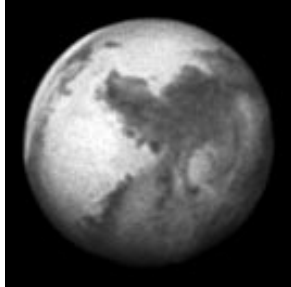
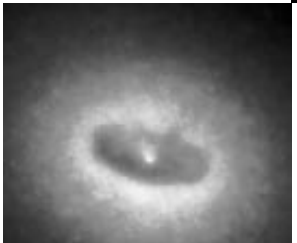
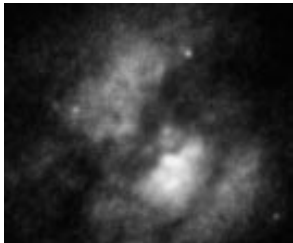
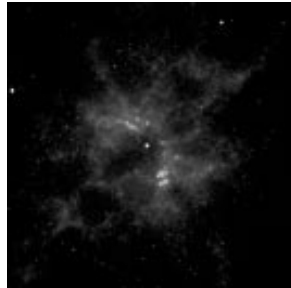


The Future of Space Imaging



The Future of Space Imaging

Report of a Community-Based Study of an Advanced Camera for the Hubble Space Telescope

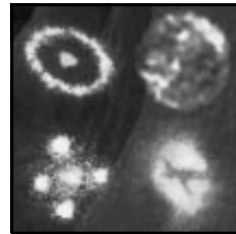


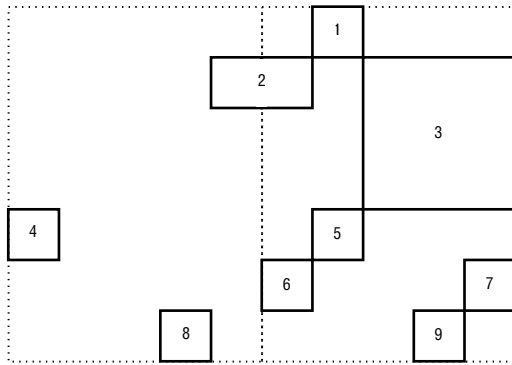


When Lyman Spitzer first proposed a great, earth-orbiting telescope in 1946, the nuclear energy source of stars had been known for just six years. Knowledge of galaxies beyond our own and the understanding that our universe is expanding were only about twenty years of age in the human consciousness. The planet Pluto was seventeen. Quasars, black holes, gravitational lenses, and detection of the Big Bang were still in the future—together with much of what constitutes our current understanding of the solar system and the cosmos beyond it. In 1993, forty-seven years after it was conceived in a forgotten milieu of thought, the Hubble Space Telescope is a reality.

Today, the science of the Hubble attests to the forward momentum of astronomical exploration from ancient times. The qualities of motion and drive for knowledge it exemplifies are not fixed in an epoch or a generation: most of the astronomers using Hubble today were not born when the idea of it was first advanced, and many were in the early stages of their education when the glass for its mirror was cast.

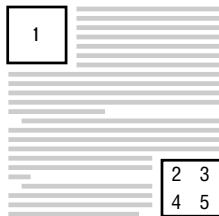
The commitments we make today to the future of the Hubble observatory will equip a new generation of young men and women to explore the astronomical frontier at the start of the 21st century.





FRONT & BACK COVER

1. Globular clusters containing young stars at the core of elliptical galaxy NGC 1275.
Credit: J. Holtzman, NASA.
2. Central core and accompanying jet of giant elliptical galaxy M87.
Credit: Tod R. Lauer (NOAO), Sandra Faber (UCSC), NASA.
3. HST Deployment.
Credit: NASA, Smithsonian Institution, IMAX.
4. Portion of Orion Nebula.
Credit: C.R. O'Dell (Rice University), NASA.
5. Mars, image taken on December 30, 1990, at a distance of 85 million km.
Credit: Dr. Philip James (University of Toledo), NASA.
6. NGC 2440 nebula with central hot star.
Credit: S. Heap, NASA/Goddard Space Flight Center.
7. Cygnus Loop Supernova Remnant.
Credit: J.J. Hester (Arizona State University), NASA.
8. Core of peculiar galaxy Arp 220.
Credit: E. Shaya, D. Dowling (University of Maryland), WFPC IDT, NASA.
9. Core of galaxy NGC 4261 with giant disk of cold gas and dust fueling a suspected black hole.
Credit: Walter Jaffe (Leiden Observatory), Holland Ford (STScI and JHU), NASA.



INSIDE FRONT COVER

1. Photograph of pre-historic petroglyphs by William K. Hartmann.
2. Gaseous ring around SN 1987A.
Credit: NASA, ESA.
3. High resolution UV image of Io.
Credit: Francesco Paresce (STScI), Paola Sartoretti (University of Padova), NASA, ESA.
4. Einstein Cross—gravitational lens G2237+0305.
Credit: FOC IDT, NASA, ESA.
5. "X" structure in core of M51, which is due to absorption by dust and gas, marks the location of a possible black hole.
Credit: H. Ford (JHU/STScI), the Faint Object Spectrograph IDT, and NASA.

Support for this research was provided by NASA under Contract NAS5-26555 and Grant NAGW-3135 through the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy. Design and production assistance was provided by Foxglove Communications.

The Future of Space Imaging

*Report of a Community-Based Study
of an
Advanced Camera
for the
Hubble Space Telescope*

Edited by Robert A. Brown

*Space Telescope Science Institute
3700 San Martin Drive
Baltimore, Maryland 21218*

October 1993

Authorship

This report is the product of the *Future of Space Imaging* study¹ and of the subsequent preparation of the proposal of the Advanced Camera to the ESA M₃ competition², in which the following people participated or to which they contributed. It was our privilege to conduct this study and author this report on behalf of the international community of astronomers. We express our particular gratitude to our many colleagues who responded to the community survey.

James Beletic ¹	Georgia Institute of Technology
Pierre Bely ¹	Space Telescope Science Institute, ESA
J. Chris Blades ¹	Space Telescope Science Institute
Robert Brown ^{1,2}	Space Telescope Science Institute
Christopher Burrows ¹	Space Telescope Science Institute, ESA
Mark Clampin ^{1,2}	Space Telescope Science Institute
France Cordova ²	Pennsylvania State University
Edward Cheng ¹	Goddard Space Flight Center
James Crocker ^{1,2}	Space Telescope Science Institute
Greg Davidson	National Aeronautics and Space Administration
Jean Michel Deharveng ²	Laboratoire d'Astronomie Spatiale, Marseille
Robert Fosbury ^{1,2}	ST-European Coordinating Facility, ESA
John Hoessel ¹	University of Wisconsin
John Huchra ^{1,2}	Center for Astrophysics
Garth Illingworth ²	Lick Observatory
Rolf-Peter Kudritzki ²	University of Munich
Tod Lauer ¹	National Optical Astronomy Observatories
Michael Lesser ^{1,2}	Steward Observatory
Edwin Loh ¹	Michigan State University
Duccio Macchetto ²	Space Telescope Science Institute, ESA
Aden & Marjorie Meinel ¹	Jet Propulsion Laboratory
George Miley ²	Leiden Observatory

Warren Moos ¹	Johns Hopkins University
Jeremy Mould ¹	California Institute of Technology
Susan Neff ¹	Goddard Space Flight Center
Francesco Paresce ²	Space Telescope Science Institute, Torino, ESA
Ronald Polidan ¹	Goddard Space Flight Center
Michel Saisse ²	Laboratoire d'Astronomie Spatiale, Marseille
Francois Schweizer ¹	Carnegie Institution of Washington
Donald York ¹	University of Chicago

Preface

In March 1992, the National Aeronautics and Space Administration (NASA) Program Scientist for Hubble Space Telescope (HST) invited the Space Telescope Science Institute (ST ScI) to conduct a community-based study of an Advanced Camera (AC), which would be installed on a scheduled HST servicing mission in 1999. The study had three phases: a broad community survey of views on the candidate science program and required performance of the AC, an analysis of technical issues relating to its implementation, and a panel of experts to formulate conclusions and prioritize recommendations. An oral report was delivered in December 1992, at which time NASA made a second request to ST ScI, to assist in preparing a proposal for the AC on behalf of the astronomical community to the European Space Agency (ESA) *Call for Mission Ideas* of 26 November 1992 (“M3 competition.”) A joint US-ESA science team was formed for this purpose, and the proposal was submitted in May 1993. This report—*The Future of Space Imaging*—is a compilation of the results of both the study and ESA proposal.

Introduction

The Hubble Space Telescope (HST) is a cooperative program of the European Space Agency (ESA) and the U.S. National Aeronautics and Space Administration (NASA) to operate a long-lived observatory in space for the benefit of the international astronomical community. The three years since the launch of HST in 1990 have been momentous, with the discovery of spherical aberration and the search for a practical solution. HST was designed for servicing in orbit by the Space Shuttle. NASA and ESA are now preparing intensively for the first shuttle repair mission scheduled for December 1993, which should obviate the effects of spherical aberration and restore the functionality expected for HST at launch.

Despite spherical aberration, hundreds of astronomers around the world are analyzing HST data and reporting their discoveries. The early results of HST are outstanding, as is illustrated by the recently published proceedings of an international conference, *Science with the Hubble Space Telescope*.¹ The HST program of astronomical research is unmatched by any ground-based telescope, and it refreshes the original vision of what could be achieved by a great telescope unfettered by the Earth's atmosphere—a vision postponed briefly but not foreclosed by spherical aberration.

HST is now at a crossroads with respect to the second half of its design lifetime—the period after 1997, when an advanced spectrograph and a near infrared camera will be installed on the second servicing mission. In 1999, a third servicing mission will be required to restore the spacecraft orbit, which decays due to atmospheric drag. Owing to the long development time for space-qualified instruments, now is the time to make critical decisions with respect to what scientific instrument can or should be installed during the 1999 servicing mission. *This instrument must be started now to be ready on time in 1999.*

Over the past year and a half, in consultation with the international

astronomical community, we have studied the issue of what best can be done with the 1999 instrument opportunity. We have asked if there is still a role for HST at the turn of the century. *The answer is clearly: Yes!* We are convinced that HST, properly instrumented, will make unique and important astronomical discoveries into the 21st century. This is because no ground-based telescope, no matter how large nor how well equipped with adaptive optics, will match HST in terms of wavelength coverage and image quality over a wide field.

We have asked what type of instrument will provide the greatest scientific benefit to astronomers. We are convinced that it should be a *camera* because otherwise an adequate optical and ultraviolet imaging capability will not be assured in 1999. Furthermore, it should be *advanced* because detector and computer technology is now capable of deriving full value from a 2.4 m diffraction-limited telescope in space.

What combination of scientific programs should define the priorities for the technical performance of the Advanced Camera (AC)? *The breadth and depth of the highly competitive HST imaging science program.* We have assessed the imaging tasks astronomers have proposed for or desired from HST, and we believe the most valuable 1999 instrument would be a camera with both near ultraviolet/optical (NUVO) and far ultraviolet (FUV) sensitivity, and with both wide field and high resolution options.

This report puts forth the scientific case and the technical basis for an AC. It lays the groundwork for a Phase A study to assure feasibility, determine cost and schedule considerations, and inform an Announcement of Opportunity for the astronomical community to participate in the development of the AC.

1. ST-ECF/STScI Workshop: *Science with the Hubble Space Telescope, Proceedings*, eds. P. Benvenuti and E. Schreier, (Munich: European Southern Observatory, 1992.)

Contents

1. ADVANCED CAMERA SCIENCE	1
2. INSTRUMENT STATUS	27
3. OPTICAL TELESCOPE ASSEMBLY	37
4. POINTING PERFORMANCE	43
5. ADAPTIVE OPTICS & HST 'UNIQUENESS SPACE'	59
6. AXIAL VS. RADIAL BAY	77
7. SAMPLING THE IMAGE	83
8. DETECTORS	99
9. FILTERS, GRISMS, POLARIZERS	115
10. INSTRUMENT COMPUTER	121
11. ONE CAMERA DESIGN	127
12. SCHEDULE URGENCY OF THE ADVANCED CAMERA	137

1. Advanced Camera Science

Today, astronomers have at their disposal—or see on the horizon—a variety of telescopes and instruments, in space and on the ground, operating at a wide range of wavelengths. Given this variety, as well as the intense competition for resources to start new projects, we have been mindful that a solid scientific case is needed to support future investments in Hubble Space Telescope (HST)—an observatory conceived in the 1940s, designed and built in the 1970s, and now operational in the 1990s. For this reason, we have developed the issue of whether the Advanced Camera (AC) should be built, and if so, what its characteristics must be to assure it will provide valuable capabilities not duplicated by other ground-based or space-based observatories.

The great potential of the Hubble Space Telescope (HST) is concentrated in two extremely important capabilities: (1) the coverage of wavelengths inaccessible from the ground and (2) superb, stable images. HST's superior image quality consists of two aspects: low image wings, and a wide field of view (FOV).

Ground-based telescopes—even 10 m class telescopes with adaptive optics—cannot rival the ensemble of advantages of space imaging at optical and ultraviolet wavelengths. (See chapter 5.) However, HST itself will achieve this potentiality only when the planned corrections for spherical aberration are in place—and then, completely, only when a camera instrument takes full advantage of the quality wavefront the HST telescope optics deliver to the focal surface.

HST's outstanding capability for imaging is essential for two of its three 'Key Projects', which are critical astronomical programs singled out by an international panel of astronomers prior to launch. Furthermore, imaging requirements that demanded what uniquely HST can provide comprised the justification for a major fraction of the initial General Observer (GO) and Guaranteed Time Observer (GTO) science programs. And even with the spherical aberration problem, a sig-

nificant fraction of the science already done in the first three years of HST operations—and the great bulk of the data volume that has accrued—have come from HST's initial complement of cameras. As of April 1, 1993, HST had acquired 8,651 astronomical images, of which 6,061 were obtained with the Wide Field Planetary Camera (WFPC) and 2,590 with the Faint Object Camera (FOC).

1.1 SURVEY OF CANDIDATE SCIENCE PROGRAM

We have examined the ultraviolet, optical, and near infrared imaging science programs that could and would be done with HST. We put particular emphasis on those scientific projects that could be best or only addressed by HST and an Advanced Camera (AC). Input was solicited from the international community and from individual experts; sample programs were selected from the community input as well as an examination of the accepted proposals from the first three years of HST operations. Each research area was examined to determine (1) what camera capabilities would be required to make the necessary observations, (2) whether ground-based capabilities might have advanced sufficiently to solve nearly or completely the astrophysical problem before the year 2000, and (3) what impact HST observations would or could have on the solution and on the field in general. The list of such projects is given in table 1.1, along with our assessment of the camera characteristics (field, preferred wavelength coverage, time resolution, apodization, filtration) required to address the scientific issues squarely.

The list of scientific programs and opportunities given in table 1.1 is extensive. It argues that the advantages of space imaging are substantial, notably because of: (1) the access to ultraviolet wavelengths, blocked at ground-based observatories, which permits the study of energetic phenomena and highly ionized regions; (2) the dark sky in the near infrared, uncontaminated by OH atmospheric emission, which permits more sensitive investigations of redshifted, reddened, or cooler objects; (3) the stable, superb image quality at all wavelengths, over a wider FOV than at ground-based sites, which generally provides a factor of 10 gain in acuity or factor of 100 gain in spatial information; and (4) the ability to sample flexibly in time, including continuous viewing in some parts of the sky, plus the certainty of clear weather, which allows variable phenomena to be discovered or critically examined.

The conclusion we have drawn from table 1.1 is that an AC would be

extremely valuable if it assured the community's continued access to the benefits of space imaging, and especially if it fully achieved the potential of the HST optics to gather and image astronomical light. Such an AC would be the enabling instrument of choice across the full range of disciplines in astrophysics and planetary science.

Table 1.1. Imaging science with HST

<u>Programs</u>	<u>Bandpasses</u>	<u>Other Char</u>
1. Solar System		
Time Variable Studies	O,U,I	H,W,N,S
Atmospheric Chemistry	U,I,O	N,W,H,S
Atmospheric Dynamics	O,I	W,N,S,0.01
Small Bodies	O	W,C,R,0.01
Extrasolar Planets	O,I	C,0.01,R
2. Stars		
Planetary Nebulae	U,O	0.01,G,N
Stellar Winds	U,O	C,N,P
High Energy Sources, CV's	U	S,W,P
Star Forming Regions, YSO'S	O,I,U	W,C,0.01,N,P
Protoplanetary Disks	O,U,I	C,0.01,P,R
Novae	U,O,I	0.01,N,C
Massive Luminous Stars	U,O,I	N,0.01
Astrometry	O,I,U	W,0.03,0.01
Pulsars, OVV's	U,O	S,N
Brown Dwarfs	I,O	W,C
3. Interstellar Medium		
H II Regions - structure	O,U	W,N,0.03
SNR's	U,O	W,N,0.01
4. Stellar Populations and Systems		
Globular Clusters	U,O	W,0.01
Galactic Center (BW)	I,O	W,0.01
The Galaxy Halo	O,U,I	W
Giant H II Regions	O,U	0.01,W,R
GC's in other Galaxies	O,U	W,0.01
Populations in other Galaxies	O,U,I	W,0.01,N,G
HR Diagrams, IMF's, UV upturn	U,O	W,0.03
5. Galaxies and Clusters		
Nuclear Structure	O,U	0.01,N
Integrated Stellar Pops	U,O,I	W
Mergers and Starbursts	O,U,I	W,0.03,N,T,P
ISM (SNR, PN)	O,U	0.03,N,W
ICM (Cooling Flows)	U,O	W,N,T,0.03
Evolution	O,U,I	W
Morphology	O,I,U	W,0.05
6. AGN's		
AGN Hosts	I,U,O	C,T,N,0.01,P
AGN Environs	O,U,I	W,N,0.03,C,T,P
AGN Evolution	O,I	W,G
Jets	U,O,I	C,P,0.01,W,T

Nearby AGN	U,O,I	T,P,0.01,C,G,H
7. Cosmology		
Distance Scale	O,I,U	W,D,H,0.03
QSO Absorbers	O,U	T,W,G,0.03
Primeval Galaxies	I,O,U	W,N,T
Galaxy Counts	O,I,U	W
Gravitational Lenses	O,U	D,M,W,0.01,G
SNe, H_0 , q_0	O	W,D
Radio/X-ray ID's/IR	O,U,I	W,D
Source Sizes	O,U	0.01
Large Scale Structure	O,I	W

Key:

Bandpasses

U* Far Ultraviolet (FUV), 1100 Å to 2500 Å

O Near UV/Optical (NUVO),
2500 Å to 10,000 Å

I Near IR, 10,000 to 28,000 Å

Camera Characteristics

W Wide field needed	C Coronagraphic mode needed
D Day or longer time resolution needed	N Narrow band filters needed
H Hour time resolution needed	T Tunable filters desirable
M Minute time resolution desirable	P Polarization measurements.
S Seconds or better time res. needed	R High dynamic range desirable
G Grism mode needed	0.01, 0.03 Resolution needed, arcsec

*Calls for a 'solar blind' detector for FUV, i.e., one that is insensitive to or blocked against longer wavelength ultraviolet and visible light. Without such blocking, the FUV signal would be swamped.

1.2 PERFORMANCE PRIORITIES FOR THE AC

We assessed what technical performance the AC should have to address the candidate science program to the fullest feasible extent. We assigned priorities to those capabilities based on judgements that sought to balance depth on some important problems with breadth corresponding to the wide range of astronomical interests. In doing so, we recognized that the HST community is widely ranging yet supportive of focused progress on fundamental problems. On this basis, we recommend these priorities:

- *Highest priority* is both (1) a wide-field near-ultraviolet/optical (NUVO) mode, 200" x 200" nominal field of view, half critical sampling at 5500 Å and (2) a high-resolution far ultraviolet (FUV) mode, preferably solar blind, 12.5" x 12.5" nominal field of view, critical sampling at 1800 Å.
- *Second priority* is a high-resolution NUVO mode, 50" x 50" nominal field of view, critical sampling at 5500 Å. This mode should include a coronagraphic option if the impact on efficiency is acceptable.

- *Third priority* is either (1) a wide-field FUV mode, 50" x 50" or greater nominal field of view, half critical sampling at 1800 Å, or (2) an independent NUVO coronagraphic mode if this mode is implemented separately from the high-resolution NUVO mode.

1.3 SPECIFIC AC SCIENTIFIC PROGRAMS

With confidence in the forgoing assertions about the general utility of an AC with the stated performance specifications, we have selected seven specific fields to illustrate the particular benefits of the AC:

- Planetary atmospheres
- Protoplanetary disks
- Stellar chromospheres
- Stellar populations
- Active galactic nuclei
- Galaxies and galaxy clusters
- The cosmological model

These scientific topics are seven important and unique areas of opportunity for observations with the AC on HST. In general, substantial progress on these critical programs will require both the imaging capabilities of the AC and complementary observations, such as spectroscopy from ground-based telescopes or measurements at longer or shorter wavelengths by other spacecraft.

1.3.1 PLANETARY ATMOSPHERES

For planetary observations, the customary advantages of space imaging are compounded with the synoptic perspective of earth orbit, which permits flexibly timed observations to address the variability of planetary phenomena and to accommodate the effects of planetary rotation. In this way, HST and the AC will provide a unique and valuable complement to space probes, which provide detailed measurements limited in time and place, and to ground-based studies, which are limited in time, wavelength, and spatial resolution.

The critically sampled NUVO mode of the AC will permit the global atmospheres of Mars, Jupiter, Saturn, Uranus, and Neptune to be studied with unprecedented acuity. Table 1.2 restates the specifications of this mode in planetary terms. Figure 1.1 illustrates the resolution in the case of Jupiter.

For wind studies on Jupiter and Saturn, these performance improve-

ments will permit (1) more reliable measurements, resulting from the ability to identify more individual tracers (improved statistics, more precise positions); (2) better access to convective features (smaller features at lower contrast levels); and (3) greater sensitivity to vertical structure (more precise coordinates especially near the limb). (See Beebe *et al.*, 1989.)

During a disturbed phase on Jupiter or Saturn, the rapid framing capability of the AC will allow the acquisition of image series through different filters, for example, centered in methane absorption bands and

Table 1.2. For planets at opposition, the number of critically-sampled resolution elements (ResEI) per diameter at 5500 Å, number of diameters (D) per 50" FOV, and the minimum time for features to shift by one spatial resolution element due to rotation.

Planet	ResEIs/D	D/FOV	Sec/ResEI
Mars	190	2.80	72
Jupiter	499	1.07	11
Saturn	208	2.56	28
Uranus	38	13.9	239
Neptune	23	23.6	610

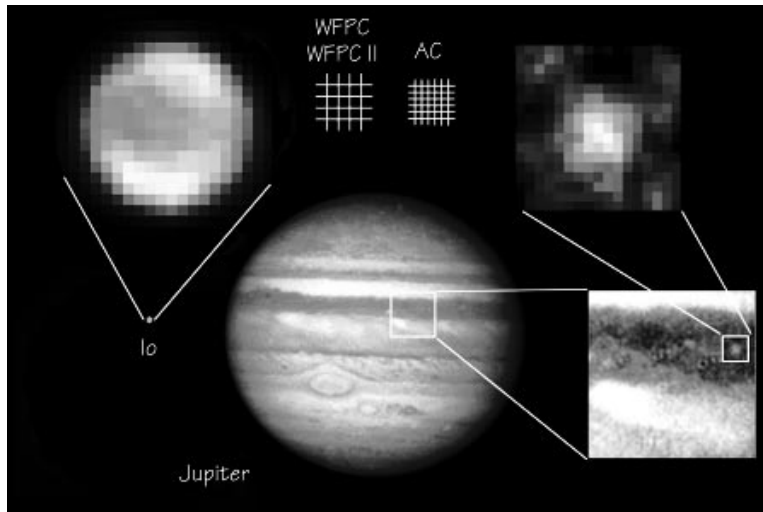


Figure 1.1. Jupiter and Io. A mosaic of images taken by the WFPC in the PC mode, on May 28, 1991. At the top of the figure, a diagram of the pixelations of the WFPC and WFPC II is compared to the expected capabilities of the AC.

nearby continuum. Such series will allow the observer to track individual features to the limb, which should overcome many of the problems associated with vertical structure in the presence of an inhomogeneous cloud deck.

Other AC giant planet studies will address the temporal variations of the ammonia clouds, the nature of convection in the upper troposphere, meridional stratospheric circulation, seasonal variability, and the dynamic coupling between the stratosphere and the troposphere (Drossart *et al.*, 1989). Studying the vertical distribution of aerosols in Jupiter's upper atmosphere requires carefully timed observations to extract center-to-limb brightness variations, which are best made both at long wavelengths (8900 Å methane band) and in the near ultraviolet (to achieve good vertical discrimination.)

To address the issue of Martian dust storm evolution as well as other questions about the meteorology and climatic change on Mars, the AC will be capable of rapidly capturing registered images through different filters. This will help us learn how dust storms on Mars arise and propagate to veil the entire planet in a few days. Models to explain them have been proposed, which are variously based on terrestrial hurricanes, planetary-scale circulation, and dust-sensitive tides. The AC may provide the first high resolution images of the critical early development stages of these dust storms.

At FUV wavelengths, the AC's UV sensitivity and absolute absence of red leak will be important improvements for studies of energetic planetary phenomena. These improvements will permit exposures with higher time and spatial resolution, uncontaminated by reflected sunlight from the bright visible disk.

The giant planets' aurora, one of the most energetic phenomenon in the solar system, is characterized by bright FUV emissions from atomic and molecular hydrogen. The estimated energy output of the Jovian aurora is 10^{14} Watts, more than a thousand times more powerful than for the Earth. This energy profoundly affects the atmosphere locally, and determines the thermal structure, aeronomy, and dynamics of the upper atmosphere. (See Clarke *et al.*, 1989.)

The FOC on HST took the first good images of the Jovian aurora (Dols *et al.*, 1992, Caldwell *et al.*, 1992), and a recent snapshot is shown in figure 1.2. The low sensitivity and high red leak of the FOC seriously limits this sort of study at the present time. The AC will take such FUV

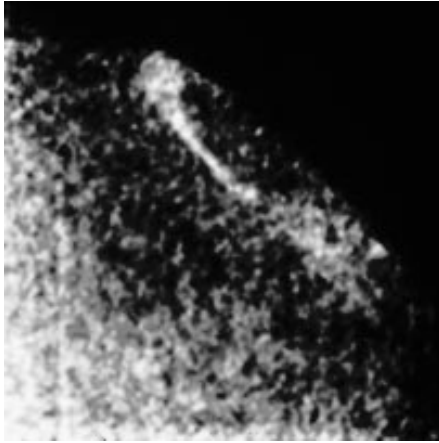


Figure 1.2. FUV image of north polar region of Jupiter taken by the FOC on Feb. 16, 1993. The image covers a 200 Å spectral range centered at 1530 Å, clearly showing the bright ring of the aurora circling the northern hemisphere. This 11" x 11" image has a spatial resolution of approximately 200 km on the planet. The AC will be able to take such FUV images with three times the signal to noise ratio at one-fifth the exposure time.

images with 3 times the signal to noise ratio in 1/5 the exposure time. Some of the outstanding issues that can be addressed with these capabilities are:

- *How does the aurora affect the giant planets' aeronomy?*
- *What are the auroral particles (electrons, protons or heavy ions) and where do they come from?*
- *What is the connection between the Io plasma torus and the Jovian aurora?*
- *Is there a connection between the aurora and the Lyman alpha bulge and dayglow emissions on Jupiter?*
- *What are the morphology and cause of the Lyman alpha bulge on Jupiter?*

1.3.2 PROTOPLANETARY DISKS

A detailed understanding of the genesis of planets is one of the major goals of astronomy. Current models of stellar formation and subsequent evolutionary stages imply the natural occurrence of flattened disks of material orbiting around young dwarf stars, which would provide the conditions and opportunity for the accumulation of planets. While the observational evidence for disks (Bertout, 1989; Backman and Paresce, 1992) is varied and convincing, the field is still in its infancy and now requires substantial improvements in technical capabilities to advance. The AC promises to make an important contribution by providing NUVO images with improved spatial sampling and reduced contami-

nation by direct starlight. This high dynamic range imaging will be provided by the high resolution NUVO capability, and will be greatly enhanced if the NUVO coronagraph is implemented.

Recent studies of large samples of pre-main sequence stars (PMS) in the Taurus-Auriga star-forming complex suggest that half the solar-type stars of age less than 3 million years (Myr) have optically thick disks. At ages greater than 10 Myr, fewer than 10% show signatures of disks. At these early ages, which are clearly critical for planet building, PMSs also display energetic stellar winds and bipolar outflows, sometimes with embedded Herbig-Haro objects. High resolution images can be expected to reveal important clues to the relative roles of accretion and stellar wind interactions in the evolution of the disks. The gaseous streams should be evident in emission lines of H, O, Si, and other species, which are excited by shock heating. Images in the continuum should reveal aspects of the disk morphology. An appropriate selection of narrow band filters and interference filters centered on the most important spectral features as well as the continuum will be a key capability of the AC. The exciting possibilities of the AC for studies of protoplanetary disks are illustrated by the recent WFPC images shown in figure 1.3.

PMS protoplanetary disks may be the precursor to optically thin

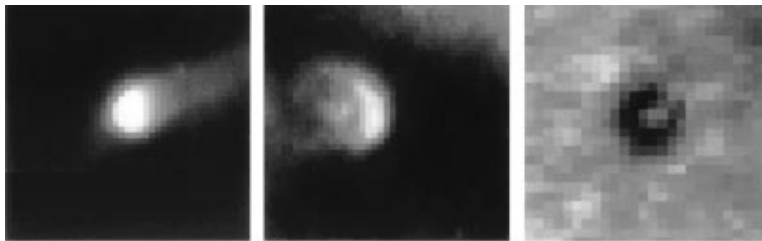


Figure 1.3. Massive gas disks around PMS stars in M42 as imaged by the WFPC (O'Dell, *et al.*, 1993). A dark accretion disk varies dramatically in its appearance according to whether it is found close to a hot star (left), which results in the outer disk material being ionized luminous, or at an intermediate (middle), or large distance (right), in which case the disk is seen in silhouette against the bright background of the nebula. Each pixel in this image represents 30 AU. NUVO images from the AC will provide improved spatial sampling and reduced contamination by direct starlight, and such images will be further enhanced if the NUVO coronagraph is implemented.

disks like the ones identified by infrared excess around the main sequence stars Vega, Fomalhaut, and Beta Pictoris, which are the only disks thus far confirmed by direct imaging (Backman and Paresce, 1992). With ground-based images, the Beta Pictoris disk can be measured in optical light in to about 40 astronomical units (AU) from the star (Golimowski, *et al.*, 1993). The zone between 10 and 50 AU is critical, for it is there that a possible planetary system either in formation or already well formed might reside (Backman, *et al.*, 1992). Preliminary calculations for the AC coronagraphic mode indicate that the scattered light from the central star will be low enough to permit reliable detection of the Beta Pictoris disk to within approximately 0.7" to 1" of the star corresponding to 12 to 17 AU. This performance will enable the detailed study of the colour, radial profile, and possible deviations from symmetry of the inner regions of the Beta Pictoris system, which could reflect the presence of planets or planetesimals interacting with the inner disk environment.

The AC should also be capable of imaging disks similar to Beta Pictoris around other IR excess main sequence stars—such as, Fomalhaut, Vega, and Epsilon Eridani—which has not been possible from ground-based observatories.

1.3.3 STELLAR CHROMOSPHERES

Key to the study of the fundamental physics of stars is the imaging of their surfaces and their interaction with the surrounding circumstellar environment. Surface imaging of stars revealing the size, structure, placement, and longevity of surface features elucidates the energetics of their atmospheres, magnetic dynamo and convection theory, and their evolution. Imaging of winds and extended atmospheres tests and defines the acceleration mechanisms of mass loss, the structure of extended atmospheres, and the origin of highly energetic processes such as jets from a surprising number of otherwise mundane systems such as symbiotic stars. Direct imaging of supergiants in the ultraviolet should also reveal nearby surrounding material indicative of previous mass loss episodes, and help to define the evolutionary state of these stars. Supergiants offer one of the most attractive classes of targets for surface imaging because of their brightness and size.

In the UV, plasmas are sampled at higher temperature than is possible in the optical region. The chromospheres and coronas of stars are

very extended making them larger targets than their photospheric dimensions imply. Recent ultraviolet and X-ray observations of cool stars in binary systems, for example, suggest that structures extend to several stellar radii, far in excess of the sizes observed on the Sun. Physical interaction between these structures and winds of each component would not be unexpected. The contrast of such structures is high when observed against the stellar disk because the photospheric continuum is weak or absent in the ultraviolet.

Excellent targets for the AC include circumstellar shells such as those around Gamma Ori. High resolution imaging coupled to ultraviolet spectroscopy would reveal the place of origin of the wind on the stellar surface, the morphology of the interaction of the wind with the surrounding medium, the energetics of the wind (fast, slow, hot, cool, etc.), and molecular formation in extended atmospheres. Two putative companions to the M supergiant Alpha Ori have been reported from speckle imaging reconstruction techniques: one at a distance of $0.06''$, the other at $0.51''$ (Karovska *et al.*, 1986). Direct imaging with a coronagraph blocking the M supergiant could easily reveal their presence.

Study of targets containing cool stars demands solar blind detectors to detect the FUV emissions without long wavelength contamination. Narrow band or interference filters will be required to isolate important strong stellar and circumstellar emission features such as C III, C IV, H₂, Mg II, etc. Search for extended emission around a central star will be carried out, in many cases, with the AC coronagraphic mode used to occult the bright central object.

The nova is another important class of object that now can be studied in greater detail. High resolution UV and optical imaging is required to study the initial development of instabilities in the shell before substantial mixing occurs with the interstellar medium. For Nova Cyg1992, for example, an optical interferometric angular diameter measurement in the first three weeks after outburst gave $0.033''$ at Balmer alpha, corresponding to material moving at about 500 km/sec, for a distance of 1 kpc. This implies that the $0.05''$ to $0.1''$ imaging capability of the AC will permit direct imaging of the early shell within the first year of a bright, nearby ejection. Ultraviolet narrow band imaging measures the different depths in the shell by using specific permitted and intercombination transitions and would permit measurement of the structure of the highest velocity material. The direct imaging of the shell

is a key to the interpretation of ultraviolet and optical spectra, providing explicit information about departures from spherical symmetry.

Comparison between UV and radio imaging is an especially important aim for this part of the study. For recurrent novae and symbiotics, imaging of jets and circumstellar material originating from the interaction between different components of binary systems, including red giants (as in T CrB and RS Oph), will be an important objective of the AC. Direct spatial resolution of the binary system and its accretion disk in nearby objects is a distinct possibility, since they are expected to have typical sizes of order $0.04''$ to $0.06''$ at 250 pc.

