The WFPC2 Archival Parallels

Principal Investigator: Dr. Stefano Casertano
Institution: Space Telescope Science Institute
USA/MD
Electronic mail: stefano@stsci.edu

Scientific category: COSMOLOGY
Total budget amount: $ 512,000
Scientific keywords: GALAXY FORMATION AND EVOLUTION, GALAXY MORPHOLOGY AND STRUCTURE, AGN PHYSICS, RESOLVED STELLAR POPULATIONS, GRAVITATIONAL LENSING
Special Proposal Types: **AR Legacy**

Abstract

We propose to process and combine with state-of-the-art tools the majority of the WFPC2 images obtained as part of Archival Pure Parallel Program, consisting of observations of over 3,000 random pointings primarily in the wide WFPC2 $UBVI$ filters. We will produce combined and cosmic-ray cleaned images, as well as object catalogs, for over 2,000 of these pointings. We will use these data to address a wide range of science topics: measuring the cosmic shear on scales from $20''$ to $2'$, discovering ~ 50 starforming galaxies at $z \sim 4$, finding optical counterparts to AGNs in wide-area radio and X-ray catalogs, improving the determination of the scale length of the Galactic disk, and studying stellar populations down to $\sim 1 M_\odot$ for about 50 separate lines of sight in the LMC. The same data will be available to the astronomical community for a wide variety of other investigations, thus helping realize the legacy of WFPC2 parallel images.

	Investigator	Institution	Country
PI:	Dr. Stefano Casertano	Space Telescope Science Institute	USA/MD
CoI:	Dr. Henry Ferguson	Space Telescope Science Institute	USA/MD
CoI:	Dr. Andrew S. Fruchter	Space Telescope Science Institute	USA/MD
CoI:	Mr. Richard Hook*	Space Telescope - European Coordinating Facility	Germany
CoI:	Dr. Paolo Padovani*	European Space Agency - Space Telescope Science Institute	USA/MD
CoI:	Prof. Nino Panagia*	European Space Agency - Space Telescope Science Institute	USA/MD
CoI:	Dr. I. Neill Reid	Space Telescope Science Institute	USA/MD

Total number of investigators: 7
Number of ESA investigators: 3 (indicated by * after name)

■ Scientific Justification

Introduction

The capability of HST for parallel observations increases its scientific productivity by enabling observations of random fields on the sky, at no additional cost in telescope time. Pure parallel observations, attached to unrelated primary observations, benefit any science program that can take advantage of what are essentially random pointings on the sky, such as statistics of stellar distributions, faint galaxies, and so on. Such observations were initially assigned to specific General Observer programs, and have formed the basis for very successful science programs, such as the Medium Deep Survey Key Project (MDS, Griffiths et al 1993; GO 5369ff) and the programs of Windhorst (GO 6609ff), MacKenty (GO 6253), and Walsh (GO 5584ff). After several years of post-Servicing Mission observations, the Archival Pure Parallel Program (APPP) was established in mid-1997 so that unassigned parallel observations would be taken in such a way as to benefit a broad range of science programs. We propose to process a substantial fraction of the WFPC2 parallel observations using improved drizzle techniques, and use those data to study weak lensing, galaxy evolution, stellar populations and Galactic structure, as well as making the processed data available to the general community.

The WFPC2 part of the APPP has used primarily the Hubble Deep Field filters (F300W, F450W, F606W, F814W), which offer the best compromise between wavelength separation and sensitivity, and some medium-width and narrow filters, such as the Hα filter F656N. The observing strategy was adapted to the duration of each parallel opportunity, in order to obtain exposures suitable to address a broad range of science questions; for more details on the strategy see `http://www.stsci.edu/instruments/parallels/HSTParallel.html`. As of August 2001, the APPP has taken 10,647 exposures in 3385 visits; most visits represent unique, non-overlapping pointings, although a small fraction repeats previous visits and thus results in partially overlapping images. About 65% of the pointings have data only in V (F606W); 20% have two filters available; 10% have three; and the remaining 5% have four or more. The spatial distribution of pointings and exposure times is given in Figure 1.

However, the enormous scientific potential of the WFPC2 APPP observations remains to date largely untapped. The primary reason is accessibility: the WFPC2 images and the resulting science products are not available in a readily usable form. Challenges that make using WFPC2 archival images difficult include the reliable rejection of image artifacts, such as cosmic rays and hot pixels; the combination of coaligned and non-coaligned exposures; and the extraction and cataloguing of source information. Some of these services are currently available for a subset of the images—the Space Telescope-European Coordinating Facility (ST-ECF) now offers identification of exactly coaligned images and their automatic combination with cosmic-ray rejection, but not improved hot-pixel identification (Micol et al. 1997). No comparable service is available for images that are not perfectly aligned.

