s and 550 s, respectively.*

The zero point of the wavelength calibration must be derived from the object positions on the direct image. Therefore the basic slitless spectroscopic dataset always consists of a set of direct images and dispersed images that were taken at the same position on the sky and the processing of the direct images is an important part of the slitless spectroscopic data reduction.

DATA PROCESSING

For the reduction of the ACS/WFC G800L data, we have modified our Pipeline for Hubble Legacy Archive Grism (PHLAG) data [4,5], which originally was built for the NICMOS HLA project. PHLAG consists of a series of modules with each performing a certain reduction step on the data. The pipeline is implemented in Python, but utilises existing software, for example for object detection or image combination, whenever possible. To estimate the mutual contamination of spectra, the processing of all available direct images and the extraction of the multicolour photometric information is important to ensure the best possible estimate of the contamination of the extracted spectra [6].

Figure 3 gives an overview of one dataset. For the direct images, we combine multi-filter images from different programmes, which usually means different guide stars and roll angles, to form deep images. The aim is to build up direct images in as many filters as possible and deep enough to

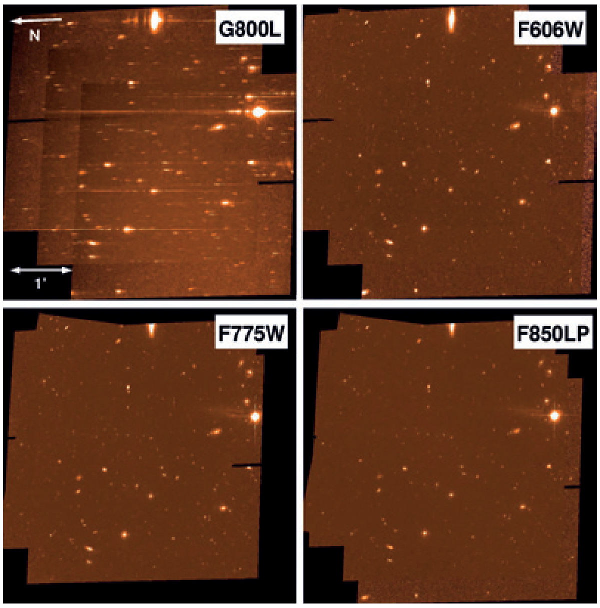


Fig 3: *Coadded direct images and slitless image for one dataset from the NICMOS HUDF parallels observations on the southern GOODS field.*

provide magnitude estimates for all objects with detectable spectra. As a consequence, the sky coverage of the coadded images differs, as can be seen in Figure 3 (comparing the F606W and F775W images). The direct image coverage can even exceed that of the grism images, which has the advantage that objects at the edge of the field, whose spectra may still fall on the grism image, can be extracted or flagged as contaminating spectra. Source detection is done on a "white" light image that is coadded from the individual filter images. The advantage of this approach is that the extraction catalogue is not biased by missing very red or very blue objects (or objects with continuum breaks that imply they have negligible flux in some bands) and all objects with spectra to a deep magnitude limit can be extracted. Then the object brightness for all sources is determined on each filter image. Since the direct imaging data is usually much deeper than the corresponding slitless data, and in order to reduce the data processing time, a sub-sample of all sources is selected, which promises to have detectable spectra in the slitless data. In the data shown in Figure 3, this sub-sample contains the ~1700 brightest from the total of ~3600 detected sources on the "white" light image. The object spectra are then automatically extracted from the slitless images with the data extraction package aXe [6] that was specifically designed for ACS slitless spectroscopy.

STATUS AND OUTLOOK

Figure 4 shows some spectra obtained with the current version of PHLAG. In the first row are two spectra of stars but all other spectra are compact emission line galaxies. The spectra appear to already be of a very good quality. A pre-release, which is likely to contain spectra from the Hubble Ultra Deep Field (HUDF) NICMOS parallels data on the GOODS fields, is expected for the next HLA data release (DR3) early in 2009.