Figure 1.4 shows an FOC image of R Aqr, which is an example of

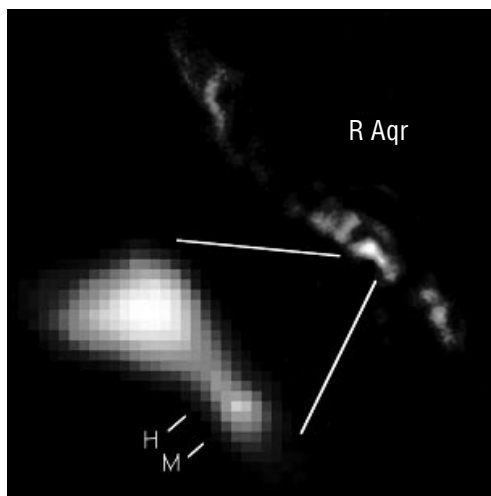


Figure 1.4. Direct image in the emission line C III at 1909 \AA of the jet emanating from R Aqr, a Mira variable that is a component of an active symbiotic binary system. Because the very bright Mira is cool and, therefore, its ultraviolet continuum weak, the FUV emissions from the jet can be isolated easily all the way down to several AU from the Mira. The insert in this figure shows that the putative hot, compact secondary (H) may be just resolved from the Mira (M). North is up, east is to the left. Each pixel in this image corresponds to $0.022''$ on the sky.

stellar imaging of this type (Paresce and Wilkins, 1992). The AC, with its superior resolution and sensitivity, should establish the nature of such interesting objects and allow studies of their evolution, since significant changes are observed to occur on time scales of less than a year. The for-

tuitous fact that R Aqr is close by (~ 200 pc) suggests the possibility of imaging in similar detail a large number of other more distant active stellar systems. This should clarify the physical processes that are common to all post MS systems especially those in which extended atmospheres interact strongly with the circumstellar environment.

1.3.4 STELLAR POPULATIONS

An understanding of the mechanisms of star formation, including the parameters that govern the process, is a critical problem in both stellar astrophysics and cosmology. A central issue is the determination of the initial mass function (IMF), which is the function describing the production rate versus stellar mass. The slope, shape and mass range of the IMF influence a wide range of processes and problems, including the early chemical history of the Galaxy, the evolution and predicted properties of high redshift galaxies, the population of stellar remnants in the Galaxy, and the missing mass problem.

Studies of resolved stellar populations in nearby galaxies have taught us a great deal about stellar evolution, including some of the parameters of star formation in particular systems: IMFs, chemical abundances, and age distributions. However, ground-based studies are severely confusion-limited and are impeded by restricted wavelength coverage. Only in the nearest systems, such as the Magellanic Clouds, have we been able to construct colour-magnitude diagrams down to 1 to 2 magnitudes below the main-sequence turnoff of the old populations. Many fundamental questions remain unanswered: How variable is the IMF over the Universe? How does it depend on the mean metallicity of the gas from which stars are formed? Is there bimodal star formation in normal and/or starburst galaxies?

The ultraviolet holds great potential for resolving these and other vexing problems concerning the age and metallicity distributions of stellar populations, including the problem of the ultraviolet upturn in old stellar populations. A single solar-blind, wide-field ultraviolet exposure with the AC will yield a complete sample of hot stars in a nearby galaxy. Such an exposure would include specimens exhibiting rapid and therefore rare phases of post-giant-branch evolution, the status of which objects is not yet fully understood.

In very crowded regions, ultraviolet images taken with the AC will isolate hot objects from the often dominant background of cool stars

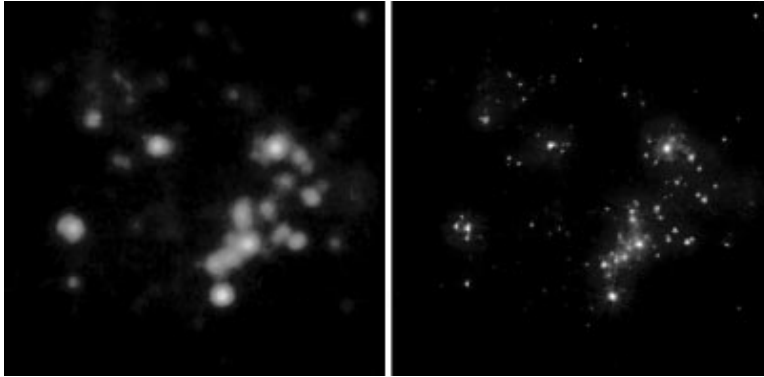


Figure 1.5. Wolf-Rayet stars in a large extragalactic starburst region, imaged in B-band, of the giant HII region NGC 604 in M33 (Drissen *et al.*, 1993). The image on the left was obtained using the 3.6 m Canada-France-Hawaii Telescope (seeing = 0.8"). The image on the right was taken with the HST WFPC in PC mode. All previously claimed 'superluminous' Wolf-Rayet stars (as inferred from ground-based images) are now found to be tight (diameter < 3 pc) aggregates containing one (or sometimes more) normal W-R star on HST images. Also, such HST images have allowed a preliminary determination of the initial mass function for stars in the 15 to 70 solar masses range. With the AC, such stellar population studies will be powerfully extended.

(e.g., in the crowded cores of globular clusters in the Magellanic Clouds and in the nuclei of nearby galaxies.) Ultraviolet images will also contain much information on interstellar dust and be aided by significantly darker sky background. In short, the AC with a wide separation of passbands from 1200 to 10,000 Å will separate factors of age, metallicity, and reddening much more effectively than previously possible.

Besides increased ultraviolet sensitivity, both high-resolution and wide-field capabilities of the AC will be important for stellar-population studies. To learn the star formation histories in Local Group galaxies, including the dwarf spheroidals around the Milky Way and M31, the AC will obtain multi-colour images (including the ultraviolet) that cover relatively wide areas (several arcminutes). Large-area coverage will also be an advantage for studying such objects as the blue S0 galaxy NGC 5102, in which the AC can establish—directly by star counts—the fraction of stars formed in a recent starburst, rather than having to rely upon interpreting the integrated spectrum via the cruder technique of population synthesis. In other studies, the highest possible resolution over smaller

areas (tens of arcseconds) will be necessary to diminish problems due to crowding. In this way, for example, high-resolution AC images of globular clusters in M₃₁ and other Local Group galaxies may help settle the question of the age spread in these cluster systems.

1.3.6 ACTIVE GALACTIC NUCLEI

Quasars and active galactic nuclei (AGN) are among the most important targets for the HST for at least two reasons. First, they include the most energetic objects known, and are believed to be powered by the gravitational energy released as matter is accreted onto a rotating black hole. They are therefore fundamental testbeds for studying high-energy astrophysical processes. Secondly, because of their enormous luminosities, they can be studied out to distances corresponding to the epoch when the Universe was only about 10% of its present age. Their attractiveness as probes of the early Universe is reinforced because they frequently emit luminous emission lines, from the ultraviolet to the infrared, which allows them to be recognized and their distances determined. Among basic physical questions of interest are:

- *Are AGN really powered by gravitational energy released by the accretion of material onto a black hole?*
- *What is the relative effect of orientation, environment (host galaxy and/or cluster), and history on determining the observational properties of AGN?*
- *How is nuclear activity related to the formation and evolution of galaxies?*

There are several spatially-extended components of active galaxies that can be studied in the optical and the FUV. These include light from stars of various ages, non-thermal radiation from synchrotron jets, scattered emissions, and emission lines from ionized gas. Observations over a range of wavelengths from the red to the far ultraviolet through broad and narrowband filters are essential to separate these various components. Imaging polarimetry will provide an important discriminant in the case of synchrotron emission and scattered light.

Several capabilities of the AC will allow important advances in the study of quasars and AGN. These include (1) high dynamic range and/or coronagraphic observations of the nuclear environments of AGN, (2) UV broad-band imaging of AGN hosts, (3) wide-field narrow-band imaging of nearby active galaxies, and (4) wide-field imaging of host envi-

rons (e.g. host clusters) surrounding AGN.

Detailed imaging of the central 1 kpc of AGN is crucial for studying the interaction of AGN with their host galaxies. By imaging in several bands in the optical and ultraviolet, morphological information on the different components of the stellar population can be obtained and related to that of the ionized gas close to the AGN. Do younger stars tend to be formed closer to the AGN? Is there morphological evidence indicating the presence of massive black holes (cusps, disks etc.)? Because such observations must be carried out near a dominant central point source, they place stringent requirements on dynamic range and spatial resolution. The AC's coronagraph, non-blooming charge coupled devices (CCD), high spatial resolution, and enhanced FUV imaging capabilities should revolutionize such studies. In particular, comparison of the optical imaging with radio interferometric imaging carried out with similar resolution will provide unique diagnostics of astrophysical processes involving jets and their interaction with the nuclear regions (Miley, 1981).

To study the evolution of the host galaxies, it is vital to image distant and nearby galaxies at similar rest wavelengths. FUV imaging of nearby active galaxies is therefore complementary to optical imaging of high-redshift AGN hosts. Radio galaxies are of particular interest for such studies. Distant radio galaxies tend to have their optical continuum



Figure 1.6. FUV image taken with the HST FOC of the long plasma jet emanating from the active nucleus of the giant elliptical galaxy M87 (Macchetto 1992). The jet's high polarization and similar appearance at radio wavelengths indicate non-thermal synchrotron emission. The improved FUV capabilities of the AC will allow the mapping of AGN features a factor of three times, and at higher resolution than previously achieved.

emission aligned with their radio axes (McCarthy, 1993; Miley, 1992). This effect is not fully understood but indicates that a fundamental relationship exists between the host galaxies and the associated radio sources. The jet may stimulate bursts of star formation as it propagates outward through the host galaxy. Alternatively, the aligned component may be scattered light from the quasi-stellar object that escapes anisotropically along the radio axis.

The optical images of the distant radio galaxies correspond to FUV rest frame images. Because the aligned component is very blue compared to the old stellar population, it is essential to carry out high-sensitivity imaging in the FUV to study this component in nearby, low redshift AGN. The enhanced FUV imaging permitted by the AC will allow the hosts of nearby AGN to be imaged at a similar rest frame. Because the vastly improved FUV sensitivity of the AC will enable features to be mapped that are more than a factor of 3 fainter than previously possible, the resultant images will allow more detailed studies to be made of the spatial distribution of stellar populations and relationships between star formation and nuclear activity. In addition, the images will provide an important nearby laboratory for studying the radio/optical alignment effect seen in distant galaxies.

One of the most informative approaches for studying AGN is to investigate how the ionized gas interacts with the other components, such as the stars and the synchrotron jets. Jets are seen to bend and decollimate as the result of such interactions, and the morphology, ionization, and kinematics of the gas knots are sometimes observed to be affected (Tadhunter *et al.*, 1988; Miley 1981). Emission line filaments are also often associated with active CD galaxies at the centers of rich X-ray clusters, and are believed to be cooling flows of gas falling into the cluster centers. Delineation of the morphologies of the gas and stars at the highest spatial resolution by the AC will provide fundamental information about the details of such interactions. Field sizes of several arc minutes are desirable for such studies. Because of the greatly decreased readout noise of the AC, detailed narrow-band imaging of the ionized gas should be able to map gas features by a factor of 2 to 3 times fainter than was previously possible with HST.

AGN, particularly radio-loud ones, can be used to pinpoint distant clusters. Because all galaxies in these clusters are at similar distances, they provide an important opportunity for studying differences between ac-

tive and passive galaxies and the evolution of nuclear activity. The simultaneous ability to provide high spatial resolution and wide field give the AC a clear advantage in carrying out such programs.

AGN research of the types mentioned here will be especially important drivers for the design of the coronagraph, the requirements of the polarizers, and the specification of the filters.

1.3.7 GALAXIES AND GALAXY CLUSTERS

One of contemporary astronomy's outstanding issues is understanding how galaxies evolve and when this evolution takes place. While we have been very successful in characterizing the present state of nearby galaxies, it has proved to be quite difficult to unambiguously define their evolutionary history and time scales. When galaxies first formed they would have been quite unlike those of today. Their sizes, shapes, and internal dynamics, as well as the age and abundance characteristics of their stellar and gaseous constituents, have changed since formation, reflecting the physical processes that have acted upon them over the last ~ 10 billion years (Gyr). Galaxies also bear the imprint of their environment; galaxies in denser regions systematically differ from those in less dense or 'field' regions. To establish what has happened we need to analyze the properties of distant galaxies and the characteristics of their environment as a function of 'lookback' time. The AC will be eminently suited for this task.

Studies of galaxy evolution also have ramifications for cosmology. Firstly, the observations carried out to characterize galaxy properties as a function of redshift typically also provide data on the spatial distribution of galaxies and on the evolution of their clustering properties with redshift. Such results provide a window on the scale and amplitude of the perturbations that seeded galaxies in the early universe. Secondly, a variety of key tests for deriving cosmological parameters (e.g., q_0 and Λ from the geometry of the Universe) require that we have galaxy samples for which the evolutionary changes with redshift have been quantified. Adequate samples do not yet exist with demonstrable, quantified constraints on evolutionary change.

Structures in galaxies (e.g., spiral arms, disk and bulge sizes), which provide a means of establishing the galaxy type and its dynamical state, typically have length scales less than about 1 kpc (and sometimes $\ll 1$ kpc). Since one arcsecond corresponds to roughly 7 kpc at a redshift $z = 0.5$ (for $q_0 = 0.5$ and $H_0 = 50$), imaging with resolutions of better

than $0.1''$ are essential. As illustrated in figure 1.7, WFC is providing a glimpse of what the required capability will accomplish: images of distant clusters and field galaxies from Dressler *et al.* (1993), Couch *et al.* (1993), and Griffiths *et al.* (1992) provide truly unique data from which we can determine morphological types, determine length scales, and characterize the environment of galaxies.

Qualitative analyses, which compare the data with the large body of morphologically defined types of nearby galaxies, are superficial but immediately productive. The reason is that the morphological appearance reflects certain physical and kinematical properties of the galaxy. The quantitative measurements extend the characterization of distant galax-

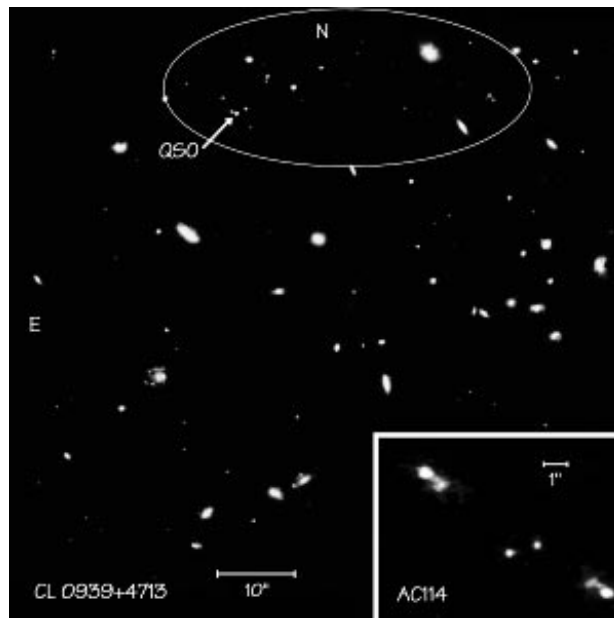


Figure 1.7. One WFC frame in R-band of the distant cluster CL 0939+4713 at $z=0.40$. Such a detailed view over a large field of a distant cluster is a unique capability of HST. The outlined area at the top, west of the QSO, shows a concentration of faint, compact galaxies that may be at the quasar redshift ($z=2.055$), and so could be the highest redshift cluster known.

The insert shows the gravitationally lensed galaxy in the cluster AC 114 at $z=0.3$ from a V-band WFC image. The remarkable symmetry in this faint image exemplifies the power of HST for studies of distant galaxies.

ies and directly relate them to their present-day counterparts.

HST AC images will allow the identification of the regions in which star formation (if any) is taking place, characterize the presence of spiral arms, bars, and other structural features, highlight the existence of interactions and mergers, identify whether AGN-like compact cores exist, and quantify the distribution of dust, gas and different stellar populations. Correlation of these properties with other data such as radio, X-ray and infrared observations will add further insight into the physics that governs galaxy evolution. HST with the AC will provide images of distant galaxies that will show more detail, *irrespective of their redshift*, than those that are typically obtained with ground-based telescopes of the Coma cluster.

The case for the wide-field, high resolution capability of the AC is enhanced by consideration of the synergism with multi-object spectroscopic observations on large ground-based telescopes. Spectroscopic observations alone are not sufficient for a comprehensive study of distant galaxies, but they do provide unique data that greatly strengthens the value of imaging data from HST. Such images allow us to relate the objects being studied directly to their present day counter-parts, while the spectroscopic results provide an estimate of their physical state. The importance of this complementary role cannot be stressed too highly. Distant galaxies are incredibly difficult objects to study and relate to well-understood, nearby systems.

By the late 1990s and into the next century, huge investments will have been made in the VLT at ESO, at Keck, and most likely in the Magellan, Columbus, and Gemini telescopes. These ground-based observatories will surely conduct major programs on distant galaxies using multi-object spectrographs. The FOVs covered by these spectroscopic systems and the HST AC are comparable, as are the magnitude limits for high resolution imaging of galaxies and multi-object spectroscopy (B ~ 24 to 25 magnitude). The synergism and scientific returns that will result from concurrent use of these 8 to 10 m class telescopes for multi-object spectroscopy with high-resolution, wide-field, NUVO images from the AC on HST cannot be overstated.

It is crucial that an imager such as the HST AC be available *when* the 8 to 10 m telescopes with their complementary multi-object spectrographs are being used to carry out deep galaxy surveys. Yet, because of lifetime considerations, the WFPC II may not be available, and thus no

relevant imaging capability is assured around the year 2000 without the AC. While WFPC II, if it is available, will clearly provide substantially improved imaging capability for distant galaxies over that of WFPC, and could provide the imaging data for the complementary spectroscopic programs, it is substantially inferior to the baseline AC. It will have a figure of merit that is $< 10\%$ of that of the AC. Its areal coverage will be $< 50\%$ of the baseline AC, its quantum efficiency $< 50\%$, and it will have poorer resolution by at least a factor of 2 (and probably worse since the potential gains with sub-pixel stepping will be less because of its poorer spatial scale: $0.1''$ in WFC mode versus $0.05''$ in the AC wide-field mode.) The gains to be made with the AC over WFPC II for distant galaxies are comparable to those of WFPC II over the aberrated images of WFPC.

It is clear that a wide-field, high-resolution AC with NUVO spectral coverage will provide a uniquely powerful capability for obtaining the large samples and detailed images of galaxies and clusters of galaxies that are needed for tackling one of the most outstanding and interesting problems of astronomy.

1.3.8 THE COSMOLOGICAL MODEL

The determination of the global value of the Hubble constant is the most important question in observational cosmology today. A close second is the determination of the mean mass density of the Universe, which means its total dark matter content. Despite over 60 years of work on the fundamental cosmological model and the substantial improvement in our ability to detect and observe galaxies and other cosmological tracers, there remains considerable uncertainty in the parameters of the currently accepted cosmological model, the Friedmann-Lemaitre model (also known as 'the hot Big Bang'). In particular, despite recent successes, including the improved calibration of the luminosities of type Ia Supernovae (SN) from HST observations of Cepheids in IC₄₁₈₂ (Sandage *et al.*, 1992), a new measurement of the distance to the LMC (Panagia *et al.*, 1991), and improved ultraviolet spectroscopy of type II SN (Kirshner *et al.*, 1993), the value of the Hubble Constant (H_0) is still uncertain by more than 50%. This is due to the lack of agreement between several powerful methods (Huchra, 1992; Jacoby *et al.*, 1992; Van den Bergh, 1992).

An accurate and unambiguous determination of H_0 is essential. The

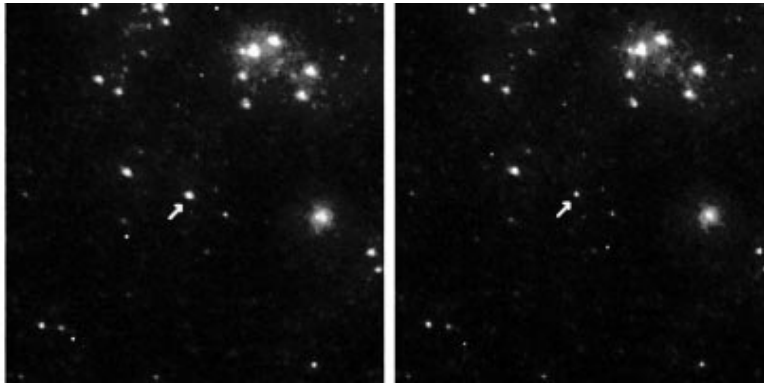


Figure 1.8. A Cepheid in the spiral galaxy IC 4182 shown in two images taken five days apart by the WFPC in the WF mode. The superior imaging capabilities of the AC will permit detection of Cepheids at further distances than currently possible, and yield more data to narrow down the value of the Hubble constant.

high values for H_0 preferred by most observers are in serious conflict with the combination of the age determinations for globular clusters (and the oldest stars in the Galaxy) and with the requirement of a flat Universe from inflationary cosmologies.

Without an accurate calibration for H_0 , we cannot even determine the fundamental parameters for extragalactic objects, much less test the cosmological model. An ambitious Key Project was planned to provide this calibration via the measurement of accurate Cepheid scale distances to a large number of galaxies using HST. This program was put on hold by spherical aberration. While WFPC II will provide some capability for proceeding with the program, the AC will provide substantially improved capability, including a significantly larger and better sampled FOV. Cepheids in Virgo cluster galaxies would have been marginally detectable with WFPC *unhindered* by spherical aberration. The AC, with 2x better sampling, the use of substepping, improved throughput, image deconvolution techniques, and a 3x larger FOV, will reach Cepheids beyond Virgo, and provide much improved accuracy in nearer galaxies.

In addition, several other key distance measuring techniques depend on the ability to survey galaxies with high spatial resolution, high sensitivity, and a wide FOV. These include measurement of brightest stars, the globular cluster luminosity function, the planetary nebulae luminos-

ity function, and even the fluctuation technique of Tonry and Schneider (1988) (because the amplitude of fluctuations in an unresolved stellar population depends inversely on the area of the spatial resolution elements). While these techniques generally depend on the Cepheid calibration, and are somewhat less accurate, with the AC they can be extended to much larger distances.

Currently, density determinations based on dynamical studies of individual galaxies or bound galaxy systems, such as clusters and groups, indicate that the mass density fails to meet the closure density by a factor of 3 to 5 (Trimble, 1988). Attempts to match large scale galaxy velocity fields and the cosmic microwave background dipole velocity with the local large scale galaxy density field, however, appear to be giving values of Ω near unity (Strauss *et al.*, 1992).

While ground-based surveys of galaxy velocities will substantially improve our knowledge of the density field over the next decade, the factor limiting the errors in flow determinations is our ability to measure accurate relative distances for large numbers of galaxies. Global techniques, such as Tully-Fisher or fundamental plane relations, provide distances to individual galaxies that are at best 20% accurate, and can only be used on a fraction of all galaxies. Furthermore, they suffer from selection biases, such as the Malmquist effect (Huchra, 1992). The substantial improvement in distance-measuring accuracy needed to map the velocity field to 10,000 km/s, which would provide an unambiguous determination of Ω on large scales, can be provided only by the AC on HST.

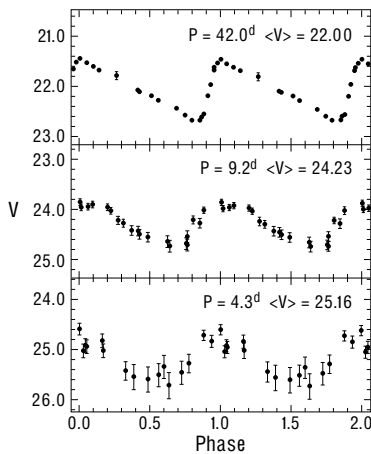


Figure 1.9. Representative light curves of the Cepheids found in IC 4182. A total of 27 Cepheids were found in this study, yielding an accurate distance, and the calibration of a 'standard candle', type Ia supernovae, one of which was observed in IC 4182 in 1937. Such supernovae are visible 1000 times farther than Cepheids, linking and extending the cosmological distance 'ladder'.

Lastly, the determination of q_0 independent of H_0 and Ω has eluded astronomers for decades. Chief among the problems associated with its determination is the effect discussed in the earlier section on galaxies and galaxy clusters—the evolution of the standard candles and standard measuring rods that might be used to measure the curvature of the universe. There is at least one route to q_0 which depends on a standard candle for which the evolution expected is small or nonexistent. The mechanism for the detonation of SN Ia is probably dependent only on the accretion processes onto white dwarfs in binary systems. SN Ia can be seen to great distances (redshifts) and can be discovered by large ground-based telescopes. However, to unambiguously classify the SN and to follow its light curve long enough to measure the expected time delay $(1+z)$ requires imaging with the superb resolution of HST. Only this capability can extract the SN from the light of the surrounding galaxy. Galaxy clusters are now being observed at redshifts near unity. A well observed sample of a few dozen supernovae Ia in these clusters will provide a galaxy and cluster evolution independent estimate of q_0 .

Similarly, the use of gravitational lenses to study cosmological parameters is just coming into its own (Schneider, Ehlers and Falco, 1992). The monitoring of existing lensed quasars requires excellent and stable imaging performance, and new lenses discovered by HST perform well will require HST follow-up observations, especially in the blue and ultraviolet, where quasar variations are largest. The improved sensitivity and spatial resolution of the AC are crucial not only for finding new lenses but also for systematic, high precision monitoring of these objects over long periods of time and for accurate determination of the mass models for the lens systems themselves (Bernstein *et al.*, 1993).

ACKNOWLEDGEMENTS

John Huchra was the lead author of chapter 1. Contributions were made by Reta Beebe, Howard Bond, Andrea Dupree, Heidi Hammel, John Hoessel, Garth Illingworth, Tod Lauer, Edwin Loh, Duccio Macchetto, Bruce Margon, John Mathis, Melissa McGrath, George Miley, David Monet, Jeremy Mould, Susan Neff, Francesco Paresce, John Raymond, Francois Schweizer, Steven Shore, Alan Stern, and Donald York.

REFERENCES

Backman, D.E., Gillett, F.C., and Witteborn, F.C. 1992. Infrared observations and thermal models of the Beta Pictoris disk. *Ap.J.*, **385**, 670-679.

- Backman, D.E., and Paresce, F. Main sequence stars with circumstellar solid material: the Vega phenomenon. In *Protostars and Planets III*, eds. E.H. Levy, J.I. Lunine, and M.S. Matthews (Tucson: Univ. of Arizona Press, 1992) 1253.
- Beebe, R.F., Orton, G.S., and West, R.A. 1989. Time-variable nature of the Jovian cloud properties and thermal structure: an observational perspective. In *Time Variable Phenomena in the Jovian System*, eds. M. Belton, R. West, and J. Rahe. NASA SP-494, 245-288.
- Bernstein, G., Tyson, J.A., and Kochanek, C. 1993. A large arc in the gravitational lens system 0957+561. *AJ*, **105**, 816-830.
- Bertout, C. 1989. T Tauri stars: wild as dust. *AnnRevA&Ap*, **27**, 351-395.
- Caldwell, J., Turgeon, B., and Hua, X.-M. 1992. Hubble Space Telescope imaging of the north polar aurora on Jupiter. *Science*, **257**, 1512-1515
- Clarke, J., Caldwell, J., Skinner, T., and Yelle, R. 1989. The aurora and airglow of Jupiter. In *Time Variable Phenomena in the Jovian System*, eds. M. Belton, R. West, and J. Rahe. NASA SP-494, 211-220.
- Couch, W.J., Ellis, R.S., Sharples, R.M., and Smail, I.R. 1993 In Texas/PASCOS Symposium, Berkeley, in press.
- Dols, V., Gérard, J.C., Paresce, F., Prangé, R., and Vidal-Madjar, A. 1992. Ultraviolet imaging of the Jovian aurora with the Hubble Space Telescope. *Geophys.Res.Lett.*, **19**, 1803-1806.
- Dressler, A., Oemler, A. Jr., Gunn, J., and Butcher, H. 1993. A cluster of nascent galaxies at $z=2$? *ApJL*, **404**, L45-L49.
- Drissen, L., Moffat, A.F.J., Shara, M.M. 1993. Hubble Space Telescope Planetary Camera view of giant H II regions: the Wolf-Rayet content of NGC 595 and NGC 604 in M33. *A.J.*, **105**, 1400-10.
- Drossart, P., Courtin, R., Atreya, S., and Tokunaga, A. 1989. Variations in the Jovian atmospheric composition and chemistry. In *Time Variable Phenomena in the Jovian System*, eds. M. Belton, R. West, and J. Rahe. NASA SP-494, 344-362.
- Ellis, R. *et al.* 1993. *MNRAS*, in press.
- Golimowski, D., Durrance, S., and Clampin, M. 1993. Coronagraphic imaging of the Beta Pictoris circumstellar disk: evidence of changing disk structure within 100 AU. *ApJL.*, in press.
- Griffiths, R., Ratnatunga, K., Doxsey, R.E., Ellis, R. *et al.* The Hubble Space Telescope medium deep survey: status report and first results. In *ST-ECF/STScI Workshop: Science with the Hubble Space Telescope, Proceedings*, eds. P. Benvenuti and E. Schreier (Munich: European Southern Observatory, 1992) 13-20.
- Huchra, J. 1992. The Hubble Constant. *Science*, **256**, 321-325.

- Jacoby, G., Branch, D., Ciardullo, R., Davies, R., *et al.* 1992. A critical review of selected techniques for measuring extragalactic distances. *PASP*, **104**, 599-662.
- Karovska, M., Nisenson, P., and Noyes, R. 1986. On the Alpha Orionis triple system. *ApJ*, **308**, 260-269.
- Macchetto, F.D. 1992. HST observations of Jets. In *ST-ECF/STScI Workshop: Science with the Hubble Space Telescope, Proceedings*, eds. P. Benvenuti and E. Schreier (Munich: European Southern Observatory, 1992) 73-81.
- McCarthy, P. 1993. High redshift radio galaxies. *AnnRevA&Ap*, in press.
- Miley, G.K. Distant galaxies with the Hubble Space Telescope. In *ST-ECF/STScI Workshop: Science with the Hubble Space Telescope, Proceedings*, eds. P. Benvenuti and E. Schreier (Munich: European Southern Observatory, 1992) 1-11.
- Miley, G.K. 1981. Jets and the Space Telescope—an introduction. Proc. ESO/ESA workshop “optical jets in galaxies.” *ESA Publications*, 22.
- O’Dell, C.R., Wen, Z., and Hu, X. 1993. Discovery of new objects in the Orion Nebula on HST images: shocks, compact sources, and proto-planetary disks. *ApJ*, in press.
- Panagia, N., Gilmozzi, R., Machetto, F., Adorf, H.-M., and Kirshner, R.P. 1991. Properties of the SN 1987A circumstellar ring and the distance to the Large Magellenic Cloud. *ApJL*, **380**, L23-L26.
- Paresce, F., and Wilkins, T. The core and jet of the symbiotic R Aquarii resolved in the UV with HST. In *ST-ECF/STScI Workshop: Science with the Hubble Space Telescope, Proceedings*, eds. P. Benvenuti and E. Schreier (Munich: European Southern Observatory, 1992) 329-331.
- Sandage, A., Saha, A., Tammann, G.A., Panagia, Nino, and Macchetto, D. 1992. The Cepheid distance to IC 4182: calibration of $M_V(\text{max})$ for SN Ia 1937C and the value of H_0 . *ApJL*, **401**, L7-L10.
- Schneider, P., Ehlers, J., and Falco, E., *Gravitational Lenses* (Berlin: Springer-Verlag, 1992).
- Strauss, M., Yahil, A., Davis, M., Huchra, J., Fisher, K. 1992. A redshift survey of *IRAS* galaxies v. the acceleration on the local group. *ApJ*, **397**, 395-419.
- Tadhunter, C.N., Fosbury, R.A.E., di Serego Alighieri, S., Bland, J., Danziger, I.J., Goss, W.M., McAdam, B. and Snijders, M.A.J. 1988. Very extended ionized gas in radio galaxies - IV PKS 2152-69. *MNRAS*, **235**, 403-423.
- Tonry, J. and Schneider, D. 1988. A new technique for measuring extragalactic distances. *AJ*, **96**, 807-815.
- Trimble, V. 1988. Existence and nature of dark matter in the Universe. In *AnnRevA&Ap*, **25**, 425-472.
- van den Bergh, S. 1992. The Hubble parameter. *PASP*, **104**, 861-883.

2. Instrument Status

What scientific instruments will be operational on Hubble Space Telescope (HST) in 1999? By mid-1999 there will have been two servicing missions and several instruments will have been replaced. We have reviewed the current instrumentation and future plans both to assure that an Advanced Camera (AC) is needed in 1999—rather than some other type of instrument—and to understand how the AC could best extend or complement previous imaging capabilities.

We confirm that an AC operating between 1000 and 10,000 Å is required to assure an adequate imaging capability on HST at the turn of the century.

From its conception, instrument replacement has been a fundamental aspect of the design and program plan of the Hubble Space Telescope (HST). Experience with ground-based observatories shows telescopes to have longer useful lives than instruments, which demand frequent repairs and can become outdated by advances in technology. HST's instrument replacement philosophy assures the continuity of basic observing capabilities while enabling the introduction of different or more powerful instruments as they become available. This approach promises to accrue full scientific benefit from the significant capital investment in the HST spacecraft and telescope optics, which are the stable foundation of an observatory with a useful life of at least 15 years.

2.1 HST SCIENTIFIC INSTRUMENTS IN 1993

Currently, in 1993, HST science instrumentation consists of two cameras, the Wide Field Planetary Camera (WFPC) and the Faint Object Camera (FOC), two spectrographs, the Goddard High Resolution Spectrograph (GHRS) and the Faint Object Spectrograph (FOS), and the High Speed Photometer (HSP). WFPC occupies a radial bay, while the other instruments occupy the four axial bays. (See chapter 6.) Fine

Guidance Sensors (FGS) in the other radial bays are used for telescope guiding and are also available for astrometric studies. (See chapter 4.)

Spherical aberration of the HST primary mirror has affected the performance of all current scientific instruments (SIs). The aberration has caused some but not all astronomical imaging programs to be delayed until after the 1993 servicing mission, the primary objective of which is to restore image quality. The science program of HST in the first three years has tended to emphasize spectroscopy over imaging. To mitigate the effects of spherical aberration, many scientists routinely use image deconvolution techniques.

Hardware problems have developed in all the instruments, which affects their scientific use. Several problems are serious. The WFPC has no useful response below 3000 Å because of internal contamination and is being used solely as an optical camera. The FOC *f*/48 imaging mode has developed an intermittent operating problem in its high voltage circuits. Early on in the mission, GHRS suffered an electronics failure, with the main loss being the echelle mode in the far ultraviolet (from 1250 to 1800 Å), this loss may be recovered during the 1993 servicing mission. FOS suffers from a loss in sensitivity in a narrow wavelength region centered near 1950 Å. Other, but less serious, instrumental problems have appeared, which can be corrected by additional calibrations or new observing procedures.

2.2 EXPECTED STATUS IN 1999

To understand the niche and role of the AC, it is necessary to construct an understanding of the status of HST and its instrumentation in 1999, based on developments underway and servicing missions planned.

On the first HST servicing mission in 1993, astronauts will replace the WFPC with an improved version (WFPC II), which will correct internally for the spherical aberration of the primary mirror. They will also install the Corrective Optics Space Telescope Axial Replacement (COSTAR), which will deploy correction optics in front of the FOS, GHRS and FOC. Also, the solar arrays, gyro packages and other failed or degraded units will be replaced.

The second servicing mission, planned for 1997, will install two second-generation instruments in HST—the Space Telescope Imaging Spectrograph (STIS) and the Near Infrared Camera (NICMOS)—and replace other spacecraft hardware units that may have failed.

Table 2.1. Expected camera instrument configurations on HST after 1997. Included are the high-priority modes of the AC to be installed in 1999.