Scope of the Project

We propose to fully process more than half, and potentially all, of the WFPC2 observations that are part of the APPP program in a uniform way, thereby providing the

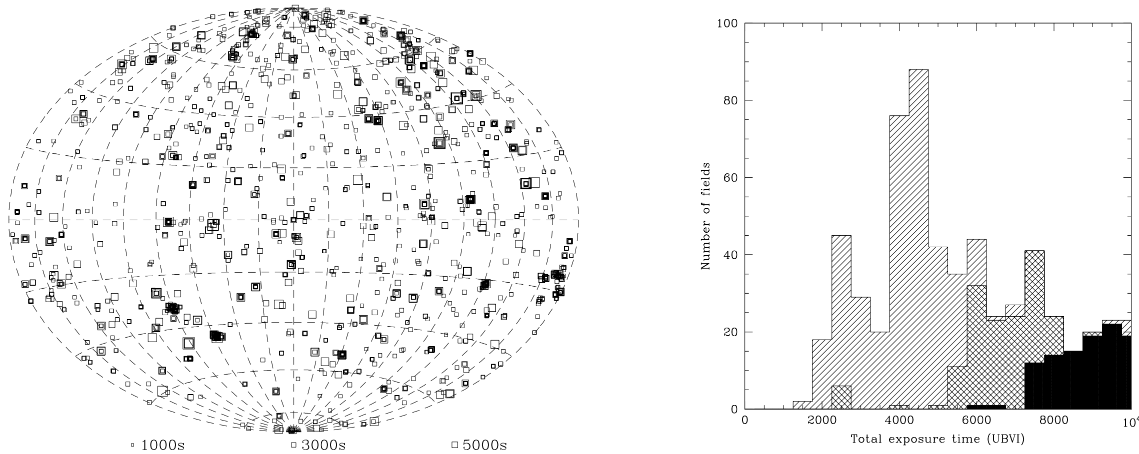

Figure 1: Left: distribution of APPP fields in galactic coordinates. The size of each point is scaled by the total time available at that location. The two concentrations in the bottom left quadrant are the Lange and Small Magellanic Cloud. Right: histogram of the number of fields vs. total available time for all fields with observations > 500s in at least two (shaded), three (cross-hatched), and four filters (filled).

astronomical community with a database of ready-to-use, quality-controlled images covering a total of about 2 square degrees to $V \sim 25.5$ at HST resolution. For each visit and filter, we will produce fully cosmic-ray and hot-pixel corrected images, noise and context images and cross-correlation noise kernels to aid in estimating the statistical properties of each image. We will also generate basic object catalogs, including star/galaxy classification and colors if available; we will produce supplemental catalogs using point-source software in regions of high stellar density, such as at low galactic latitude or in sightlines that intersect narby galaxies. This massive processing will be made possible by automated combination scripts based either on cosmic-ray rejection or on the DRIZZLE algorithm (Fruchter and Hook 2001); details are given in the Data Analysis plan. We will inspect each input image and processed product carefully, applying both visual and automated quality controls; this will most likely be the most labor-intensive part of the process. The resources we ask will enable us to fully process at least 2,000 of the visits, selected for maximum scientific return on the basis of total exposure time, number of filters, and overlap with regions of interest. All other visits will receive at least minimal processing, such as cosmic-ray and hot-pixel rejection; full processing will be carried out to the extent of the available time.

All this information—final processed images, improved image information, and object catalogs—will be made available to the community via a direct web interface, and possibly as Prepared Datasets via the HST Archive (see `http://archive.stsci.edu/prep_ds.html`). One of the Co-Investigators (PP) is a member of the STScI Archive division and will facilitate the process of making our finished data available to the community.

Our ambitious program will be made possible by the experience accumulated by the proponents in the analysis of WFPC2 images and in the handling of large databases. Our team includes the leaders of the WFPC2 data analysis for both Hubble Deep Fields and of the object cataloging for the HDF-South; the inventors of the DRIZZLE technique; a collaborator of the MDS; and a member of the STScI Archive staff with extensive experience in working with large catalogs. Our program will truly enable the legacy of the vast database of WFPC2 archival parallels, identified as a major goal of Archival Legacy proposals in the Call for Proposals; it falls fully within the recommendations of the Hubble Second Decade Committee (`http://sso.stsci.edu/second_decade/recommendations/index.html`); and represents a necessary step in fully integrating WFPC2 archival observations in the National Virtual Observatory concept.

