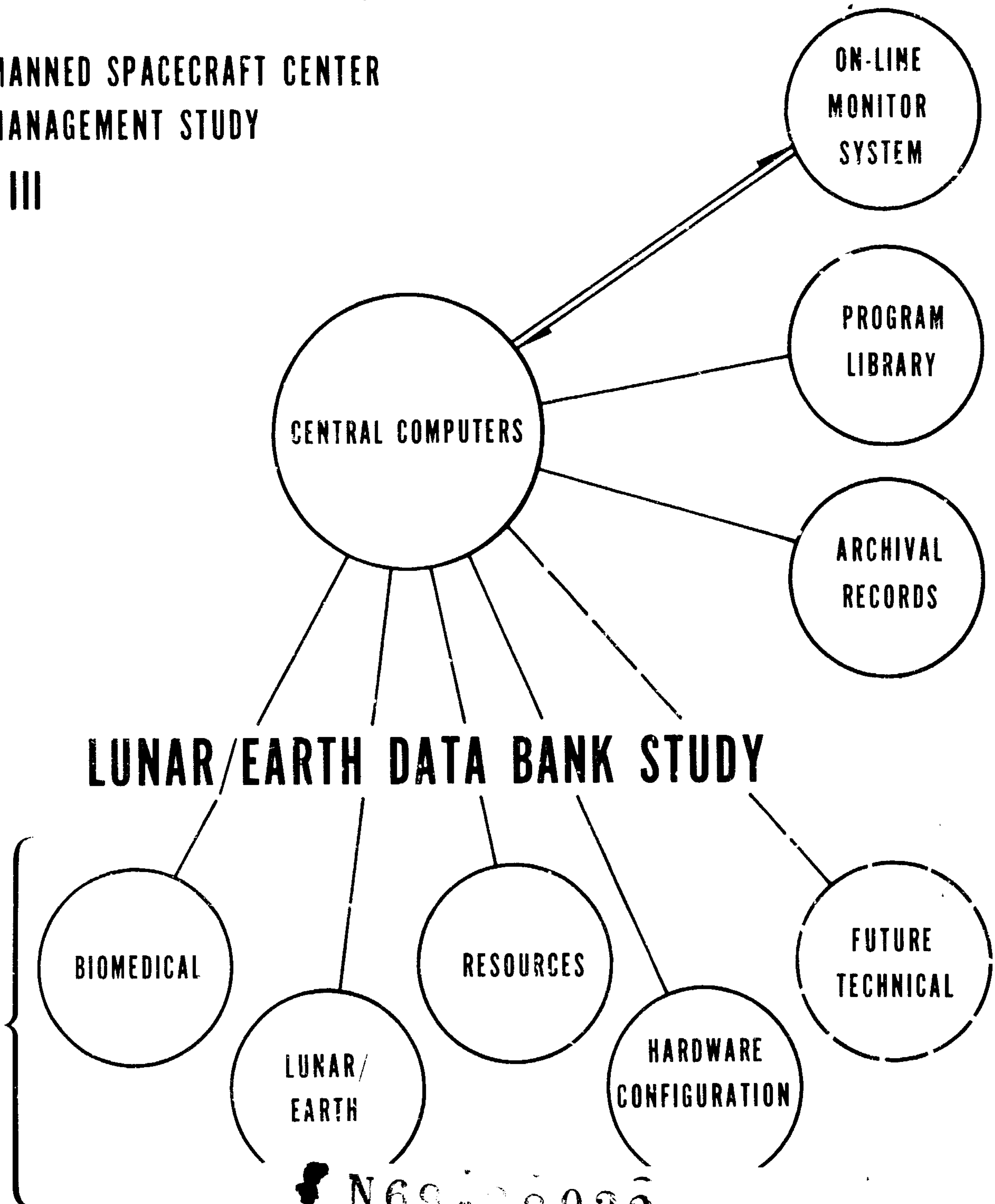
NASA MANNED SPACECRAFT CENTER
DATA MANAGEMENT STUDY
VOL III

LUNAR/EARTH DATA BANK STUDY



Lockheed

PALO ALTO RESEARCH LABORATORY

LOCKHEED MISSILES & SPACE COMPANY • A GROUP DIVISION OF LOCKHEED AIRCRAFT CORPORATION
PALO ALTO, CALIFORNIA

NASA/MSC Data Management Study
Contract Lockheed Electronics Company
ITA 7-0538 for NASA/MSC Information
Systems Division

LUNAR/EARTH
DATA BANK STUDY

N-71-68-1 January 1968

Prepared by:

R. K. Summit
Lunar/Earth Task Leader

Approved by:

C. E. Duncan
Project Technical Director

Information Sciences
Electronic Sciences Laboratory
Lockheed Palo Alto Research Laboratory
LOCKHEED MISSILES & SPACE COMPANY
Palo Alto, California

FOREWORD

This document, which discusses the requirements and specifications for a Lunar/Earth Data Bank System, is one of a series of reports written as part of the Manned Spacecraft Center Data Management Study under Contract NAS 95191. To enable the reader to consider this report in the context of the entire study, the reports are listed here:

MSC Data Management Guidelines Study
Biomedical Data Management Study
Lunar/Earth Data Bank Study
State-of-the-Art Survey

The reader desiring greater insight into any of the above areas is referred to the appropriate report.

The Lunar/Earth Data Bank Study team included R. K. Summit, Task Leader, S. R. Levine, S. Elster, C. H. Marshall, W. Simpson, L. J. Raffle, and M. Bernstein. The study team would like to acknowledge the friendly, cooperative, and informative nature of the interviews granted by the Manned Spacecraft Center personnel, and the helpful assistance provided by the personnel of Lockheed Electronics Company.

PRECEDING PAGE BLANK NOT FILMED.

SUMMARY

The contemplated use of Earth-orbital satellites in an Earth resources survey program to observe surface phenomena suggests a variety of potential applications to study natural resources in water, air, and land environments. Data thus far gathered from remote sensing of the lunar surface will greatly assist in the final selection of landing sites for the planned manned lunar landings. In both cases the variety and complexity of the data collection media and the data content require that formal procedures of data handling and control be established to ensure maximum data utilization. Too frequently, past emphasis has been directed at the collection of data with inadequate planning for the processing, analysis, storage, retrieval, and dissemination of such data after collection. This report represents an investigation of the information requirements of potential lunar and Earth data users, and the conceptual design of a Lunar/Earth Data Bank system. The described system incorporates the latest advances in information technology to achieve maximum data utilization with minimum user inconvenience.

Current NASA remote sensing programs include both aircraft and spacecraft activities, the former being conducted to test and evaluate equipment and techniques for later planned spacecraft missions. At present, principal investigators and other agencies requesting data deal directly with the program office or a designated service facility to obtain needed data. Responses are manually processed, and are provided on a no cost or minimum cost basis. The potential user must request specific, known material with few exceptions. The volume and potential usefulness of the data from future missions will preclude such specialized handling and distribution.

If information so collected is to be effectively utilized, the potential user must be informed of its existence, and must be provided a means to identify and obtain the data he requires. With source data as diverse in content, form, and format as those anticipated from the NASA programs, a storage and retrieval system is indicated which will include provision for physical storage of source data, descriptive cataloging of source data, dissemination of information describing the availability of source data, means for identifying relevant source data, and some means for reproducing and distributing requested source data in response to users' needs.