<u>Instrument</u>	<u>FOV</u>	<u>Pixel size</u>	<u>Wavelength</u>	<u>Detector</u>
WFPC II (<i>f</i> /12.9)	75" x 75" (3)	0.1"	1200 to 10,000 Å	3 CCDs combined in 'L' shape. 800 x 800, 15 µm pixels
WFPC II (<i>f</i> /28.3)	34" x 34"	0.046"	1200 to 10,000 Å	CCD, 800 x 800, 15 µm pixels
FOC & CO-STAR (<i>f</i> /151)	7" x 7"	0.014"	1200 to 6000 Å	Photocathode, 3-stage intensifier, 25 µm pixels
STIS (UV camera)	25" x 25"	0.025"	1150 to 3100 Å	MAMA, 1024 x 1024, 25 µm pixels
STIS (optical camera)	50" x 50"	0.05"	3000 to 10,000 Å	CCD, 1024 x 1024, 21 µm pixels
NICMOS (camera 1)	11" x 11"	0.043"	1 to 2.5 µm	HgCdTe, 256 x 256, 40 µm pixels
NICMOS (camera 2)	19" x 19"	0.075"	1 to 2.5 µm	HgCdTe, same format
NICMOS (camera 3)	51" x 51"	0.2"	1 to 2.5 µm	HgCdTe, same format
AC WF-NUVO	200" x 200"	0.047"	2000 to 10,000 Å	CCD, half critically sampled
AC Hi Res-NUVO	50" x 50"	0.024"	2000 to 10,000 Å	CCD, critically sampled
AC Hi Res-FUV	12.5" x 12.5"	0.01"	1150 to 2000 Å	Solar-blind, critically sampled
AC WF-FUV	50" x 50"	0.02"	1150 to 2000 Å	Solar-blind, half critically sampled

Key: NUVO is near ultraviolet/optical, 2500 - 10,000 Å.
FUV is far ultraviolet, 1100 - 2500 Å.

Table 2.1 summarizes the image-taking instrumentation expected to be on-board HST at the time of the 1999 servicing mission.

The FOC will be 15 years old with about nine years of use on-orbit, six in conjunction with COSTAR. Not only will the FOC be beyond any reasonable expectation of operability, but the scientific utility of the instrument may have been exhausted. For example, the FOC filters have substantial red leaks and the sensitivity in the ultraviolet is relatively low.

The WFPC II will be seven years old, over five years on-orbit and beyond its design lifetime. If it is still operational in 1999, the utility of the WFPC II may have declined due to its coarse detector pixelation and low ultraviolet response.

The STIS cameras are designed for target acquisition and have a small number of filters. These filters are located close to the focal plane of the instrument, which is not optimal for scientific imaging as varia-

tions in the filter will produce field-dependent effects. STIS is by no means an adequate stand-in for either the FOC or the WFPC II—and certainly not for the AC.

NICMOS will serve the role of the WFPC or WFPC II for wavelengths 1 to 2.5 μm .

In 1999, STIS and NICMOS will be operating well within their design lifetimes, and are likely to be returning new scientific results in their domains of ultraviolet and optical spectroscopy and near infrared imaging.

2.3 DESCRIPTION OF THE INSTRUMENTS

2.3.1 WFPC

The WFPC is a dual, 2-dimensional spectrophotometer, operating from 3000 to 10,000 \AA with spatial resolutions of 0.1 arcsec per pixel (Wide Field Camera (WFC) $f/12.9$) and 0.043 arcsec per pixel (Planetary Camera (PC) $f/30$), using a 2 x 2 mosaic of four optically contiguous Charge-Coupled Device (CCD) detectors.¹ The WFC provides the largest field of view (FOV) available on the telescope, 160" x 160", however the WFC pixel spacing undersamples the image at all wavelengths with reference to the Nyquist criterion. (As explained in chapter 7, Nyquist sampling—which may be achieved either by pixelation or by substepping multiple exposures—is required to achieve full spatial resolution.) The pixelation of the PC is sufficiently fine to sample critically sample its 64" x 64" FOV at wavelengths longer than 9300 \AA .

The WFPC field of view is divided into four beams by a pyramid-shaped mirror near the HST focal plane. There are eight sets of relay optics and CCDs, four of which—either those of the PC or the WFC—are active simultaneously, depending on the rotation of the pyramid.

The CCDs, which were fabricated by Texas Instruments (TI), are thinned, backside-illuminated devices with an 800 x 800 array of 15 μm pixels. They are coated with the phosphor coronene, which is meant to convert ultraviolet (UV) light to visible light for more effective detection. Each CCD is hermetically packaged in a ceramic tube body, sealed with a MgF_2 field flattener, and cooled to an operating temperature of -87°C . The contaminants that cause the currently experienced loss of UV response have collected on the field flattener.

Optical filters are used to define the spectral bandpass imaged on the WFPC detectors. The filter assembly contains 12 filter wheels each with four filters, consisting of a variety of narrow, medium, wide, and long-

pass filters.

The optical efficiency of the combined WFPC and Optical Telescope Assembly (OTA) has a peak at 6000 Å, where the detective quantum efficiency (DQE) is about 0.18 electrons (e^-) per photon. Toward longer and shorter wavelengths, the DQE drops to roughly 0.05 near 4000 and 10,000 Å.¹

2.3.2 WFPC II

This camera² was designed to be a copy of the WFPC. Accordingly, the two SIs have similar overall design, electronics, and mechanisms. However, a number of changes have been introduced in the design in order to correct for spherical aberration, to incorporate improvements in CCDs and filters, to maintain sensitivity in the ultraviolet, and to save costs as a result of a reduction of scope in 1992.³

The WFPC II FOV is divided into four cameras by a stationary, four-faceted pyramid mirror. There are four sets of relay optics and CCDs in WFPC II rather than the eight in WFPC. Instead of four WF camera relay channels as in WFPC, there are now three, and instead of four PC channels, there is now one, with the PC being mounted into a position formerly occupied by the absent WFC channel. The plate scales are very similar between WFPC and WFPC II, although the overall FOVs are smaller. Each shutter opening exposes a mosaic of one PC and three WFC images.

The WFPC II CCDs, fabricated by Loral, are unthinned, front-illuminated, phosphor-coated devices with the same 800 x 800, 15 μm pixel format as the WFPC devices. Various deficiencies of the TI CCDs, such as quantum efficiency hysteresis (QEH) and the deferred charge problem are rectified in the new detectors. The Loral CCDs also exhibit considerably lower dark current and read noise, larger full well capacities, and are substantially more uniform in response. The expected DQE of the WFPC II and OTA is expected to exhibit a peak of 0.40 e^- per photon at 6500 Å. The phosphor coating should maintain a response of about $0.14 \pm 0.02 e^-$ per photon 1600 to 4500 Å, declining toward shorter wavelengths to about 0.09 e^- per photon at 1200 Å.⁴

An important aim of WFPC II is to provide a UV imaging capability. As stated earlier, the cold front windows of the CCD assemblies in the WFPC trap molecular condensates, which has destroyed the WFPC UV sensitivity. WFPC II follows a similar design. However, condensa-

tion rates are expected to be reduced by three to four orders of magnitude through a combination of improved manufacturing quality control, improved venting, the addition of adsorbent material, and a warmer (-70°C) operating temperature for the WFPC II CCDs. The engineering goal for WFPC II is a condensation rate that would degrade the system sensitivity at 1470 \AA by no more than 1 percent per 30 days. If that goal is achieved on-orbit, then a WFPC II UV imaging program will be feasible.

Like its predecessor, WFPC II has a set of 48 filters but with a number of modifications and improvements. The core set includes identical filters to WFPC covering standard UBVR bands and extends the sequence of wide filters into the UV. New filters include a set of Stromgren *uvby* filters. Improved narrow band spectral coverage has been achieved with the introduction of both quad and ramp filters. Additionally, there is a Wood's filter that is a broad band, short-pass ultraviolet filter. At 1600 \AA , the transmission of the Wood's filter is about 16 percent, cutting off sharply at 2100 \AA , with transmission below 10^{-6} at wavelengths longer than 3000 \AA . This filter will generally be used as a blocking filter in conjunction with far ultraviolet filters, which by themselves would transmit some long-wavelength light to the red-sensitive CCD detectors.

2.3.3 FOC PLUS COSTAR

The European Space Agency (ESA) FOC is one of the current four axial SIs and is affected badly by the spherical aberration of the HST primary. COSTAR should restore full use of the FOC in both the high spatial resolution, $f/96$ mode and in the long slit facility.

There are two optical chains, currently $f/96$ and $f/48$, which have independent photon-counting detectors and electronics. Each detector consists of a three-stage image intensifier, which is optically coupled by a relay lens to an electron-bombardment, silicon-target TV tube. The photocathodes are sensitive to radiation between 1150 and 6000 \AA . Optical filters are used to define the spectral bandpass imaged on the detectors. The filter assembly in the $f/96$ mode contains four filter wheels with 11 filters each, which consist of a variety of narrow, medium, wide, and long-pass filters, as well as an objective prism and neutral density filters.

In combination with COSTAR, the $f/96$ chain will become $f/151$ and will be the preferred channel for imaging with the FOC. The nor-

mal field size will become 7" x 7" with 0.014 arcsec pixels (25 μm). A field of view as large as 14" x 14" can be obtained but will require a larger pixel (25 x 50 μm .)

The design performance of the $f/151$ chain is 60 percent encircled energy within 0.1 arcsec radius at the field center and within 0.15 arcsec radius at the edge of the FOV of the full format imaging mode. The design also predicts a Strehl ratio of 0.55 at the field center.

Based on actual reflectivity measurements, the additional reflections introduced by COSTAR are expected to reduce the throughput of the FOC by factors 0.56 at 1200 \AA and 0.78 at 6300 \AA . The DQE of the combined FOC, COSTAR, and the OTA will peak near 4000 \AA at 0.08 counts per photon and drop to 0.01 at 1200 \AA .⁵

2.3.4 FOS AND GHRS

The FOS performs low-resolution (resolving power (R) ~ 250 and ~ 1300) spectroscopy of faint sources over the wavelength range 1150 to 8500 \AA . A variety of apertures of different sizes and shapes is available to optimize either the resolution or the throughput. The GHRS operates with resolving powers of R $\sim 20,000$ (1150 to 3300 \AA) and R $\sim 90,000$ (1700 to 3200 \AA) and has two entrance apertures. These first generation SIs do not have any useful imaging capabilities and are expected to be replaced on the 1997 servicing mission with NICMOS and STIS.

2.3.5 NICMOS

NICMOS contains three cameras of different spatial resolutions, which view different portions of the image plane and operate simultaneously. NICMOS serves the role of the WFPC or WFPC II for near infrared (IR) imaging, but with the added advantage of fully-sampled, diffraction-limited images in two of the three cameras.⁶

Each of the NICMOS cameras has a 256 x 256 HgCdTe detector array inside a cryogenic dewar. The NICMOS fore-optics focus the HST focal plane on field-splitting mirrors, which send the appropriate fields to individual optical systems that determine the magnification for each camera. Camera 2 has a coronagraphic mask and a filter wheel that contains polarizers for polarization studies. The nominal lifetime of the NICMOS cryogenic system is 5 years of operating time with a 10 percent margin.

The NICMOS cameras 1 and 2 are critically sampled according to

the Nyquist criterion for wavelengths 1.0 μm and 1.75 μm respectively, which ensures diffraction-limited imaging at any wavelength in NICMOS's full 1.0 to 2.5 μm imaging spectral range. NICMOS camera 3 provides wider field images for survey work and extended objects such as galaxies and planets.

All of the optical elements of the NICMOS cameras are outside of the dewar for easy alignment and access. Thermal control maintains the camera optics at a maximum stable temperature of 2° C to reduce thermal background emission and variation. The cryogenic dewar contains the second pupil apertures, filter wheels and detector arrays for all of the cameras. The cold filter wheels contain the polarization analyzers for camera 2.

The detectors are 256 x 256 pixel hybrid arrays of HgCdTe photo-diodes indium bump bonded to silicon direct-readout multiplexers. All of the arrays have long wavelength cutoffs of 2.5 μm .

2.3.6 STIS

The STIS provides 2-dimensional spectroscopy and limited imaging in a variety of different modes using 1024 x 1024 Multi-Anode Microchannel Array (MAMA) and CCD detectors. The STIS spectral range is 1150 to 10,000 Å. The purpose of STIS is to provide a wide range of spectral resolutions, very low ($R \sim 1000$) to very high ($R \sim 100,000$), to cover all the spectroscopic functions currently provided by either the FOS or the GHRS. Further, the two-dimensional aspect of the detectors allows spatially resolved spectra to be obtained for extended astronomical targets. STIS offers advances over the FOS through the lack of red leak in the ultraviolet and over GHRS through reduced scattered light due to improvements in the grating and the detector design and due to a coronagraphic mode.⁷

An earlier version of the STIS included a high-resolution, UV imaging mode with a FOV of 6" x 22" and pixel size of 0.011 arcsec, using the far-UV(FUV) MAMA detector (1150 to 1700 Å). This mode was deleted in an effort to simplify the instrument and contain the overall cost. Nevertheless, STIS does retain a target acquisition camera mode with basic imaging characteristics. This camera mode has a 25" x 25" FOV with 0.025 arcsec pixel size in the UV (MAMA detector), and a 50" x 50" FOV with 0.05 arcsec pixels in the optical (CCD detector).

The STIS cameras are designed for target acquisition and presently have a small number of filters, which include neutral density and narrow

band filters. The filters are located close to the focal plane of the instrument, which is not optimal for scientific imaging as variations in the filter produce field-dependent effects.

The far-UV MAMA detector has a CsI photocathode, and early results from a demonstration of this detector show an excellent DQE of 19 percent at 1220 Å. Thus, this camera mode should have better sensitivity in the far-UV than FOC with COSTAR.

Adding a small set of filters into existing vacancies in the filter wheel would be a modest change in STIS with significant scientific benefits. First, this enhancement would provide insurance for imaging capability with HST, particularly in the ultraviolet: STIS would retain capabilities similar to the FOC should that SI experience further problems. Second, adding a few filters would exploit the value of the solar-blind aspect of the STIS UV detector, which neither WFPC II nor FOC provide. Operating as a scientific imaging system, the STIS UV camera has better ultraviolet sensitivity and dynamic range than the FOC. STIS would also have better ultraviolet sensitivity than the WFPC II with its combination of a Wood's filter and CCD detector, and the STIS visible camera mode could serve as a backup to WFPC II in the optical. However, STIS is by no means an adequate replacement for either the FOC or the WFPC II—and certainly not for the AC. Providing STIS with sufficient filters for its existing camera modes is simply a prudent option.

2.4 RECOMMENDATIONS

Based on the foregoing considerations, we foresee an overriding need in 1999 for a modern camera with high spatial resolution and high throughput operating in the wavelength range 1150 to 10,000 Å. *It seems unthinkable that HST should have no imaging capability in that wavelength range, yet without the AC the risk of that eventuality is significant.* Also, it is a great shortcoming—one which the AC can correct—that HST will otherwise never have had a camera that takes full advantage of its imaging quality in the near ultraviolet/optical (NUVO) and of its far FUV efficiency. (The AC will be operating in parallel with NICMOS for three years on orbit and will provide a valuable, complementary ultraviolet-optical capability.)

We recommend that the STIS team expand the filter set on STIS to provide a limited, back-up, far-FUV imaging mode, provided this addition has no impact to STIS or NICMOS development nor any delay in

the 1997 flight schedule.

ACKNOWLEDGEMENT

J. Chris Blades was the lead author of chapter 2.

ENDNOTES

1. Mackenty, J.W. *et al.*, 1992, WFPC Handbook, ST ScI Publication.
2. Burrows, C.J. *et al.*, 1993, WFPC II Handbook, ST ScI Publication.
3. Trauger, J.T. *et al.*, 1992, in Science with the Hubble Space Telescope.
4. Janesick, J.R. and Collins, S.R. 1992, Jet Propulsion Laboratory (JPL) memo.
5. Paresce, F. 1992, FOC Handbook, ST ScI Publication.
6. A detailed description of the instrument is given by Thompson, R.I. (1991, University of Arizona report)
7. Woodgate, B.E. 1991, GSFC Report.

3. Optical Telescope Assembly

What are the qualities of the astronomical image at the entrance of the Advanced Camera (AC) as determined by the Optical Telescope Assembly (OTA)?

The OTA image exhibits spherical aberration, which can be corrected internally by the AC. OTA “breathing” may require an autonomous focussing mechanism in the AC. The OTA image wings are diffraction limited at wider angles and shorter wavelengths than expected before launch, which indicates that an apodizing coronagraphic mode in the AC would be particularly effective.

The Hubble Space Telescope (HST) Optical Telescope Assembly (OTA) was designed as a 2.4 m diameter Ritchey-Chretien telescope with an $f/2.3$ primary mirror, final focal ratio of $f/24$, and back focus of 1.5 m.

The OTA collects and delivers light to the scientific instruments positioned close to the optical axis below the central hole in the primary mirror. HST has four “axial” bays for instruments behind the focal plane and one “radial” bay for a scientific instrument in front of the focal plane. The radial bay instrument views the center of the telescope field of view through a pick-off mirror. The four quadrants of the focal plane, outside the zone vignetted by this mirror, are accessible to the four axial bay instruments. The Advanced Camera (AC) could occupy either the radial or axial bays. (Chapter 2 discusses the instruments presently installed in the telescope, and chapter 6 discusses the choice of bay type.)

3.1 OPTICAL PRESCRIPTION

The main difference between the OTA design and as-built optical prescription for the telescope, is that the primary mirror was built with a conic constant different from that specified. It should have been -1.0022985 , and instead was $-1.0139(5)$. As a result, the edge of the pri-

mary mirror is two microns too low, and rays hitting that region of the mirror focus 38 millimeters behind rays striking near the center of the mirror. (See figure 3.1.)

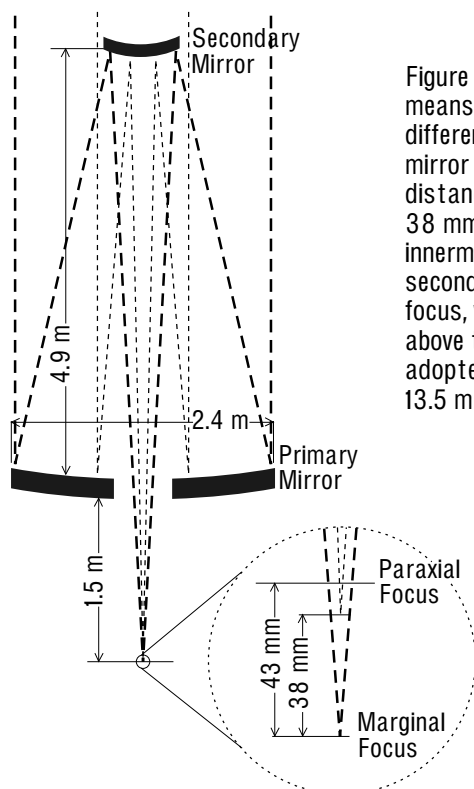


Figure 3.1. Spherical aberration means that light rays from different radii on the primary mirror come to focus at different distances. The marginal focus is 38 mm below the focus of the innermost rays, which graze the secondary mirror. The paraxial focus, which is obscured, is 43 mm above the marginal focus. The adopted focus (not shown) is 13.5 mm below the paraxial focus.

One effect of spherical aberration on the image is to cause a loss in limiting magnitude when background limited of approximately two magnitudes. It also causes a loss of contrast in crowded or complex fields. It is often impossible to tell which structures are caused by faint objects and which are parts of the image blur associated with brighter areas or stars in the field.

The first HST servicing mission, currently scheduled for December 1993, will effectively solve the spherical aberration problem with respect to the first generation spectrographs and cameras. The Corrective Optics Space Telescope Axial Replacement (COSTAR) will deploy pairs of corrective mirrors in front of the axial Faint Object Camera (FOC), Faint Object Spectrograph (FOS), and Goddard High Resolution Spectrograph (GHRS). (COSTAR will replace the High Speed Photometer

(HSP.) The Wide Field Planetary Camera (WFPC) will be replaced by WFPC II, a similar instrument with internal correction for spherical aberration.

The AC must correct spherical aberration—probably by including the reverse error on an internal mirror on which the HST primary mirror is imaged, as is the approach of COSTAR and WFPC II. The degree of correction possible in the absence of any active (deformable) optical elements depends on the accuracy with which the error has been characterized. The quoted error of $\pm .0005$ in the conic constant measurement corresponds to about two hundredths of a wave of spherical aberration. This figure may be improved after the servicing mission, however even at this level, the errors in the prescription are smaller than other errors in the original polishing and figuring of the mirrors unrelated to the spherical aberration.

In all corrective schemes, it is necessary to steer the beam onto the corrective element because the positions of the latches that hold both axial and radial bay instruments in position is not sufficiently well known. The uncertainty corresponds to a twisting of the instrument up to about 90 arcsec, which is one and a half percent of the beam diameter. If the beam is mispositioned on the pupil image by this amount, the result would be approximately a tenth of a wave of coma in the resulting image, which would be unacceptable. The only way this coma could be corrected without active optics within the instrument would be by repositioning the OTA secondary mirror, which is undesirable as it would affect all instruments and Fine Guidance Systems (FGS). Thus, the AC must have internal steerable elements to correctly position the re-imaged pupil relative to the optical train within the camera.

3.2 POINT SPREAD FUNCTION

A corrected HST image has three important attributes: (1) a high Strehl ratio (small low-order aberration, which results in concentrated image cores), (2) deep, diffraction-limited image wings (the mirror is very smooth, which results in small scattering), and (3) a wide field of view (FOV) where aspects (1) and (2) apply. The Strehl ratio is the ratio of the intensity of the actual image peak to that of the theoretically perfect image for that system. An optical system is commonly called 'diffraction-limited' if its Strehl ratio is greater than 0.8 (Maréchal's condition), which corresponds to wavefront errors totaling less than $\lambda/14$ if

they are the only source of image degradation.

According to the original optical error budget, HST on-orbit should produce a wavefront with a quality of approximately $1/20^{\text{th}}$ wave. The mirrors by themselves were to have produced an image with a quality of about $1/30^{\text{th}}$ wave, and alignment errors were to account for the rest. The known errors in the mirrors occur at all spatial frequencies but are dominated by low to mid spatial frequencies.

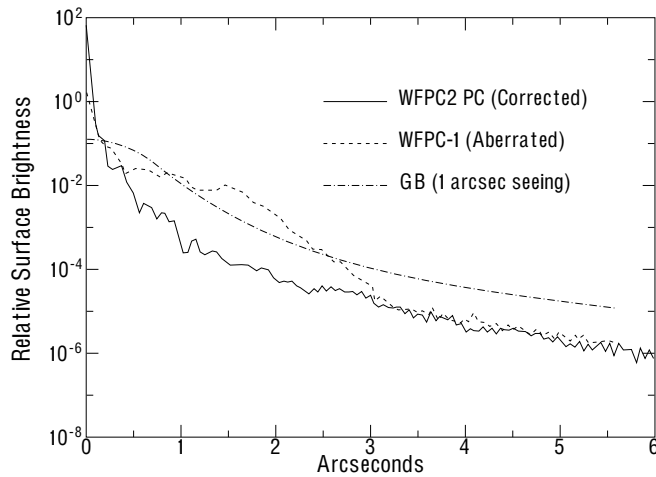


Figure 3.2. Comparison of optical-wavelength point spread functions.

Figure 3.2 illustrates how the AC will improve over the present HST and ground-based images. It shows the current point spread function (PSF), the PSF obtained in a well-corrected instrument, and the PSF one gets from the ground. It can be seen that the current PSF has a higher resolution core than the PSF from the ground. However, its wing is an order of magnitude higher than the ground based profile at 2 arcsec from the center, although the wing falls an order of magnitude lower than the ground base-seeing profile beyond about 3.5 arcsec. The corrected PSF is over an order of magnitude below the ground-based seeing disk at all radii beyond 0.5 arcsec.

There are three regions in the halo of a star image where different physical mechanisms cause the observed intensity. For small angles and long wavelengths the dominant effect is simply the diffraction pattern of the aperture. This gives a light level that increases linearly with the wavelength and goes down as the cube of the distance away from the center

of a point image. At larger distances from the star or at shorter wavelength, we expect to start seeing the effects of the mirror microroughness structure and polishing imperfections, which vary approximately as the inverse square of the wavelength and as the inverse of the square of the scattering angle. Both of these effects fall off rapidly enough that beyond about 2 arc minutes we expect the halo in the image to be dominated by dust scattering from the primary mirror. (Approximately 5% of the area is covered by particles 5 to 200 μm in size.) Note that this region of the PSF mode of the AC is outside that observed directly even by the proposed wide field, so only sources outside the field of view of the detector would lead to scattered light dominated by the dust on the optics.

There are some effects that are intrinsic to the present PSF and that might be avoided or mitigated in the AC by design. One is the simple fact that the surface of a CCD or photocathode is shiny and reflects light back up through the optical system. In the case of WFPC and in some modes of WFPC II, this light is reflected almost back upon itself: it hits the filter and then returns to the detector as an expanding cone of light. The result is a large halo around any point source with about the diameter of the chip and containing about five percent of the light.

This ghost is clearly unacceptable in high-contrast fields. For example, it could mask an extended source to be imaged in the neighborhood of a bright object. The problem can easily be controlled either by tilting the detector slightly within the constraints of maintaining focus, or better, by tilting the filters so that the reflected light is steered away from the detector into baffles.

For the current instruments, the ghost dominates the wings of the image profile. This was not expected before launch. This region of the image of the PSF was expected to be dominated, at short wavelengths at least, by scattering from the mirror surface. In fact we find that the scatter from the mirror surface is between five and ten times lower than predicted and may not even have been detected yet. The level of light that is observed in the wings outside of the ghost is consistent with a ghost of the ghost. If this ghost is eliminated, considerable benefit may be obtained by the installation of a apodizing coronagraph to suppress the PSF wings.

3.3 FOCUS STABILITY

The telescope is known to gradually shrink as a result of desorption of water from the graphite-epoxy metering truss that connects the pri-

mary and secondary mirrors. The rate of desorption is now of order one micron per month, and appears to be approximately constant. This is not a problem for the AC because it can be corrected by periodic adjustments of the secondary mirror.

Additionally, there appears to be a “breathing” variation in focus with an orbital time scale. Observations both in the FOC and the WFPC taken within the same orbit show changes in focus with a peak-to-peak amplitude corresponding to about five microns on the secondary mirror, or half a millimeter in focus. This difference is $1/20^{\text{th}}$ of a wave. If the effect is limited to such changes, and we are able to focus at the middle position, the largest focus error is $1/40^{\text{th}}$ of a wave, which would represent an acceptable amount relative to an overall diffraction-limited budget of $1/13^{\text{th}}$ wave. This assumes that slow focus changes can be taken out accurately by moving the secondary mirror.

A peak-to-peak amplitude 2x larger has been observed in one case in August 1992. Frequent changes of this magnitude would represent a serious threat to the error budget of any instrument that is not capable of autonomous focus. At this stage we see no compelling need for an autonomous focusing mechanism in the AC. If the effect develops a larger amplitude and frequency than we presently believe, it will become obvious in observations by the WFPC II.

The amplitude of the short term component of the breathing has been correlated with the forward light shield temperature (present value minus the average over the last orbit). This correlation could be used for an internal open loop control of AC focus if autonomous focussing should prove necessary.

WFPC II does not have any internal focus mechanism and relies on the telescope’s secondary mirror to provide all of its focusing capability. An advanced radial camera could operate on the same assumption. On the other hand, if the advanced camera is an axial instrument, it will require its own focus mechanism in order to assure confocality with WFPC II. The absolute focus of the telescope will be known to about 1 mm when uncertainties in the COSTAR and WFPC II focus are taken into account.

ACKNOWLEDGEMENT

Christopher Burrows was the lead author of chapter 3.

4. Pointing Performance

Unlike with ground-based telescopes, for which atmospheric ‘seeing’ is the major source of image degradation, pointing errors are the dominant source for diffraction-limited telescopes in space. We find that except for disturbances caused by the solar arrays, which are due to be fixed, and except for requiring relatively bright guide stars, the current pointing performance of Hubble Space Telescope (HST) is much better than expected before launch. The current line-of-sight stability is adequate for the Advanced Camera (AC), assuming the solar array problem is eliminated by replacing the arrays.

Adequate guide stars will be available for the AC only if it is installed in an axial bay.

Although not strictly required, an autonomous guiding feature in the AC, would offer insurance against the degradation of the existing pointing system, allow higher time resolution by synchronous scanning, and provide efficient and precise sub-stepping for recovering spatial resolution in a wide-field camera mode with coarsely-spaced pixels.

As shown in figure 4.1, pointing errors distribute the light from a point source over a larger area and reduce the intensity of the image peak, which is to say its Strehl ratio. (See chapter 3.2.) Figure 4.2, which shows the Strehl ratio of an optically perfect Hubble Space Telescope (HST) as a function of the magnitude of the pointing errors, indicates that pointing errors should not be more than a few milliarcseconds (mas) in order to maintain diffraction-limited quality in the ultraviolet (UV). The design specifications for the HST Pointing Control System (PCS) called for a pointing error over 24 hours of less than 7 mas rms (root mean square). At the time, this level of image stability was unprecedented in either ground telescopes or space platforms, and achieving it was one of the major challenges in the HST development project. In this chapter we examine how this requirement has been met—and exceeded—on orbit and why we find the pointing quality

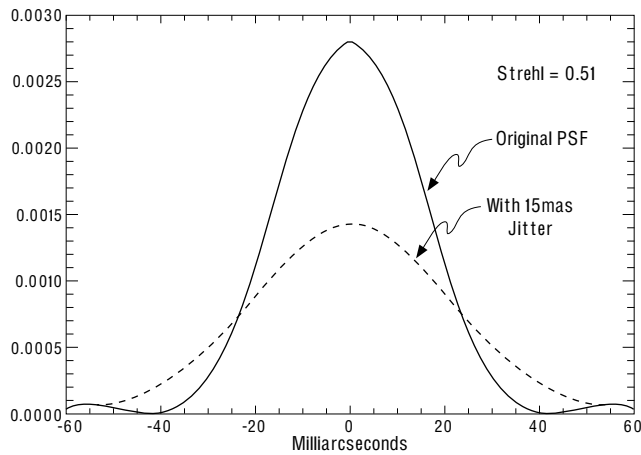


Figure 4.1. Effects of jitter on a perfect diffraction limited image.

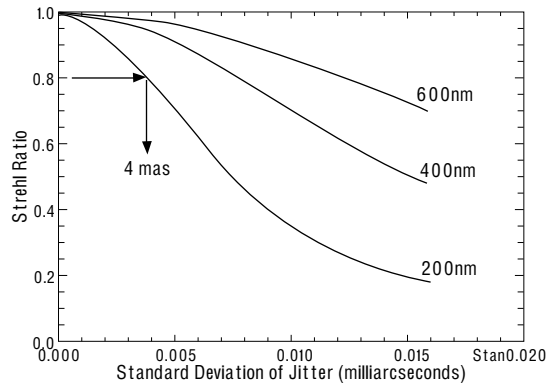


Figure 4.2. Effects of jitter on the Strehl ratio of a perfect image as a function of wave length.

appears adequate for the proposed Advanced Camera (AC). We also examine the availability of guide stars and the potential benefits of an autonomous guiding function for the AC to provide insurance and enhanced scientific capability.

4.1 DESCRIPTION OF THE GUIDING SYSTEM

The PCS uses several sources to determine the attitude of HST: sun sensors, magnetic sensors, external star trackers, field star trackers and gyroscopes. During observations, however, the stabilization of the line-of-sight (in other words, 'guiding') relies solely on the gyroscopes with a

periodic position update supplied by two (of three) Fine Guidance Sensors (FGS), which track two guide stars in the field of view (FOV) of the telescope. (Very short observations can be guided using the gyroscopes alone, i.e. without FGS update; short term drift is typically less than 1 mas per second.) Attitude corrections are applied to the entire spacecraft by varying the speed of spinning flywheels called 'reaction wheels'.

An FGS has two operational modes: a coarse mode ('coarse track') based on image scanning and a fine mode ('fine lock') based on interferometric fringe tracking. Fine lock is the only mode capable of providing the image stability required for high spatial resolution. Because fine lock has a limited capture range, coarse track is used as a preliminary step in the acquisition sequence to center the guide star before initiating the fine lock mode. Coarse track is more robust than fine lock, and can also be used as the primary guiding mode in observations not requiring the ultimate in pointing performance.

4.1.1 GUIDE STAR ACQUISITION PROCEDURE

A typical guide star acquisition begins with the telescope slewing to the desired field, usually followed by refining the attitude by a fixed-head star tracker update. At that point, the residual attitude error is about 12". Next, one FGS starts to search for the first guide star using a spiral pattern with its 5" x 5" instantaneous FOV. When it finds a star of the correct brightness, the FGS tracks it in coarse track mode, and the second FGS searches for its guide star. When the second star is found, its position relative to the first provides a check. If the guide star pair is not verified, the search resumes. If the pair is confirmed, the FGSs continue in coarse track if that is the desired mode, otherwise they transition to fine-lock mode.

The actual duration of the guide star acquisition process depends on the guiding mode and on how long the star search lasts. For first-time acquisitions, typical durations are 180 seconds for coarse track and 280 seconds for fine lock. Subsequent re-acquisitions in either coarse track or fine lock typically require 40 seconds less than first-time acquisitions.

If the guide star acquisition process fails because one of the two guide stars cannot be found or tracked, a second attempt is made on a different pair of guide stars. Up to three guide star pairs can be tried in one acquisition sequence. In practice, failed acquisitions are rare, and generally only one guide star pair is planned for each observation.

Once the guide stars have been acquired in fine lock, the error signals are fed at the rate of 40 Hertz (Hz) into the FGS control systems to keep the stars centered in the FGS instantaneous FOV. (In other words, the FGSs track their guide stars independently of what the spacecraft does.) The telescope attitude errors detected by the FGSs are fed at the rate of 1 Hz into the spacecraft PCS, which corrects the attitude using the reaction wheels.

4.1.2 COARSE TRACK MODE

In coarse track mode, the instantaneous FOV of the FGS is commanded to revolve (or 'nutate') around the guide star once per second at a radius of about 2.7 arcsec. (See figure 4.3.) The intensity of light collected in the aperture is measured by photomultiplier tubes (PMTs) at a rate of 40 Hz. The 40 values obtained in one second are summed separately for each of the four quadrants of the nutation to obtain the current star position error in two perpendicular directions.

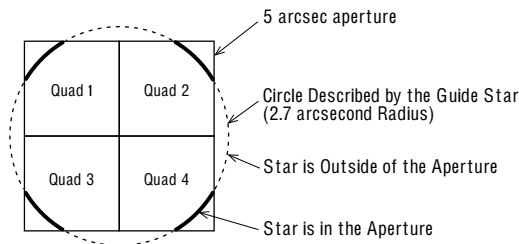


Figure 4.3. Coarse track mode

4.1.3 FINE LOCK

In fine lock mode, the collimated light of the guide star in the FGS instantaneous FOV is split into two beams in perpendicular directions, and wave front tilts in each beam are measured using Koester-prism interferometers. A Koester prism consists of two halves of an equilateral prism with a quarter-wave dielectric film sandwiched in between. Each beam incident on the Koester prism is further divided into two beams, and each is interfered with the mirror image of the other, now $\lambda/4$ out of phase. The intensity of each output beam—four in all—is measured by PMTs to determine the error signal.

As shown in figure 4.4, left, for a plane wave front parallel to the input face of a Koester prism, the two output beams have equal intensity

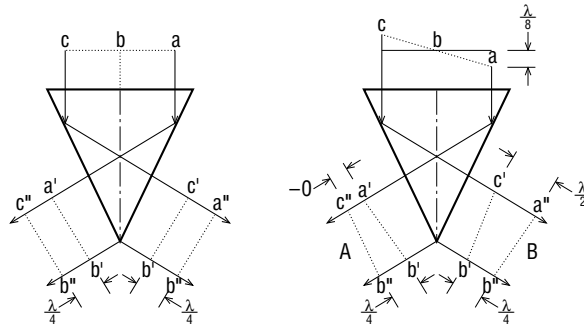


Figure 4.4. Fine lock mode. At left, wavefront parallel to input face, at right the wavefront is tilted.