Scientific value of WFPC2 Parallel images

The processed images and catalogs we propose will enable scientific investigations on topics ranging from local Galactic structure to fundamental cosmology questions. When defining the APPP, the Parallel Working Groups identified specific science topics, such as searching for low-mass stars, studying properties of faint galaxies in parameter space insufficiently covered by the small area of the Hubble Deep Fields, identifying galaxies at specific redshifts ($z \sim 2.4$ and 2.9) via medium-width filters, tracing resolved stellar populations in and just ourside nearby galaxies, measuring gravitational shear, finding rare strong gravitational lenses, and collecting an atlas of UV morphology of galaxies.

Our team will focus on five specific investigations which span a broad range of topics: 1) an improved measurement of the cosmic shear on small scales and of the mass of galaxy haloes; 2) the determination of the frequency, morphology, and cosmic star formation rate for bright galaxies at redshifts 3-5 (U- and B-band dropouts); 3) the optical identification of AGN and other sources detected at other wavelengths (X, radio, etc); 4) the structure of the Milky Way; and 5) resolved stellar populations in nearby galaxies. These topics reflect our specific scientific interests, and illustrate the broad range of scientific applications of WFPC2 parallel data; in truth, any project that can benefit from high-resolution images reaching $V \sim 25.5$ or $U \sim 24$ in random locations on the sky, or in locations within a few arcmin of likely targts for HST programs, will benefit from the WFPC2 APPP and thus from the images and data we will produce.

Weak Lensing

Weak gravitational lensing has recently become one of the most direct ways to probe for the existence and properties of dark matter structures, both around individual galaxies and unassociated with light. Cosmic shear, the signature of large-scale mass structures, has now been measured convincingly both from the ground and from space. Ground-based measurements cover very large areas and can typically measure shear on scales $> 1'$; the relatively low density of suitable sources makes it difficult to measure shear on smaller scales. HST can measure the shape of faint, small galaxies and thus reach source densities exceeding 100 per square arcminute, thus enabling measurements of cosmic shear on scales $\sim 20''$. A new, careful analysis of the 350 fields included in the Medium Deep Survey images yield a high-quality measurement of cosmic shear from $20''$ to $2'$ (Figure 2). With about five times as many fields, although of lower depth, we expect the Archival Parallel data to yield a factor of two improvement in the quality of the measurement. Note that, aside from

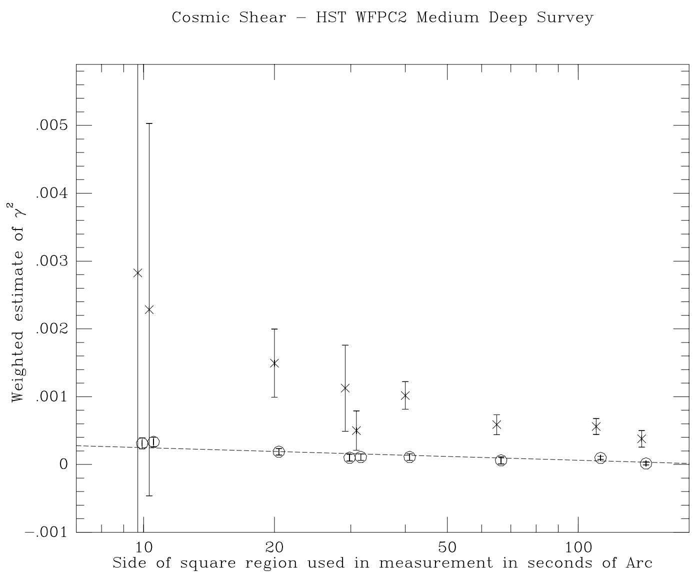

Figure 2: Measurement of the cosmic shear from 350 MDS parallel images. The points with error bars indicate the measured shear variance averaged over square regions of the given size, with errors estimated from Monte Carlo simulations. The lower set of points give the result of null experiments on the same data. We expect that the APPP images will reduce the error bars on this measurement by a factor of 2 to 2.5.

the ability to measure smaller scales, the Archival Parallel data also offer the advantage of a very large number of independent sightlines, thus providing perhaps the best dataset in existence in terms of reducing the impact of cosmic variance.

The APPP data will also provide a substantial improvement on the current measurement of halo masses via galaxy-galaxy lensing; as for cosmic shear, space-based data offer a significant advantage on small angular scales over the large amount of ground-based measurements The latter constrain primarily the cumulative mass distribution at $\gtrsim 200$ kpc, which typically includes other galaxies physically associated with the lens, while the HST measurements can be used to measure the mass profile of individual haloes in the range 30–300 kpc.

