ICRR−Report−403−97−26

Presented at the Joint Discussion (JD11) of the 23rd General Assembly of the International Astronomical Union, Kyoto, 17-30 August, 1997; to be published in Highlights of Astronomy, vol. 11

JAN 19 1998

THE SLOAN DIGITAL SKY SURVEY

MASATAKA FUKUGITA
Institute for Cosmic Ray Research, University of Tokyo
Tanashi, Tokyo, 188 Japan

ICRR-REPORT-403-97-26

Fermilab Library
0 1160 0059649 8

Abstract.

The concept and the present status of the Sloan Digital Sky Survey are described, with emphasis on the instrumentation aspect of the project.

1. Introduction

The Sloan Digital Sky Survey (SDSS) is a project undertaking a photometric survey of a half of the northern sky (π ster) and a follow-up spectroscopic survey of over one million astronomical objects, one million galaxies and a hundred-thousand quasars, complete within precisely defined selection criteria. This project aims at producing a large homogeneous data sample of the northern sky with high accuracy photometry. The data base should provide a basis for quantitative cosmology in the 21st century.

The most important feature of the project lies in a careful design of the system and a strategy to precontrol the quality of data and to enable us to understand the selection effects. No surveys are free from selection biases, and SDSS is not an exception. The important requirement is to know quantitatively what is observed and what is missed in the survey.

The photometric survey will be made in five colour bands to a limiting magnitude of $r' \simeq 22.5$ mag. We expect the photometric catalogue to contain 5×10^7 galaxies and a comparable number of stars. The follow up spectroscopy is wide band extending from 3900Å to 9100Å with a moderate dispersion of $\lambda/\Delta\lambda \sim 2000$. The spectroscopic sample will include 10^6 galaxies to $r' \simeq 18.2$ and 10^5 quasars to $g' \simeq 19.5$.

We have been building a dedicated wide-field telescope of 2.5m diameter (f/5), a large format mosaicked CCD camera that fits the focal plane diameter of 65 cm, and two double spectrographs fed with 640 fibres altogether. The photometric data will be calibrated using an ancillary 60cm telescope installed for this purpose, and all detected objects will be catalogued. Target selection is made based on this catalogue and selected objects are then subjected to spectroscopic studies. A fast computing facility and supporting software are also an important and challenging ingredient of this project.

The system parameters were determined to minimize the cost of instrumentation under the requirement that 10^6 spectroscopic objects are obtained in 5 years, and under the assumption of the availability of 2048×2048 thinned CCD chips with 24 μm pixels (physical size is 4.9 cm square), the state-of-the-art technology at the time of the system design. These two constraints fix the basic parameters of the instruments.

This survey is being carried out by a collaboration of University of Chicago, Princeton University, Institute for Advanced Study, Fermilab, JPG (a group of Japanese astronomers), Johns Hopkins University, University of Washington and US Naval Observatory. We refer to Gunn & Weinberg (1995) for an earlier publication that describes this project.

2. Telescope

It is crucial to realize a wide field of view for efficient photometric imaging. It is known that a wide, flat field of view can be achieved with a Ritchey-Chrétien design with two hyperboloids with the same curvature (Bowen & Vaughan 1973). The most successful telescope of this type is the du Pont 2.54m at Las Campanas Observatory, which has a 1.5° field of view (FoV) at f/7.5. The SDSS telescope has a diameter of 2.5m with a faster focal ratio f/5 and a FoV of 3°. Our survey strategy requires very low image distortion across the field. The two mirrors (2.5m f/2.25 primary and 1.08m secondary) have 6th and 8th order polynomial corrections of the order up to a few wavelengths to offset a positive curvature introduced by a Gascoigne corrector that corrects for astigmatism. The mirror structure function has been tested using the null lens technique with the aid of a computer generated hologram, and verified that it is equivalent of an atmosphere that adds 0.2 arcsec FWHM to otherwise perfect seeing. Two more correctors are prepared, one optimized for the imaging camera and the other for the spectrographs. The 3° field is 65 cm in diameter and the net focal surface of the camera is still slightly quartic in the radius. 27% of the incoming beam is obscured by the secondary mirror and baffles.

The image scale is $16.''5$ mm^{-1}, or $0.''4$ per 24 micron CCD pixel. We want a $3''$ fibre aperture for efficient spectroscopy of galaxies at our limiting magnitude; this aperture corresponds on the focal plane to 180 μm, for which good quality fibres that do not cause serious f degradation are available.

The telescope is of an altitude-azimuth design, which allows for a light weight supporting structure, resulting in low inertia, low friction and mechanical simplicity. It is housed in a small shelter which rolls away during observations, and has independently tracking wind and light baffles.

The observing site, Apache Point Observatory (APO), is situated on Sacramento Peak, New Mexico (32°46′50″N, 105°49′12″W), at 2800 m altitude, one killometre away from National Solar Observatory. A 3.5m telescope of Astronomical Research Consortium already exists at APO. The Observatory is 30km east of Alamogordo. We expect the median seeing to be about $0.8''$ in good photometric conditions. Surface brightness of the dark night sky is $V = 21.7$ mag (arcsec)$^{-2}$.