To meet these requirements, a Lunar/Earth Data Bank system is suggested which includes subsystems for input processing, information storage, inquiry processing, and reproduction and dissemination. System inputs originating from NASA programs as well as outside sources take various forms including magnetic tape, film, maps, reports, data forms, and prints. The Input Processing Subsystem is responsible for review, acceptance, and initial processing of all data to be included in the proposed data bank. After acceptance, source data are cataloged, and selected items are reproduced as microimages in the form of unit records. The cataloging process produces citations which characterize the contents of the source data with descriptors and keywords which become the basis for indexes used in the inquiry subsystem. These three types of items – source data, unit records, and citations – are passed to the Information Storage Subsystem for physical, image, and digital storage, respectively. The Inquiry Subsystem allows the user to specify his information needs relative to the contents of the records he wishes retrieved, utilizing an input/output terminal coupled to a computer. This interactive mode of operation allows the man to couple his intellectual powers of recognition and selection to the data processing power of the computer in the identification and selection of relevant source data. The Reproduction and Dissemination Subsystem distributes announcement information regarding the availability of source data, produces specified unit records, and responds to requests for specific material by supplying original source data or reproduced copies thereof to requestors. The system as proposed allows for centralization of storage and decentralization of inquiry terminals to provide a maximum of access convenience with a minimum of file maintenance problems.

Because of the state-of-the-art nature of the proposed system, it is recommended that an implementation program be initiated. This program would include a prototype phase followed by an operations phase. Knowledge and experience gained in the prototype environment will help to crystallize the requirements for the large-scale operational system. It is recommended that the prototype phase be implemented on a data base of existing source data.

The final section discusses a number of additional considerations and studies which should be conducted in support of the general problems associated with remote sensing technology. In particular, functions involved in the acquisition, processing, and analysis of remote sensor data are stressed. Two additional recommendations are made: one for the establishment of a Remote Sensing Technology Laboratory which would develop integrated, highly automated acquisition, processing, and analysis systems; and another for the development of techniques for dissemination of Lunar/Earth data and the education of potential users in the exploitation of these data.

PRECEDING PAGE BLANK NOT FILMED.

CONTENTS

ILLUSTRATIONS

PRECEDING PAGE BLANK NOT FILMED.

TABLES

Part I
DETERMINATION OF REQUIREMENTS

Part I of this report describes the flow of lunar sciences and Earth resources data within NASA, MSC; and between NASA, MSC, and principal investigators performing research in various application areas such as geology, hydrology, geodesy, agriculture, and forestry. Emphasis is placed on the information storage and retrieval aspects of the data flow in order to develop a set of requirements and general design criteria for a Lunar/Earth Data Bank.

Section 1
INTRODUCTION

1.1 BACKGROUND

The rapidly burgeoning science and technology of remote sensing is being used in the lunar sciences to acquire data relative to the composition of the lunar surface. This information will greatly assist in the final selection of landing sites for the planned manned lunar landings. Spacecraft have also been used to collect Earth-related data in the areas of meteorology, communications, and navigation. The contemplated use of Earth-orbital satellites in an Earth resources survey program to observe surface phenomena suggests a variety of potential applications to study natural resources in water, air, and land environments. Data from a spacecraft observation vantage point can provide substantial scientific and economic benefits in the areas of geography, cartography, and cultural resources; agriculture and forestry resources; geology and mineral resources; hydrology and fresh water resources; and oceanography and marine resources. Some specific applications which have been suggested include earth mapping, collection of agricultural census data, global surveys on crops, forest inventory, detection of sea ice, and the discovery of new mineral and fuel deposits.

The objectives of the various Lunar/Earth resources programs were assigned to NASA under the National Aeronautics and Space Act of 1958. These objectives are (Ref. 1):

- To develop and test procedures, instruments, subsystems, spacecraft, and interpretative techniques in the various applications areas
- To accomplish long-range studies of the problems and potential benefits resulting from utilization of space activities for peaceful and scientific purposes

To examine the extent of applicability of remote sensing technology, and to develop the technology itself, the Earth Resources Aircraft Program currently in progress is investigating and evaluating alternative operating conditions for remote sensing devices. In upcoming manned Earth-orbiting experiments, individual sensors will be calibrated, sensor performance will be evaluated, and imagery and other data will be collected over various established test sites. The collected data will be compared with surface data to determine feasibility and methodology of identifying surface phenomena. In later missions, data will be collected by aircraft and satellites as well as by surface stations using several different remote-sensing instruments ranging from high-resolution cameras to imaging radars. These data will be extremely useful to many Earth and lunar science disciplines and will form the basis for continuing operational programs.

1.2 PURPOSE AND SCOPE OF REPORT

Program emphasis thus far has been on the collection and interpretation of data with relatively less attention being paid to data organization, control, and wide dissemination. Many of the missions have contained specific experiments designed by a principal investigator who has sole responsibility for the initial interpretation of the data, and who controls the availability of the data. Other missions such as Ranger, Surveyor, Lunar Orbiter, and Gemini, however, have collected data of wide interest and varied application. Demand for these data has indicated that control using manual filing techniques is inefficient and leads to under-utilization of the data. The quantities and variety of data anticipated from currently planned programs further indicate the need for a formal system to acquire, organize, store, retrieve, announce, and disseminate needed data if full benefit is to be realized from the programs and if the requirements of potential users are to be served. It is the purpose of this report to provide design guidelines for a Lunar/Earth data bank system which will incorporate the latest advances in information technology to achieve maximum data utilization with minimum user inconvenience.

Figure 1-1 depicts the scope of this study (shaded portion) relative to the overall data flow from NASA programs. It can be seen that data collected from satellite missions, aircraft flights, and ground-based stations are passed to the data bank. The Input