Galaxy evolution and star formation rate

The last several years have witnessed an explosion of information on galaxies at high redshift $z > 2$. Thanks in large part to observations from Keck and HST we now have reasonable constraints on the luminosity function, of galaxies at $z \sim 3$ and $z \sim 4$. For the $z \sim 3$ sample, the observations provide estimates of stellar masses, dynamical masses, and clustering bias. While this flood of information is improving our knowledge of this high-redshift population, constraints on the evolution of galaxies to lower redshift $z < 3$ and to higher $z > 4$ are still largely qualitative. Larger statistical samples of high$-z$ galaxies are needed, especially those with HST images.

The HDF observations provide information on sizes and morphologies in the rest-frame UV for samples of order 10 galaxies with $L > L^*$ and for of order 200 galaxies up to 2 magnitudes fainter. Comparisons of the size distribution at higher and lower redshift are severely hampered by the small number statistics. One of the motivations for planning F300W observations as part of the APPP survey was to improve on these statistics. There are ~ 61 fields in the archive reaching to limiting depth $U_{300} > 23.8$. These images cover 30 times the area of the HDF-N and HDF-S fields combined and probe to a depth $1.7L^*$ at $< z >= 2.75$. We expect secure identifications of roughly 70 Lyman-break galaxies at this redshift $\pm\Delta z = 0.5$.

At higher redshift $< z >\sim 4$ the combined HDF observations also yielded about 8 galaxies $L > L*$. In the APPP database there are 126 fields with limiting B-band depth greater than $B_{450} = 24.8$ (with at least 1ks each in V and I) , yielding an expected sample of 50 $L > L^*$ galaxies at $< z >= 4 \pm 0.5$.

These sample sizes are large enough to provide an interesting differential comparison of the rest-frame UV scale lengths of galaxies between two samples roughly a factor of 2 different in age (assuming $z_{formation} \sim 7$). The samples are large enough to determine also if there are systematic differences in morphology (e.g. asymmetry or concentration) between $L \sim L^*$ galaxies at the two different redshifts.

This morphological information from HST will be a vital complement to groundbased studies and to studies from ACS. Because ACS has only a small field of view in the ultraviolet, coverage of significantly greater depth×area at HST resolution is likely to be possible only after WFC3 is installed. Meantime the APPP has a significant store of morphological information waiting to be explored.

Optical Counterparts of Catalogs at Various Wavelengths

This proposal will produce a homogeneous catalog of optical positions and magnitudes down to $V \sim 25.5$ over an area of about 2 square degrees. This resource will allow optical identification of sources detected in large-area surveys at other wavelengths (radio, IR, X-ray). Optical identification represents the first, necessary step towards source classification, and is also used to screen out less interesting sources; for example, AGNs can easily be separated from stars detected in X-ray surveys because their X-ray-to-optical flux ratio is much larger.

Currently available optical catalogs (e.g., APM, GSC-II) reach relatively bright ($\sim 20-21$) magnitudes and therefore are inadequate to determine the optical counterparts of deep surveys, which usually require dedicated observations. For example, at 1 mJy, the limit of the FIRST radio survey (e.g., Becker, White, & Helfand 1995), a radio-loud quasar will have $V \approx 24$. The same magnitude corresponds also to the optical counterpart of a typical radio-quiet quasar with soft X-ray flux $\sim 10^{-15}$ erg cm^{-2} s^{-1}, corresponding to the flux limit of a typical XMM exposure.

Given our relatively small survey area (~ 2 square degrees), our products will be most useful when combined with all-sky/large-area surveys. For example, our catalog will be cross-correlated with the VLA-FIRST and the NVSS surveys. The VLA-FIRST survey will cover over 10,000 sq. deg. of the North and South Galactic Caps down to 1 mJy at 1.4 GHz; the NVSS covers all the sky north of $\delta = -40°$, reaching 5 mJy at 5 GHz. A cross-correlation of the FIRST catalog with WFPC2 images has been started (Rick White, private communication), but identifying optical counterparts of radio sources and assigning optical magnitudes has proven extremely labor-intensive. Our catalog will make such a task trivial.

Further applications include the identification of IR-bright/optically-faint sources (i.e., very red or obscured sources) by cross-correlating it with the 2MASS catalog (`http://www.ipac.caltech.edu/2mass/`), and optical counterparts to X-ray sources found in the XMM serendipitous survey (Watson et al. 2001) and in the WGA catalog (White, Giommi, &

Angelini 1995). The XMM serendipitous survey will cover a relatively small area ($\sim$ 100 sq. deg. per year), with a small overlap ith our catalog, but to deep flux limits ($\approx 10^{-15}$ erg cm^{-2} s^{-1}). Even at the depth of the WGA catalog (a few $\times 10^{-14}$ erg cm^{-2} s^{-1}), the typical AGN magnitude ranges between $\sim$ 21 and $\sim$ 23.