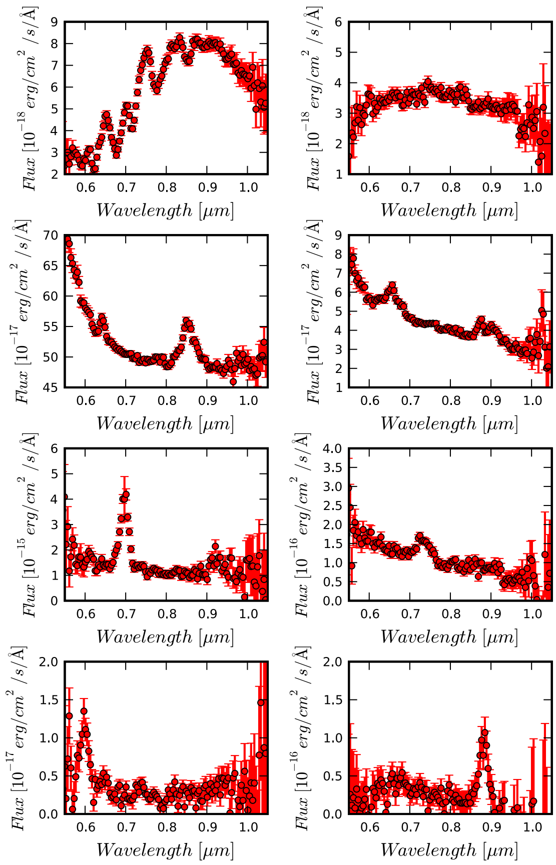


Fig 4: *Selection of ACS/WFC G800L spectra extracted from the NICMOS HUDF parallels observations in Figure 3 and parallel observations to STIS at RA: 11h:17m:6.6s Dec: 18d:13m:54.9s.*

A full data release is projected for summer 2009. For the data releases, improvements in the astrometry (cross-matching with 2MASS, USNO-A2.0 and SDSS) and the quality control (identification of fake sources and selection of high quality spectra) are to be expected. In total ~20 000 fully calibrated spectra will be published via the HLA archive interface [7] and made available through the Simple Spectrum Access Protocol (SSAP) server [8] at the ST-ECF as well as the HLA portal at the STScI [9].

REFERENCES

1. *Walsh, J. R. & Hook R. N. 2006, ST-ECF Newsletter, 40, 6*
2. *Freudling, W. et al. 2008, A&A, 490, 1165*
3. *Freudling, W. et al. 2008, ST-ECF Newsletter, 44, 3*
4. *Kümmel, M. et al. 2006, ST-ECF Newsletter, 41, 6*
5. *Kümmel, M. et al. 2008, in Kaufer, A. & Kerber, F. (Eds.), The 2007 ESO Instrument Calibration Workshop, ESO Astrophysics Symposia XXVIII, 185*
6. *Kümmel M. et al. 2008, submitted to PASP*
7. *The HLA archive: archive.eso.org/wdb/wdb/hla/product_science/form*
8. *The ST-ECF SSAP server: www.stecf.org/hla-vo*
9. *The HLA portal at the STScI: hla.stsci.edu*

ZwCl 1358+62 is located 3.7 billion light-years from Earth (z=0.33) and is made up of at least 150 individual galaxies. This image depicts multiple blue, red and orange arcs scattered across the image, which represent amplified and stretched images of the galaxies behind the cluster's core. The colours displayed by the various lensed galaxies vary according to their distance and galaxy types. The natural gravitational lensing effect in combination with Hubble's potent mirrors provide astronomers with a powerful set of tools to gather information on the nature of distant galaxies and the workings of the "hidden" world around us.

LENSES GALORE - HUBBLE FINDS LARGE SAMPLE OF VERY DISTANT GALAXIES

[heic 0814]

New Hubble Space Telescope observations of six spectacular galaxy clusters acting as gravitational lenses have given significant insights into the early stages of the Universe. Scientists have found the largest sample of very distant galaxies seen to date: ten promising candidates thought to lie at a distance of 13 billion light-years (~redshift 7.5).