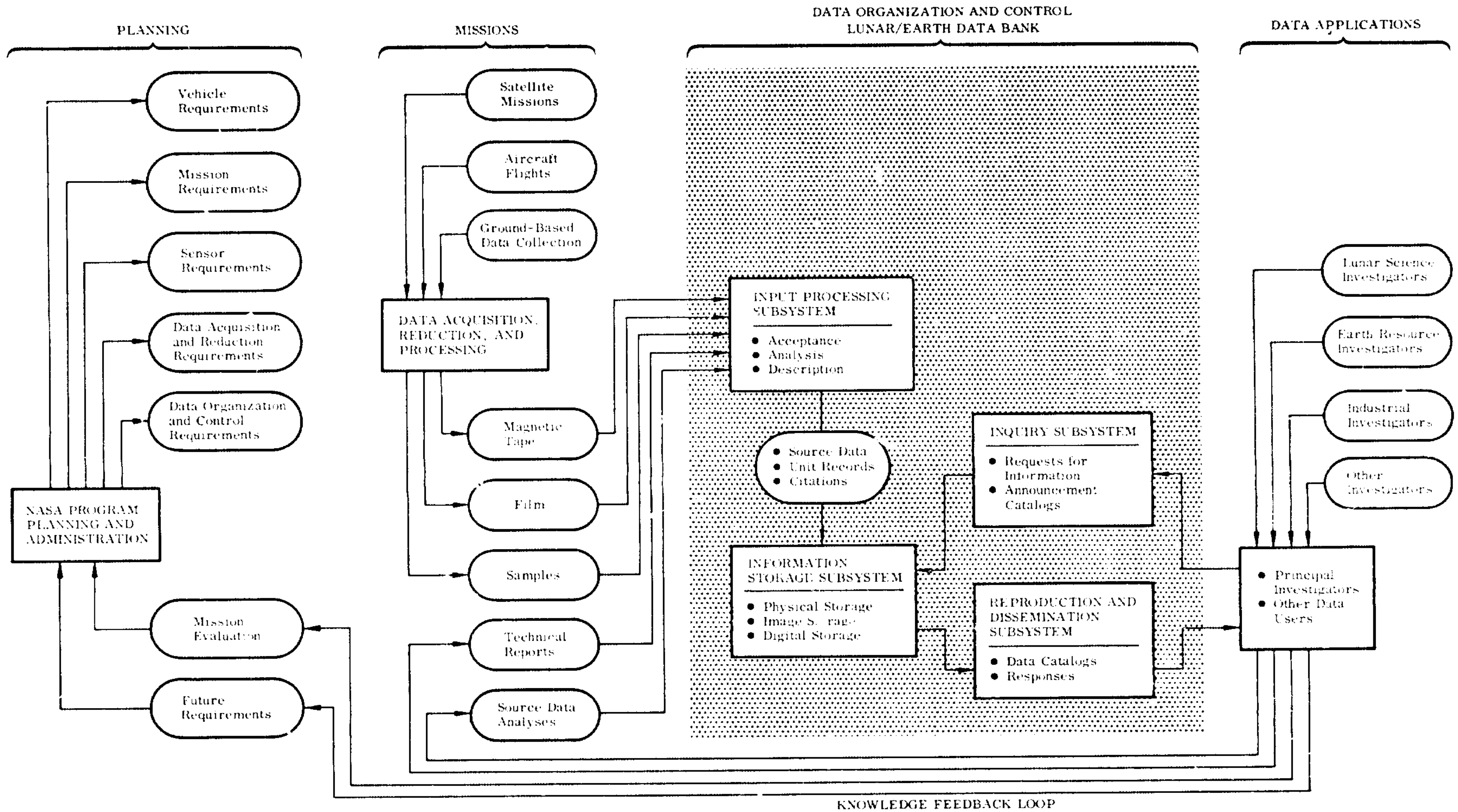


Fig. 1-1 Overall Data Flow Depicting Scope and Relationship of Lunar/Earth Data Bank System

Processing Subsystem is responsible for accepting, identifying, formatting, and cataloging these data. Secondary records consisting of descriptive citations and unit records (microfiche copies of formatted source data) are generated by this section of the data bank. Source data, unit records, and descriptive citations are passed to the Information Storage Subsystem for physical, image, and digital storage, respectively. Inquiries from users are processed in the Inquiry Subsystem and responses to requests are processed through the Reproduction and Dissemination Subsystem. This latter section is also responsible for disseminating announcement information concerning the availability of source data to prospective users. Results of user analyses of source data as well as mission evaluations and requirements summaries are fed back to the data bank, and also to the NASA planning and administration function. Important features of the proposed data bank system include:

- Centralized storage and decentralized inquiry
- Data structuring for multiple, cross-disciplinary usage
- User knowledge feedback to upgrade the data store
- Provision for coordinate rather than faceted retrieval
- Compatibility with existing NASA scientific and technical information system

Not included within the direct scope of the current study are several other important data handling areas associated with NASA remote sensing programs which are outlined in the last section of the report. The areas requiring additional consideration include data acquisition, time and feature correlation of multiple records, data compression, image enhancement, pattern recognition, automated photogrammetry, and photometrics. Results of research and experimentation in these areas of remote sensing technology can substantially improve the quality and utility of the data controlled by the proposed data bank system.

1.3 METHODOLOGY AND REPORT ORGANIZATION

To perform the study, a team of specialists was assembled which included a research scientist in information technology, a staff scientist in remote sensing and photo technology, a data systems engineer, a geologist, an astronomer, and two members with general operations research backgrounds. It was decided that the study would consist of two phases: definition of user requirements, and conceptual design of a Lunar/Earth data bank system. The basic tool used in the definition of user requirements was the interview, with interview sessions limited to two or three persons whenever possible to keep the inverviews informal and to avoid digressive conversation. Interviews were conducted at MSC with representatives of the Management Services Division, Photographic Technology Laboratory, Test and Operations Office, Applications Project Office, Mapping Science Branch, Lunar Receiving Laboratory, and the Information Systems Division. These groups are included in the Science and Applications Directorate, Engineering and Development Directorate, and the Administration Directorate. From these interviews, a description of current NASA MSC Lunar/ Earth data flows was developed which appears as Section 2.

Section 3 synthesizes Lunar/Earth information requirements into a set of overall guidelines which serve as the basis for the Lunar/Earth Data Bank System design concept presented in Section 4. An implementation program is outlined in Section 5. Section 6 concludes the report by discussing additional considerations and investigations relating to the total data handling problem associated with remote sensing programs. A reference list is included as Section 7. Appendices provide detailed discussion of the information requirements of Lunar/Earth data bank users (Appendix A), indexing concepts (Appendix B), and DIALOG, the Lockheed on-line information retrieval language (Appendix C).

Section 2
NASA LUNAR/EARTH DATA FLOWS

2.1 INTRODUCTION

The description of the NASA Lunar/Earth data flows is presented here to establish the present status of the available data services in support of the lunar sciences and earth resources programs. Appendix A contains a supplemental discussion of Lunar/Earth disciplines and users' information requirements. No evaluation of the adequacy or efficiency of these operations is intended.

The scope of the Lunar/Earth Data System study is to determine the data requirements for a Lunar Sciences and Earth Resources data bank. Additional requirements arise from the experimental payloads which support the activities of principal investigators in the scientific community at large. The MSC at Houston has cognizance of data generated by the Lunar/Earth applications programs. This jurisdiction presently includes operations and services supplied by Langley Research Center (LRC), Aeronautical Chart and Information Center (ACIC), Army Mapping Service (AMS), the United States Geological Survey (USGS), and laboratories operated by the principal investigators. The facilities at MSC provide parallel processing support in most areas plus detailed services for agencies requiring data.

The present data-processing functions at MSC principally support the program offices for mission evaluation and analysis. The organizational structure provides for program management through the office of the Director of Science and Applications. The office of the Director of Engineering and Development provides the principal support in electronic data processing and gives computational support.

Principal investigators, and other agencies requesting data, must deal directly with the program office to obtain support from the designated service facility, i.e., the photo laboratory or the data reduction center.

2.2 DATA INPUT INTERFACES

The data inputs for processing and handling encompass many programs. Data flows described here fall into two general categories – the Earth Resources Program and the Lunar Science Program. The objectives in these two areas differ widely; however, a large overlap exists in the techniques utilized for data acquisition and processing.

The Earth Resources Program generates data from four basic sources. These sources are shown in Fig. 2-1 as aircraft, spacecraft, test site, and laboratory calibrations. Data acquisition by the Earth-orbiting satellites consists of telemetry data recorded on magnetic tape at the various recording sites plus the video transmission of photographic data at those receiving stations equipped for recording. In the case of recoverable instrumentation packages and manned spacecraft systems, the data are processed as shown in Fig. 2-1.