Galactic structure

Two centuries after Herschel, starcounts remain a highly effective means of probing the structure of our Galaxy. Ground-based surveys have revolutionised our knowledge of the Disk, notably through the identification of the extended thick-disk component, but those studies are limited to relatively bright magnitudes by the fundamental issue of star/galaxy separation. Fainter than R 20, galaxies increasingly outnumber stars at moderate and high latitudes, so small systematic classification errors can seriously bias analysis. HSTs unparalleled resolution offers the most efficient method of circumventing this problem, but the 5.3 square arcmin field of view of WFPC2 limits the probative power of individual programs. Our project, combining data from several hundred pointings, provides the necessary areal coverage to tackle large-scale structural questions.

Starcount analysis requires observations in at least two passbands, to permit discrimination between disk and halo. Our goals are to:

1. Examine the vertical structure at the interface between the thick disk and halo, 1-3 kpc above the Plane;
2. Test the asymmetry of disk and halo by comparing numbercounts at positive and negative longitudes;
3. Extend the analysis by Zheng et al. (2001) of the stellar luminosity function at large distances above the Plane;
4. Probe the radial structure of the disk through counts at low latitudes, notably the disk scalelength. Recent estimates (Zheng et al. 2001, Drimmel & Spergel 2001) favour 2-3 kpc for the latter, relatively short values, but are based largely on data at moderate latitudes. HSTs resolution, coupled with reddening estimates, allows us to probe nearer the mid-Plane (Figure 3).

These studies require better than 10% photometry to B$\sim$25, V$\sim$24 and I$\sim$23, which in turn demand exposure times exceeding 500 seconds for V and I, or 1000 seconds in B. At least two of the three filters must be available to the required depth, in order to provide a broad-band color that can be used to discriminate between disk and halo stars. The 370 pointings that satisfy these conditions are well distributed over the celestial sphere, and can be grouped to give sufficient solid angle coverage to provide sufficient statistics to address these four issues.

Resolved Stellar Populations 1: The Local Group

The study of stellar populations has greatly benefited of the superb performance of HST in angular resolution and photometric stability. Thanks to these capabilities it is now possible to resolve and characterize stars of masses as low as 0.9 $M_\odot$ in the Magellanic Clouds and 3 $M_\odot$ in the Local Group. Note that the main problem to overcome in the study of stellar clusters are confusion and crowding that set the practical limit to completeness at brighter magnitudes than the telescope size and the exposure time may allow one to reach (e.g. Romaniello 1998, Panagia et al. 2000, Romaniello et al 2001).

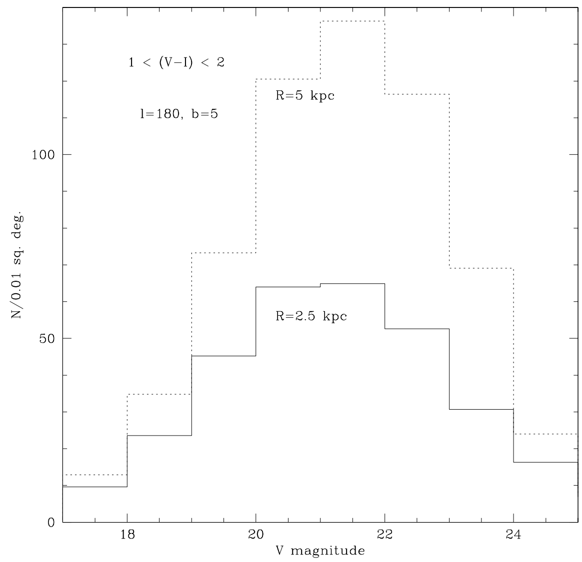

Figure 3: Expected number count distribution for a low galactic latitude field near the anticenter direction, for two values of the disk scale length, showing the sensitivity of the counts to the structure of the Galactic disk. The counts are for the color interval $1 < (V - I) < 2$ and are scaled to an area of 36 square arcmin, about 7 WFPC2 fields. Nine fields with at least two broadband filters exist within 15° of the anticenter; 34 within 20°. In practice, the counts will be compared against a Galaxy model computed separately for the coordinates of each field.