3. Survey Strategy

We want target the largest possible contiguous volume of the Universe. A naive solution is to choose a 120° cone centred on the north Galactic pole, i.e., the $b > 30°$ region. This is not optimal for the following reasons: (i) Galactic extinction is not minimum for all the region of $b > 30°$. (ii) This choice requires the telescope to be inclined down to 25°, which causes a significant differential chromatic refraction for fibre-fed spectroscopy. We have thus decided to squeeze the region to 110° in the north-south direction but 130° in the east-west. The central meridian is 12^h20^m with the survey centre at $\delta = +32.8°$. The selective extinction is reduced to $E(B-V) = 0.08$ at the survey boundary. With this choice we need to incline the telescope only to 35°, for which the loss of light due to differential chromatic refraction is at most 16% at the end of the spectrum. An extra advantage of this choice is that it makes the prime time of the observation earlier in the year, which is favourable from the weather point of view at APO, and slightly extends the observing season.

The photometric survey will be carried out with the drift scan or time-delay-and-integration (TDI) mode (Schneider et al. 1989) at the sidereal rate of 15 arcsec/sec. Off the equator, the telescope must be tracked in great circles. The effective exposure time is 55 sec. An important advantage of this technique is that we can save the readout time of CCD, which is comparable to the exposure time for our shallow imaging. Another advantage is that we can obtain excellent flat-fielding, since objects directly pass through 2048 pixels: extra flat-fielding is required only for one direction. The cost is that specifications of the optics and camera are significantly more demanding. As we discuss later we have already solved this problem for the camera. For the telescope optics, calculations show that the requirement is well satisfied.

We cover the survey area with 45 stripes of 2.5° wide and 130° long. Because the CCDs are not contiguous in the focal plane, one stripe is filled with two swaths of TDI scan. This requires 780 hr, or about 1.5 yr taking account of the fraction of clear, dark, and photometric nights. With 55

sec exposure, we expect $S/N = 5$ images at $r' = 22.5$ mag and $S/N = 35$ at $r' = 18.2$ mag, the limit of the spectroscopic survey.

We note that the survey efficiency is given by

$$\epsilon = \Omega D^2 q, \tag{1}$$

where Ω is the solid angle of FoV, D the diameter of the telescope and q is the quantum efficiency of the detector. The SDSS ensures $\epsilon = 4.8$ m^{2}deg^2 compared to 0.30 m^{2}deg^2 for 48″ Schmidt photography. This gain enables us to carry out multi-colour photometry simultaneously; we decided to carry out five colour photometry. This not only brings us important information on the stellar content for a huge number of galaxies, but also provides the essential means for the quasar target selection for a wide range of redshift.

When we cannot observe the north pole in autumn, we intend to carry out a survey of the southern sky. We plan to repeat scans of long thin strips (probably three strips): i.e., the region of 200 square degree will be surveyed 35-45 times. This makes the limiting magnitude 2 mag deeper and should bring us new insight on variable objects.

The strategy for spectroscopy is more conventional. We take 1800 3° fields, chosen with the aid of a tiling algorithm to minimize the loss of targets. Each field contains 640 objects and average exposure time is 45 min. With an appropriate overhead, we need 2000 hr or about 3.8 yr for this survey. We expect 100 galaxies per square degree at $r' = 18.2$ mag, and hence one million galaxies for π steradian of the sky. $S/N \geq 13/$Å is expected for the galaxies at our limiting magnitude.

An option for the spectroscopic survey is that we may select 0.1 million bright red galaxies and 0.9 million normal galaxies. The depth of the survey would be significantly increased with such a selection.

We will need to interchange quickly the imager and the spectrograph when the weather condition changes. So, these instruments are built in the form of cartridges.

4. Imaging Camera

The camera should fit the focal plane with a diameter of 65cm. We decided to make an array of $5\times6 = 30$ CCDs mosiacked on the focal plane. Extra space is used to arrange 24 smaller CCD chips for astrometry and automated focusing. A schematic view is given in Fig. 1. The central 30 CCD chips constitute the main photometric camera where 5 chips in one column are housed in one dewar. The u' imaging that requires the most stringent optics is placed in the central row. For g' and r' this requirement is less stringent and they are placed in the two edges. The space of the adjacent CCD in the row direction is slightly smaller than the size of the chip, so that it results in an 8% overlap for two swaths.

All photometric CCD chips have 2048×2048 imaging pixels procured from Science Imager and Technologies Inc (SITe, formally Tektronics). The CCD for u' imaging is a thinned back-illuminated chip with UV antireflection coating which attains UV sensitivity of 40% on average. We use thinned back-illuminated chips with normal antireflection coating for the $g'r'i'$ colour bands. Typical quantum efficiency from 5000Å to 7000Å is 80%. For z' band we use thick front-illuminated devices without coating. The filters are described in the next section, but here we only mention that they are plano-concave lenses serving as field-flatteners to match a slightly convex surface of the chip. The front side of the corrector (made of quartz) is coated differently in 5 strips to match the colour bands.