The Earth Resources Program data flows are extremely complex because of the relative ease with which sensor data may be obtained. Under special circumstances, principal investigators maintain their own processing facilities. MSC is coordinating these activities, and has cognizance over these data, particularly where the external data are obtained over NASA-designated sites. MSC provides a clearing-house service for the data handling, processing, duplication, and storage in cooperation with the principal investigators.

MSC maintains data processing facilities for magnetic tape recordings – voice, video, or telemetry – and a photographic laboratory capable of processing an almost infinite number of film sizes and film types. The data unit coordinates these activities and is responsible for fulfilling the distribution requirements and services of special requests. The Technical Information Center provides an indexing and cataloging system for the reports and supporting documents.

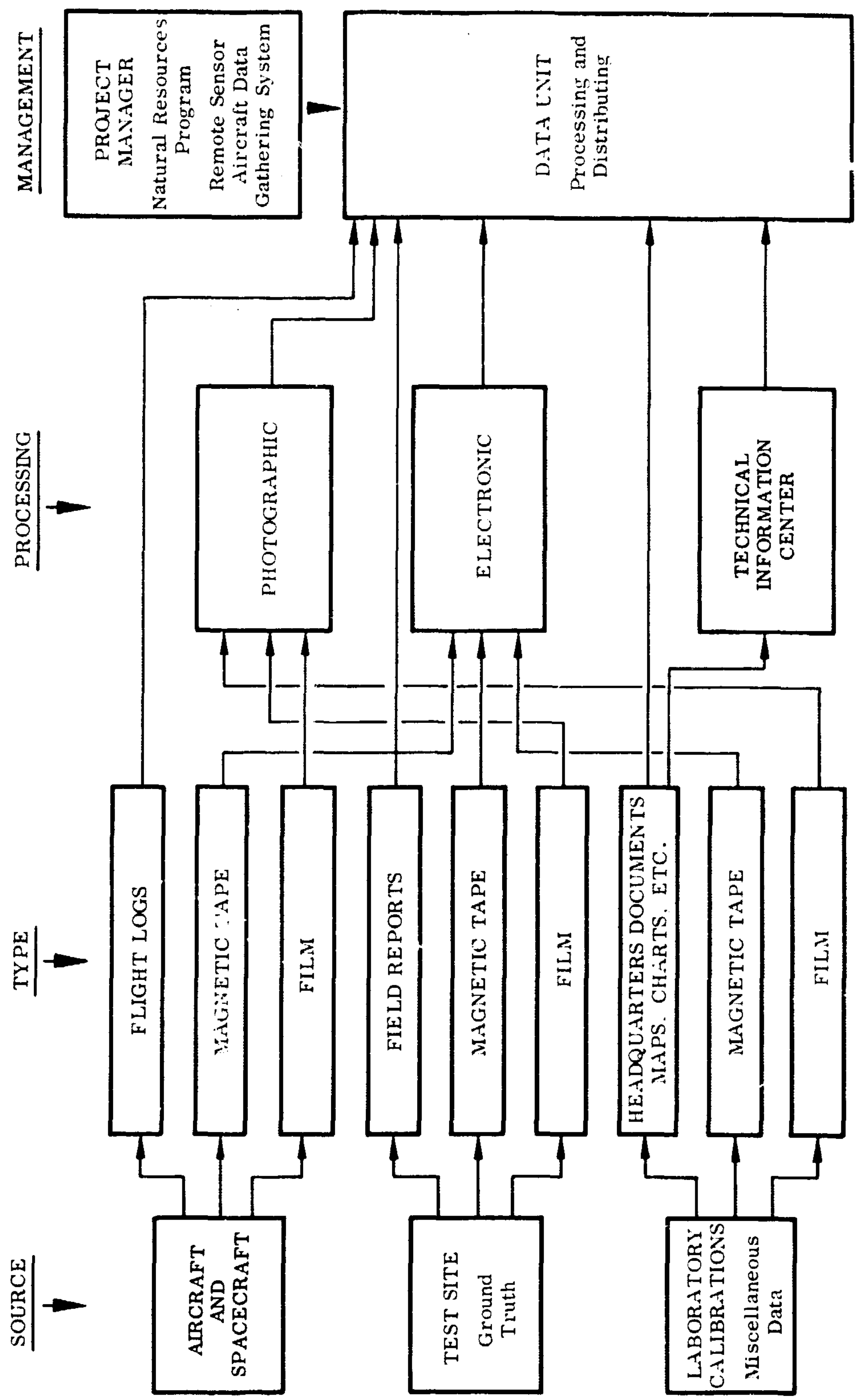


Fig. 2-1 Earth Resources Data Flowchart

Lunar Science Program data flows are more disciplined because of the necessary rigorous attention to the constraints of the communications network furnished by the Deep Space Instrumentation Facility (DSIF). Further, the ground reconstruction equipment (GRE) produces a single size output film which minimizes the number of film formats and types handled. Future requirements for supporting manned spacecraft operations will impose additional processing requirements. The present flows support the Lunar Orbiter Program.

Figure 2-2 shows the lunar orbiter photographic system. The spacecraft photographic system flow is useful in identifying the environmental features at the time of exposure. This information is annotated to the film during the ground reconstruction of the photographs. For photographic data, the DSIF at Goldstone, California, Madrid, Spain, and Woomera, Australia, are equipped with dual GRE. The output of the GRE is a 35-mm filmstrip which is processed by Eastman Kodak onto 9.5-in. film. The initial data reduction or "screening" has been performed at LRC by screening teams composed of members from participating agencies. Simultaneously, other support functions were performed at MSC, ACIC, AMS, and the USGS.

The Lunar Science Program inputs include data from other programs such as the Explorer and Surveyor series and will include the data from the Apollo program. The MSC activities provide the basic support for each program office.

2.3 DATA OUTPUT INTERFACE

The data outputs can be classified into three groups independent of the nature of the data. The first group is represented by the specific Program Office. The Program Office is responsible for the mission analysis and data quality assessment. The Program Office also supplies team members to the screening and analysis activity. The second group receiving data outputs is the principal investigator or coinvestigator. In some cases the Program Office and the principal investigator may be the same, particularly for the experimental data. The third group of outputs represents all other interested parties. In this area there are two principal groups, the first being