(‘grey’ because of the $\lambda/4$ retardation). If the plane wave is tilted, however, figure 4.4 right the outer portion of one of the output beams will tend to experience constructive interference (higher overall intensity), while the outer portion of the other output beam will tend to experience destructive interference (lower overall intensity). The position error signal is formed by combining and normalizing the intensity in the two channels $((A-B)/(A+B))$. The error signal as a function of wave front tilt, commonly called an ‘S-curve’, is illustrated in figure 4.5. The theoretical modulation is ± 0.7 for a peak-to-peak separation of 43 mas , which is the diffraction limit of the telescope at wavelength 6000 \AA .

The problem with fine lock is its limited dynamic range—about 80 mas —which renders this mode very sensitive to vehicle jitter. Attitude disturbances can throw the pointing out into the S-curve aprons, where the error signal does not provide the information required to recover the central null, i.e., lock is lost.

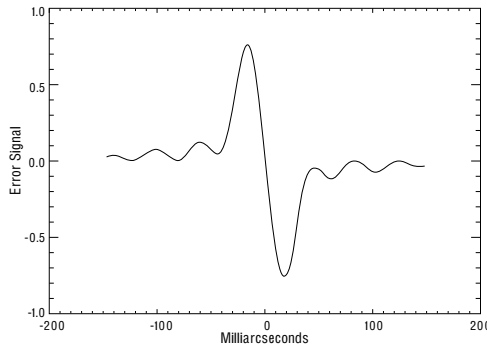


Figure 4.5. Theoretical error signal of the fine lock interferometer.

4.1.4 EFFECT OF SPHERICAL ABERRATION ON THE GUIDING SYSTEM

The spherical aberration in the HST primary mirror redistributes light from the core of a star image into a surrounding halo. In coarse track, this means starlight falls outside the FGS instantaneous FOV—60% or about half a magnitude loss averaged over a nutation cycle. Also, the coarse track modulation (sensor gain) is reduced as a result of image wings still present in the aperture when the image core is outside. Hence, even for brighter stars, the error signal is not as strong as expected before launch, and the control system response is degraded.

In the case of fine lock, the effect of the spherical aberration is more subtle. In principle, the ‘mirror image’ interferometric system is insensitive to any symmetric aberration such as focus or spherical aberration. However, this is not true in practice due to small misalignments between the incoming beam and the Koester prism axis. For a focus change, these misalignments cause a shift in the position of the S-curve null, and for spherical aberration, they reduce the error signal modulation. A typical real S-curve is shown in figure 4.6, which, when compared to the ideal one in figure 4.5, indicates how much internal misalignments affect both the amplitude and the shape of the S-curves. Due to the reduction of modulation and the secondary maxima, fine lock cannot be reliably obtained with the full pupil, and the use of the pupil stop ($2/3$ full aperture) to reduce the effect of spherical aberration is mandatory on an operational basis.

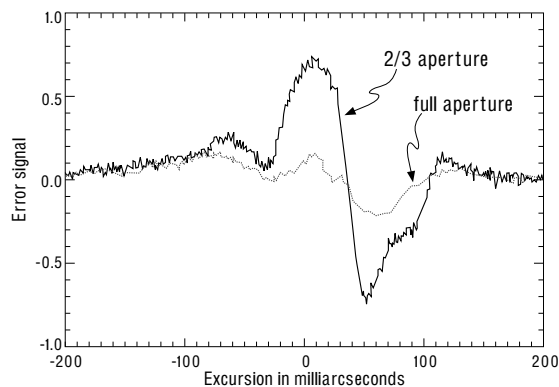


Figure 4.6. Typical actual S-curves. Dotted, an S-curve obtained with the full pupil. It is so degraded that reliable guiding is not possible. Solid, an S-curve obtained with a $2/3$ aperture stop shows how proper modulation can be restored.

The HST guiding system was designed to operate on stars down to 14.5 magnitude in order to ensure an 85% probability of finding guide stars near the galactic pole, where the distribution of field stars is most sparse. This specification was based both on the statistical occurrence of stars as a function of apparent magnitude and on the expected frequency of unidentified binary stars that would deny fine lock. Due to the necessity to mask the pupil used by the FGS to 2/3 of its full diameter for adequate modulation in fine lock, guidance is not reliable on guide stars fainter than 13th magnitude in fine lock. When an adequate pair of guide stars—stars bright enough and properly separated—cannot be found, the observation is either done in coarse track (if acceptable to the observer) or rescheduled later in the year for a different roll configuration. Currently, out of approximately 700 observations per year that are requested in fine-lock guidance, 30 are executed in coarse track, 100 are rescheduled later in the year, and only 2 or 3 are not executed at all for lack of guide stars, whether in fine lock or coarse track.

4.1.5 BINARY GUIDE STARS

Somewhat counteracting the loss in FGS sensitivity due to spherical aberration, it was a pleasant surprise to find after launch that binary guide stars have much less effect on fine-lock guidance than expected. Missed acquisitions or repetitive losses of lock attributable to the possibility that guide star is a binary represent only about 10% of all acquisitions. Binary stars can distort S-curve significantly, but the effects on guidance has now been shown to have been grossly overestimated. This development has obviated the need to uplink several pairs of guide stars to ensure acquisition success, and currently essentially all star acquisitions use only one pair.

4.2 CURRENT GUIDING PERFORMANCE

4.2.1 MEASURING HST'S LINE-OF-SIGHT JITTER

HST pointing jitter can be analyzed routinely using telemetry data from the gyroscopes and the FGSs, which in principle are the best indicators as they sense the line-of-sight directly through the entire optical train. The gyroscopes sense only the overall body inertial attitude, which may differ from the line-of-sight orientation due to the motion of optical components. In practice, however, the optical components contribute very little to jitter, and the gyroscopes and FGSs tend to agree on the

pointing performance, at least in fine lock.

The highest sampling rate of both the gyroscopes and FGSs is 40 Hz, which limits the analysis of the jitter spectrum to frequencies below 20 Hz. The line-of-sight jitter at higher frequencies has been measured with a dedicated test using the High Speed Photometer (HSP), which covered frequencies up to 500 Hz with a sensitivity of about 0.1 mas. The test demonstrated that except for very momentary and low amplitude excitations at high frequencies, most of the power is at less than 4 Hz. The mean frequency spectrum of jitter up to 10 Hz is shown in figure 4.7.

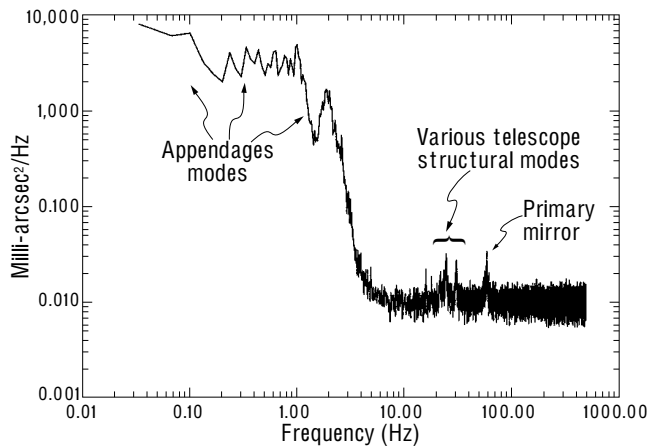


Figure 4.7. Power spectrum of jitter. Most of the power occurs at less than 4 Hz and is due to the oscillation of the appendages (solar arrays, high gain antenna and aperture door).

4.2.2 ORIGIN OF HST'S LINE-OF-SIGHT JITTER

Analysis of both routine and special-test data reveals that line-of-sight jitter has the following three components: rigid body motion of the telescope, thermally generated components, and components excited by moving mechanisms within the spacecraft.

- *Rigid body motion of the telescope.* The main body of the spacecraft moves in reaction to the oscillation of the flexible solar arrays, which in some orientations can contribute up to 6% of the total moment of inertia of the satellite. This is by far the largest contributor to jitter. Other appendages—the aperture door and the high gain antennas—also contribute. The spectrum of rigid body motion extends from 0.1 Hz to

about 4 Hz. Rigid body motion of the telescope is especially strong during earth terminator passages, when the solar arrays experience a thermal shock. During the rest of an orbit, it results from secondary solar array disturbances, thermal creaks in the outer body of the spacecraft, and PCS instability residuals.

- *Thermally-generated excitations propagating to the optical train.*

The tilting vibration of the primary mirror on its support is particularly noticeable at about 60 Hz, and various modes in the telescope support structure contribute in the range of 13 to 30 Hz. However, no vibration of the secondary mirror has been identified. These vibrations in the optical train and their supporting structure appear to be excited by shocks induced by the solar array disturbances and/or by thermal creaks within the outer body. The resulting image motion has small amplitude—a few mas—and damps quickly with no significant effect on observations.

- *Internally-generated excitations propagating to the optical train or the science instruments.* In principle, every active device on board the spacecraft—reaction wheels, antennas, and tape recorders, as well as the mechanisms in the guiding systems and the SIs—can potentially excite jitter. Because they are always operative and constantly changing speed, the influence of the reaction wheels is difficult to separate from the background excitation created by the solar arrays. Of the other devices, only the tape recorder produces a clearly identified effect. While in use at the normal record/playback tape speed, the tape recorder induces a continuous vibration of the line of sight at 65 and 300 Hz with an amplitude of 1 to 2 mas.

4.2.3 MAIN POINTING DISTURBANCES

The largest pointing disturbances occur twice per orbit at the passage between sunlight and darkness. The thermal shock to the solar arrays had been grossly underestimated before launch. The main body of the spacecraft reacts with an attitude displacement of several tenths of an arcsecond and oscillates for about two minutes at 0.1 and 0.6 Hz, which are the frequencies of the main torsional and in-plane bending modes of the solar arrays. The PCS has now been tuned to minimize this effect, but the limited bandwidth of the PCS does not permit full cancellation.

In addition to the terminator passages, the telescope pointing exhibits relatively frequent, abrupt, non-oscillatory glitches. These disturbances are 10 to 50 mas in amplitude, last 3 to 5 seconds and occur at ap-

parently random intervals with an average rate of one per 14 minutes. These attitude disturbances are all in pitch, and they affect the focal plane as a whole, which is to say that both the FGSs and the SI record them simultaneously. Detailed analysis of these disturbances and modeling of PCS seems to indicate that the cause is either micro-displacements in the shelf supporting the gyroscopes or temporary glitches in the gyroscopes themselves.

4.2.4 MEAN AMPLITUDE OF THE LINE-OF-SIGHT JITTER

A typical line-of-sight jitter time profile is shown in figure 4.8. This example is particularly meaningful because it was obtained during the HSP jitter test and represents the true motion of the image in the aperture of an SI (although, as mentioned earlier, the FGS data is closely correlated). Aside from the recognizable disturbances due to the terminator passage and a case of tracking disturbance, the tracking here is superb: the standard deviation over the final 800 seconds is 2.9 mas.

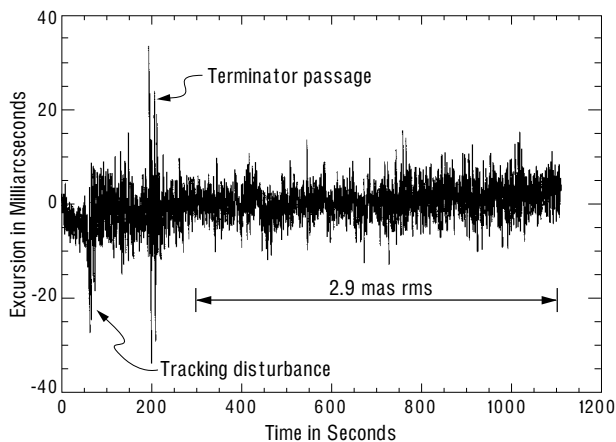


Figure 4.8. A typical jitter time profile while guiding in fine lock.

The effect of HST jitter is customarily measured by the full width at half maximum (FWHM) or standard deviation (rms) of its distribution averaged over a 1-minute interval. This time scale is both short enough to capture the effect of the large terminator-crossing disturbances, which typically last for several minutes, and long enough compared to the 10-second period of main solar array oscillation to provide meaningful

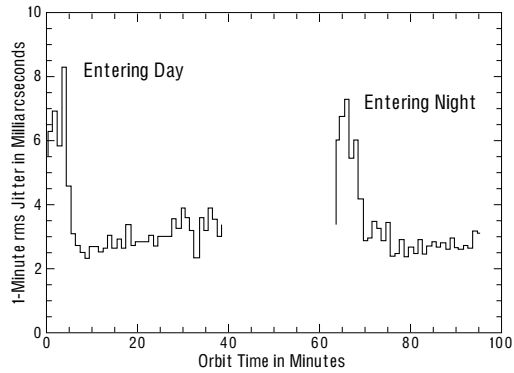


Figure 4.9. Average jitter during an orbit

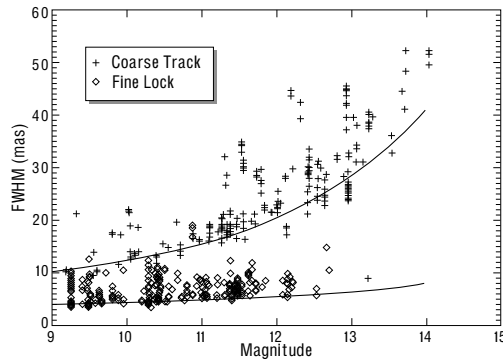


Figure 4.10. Typical jitter in coarse track and fine lock as a function of guide star magnitude. The two curves are representative of best guiding performance that can be obtained when vehicle disturbance level is minimal.

statistics. Figure 4.9 shows the time distribution of mean fine-lock jitter values over an orbit. (These measurements were obtained after the PCS software upgrade installed on day 92.107, which has been very successful in reducing the effect of the solar array terminator disturbances.) Averaged over an entire orbit, the line-of sight jitter while guiding in fine-lock is 3.5 mas—twice as good as the design specification!

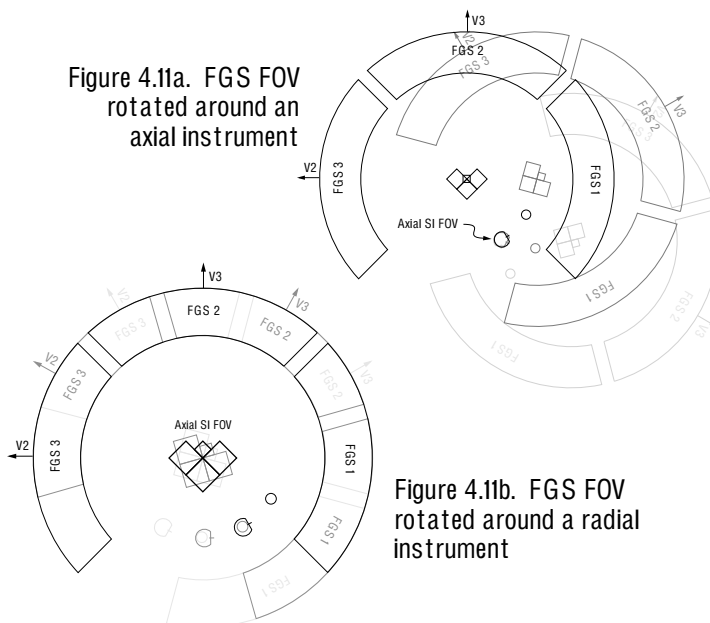
The influence of guide star magnitude on tracking performance is shown in figure 4.10. As expected, jitter varies as a function of an inverse square root law of the main guide star photon connects, but the effect is most sensitive for coarse track.

4.2.5 LOSS OF LOCK

As indicated earlier, each day/night terminator passage is accompanied by a large vehicle disturbance. In the coarse track mode, the guiding system never loses a guide star by virtue of its large dynamic range (ζ arc-sec). On the other hand, the fine-lock system, which has a very narrow dynamic range (± 40 mas), is susceptible to large, rapid excursions of a guide star. A recent modification of PCS, in which the FGSs are essentially turned off during the largest line-of-sight excursions, has been very successful in eliminating loss of lock. Losses of lock which used to occur about half of the time during terminator passages now practically never happen. This has not suppressed the higher jitter prevailing during two or three minutes at each terminator passage, but the impact on observation is really negligible.

4.2.6 GUIDE STAR AVAILABILITY

Due to an interesting geometrical factor, the number of available guide stars is different for axial and radial SIs. The nominal roll orientation of the spacecraft places the sun in the V_1 - V_3 plane. For a given astronomical target, this roll orientation fixes the FGS FOVs on the plane of the sky. As the spacecraft rolls around a target—either by a commanded off-nominal roll or as the sun-target orientation changes



through the seasons—the FGS FOVs move relative to the target, which makes more guide stars available to support the observation. As illustrated in figures 4.11a and 4.11b, this effect is much more dramatic for a roll around an off-axis than an on-axis target, which means that more guide stars are available to an axial than to a radial SI. This is shown quantitatively in figures 4.12 and 4.13, which compares guide star probabilities for typical axial and radial SIs. As summarized in table 4.1, the typical axial SI has a significantly higher probability of finding a pair of 13th magnitude guide stars at mid- and high galactic latitudes.

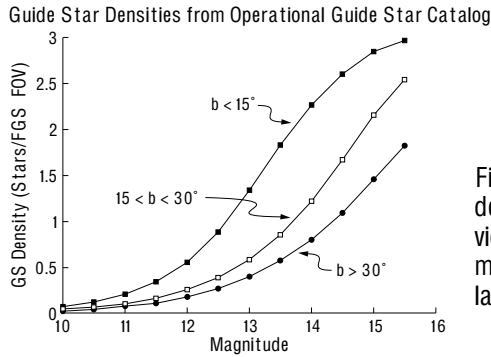


Figure 4.12. Guide star densities per Fgs field of view as a function of magnitude and galactic latitude (b)

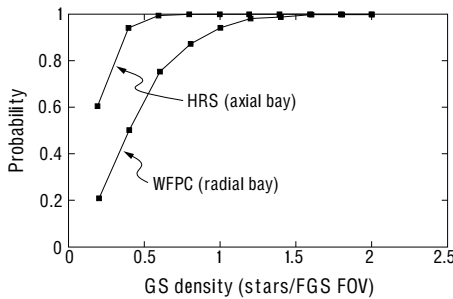


Figure 4.13. Probability of finding an adequate pair of guide stars as a function of magnitude for an axial or a radial bay instrument

Table 4.1. Effect of limiting guide stars to 13th magnitude

	Galactic Latitude	
	$15^\circ < b < 30^\circ$	$b > 30^\circ$
Guide star density (13 th magnitude guide stars/FGS FOV)	0.55	0.4
Axial SI success probability (1 pair of guide stars)	99%	95%
Radial SI success probability (1 pair of guide stars)	75%	55%

4.3 IMPROVEMENTS EXPECTED FROM THE REPAIR MISSION

On the first servicing mission, new solar arrays will be installed that have been designed to be insensitive to the thermal shock of the day/night transition. This should eliminate all terminator-induced high jitter as well as losses of lock. If this is achieved, we expect the amplitude of line-of-sight jitter to be about 3 to 4 mas routinely—twice as good as the original specification and adequate not to degrade the point-spread function significantly even at $\lambda = 1100 \text{ \AA}$, the shortest wavelength we envision for the AC. The small and random tracking discontinuities will likely still remain, however, but their scientific impact on AC observations will be minimal.

4.4 FOCAL PLANE STABILITY AND CONTRIBUTION BY THE SCIENCE INSTRUMENTS

In addition to pointing errors which strictly refer to the motion of optical images delivered at the focal plane, focal plane or FGS instabilities can add their own contribution. Indeed, there is evidence that for long exposures, the image can move by somewhat more than what the simple line-of-sight jitter would indicate. Whether this is due to effects internal to the instruments, drift in the FGS, changes in the telescope optical train, or instabilities in the focal plane is not yet clear.

For example, there is the suspicion that the FGS may experience slight thermally-induced drift at the milliarcsecond level. There is also one HSP long observation which exhibited a flux variation which could be attributed to a motion of the HSP relative to the FGS, although there are alternate explanations (focus changes). Concerning the cameras, image size measurements do not quite match the values predicted from the simple line-of-sight jitter. Overall these effects appear to be on the order of 10 mas at the most.

In the current design of HST, there is no provision for the pointing system to guide these effects out since the FGS can only correct for errors that affect the focal plane image as a whole. Only a guiding system internal to an instrument would be able to correct for these focal plane drifts or SI internal instabilities.

4.5 ADVANTAGES OF AN AUTONOMOUS GUIDING FUNCTION IN THE AC

Even if line-of-sight jitter is reduced to a fully acceptable level with

the repair mission, an autonomous guiding function in the AC would have multiple advantages:

1. A local guiding would guarantee a *short term* line-of-sight stability to a level better than what can be provided by the FGS/PCS, because any PCS residual errors would thus be corrected. Although these residual errors may be brought down to within specifications with the installation of the modified solar arrays it is likely that the spacecraft will still be sensitive to some extent to thermal creak affecting either still the solar arrays or the spacecraft body. Sudden disturbances of the type described above in section 4.2.3 would also be eliminated.

2. Local guiding would guarantee the *long term* line-of-sight stability which is affected by the differential motion of the FGS and the SI. This thermally-driven effect is not fully characterized yet but thought to be on the order of several milli-arcseconds over a time scale of several orbits (due to the thermal inertia of the spacecraft). The current FGS based guiding system is, by design, incapable of detecting these errors and only a system internal to an SI can compensate for them.

3. An autonomous guiding system would serve as a backup, in case of failure or degradation of the primary FGS-based system. Our current system is resilient to failures, and there exist several levels of backup modes which would maintain some sort of guiding capability. Examples of these backup modes are: only two FGS, FGS with only one axis control, or one FGS combined with FHST or gyros. Most backup cases would lead to a guiding degradation. An autonomous guiding capability would act as an insurance, and guarantee that even in the case of backup operation, the line-of-sight stability of the AC is kept at the highest level.

4. It may be possible to feed the error signal back to the PCS in order to control the entire spacecraft. This could be advantageous in order to maintain a high guiding quality for *all the SIs* in case of degradation of the primary (FGS-based) guiding mode. The AC being then only used a guider.

5. An autonomous system which allows precise and rapid control of the line-of-sight beyond the capability of the normal guiding system, has many potential astronomical applications. One example is increasing the time resolution beyond the nominal frame rate of the camera by synchronous scanning. In this technique, the line of sight is rotated on the imaging camera and variable sources can be detected by measuring the change in intensity over the circular trace.

6. Increasing the efficiency of the observatory by allowing the systematic use of the FGS coarse track mode instead of fine lock, the higher associated jitter being compensated by the autonomous system. Coarse track leads to essentially full sky coverage, and saves about 100 seconds in acquisition time.

7. Increasing the efficiency of the observatory by allowing faster step and dwell than what the current pointing system is capable of (2 arcsecond per second).

4.6 CONCLUSION

In conclusion, we consider that once the terminator passage disturbances will have been eliminated by the replacement of the solar arrays, pointing should be well within the specification for the AC. Still, although not strictly required, a guiding system integral to the AC would offer an insurance against pointing degradation, improve overall observing efficiency and allow sophisticated observational techniques that the current slow and somewhat imprecise pointing system cannot offer. Concerning guide star availability, full sky coverage can only be guaranteed if the AC is in an axial bay.

ACKNOWLEDGEMENT

Pierre Bely was the lead author of chapter 4.

5. Adaptive Optics & HST 'Uniqueness Space'

Adaptive optics at ground-based observatories promises enhanced resolution and raises the question: What are the unique imaging aspects of the Hubble Space Telescope (HST)? Upon review, we find adaptive optics is an inherently limited astronomical solution compared with space imaging. An advanced camera can exploit the advantages of space to provide scientific benefits unchallenged by ground-based telescopes.

Adaptive optics offers greatest gains in the infrared, with progressively smaller gains in the green and blue. Adaptive optics provides only small fields of view, which rules out crisp images for accurate photometry of extended objects like planets and clusters of stars and galaxies. Adaptive optics improves only the central part of the image, and the remaining halo of scattered light degrades the obtainable contrast and dynamic range for detecting faint companions and studying subtle features on extended objects.

HST performs better than adaptive optics in providing a stable, superb image over a wide field of view from far ultraviolet to near infrared wavelengths.

Adaptive optics (AO) is a technique that aims to compensate for the aberrations caused by the Earth's atmosphere in real-time. The measurement and compensation cycle is 1 to 10 milliseconds (msec) of time. This approach is different from active optics, which aims to provide corrections to *mirror* figure errors on a slower time scale.

AO is intriguing because it has the potential to completely remove atmospheric distortions and to achieve diffraction-limited imagery. The outlook for AO is promising: it is making technical strides, its funding levels are increasing, and it should attain the status of facility instrumentation at many ground-based observatories by the end of the 1990s. In view of these developments, this study of the Advanced Camera (AC) for the Hubble Space Telescope (HST) must confront the following is-

sue: How will ground-based AO restrict the ‘uniqueness space’ of the AC? It would be foolish to implement a capability in space that could be achieved at lower cost on the ground.

To approach the uniqueness issue, we have synthesized several sources of information. We have reviewed the responses from the community to our broadcast letter soliciting ideas about the AC, current adaptive optics research results, and recently declassified military technology. Our initial assessment was that the role of AO would depend greatly on the funding and focus of the major research groups, which prompted us to conduct a supplemental survey of the AO community before forming our final prognosis.

In this chapter, we begin by stating the imaging attributes of HST that define its potential uniqueness space. We then explain the atmospheric distortion problem, present the generic AO approach, identify its physical limitations, and discuss the cost drivers for its implementation. We summarize the major adaptive optics projects and their current goals. On the basis of all this information, we give our prognosis for AO and delineate the performance aspects of the AC that will remain unquestionably unique and valuable for science.

5.1 THE ADVANTAGES OF SPACE-BASED IMAGING

Two of the main advantages of a space-based telescope are wavelength coverage and the potential for a point-spread function (PSF) characterized by three attributes: (1) high resolution, i.e. a diffraction-limited core, (2) low scattering, i.e. diffraction-limited wings, and (3) wide field of view (FOV) over which the PSF does not vary.

Useful ancillary descriptors of the PSF are (a) the encircled energy function (EEF), which is the definite integral of the PSF as a function of the radial distance from the PSF centroid, (b) the full-width at half-maximum (FWHM), which for a diffraction-limited system is approximately λ/D where λ is the wavelength of light and D is the diameter of the telescope aperture, and (c) the Strehl ratio, which is the ratio of the intensity of the actual image peak to that of the theoretically perfect image for that system. We will use the Strehl ratio rather than the FWHM to measure spatial resolution because it is more sensitive for nearly diffraction-limited systems. The EEF is a useful shorthand for comparing PSF wings.

The relative importance of resolution (Strehl ratio) and PSF wings (EEF) depends on the imaging scenario. For instance, measurement of

the separation of two equal-magnitude stars is primarily limited by resolution because the stochastic uncertainty is dominated by the statistical noise in the cores of the two source images. On the other hand, for faint companions or low-contrast features in extended objects, *both* high Strehl ratio and tight EEF may be necessary conditions for detection.

For surveys and sensitive photometry of scattered sources, the variation of PSF over the FOV is an important consideration, which for various observations combine with high Strehl ratio or tight EEF or both as performance requirements. Ideally, the PSF is the same over the entire image plane, that is, the PSF is 'isoplanatic'.

Interestingly, the following sections will show that comparing the expected 1999 capabilities of ground-based AO with HST is similar to comparing performance before and after correcting HST spherical aberration. Ground-based AO will reach very high resolution, but its PSF will be spatially varying and will exhibit a significant halo, which is analogous to the present HST performance. Thus, our assessment of the uniqueness of HST in 1999—and the scientific benefits of capitalizing on those strengths by building the AC—echo the prospects of the HST as planned and the arguments for correcting the spherical aberration, as is planned in late 1993 on the first servicing mission.

5.2 OVERVIEW OF ATMOSPHERIC DISTORTIONS AND THE AO APPROACH

During the propagation of light from an astronomical object to a ground-based telescope, the wave front of the electromagnetic radiation is effectively unperturbed in its transit through space (neglecting dust obscuration and gravitational lenses). However, during the last 30 microseconds of the light's travel it encounters a distorting medium, the earth's atmosphere. Fluctuations in the atmosphere's index of refraction, due to temperature inhomogeneities, induce phase distortions on the wave front. These effects occur throughout the troposphere, from the ground up to an altitude of about 10 kilometers (km). The aggregate effect is a phase-distorted wave-front surface that has a fractal-like appearance. The wave-front surface is nearly same for all wavelengths due to the weak wavelength dependence of the index of refraction of air. There are three main parameters used to describe the atmospheric distortions: the lateral coherence length, ρ_o , the atmospheric correlation time, τ_o , and the isoplanatic angle, θ_o .

5.2.1 LATERAL COHERENCE LENGTH OF THE ATMOSPHERE, ρ_0

The distortions in the wave-front surface can be described by the phase structure function, which is the mean-square difference of the phase between two points on a plane that is orthogonal to the propagation vector. The two points are considered 'out of phase' when the mean-square difference is greater than 1 radian of phase angle; the separation of the out-of-phase points is termed the lateral coherence length of the wave front. For simplicity, the distortions are considered to be isotropic and homogeneous, and the lateral coherence length is stated as a scalar quantity, ρ_0 . A practical and intuitive definition of ρ_0 is the maximum aperture diameter that can attain a diffraction-limited PSF.

The phase of the wave front depends on wavelength, and longer wavelengths will have larger ρ_0 . Due to the random-walk nature of phase surface, ρ_0 grows faster than the first power of wavelength. Using the most widely accepted model of atmospheric turbulence (due to Kolmogorov), one finds that ρ_0 is proportional to $\lambda^{1.2}$. For one arcsec seeing, ρ_0 is approximately 10 centimeters (cm) in the visible spectrum near 5500 Å.

5.2.2 ATMOSPHERIC CORRELATION TIME, τ_0

The wave-front distortions evolve in two ways. On short time scales, the distortions can be considered to be frozen and blowing past the telescope at some given velocity; this is termed the Rayleigh model of atmosphere turbulence. In reality, the atmosphere is not frozen, and if one could fly along with the wave front at its mean velocity, the wave front would appear to evolve in a boiling fashion; this effect is Mintzer turbulence. For AO, Rayleigh turbulence defines the correlation time of phase variation over an ρ_0 -sized subaperture or zone of the telescope entrance aperture. Thus, the time over which the wave-front phase does not vary by more than one radian, τ_0 , is simply ρ_0 divided by the wind velocity, and $\tau_0 \propto \lambda^{1.2}$. In the visible, at a good site with good seeing, $\tau_0 \sim 2$ to 5 msec.

5.2.3 ISOPLANATIC ANGLE, θ_0

If the atmospheric distortions were concentrated in a thin layer immediately in front of the telescope, light from any angle in the FOV would get the same distortion. Neglecting the effects of obscuration, the PSF would be the same for all angles (isoplanatic).

Unfortunately, for ground-based imaging, the atmosphere is 3-dimensional and the distortions occur over a 10 km tall volume above the tele-

scope, which causes the distortions to be non-isoplanatic. To draw a simple picture that gives a physical understanding of non-isoplanatism, let us assume that all distortions occur in two thin layers—one immediately above the telescope and a second layer that is 10 km above the telescope aperture. Also assume that $\rho_o = 15$ cm in both layers. If the telescope looks directly overhead, a star that is on-axis will experience a total distortion that is the sum of the lower and upper distortions (after propagation of the upper atmosphere distortion) and the effective ρ_o will be less than 15 cm. For a star off the zenith, the upper atmosphere distortion is shifted relative to the lower atmosphere distortion and the overall distortion is different. At an angle of 3 arcsec (3 arcsec $\approx 1.5 \times 10^{-5}$ radian = 15 cm/10 km), the projected upper atmospheric distortion has shifted by a distance ρ_o . Thus, in this simple model, the isoplanatic angle is 3 arcsec. In practice for the 3-D atmosphere, the isoplanatic angle at visible wavelengths is typically about 5 arcsec; in the infrared (IR), θ_o ranges from 10 arcsec at 1.0 μm to 25 arcsec at 2.5 μm . As for ρ_o and τ_o , θ_o is proportional to $\lambda^{1.2}$.

We note that at least one AO group is attempting to expand the isoplanatic angle by using two or more adaptive mirrors, each conjugate with a layer that may be producing a distortion. While this approach may show promise if the layered atmosphere model has validity, the efforts are embryonic and largely theoretical; their costs and limitations are not yet well-defined, and we do not address this approach further in this report.

5.2.4 SCHEMATIC DIAGRAM OF ADAPTIVE OPTICS SYSTEM

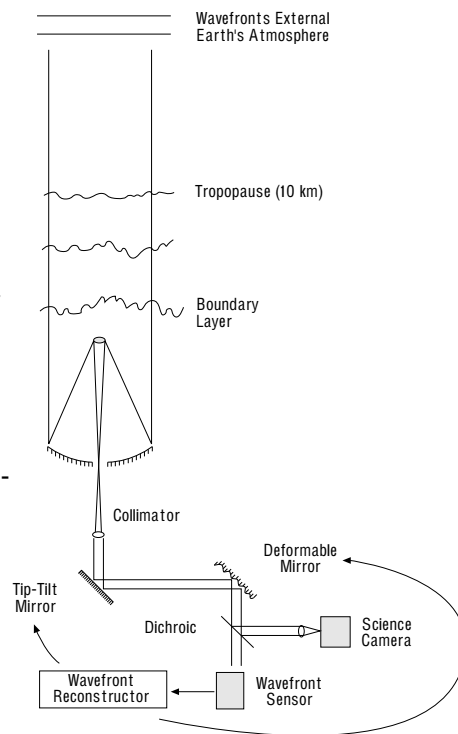
AO aims to correct for the atmospheric distortions by measuring and compensating ρ_o size subapertures on a time scale shorter than τ_o . The five main components of an AO system are shown in figure 5.1: (1) a wave-front sensor that measures the atmospheric distortions using light from a reference source, (2) a wave-front reconstructor that converts the wave-front sensor measurements into a phase surface and sends information on the required compensation to the adaptive elements, (3) a deformable mirror that bends to correct the distortion, (4) a tip/tilt mirror that is usually required to compensate the large image translations that are outside the dynamic range of the deformable mirror, (5) the imaging (science) camera.

As seen in figure 5.1, these systems are usually arranged so that the wave-front sensor looks at the wave front after correction; that is, it measures the residual wave-front error and commands the tip/tilt and

deformable mirrors to achieve a flat wave front.

There are various ways to measure the wave front, a variety of types of deformable mirrors, and alternate approaches to wave-front reconstruction. However, for current purposes, one only needs to be aware that these differences are not drivers in terms of ultimate system performance. While some approaches are more appropriate for more elementary AO, all systems have the same fundamental physical limitations, which we discuss in the next section.

Fig. 5.1 Main components of adaptive optics systems. The AO system attempts to correct in real-time the distortions that occur in the troposphere, dominated by boundary layer effects and the jet stream (at the tropopause, about 10 km above sea level). The critical components of an AO system are the deformable mirror (DM) and the wavefront sensor (WFS). The most advanced DM in use today has about 500 actuators and can distort the surface $\pm 4 \mu\text{m}$. The WFS either measures the first order (slope—Hartmann-Shack design) or second order (curvature—Roddier approach) properties of the wavefront. A critical aspect of wavefront sensing is the efficiency and noise of the detector; new generation high speed, low noise CCD devices define the state-of-the-art.