The APPP program for nearby galaxies has been designed to reach S/N $\gtrsim$ 10 for stars as faint as V=24 for all spectral types in BVI, and for early B and O stars in U; relatively deep Hα exposures (F656N) were taken for multi-orbit visits near Local Group galaxies. Each visit thus offers accurate four-band photometry extending from shortward of the Balmer discontinuity to around the Paschen discontinuity. Comparison with model atmospheres allows the measurement of luminosity and temperature for all stars, and of reddening corrections for stars hotter than about 10000K and in the range 6000-8000K (Romaniello 1998, Romaniello et al 2001). For each such field we can obtain a log(T_{eff})-log(L) HR diagram including several thousand stars down to the confusion limit (0.9 $M_\odot$ in the Magellanic Clouds, and $\sim 3M_\odot$ in the rest of the Local Group. The Hα equivalent width can be used to identify OB stars with strong mass loss, Be and Ae stars, and pre-main-sequence (PMS) stars of moderate masses (strong line T Tau stars, M$\sim 1 - 2$ $M_\odot$), thus seaprating PMS stars of recent stellar generations from old, evolved stars of the general field (Romaniello 1998, and in preparation). This potential is shown in Figure 4, based on data derived from GO observations of SN1987A but very similar to APPP observations near the LMC. About 22,000 stars were identified within 30pc of the supernova, and their temperatures and luminosities measured; individual reddening corrections were derived for about 2500 stars (Romaniello 1998, Panagia et al 2000).

Resolved Stellar Populations 2: Stellar Halos of Nearby Galaxies

Within the APPP database there are dozens of lines of sight that pass through or next to nearby galaxies $D < 7Mpc$. For galaxies with distance moduli $(M - m) < 28$ 1000s in F814W is sufficient to reach 1 magnitude fainter than the tip of the RGB of a metal-poor population ([Fe/H] < -0.7). For fields with 2 exposures in F814W, the RGB can be resolved for galaxies with $(M - m) < 29$ (6.3 Mpc).

The run of star-density with radius from the host galaxy will provide statistics on the radial density profile of galaxy stellar halos. The F606W-F814W color distribution

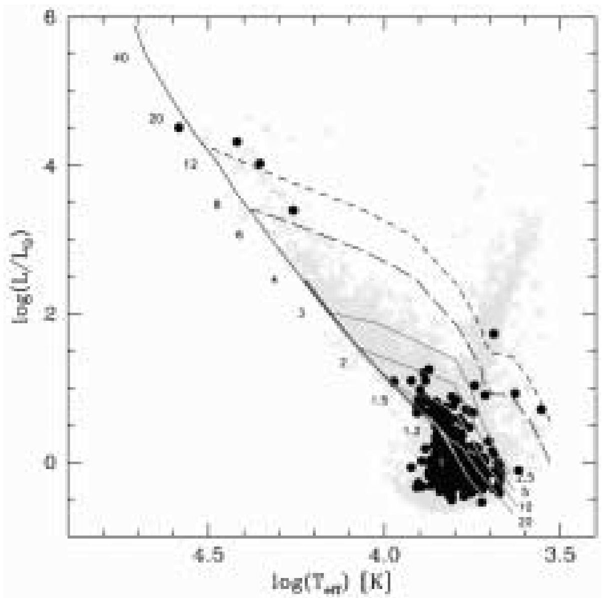

Figure 4: The HR diagram for SN 1987A field with the positions of stars with strong Hαexcess (dots) overlayed on the general stellar population (grey squares) found in the WFPC2 field. For reference, we show the theoretical ZAMS with marked position for stars of various masses. Also shown are 2.5-20 $Myrs$ PMS isochrones and the birthlines for accretion rates of 10^{-4} (short-dashed line) and 10^{-5} $M_\odot$ yr^{-1} (long-dashed line).

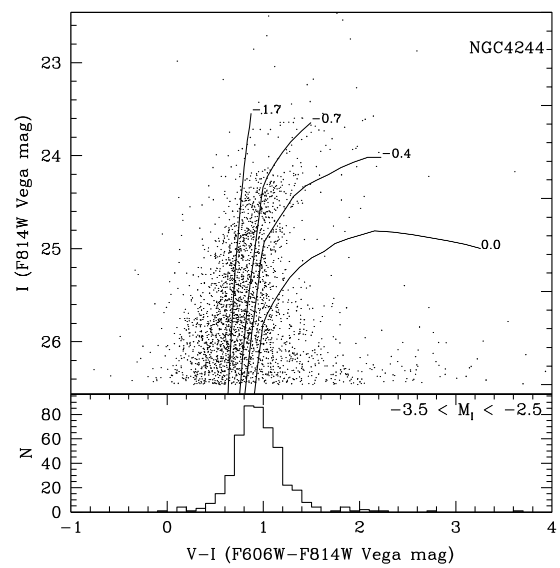

Figure 5: Preliminary color-magnitude diagram of the halo of NGC4244 obtained with 1.5 orbits each in F814W and F606W. The observations reach $M_I = -1$ at $S/N = 5$. Isochrones of age 14 Gyr from Girardi et al. 2000 are superimposed on the data. The distribution in V-I (bottom panel) provides constraints on the metallicity distribution. The data shown here are from GO-9096 (pointed observations by Ferguson and collaborators) but similar data sets exist as archival pure parallels.