One of the major challenges for constructing this camera was to satisfy the stringent requirement for the accuracy of the position of the CCD chips. The direction of the rows of CCD must agree precisely with that of the tracking: we achieved the rotation of the chip with respect to the tracking line only $< 3\mu$m from the centre to the edge, which satisfies well the requirement. The departure of the CCD surface from the focal plane is only 15μm, much smaller than $24\mu\text{m}\times5$ for the f/5 beam, image degrading being negligible.

The auxiliary 24 chips are 2048×400 thick front-illuminated devices with r' and neutral density filters, and are divided into an upper and a lower group housed by two dewars. 22 chips are for

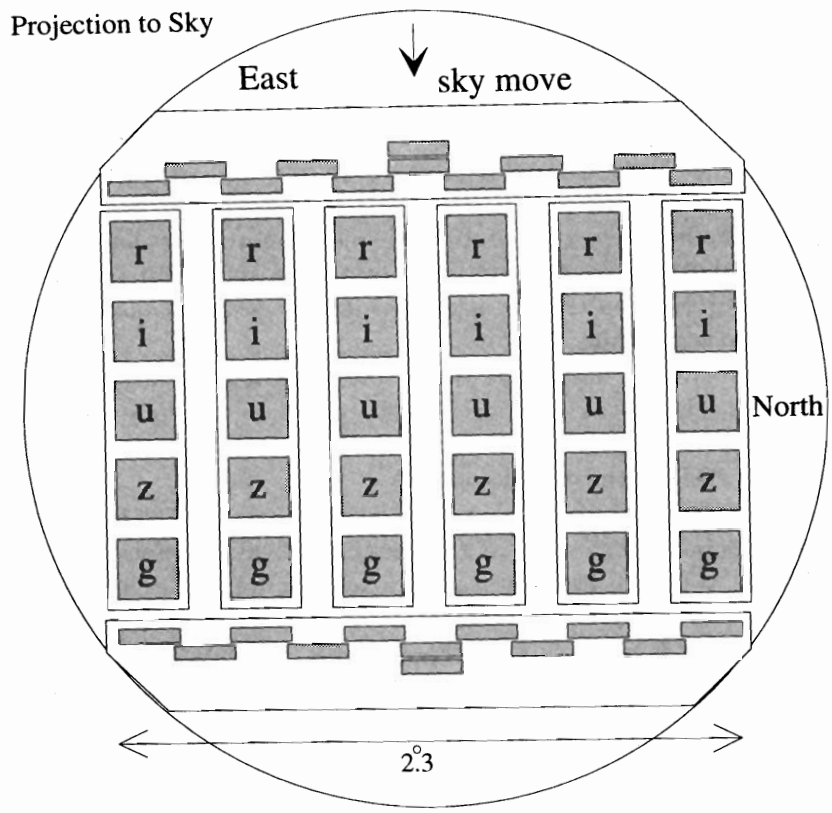

Figure 1. Schematic view of the photometric imager.

astrometry and two chips are for focusing. These chips allow observation of the bright stars in astrometric star catalogues to tie the SDSS system into these catalogues (see below).

A number of technically demanding requirements are satisfied by mounting the camera directly on a thick (about 5cm) quartz corrector, which provides a very stiff support. The filters are cemented directly onto the backside of the corrector to reduce the number of surfaces that would cause a loss of light. Holes are drilled on the corrector and screw anchors are embedded into these holes to support both optical benches, on which CCD chips are mounted, and dewars that contain electronics and liquid nitrogen containers.

High modularity of electronics is another notable feature of this camera. The camera contains altogether about 80 channels. 42 clock driver and 42 preamp cards (each contains two channel processors) are packed in the dewars to minimize noise. The signal from preamps goes into dual slope integrator and then digitized with a 16 bit ADC on signal-bias chain cards (also 42 units). The control signal, received by 8 bus-receiver cards, drives in-dewar clock drivers. The clock of the CCD is semi-synchronous (for some of the CCDs it is a factor 2 faster for technical reasons). One pixel is read in 24μs and it takes 52 sec per frame including overscan columns. The output of the on-chip charge amplifier is 1 μV per electron and the readout noise is typically 5 e. The photometric chips are cooled to $-80°$C by LN_2 to suppress the dark current.

5. Photometric System and Photometric Calibration

Our photometric system is five colour bands (u', g', r', i', z') that divide the entire range from the atmospheric ultraviolet cutoff at 3000 Å to the sensitivity limit of silicon CCDs at 11000 Å into five essentially non-overlapping pass bands. The SDSS system is substantially different from the Thuan-Gunn (1976) photometric system. The most salient feature is the very wide bandpasses used, even significantly wider than that of the standard Johnson-Morgan-Cousins system to ensure