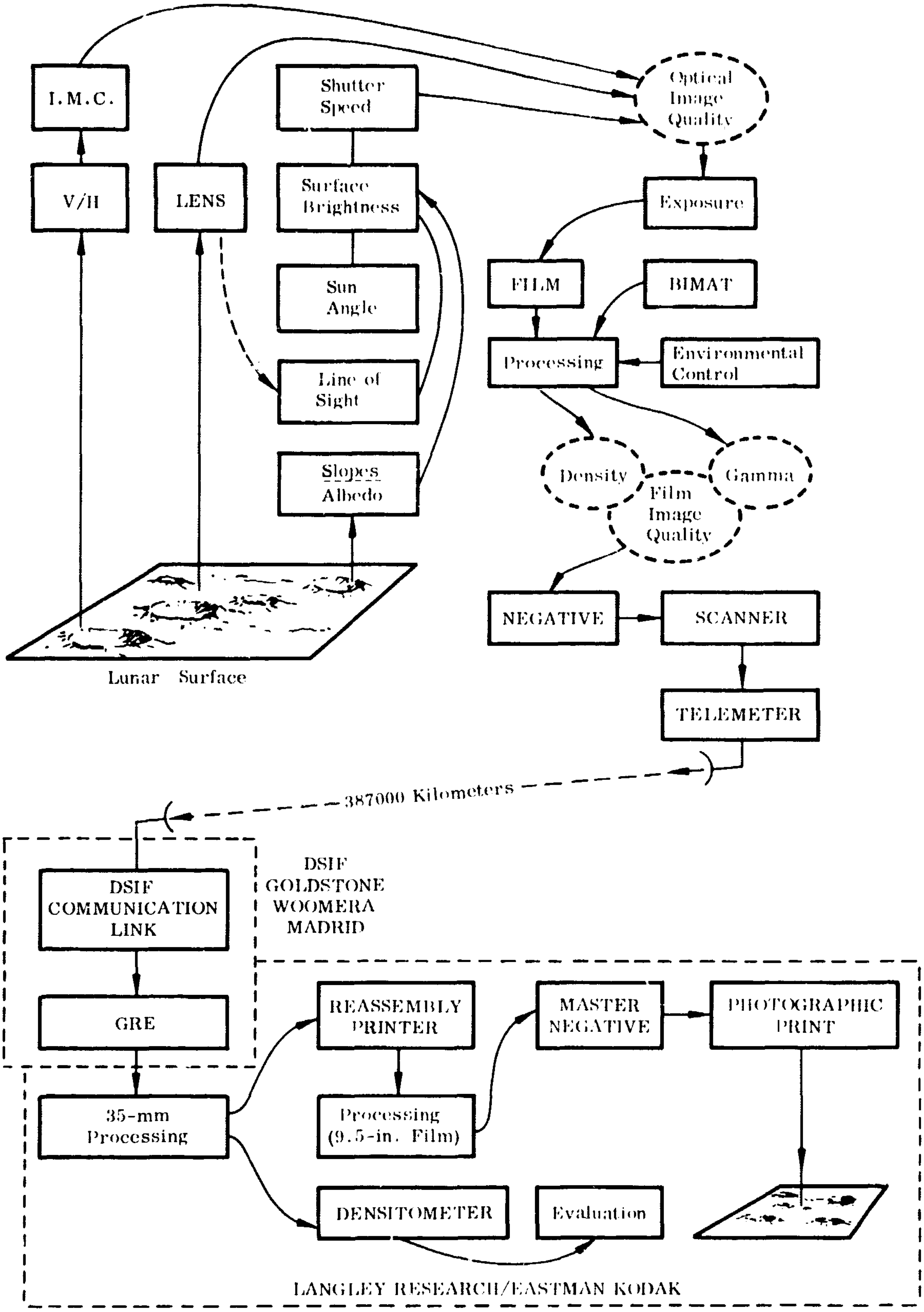


Fig. 2-2 Lunar Orbiter Photographic System

other interested investigators and the second the general public. The general public is serviced through the Public Affairs Office (PAO).

2.4 DATA PROCESSING ACTIVITIES

The available input data and the specified outputs generally define the required data flows in support of the Lunar/Earth data system. The existing flows reflect an organization which has evolved based on the availability of facilities and personnel of the required disciplines. The description of these data flows is shown in Fig. 2-3. The functions shown are located at MSC; however, duplicate capabilities for selected functions are operated by principal investigators and by other research centers. The Director of Science and Applications has the responsibility for the Lunar Sciences and Earth Resources data management. The nature of the data has resulted in the Mapping Sciences Branch assuming the data analysis and control responsibility for the Lunar programs, with the Mission and Data Management Office providing the basic support to the Earth Resources programs.

The Science and Applications Directorate utilizes the services provided by the Administration Directorate, namely, the Technical Information Center (TIC) or library and the Photographic Technology Laboratory, and the service provided by the Engineering and Development Directorate for data reduction support and storage facilities for magnetic tapes and associated documentation.

The Photographic Technology Laboratory and the Central Metric Data Files operations require further description. These two centers provide the interface points for the data inputs. The Photographic Technology Laboratory (PTL) processes the film from the Earth Resources, ground truth, stationary remote sensors, aircraft, or spacecraft operations. The PTL also receives processed film or copies of film from the principal investigators' laboratories or from the LRC/EK produced outputs. Similarly, the PTL also stores reproducible photo data from Gemini, Ranger, Surveyor and other NASA programs.