By using the gravitational magnification from six massive lensing galaxy clusters, the NASA/ESA Hubble Space Telescope has provided scientists with the largest sample of very distant galaxies seen to date. Some of the newly found magnified objects are dimmer than the faintest ones seen in the legendary Hubble Ultra Deep Field, which is usually considered the deepest image of the Universe.

By combining both visible and near-infrared observations from Hubble's Advanced Camera for Surveys (ACS) and Near Infrared Camera and Multi-Object Spectrometer (NICMOS), scientists searched for galaxies that are only visible in near-infrared light. They uncovered ten candidates believed to lie about 13 billion light-years away (a redshift of approximately 7.5), which means that the light gathered was emitted by the stars when the Universe was still very young — a mere 700 million years old.

Credit: NASA, ESA, and Johan Richard (Caltech, USA)

The picture shows Abell 2218, a rich galaxy cluster composed of thousands of individual galaxies. It sits about 2.1 billion light-years from the Earth (redshift 0.17) in the northern constellation of Draco. When used by astronomers as a powerful gravitational lens to magnify distant galaxies, the cluster allows them to peer far into the Universe. However, it not only magnifies the images of hidden galaxies, but also distorts them into long, thin arcs.

CONTENTS

ST-ECF

Head
Robert A.E. Fosbury
+49 89 320 06 235
Robert.Fosbury@stecf.org

Advanced Data Products
Jeremy R. Walsh
+49 89 320 06 248
Jeremy.Walsh@stecf.org

Public Outreach
ESO education and Public Outreach:
Lars Lindberg Christensen
+49 89 320 06 291
lars@eso.org

The Space Telescope-European Coordination Facility
Karl-Schwarzschild-Str.2
D-85748 Garching bei München, Germany

Websites
www.stecf.org
www.spacetelescope.org

Phone
+49 89 320 06 291

FAX
+49 89 320 06 703

Hotline (e-mail)
stdesk@stecf.org

E-mail
<user>@stecf.org

ST-ECF Newsletter

Editor
Richard Hook, Richard.Hook@stecf.org

Editorial Assistant
Britt Sjöberg, Britt.Sjoeberg@stecf.org

Layout & Illustrations
Martin Kornmesser
Katarina Sopcic

Production Coordination
Lars Lindberg Christensen

Printed by
TypeSet, München

Published by
ST-ECF/ESO ePOD

SCISOFT 7.2 — PATCH RELEASE

Richard Hook

Since the last release major release of the ESO Scisoft software collection of astronomical software in June 2007 several updates have been needed and a few problem identified and fixed. Although new DVDs are not made, we update the distribution of the collections that are available through the web, as well as installing the updated version internally at ESO.

Two patch releases (7.1 and 7.2) have been made and the more significant updates include:

- IRAF updated to 2.14.1.
- STSDAS/TABLES updated to 3.8.
- MIDAS updated to 08SEPpl1.0.
- Updates to many IRAF and Python packages.
- A newer and more powerful version of GAIA (4.2.2).
- Tiny Tim updated to V7.0 – with WFC3 support.

In addition, the new items added to the collection include:

fitscut, VisIVO, graphviz, funtools, minicrush, BoA and VirGO.

Scisoft 7.2 remains on Fedora Core 6 Linux.

For more details, and to download the collection, please go to the Scisoft web pages (www.eso.org/scisoft).

Cover Image: *This remarkable Hubble picture is the first visible light image ever taken of a planet circling another star. Fomalhaut b (in small box) is named after the bright naked-eye star that it orbits. The planet is a billion times fainter than its parent star and is almost drowned out by scattered light. This picture was taken using the coronagraph in the High Resolution Camera on Hubble's Advanced Camera for Surveys. Fomalhaut itself is at the centre of the image but was hidden behind an obscuring finger within the camera to allow the planet, and the ring of dust with which it interacts, to be seen. Observations over the last few years (insert) show that Fomalhaut b is clearly orbiting the parent star with a period of 872 years. Credit: NASA, ESA and P. Kalas (University of California, Berkeley, USA).*