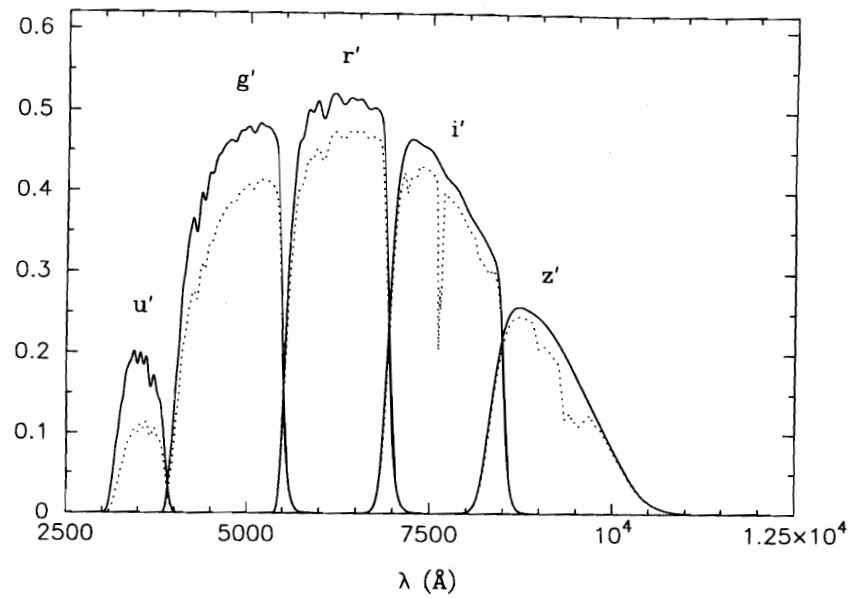

Figure 2. Response function of the SDSS photometric system (solid curves are without atmospheric extinction; dashed curves are at 1.2 airmass).

high efficiency for faint object detection.

The division of the pass bands is designed to exclude the strongest night-sky lines of O I λ5577 and Hg I λ5460. The u' band response is designed so that the bulk of the response is shortward of the Balmer discontinuity; this produces a much higher sensitivity (combining with g') to the Balmer jump at the cost of lower total throughput. The i' band terminates at 8500Å before a rapid increase of OH brightening. The filters are made from colour glass elements and a multi-layer interference film coating as detailed in Table 1. The transmission is > 90% except for the u' band.

TABLE 1. Elements of the SDSS filters

filter	glass	coating
u'	1mm UG11 + 1mm BG38 + 3mm quartz	suppression 6600-8200Å
g'	2mm GG400 + 3mm BG38	short-pass, cut off at 5500Å
r'	4mm OG550 + 1mm BK7	short-pass, cut off at 7000Å
i'	4mm RG695 + 1mm BK7	short-pass, cut off at 8500Å
z'	4mm RG830 + 1mm BK7	

For the calibration of the photometric system we have installed a 61 cm monitor telescope (Ritchey-Chrétien, f/10), and the precise definition of the SDSS photometric system is given by the response function of this monitor-telescope detector system. A thinned UV-antireflection coated 2k×2k chip is installed in this detector (FoV $27' \times 27'$). The system response functions are presented in Fig. 2 and the characteristic numbers are given in Table 2. The 2.5 m survey telescope uses the same filters, but their response is slightly different from that of the monitor telescope. Instrumental magnitudes obtained with the survey telescope are then corrected into the monitor telescope magnitude. Details are given by Fukugita et al. (1996)

Another new feature is the use of the AB magnitude, which can directly be translated into a flux in physical units, with the additional advantage of making the translation from spectroscopic flux to photometric wide band flux straightforward. The AB system of Oke & Gunn (1983), designated as AB_{79}, is now somewhat outdated; we have developed the AB_{95} system to meet the demand (Fukugita et al. 1996). In the new scheme we have retained the definition:

$$AB_\nu = -2.5 \log f_\nu \ (\mathrm{erg\ s^{-1} cm^{-2} Hz^{-1}}) - 48.60 \ , \qquad (2)$$

where f_ν is the flux per unit frequency from the object. The broad band magnitude is written

$$m = -2.5 \log \frac{\int d(\log \nu) f_\nu S_\nu}{\int d(\log \nu) S_\nu} - 48.60, \tag{3}$$

where S_ν is the response function including atmospheric extinction. A transformation into the conventional zero point (zero mag for A0 stars) is approximately $u' = u + 0.93$, $g' = g - 0.09$, $r' = r + 0.16$ and $i' = i + 0.40$, where primed refer to AB magnitude. To utilize the AB system, it is essential to know the accurate system response. A spectrophotometer was built specifically to measure the quantum efficiency of the CCD.

TABLE 2. Characteristics of the response function at 1.2 air mass

	u'	g'	r'	i'	z'
$\lambda_{\rm eff}$ (Å)	3557	4825	6261	7672	9097
FWHM (Å)	599	1379	1382	1535	1370

Photometric Calibration

We take as the fundamental photometry calibrators two spectrophotometric standard stars (metal deficient F subdwarf) BD+17°4708 and BD+26°2606 (Oke & Gunn 1983), which are given accurate spectrophotometry with respect to α Lyr. We want to add one more metal deficient F subdwarf BD+21°607 (Kent 1985) so that two stars are always accessible at any time, although this star lacks spectrophotometry. Some 100 stars ($V \leq 14$ mag) have been selected as candidates for primary standard stars which are to be calibrated against the fundamental triad to set up the network of standard stars over the sky.