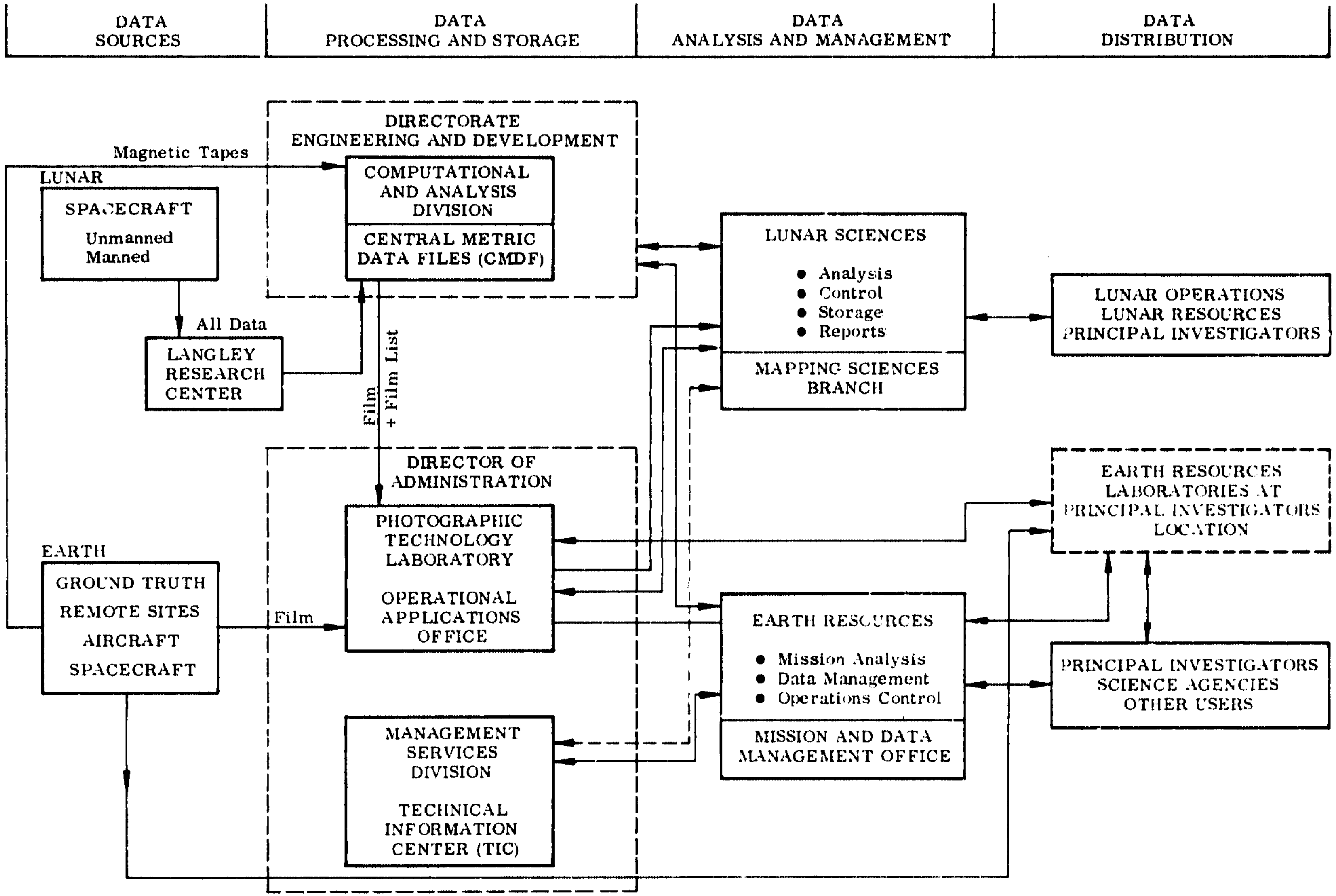


Fig. 2-3 Lunar/Earth Data Support Systems

The Central Metric Data Files (CMDF) are maintained by the Computation And Analysis Division (CAAD) to support the data processing functions. The data processing facility provides support data for mission evaluation. As a byproduct, the CMDF has established an efficient method for the logging of source data and maintains an accession list of these data. As a result, the CMDF receives the Lunar Orbiter source data from LRC for logging. The magnetic tapes are processed to extract event data and the film and annotation list identifying the film is forwarded to the PTL. The CAAD provides video and telemetry data reduction services for Lunar/Earth Resources operations.

The Test and Operations Office publishes an accession list for the Earth Resources Aircraft Program Data Bank. The information so cataloged is divided into three major categories as follows:

(1) Technical Documents
(2) Photographic Data
(3) Magnetic Tape Data

The three categories correspond to the TIC, PTL, and the CMDF data types.

The Data Management Office is presently implementing an Interim Information Storage and Retrieval System for the natural resources data.

2.5 DATA ACTIVITIES

The total data activity included in support of the Lunar/Earth programs can only be approximated. This is due to the fact that the instrumentation requirements must be adapted to changing requirements. The many principal investigators representing many areas of research continually establish new requirements. The instrumentation to be used in spacecraft operations has not been determined at this time. These instruments will be similar to those used in the aircraft program; however, they must be adapted to space environment. New instruments are being designed to meet the requirements of the space program. They are designed to take maximum advantage of the space vantage point and the larger surface area coverage contained in a single image.

2.5.1 Aircraft Program

The list of aircraft program sensors and instruments shown in Table 2-1 is representative of the variety of systems employed. As the Earth Resources program is expanded, new techniques will be employed necessitating more aircraft and/or Earth-orbiting vehicles.

The data acquired and the types of products to be derived from the aircraft program acquisition equipment are as follows:

(1) Film imagery from the RC-8 and T-11 cameras will be reproduced as duplicate positives and duplicate negative and/or prints as requested.

(2) Film imagery, from the multiband cameras, will be produced as black and white prints and duplicate positives or negatives. After analysis, certain frames may be enhanced electronically, optically, or during additional photographic processing.

(3) Film imagery (70 mm) from the dual-channel IR imager, the Reconofax IV IR scanner, and the RS-7 IR imager will be produced as black and white duplicate positives. After analysis, film strips of selected areas will receive electronic enhancement and/or selected paper enlargements will be made.

(4) Film imagery (35 mm) from the AAS-5 (IR/UV) will be produced as black and white duplicate positives. After analysis, film strips of selected areas will receive electronic enhancement and/or selected paper enlargements will be made.

(5) DPD-2 side-looking radar slant-range imagery is electronically converted into optical imagery while in the aircraft, and recorded on 5-in. roll film. Black and white duplicate positive rolls will be produced from this distorted slant-range imagery. Occasional electronically rectified (ground range) prints will be produced as requested.

(6) The MR-62 and NR-64 radiometers will record amplitude versus time data on one or more channels of magnetic tape, as well as on a stripchart. The magnetic tapes and the graphs will be the end products delivered to principal investigators.

Table 2-1

AIRCRAFT PROGRAM SENSORS AND INSTRUMENTS

Equipment	Recording Medium
Sensors on Convair 240A	
RC-8 Camera	9.5-in. film
Multiband Camera (Itek 9 Lens)	70-mm film
Reconofax IV IR Scanner	70-mm film
AAS-5 (IR/UV) Imager	35-mm film
T-11 Camera	9.5-in. film
13.3 GHz Scatterometer Radar	Magnetic tape
MR-62 Microwave Radiometer, 2 Channel	Magnetic tape and strip chart
MR-64 Microwave Radiometer, 2 Channel	Magnetic tape and strip chart
Instruments on Convair 240A	
Automatic Data Annotation Set (AN/ASQ-90) (When Acquired)	Magnetic tape and film
Precision Instruments Tape Recorder	
Photo-Data Panel	35-mm film
Haydon Time-Code Generator (IRIG)	Magnetic tape
Sensors on NP3-A Aircraft	
RC-8 Cameras (2)	9.5-in. film
Multiband Cameras (4 KA 62's)	5-in. film
Boresight Cameras (4)	35-mm film
Microwave Imager	70-mm film
Single-Channel IR Imager, RS-7	70-mm film (eventually magnetic tape)
DPD-2 Side-Looking Radar	5-in. film
400 MHz Scatterometer Radar	Magnetic tape[a]
13.3 GHz Scatterometer Radar	Magnetic tape[a]
Infrared Spectrometer	Magnetic tape[a]
Dual-Channel IR Imager	70-mm film
IR Radiometer	Magnetic tape[a]
Barnes IR Thermometer	Magnetic tape[a]
Microwave Radiometer	Magnetic tape[a]
Laser Altimeter	Magnetic tape[a]
Instruments on NP3-A Aircraft	
ADAS System (AN/ASQ-90)	Film and magnetic tape
Ampex AR-1600 Tape Recorder	Magnetic tape[a]
Haydon Time-Code Generator	Magnetic tape[a]

(a) Common 14-channel tape.

(7) The four bore-sight cameras will record on 35-mm film the direction that the aircraft sensor antennas are pointing. This black and white film record will be produced as duplicate positives. The 35-mm recording will accompany each magnetic tape of the sensor that was recording through that particular antenna.

(8) The 13.3 GHz and the 400 MHz scatterometers record the phase and amplitude data on one or more channels of magnetic tape. Stripchart profiles, 4020 CRT XY plots, and/or digitized magnetic tape will be produced from the original record.

(9) The infrared spectrometer collects radiometric values versus wavelength, which are recorded as analog signals on magnetic tape. The final product would be CRT plots and/or digital magnetic tapes.

(10) The output of one or more of the following data recording instruments on the aircraft is required for data reduction and correlation with the other sensors. Their output and a time code will be recorded on the film imagery or on one or more of the 14-channel magnetic tapes. After reproduction and duplication they will accompany the data from the other sensors.

   (a) The Automatic Data Annotation Set, AN/ASQ-90, records coded time, identification, and navigation signals on the film and/or on one or more channels of the 14-channel magnetic tape.
   (b) Photo-Data Panel records aircraft sensor instrumentation panel on 35-mm film.
   (c) The laser altimeter records aircraft altitude above ground on magnetic tape.
   (d) The Barnes IR Thermometer records IR background temperature.
   (e) The Haydon Time Code Generator produces a time code, which is placed on magnetic tape for time correlation of the various sensors. This is a highly desirable feature for principal investigators, since no sophisticated equipment is required to read it out.