5.3 PHYSICAL LIMITATIONS OF AO AND COST DRIVERS

The fundamental physical limitation to AO performance is incomplete information about the wave-front distortions. The wave-front measurement is limited by sampling in both space and time, and it is subject to errors due to photon noise. Wave-front sensors measure the tilt (or curvature) of the wave front in individual subapertures averaged over the sample time. The subapertures cannot be made arbitrarily small, nor the sample times too short because the reference source is not

infinitely bright. The ratio of the maximum number of subapertures to the sample time proportional to the brightness of the reference source.

The distortion measurements are undersampled in both time and space, and they cannot be used to compensate for distortions at spatial or temporal frequencies than those for which information has been obtained. Distortion power at higher spatial frequencies will scatter into a halo of the size of the ambient seeing disk; that at higher temporal frequencies will appear as fluctuations in the image.

The optimal trade-off between subaperture size and sample time for given light level is the most basic design choice for AO systems. If one wishes to provide high-order correction, the subapertures must be small and sample times short, with the requirement that the wave-front reference source be bright. In order to utilize fainter reference sources, the subapertures can be made larger and cycle times longer; this is termed low-order AO. A minimal number of photons must be detected in each subaperture for the distortion measurement to be useful; most AO system designers require that at least 100 photons be detected for a useful measurement.

One approach to system design is to match the subaperture size to ρ_0 , estimate the maximum sampling time that is useful and then calculate the faintest reference source that will provide 100 detected photons. As an example, let us assume 15 cm square subapertures and 2 msec sample time. In the best case, with 90% atmospheric transmission, 50% optical efficiency, and 0.80 electrons (e^-) per photon detective quantum efficiency (DQE) in the detector over the band 4000 to 10,000 Å. Using the flux calibration of 940 photons/sec/cm²/Å for an $m_v = 0$ star, the number of photons detected per subaperture per frame time is:

$$\begin{aligned} \text{photons/subaperture/frame time} &= (940)(15^2)(0.002)(6000)(0.9)(0.5)(0.8) \\ &= 920,000 \quad (m_v = 0 \text{ star}) \\ &\approx 100 \quad (m_v = 10 \text{ star}). \end{aligned}$$

Equation 5.1. Typical calculation of the required reference star brightness.

The presently operating AO systems do not have the high efficiency assumed in equation 5.1 and thus are limited to $m_v = 6$ to 7. By 1999, however, reference stars as faint as 10th magnitude may be accessible to high-order AO systems. A low-order AO system—adequate for IR imaging at a good site—could use 60 cm subapertures with a 10 msec cycle time, which would make 15th magnitude reference stars useful.

5.3.1 LASER REFERENCE STARS

The magnitude limit for AO with natural starlight can be overcome by using an artificial reference star produced by a laser which is collocated with the imaging telescope. Successful use has been made of Nd:YAG and copper vapor lasers to produce reference stars from Rayleigh scatter in the lower atmosphere at 6 to 10 km range; eventually these systems may be able to work up to 20 km range.

An artificial reference star located at this close distance is a poor probe of the upper atmosphere distortions. This is primarily due to the geometric 'cone effect' shown in figure 5.2, also known as focal anisoplanatism. A better solution would be to place the reference star at a greater distance to minimize the cone effect.

Due to a unique conjunction of disparate physical phenomena, it is possible to make reference stars in the mesosphere between 85 and 100 km (MSL), where there is a tenuous layer of sodium atoms, a few thousand atoms per cm^3 . Sodium exhibits strong resonance fluorescence at 5890 \AA , a wavelength for which it has proven difficult to make a high power monochromatic laser. By a fortunate coincidence of nature, however, a laser source for sodium layer excitation can be made by mixing two lines of a Nd:YAG laser; the $10,060 \text{ \AA}$ line mixed with the $13,190 \text{ \AA}$ line will produce 5890 \AA radiation. While lasers of adequate power for producing a sodium layer reference stars are presently under development, no AO system has yet become operational with this technology.

Even though a sodium layer reference star provides a tenfold geometrical improvement over present day Rayleigh reference stars, the cone effect still cannot be neglected for large apertures. It is commonly held that a single sodium layer reference star is only adequate for correction of 3 to 4 m apertures; for larger apertures an array of reference stars is required and the wave-front measurements need to be stitched together to provide the full aperture correction signals. Thus, a number of technical problems need to be overcome to provide full aperture correction to the largest telescopes.

The artificial reference stars also have a peculiar problem. Although they hold promise for high-order correction, they are not useful for the most basic measurement of the atmosphere: the tilt of the wave front. This is due to the reciprocity of light propagation. If the atmosphere is thought of as behaving as a wedge of refractive material, the light will retrace its path upon return through the atmosphere. There is no proven

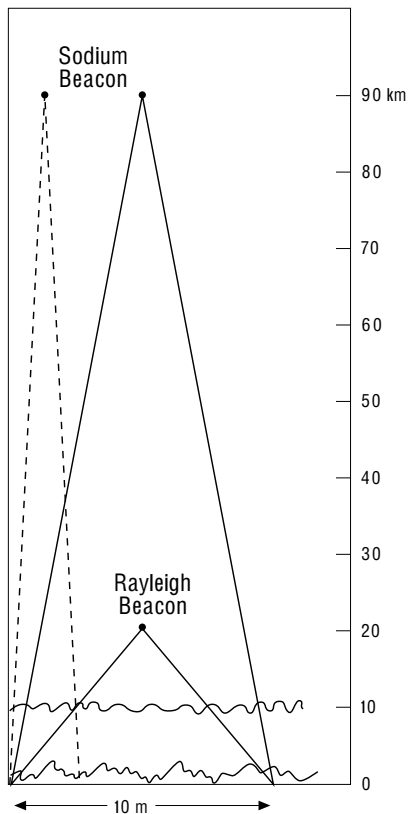


Fig. 5.2 Cone effect of laser guide star. The light rays from an astronomical object, which is effectively at infinity, are all parallel to each other. However, a laser guide star generated in the Earth's atmosphere is at a finite distance, and its light rays do not propagate through the same path to the telescope mirror. Thus, the distortion to the wavefront phase at the mirror surface is not the same for laser guide stars and astronomical objects. This effect is greater for larger telescopes and for close range guide stars. For a 10-meter class telescope with a Rayleigh guide star, the wavefront errors induced by this so-called "focal anisoplanatism" would dominate all other sources of error. Thus, there is great interest in development of inexpensive, reliable sodium wavelength lasers with the power required by astronomers (2 to 20 W).

solution for this effect other than to utilize some natural light, from either the object of interest or a nearby source to track the wave-front tilt. Advances on this problem are hard to predict, and until solutions appear ground-based AO will have a disadvantage compared to HST in terms of magnitude limit (where the target must supply tilt information) and sky coverage (where a nearby reference star must be used).

Even without any cost constraints, there are fundamental physical limits to the performance of AO systems using laser reference stars. The cone effect is one, but more basic are the same wave-front measurement errors that are associated with natural starlight AO systems. The subapertures cannot be made arbitrarily small, nor the cycle times too short, because not enough photons will be detected for a useful measurement. There are limits to the brightness of the sodium layer bright star because the sodium layer will saturate if pumped with too much power.

5.3.2 LASER REFERENCE STAR IMPLEMENTATION CONSTRAINTS

The physical limits notwithstanding, the performance of AO in the green and blue during the next decade will be limited primarily by the funding level for developing AO and focus of the major AO groups.

The use of a larger number of subapertures requires that all parts of the AO system scale up in a comparable fashion. There must be a larger number of wave-front sensor detector pixels, a correspondingly larger and faster computer, and a deformable mirror with a large number of degrees of freedom. Each of these components is quite expensive for large scale systems; the military systems now in use cost many millions of dollars. Cost is a more important issue for astronomical than for national security applications.

The high-performance sodium-layer reference star systems face a number of potentially costly technical challenges. Bringing an appropriate working laser on-line will require both the initial acquisition cost and a sizable operational cost estimated at several hundred dollars per hour. Also, a sizable staff may be required to run and maintain the system and to provide the safety precautions required by the FAA. (Interestingly, the respondents to our AO survey did not feel that regulatory approval for high power laser operation was a significant issue. Some respondents had attained limited approvals and found that the regulatory process was not unduly restrictive.)

5.4 PROGNOSIS FOR AO PERFORMANCE

We have developed our prognosis for the optical performance of AO in 1999 from the expectations of current ground-based AO programs. The main progress will be in achieved resolution (Strehl ratio) and operational duty cycle. We have accepted AO limitations based on physical principles as fixed: FOV size equals isoplanatic angle, halo due to seeing image being uncompensated outside subaperture diffraction FWHM, and magnitude limitation due to tip/tilt.

5.4.1 CURRENT GROUND-BASED AO PROGRAMS

There are several strong efforts aimed at the advancement and implementation of adaptive optics. In the U.S., the National Science Foundation (NSF) is supporting AO system development at the University of Chicago, University of Arizona and University of Hawaii. Private foundation, other federal agencies, and state funds are also expected to play

an important role in AO: Mt. Wilson Observatory, the Carnegie Institution, Lawrence Livermore National Laboratory and the California Association for Research in Astronomy have AO programs that are either underway or in the nascent stages.

In Europe, the Meudon Observatory and collaborators is one of the leaders in AO. The European Southern Observatory has plans to make AO an integral part of the Very Large Telescope (VLT), and the Max Planck Institutes in Germany have several groups planning to implement some aspect of AO.

To date, the largest AO program has been that of the U.S. military establishment. The U.S. Air Force operates systems in Maui and Albuquerque and supports a Rayleigh reference star system at the Massachusetts Institute of Technology's Lincoln Laboratory. U.S. Department of Defense funds have also supported the National Solar Observatory AO system and a number of industry efforts, including those at Litton Itek Optical Systems and ThermoTrex (formerly Thermolectron Technologies Corporation).

This list of AO projects is not all-inclusive; combined with other efforts and proposals presently in review, the AO community has apparently achieved the vitality and critical mass necessary for fulfilling the predictions made in this report—given adequate funding. These groups represent the scope of efforts upon which we have based our prognosis.

Through our communications with these groups, we find that ground-based AO systems are proceeding upon two distinct paths as defined by funding levels and group interest. Perhaps not surprisingly, the groups divide into those focused on astronomical research and those concerned with military applications.

As discussed in section 5.3, AO becomes much more difficult—and expensive—as one attempts to correct shorter wavelengths; the cost is driven by the associated demands for shorter response time, more pixels on the wave-front sensor, more actuators on the deformable mirror, higher power lasers (for reference stars), and larger and faster computers.

The cost of implementing visible light correction on an HST-sized aperture can be estimated from the military systems. The two most prominent are under the purview of Phillips Laboratory of the U.S. Air Force; these are the 3.5 meter AO telescope being completed in Albuquerque and a similar system planned for Maui, which will have darker skies and probably better seeing. The AO system at Albuquerque is esti-

mated to cost \$2 million per year to upgrade and operate, thus the total AO system cost is about \$15 million through 1999. The future availability of these systems for astronomy is uncertain.

It is easier and less costly to implement AO in the IR on a larger telescope than to attempt visible light correction on smaller telescopes. Since astronomers wish to get productive systems on-line with minimal cost, nearly all astronomical groups are focusing on correction in the IR. The performance of astronomical AO systems optimized for IR on a large telescope will degrade gracefully toward shorter wavelengths.

Our performance projections for the two types of ground-based AO systems are shown in figure 5.3 as two zones on a plot of Strehl ratio as a function of wavelength. We derived the performance projection from the replies from the our AO survey, a number of articles, and conversations with AO researchers. The exact shape of the curve can be debated, but unless the performance is significantly enhanced for $\lambda < 6000 \text{ \AA}$, the conclusions reached in the next section are fundamentally unchanged.

It is interesting to note that achieved Strehl ratio is largely independent of telescope aperture. This is because the performance depends most strongly on the photon noise and sampling errors of individual subapertures. Of course, there is a geometric increase in cost for enlarg-

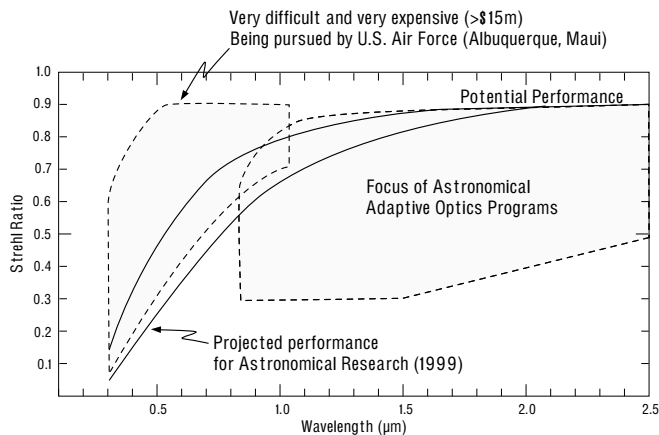


Fig. 5.3 Projected performance of AO systems: Strehl ratio. High Strehl will be difficult and expensive to attain in the visible. Thus, all major astronomical AO programs have targeted the infrared for application of this technology. Only the U.S. Air Force is presently concentrating in earnest on high order correction in the visible. Since the U.S. Air Force telescopes will not be routinely used for astronomical investigations, the HST should retain uniqueness space for high resolution visible imaging.

ing telescope aperture. One should also note that the resolution of the large telescope astronomical IR systems will be comparable to that of the military visible light systems due to the difference in aperture.

5.5 CORE:HALO INTENSITY RATIO

An important parameter to assess in an imaging system is the core:halo intensity ratio. A simple first order model of the AO PSF is a diffraction-limited core combined with a uniform halo the size of the ambient seeing disk. In practice the halo will extend beyond the seeing angle (usually defined by FWHM) and it will be closer to a Gaussian profile than a uniform disk. However, the simple "thumbtack" model is useful for providing a first order estimate of system performance. The fraction of light that is concentrated in the core is approximately equal to the Strehl ratio, and the remaining light is spread uniformly in the seeing disk which has an angular width of λ/ρ_0 . The core:halo intensity ratio is thus,

$$\frac{\text{core}}{\text{Strehl}} = \frac{\text{Strehl}}{(1-\text{Strehl})} \left(\frac{D}{r_0}\right)^2$$

Using the projected performance shown in figure 5.3, the core:halo ratio is plotted in figure 5.4. Interestingly, this ratio is about 100 for most of the visible band; the increase in Strehl ratio at longer wavelengths is compensated by the smaller number of resolution elements within the seeing disk.

At a 0.25 arcsec radius the aberration-corrected HST core:halo ratio is 1000 to 4000 in the visible, depending on the location and strength of the PSF sidelobes. We will adopt a value of 1000 so as to be certain that no sidelobe of the HST PSF can be confused with presence of a secondary source.

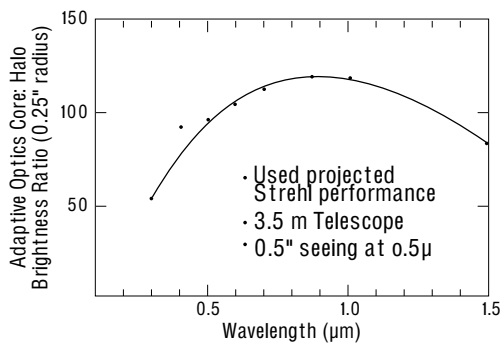


Fig. 5.4 The projected AO core:halo contrast ratio. For imaging faint objects near bright objects or imaging low contrast features on an extended object, the core:halo contrast ratio is more important than FWHM of the PSF. This ratio is projected to be about 100 for 0.5-1.5 μm imaging.

5.6 DYNAMIC RANGE AND CONTRAST

A larger core:halo intensity ratio improves the ability to detect faint companion sources and image low contrast extended objects at high resolution. To demonstrate these effects, we show two extremes of imaging in figure 5.5: detecting a binary source of very different brightnesses and measuring a low contrast feature on a uniform background.

Both of these imaging scenarios will be limited primarily by photon noise. Let us assume that we use a CCD with a full well capacity of $100,000 e^-$ and that we do not allow any pixel to saturate in order to obtain photometric information.

For the AO system, if we set the exposure such that the primary component fills its core pixel to full well, the halo of the primary will have a level of $1000 e^-$. The standard deviation of this halo (due to photon noise) will be about $32 e^-$. If we wish to detect a secondary at the 3σ level, the secondary must contribute at least $96 e^-$. This means the AO system has a dynamic range of ~ 1000 or 7.5 stellar magnitudes. A similar calculation for the HST gives a dynamic range of 33,000 or 11.3 stellar magnitudes.

It should be emphasized that the halo of the HST PSF is due primarily to diffraction, which can be substantially reduced by implementation of a coronagraphic mode. The wings may be reduced by as much as a factor of 100, providing an effective core:halo ratio of greater than 100,000 for detection of a faint secondary near a bright source.

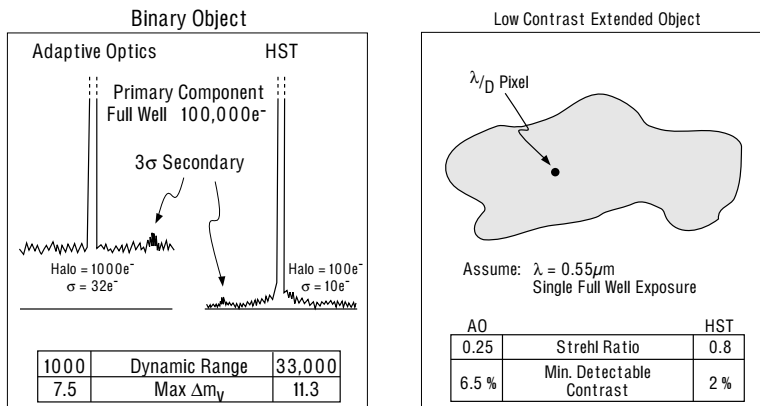


Fig. 5.5 Two extreme imaging problems for which the core:halo intensity ratio is critical. On the left: a high-contrast binary. On the right: a low contrast feature on an extended source. The aberration-corrected HST will enable detection of much larger contrast ratio binary objects and higher quality imaging of extended low contrast sources.

The second imaging scenario involves detecting low contrast features on an extended object. Here, too, sensitivity is critically dependent on the core:halo ratio. Using Strehl ratios of 0.25 for AO and 0.8 for HST, we find that a uniform extended disk will have 3σ photon noise of 6.5% of the mean intensity of the disk for AO imaging whereas the HST will be able to be able to make a 3σ detection for intensity variation of 2%.

5.7 PSF STABILITY

The uniqueness space of HST is further defined by the spatial and temporal stability of the PSF.

The FOV limitation of AO is well known. Possibly a set of imagery could be patched together to obtain a larger FOV. This approach has not been explored and would be difficult for objects that change rapidly, such as planets in our Solar System. (see section 1.3.1) Faint large field objects that require long exposures would also be difficult for AO. The relative fields of view of ground-based AO and an axial bay HST AC are shown in figure 5.6. Note that isoplanatism is best preserved by an advanced camera that uses non-vignetting relay optics.

The ground-based AO PSF will exhibit a time-variable halo of light, due to the remnant seeing disk after correcting the lower spatial and temporal frequencies of the atmospheric distortion.

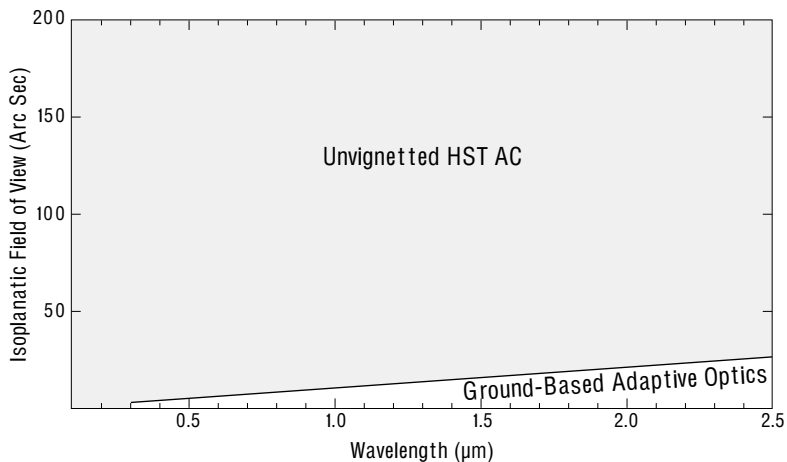


Figure 5.6. The field of view over which diffraction-limited image quality is available. Comparison of HST and ground-based AO. One of the best qualities of HST imagery will be the wide field over which high quality imaging and photometry can be obtained.

An issue which has not been discussed much in the AO literature are limitations to the length of AO exposures. The AO systems we are familiar with have difficulty maintaining lock, and the showpiece AO imagery often are 10 msec exposures during which the system was "locked in". We expect significant strides to be made in temporal stability of AO systems, but establishing the capability to routinely take multi-hour exposures may prove to be a tough task.

5.8 CONCLUSIONS

We expect that ground-based AO will make significant technical strides during the 1990's. The advancements should include development of reliable lasers for generation of sodium layer reference stars, very low noise wave-front sensor cameras, large degree of freedom deformable mirrors at lower expense and special purpose computers to run the AO systems.

The level of funding and diversity of funding sources appears to be adequate to accomplish the goal of wave-front compensation in the IR. However, due to the limits on funding levels, the astronomical community does not plan to pursue high quality correction for $\lambda < 6000 \text{ \AA}$.

Ground-based AO will achieve higher resolution than HST in the visible and near IR, due to diffraction-limited performance with the larger apertures of ground-based telescopes. HST will provide a much better core: halo ratio than AO, which will be limited in quality due to the residual halo from lack of full compensation. The AO halo, which will be significant in the visible, limits the dynamic range of imagery and the detection of low contrast features on an extended object.

HST will provide a much larger FOV than ground-based AO: 200 arcsec for the HST AC in an axial bay, compared with 5 arcsec in the visible and 10 to 25 arcsec in the IR for AO. Thus, wide field photometry and wide field imaging of faint sources will best be done with the HST.

Other difficulties of AO, such as the tilt determination problem of laser reference stars and the long term stability of the lock of an AO system, may lead to an advantage of HST in terms of sky coverage and magnitude limit.

All of these performance projections when combined with the UV capability of HST, lead us to conclude that HST will have a unique portion of parameter space within which the advanced camera can pursue unequalled astronomical research.

ACKNOWLEDGMENTS

James Beletic was the lead author of chapter 5.

We wish to acknowledge and thank those who took the time to make thoughtful and insightful responses to the AO survey: Roger Angel, Horace Babcock, Jacques Beckers, Steven Beckwith, Chris Dainty, Robert Fugate, Brent Ellerbrock, Edward Kibblewhite, Tadashi Nakajima, Malcolm Northcott, Ronald Parenti, Rene Racine, Bruce Sams, David Sandler, Donald McCarthy, Laird Thompson, and Peter Wizinowich. The survey responses were critical to the validity of this document.

Mark Colavita and Mike Shao wrote an excellent white paper for the panel, as did Chris Burrows and Pierre Bely. Richard Goody gave us some valuable mid-course guidance during our synthesis process.

REFERENCES

- Beckers, J.M. (1993). "Adaptive optics for astronomy: principles, performance and applications," to be published in *Annual Reviews of Astronomy and Astrophysics*, 31.
- Fugate, R.Q., editor (1992). *Laser Guide Star Adaptive Optics Workshop, Proceedings 1 & 2*, 10-12 March 1992, Albuquerque. These volumes contain over 50 articles that present most of the latest research in this area.
- Parenti, R.P. (1992). "Adaptive optics for astronomy," *The Lincoln Laboratory Journal*, 5, no. 1, pp. 93-114. The entire issue of this journal is devoted to adaptive optics work at MIT Lincoln Laboratory and the interested reader should obtain the full issue for review.

6. Axial vs. Radial Bay

The Advanced Camera (AC) must be designed for installation into either an axial or a radial bay. We have examined the scientific, technical, and logistical considerations that bear on this issue. We recommend that the AC be designed as an axial instrument.

Hubble Space Telescope (HST) contains five instrument bays to accommodate scientific instruments (SIs), as illustrated in figure 6.1. Four of these bays view the focal plane directly and are called 'axial bays'. One instrument bay—the 'radial bay'—is located at right angles to the telescope beam and receives the beam via a 45° pickoff mirror. Whether to build the Advanced Camera (AC) for an axial or the radial bay position is a decision that must be made early in the development process.

A new camera developed for either bay type would benefit from designs and hardware inherited from previous SIs. For example, ground test equipment exists to test and calibrate either option. Appropriate car-

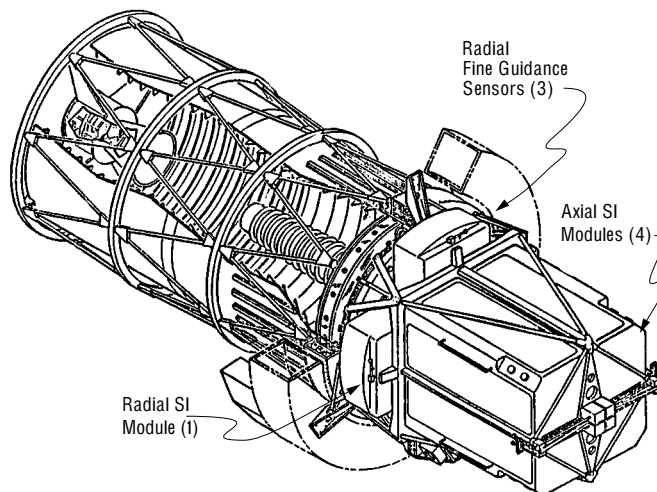


Figure 6.1. Axial and radial instrument bays

riers and tools exist to transport either an axial bay camera (ABC) or a radial bay camera (RBC) to orbit and install it in HST. The ABC and RBC would have similar mechanical stability, thermal stability, and heat rejection. Nevertheless, the axial and radial bay options are quite different in other important ways, due primarily to differences in shape, volume, and orientation as illustrated in figure 6.2. These factors constrain the optical design and the physical layout of the AC and influence its complexity, risk assessment, and scientific performance. There are also logistical and operational comparisons to consider. After an investigation summarized in table 6.1, we have concluded that the axial bay position for the AC is clearly superior to the radial bay option.

Table 6.1. Tradeoffs between axial and radial bays. Check marks indicate factors favoring axial bay.

	<u>Radial</u>	<u>Axial</u>
Existing design, hardware	Yes	Yes
Carrier compatibility	Yes	Yes
Mechanical stability	Similar	Similar
Thermal stability	Similar	Similar
Heat rejection	Similar	Similar
Relative UV throughput	<0.5	1.0 ✓
Risk of contamination	Higher	Lower ✓
Bay availability	Working WFPC II?	Likely 2 - 3 free ✓
Guide star availability	Marginal	Adequate ✓
Detector selection	Biased to CCD in UV	Neutral ✓
Parallel science	Not favored	Yes ✓
Polarimetry	Bad 45° reflection	Yes ✓
FOV size	160" x 160"	200" x 200" ✓
Room for filters	48	More ✓
Pickoff mirror concerns	Yes	No ✓
On-orbit installation	More difficult	Less difficult ✓

6.1 FAR ULTRAVIOLET (FUV) THROUGHPUT

As discussed in chapter 6, the optical designs we explored for the ABC require a minimum of three reflections in each of three beams. To avoid collimation problems, two of the three ABC beams must have one additional, oblique reflection. Because of the pickoff mirror, our RBC designs require a minimum of four reflections in each of three beams, and—due to space considerations—the fixed triad RBC option calls for two additional folding reflections in two of the three beams. In general terms, these two folding mirrors are a consequence of the fact that the RBC volume of the is 31% less than the ABC, and the maximum throw behind focus is 25% shorter, as illustrated in figure 6.2.

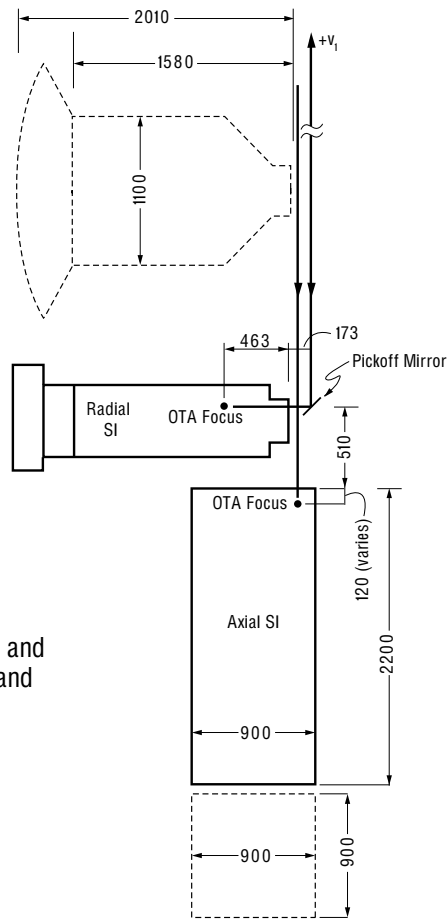


Figure 6.2. Dimensions in mm, and relative orientations of radial and axial bay instruments

Whereas the ABC oblique reflections would be without appreciable losses, the one (or three) additional RBC mirror(s) would reduce the far ultraviolet (FUV) throughput of the RBC by a factor 0.7 (or 0.34) relative to the ABC, assuming 70% reflection efficiency per mirror. The axial bay is favored.

6.2 RISK OF CONTAMINATION

The greater design flexibility of the ABC volume permits precautions to control contamination that are more difficult for a RBC, such as dewars to isolate the cold charge-coupled device (CCD) detectors. Also, the optics are all contained within the protective envelope of the instrument for an ABC, whereas the pickoff mirror is in a more vulnerable outside position for the RBC. The axial bay is favored.

6.3 BAY AVAILABILITY

Because one SI must be removed to make a place for the AC, the expected status of the on-board SIs around the year 2000 is an important consideration, as it would be unfortunate to replace a functioning, productive instrument. As discussed in chapter 2, the Faint Object Camera (FOC) will be far beyond its expected lifetime and has a low probability of useful operation at the time in question—due mostly to its high-voltage detectors. Also, after the second servicing mission, the FOC will be the only SI still dependent on the Corrective Optics Space Telescope Axial Replacement (COSTAR) for its optical correction.

Depending on the date of the AC installation, another axial SI—the Near Infrared Camera (NICMOS)—might be due for change out. The NICMOS detector will be cooled by a cryogen, which will be depleted with time and operations on orbit. The lifetime of the cryogen will be 3 to 5 years depending on the schedule of use. Thus, the opportunity may exist to replace the NICMOS with the AC, return the NICMOS to the ground, replenish the cryogen, and re-fly the NICMOS. In this instance, the NICMOS would provide a very low cost instrument for the next service mission.

The second Wide Field Planetary Camera (WFPC II) will occupy the radial instrument bay when the AC is installed. While it will be beyond its design life, the WFPC II contains no consumable materials and employs low-voltage solid state detectors. For these reasons, it is likely that the WFPC II will be at least partially operational in the post-2000 period.

Thus, at least two and perhaps three of the four axial bays will likely contain inoperative SIs at the time the AC is installed, whereas the one radial bay will likely contain a working WFPC II. If the radial bay is chosen for the AC, the almost certain result would be just one operational camera on HST. On the other hand, selecting a likely-available axial bay would increase the probability of two functional cameras. The axial bay is favored.

6.4 GUIDE STAR AVAILABILITY

As discussed in section 4.2, more guide stars are available to an axial SI than to a radial SI due to increased area of the sky swept out by the FGS field of view (FOV) as the spacecraft rolls. As shown in table 4.2, an estimated 25% of mid-galactic latitude targets would be unavailable to

the AC installed in a radial bay, a loss that increases to 45% at high latitudes. The corresponding losses are 1% and 5% for the AC in an axial bay, and hence the axial bay is favored.

6.5 DETECTOR SELECTION AND PARALLEL SCIENCE

Three of the four camera designs we have explored have independent detectors for each beam, and in principle these detectors can operate simultaneously to provide a parallel science capability. The exception, the RBC carousel option, requires one detector, which would necessarily be a CCD to satisfy both long- and short-wavelength requirements. Selecting the RBC carousel option would resolve the issue of the FUV detector not on the basis of performance but of versatility. Of the two RBC options, the carousel design is somewhat superior due to higher UV throughput. Therefore, because it better assures that detector selection will be based on merit and that parallel science operations will be feasible, the axial bay is favored

6.6 POLARIMETRY

The large angle of the pickoff mirror eliminates the possibility of polarimetry with the RBC. The axial bay is favored.

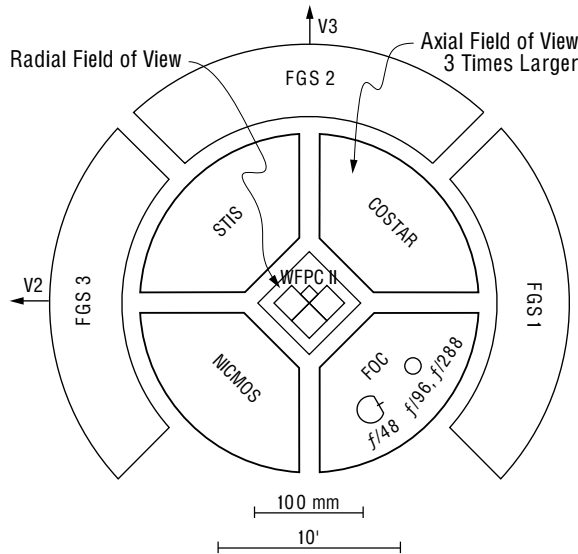


Figure 6.3. Instrument FOVs as projected onto the sky. Expected configuration after 1997 servicing mission.

6.7 FOV SIZE

The maximum RBC FOV is about 160 x 160 arcsec to avoid vignetting the FOVs of the axial SIs. As shown in figure 6.3, the axial bays have a three times larger area available to accommodate larger FOVs. As discussed in chapter 11, ABC optical designs can achieve full correction of spherical aberration and diffraction-limited performance over at least 200 x 200 arcsec FOVs. The axial bay is favored.

6.8 ROOM FOR FILTERS

The WFPC II has a total of 48 filters, limited mainly by the space constraints of the radial bay. The larger volume of the axial bay increases the space available for the filter complement, perhaps by as much as 50%. The axial bay is favored.

6.9 PICKOFF MIRROR CONCERNS.

The RBC's pickoff mirror appears to be an unavoidable single point of failure, located as it is in a vulnerable, extended position. It appears to be both at physical risk and susceptible to contamination, as discussed in section 6.2. Also, the pickoff arm may have reduced dynamic stability: its position must be stable to about 1 micron to avoid detectable image degradation. The axial bay is favored.

6.10 ON-ORBIT INSTALLATION

The delicate light seal of the RBC radiator may result in a more difficult on-orbit installation as compared with the ABC, which is within the light-tight aft shroud. The axial bay is favored.

ACKNOWLEDGEMENT

James Crocker was the lead author of chapter 6.

7. Sampling the Image

What image sampling strategy is appropriate for the Advanced Camera (AC)? The ability to critically (Nyquist) sample the image is required at all wavelengths. This may be accomplished either by sufficiently fine pixelation in the detectors or by substepping: moving the telescope slightly and taking another exposure. Substepping may be required in the AC's wide-field mode.

The choice of how finely to sample the image produced by the Optical Telescope Assembly (OTA) and Advanced Camera (AC) optics has profound effects on the instrument design. That is, the process of decision making involves technical compromises depending on the characteristics of the detector(s) and the physical constraints imposed by the spacecraft. Here we explore the sampling requirements imposed by the scientific case for the AC to see how these can be achieved given our projections of optical and detector performance. The expectation is that the choice of sampling depends critically on the parameters that are to be estimated from the data and on the signal-to-noise ratio attained by the observations. The availability of suitable image processing tools also bears on our degree of confidence in different approaches. It is desirable to search for a strategy that can, by suitable choice from an array of available observation modes, achieve a wide range of goals in an optimal manner.