The monitor telescope is used to observe several primary standard stars every hour to determine atmospheric extinction. In addition, we observe 'secondary standard stars' ($13 < V < 15$ mag) located in a few patches set in the survey strips with almost one hour interval to tie the survey magnitude with monitor telescope magnitude.

Our goal of accuracy is 0.01 mag for $g'r'i'$ and 0.02 for $u'z'$ for the primary standard star network, and 0.02 for $g'r'i'$ and 0.03 for $u'z'$ for the stellar objects in the survey.

6. Astrometric Calibration

The purpose of astrometry is two-fold: (i) to locate the fibre positions for drilling plates for spectroscopy, and (ii) for a good catalogue. For (i) an accuracy budget of < 270 mas is required not to degrade the spectroscopic survey.

We will use the Tycho Catalogue for astrometry, which will be supplemented with the proper motion data from Twin Astrograph Catalogue of USNO. The stars that appear in Tycho or TAC are as bright as $V = 12$ mag, which saturates the photometric camera. The astrometric camera installed in the imager assembly, however, can observe these bright stars, and the positions derived from these stars are used to give the positions of fainter 'secondary standard' stars of $14 < V < 17$ mag, imaged in the main camera.

We expect the error budget to be: errors in primary standards (40 mas); seeing in primary standards (25 mas); seeing in secondary standards (42 mas); shot noise in secondary standards (20 mas), and uncorrelated tracking errors (55 mas). These altogether yield 90 mas as the total error in the position.

7. Photometric pipelines

All data processing must be fully automated to deal with huge data volume. This software is called a "pipeline" and there are many pipelines in the survey system. For instance for the photometric data reduction, the camera data are cut into "frames" and each set of five frames sent to the "frames

pipeline" to detect and measure objects. To obtain the point spread function and flat-field vectors for a whole run, which are used by the frames pipeline, a reduced amount of data from the camera is sent to the "postage stamp pipeline". In this way, the code is parallelized, and the frames can be reduced in any order. The "postage stamp pipeline" also receives data from "monitor telescope pipeline" (where data from the monitor telescope are processed) and the "astrometric pipeline" (where data from the astrometric camera are processed), and forward them to the frames pipeline for calibrations.

The output of the frames pipeline is written into the "operational data base" and these data are sent to the "target selection pipeline" to give coordinates for plate drilling. The "science data base", which will be used for science, is also derived from the operational data base.

Apparently the most demanding element is the frames pipeline, which is a package for automated surface photometry, but which is significantly more sophisticated than existing software. The main part consists of "correct frame" (flat-fielding; interpolating bad columns or bleed trails of saturated pixels), "bright star subtraction", "find object", and "measure object" modules. The find object module is basically peak finding after smearing the image with the point spread function. A verification with simulated data ensures almost complete detection of normal objects to $r' = 22.5$ mag. What deserves some discussion is the choice of photometric parameters for galaxies derived by the measure object module. It is particularly difficult to measure a total flux for galaxies; isophotal magnitudes are clearly inappropriate for our purpose. We measure magnitudes at many apertures, and in addition measure the Petrosian magnitude (Petrosian 1976) as a pseudototal magnitude. Our trial parameters are the radius r_P at which surface brightness is 1/8 the average, and the flux encircled within $2r_P$. For the exponential profile this gives virtually the total magnitude, and for the de Vaucouleurs profile 92% of flux is included.

Target selection

We select objects that are subject to spectroscopy from the photometric catalogue. This process crucially determines the selection function of the spectroscopic sample. There are severe requirements from both science goals and practical feasibility; some compromises are required.

For the galaxies, we plan to use the Petrosian magnitude in the r' band for the flux limit, but we also apply the constraint that average surface brightness (or $3''$ aperture magnitude) be brighter than some limiting value to avoid observing objects for which we cannot obtain a redshift. This parameter introduces complication into the selection function but is necessary in order not to waste fibres.

We can take advantage of multi-colour photometry for quasar target selection, with which one expects 90% completeness and an efficiency of 65%. While we have set up a default selection criteria in multi-colour space, we need extensive testing of our tactics before the survey begins.

We also allocate a small number of fibres for stars, blank sky, calibrations and for the purpose of data quality assurance. Our provisional partition of fibres is 470-500 for galaxies, 100-110 for quasars, 5 for stars, 10–25 for blank sky and 5 for other purposes. All selection criteria are set in the target selection pipeline.

Adjacent fields of exposures necessarily have some overlaps. We can utilize this overlap to cover more galaxies in dense regions with appropriate tiling. Allowing for some regions with three fields overlapping, we can take virtually all galaxies to $B = 19$ for the APM catalogue (Maddox et al. 1990), if infinitesimally small spacing were allowed for adjacent fibres. In practice, however, we cannot allocate two fibres closer than 3.2mm in the focal plane, which corresponds to 55 arcsec. With this constraint 95.6% of the APM objects are taken as spectroscopic targets with 610 fibres.

In our project we plan to have an extended test period to fix all detailed parameters and tactics, no changes being allowed after the start of the survey. Target selection must be particularly carefully tested and verified.