2.5.2 Spacecraft Program

The exact instrumentation or specifications for the various spacecraft sensor packages has not been determined firmly enough as of this date to list separately. They may be covered in general as follows:

(1) UHF audiofrequency wavelengths will be recorded on magnetic tapes, which will be telemetered to Earth and reproduced as 4020 CRT X-Y plotter profiles or as digitized magnetic tape.

(2) The IR spectrometer will record preselected spectra on magnetic tape. Data will probably be recorded on 10 bands in the 15 or 4.3-μ region over designated portions of the Earth. On unmanned missions the data will be telemetered to Earth and produced as either digital tape or ADP Computer readouts. Manned missions will return the original tapes to the Earth, where they will be converted to digital tapes on ADP Computer readouts.

(3) The Dielectric Camera is essentially a camera with a dielectric recording feature wherein imagery is stored on dielectric plate in a vacuum. The data are telemetered to Earth when the satellite is within transmitting range of a NASA tracking facility. The telemetered imagery will be duplicated, stored, and distributed as magnetic tape. When processed through GRE it may be converted to photography before distribution.

(4) A UV camera will be carried on manned missions and returned at the same time as the astronaut. Except for special processing, the film will be handled in much the same manner as other camera film.

(5) Scatterometer, radiometer, laser altimeter, microwave radiometer, and spectrometer data will be collected as analog signals on tape. Stored information will be telemetered to Earth where it will be converted to digital tape. The final product to be given principal investigators will consist of digitized magnetic tapes, CRT X-Y plotter profiles, or strip charts along with appropriate spacecraft support data.

(6) Metric cameras will be carried on manned missions and the recorded imagery will be returned to Earth in film cassettes at the time the astronauts are returned. This film will be processed and distributed in the normal manner for high-resolution metric photography.

(7) A stellar camera will be carried on manned missions and will be returned at the same time the metric camera film is returned. The final product from this imagery will be camera/spacecraft orientation data for each exposure of the metric camera.

(8) Hassleblad and/or other handheld cameras will produce color photos of unusual phenomena or targets of opportunity. They will be processed and distributed in a manner common to this type of imagery.

(9) A multiband camera will be carried on manned missions. Multiband and 70-mm imagery will be returned at the same time as the astronauts. Film will be processed under a regimen designed for highly calibrated imagery.

(10) Sidelooking imaging radar can be flown on either manned or unmanned missions. The imagery from manned missions would be returned in the usual manner. For unmanned missions, the data must be telemetered to Earth or recovered from orbit and reconstructed as slant-range imagery. The imagery will be duplicated and distributed as roll film and/or selected image-enhanced photographs.

(11) IR and/or UV scanner imaging systems will record imagery on film and/or analog magnetic tape. The film and the magnetic tape will be returned in the usual manner. Analog data on unmanned missions may also be telemetered to Earth and reconstructed on ground reconstruction equipment. The final product will be film images with selected electronic enhancement. Digitized tapes of selected areas may be required for special purposes.

(12) The day-night camera is essentially a low light level TV camera in which imagery is stored as an analog signal on magnetic tape. Data are telemetered to Earth at the appropriate NASA facility. The imagery will be reconstructed in GRE as digital magnetic tape or photo imagery.

(13) The micrometeorite detection system records the occurrence of micrometeorite impacts. This system stores the speed and direction of the particles, with respect to an array element of sensor panels modeled externally to the spacecraft. These data are stored on digital magnetic tape until they are telemetered to Earth on command from a NASA tracking station. The resulting data will be tabulations of two 24-character words for each impact recorded.

Estimates for the volume of data requiring processing must be made in order to determine the necessary supporting facilities. Realistic estimates have only been made for the Earth resources programs reflecting data from the aircraft and Earth spacecraft missions. These estimates are summarized in Table 2-2 for the aircraft program and Table 2-3 for the spacecraft program. Similiar estimates must be made for all Lunar data requirements.

Table 2-2

ESTIMATES OF ANNUAL DATA FROM AIRCRAFT PROGRAM

Recording Format	Footage Reel	Number of Reels	Total Footage (thousand ft)
1-in. Magnetic Tape	3600	1500	5400.00
35-mm Film	180	63	11.34
70-mm Film	150	1200	180.00
70-mm Film	180	700	126.00
5-in. Film	200	360	72.00
9.5-in. Color Film	180	375	67.50
9.5-in. Color IR Film	180	375	67.50
9.5-in. B/W Film	180	54	9.72

Table 2-3

ESTIMATE OF ANNUAL DATA FROM THE EARTH RESOURCES SPACECRAFT PROGRAM

Electronic Processing
(All Data on 1-in. Magnetic Tape)

Sensor Data	Footage/Reel	Number of Reels	Total Footage (thousand ft)
IR Spectrometry	3300	6	19.8
Absorption Spectroscopy	3600	3	10.8
Passive Microwave Emission	3300	2	6.6
Scatterometer/Altimeter	3600	13	46.8
RF Reflectivity	1000	1	1.0
Micrometeorite Detection	3000	2	6.0

Major Photographic Processing Requirements

Sensor	Film Width	Footage/ Reel	Number of Reels	Total Footage (thousand ft)
Metric Camera	9.5 in.	1100	8	8.8
Multispectral Camera	9.5 in.	100	4	0.44
UHR Camera	70 mm	2200	2	4.4
Stellar Camera	70 mm	1500	2	3.0
IR Imager	70 mm	400	1	0.4
Panoramic Camera	5.5 in.	5500	2	11.0
Passive Microwave Imager	70 mm	50	1	0.05
Radar Imager	5.0 in.	5500	10	55.0
UV Absorption/Luminescence	35 mm	2100	2	4.2

Section 3
DATA BANK DESIGN GUIDELINES

3.1 INTRODUCTION

Section 2 and Appendix A discuss current Lunar/Earth data flows and users' information requirements. This section presents users' needs in terms of information retrieval; i.e., the identification and procurement of relevant information from a large store of potentially useful data. There are several characteristics that complicate retrieval of Lunar/Earth data such as:

- Source data are received and stored in a variety of forms, including rolls of film, positive prints, magnetic tape, data sheets, and documents.
- Content variety includes text, numeric data, images, and recorded signals.
- Potential users include professionals who represent a variety of disciplines and interests within disciplines.
- Potential users are geographically dispersed, yet require access to similar source data.
- Useful descriptive attributes include many unconventional types such as geographical position, environmental characteristics, and topographic features.

If the information from Lunar and Earth programs is to be effectively exploited, the user must be informed of its existence, and must be provided efficient means to identify and obtain the data he requires. Traditional library tools such as the card catalog and the reference librarian are inadequate to cope with this complex assortment of data and diversity of user requirements. Recent developments in information technology, utilizing third-generation computers, allow the design of powerful, interactive, man-machine systems which can substantially increase the utility of data bank information at a fraction of the cost of originally acquiring the data.

This section develops a set of guidelines or criteria which are used in the system design and implementation concepts presented in Sections 4 and 5. These guidelines are discussed within the following categories which relate to the necessary processing steps in a Lunar/Earth data bank system:

- Descriptive indexing of source data
- Physical storage of source data
- Announcement of the availability of source data
- Automatic or interactive processing of requests for data
- Distribution of information items required by user
- User feedback

3.2 DESCRIPTIVE INDEXING REQUIREMENTS

For purposes of both dissemination and retrieval of source items, it is useful to develop a standardized descriptive citation for each important unit of source data. Because these citations use consistent format and terminology, standardized announcement and retrieval procedures can be developed to control the diversity of content and physical form represented in the Lunar/Earth sciences source data.