provides information on the spread in metallicity in the galaxy halo, and comparisons of this distribution as a function of radius within individual galaxies and between different galaxies will provide unique constraints on the formation mechanisms of galaxy halos. The data can also be searched for clumps in the spatial distribution of halo stars that could be remnants of disrupted dwarf galaxies. For some galaxies it will be possible to measure the apparent magnitude of the tip of the red giant branch (TRGB) providing an independent distance estimate for comparison to estimates from Cepheids, PNLF and the Tully-Fisher relation.

References

Becker, R. H., White, R. L., & Helfand, D. J. 1995, ApJ, 450, 559
Condon, J. J., et al. 1998, AJ, 115, 169
Drimmel & Spergel, 2001, ApJ 556, 181
Micol, A., Pirenne, B., & Bristow, P., 1998, in ADASS VII, ASP 145, p. 45
Romaniello, M. 1998, PhD Thesis, Pisa
Panagia, N., et al 2000, ApJ 539, 197
Romaniello, M., et al 2001, AJ, in press
Watson, M. G., et al. 2001, A&A, 365, L51
White, N. E., Giommi, P., Angelini, L., 1995, http://wgacat.gsfc.nasa.gov/
Zheng et al. 2001, ApJ 555, 393

Analysis Plan

Data processing

Our data processing will be divided into three major steps: identification and preparation of the images to be combined; combination and quality verification; and object cataloging. The procedures described in the first two steps have been recently field-tested on the data used in the STScI project to search for Supernovae at high redshift (SNaZ; STScI newsletter). We found that in practice, full processing of an eight-image data set requires approximately four hours, including quality verification. We expect therefore to be able to process about two pointings/person/day, or about 2,000 pointings with two people over two years. This is a somewhat conservative goal, as we expect that our efficiency will increase as we gain experience on the processing; we might thus be able to process a larger number of datasets, up to the full set of WFPC2 APPP data, in the stated time period.

1) Identification and preparation of images

Exposures taken in about 30% of the visits—and 40% of the visits with more than one filter—are not nominally coaligned, so that rejection of cosmic-ray hits requires special procedures based on the DRIZZLE algorithm (Fruchter and Hook 2001). Even those that are nominally coaligned can present small accidental shifts that require special consideration before cosmic ray rejection. The ST-ECF and the Canadian Astronomical Data Center have approached the task of identifying coaligned images, and measuring accurately the shifts between non-coaligned images, as part of general development of value-added archive services and are well advanced with automatic procedures available through archive web interfaces (Micol, Pirenne & Bristow, 1997). Their experiments show that WCS information in image headers is less reliable than information obtained from the jitter files which form part of the HST engineering stream and the latter have been used, along with related data quality information, to populate offset values for groups of dithered WFPC2 images. Work currently in progress will apply cross-correlation methods to determine shifts between associated images in the cases where the jitter information is inadequate. Where possible, we will use their results. We will also identify those cases in which multiple visits overlap, and determine whether it is desirable to combine them, depending on the total exposure time within each visit and on the degree of overlap.

2) Image combination and quality verification

Image combination will follow different paths depending on whether individual eposures are exactly coaligned (within 0.1 pixels) or not. Exactly coaligned images will be cosmic-ray rejected using standard coincidence algorithms; indeed, we expect that the ST-ECF on-the-fly combination procedure will provide an optimal cosmic-ray rejection. Improved hot pixel rejection will be carried out using the Supplemental Darks that have been obtained since 1997 at the rate of three per day, but are not incorporated in the routine pipeline calibration and processing.

However, dithered data sets present both an opportunity and a challenge. They are an opportunity because dithering frequently allows identification of image defects (hot or bad pixels), and generally improves the spatial quality of the final WFPC2 image. At the same time, reducing dithered data requires significantly more computation, and until recently more human attention. However, one of us (Fruchter) in collaboration with his graduate student (P. Vreeswijk) has recently created a new script, "METADRIZZLE", which combines many of the tasks that have been created by Fruchter and collaborators to handle dithered data. The present version of METADRIZZLE automatically determines the offsets from the image headers and combines the images twice; the first time via median filtering to produce a comparison image for the identification of cosmic rays, the second to produce the final image. An intermediate step can be added to allow an iterative refinement of the image offsets. The software has been field-tested in an STScI program for supernova searches, and works very well. We will upgrade the software to use pointing information from the jitter file or from the ST-ECF association information.