8. Spectrograph

The requirements on the spectrograph are (i) a resolution $\lambda/\Delta\lambda \sim 2000$, corresponding to $\Delta\sigma \sim$ 140 km s^{-1}, a typical velocity dispersion for normal galaxies, to avoid the loss of S/N due to contamination of absorption features by continuum light. (ii) Wavelength coverage extends from

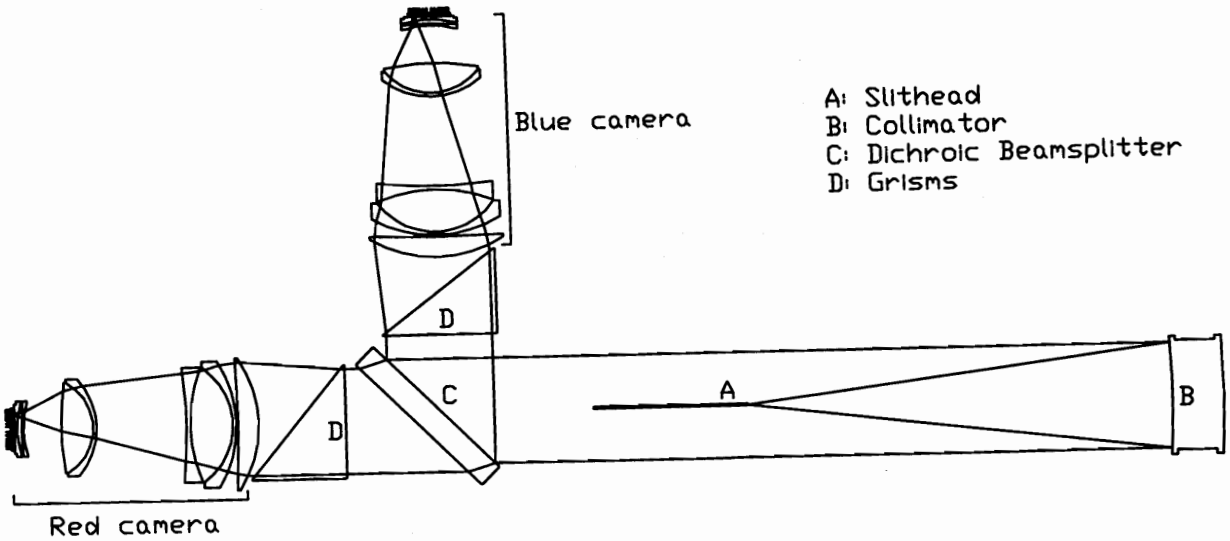

Figure 3. Schematic view of the spectrograph. See explanations in the text.

3900 to 9100Å. The minimum requirements are to cover Ca II H+K and Hα for all redshift up to $z = 0.25$; the above range extends a little further. G band and Mg *b* features are also covered. We miss [O II]λ3727 for nearby galaxies ($z < 0.05$), which, however, are a small fraction of the sample. (iii) we need > 550 fibres for the survey to be finished in 5 years.

These requirements are satisfied by constructing two double spectrographs each with dichroic splitter and 2k×2k CCD chips. The schematic design is shown in Fig. 3: A is a slit head with a slit 4.9 inches long and with 320 fibres installed perpendicular to the sheet, emitting f/4 light; B is a collimator of the Schmidt design (without a corrector) with a 6° FoV, movable with a DC servo motor. A 150 mm diameter beam is reflected from the collimator; C is a dichroic splitter, blue light (3900-6100Å) being reflected, while red light (5900-9100Å) going straight. Almost no loss of light is expected for reflection and 90% efficiency for transmission; D stands for two right angle grisms with grating of ruling densities of 640 (blue) and 440 (red) lines/mm. Zero angular deviation is at 4960Å and 7400Å . The use of a transmitting (rather than reflecting) grating allows us to install cameras close to the system pupil, so that the size of the camera can be reasonably small. The efficiency is 50-60% for all wavelengths. The flexure of the spectrograph is designed to be < 0.1 pixel for one hour exposure.

Behind the optics are two CCD cameras. In order to exploit the advantage of 2k×2k CCD chips with 24μm pixel, we need a wide field camera with a FoV of 16.5°. The focal length is 240 mm with f/1.3. The camera is of Epps' design; it is similar to that of LRIS at Keck (Oke et al. 1995) but with smaller aperture and larger FoV. The optics is composed of one singlet, one triplet and one doublet Petzval lenses plus a flattener. All but one of the surfaces are spherical and the asphere on the backside of the doublet is only mild. The detector is a thinned back-illuminated CCD with normal antireflection coating. The readout electronics is the same as that for the imager. The output flux spans 72 μm, i.e., 3 pixels, and hence the demagnification of the spectrograph is modest, only a factor of 2.5. We place individual fibres every 360μm (twice the inner aperture of fibres) at the slit head to avoid cross talk. So each camera can image >320 spectra. The net image quality is 20μm, which is about 1 pixel. The total throughput is expected to be 15-20% including the quantum efficiency of CCD and telescope losses.