It is important in the following discussion to distinguish clearly between the 'pixel size'—which is a characteristic of the detector and camera focal ratio—and the 'sampling'—which can be determined by multiple exposures with different pointings on the sky. A space-borne camera with current detectors, can achieve significant advantages from 'interleaved sampling' (sub-stepping) strategies coupled with suitable image combination algorithms.

We first sketch out in general terms the impacts of the sampling choice and show that some of these are substantially different with dif-

ferent types of detector. The discussion is based on critical (Nyquist) sampling of the Airy function for a circular aperture. The necessity for sub- and super-Nyquist sampling is investigated using model calculations and simulations based on realistic estimates of camera and detector parameters. The experience of restoring the currently aberrated images is relevant since the aberrated point spread function (PSF) still shows characteristics of diffraction-limited performance in the core. Finally, we give some recommendations for the camera design based on these considerations, including suggestions for those parameters that need to be well measured and understood before final design decisions are made.

7.1 GENERAL IMPLICATIONS

The set of closely linked characteristics that are affected by or will determine the image sampling are sketched in figure 7.1. As a reference for the discussion, figure 7.2 gives a geometrical overview of the sampling parameters on which are marked the first and second generation of Hubble Space Telescope (HST) imaging instruments expected to be in operation after the 1997 servicing mission (see chapter 2). The lower part of the figure gives the locus for critical (Nyquist) sampling by a detector array as a function of wavelength. We have used a critical interval of half the distance to the first minimum of the Airy pattern for a circular aperture, $0.61 \lambda/d$, where d is the diameter of the aperture. Since sky background and

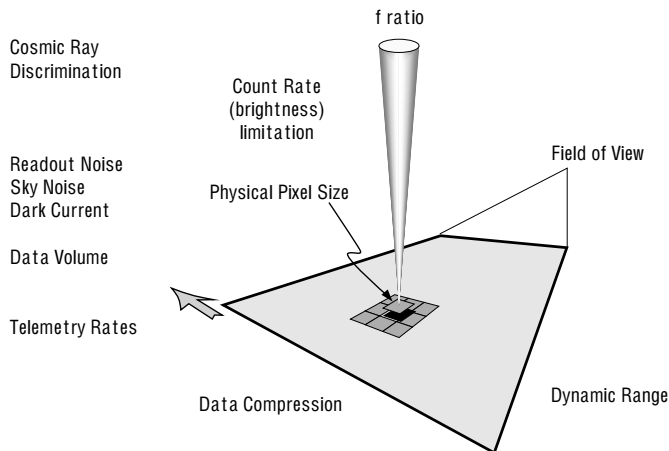


Figure 7.1. Some of the implications of the choice of how finely to sample the image.

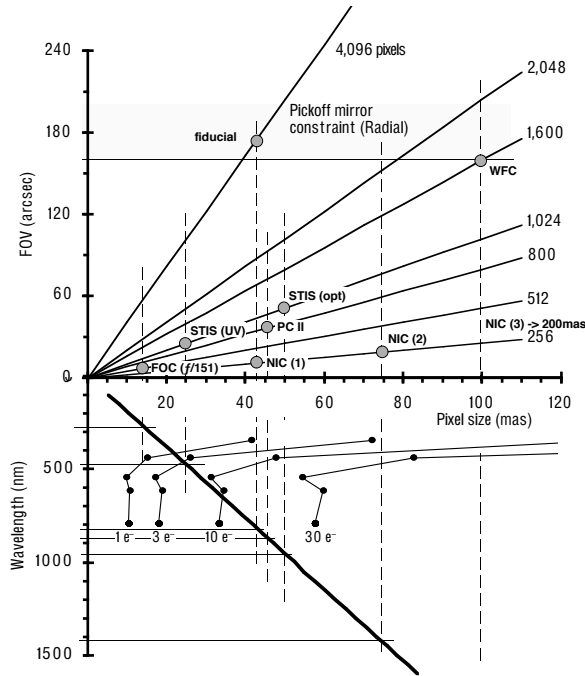


Figure 7.2. The relationship between wavelength for critical (Nyquist) sampling and camera field of view for detectors with a set of different pixel array dimensions recording the image without sub-stepping. The projected set of post-1997 HST imaging instruments is marked along with the 'fiducial' camera discussed in the text and described in table 1. The FOV limit imposed by the pickoff mirror for a radial camera is shown. The lines (with •) in the lower panel show the loci where sky equals readout noise for a CCD detector working in the current WFPC wide filters (F336W, 439W, 555W, 622W, 791W) with readout noise from 1 to 30 e^- .

detector noise will, in practice, limit the size of the pixels, we show in figure 7.2 some loci on which readout noise of a Charged Coupled Device (CCD) equals sky noise for some broad-band filters. This gives a relatively crude picture of the very limited regimes—broad-band imaging near the peak of the detector spectral sensitivity—in which a CCD detector can be used effectively with critically sampling pixels and suggests strongly that we have to search for alternative strategies.

Figure 7.3 shows the output focal ratio required to critically sample the image with the basic pixel array as a function of wavelength. Curves are plotted for four different physical pixel sizes from 7.5 μm (the minimum size expected from available technologies) to 25 μm . The post-1997 instruments have pixels of 40 μm Near-Infrared Camera (NICMOS, all

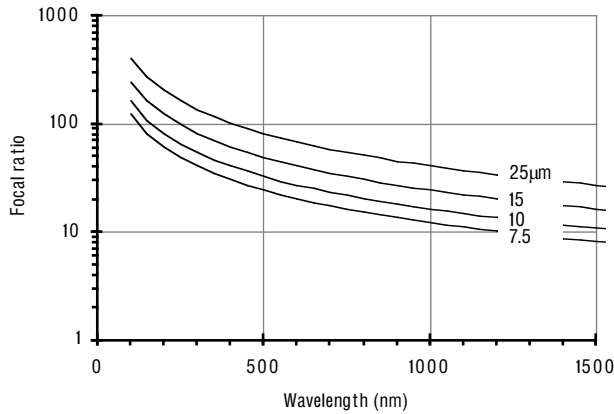


Figure 7.3. The required camera output focal ratio to achieve critical sampling for a set of four physical pixel sizes from 7.5 to 25 μm .

cameras), 25 μm Faint Object Camera (FOC), 21 μm Space Telescope Imaging Spectrograph (STIS optical) and 15 μm Wide Field Planetary Camera (WFPC II, STIS UV).

The simultaneous desire for wide field of view and appropriately fine sampling leads to very large data volumes, especially for the shorter wavelengths. This carries immediate implications for the type of on-board computing needed in order to overcome the limited spacecraft system telemetry rate of 1 Mbit/sec (see chapter 10). To put these concerns into perspective, we consider a 'fiducial' camera/detector system with a focal ratio of 20 and a detector with an array of 4096 x 4096 10 μm pixels. Table 7.1 shows the characteristics of such a device which would critically sample only in the near-infrared.

Table 7.1. A fiducial camera/detector to illustrate some basic quantities.

Pixel pitch	10	μm
Detector size	41	mm
f /ratio	20	
Pixel	43	mas
FOV	176	arcsec
	2.94	arcmin
number of pixels	1.68×10^7	
pixels per row & column	4096	
data volume (2 byte/pix)	32	megabytes
1 dumptime at megabit/sec	4.4	min
50 kHz read time	5.6	min

7.2 DETECTOR SPECIFIC IMPLICATIONS

Further detailed implications depend on the nature of the detector operation and we distinguish two basic types for discussion in turn. (See chapter 8.)

7.2.1 INTEGRATING DETECTOR (CCD)

CCDs can currently be fabricated with physical pixel sizes as small as $7.5\ \mu\text{m}$ although these have full-well capacities limited to around $25,000\ e^-$. The capacity increases approximately with the pixel area and so a balance between the achievable dynamic range in a single exposure and the optical relay magnification will have to be achieved.

It may not always be desirable to operate this type of detector at full resolution because of the effect of readout noise or data-rate limitations. These parameters are coupled by the need to read the detector relatively slowly for the best noise performance ($50\ \text{kHz}$ for 1 to $2\ e^-$ rms noise). While pixels can be binned on chip to achieve greater surface brightness sensitivity, there is still clearly a field of view (FOV) penalty incurred when the pixels have a small angular size.

An important issue for these detectors is the very high rate of cosmic ray induced charge generation. Fine sampling of the PSF in a single exposure allows for considerably easier discrimination of such events from real image features. If on-board cosmic ray cleaning is considered, full sampling by the pixel array would be an advantage.

At the long focal ratios needed for a space-based camera and typical detector sizes, the thermally generated dark current in a CCD can become a significant source of noise compared with the sky and readout contributions. Care has to be taken, therefore, to reduce this to an acceptable level by adjustment of the operating temperature and/or keeping the physical area of the resolution element projected onto the detector sufficiently small. The balance between the various noise sources for the current generation of CCD detectors used in a space-based camera is discussed in the appendix. There, we make a specific comparison between a single, one-orbit observation made with a critically sampling array in the V-band with four, half-pixel stepped sub-images taken at half the focal ratio with the same detector. It is this example which is illustrated in the simulation shown in figure 7.4.

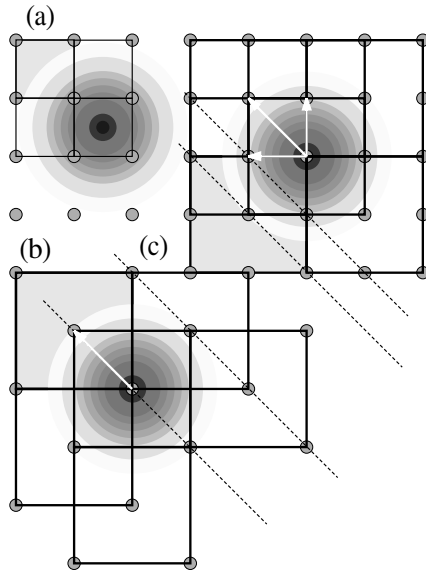


Figure 7.4. Three sampling/pixel size scenarios appropriate to the examples discussed in the appendix. The schematic stellar image is scaled to a 22 mas grid spacing in the V-band. (a) small (22 mas) pixels, no interleaving. (b) 44 mas pixels, 2 sub-steps displaced along a diagonal. (c) Four-fold interleaving of 44 mas pixels by one grid-spacing.

7.2.2 PHOTON COUNTING DETECTOR

The photon counting detectors do not generally suffer a readout noise and so a fine image pixelation does not carry a signal/noise penalty with it. The most severe constraint comes from the total number of pixels since the array has to be read faster than the photon arrival rate. The details of this saturation depend on the particular technology and can result in either a limited rate per pixel or a limited rate integrated over the whole array. In the former case, there is clearly an advantage in a high magnification since this will reduce the count rate per pixel for a given source. This means that the best performance in terms of dynamic range, astrometric precision and image resolution can be obtained with an array which critically samples, or even oversamples the image.

7.3 A WIDE-FIELD CHANNEL

A wide-field channel working in the optical/near infrared bands will almost certainly use a CCD detector. The noise characteristic of such a device—dark current, susceptibility to cosmic rays, readout noise—will determine the optimum strategy for a particular kind of observation. The range of application for this channel will, of course, be very wide—from spatial resolution of bright, complex structures to the detection of

very faint extended objects against the sky background. Therein lies the difficulty in choosing design parameters. The goal will be a solution which is flexible in allowing both of these extremes while maintaining the widest possible field of view.

High astrometric precision and the resolution of structure at high s/n demands critical sampling. Photometric performance, we will show, is better—at least for faint sources—with pixels which themselves undersample the image. The limit in this case is likely to be the knowledge of the pixel-scale variations in the flat-field response. For a given detector pixel grid the spatial sampling can be improved by combining multiple ‘interlaced’ (or ‘interleaved’ or ‘sub-stepped’) frames (Adorf 1989, 1990). The method of interlaced frames is very flexible since it allows different sampling geometries to be realized simply by observational design. Doublets of frames, shifted by half a pixel-diagonal with respect to each other, increase the spatial sampling by a factor of 2 in both the principal detector directions. Quartets of subframes can be combined either in a square grid or, perhaps more interestingly, in a hexagonal grid resulting in an improved spatial sampling of the isotropic optical PSF generated by a circular aperture. The scenario described in the appendix considers the case of broad-band V imaging of point sources using a total exposure time of one orbit (2,400 s). It compares the photometric and astrometric precision which can be obtained in three different modes:

1. a single exposure with critically sampling pixels (figure 7.4a)
2. a single exposure with pixels which two-times undersample
3. four $1/2$ pixel interleaved exposures with the larger pixels—identical *sampling* to case 1 (figure 7.4b).

Cases 2 and 3 result, of course, in an image field which is four times as large as that in case 1.

Calculations of the minimum-variance (Cramer-Rao) bound for these cases (Adorf private communication, see also King 1983) show the variation of the photometric and astrometric information content in the observation of a point source with sky and detector background in these three and other cases. The realization of this lower bound in the data analysis demands an appropriate combination of the interleaved sub-images. With the simple interleaving of the sub-images onto a finer grid, the resolution is, of course, determined by the convolution of the optical PSF with the large pixel response function. A conservative application of

the iterative co-addition technique described by Lucy & Hook (1992) allows the combined data to be presented with an effective PSF equal to that of the optical PSF alone. This is what is shown in the simulations presented in figure 7.5a.

A simplified analytical view of the effect of increasing the pixel size while keeping the sampling interval constant is shown in figure 7.6 where the information content for photometry and astrometry is plotted separately as a function of pixel size. This calculation is idealized in the sense that it assumes a 'top hat' pixel response function. Another calculation (not shown) with a Gaussian pixel response function (PRF) shows a similar behavior but with the astrometric information dropping somewhat more rapidly. This is a result of the PRF acting as a low-pass filter with a severity which scales with pixel size. A good knowledge of the real PRF of the detector will be necessary for choosing an optimum observation design.

Interleaved sampling, while offering advantages in resolution for a given FOV, does not solve the data volume problem. In particular,

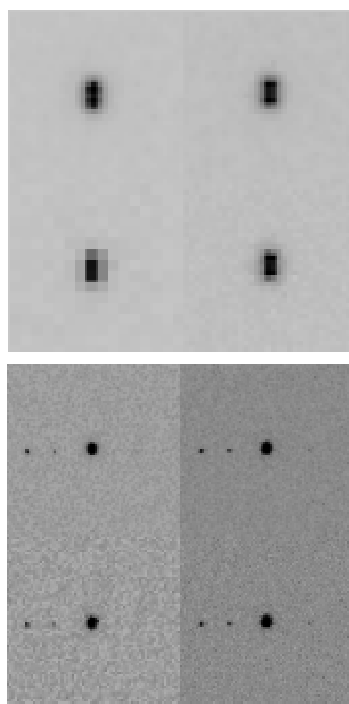
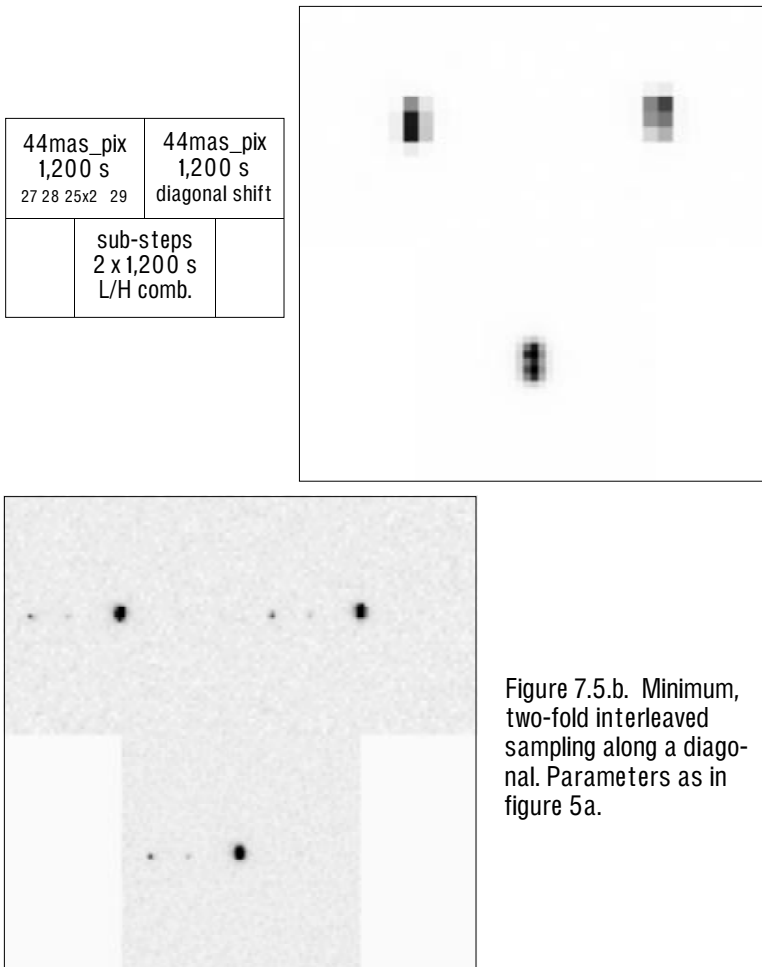


Figure 7.5.a. An illustration of the effect of 4-fold sub-stepping compared with a single one-orbit exposure with half the pixel size. The four frames show unprocessed simulations on the bottom row with a Lucy-Hook combination of four sub-stepped frames on the top left. A 'null' L/H version of the single exposure is at the top right. The bright object in the center, shown enlarged above, is a pair of 25th magnitude stars separated by the Rayleigh limit. The magnitudes of the other, fainter, stars are indicated in the legend.

sub-steps 4 x 600 s L/H comb.	2,400s L/H
43 mas_pix 600 s 27 28 25x2 29	22 mas_pix 2,400 s

when making the above comparisons, we have ignored the question of the operational overhead of multiple exposures. With the long readout times from large array CCDs, it will often be a requirement to minimize the number of exposures. Given that multiple exposures will be necessary for cosmic-ray discrimination, the minimalist interleaving strategy of just two exposures displaced by $1/2$ a pixel in each coordinate still offers significant resolution advantages over two identical exposures (figures 7.4b and 7.5b). An image processing algorithm can be envisaged which, by using a precise knowledge of the displacement, could use just two frames for simultaneous resolution enhancement and cosmic ray event cleaning.



This discussion pre-supposes that the optical performance of the camera, working at the faster focal-ratio implied by the undersampling pixel array, is capable of providing the image quality necessary to allow the achievement of the benefits offered by sub-stepping. This may prove

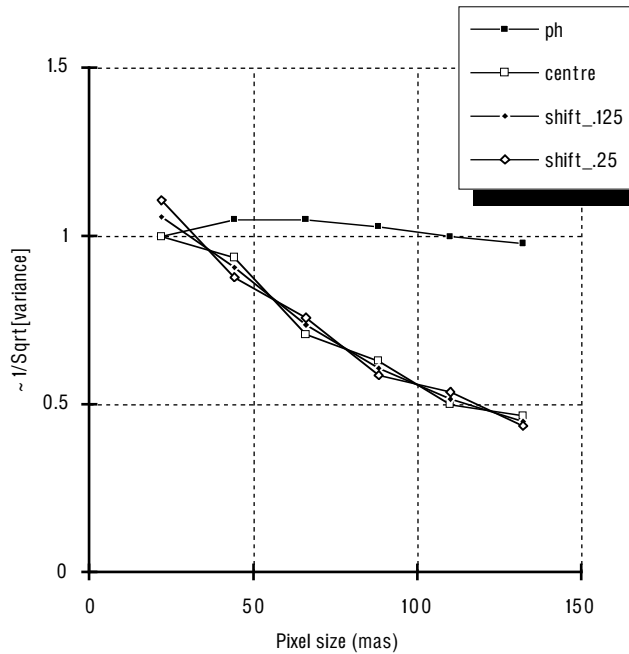


Figure 7.6. The photometric (ph) and astrometric information content of a stellar image (29th magnitude in V) calculated using the minimum-variance bound (see the appendix). The image sampling is 22mas for all points but the pixel size is varied as shown. The ordinate is the normalized Fisher information ($\sim 1/\text{Sqrt}[\text{variance}]$, i.e. proportional to sigma). The three astrometric curves are for different positions of the peak of the star image on the sampling grid. From these analytic calculations, reinforced by the qualitative impression gained from the simulations, we can conclude that, for the astrometric and photometric measurements of faint stars, sub-stepping with a pixel array which undersamples by a factor of about two is a very effective compromise.

impractical and drive the design towards physically larger detectors (pixels) where a careful watch on the dark current performance is necessary.

7.4 A HIGH RESOLUTION CHANNEL

The choice of sampling for such a channel will depend on the choice

of detector. For photon counting devices, as we have already pointed out, there are advantages in having pixels which critically or super-sample the PSF. For integrating detectors on the other hand, it is clear that the pixel readout will be the dominant noise source for faint objects since such a device will not see the sky background even in broad-band filters. Critical sampling in this case can only be recommended to satisfy scientific requirements other than reaching the faintest objects. Apart from straightforward high-resolution studies of bright objects, the application of coronagraphic techniques and/or the use of apodizers may demand that the image is sampled directly with a fine grid. It is clear that telescope repointing cannot be used for sub-stepping the image in this case although an internal tip-tilt mirror system could, possibly, be devised.

7.5 CALIBRATION AND IMAGE PROCESSING REQUIREMENTS

The existence and maintenance of good calibration is essential for high quality data analysis. This rather obvious statement includes the need for a detailed (stochastic) model of both the deterministic and noise characteristics of the instrument. Optimum estimators can then be constructed for a particular measurement task. Such a model has been derived for the WFPC and is described by Snyder, Hammoud & White (1992).

In practice, with a high dynamic range detector, uncertainties in flat-fielding and knowledge of the PSF can become the limiting factor in the accuracy of measurement.

7.6 THE PSF

Photometric and astrometric experience with ground-based observations has generally been based on simple ‘unimodal’ PSFs with a nearly Gaussian shape (e.g. King-profile, Moffat-profile etc.) produced by atmospheric distortions. These PSFs, characterized by a sharp central peak and featureless wings, are usually sufficiently well sampled that knowledge of the exact shape of the pixel and its sensitivity response function is not important. In particular, there is practically no positional information in the outer PSF-wings.

The nearly diffraction-limited PSF of a space telescope has quite different characteristics. Apart from the central peak there are diffraction rings, into which the flux of the wings is concentrated. Thus more photometric/flux information is contained in the wings of a diffraction-limited PSF than in that of a structureless PSF. For a centrally obscured ap-

erture like HST, more than a quarter of the photometric, and a significant fraction of the astrometric, information is carried outside the PSF core.

When the PSF is near-critically sampled the exact shape and response function of the pixels and the geometry of the sampling grid become important. The available astrometric information resides in the slopes of the effective PSF. Assuming an ideal (hat-shaped) PRF, appropriate for a thinned CCD-detector, each pixel contributes astrometric information (in the appendix we use the 'Fisher' information from statistics—the inverse of the variance) proportional to the square of the difference of the optical PSF value at its boundaries (Adorf, priv. comm.). Thus two pixels contribute relatively more astrometric information when a mode of the PSF coincides with their mutual boundaries, compared with the situation when the mode coincides with one of the pixel centres.

Thus the rings of a diffraction-limited PSF contribute a significant fraction of the overall astrometric information, particularly since their modes cross pixel boundaries even when the core is centered on a pixel. For bright objects the diffraction spikes due to the spiders in the optical path may hold considerable additional spatial information, particularly when their modes coincide with pixel boundaries. Only a full 2D-calculation of the relevant minimum variance bound can clarify the quantitative aspects of these considerations.

The spherical aberration corrected OTA will deliver images which are close to being diffraction limited in the optical. In the UV, the PSF will contain more structure with some fixed 'speckles' around a sharp core. The image restoration techniques, whose accelerated development was necessary to deal with spherical aberration, will remain fully applicable to the corrected images. This implies that a high precision knowledge of the PSF will be essential.

Currently, HST PSFs are being derived from a combination of measurement and modelling. For the treatment of high dynamic range data, observed PSFs of sufficient quality are very hard to obtain and even the best computed functions are inadequate. In the case of the WFPC, the situation is exacerbated by the strong spatial dependence introduced by asymmetric pupil shadowing from the Cassegrain repeater secondary mirror.

The lesson we learn from this is that the PSF of a new camera needs to be precisely known from measurement or accurately computable for

the appropriate conditions or, preferably, both. The modelling will always be necessary at some level, even if only to provide interpolating functions, since the PSF needs to be known at finer sampling than the data and the color dependence will produce object dependent results for broad filter observations.

The spatial variation of the PSF introduces considerable computational complexity into the restoration task and the best solution would be an optical design which minimized it. If spatial variation is unavoidable, it is important that we are able to compute it accurately.

Spatial distortions in the pixel array, unless well calibrated and treated in a sophisticated way, can cause loss of resolution and undesirable noise correlations as a result of conventional resampling algorithms. Such an effect is seen clearly in geometrically corrected images from the FOC which has substantial geometric distortion arising within its electron optical components.

With the high optical magnifications needed to sample the PSF in the far-UV, it is possible to run into difficulties with 'edge-effects'. The sidelobes of the Airy function spread a long way from the core and restoring high dynamic range data can be a problem if the detector has only a small area. This problem is seen with FOC $f/96$ data at the moment with the very wide aberrated PSF.

For the ultimate resolution on high dynamic range objects to be achieved using techniques such as 'roll-deconvolution', Mueller and Weigelt (1987) recommend pixel sizes as small as 4 to 6 mas in the far-UV. Similar results could be achieved by an active control of the PSF, e.g. by small focus changes, which are similarly able to move the zeroes in the Modulation Transfer Function (MTF) of the camera. Since changing the OTA focus as part of an observing program is undesirable and unlikely to be allowed by the project, the ability to change the camera focus internally on a regular and highly repeatable basis could be considered important. For well-sampled images, the limits to the resolution which can be obtained using non-linear restoration algorithms have been derived as a function of signal-to-noise ratio by Lucy (1992a,b).

7.7 POINTING AND TRACKING

The projected tracking performance of the HST is discussed in chapter 4, and we make a few remarks here which are pertinent to image processing.

An rms jitter of 3 to 4 mas will not have a profound effect on resolution, even in the far-UV. Since it is not clear, however, that this jitter will have a Gaussian distribution and be circularly symmetric, it is important that the data carry with them the record of the actual telescope pointing during an observation. This can be used in the construction of the appropriate PSF.

If undersampled frames are to be obtained and processed, the ability to offset the telescope by non-integral pixel steps can be used, with subsequent combination, to recover resolution. While this technique can be used for random, non identical, pointings (Adorf 1992), the ability to specify and control sub-pixel pointings is desirable. The ability to do this quickly and repeatedly without incurring the overhead of telescope re-pointing, could be important for operational efficiency.

7.8 CONCLUSIONS

1. The pixel size of the detector and the frequency of the grid used to sample the image should be considered separately since their relationship to one another depends on the detector noise characteristics.

2. For high signal/noise observations, the ability to critically *sample* the image is essential in order to recover and, in some cases, exceed the resolution of the optical system.

3. Given the very wide wavelength coverage available to the camera, this critical sampling requirement, coupled with the desire for the widest FOV, will necessitate more than one optical magnification or channel.

4. To achieve the highest surface brightness sensitivity in the presence of detector readout noise, it will be necessary to use relatively large pixels. This can be achieved, in the case of a CCD, by 'noiseless' on-chip binning techniques.

5. Sub-critical sampling may be necessary in some cases to alleviate data volume and telemetry constraints and to achieve fast exposure repeats.

6. The use of 'interleaved sampling' (sub-stepping) strategies offers the prospect of obtaining a large FOV with available CCD detector arrays while maintaining good astrometric precision and image resolution. The effective and efficient use of these techniques may require that the camera have a precise internal pointing/guiding system for producing sub-pixel offsets. Very effective sub-frame combination algorithms already exist and can be optimized for this application.

7. The resolution of an image is inherent in the sampling grid, The increase in pixel size above the grid spacing by up to a factor of about two produces only a slow loss of astrometric/resolution information.

8. A ‘minimalist’ sub-stepping strategy, using just two exposures, should be developed for simultaneous resolution recovery and cosmic ray detection and removal.

9. Image restoration and combination techniques require an accurate model of the deterministic and the noise characteristics of the instrument. In particular, the PSF needs to be precisely measured and/or accurately computable. A precise knowledge of the individual pixel response function of the detector(s) will be necessary for devising appropriate sub-stepping strategies.

10. A high-resolution channel, which may provide coronagraphic and/or apodizing capabilities, will probably require a detector array which critically or supercritically samples the image directly.

11. For very high precision observations, active control of the PSF (i.e., internal focusing) is desirable.

On this basis, we recommend:

1. The ability to critically (Nyquist) *sample* the image is required at all wavelengths.

2. In the wide-field mode, it may be neither feasible nor desirable to achieve critical sampling with the pixel array in a single exposure.

ACKNOWLEDGEMENTS

Robert Fosbury was the lead author of chapter 7. Hans-Martin Adorf, Richard Hook, Leon Lucy, Wolfram Freudling, and Adeline Caulet contributed.

REFERENCES

- Adorf, H.-M., 1989, “On the HST Wide Field and Planetary Camera Undersampling Problem”, in: *1st ESO/ST-ECF Data Analysis Workshop, Apr 1989*, Garching b. München, FRG, P. Grosbøl, R.H. Warmels, and F. Murtagh (eds.), European Southern Observatory, Garching b. München, FRG, pp. 215–218.
- Adorf, H.-M., 1990, “WFPC: Options for Overcoming Undersampling”, in: *Workshop “Errors, Bias, and Uncertainties in Astronomy”, Sep 1989*, Strasbourg, France, C. Jaschek and F. Murtagh (eds.), Cambridge University Press, Cambridge, UK, pp. 71–78.

- Adorf, H.-M., 1992, "HST image restoration—recent developments", in: *Science with the Hubble Space Telescope*, Chia Laguna, Sardinia, 29 Jun–7 Jul 1992, P. Benvenuti and E. Schreier (ed.), European Southern Observatory, Garching b. München, FRG.
- Cheung, K.F., Marks, R.J., II: 1990, "Imaging sampling below the Nyquist density without aliasing", *J. Opt. Soc. America A* 7, No. 1, 92.
- Ghiglia, D.C.: 1984, "Space-invariant deblurring given N independently blurred images of a common object", *J. Opt. Soc. America A* 1, No. 4, 398–402.
- Hildebrant, J.W., Newsam, G.N.: 1990, "Subpixel Resolution by Image Registration", *IEEE ISSPA*, 551–518.
- Irani, M., Peleg, S.: 1991, "Improving resolution by image registration", *CVGIP: Graph. Models Image Proc.* 53, No. 3, 231–239.
- Kim, S.P., Bose, N.K., Valenzuela, H.M.: 1990, "Recursive Reconstruction of High Resolution Image From Noisy Undersampled Multiframe", *IEEE Trans. Acoust. Speech Signal Proc.* ASSP-38, No. 6, 1013–1027.
- King, I. J., 1983. "Accuracy of Measurement of Star Images on a Pixel Array". *PASP*, 95, 163.
- Lucy, L.B., 1992a. "Resolution limits for deconvolved images". *Astronomical Journal*, in press.
- Lucy, L.B., 1992b. "Statistical limits to superresolution". *Astronomy & Astrophysics*, in press.
- Lucy, L.B. & Hook, R.N., 1992. In: *Proceedings of the 1st Annual Conference on Astronomical Data Analysis and Systems*, Tucson, November 1991, eds. Worrall et al., ASP Conference Series, 25, 277.
- Mueller, M. & Weigelt, G., 1987, *Astron. Astrophys.* 175, 312.
- Namura, Y., Naruse, H.: 1988, "Reduction of Obscuration Noise Using Multiple Images", *Pattern Analysis and Machine Intelligence* 10, 267.
- Snyder, D.L., Hammoud, A.M. & White, R.L., 1993. "Image recovery from data acquired with a charge-coupled-device camera". *J. Opt. Soc. America A*, 10, p. 1014.
- Tsai, R.Y., Huang, T.S.: 1984, "Multiframe image restoration and registration", in: *Advances in Computer Vision and Image Processing*, JAI Press Inc., pp. 317–339.
- Ur, H., Gross, D.: 1992, "Improved resolution from subpixel shifted pictures", *CVGIP: Graph. Models Image Proc.* 54, No. 2, 181–186.

APPENDIX

A) Simulated examples referred to in the text

This appendix outlines a quantitative comparison which has been

performed of imaging performance at different focal ratios for a CCD detector operating in the WFPC V-band. The CCD parameters are thought to be representative of current state-of-the-art devices: read noise: $3 e^-$ rms; dark current: $4.4 \times 10^{-5} e^-/\mu\text{m}^2/\text{s}$; pixel size: $10 \mu\text{m}$; sky brightness: $24.2 \text{ mag}/\text{arcsec}^2$; total (detector+optics) efficiency: 45%. These parameters are used in the simulations illustrated in figure 7.5a and b.

The comparison between a single one-orbit exposure at $f/40$ and $4 \times 600 \text{ s}$ sub-stepped exposures at $f/20$ illustrates the effect that a significant dark current can have on exposures with a large f -ratio—because a larger physical area of silicon is used to cover a resolution element. This may not be a serious issue if the dark current can be kept to a sufficiently low level. Narrow band observations will, of course, be more affected by the detector noise sources and so, for photometry, larger pixels are an advantage. The practical limit to increasing pixel size will come from considerations of astrometric precision, even where sub-stepping is practised. For effective astrometry, the PSF gradient must be sampled by pixel boundaries.

For the final design of the camera, it will be important to know the spatial response function of a pixel in order that the 'effective PSF'—the convolution of the optical PSF with the pixel response function—can be used in modelling the process of image formation.

In the simple model used for the examples, the photometric performance has been calculated using a King profile (King 1983) with a scale factor appropriate to the V-band. 'Aperture' photometry covering 60% of the flux in the PSF is used in calculating the variance and the signal-to-noise ratio. The readout noise has, for simplicity, been included in the variance calculation as a Poisson rather than a Gaussian random variable.

B) The Minimum-variance (Cramer-Rao) bound: astrometry with sub-stepping

Many considerations including sky-coverage and flexibility speak in favour of 'large' pixels on an array that insufficiently samples the effective point-spread function. Also, depending on detector characteristics, the photometric precision obtainable from four combined interlaced frames with 'big' pixels may be better than that obtainable from a single frame with 'small' pixels exposed four times as long. The only considerations

speaking in favour of small pixels relate to astrometry and resolution.

Calculations of the astrometric minimum variance bounds of an unbiased estimator for the flux and position of a point source (see figure 7.6) have been carried out under the following assumptions:

1. The PSF has a King profile.
2. The spatial sampling is critical (Nyquist).
3. The pixel size (i.e., side length) is varied from the sampling distance (22 μm) upwards.
4. The same total exposure time is divided amongst sub-steps at all points.
5. The quantities for the source and background counts are taken from the example in Appendix A.

C) Combination of undersampled multiframe—bibliography

The method of combining several individually undersampled interleaved ('dithered') image frames into a common high resolution image—with or without an internal image restoration step—has been investigated in recent years (Ghiglia 1984; Tsai & Hunag 1984; Namura & Naruse 1988; Cheung & Marks 1990; Hildebrandt & Newsam 1990; Kim, Bose & Valenzuela 1990; Irani & Peleg 1991; Ur & Gross 1992).