Processed images will also be registered to a common astrometric system, using deep stellar catalogs such as the USNO and the recently released GSC2. These catalogs contain approximately 3 stars per WFPC2 field at high galactic latitude, and should allow a local accuracy of about $0\farcs1$ to $0\farcs2$.

3) Object cataloging

The cataloging will carried out with a carefully supervised automated pipeline. Sources will identified using SExtractor (Bertin & Arnouts 1996). Different passes through the software (with different convolution kernels and thresholds) will be used for optimal identification of point sources and galaxies. The final merged catalog will thus have reliable star-galaxy separation and a reasonable level of source deblending. For point sources the catalogs will include aperture photometry, along with aperture-corrected data corrected for instrumental effects such as CTE. The extended source parameters will include the standard SExtractor magnitudes and shape parameters, as well as optimal Gaussian-weighted complex ellipticities useful for lensing.

For fields at low Galactic latitude, where the vast majority of the objects are stars and source confusion can be a problem, we will produce alternate point source catalogs using standard stellar photometry packages such as DAOPHOT.

Effect of DRIZZLE on PSF and on measurement of the cosmic shear

We have not yet tested DRIZZLE for its effects on PSF shape, but based on the work that has already been done on drizzled images (Fruchter and Hook 2001) we expect it most probable that these will be superior to images based on single pointings. Should tests we will undertake not support this claim, we will nonetheless have no difficulty using dithered data. The METADRIZZLE script can be asked to create individual "corrected" images. These are images which are identical to the invidual input images, except where a cosmic ray or chip defect has been flagged. At these positions the data from the individual image is replaced by an estimate obtained from the combined dithered data. The lensing signature can then be obtained from each of these images separately and combined statistically. As typically somewhere between 2 and 10% of pixels are affected by cosmic rays or defects, and as the cosmic rays in individual images are uncorrelated (and the chip defects are offset by dithers), any difference between the estimate and the "truth" should not have a measurable effect on the final lensing signature.

Data products

As a result of our proposed Legacy program, we will make available to the astronomical community the following data products:

- Corrected, CR-rejected, combined images, using DRIZZLE if exposures are not aligned, with improved astrometry and hot pixel rejection
- Exposure and noise images, cosmic ray masks—which enable other users to reprocess the images without repeating the complex cosmic-ray rejection—and context images, which identify which input images contribute to each output pixel
- A master database identifying all images, with improved pointings and other quality remarks
- Merged object catalogs for all fields, with magnitudes, size and shape parameters, and basic star/galaxy classification, and special point-source catalogs for high-density regions

These products will be delivered progressively as processing continues, after allowing approximately six months to refine the processing algorithms and develop standard procedures. They will be made available to the community via web interfaces and, to the extent possible, via established archives.

Budget Narrative

The massive processing effort envisioned as part of this Legacy program will require two dedicated persons for two years: a Postdoctoral Fellow to develop the necessary programs and procedures and exercise daily scientific supervision of the process, and a Data Analyst to carry out the processing itself. Our procedures will be largely automated, making the program feasible within the stated time period, but we will maintain a high level of quality testing on all input and combined images. We therefore request support for one Postdoctoral Fellow for two years, at $ 100,000/year, and for one Data Analyst for two years, at $ 100,000/year. We also request support to buy back 10% of the PI time over two years, for a total of $ 17,000/year. The total funding request for salaries and personnel costs is thus $ 217,000 /year, or $ 434,000 over two years.

The project will also require fast workstations for processing and a large amount of storage space, both for processing and to store finished products. We envision two top-level, Sun Blade 1000 workstations with dual 750MHz processors, at a configured price of approximately $ 15,000 each, and a total of 650GB of mass storage space, in three 218 GB Sun StorEdge disk packs at $ 11,000 each. We request in addition funds to purchase one portable computer for a co-investigator, at $ 5,000. The total cost for computer equipment is thus $ 68,000 for the duration of the project. Finally, we request $ 5,000 per year for travel for the Postdoctoral Fellow and contribution to publication costs, for a total of $ 10,000.

The total request for the project, in all categories, is $ 512,000.

Previous HST Programs

GO 7293: 12 orbits. Observations completed in 2000, analysis ongoing.