We adopt the old technology of manual plugging of fibres (Limmongkol et al. 1993). A plug plate of diameter 81 cm is plugged with 640 fibres and 10 coherent fibre bundles, the latter used for a TV guider mounted on the spectrograph. There is one more larger hole with a wide-field fibre bundle for sky monitoring. Holes are drilled by applying a bending moment to the edge of the plate to match the best focus surface. We have confirmed that the engineering accuracy of the hole position

is 6-10μm (100-170 mas) and the deviation from the focal plane is 60μm rms (the maximum is 200μm at the centre where the optics is the best). We require the total de-centreing error to be smaller than 500 mas, and this accuracy is sufficient taking account of errors from astrometry and differential refraction. Fibres with a core diameter of 180 μm are made of silica and UV enhanced with a total flux loss of 10% rms (max 13%). The output beam is f/4 for the f/5 input beam. Fibres of 2 m long are bundled in units of 20 to ease handling. These fibres are plugged in day time, and installed on a cartridge, which contains a plug plate holder and a slit head of the spectrograph. We will prepare 10 cartridges which can be changed quickly from one to another. We do not need to mind which fibres are plugged into which holes. All connections are mapped by illuminating fibres from the spectrograph side using a fibre mapper.

9. Spectroscopy

Spectra of each target will be obtained with three 15 minutes exposures. This exposure time will be modified according to airmass, Galactic extinction and sky brightness to maintain the homogeneity of the spectroscopic sample. The overhead is expected to be 18 min per target field, including wavelength calibration and flat-fielding exposures each time.

The data reduction is made by the "spectroscopy pipeline" which is divided into two parts. The first is the "2D spectro pipeline", which processes the raw 2D spectral frames taken with three exposures into single 1D calibrated spectra for each object. The process includes bias subtraction, flat-fielding, wavelength calibration and sky subtraction. We use modified IRAF routines. The second ("1D spectro pipeline") processes the 1D spectra to extract scientific information. After merging red and blue halves of spectra together and masking pixels contaminated with sky emission lines, emission lines are identified and classified. This is followed by an analysis of absorption features with the cross-correlation technique (Tonry & Davis 1979). We also attempt to classify the spectra using a set of template spectra, and carry out a principal component analysis. Spectrophotometry with an accuracy of $\sim$ 15% is also one of the aims of the analysis. We must process 5000 spectra per day.

10. Data Base

A conspicuous feature of the SDSS is the huge amount of data it produces. The photometric imager produces data at a rate of 4.6 MB/sec and 12 TB altogether by the end of the survey. The astrometric camera also yields 3.7 MB/sec, but these data will not be archived. Postage stamp images of bright stars and quartile data of image frames, used to generate flat-field vectors, amount to 99 kB/sec (254 GB altogether) and 10 kB/sec (25 GB), respectively. The spectroscopic survey yields 153 kB/sec or 360 GB. We need to keep all data in the operational data base. We use the commercial object-oriented data base system Objectivity.

From these data bases we intend to produce the "science data base" which contains the data fully calibrated, indexed and read only, open to the public. We are planning a three level structure referring the frequency of use and amount of storage required, as presented in Table 3. The most memory consuming is the archive of atlas image, small regions that comprise all detected objects and also those where other surveys (e.g., X-ray, radio, IR surveys) found objects.

TABLE 3. Science data base

Level	data	storage
Level 1	object list (w/ spectroscopy)	2.5 GB
Level 2	object list (all)	250 GB
Level 1	atlas image	700 GB
	spectra (1D)	20 GB
	4×4 pixel all sky map	50 GB

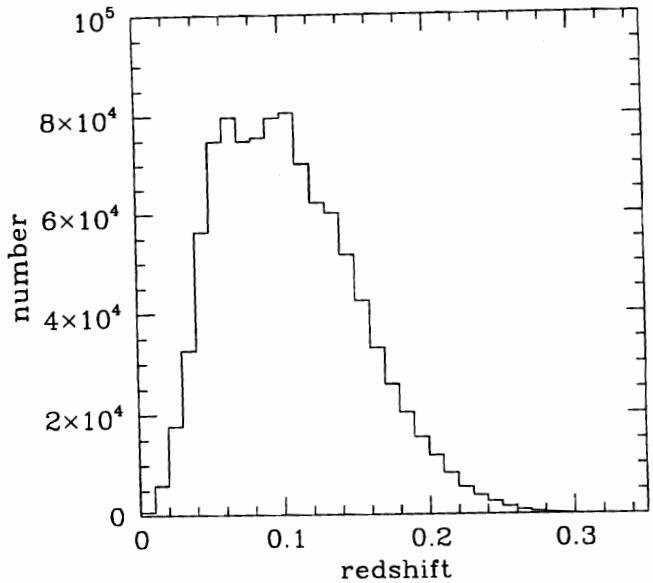

Figure 4. Redshift distribution of one million galaxies expected with the selection criteria discussed in the text. In this example a half-light surface brightness cut is imposed at < 22 mag (arcsec)$^{-2}$ in addition to the Petrosian-magnitude flux limit.