8. Detectors

History has shown detectors are a critical and problematic aspect of developing astronomical instrumentation. We have examined the issue of detectors for the Advanced Camera (AC): What detectors are appropriate? What is their current state of readiness? How can we assure that the proper detectors will be available?

We find that for wavelengths 3000 to 10,000 Å, current charge-coupled device (CCD) technology can produce nearly ideal, fully-featured detectors for the AC. However, immediate investment is required to assure the availability of these CCDs.

For wavelengths 1000 to 2000 Å, no satisfactory detector/filter combination for the AC is currently assured. To develop one possible solution, technology development is required in anti-reflection (AR) coatings and backside passivation to enhance the ultraviolet (UV) performance of CCDs in the 1000 to 3000 Å range. Coupled with this, reliable Wood's filters are needed to render the CCDs blind to red light in UV applications. To assure a viable alternative to UV-sensitive CCDs, further development is required of large-format micro-channel plates (MCPs) and appropriate readout technology for photocathode detectors in the wavelength range 1000 to 2000 Å.

The scientific goals of the Hubble Space Telescope (HST) Advanced Camera (AC) translate into detector performance requirements such as pixel size and format, quantum efficiency (QE), resolution, noise level, photometric stability, controller electronics, and data capacity for the instrument computer. The candidate science programs for the AC—which call for detecting light over the decade of wavelengths from 1000 to 10,000 Å—will require at most two detector technologies. These technologies are charge-coupled devices (CCDs) and photocathode-based, photon-counting detectors. In this chapter we discuss the current status of these detector technologies and present our assessment of the feasibility of developing them for the AC in the time frame required.

8.1 CHARGE-COUPLED DEVICES (CCDs)

CCDs are the imaging detector of choice for nearly all scientific investigations in the near ultraviolet/optical (NUVO) spectral region.¹ CCDs are thin silicon structures that are divided into an array of pixels by electronic circuits on the 'frontside' surface. When a quantum of light (a photon) enters the silicon and is absorbed, it produces a photoelectron, which persists until transferred to an amplifier at the edge of the CCD array by the controlling electronics. As the charge is transferred, the amplifier output signal is proportional to the accumulated photoelectrons in each pixel of the CCD array transferred in turn according to a programmed clocking sequence. Finally, a computer receiving the digitized amplifier output organizes the time record of the amplifier output into an array of numbers that are proportional to the number photons absorbed in each pixel of the array, and produces an image.

They have been used on many space missions, including the Galileo spacecraft bound for Jupiter and the Wide Field Planetary Camera (WFPC) launched with HST. The CCD's wide dynamic range, high sensitivity, linearity, stability, and low noise has made it an excellent choice for space applications. The major problems associated with CCDs are the great difficulty of achieving high sensitivity in the ultraviolet (UV), especially in the 1000 to 3000 Å spectral region, their intrinsic read noise, and their high sensitivity to visible/near infrared (IR) light when used for UV observations.

Most scientific quality CCDs made today have similar characteristics, and their differences are due mainly to different approaches to optimization on the part of the manufacturers. Photoelectrons are transferred to the output amplifiers with 0.999999 efficiency. Formats of 2048 x 2048 pixels are standard and 4096 x 4096 is experimental at the current time. The maximum number of photoelectrons per pixel ('full well capacity') is proportional to pixel area. For 7.5 x 7.5 μm pixels, the full well capacity is about 25,000 electrons (e⁻). The photometric non-linearity is less than 1%, and the spatial response is uniform to within 5%. The amplifier readout noise is 2 to 4 e⁻ at operating temperatures -60 to -100° C.

8.1.1 FRONT ILLUMINATED CCDS

Normal, off-the-shelf CCDs are used in the front illuminated mode. They have inferior quantum efficiency (QE) in the visible (<50%) due

to reflection and absorption in the frontside electronic structures. Because light must pass through these structures that define the pixels (composed mainly of polysilicon), most radiation of $\lambda < 4000 \text{ \AA}$ is absorbed or reflected before detection. Since the frontside is usually passivated with a protective glass coating, anti-reflection (AR) coatings cannot be applied to reduce the significant reflection loss. The refractive index of the passivation is about 1.5, whereas that for silicon is as high as seven in the near-UV. Therefore, an AR coating on the glass will improve QE by only a few percent, and substantial losses will remain at the inaccessible interfaces between the passivation glass and the polysilicon gates, and between the gates and the underlying gate oxides.

New front illuminated CCD concepts exist that could reduce these UV problems, as discussed below.

Phosphors

By coating the front side of a CCD with a phosphor, which absorbs UV photons and re-emits them in the visible, it is possible to obtain 20% QE in the UV with a peak response of 50% in the visible. Such detectors will be used in the HST second-generation Wide Field Planetary Camera (WFPC II). The major problems with phosphors are the possible degradation in UV efficiency due to coating changes, the risk of condensing contaminants on the cold windows, and the low overall UV response.

Thin polysilicon

It is also possible to make CCDs with thinner polysilicon than is traditionally used to decrease the absorption in the frontside structures. In this case, the UV response with a phosphor overcoat is expected to be 25%, with a peak response in the visible of 65%.

Open-phase devices

It is also possible to produce devices with “exposed” silicon on the frontside by adjusting the relative sizes of the pixel phases. In this case, bare silicon can be used for some fraction of the detecting area, and the benefits of an AR coating can be realized. Such an open-phase CCD will also not suffer from frontside UV absorption over that fraction of a pixel which is bare silicon.

8.1.2 BACK-ILLUMINATED CCDs

The most efficient method of making UV sensitive CCDs is to thin the devices to a total thickness of less than 20 μm and illuminate the device from the back side. In this case the light passes directly into the silicon for efficient detection. Although these devices are difficult and expensive to develop as compared with front-illuminated detectors, they are critically important for efficient UV observations.

Since the surface of back-illuminated CCDs is silicon, an AR coating is required, which can increase QE to nearly 100%. Careful control of the surface properties can allow radiation sensitivity from X-rays to the near-IR. Currently, Tektronix is the only domestic commercial supplier of such devices, although Lincoln Labs at the Massachusetts Institute of Technology (MIT) and Reticon have made thinned devices, and Steward Observatory at the University of Arizona is thinning devices from various manufacturers.

Thinning and packaging

A concern with using existing thinned CCDs for astronomy is mounting the devices flat in the final package. For example, the Texas Instruments 800 x 800 CCDs used in the HST WFPC have peak-to-valley non-flatness deviations up to 100 μm , which results in defocus dependent on field position and therefore loss of resolution. Thinned CCDs must be mounted in a manner that assures flatness and mechanically stability.

Backside charging

Thinning CCDs leaves a large backside potential well, which is formed by positive charge. Without corrective measures, this potential well lowers the QE of the thinned CCD below even that of a front-illuminated device. The process of eliminating this effect is referred to as 'backside charging'.¹² For the AC CCDs we expect backside charging will be accomplished either by 'flash gate', 'bias gate', or ion implantation. A flash gate³ employs a very thin layer of a metal with a larger work function than silicon deposited on p-type silicon, which causes electrons to flow from the silicon to metal and produce a backside charge. A bias gate employs a thin metal layer separated from the silicon by a deposited insulating layer, effectively creating a capacitor; applying a potential to the metal then induces a charge in the silicon. Ion implantation changes

the doping profile in the silicon near the backside to create an internal electric field.

Anti-reflection (AR) coatings

A large QE gain in back-illuminated CCDs can be achieved by overcoming the reflection losses due to the high index of refraction of silicon.⁴ The losses approach 60% in the ultraviolet. Reducing this loss produces a secondary benefit in reducing the ‘ghost’ images often formed when the silicon-reflected light is returned back to the detector after reflection from a dewar window or other optical element. The application of a thin film AR coating directly onto the CCD back surface can effectively solve the problem.

AR coating produces the greatest QE gain in the blue and near ultraviolet, but significant increases are expected at all wavelengths. In fact, since AR coatings can be tailored for particular observations, a 50% to 100% QE enhancement can be realized at any specified wavelength, as illustrated in figure 8.1.

Although the back surface of a thinned CCD is well suited to thin film deposition, careful consideration must be given to selecting appropriate materials and application methods. Depending on the refractive index and absorption coefficient, some materials perform better in the red or blue than others. Additional research is needed to optimize AR coatings for thinned CCDs at wavelengths below 3000 Å.

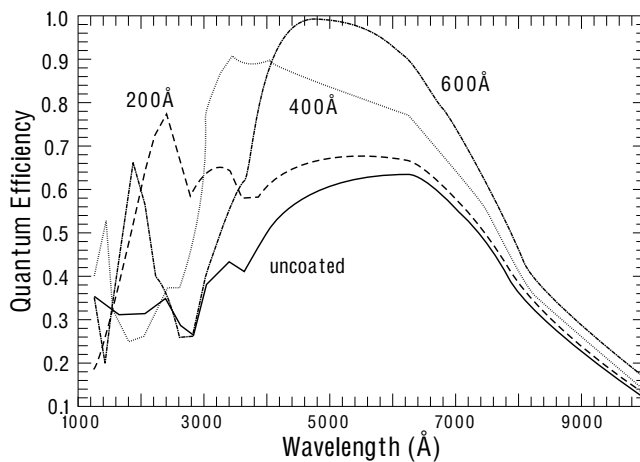


Figure 8.1. Predicted QE of a thinned CCD with various thicknesses of HfO_2 deposited as an AR coating.

8.1.3 CONTROLLER REQUIREMENTS

Major gains have been made in recent years in the performance and operational flexibility of CCDs by the addition of electrical circuits that make the CCD highly programmable. Accessed through the instrument computer (chapter 10), these controller capabilities can be expected to render the AC CCDs robust, efficient, and far more capable than the CCDs currently installed on HST.

On-chip binning

Binning allows multiple pixels to be summed together before readout in a noiseless manner. This can significantly reduce readout time when full spatial resolution is not needed.

Selectable sub-array readout

Sub-array readout lets the observer specify a subsection of the CCD pixels to be readout, disregarding other data on the detector. This is especially useful when data on a small object is desired.

Partial inversion/multi-pinned phase (MPP)

MPP operation allows reduced dark current at a given temperature of operation, at the cost of reduced full well capacity.

Anti-blooming

Anti-blooming mode keeps bright objects in the imaging area from saturating the entire image by blooming up and down columns when the full well capacity is exceeded.

High speed erasure

High speed erasure allows flush the entire CCD in just seconds before each exposure.

Skipper clocking

If the CCD has a skipper amplifier, it can be multiply sampled, to allow N multiply reads of each pixel value, decreasing the effective read noise by \sqrt{N} .

Selectable gain

Depending on the observing requirements, different observers may desire different gains to improve dynamic range.

Commandable voltage settings

Commandable voltage settings would allow CCD clock voltages to be changed in flight for device optimization.

Programmable clocking sequences

Programmable clocking would also allow the clocking wave forms to be changed in-flight to allow device optimization and help guard against various failure modes.

8.1.4 MOSAICS

CCD mosaics have several possible advantages for the AC. They allow completely redundant detectors in case of device failure, an effectively large image area can be obtained with any desired pixel size, and they allow many amplifiers to be used for simultaneous readout for higher effective data rates. Several groups have now produced CCD mosaics of up to four 2048×2048 pixel CCDs. If mosaics are required in the AC, they will be straightforward to design and produce.

8.1.5 RED LEAK

Most astronomical objects of interest emit 10^4 to 10^8 visible photons for every far-UV (FUV) photon in the 1000 to 2000 Å wavelength region. Since typical FUV filters have visible transmissions of the order of 10^{-3} to 10^{-4} , FUV images would normally be hopelessly contaminated by red-leak photons.

A limited amount of research has gone into making Wood's filters, which use alkali metal films to block red light at frequencies below the plasma frequency of the free electrons in the metal. Wood's filters reject red light by factors about 10^{-8} . In comparison, so-called 'solar-blind' CsI photocathodes have visible light quantum efficiency of the order of 10^{-6} , which when combined with standard FUV filters provides visible light rejection better than or comparable to the Wood's filter. The combination of relatively low throughput and concerns regarding their longterm stability means that further development is probably required for Wood's filters to compete with photocathode-based detectors for UV observations.

8.1.6 RESEARCH REQUIRED FOR AC CCDS

The basic manufacturing processes required for CCD technology

are relatively standard for the semiconductor industry and will hopefully continue to be available at fabrication facilities. The areas of CCD expertise that the scientific community has been in danger of losing are those of very low noise CCD amplifier design, techniques to obtain excellent low-level charge transfer efficiency, thinning technology to produce back-illuminated devices, coating technology for AR coatings and backside charging, and packaging techniques for flat imaging surfaces and large area mosaics. In addition, space missions require UV transmitting, visible/red light rejection filters. These technologies are unprofitable for industry but are required for NASA instrumentation like the AC. We strongly recommend that NASA invest in these areas of technology development in order to assure that the appropriate CCD detectors be available for the AC.

8.2 PHOTOCATHODE-BASED DETECTORS

Photocathode-based detectors provide excellent imaging capabilities in the FUV spectral region. If available, they are preferred over CCDs because of their noiseless photon counting ability, relatively high QE, and lack of sensitivity to long-wavelength light.

8.2.1 PHOTOCATHODES

Photocathodes with high sensitivity in the FUV fall into two categories, those with spectral response that extends into the optical, and those that have a cutoff in the UV and are “solar blind” (see Coleman⁵). Examples of FUV photocathodes may be found in the current HST instrument complement. The Faint Object Camera (FOC) and Faint Object Spectrograph (FOS) have bi-alkali photocathodes with a spectral response that extends from the FUV through the optical. The Goddard High Resolution Spectrograph (GHRS) has solar blind CsTe and CsI photocathodes. CsI is the better solar blind photocathode, since CsTe has a red leak beyond its nominal long wavelength cutoff at 2500 Å, extending to 8000 Å.

8.2.2 GAIN STAGES

Photocathode-based detectors suitable for high spatial resolution require a gain stage to amplify the photoelectron to a signal level suitable for photon counting, followed by a position encoding anode. There are three classes of photocathode-based UV detector, distinguished prima-

rily by one of the following gain mechanisms: multistage intensification, micro-channel plates (MCPs) and direct electron bombardment. For each class of detector there are a number of possible position-encoding methods that can be employed.

Multistage intensifiers

Multistage intensifiers have been in use in space and ground-based astronomy for the last 20 years.^{6,7,8} The most obvious example is the FOC, a multi-stage intensifier optically coupled to a TV-based event location system.⁸ The main drawbacks of these systems are their limited dynamic range, power consumption and the size and complexity of the intensifier assemblies.

Micro-channel plates

An MCP is a compact array of channel electron multipliers that can provide high electron gains (10^5 to 10^6 e^-) and high spatial resolution (typical channel size is 12 μm .) The CsI photocathode can be deposited directly onto the front surface of the MCP. In this configuration, the photoelectron is directed into a channel opening by a potential gradient due to a repelling electrode located in front of the MCP. The position encoding device, positioned close to the other side of the MCP ('proximity focused'), receives the charge cloud, which has been amplified by passing down a channel.

The dynamic range of an MCP is dependent upon the extracted charge per photoelectron event. Higher detector operating gains result in an overall decrease in dynamic range. In practice, most position encoding schemes for MCP-based detectors cannot process more than 10^5 to 10^6 cps and so the global dynamic range of a given MCP configuration is seldom an issue. Local dynamic range, over a small number of MCP channels, drops when more charge is extracted from a channel than can be replenished before the next event.

Optimum local dynamic range generally comes from minimizing the number of MCPs and their operating gain, combined with a choice of low resistivity MCP. Typical local dynamic range limits for a stacked MCPs are of the order of 5 to 10 counts/second/pixel (Siegmond *et al.*⁹). The multi-anode microchannel plate array (MAMA), which employs a single curved-MCP, exhibits a local dynamic range in excess of 50 counts/second/pixel (25 μm pixel). Recently, low resistance MCPs,

known as “hot” plates, have allowed dynamic range to be extended. Hot-plates are made of low resistance glass, which results in a higher MCP wall strip-current. Siegmund¹⁰ has demonstrated a local dynamic range of 20 counts/second/pore (i.e. 80 counts/second/pixel for a 25 μm pixel, modal gain 1×10^7). An concern with low resistance plates is that thermal runaway is possible unless the plate is current limited. An alternative method of increasing MCP local dynamic range is to decrease pore size so that more MCP channels fall within the boundary of a detector pixel.

Curved channel MCP's are able to provide high electron gains ($\sim 10^6$) in a single plate configuration. The curved channels minimize ion feedback—ions returning up the MCP channels. Two technologies exist for producing the curvature in large MCPs that the D₂ detector will require. It can be done by heating the MCP while spinning it at high speed. Or, the plate can be cut at a bias, held by a fixture at the bottom, then, while heated uniformly from the top, the upper half is pushed under computer control to produce the desired ‘C’ curve. The challenge in manufacturing the larger MCP has been process control: obtaining the proper shear time and temperature to produce the optimum curvature without introducing mechanical stress or unwanted distortions.

8.2.3 MCP READOUT

The most promising MCP readout technologies are those that combine high spatial resolution with low MCP gain requirements, high photon-counting dynamic range, geometric and spatial linearity and the potential for expansion to large formats. The spatial resolution is determined by a combination of factors: MCP pore size, quality of the proximity focus, and the sampling of the readout array. A variety of technologies exist, here we highlight the MAMA, the wedge and strip anode, the delay line system, and the intensified CCD.

Multi-anode micro-channel plate array (MAMA)

Timothy¹¹ has reviewed the MAMA which is, currently under development for a number of space instruments, including the Space Telescope Imaging Spectrograph (STIS). The MAMA package typically comprises a single curved MCP and a multi-anode array that consists of two orthogonal layers of interwoven anode elements that define an array of discrete pixels. Spatial and geometric linearity are determined by the

individual pixels and, therefore, excellent.

Spatial resolution/MCP gain : $25 \mu\text{m}/4 \times 10^5$ to $1 \times 10^6 e^-$

Global dynamic range : 3×10^5 counts/second.

Local dynamic range : 50 counts/second/pixel.

Format : 2048×2048 pixels

(80 mm active area MCP).

Wedge and strip anode

The wedge and strip anode consists of an array of interwoven wedge and strip electrodes on a metallic anode, where the relative area of the electrodes varies with location on the anode.¹² A charge cloud incident on the anode is partitioned between the three anode elements to provide a unique diagnostic of the event's spatial location. Spatial resolution of wedge and strip detectors is determined by the charge partition noise and the charge amplifier's noise. Since the charge amplifier's noise contribution is sensitive to the anode's capacitance spatial resolution is determined by the size of the anode array, unless the MCP gain is increased to compensate, or anodes are arranged in a mosaic. Wedge and strip detectors also require relatively large gaps between the rear MCP and the anode (3 to $5 \mu\text{m}$) so that the charge cloud can spread out enough to cover several elements at the anode. This large gap can make the detectors susceptible to electromagnetic fields affecting geometric and spatial linearity. Vallerga *et al.*¹³ have investigated these issues for the case of the EUVE detector and shown that they are relatively low level (<1.3%), can be accurately modelled and, therefore, calibrated.

Spatial resolution/MCP gain : $50 \mu\text{m FWHM}/10^7$

Global dynamic range : 10^5 counts/second

(signal processing limited)

Local dynamic range : 5 to 10 counts/second/pixel

(MCP stack limited)

Format : 1000×1000 pixels

Delay-line

An alternative to encoding spatial location from charge partition is to use the pulse rise-time. This technique, which was originally developed for gas proportional counters, has recently been applied to MCP-based detectors.^{14,15,16} One of the major advantages of this technology is

that, in contrast to charge partition methods, increasing the format size does not degrade spatial resolution and may actually improve it. Two different approaches are currently under development. Siegmund *et al.*¹⁴ are developing a hybrid two-dimensional planar delay line for spectroscopy, which consists of a double delay line structure where the C coordinate is determined from pulse arrival times and the Y coordinate from charge partition. An alternative approach is the 2-dimensional helical delay line, which consists of an orthogonal pair of flattened-helical bifilar transmission lines.^{15,16} Pulse arrival times at the end of each wire each line give the event location, with the noise pickup and x-y cross talk being eliminated by differential-mode amplification for each pair of lines.

Spatial resolution/MCP Gain : $18 \mu\text{m FWHM}/2 \times 10^7$
 Global dynamic range : 4×10^5 counts/second
 (signal processing limited)
 Local dynamic range : 5 to 10 counts/second/pixel
 (MCP stack limited)
 Format : 2000 x 2000 pixels
 (65 mm diameter)

Intensified CCD

The intensified CCD employs an MCP based intensifier coupled to a CCD detector, usually by means of a fiber-optic taper. The CCD is usually operated in a fast-scanning video mode with analysis of each frame in hardware to determine the centroid of each event.¹⁷ Analog modes of operation have also been proposed to overcome dynamic range problems in some situations, though this technique still needs to be fully explored.

Spatial resolution/MCP Gain : $11 \mu\text{m FWHM}/10^7$
 Global dynamic range : 3×10^4 (CCD frame rate)
 Local dynamic range : 5 counts/second/pixel
 (2 MCP stack)
 Format : 2048 x 2048 pixels
 (75 mm diameter)

In addition to the MCP readout schemes we have discussed, there exist a multitude of methods including the resistive anode,¹⁸ the quadrant anode,¹⁹ the coded-anode array (CODACON)²⁰ and the

multi-wire array.²¹ An alternative to MCPs, proposed by Vallerga *et al.*²² is the focused-mesh electron multiplier (FMEM), which comprises a stack of perforated metal dynodes. The FMEM is used in a similar way to MCPs, with the same readout systems. Its main advantage is that it does not suffer from dynamic range limitations and, in principle, it can be applied to very large format detectors where MCPs would be unsuitable. Vallerga *et al.*²² demonstrated its operation with a simple wedge and strip readout system, but further development work would be required to overcome its main limited spatial resolution, set by the dynode mesh pore spacing and size.

Direct electron bombardment

Electron bombardment is also a well proven readout technique. The spectrographs currently installed on HST, the FOS and GHRS, both employ Digicon tubes with electron bombarded linear arrays. The concept has been investigated for 2-dimensional applications, most recently by Cuby *et al.*²³ who use a thinned, CCD for backside electron bombardment, avoiding damage to the frontside electrode structure. Frontside bombardment had led to rapid degradation of the CCD in previous experiments. The 2-dimensional Digicon has achieved good photon counting performance, with sub-CCD pixel resolution. Dynamic range is similar to that of the intensified-CCD, since the same frame rate and event processing limitations apply. Further work with larger format CCDs is needed to demonstrate the 2-dimensional Digicon's potential for large format UV imaging. Geometric stability is a concern for these detectors since experience with the GHRS and FOS digicons indicates that the influence of the earth's magnetic field is sufficient to cause image drift.

8.3 CONCLUSIONS

Based on our assessment of today's detector technology and upon our understanding of the resources being invested into this technology, we have arrived at the following three conclusions relating to the HST AC detectors.

First, the state-of-the-art CCDs being manufactured today are suitable for 3000 to 10,000 Å observations *provided such detectors are still available when required for the AC*. Because of the limited number of manufacturers who make suitable scientific grade large format CCDs,

we believe it to be in NASA's best interest to procure the required detectors in this wavelength region as soon as possible. If suitable CCDs are not being commercially manufactured at the time required for the AC, the cost may be several million dollars (compared to several hundred thousand at the present time).

Second, CCDs are well suited for 2000 to 3000 Å observations but they have not been suitably optimized for the AC requirements. We encourage NASA to support the development of CCDs optimized for space flight observations in this spectral region, especially by supporting the development of thinning, backside charging, and anti-reflection coating techniques.

Third, the appropriate detector technology for 1000 to 2000 Å observations is less certain than for the longer wavelength regions and requires considerable development for the AC. We encourage NASA to continue to support the photocathode-based technology that insures solar-blind operation, especially the use of large-format MCPs and appropriate readout technology. We also encourage NASA to support the development of the appropriate filter technology (such as Wood's filters), which might allow efficient solar-blind observations in this spectral region using CCDs. The combination of efficient Wood's filters and CCDs optimized in the 1000 to 2000 Å region would be a significant advance in detector technology.

ACKNOWLEDGEMENTS

Michael Lesser and Mark Clampin were the principal authors of chapter 8.

REFERENCES

1. Janesick, J., Elliott, T., Collins, S., Marsh, H., Blouke, M., and Freeman, J. 1984, SPIE 501, 1.
2. Janesick, J., Elliott, T., Daud, T., and McCarthy, J. 1985, SPIE 570, 46.
3. Janesick, J., Elliott, T., Daud, T., and Campbell, D. 1986, SPIE 627, 543.
4. Lesser, M. P. 1987, *Optical Engineering*, 26, 911.
5. Coleman, C. I. 1981, *App. Opt.* 20,3693.
6. A. Boggess *et al.* 1978, *Nature* 275, 372.
7. A. Boksenberg and C. I. Coleman 1973, *Adv. Elect. Electron. Physics* 52, 355.
8. F. Macchetto 1982, in *The Space Telescope Observatory*, NASA CP-2244, 40.

9. O. H. W. Siegmund, K. Coburn and R. F. Malina 1985, *IEEE Trans. Nucl. Sci.* NS-32 443.
10. O. H. W. Siegmund and J. Stock, *SPIE* 1549, 81.
11. J. G. Timothy 1981, *Rev. Sci. Instr.* 52, 1331.
12. C. Martin, P. Jelinsky, M. Lampton, R. F. Malina and H. O. Anger 1981, *Rev. Sci. Instr.* 52, 1067.
13. J. V. Vallerga, G. C. Kaplan, O. H. W. Siegmund, M. Lampton and R. & F. Malina 1989, *IEEE Trans. Nucl. Sci.* 36 881.
14. O. H. W. Siegmund 1992, Private Communication.
15. S. E. Sobotka and M. B. Williams, 1988, *IEEE Trans. Nucl. Sci.* NS-35, 348.
16. Friedman, P. G., Martin, C. and Rasmussen, A. 1990, *SPIE Proc.* 1344, 183.
17. J. L. Fordham, D. A. Bone, T. J. Norton and P. D. Read 1990, *SPIE* 1235, 636.
18. M. Lampton and F. Paresce, 1974, *Rev. Sci. Instr.* 45, 1098.
19. M. Lampton and R. F. Malina 1976, *Rev. Sci. Instr.* 47, 1360.
20. W. E. McClintock, C. A. Bazhi, R. E. Steele, G. M. Lawrence and J. G. Timothy 1982, *Appl. Opt.* 21, 3071.
21. E. Kellogg, P. Henry, S. Murray, L. Van Speybroeck and P. Bjorkholm 1976, *Rev. Sci. Instrum.* 47, 282.
22. J. V. Vallerga, J. Hull and M. Lampton 19, *IEEE Trans. Nucl. Sci.*
23. J. Cuby, J. Richard and M. Lemonier 1990, *SPIE* 1235, 294.

9. Filters, Grisms, Polarizers

We have considered in general terms what complement of filters, grisms, and polarizers the Advanced Camera (AC) will require.

In the far ultraviolet (FUV) channel, about 24 filtration elements will be needed, and in the near ultraviolet optical (NUVO) channel, more than about 48. We recommend two or three grisms in the filter wheels of the NUVO wide-field channel, and a set of polarizing filters for at least the UV and NUVO high-resolution channels. Further development of Wood's filters is required to assure the utility of FUV sensitive charge coupled devices (CCD)s. We recommend an internal flat field system capable of providing reference flat field images over the full spectral range. A capability to recalibrate filters on orbit is desirable.

The Advanced Camera (AC) will require filters, grisms, and polarizers to define scientifically useful spectral passbands and permit optical analysis of the images over the full wavelength range 1000 to 10,000 Å. A relatively large filter set will be required to reproduce the selection of passbands that both Hubble Space Telescope (HST) and groundbased observers require. The selection of this filter set will be driven by the science priorities of the AC, and limited both by the state of technology and the space available within the AC. Later study phases will prioritize the actual filter choices based on detailed scientific consideration. In this chapter, we outline some general considerations.

9.1 CURRENT CAMERA FILTERS

The three cameras built for HST as of late 1993 serve as background for this report: the Faint-Object Camera (FOC), the present Wide Field Planetary Camera (WFPC), and the second Wide Field Planetary Camera (WFPC II) with corrective optics to be installed during the HST repair mission. All three use stacked filter wheels populated by a large

Table 9.1. Filters, grisms, & polarizers in pre-1999 cameras

	FOC	WFPC	WFPC II
# filters, etc.	58	48	48
#filter wheels	4+2	12	12
Red leaks in UV	Yes	Yes!!	Yes!
#Grisms	[5] _{Sph Ab}	[3] _{Sph Ab}	—
Polarization	Yes	[Yes] _{Radial}	[Yes] _{Radial}
# narrow-band	4	~14	15+linear+quads

Note: Quantities in square brackets indicate impaired performance for the reasons indicated.

number of individual narrow-, medium-, and broad-passband filters, as well as other elements such as grisms and polarizers. The number of filters in WFPC and WFPC II is limited by the volume available within the radial bay envelope, is 48 separate elements installed in 12 wheels. The FOC contains a total of 58 optical elements installed in four wheels of the $f/96$ channel and two wheels of the $f/48$ channel (see table 9.1 for a summary).

In the first generation cameras, conventional filter technology was pressed to its limits. The broad-band visible filters are generally colored glass combined with low-pass, long-wavelength blockers. The ultraviolet (UV) filters are generally metallic. The narrow-band and interference filters are multi-layer or multi-element.

9.1.1 WOOD'S FILTERS

In both the WFPC and the FOC, the UV filters have undesirable red leaks, the effect of which is particularly bad for WFPC because of the high red and near-infrared quantum efficiency of its charge-coupled device (CCD) detectors. Because most astronomical sources emit from 10^4 to 10^8 visible photons per far ultraviolet (FUV) photon, the contamination from the red leak is significant for anything other than the very hottest sources.

There are two ways to solve this red-leak problem for FUV observations with the AC. The AC can use a solar-blind detector for the FUV channel or use Wood's filters, which transmit at FUV wavelengths but block longer wavelengths. The two successful types of Wood's filters are based on thin sodium and potassium films. Attempts at building Wood's filters based on lithium have failed thus far. Table 9.2 summa-

izes the cutoff wavelengths λ_{cutoff} and remaining transmissions $T_{6000\text{\AA}}$ at 6000 Å for these three types of Wood's filters. A sodium Wood's filter developed for WFPC II has been shown to provide excellent red-leak suppression, but relatively low, 18% peak transmission. Sodium Wood's filters are now being fabricated by Ball Aerospace as well. To assure the possible future utility of FUV-sensitive CCDs, we recommend that Wood's filters be further developed.

Table 9.2. Three types of Wood's filters

Type	Thickness	λ_{cutoff}	$T_{6000\text{\AA}}$
Li	0.5 μ	1600 Å	10^{-7}
Na	0.5 μ	2000 Å	10^{-12}
K	0.5 μ	3200 Å	10^{-8}

Note: No Lithium filter has been successfully built yet.

9.1.2 NARROW-BAND FILTERS

Whereas the FOC and WFPC have only a limited set of narrow-band filters (see table 9.1), WFPC II has an enhanced narrow-band capability, including 'quad' filters, which consist of four adjacent filters lying in the beam in front of a single CCD chip. The resulting subdivision of the field of view into four smaller fields with different passbands offers four times more passbands, but presents a disadvantage for wide-field applications. WFPC II also uses a set of linearly variable ramp filters, which provide a narrow passband of $\sim 1.5\%$ of the central wavelength from 4000 Å to 10,000 Å over a small, 13" field of view. This wavelength range is covered by four strips on each of four filter substrates. The wavelength to be observed is then selected by choosing the proper CCD, filter, strip, and position along the strip. The wavelength scale of the ramp filters can be recalibrated in orbit by observing emission-line objects. In a second step, the conventional narrow-band filters can then be recalibrated by making observations by crossing them with the calibrated ramp filters.

To ensure similar or improved broad- and narrow-band capability in the AC, we recommend that a FUV channel accommodate at least 24 filters, and the high-resolution and wide-field channels at least 48 filters each. Preliminary design considerations suggest that an axial-bay camera could accommodate filter wheels with at least 60 positions. Tunable narrow-

band filters for the full FOV would be ideal (e.g., based on Fabry-Perot etalons), but peak transmission would need improvement over the low values that are customary. Therefore, we recommend that whatever narrow-band filter solution is finally adopted, it should not impact negatively on the wide- and intermediate-band filter capability, on which the majority of science projects with HST rely.

9.1.3 GRISMS

Both the FOC and WFPC provide sets of gratings for slitless spectroscopy, but because of the spherical aberration of HST this capability has been virtually useless. The replacement camera WFPC II does not contain any gratings, and thus the capability for low-resolution, large-field spectral surveys will be very limited at least until the AC. Arguing in favor of AC gratings is the utility they have shown on ground-based telescopes for surveys of point sources, such as planetary nebulae, and QSOs. The resolution gain of the corrected HST point-spread function will translate into faint detection limits for grating work, which will be further helped by the darker sky in orbit. Table 9.3 shows signal-to-noise ratios that would have been attainable with WFPC, had HST not had spherical aberration. With detector and throughput improvements to be expected for the AC, it seems likely that HST will remain competitive with even the largest ground-based telescopes, with a clear advantage in crowded fields and in the UV. Therefore, we recommend a limited set of about two to three gratings in the filter wheels of the wide-field channel.

Table 9.3. Predicted grating performance for WFPC

V	S/N(blue)	S/N(red)
21	21	44
22	11	25
23	5	13
24	2	6

Note: Above signal-to-noise ratios computed for WFPC without spherical aberration, $\Delta\lambda = 60 - 120 \text{ \AA}$, and exposure time of 1800 sec.

9.1.4 POLARIZERS

All three pre-1999 cameras feature some polarizing filters to measure polarization. Because of the near 45° reflection necessary to feed the WFPC and WFPC II in the radial bay, only relatively strong polarizations can be measured by these two cameras. On the other hand, the recommended axial location of the AC may permit the measurement of weaker polarization in straight-through optical configurants. To take advantage of this opportunity, we recommend the inclusion of a set of polarizing filters in the filter wheels of at least the FUV channel and, if possible, all the channels without off-normal reflections.

9.2 CALIBRATION

The AC should incorporate by design an internal flat-fielding system capable of providing reference flat-field images over the full 1100 to 10,000 Å spectral range of the camera. In many science applications, the degree to which images in different passbands can be flat-fielded determines the effective limit for detection or measurement (e.g., color gradients of extended sources, stellar photometry over wide fields, etc). We have learned that flat-fielding by sweeping across the Earth yields too many photons for the broad-band filters of WFPC. Therefore, the AC must have internal flat-fielding capability, as indeed will be the case in WFPC II. Also, a capability to recalibrate all filters in orbit is desirable, but may be only partially achievable.

ACKNOWLEDGEMENTS

Francois Schweizer was the lead author of chapter 9, and John Hoessel contributed.

10. Instrument Computer

We have examined the functional requirements for the instrument computer, and find that a robust computer will be required to assure adequate detector control, data compression, self-documented data, and possible autonomous guiding. A large memory will be required to hold at least one or possibly more frames for on-board processing.

In considering the scientific requirements for an Hubble Space Telescope (HST) Advanced Camera (AC), it is clear that the capability of the on-board data system is a crucial element, one that can potentially limit the scientific success of the instrument. In this context, the on-board data system includes the detector outputs, readout controllers, on-board computer, and ground system. (See chapter 8.1.3 for a discussion of the charge coupled device (CCD) detector controller requirements.) For downlink, the concerns are mostly in the area of data bandwidth and compression. For uplink, the issue is the predictability of the state of a highly capable instrument so that efficient commanding is possible.

This chapter discusses six specific issues pertaining to the data system for the AC:

- Data volume
- Data compression
- Memory requirements
- Use of a common computer system
- Processor speed
- Operational concerns

10.1 DATA VOLUME

To improve on the 1600 x 1600 pixel Wide Field Planetary Camera (WFPC) design, which digitizes to 12 bits per pixel, the AC will employ at least a 2048 x 2048 pixel detector with at least 16 bits per pixel. How-

ever, in order to accommodate the high resolution, high contrast, wide field imaging capability that is unique to the HST, the AC should contain one 4096×4096 pixel detector with 18 bit digitization. This implies a data volume of 302 Mbits / image readout. At the spacecraft-limited rate of one Mbit/second, each readout requires roughly five minutes to dump to the Science Tape Recorder (STR). As will be shown below, this rate would be a severe limitation for a moderate speed, multiport CCD unless a significant amount of local data memory is available. Since the STR can hold and dump only one 20-minute segment of data (absolute maximum) per orbit, without data compression there will be a maximum of four AC readouts per orbit. Even if we were dumping directly from instrument memory, the amount of Tracking and Data Relay Satellite System (TDRSS) downlink access time is well matched to the STR limitations, and the result is not substantially different. Since we want to operate the AC in fast framing and parallel observation modes, an effort must be made to reduce the data volume.

10.2 DATA COMPRESSION

Direct, lossless data compression today can routinely achieve factors of 2 to 4 in compression. More advanced (and probably more compute intensive) methods can provide up to a factor of ten. However, when discussing the feasibility of data compression strategies, it is important to look at the overall data transmission system. For instance, a particularly difficult possibility is that when data is compressed twice (say, once by the instrument, and again during the TDRSS transmission), some algorithm combinations actually lead to data expansion. The benefit of such direct compression schemes is that they are straightforward to implement and, when properly integrated, work transparently.

Other forms of compression that do not involve direct interaction with the data stream may be more useful to consider. These require intelligent on-board control of data collection and, in many instances, can be much more effective. Readout of a subset of the available pixels is an example, although it may not be applicable for wide-field science. For fields where photometry is required at known positions, on-board selection and readout of subimages can significantly reduce the data rate. For CCD's using long exposures, the possibility arises for taking multiple frames and doing cosmic ray filtering before transmission. Such strategies could reduce bandwidth significantly with almost no loss of capability, but would re-

quire a relatively powerful on-board processing capability.

Finally, it is also possible that some of the compression can take place using specialized readout modes of the detectors. In particular, if CCD's are used, then on-chip pixel summation is a standard strategy. (See chapter 8 for details on these modes.)

10.3 MEMORY REQUIREMENTS

For large format detectors, an important factor is the memory capacity of the processor since we are assuming that full frame(s) can be stored and operated on. Assume also that 8-bit error correction is used so that each 32-bit word actually requires 40 bits of physical memory. A full frame has one Mpixels. For 18-bit pixels this translates into 302 Mbits (with data packing) or 38 Mbytes. With error correction, the real physical memory required is 48 Mbytes.

For planetary imaging, the data memory requirements are much greater due to ephemeral nature of planetary objects. Many investigations will benefit from three color imagery; for instance dust clouds on Mars are best differentiated from ice clouds by comparison of yellow (green+red) and blue images. For such planetary imaging, the acquisition of imagery in three colors must be accomplished within a relatively short time period, else rotation will occur between the frames. (See table 1.2)

It will be possible for the AC to use a moderate speed, multi-port CCD, so that an image could be read out in less than a second, and the time between frames could be simply the time required to rotate filters.

The frame transfer region of CCD chips is a relatively inexpensive form of local memory. A frame transfer chip can contain two full frames, so that the total acquisition time for three frames is equal to the readout time of one frame (plus cycle time of filter rotation, etc.). A double frame transfer CCD could be used to accumulate three images "on chip." Unfortunately, cosmic rays also accumulate as the image waits for readout and, in the case of bright extended planetary objects, it will be difficult to detect such cosmic ray events.

Due to the cost of memory, it may be necessary to utilize sub-array readout for "fast framing" operation, but the data volume is still quite substantial. In order to utilize the high resolution capability of the corrected HST for imaging Jupiter, at least a 1000 x 1000 pixel array must read out for each frame, equal to 48 Mbits of uncompressed (and without error correction) data for three frames.

It is important to note that, internal to the instrument, it is data volume and not data rate that is the limiting factor. Large format detectors can be arranged to have parallel readout channels that can substantially increase the aggregate data. Local memory is then used to match this internal rate to the external interfaces.

10.4 COMMON COMPUTER SYSTEM

The HST Project is currently discussing the possibility and merits of common computer systems in the future Science Instruments (SI's) and for spacecraft functions (e.g., coprocessor for DF-224). The benefits are decreased maintenance costs and perhaps even development costs. Such common systems, if feasible, would require only a single type of simulator and skill to maintain, and would therefore contribute to the cost efficiency of the long-term HST program. However, there is also a loss of flexibility for the instruments when such a critical subsystem cannot be fully optimized. These discussions are not yet complete, but a number of results are useful for this study.

The systems being considered for Near-Infrared Camera (NICMOS) and Space Telescope Imaging Spectrograph (STIS) are based on either 386 or TI C30 hardware (or both). The coprocessor being developed for the flight DF-224 is also 386 based. This provides a guideline for the computing capability that can easily be envisioned for the AC. While a 386 provides a nominal amount of computing capability by today's standards, it will no doubt be woefully obsolete in the AC time frame, if not sooner.

The NICMOS/STIS experience can be applied to the AC in the sense that flight computers of these instruments are expected to be able to support the kinds of commanding and data processing capabilities that have been discussed for the AC. This includes 'macro'-type commanding capabilities, some degree of acquisition autonomy (limited more by the ground system than on-board capability), data management in on-board memory, interactions with the STR, data formatting, as well as the usual mechanism and detector control functions. The main difference for the AC is in data volume; the computers' capabilities hopefully will be substantially increased. Note that on-board data processing for the STIS using the currently envisioned data system was slightly marginal when 2048 x 2048 detectors were still being considered.

10.5 PROCESSOR SPEED

In addition to the on-board processing schemes mentioned previously, it is also possible to use this capability to improve image quality in some cases. For example, when using photon counting detectors, it is desirable to create a synthetic image in real-time, while monitoring the telescope pointing, in order to correct for pointing errors. However, such capabilities require substantial computing power.

As a crude estimate of 386 processing power, assume that each pixel requires anywhere from 4 to 20 μ sec of central processing unit (CPU) time. This works out to a full frame processing time of one to six minutes. These numbers are somewhat high, considering that the CPU must also deal with real-time events and multiple detectors. However, proper event-driven design may still make this feasible.

10.6 OPERATIONAL CONCERNS

For a camera that has a powerful on-board computing capability, the possibility arises for autonomous operation that can substantially increase efficiency for certain classes of observations. This capability should be explored and exploited to the maximum extent possible in order to keep the AC competitive. For example, automatic sequencing of multiple color exposures, self-guiding capabilities, and dynamic subarray selection for processing are useful for meeting the proposed minimum scientific requirements. However, detailed interaction with the ground system needs to be considered, particularly for command generation and data interpretation. It seems that a self-documenting protocol for the state of the camera needs to be developed for the downlink.

10.7 SUMMARY

Given the detector improvements that will be incorporated into an AC, this instrument will need a very capable on-board data processing system that will allow for efficient observations. A flexible design will also enable future optimizations in AC operation to exploit unforeseen opportunities.

It appears that unless some development is done on faster processors and cheaper memory, we will be running into practical data processing limitations with the 4096 x 4096 pixel class detectors. While flight qualification of better processors cannot be assured for the development time frame of the AC, every effort should be made to pursue improving this

capability. Greater processing capability will enable using data compression and other more sophisticated data management techniques to handle the large data volume required.

We recommend that more sophisticated data handling techniques be vigorously pursued in order to enable the most efficient and cost-effective science capabilities for the next generation of imaging systems.

ACKNOWLEDGEMENTS

Edward Cheng was the lead author of chapter 10. Richard White contributed the section on data compression.

11. One Camera Design

This section describes one conceptual design for the Advanced Camera (AC), which was developed in Spring 1993 for a proposal to European Space Agency (ESA). It is described here to illustrate major technical issues, show how the science priorities can be addressed, and render a preliminary argument that the AC is feasible in technical and programmatic terms. This design also provided the starting point for the NASA Phase A study of Summer 1993, which looked at trades and options, and assessed the required resources of time and money for the AC.

The essential features of this conceptual design are defined by our preliminary choices for the detectors. The f -numbers of the camera modes are chosen to reconcile the performance priorities to the detector formats. The performance priorities (see chapter 1.2) are stated in terms of the desired field of view (FOV) and fineness of sampling, which are essentially field angles on the sky. Detector formats are given by the pixel spacing in microns (μm) and the number of rows and columns in the pixel array. The f -number of a camera mode determines the image magnification—in particular, how many detector pixels receive the light from one spatial resolution element on the sky.

11.1 CAMERA MODES

In response to our survey, the astronomical community expressed a strong desire for the AC to provide critical sampling at both long and short wavelengths. (See chapter 7.) Critical sampling means the detector subdivides just finely enough to capture all the spatial information the image contains.

For a diffraction limited telescope, one resolution element corresponds to the field angle λ/d , where λ is the wavelength of light and d is the diameter of the aperture. According to the Nyquist criterion, critical sampling calls for two samples per resolution element, which implies a

pixel spacing of $\lambda/2$ times the f -number of the camera. Even if the detector pixels undersample the image by as much as a factor of two, it may be possible to recover nearly the full information operationally. This would involve ‘substepping’—or taking other images displaced by fractions of a pixel—and analyzing all the images as an ensemble. (See chapter 7.)

For each of the two wavelength regions of the AC—the near ultraviolet/optical (NUVO) and the far ultraviolet (FUV)—we have baselined a single detector in this design. (Detectors are discussed in chapter 8.) D_1 , the baseline NUVO detector, is a charge-coupled device (CCD) with an $8192 \times 8192 \times 7.5 \mu\text{m}$ format. It is a mosaic of four 4096×4096 arrays, each of which can be reformatted independently by on-chip binning into pixels that are larger in size and smaller in number by programmable powers of two. In the FUV range, we have baselined a photocathode detector, D_2 , with $2048 \times 2048 \times 25 \mu\text{m}$ pixels.

Table 11.1 shows the modes in this design. Modes 1 and 2 satisfy the highest two of the three NUVO science priorities, except that the specified sampling is achieved at a 14% longer wavelength. Mode 2 provides a FOV that is twice as large in linear dimension as the science recommendation in chapter 1.

Table 11.1. Modes of the AC.

Mode #	f/l	Detector	Binning	Type	Band ²	Format	FOV ³	Pixel ⁴	Critical Sampling
1	24	D_1	2 x 2	CCD	NUVO	$4096 \times 4096 \times 15 \mu\text{m}$	221"	0.054"	$\lambda > 12,600 \text{ \AA}$
2	24	1/4 D_1	1 x 1	CCD	NUVO	$4096 \times 4096 \times 7.5 \mu\text{m}$	110"	0.027"	$\lambda > 6250 \text{ \AA}$
3	192	D_2	n/a	PC	FUV	$2048 \times 2048 \times 25 \mu\text{m}$	23"	0.011"	$\lambda > 2600 \text{ \AA}$
4 ¹	—	$D_3(?)$	—	CCD	NUVO	—	—	—	—

¹Coronagraph: under study, not in baseline.

²Near ultraviolet/optical (NUVO): 2500 - 10,000 \AA . Far ultraviolet (FUV): 1000 - 2500 \AA .

³Maximum due to detector and f -number. Linear dimension.

⁴Angular subtense of one pixel.

Mode 3 addresses—but does not fully satisfy—the two FUV priorities; the sampling is too coarse for the high resolution FUV specification and the field is too small for the wide field FUV specification. Thus, Mode 3 should be regarded as one possible compromise with two problematic issues for the FUV, (1) uncertainty about the shortest wavelength for which critical sampling is actually beneficial in the case of the

as-built Hubble Space Telescope (HST), and (2) availability of a FUV detector with *either* a larger format than 1024 x 1024 *or* smaller pixels than 25 μm (see chapter 8).

Mode 4—the NUVO coronagraph—is a desirable option, but it is not included in this design. Further effort is needed to study the implications of either making it a sub-mode of Mode 2 with D_1 as the detector, putting a second set of optics behind the $f/192$ aperture and using a new NUVO detector (D_3), or other possible implementations.

11.2 FOCAL PLANE CONSTRAINTS

Based on the analysis and recommendation in chapter 6, this design is for the AC installed in an axial bay. Figure 11.1 shows the HST focal plane with the AC and three other axial instruments in place. The arrangement of the $f/24$ and $f/192$ relay entrance apertures are constrained by the proximity of the corner of the box and by the mechanical interface with the Optical Telescope Assembly (OTA), as shown in figure 11.2. The 200" x 200" (nominal) aperture of the AC is the largest unvignetted FOV available. A grazing-angle reflection, shown in figure 11.3, is required to divert the beam of the $f/24$ relay away from the corner of the box in order to implement the large primary mirror. This angle should not be larger than 80° if one wants to keep the length of the diverting mirror within reasonable limits (415 mm for 80°). The $f/192$ relay is a straight-through configuration. The positions of both FOVs are completely determined by the need to fit the primary mirror of the $f/192$ channel within the inner corner of the box and to clear the $f/192$ beam past the $f/24$ diverting mirror.

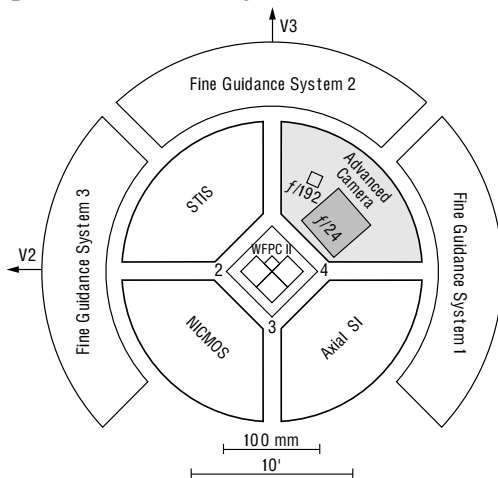


Figure 11.1. HST focal plane showing the AC and its two entrance apertures

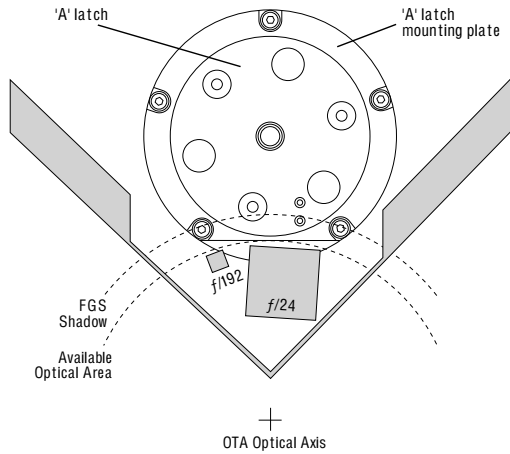


Figure 11.2. Available space of the HST focal plane for the AC entrance apertures.

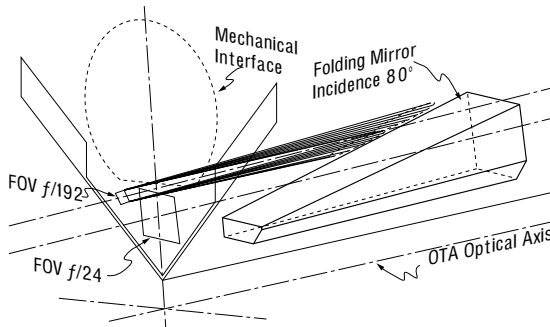


Figure 11.3. The optical design is physically realizable with the dual aperture layout shown in Fig. 12.1.

11.3 NOMINAL OPTICAL DESIGN

As was learned from the optical studies for the Corrective Optics Space Telescope Axial Replacement (COSTAR), the large OTA spherical aberration must be corrected at an image of the telescope exit pupil in order to provide a correction over the entire field of view of the camera. Therefore, the AC nominal design includes a large primary mirror (PM) that reimages the OTA exit pupil onto an aspheric secondary mirror (SM) that implements the correction. A third mirror (MM₁) permits the desired magnification and detector location. The mirror positions and tilt are chosen to correct for the HST astigmatism, image tilt, and camera field aberrations. Some residual axial coma can be compensated by slightly decentering the pupil image with respect to the secondary mirror.

11.4 ADOPTED OPTICAL SOLUTIONS

11.4.1 $f/24$ CHANNEL

As shown in figure 11.4, the first element in the $f/24$ channel is the 80° incidence folding mirror (FM) required in order to fit the channel within the axial bay envelope. The field of view is centered at 6.6 arcmin from the telescope axis. The primary mirror distance from the OTA focal surface is 1900 mm, the PM to SM distance is 905 mm, and the pupil diameter is 29.7 mm. All three mirrors with power have axially symmetric spherical or aspherical figures.

In the $f/24$ channel, the main difficulty is to compensate fully for the astigmatism variation over the 200×200 arcsec field of view, which is very large. Based on optical analysis, we can achieve this compensation and expect an excellent Strehl ratio at 5500 \AA over the whole field of view.

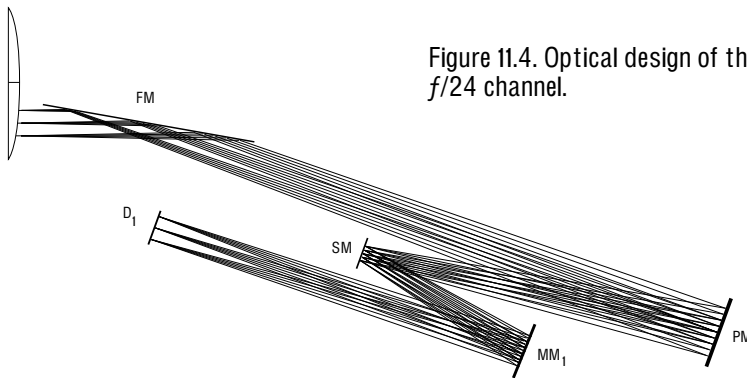


Figure 11.4. Optical design of the $f/24$ channel.

11.4.2 $f/192$ CHANNEL

As shown in figure 11.5, the $f/192$ straight-through channel is centered at 8 arcmin from the telescope axis. The PM and SM have been placed to obtain a sufficiently large image of the exit pupil that the tolerances for maintaining the exit pupil image on the secondary and the manufacturing difficulties of this mirror are kept within reasonable limits. The PM distance from the OTA focal surface is 1900 mm, the PM to SM distance is 640 mm, and the pupil diameter is 21 mm. (For comparison, the already-fabricated COSTAR mirror for the $f/96$ FOC relay— $f/151$ after correction by COSTAR—is a 15 mm diameter anamorphic asphere.) The PM is spherical and the SM is an axially symmetric asphere.

The $f/192$ high resolution mode is obtained by a toroidal MM_1 and a folding flat MM_2 , which feeds the image to D_2 , the solar-blind, FUV detector. The optics of this mode are diffraction limited at 1200 \AA over the whole field of view (Strehl ratio > 0.8).

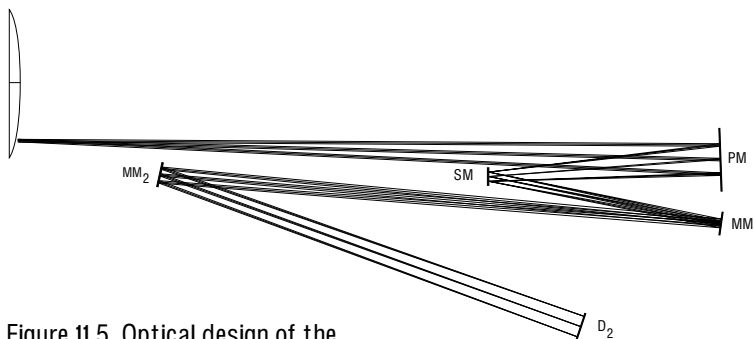


Figure 11.5. Optical design of the $f/192$ channel.

11.4.3 ADDITIONAL $f/192$ CHANNEL OPTIONS

The $f/192$ channel can accommodate the coronagraphic option, Mode 4, which is not part of our baseline design, but which is under study. Figure 11.6 illustrates this $f/86$ coronagraphic mode, which makes use of a separate CCD detector (D_3). MM_1 and MM_2 would be changed to different toroidal and aspherical prescriptions. A mechanism to change the fourth mirror could probably be avoided by using two different, fixed mirrors MM_2 at slightly off-plane positions. As one of these latter mirrors is a flat, flexibility is provided in the relative location of the detectors D_2 and D_3 .

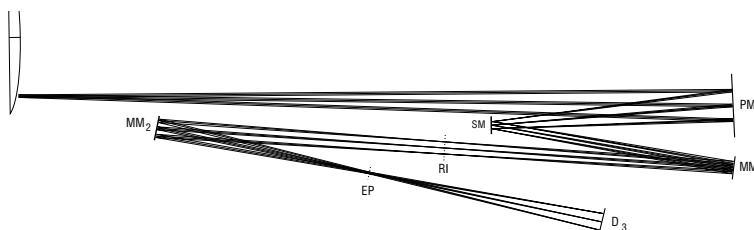


Figure 11.6. One concept for the optical coronagraph. Not in the baseline design.

To implement the apodization, a field occulter would be located at the intermediate real (and corrected) image (RI), which is located between MM_1 and MM_2 , and the Lyot stop would be placed at the second reimaging of the exit pupil (EP), which is located between MM_2 and D_3 . This mode is diffraction limited over a $50'' \times 50''$ FOV at 5000 \AA (Strehl ratio > 0.93) and near-diffraction limited at 1200 \AA (Strehl ratio 0.9 at center and 0.3 at the extreme field of view corner.)

As a whole, the current optical design of the $f/192$ channel allows much flexibility to respond to developments during the early study phases of the AC. For example, were developments to permit us to consider a $7.5 \mu\text{m}$ pixel CCD with a Wood's filter as a feasible option for the FUV detector D_2 , an $f/96$ mode satisfying the FUV critical sampling requirement could be easily implemented by only changing MM_1 of the current $f/192$ mode to a flat feeding this new detector directly. The same CCD as in the NUVO mode could probably be used, though with a different surface treatment for the FUV.

11.5 ADOPTED DETECTORS

Tables 11.2 and 11.3 give the detector performance specifications we have baselined for this AC design. In chapter 8, we discuss the issues related to achieving these specifications using the two major classes of detectors: CCD and photocathode-based photon counters (PC).

Table 11.2. Functional specifications for near ultraviolet/optical (NUVO) detector.

Format	: 4096 x 4096 x $7.5 \mu\text{m}$ on chip binning for normal 2048 x 2048 mode
Design	: 2-side buttable to make 2 x 2 array
Wavelength range	: AR coatings optimized for 2000 to 10,000 \AA Thinnable for back illumination with stable backside charging.
Imaging surface flatness	: $10 \mu\text{m}$ p-v
Read noise	: 2 to 4 electrons rms @ 50 kHz or greater
Dark current	: < 20 electrons/pixel/hour
Charge transfer efficiency	: > 0.999999
Quantum efficiency	: 50 to 100% in 2000 to 10,000 \AA region.
Dynamic range	: $> 10^5$ electrons ($15 \mu\text{m}$ -equivalent pixels) and 2.5×10^4 ($7.5 \mu\text{m}$ pixels)

Table 11.3. Functional specifications for the far ultraviolet (FUV) detector.

Format	: minimum 2048 x 2048 resolution elements
Wavelength range	: 1000 to 2000 Å
Spatial resolution	: less than 25 μm FWHM
Dark count rate	: less than 20 counts/sec/cm ²
Flat field uniformity	: at least 7% RMS
Quantum efficiency	: ~20% at 1216 Å and 10% to 15% at 1400 Å
Visible light rejection	: (1) CCD detector with Woods filter requires 10 ⁻⁸ (2) MCP detector requires 10 ⁻⁶
Dynamic range	: minimum 5 x 10 ⁵ counts/sec over detector
Local dynamic range	: better than 50 counts/sec/pixel

11.5.1 NUVO DETECTOR D₁

The AC will require thinned, back-illuminated CCDs to achieve the desired ultraviolet sensitivity. In the front-illuminated mode, which is how most off-the-shelf CCDs operate, light must pass through the electronic structures that define the pixels. This absorbs or reflects most radiation of $\lambda < 4000$ Å before detection. Back illumination avoids this problem, as the light can pass directly into the silicon. However, to assure that the electrons produced can be collected, it is necessary to thin the silicon layer to allow the photo-generated electrons to be detected at the CCD frontside.

In this design, the NUVO subsystem is implemented in a single optical configuration using a 2 x 2 mosaic of back illuminated CCDs. The individual CCDs will be 4096 x 4096 x 7.5 μm pixel devices. In the wide field mode, Mode 1, all four detectors are read out simultaneously, with each CCD being binned 2 x 2 on-chip to provide effectively four 2048 x 2048 x 15 μm devices. The optical sampling is therefore equivalent to a single 4096 x 4096 x 15 μm pixel CCD, with two small gaps running through the center of the image where the CCDs butt together. These gaps can be made less than 10 pixels wide.

In the high resolution Mode 2, only one 4096 x 4096 pixel CCD is read out. This provides twice the resolution of the wide field mode over one-fourth the area of the sky. In the baseline optical design, this mode provides full Nyquist sampling at 6250 Å. (The other three 2048 x 2048 CCDs are not used.) This mode is multiply redundant, since any of the four CCD detectors could be used as the Mode 2 detector.

11.5.2 FUV DETECTOR D₂

Based on emphatic community input, the response of the FUV detector should be strictly limited to FUV radiation and block contamina-

tion of the images by light from normal stars. To achieve this ‘solar blind’ performance, we must choose between two entirely different technologies for the FUV detector D_2 . The first type, PC detectors, are insensitive to red light without a blocking filter. The other, the CCD, is *very* sensitive to red light, and requires a blocking filter in front of it, known as a ‘Wood’s filter.’ A Wood’s filter employs a thin layer of an alkali metal to absorb light at frequencies higher than the plasma frequency of the free electrons. The result is that FUV light is transmitted to the CCD, and red light is absorbed.

Our choice to use the PC instead of the CCD/Wood’s filter combination in this conceptual design was not simple, nor can it be regarded as final until the availability of a PC detector is assured. Both technologies require development to assure the required performance. We believe the PC choice is currently the more solid option.

The PC has excellent imaging capabilities in the FUV region, and the performance factors that make them preferable over CCDs are their noiseless photon counting ability, relatively high quantum efficiency (QE), lack of sensitivity to red light, and high-frequency timing capability, which benefits observations of time-variable phenomena. (It may also provide an operational way to shutter out pointing jitter to improve image quality under some conditions.)

We currently envision that D_2 will consist of a CsI photocathode, which converts the FUV photon into a photoelectron, a micro-channel plate (MCP), which amplifies the photo-electron into a ‘cloud’ of 10^5 to 10^6 electrons, and a readout device, which encodes the position of the electron cloud within the detector format. FUV detectors of this general design have been used on a number of recent extreme ultraviolet and x-ray instruments in space, including the Roentgen Observatory Satellite (ROSAT), the Extreme Ultraviolet Explorer (EUVE), and the Hopkins Ultraviolet Telescope (HUT). Our current concept for D_2 is based on the FUV detector for STIS, a second generation HST instrument.

The original STIS MAMA detectors had the format specified in table 11.3 until they were downsized from a 2048×2048 to a 1024×1024 format in a descoping exercise to reduce cost and schedule risk. The MAMA package typically comprises a single curved MCP and a multi-anode array that consists of two orthogonal layers of interwoven anode elements, which define the array of discrete pixels.

ACKNOWLEDGEMENTS

Aden and Marjorie Meinel contributed the initial optical designs and analysis of the AC for this study. Starting with their work, Michel Saisse further developed the camera design presented here.

12. Schedule Urgency of the Advanced Camera

When is the Advanced Camera (AC) needed, and when must it be started? The 1999 Hubble Space Telescope (HST) servicing mission is required for other reasons—to reboost the orbit to prevent reentry. The AC is needed in 1999 to assure HST can provide an adequate imaging capability. The AC must be started now and developed on an aggressive schedule to assure it can be installed on the 1999 flight opportunity.

The urgency for initiating effort on the Advanced Camera (AC) is determined by the National Aeronautics and Space Administration (NASA)'s estimation of when the instrument will be required for installation, and then backing off from that future date by the length of time needed to develop the AC. To investigate the urgency of starting work now we must question those two assumptions.

12.1 INSTRUMENT DEVELOPMENT TIME

The development of instruments has emerged as the most difficult element of space science, at least in terms of the cost growth and schedule delays which have been experienced. Historically, on the Hubble Space Telescope (HST) Program, instruments have suffered many years of delays due to problems with the initial development effort and the Challenger accident. Second generation instruments, such as the second Wide Field Planetary Camera (WFPC II), Space Telescope Imaging Spectrograph (STIS), and Near-Infrared Camera (NICMOS) were repeatedly delayed by funding cuts and reprioritizations dictated by the difficulties facing the HST program before and after launch. If those trends continued, it would not be possible to build an instrument in twelve years, let alone six.

After the discovery of spherical aberration, plans were laid out to develop corrective optics, both through modifications to the WFPC II and

with the Corrective Optics Space Telescope Axial Replacement (COSTAR). Despite the challenge and importance of this effort, and lower financial reserves than usual for building instruments, these development activities have deviated substantially from previous NASA and HST instrument experience and are still on schedule for the December 1993 launch date. Part of this good performance is due to getting the extra attention and priority that the HST first servicing mission receives, but some inroads have also been made in reducing the turn-around time for decisions and streamlining some of the bureaucracy that has evolved. The lessons learned from WFPC II and COSTAR are also being applied to later instruments. Although the funding for COSTAR came right out of the follow-on instruments, the STIS and NICMOS teams have conspicuously departed from historical trends in managing to maintain an early 1997 launch date for the second servicing mission. The HST Instrument Program is continuing to investigate ways to accelerate the instrument development process.

In summary, although general NASA experience has been trending towards longer development schedules for instruments, for the last two years HST instruments have been bucking that trend. The schedule for the AC that begins in early 1993 is based on the more successful trends of the last few years, and indeed is based on a faster-than-average development schedule. Since there is always some risk that the schedule will slip due to technical problems, we should plan for completion as soon as possible in order to have the maximum schedule margin if the April 1999 launch date is a hard cutoff. The AC schedule is ambitious, and this should be reflected both in the need to start expeditiously, and in terms of the relative simplicity that will be required if the instrument is to be delivered on time.

12.2 URGENCY OF APRIL 1999 FOR THE THIRD SERVICING MISSION

The life cycle schedule for the HST Program is based on servicing missions scheduled for December 1993, March 1997, April 1999, April 2002, and a flight to safely de-orbit HST in April 2005. This schedule is embodied in NASA's formal planning. The AC is scheduled for the April 1999 mission, and this study is timed in accordance with that date. Therefore, to examine the urgency of the schedule for the AC we need to examine the rationale for the April 1999 launch date.

The three factors which influence the timing of the servicing mission in the life cycle plan are the solar cycle, the need for repairs, and the importance of upgrading scientific capabilities.

12.2.1 THE SOLAR CYCLE

Every eleven years, peak solar activity heats the atmosphere. As temperatures rise, the density of the atmosphere at a given altitude also increases. The effect of this increased atmospheric density can be profound on a satellite in low Earth orbit, such as HST. The increased aerodynamic drag from the 1978/79 solar cycle slowed Skylab, which lowered its altitude and thereby exposed it to even denser atmosphere, leading to its unanticipated reentry in July 1979.

Solar activity is very difficult to predict, particularly in light of the fact that four of the last five solar cycles have been the most intense since reliable records have been kept back to the 1850's. Solar cycles can vary a year or more in terms of when they start or end, and the intensity is similarly difficult to predict.

HST was deployed at an altitude of 611 kilometers (km), and over the first two and a half years during the current solar cycle peak, the orbit has decayed to 590 km. The HST plan calls for reboost on every mission when there is sufficient residual fuel for the shuttle after rendezvous. Particular efforts will be made to reboost HST as high as possible prior to the next solar cycle in 1999. (During the period of minimum solar activity over the next five years or so, orbital decay will be relatively minor.) The goal of this reboost strategy is to minimize the likelihood that an additional shuttle reboost flight would be required in the 2000 to 2001 time frame simply to maintain a viable orbit. Before the initial launch of HST, analysis indicated that the optimal strategy would be to reboost as early as possible to minimize HST's exposure to the denser atmosphere during peak solar activity.

Even though HST is unlikely to be facing the immediate threat of orbital re-entry in April 1999, solar activity is likely to be building towards a peak by that time. The longer that HST is exposed to the denser atmosphere during peak solar activity, the more its orbit will decay, and thus the lower its subsequent altitude after a shuttle reboost. This currently appears to be a strong driver for an April 1999 launch date.

12.2.2 THE NEED FOR ON-ORBIT REPAIR

HST was designed with recognition that spacecraft subsystems fail or degrade over time, some through known chemical processes (such as batteries and solar arrays) and others through unforeseen technical difficulties. HST has substantial on-board redundancy to allow it to function even after one or more elements have failed. Most of the major HST subsystems are designed for replacement on-orbit, and part of the planned servicing missions are dedicated to replenishing redundancy and replacing degraded equipment.

Of the hardware that is expected to degrade over time, batteries are tentatively scheduled for replacement in 1997 and solar arrays in 1999 (six years after a new set is installed in 1993.)

Predicting specific failures is exceptionally difficult, but examination of historical data indicates that, after a period of infant mortality early in mission life, the distribution of failures with respect to time is random. In other words, the longer the gap between HST servicing missions, the more hardware there will be to repair. In the current plans, the longest gap between servicing missions is that between the initial launch and the first servicing mission. The shortest gap is between the March 1997 mission and the April 1999 mission. Therefore, the predicted level of failures requiring replacement does not provide a strong rationale for a servicing mission as early as April 1999.

12.2.3 MAINTAINING AND ENHANCING SCIENTIFIC CAPABILITIES

The HST life cycle plan is based on regular replacement of science instruments to insure continuity of service and to enhance scientific capabilities. Originally, two new instruments were to be installed every 2 1/2 years. The current baseline supports installation of two instruments in 1997, one in 1999, and one in 2002.

The difficulty in predicting subsystem lifetimes is particularly troublesome for scientific instruments, which historically have been among the most likely spacecraft components to degrade or completely fail. The significant risk of suffering partial or complete failures in one or more of these instruments on-board after the 1997 servicing mission argues against delay of the April 1999 servicing mission.

To the extent that an AC is more efficient or more sensitive than the instruments that it replaces, a delay in installation also reduces the overall scientific productivity of the HST system until the replacement is

made. Even more important, chapter 2 explains that, without the AC, no adequate imaging capability may be functional on HST in 1999. The also argues against delaying the April 1999 servicing mission.

SUMMARY OF SCHEDULE URGENCY

The rationale for proceeding briskly with the development of the AC is based primarily on the solar cycle and the risk of development schedule slips. Delaying the April 1999 reboost will increase exposure to denser atmosphere, leading to a lower orbital altitude both before and after reboost, and thus increasing the possibility that a “reboost-only” mission would be required in the 2000 to 2001 time frame. Delaying the development schedule leading to readiness for that April 1999 flight increases the risk that the instrument will not be ready because of technical problems at the absolutely last “drop-dead” point at which a shuttle must reboost HST, which will be in late 1999 or 2000. An additional factor motivating against delay of the April 1999 date is the desire to reap the scientific benefits of the AC for as much of the HST lifetime as possible.

ACKNOWLEDGEMENT

Greg Davidson was the lead author of chapter 12.