The requirement is that we must be able to sort out objects with desired properties, such as "blue galaxies fainter than 21 mag which are located within $10''$ of quasars of $z > 2.5$", as well as allowing easy access without any special facilities. Our simulations show that a brute force approach does not work any more, if a data base comprises 100 GB or more. We plan to use an object-oriented data base with a hierarchical structure, such as ODBMS which is commercially supported. Progress in data processing is rapid, and we still have some time to choose before the survey starts

11. Science

The science that can be contemplated with the SDSS data base is immense, and I do not intend to describe it here. Rather I mention only a few basic features anticipated for the study of galaxies, large-scale structure and quasars.

The depth of the survey reaches $z \simeq 0.2$, and the median of the redshift distribution is $z \simeq 0.1$ (see Fig. 4). The high redshift tail is cut off by the selection criterion for surface brightness compared to what is expected without this selection. The survey is dense and contiguous for the northern sky. We can, in principle, obtain power spectrum to $\lambda \geq 500$ Mpc, which overlaps with the region surveyed by COBE, to good accuracy. What we found crucial during our preliminary study is the necessity of accounting for Galactic extinction. Patchy Galactic extinction induces mock large-scale power in excess of the real power spectrum beyond 250 Mpc. This means that we cannot obtain meaningful power spectrum for $\lambda \geq 250$ Mpc unless the Galactic extinction is accurately known. In our survey, we plan to cross correlate an extinction map obtained from differential extinction of hot halo stars with those from galaxy number counts, star colour map and also with those from external sources, e.g., from HI and infrared data. If we succeed to control the extinction effects down to 0.05 mag in r', we can obtain a reliable power spectrum up to 400 Mpc or so.

The quasar survey is another important item. With a "normal selection" we expect to have 84,000 quasars and 250 with $z > 4$, median redshift being $z \sim 1$. The spectra will contain ~200,000 Lyman α absorbers with $2.3 < z < 3.0$ and ~2000 CIV absorbers with $2.8 < z < 5.0$. About 400 damped Lyman α absorbers will be obtained for $z > 3.4$. We also plan to produce a "high z quasar sample" using different selection criteria, resulting in 4500 $z > 4$ quasars and hopefully 300 $z > 5$ quasars. Roughly 400 lens candidates will be contained in the quasar sample, although we will need a separate follow-up study to confirm each lens.

12. Present Status and Schedule

The imaging camera has been completed and shipped to APO (October 1997). One of the spectrograph is finished except for the dewar in the red channel. The optics has been verified to work as designed, and readout is being tested. The blue camera is also being tested now, and will be installed in a month or so. All parts of the second spectrograph are ready except for a collimator that is not verified so far, but they yet have to be assembled. We except one spectrograph will be ready for verification in early 1998. The telescope still need more work. Basically all components are in hand, but the present image rotator is not adequate, and we have to wait until summer of 1998 for a new image rotator to be installed. The work on the telescope control will also take some more time. We expect, however, first light on the imager in the early summer of 1998. We plan to start the verification of the imager and tests for photometry with the telescope parked shortly after the first light. The monitor telescope system is in a commissioning phase, and expected to start data taking from the spring of 1998 (this work must be done in advance of the main survey). Spectroscopic first light will be achieved only in 1999. So, the year 2000 is perhaps a reasonable estimate for the survey to start.

I would like to thank all my colleagues working for this Project, in particular M. Carr, M. Doi, J. Gunn, T. Ichikawa, S. Kent, G. Knapp, R. Lupton, T. Mckay, E. Mannery, G. Pauls, M. Sekiguchi, W. Siegmund, D. Schneider, A. Szalay, A. Uomoto, D. Weinberg, N. Yasuda and D. York for discussions and information. I am grateful to G. Knapp and E. Turner for comments improving the manuscript.

References

Bowen, I. S. & Vaughan, A. H. 1973, *Applied Optics*, **12**, 1430
Fukugita, M., Ichikawa, T., Gunn, J. E., Doi, M., Shimasaku, K. & Schneider, D. P. 1996, *AJ*, **111**, 1748
Gunn, J. E. & Weinberg, D. H. in *Wide Field Spectroscopy and the Distant Universe*, ed by S. J. Maddox and A. Aragón-Salamanca (World Scientific, Singapore, 1995), p.3
Kent, S. M. 1985, *PASP*, **97**, 165
Limmongkol, S., Owen, R. E., Siegmund, W. A. & Hull, C. 1993, in *Fiber Optics in Astronomy II*, ASP Conference Series vol. 37, p. 127
Maddox, S. J., Sutherland, W. J., Efstathiou, G. & Loveday, J. 1990, *MNRAS*, **243**, 692
Oke, J. B. & Gunn, J. E. 1983, *ApJ*, **266**, 713
Oke, J. B., Cohen, J. G., Carr, M. Cromer, J., Dingizian, A., Harris, F. H., Labrecque, S., Lucinio, R., Schaal, W., Epps, H. & Miller, J. 1995, *PASP*, **107**, 375
Petrosian, V. 1976, *ApJ*, **209**, L1
Schneider, D. P., Schmidt, M. & Gunn, J. 1989, *AJ*, **98**, 1507
Thuan, T. X. & Gunn, J. E. 1976, *PASP*, **88**, 543
Tonry, J & Davis, M. 1979, *AJ*, **84**, 1511