The citation resulting from each unit of source data will include, as appropriate, the following descriptions:

- Physical form – magnetic tape, film transparency, positive print, document table of values, etc.
- Content – features identified and substantive information or data contained in the item
- Format – an indication of the general structure or layout of the item
- Collecting environment – instrument characteristics, vehicle characteristics, flight number, orbit pass
- Information environment – geographical location, time of observation, temperature, sun angle
- Related information – indication of ground truth and technical reports which refer or relate to the item

- Physical location – where and in what form the item can be obtained
- Reference or accession code – reference or identification nomenclature by which the item can be unambiguously identified

Each of these types of information requires the development of standardized descriptive terminology. The development, control, and use of such terminology is discussed in the DOD Manual for Building a Technical Thesaurus (Ref. 2). Representative terminology can be found in the Space Data Thesaurus (Ref. 3).

There are several problems relating to citation generation which require solution during a system design phase.

- Should descriptive cataloging and indexing be performed centrally and initially, or should the users be responsible for this function?
- To what extent can computer techniques be used for pattern recognition and feature description?
- Should all information be formally cataloged (described) and, if not, what are the criteria for inclusion of items?
- Which types of descriptive information should be included in indices and thus be directly retrievable?
- How can the user be utilized to develop cross referencing of information and descriptive terminology?
- Should the cataloging of Lunar/Earth sciences data be accomplished through the central NASA documentation control, or should it be done locally?
- What is the information item to be described, the report or each chapter in the report, the film roll or each frame in the roll, etc. ?

3.3 PHYSICAL STORAGE OF LUNAR/EARTH SCIENCE SOURCE DATA

Because of the variety of physical forms of information collected by remote sensing devices, the normal problems of physical storage of information are compounded. Traditionally, physical storage of information is accomplished either by grouping source material according to a particular characteristic of the material (e.g., subject category, document size or type, author), or merely by storing it in order of receipt. In the latter case, retrieval is solely dependent on a descriptive catalog of the information, whereas in the former case unidimensional retrieval can be accomplished by locating a desired subdivision of the physical store.

Although access to Lunar/Earth sciences data will frequently be by geographical location, the variety of source data media precludes physical organization on this basis. It thus seems that source data should be physically stored by media type. Retrieval then becomes largely dependent on a union catalog which contains descriptions of the source data material.

Problems associated with physical storage of source data include:

- Physical storage environment requirements for the various media (e.g., film, magnetic tape, etc.)
- Should physical storage be centralized or should it be associated with the means of reproduction or dissemination?
- Which generation(s) of source material should be saved (i.e., original telemetry versus reconstituted imagery)?
- On what basis should the store be purged?
- Should there be duplication of source items for archival purposes?
- Should continuous records (analog traces, film strips, etc.) be divided into unit records for control and distribution?

3.4 ANNOUNCEMENT OF INFORMATION AVAILABILITY

As long as there are relatively few data, or relatively few users, dissemination of the availability of information is accomplished by word-of-mouth and other informal means. The growing realization of the broad interest in imagery collected from orbiting vehicles resulted in the publication of a book of over 200 earth photographs collected during various Gemini missions (Ref. 4). Each photograph is annotated according to geographic location, camera type, gross topographic features, and time. Although the quality of reproduction is likely to be too crude for most scientific applications, the book assists the potential user in identifying which items merit further review. Without some means of description and announcement, each user must review the original source photography.

There are many system design decisions with regard to the dissemination problem.

- Should dissemination be accomplished through the NASA Scientific and Technical Information Division in parallel with technical literature announcement?
- Could micro-images serve to assist the potential user in selection of relevant imagery?
- Can announcement effectively be accomplished as with the Gemini photographs?
- Should all items in the data bank be generally announced; if not, what selection criteria should be employed?
- Should there be one standard announcement medium, or can the announcements be tailored to individual or group interests?

3.5 AUTOMATIC PROCESSING OF REQUESTS

With no indices and no organization to the information store, a user must exhaustively review each item to be confident he was not missing something relevant to a particular problem. It is the purpose of information retrieval systems to reduce the number of items which must be reviewed to satisfy a requirement.

To be responsive to potential users, an information retrieval subsystem should include the following features:

- The system should be reasonably easy for the nonspecialist to use.
- The system should assist the user in identifying the proper terminology to use in specifying his interest.
- The retrieval process should be flexible enough to allow a variety of search strategies to be employed.
- Given a particular item of interest, the system should assist the user in identifying related items in the collection.

- System response should be immediate and interactive rather than delayed and remote.
- All associated information relating to a particular need should be retrievable through a single information retrieval subsystem.

3.6 DISTRIBUTION REQUIREMENTS

Because of the variety of physical forms and formats of data anticipated from Lunar/Earth sciences programs, the final output from an information retrieval system can vary from source imagery in the form of film magazines, through analog or digital magnetic tape reels, to specially formatted compendia of data extracted from various source data records. At present there is little policy relative to distribution of informational items, and such policy should be developed as soon as possible.

At present, Gemini photographs can be ordered either from the Public Information Office of NASA Headquarters or the Public Affairs Office of the Manned Spacecraft Center. In the future, distribution and control of this information should probably be centralized and coordinated with the information retrieval subsystem.

A variety of distribution problems must be considered in the design of a Lunar/Earth sciences data handling concept.

- Under what circumstances (if any) will original or source data be supplied (as opposed to copies of source data)?
- What should be the pricing policy for supplying data and what physical forms should be offered?
- What copyright provisions should be made?
- Should material be loaned (with the accompanying problems of control) or should copies be made in every case?
- Should there be unrestricted distribution?
- Should distribution be coordinated with technical report distribution?
- To what extent can image transmission equipment be used in place of hard copy distribution?

3.7 UTILIZATION OF USER FEEDBACK

A Lunar/Earth data bank system should be designed to utilize the user both as a contributor and as a critic. As a contributor, the user could supply supplemental feature identification for images and technical reports resulting from his use of data bank information. As a critic, he can supply system evaluation information.

Problems related to utilization of users include:

- How can the user be effectively required to submit feature identification data on images supplied to him?
- Can the commercially oriented user be expected to identify features of a proprietary nature?
- Should Lunar/Earth sciences reports be processed through formal NASA documentation channels, or should they be specially processed at MSC?
- How can the user be utilized in developing descriptive terminology for features and annotation?
- How can user acceptance of the system be measured?

3.8 CONCLUSIONS

If information is to be effectively utilized, the potential user must be informed of the existence of the information, and must be provided with a means to identify and obtain the data he requires. With source data as diverse in content, form, and format as those anticipated from the NASA programs, a formal storage and retrieval system is indicated which will include provision for physical storage of source data, descriptive cataloging of source data, dissemination of information describing availability of source data, identification of subsets of source data items relevant to a particular need, and distribution of data items so identified. These problem areas are given consideration in the following sections of this report.

Part II

CONCEPTUAL DESIGN

Part II of the report indicates how the data requirements identified in Part I can be satisfied. First a Lunar/Earth Data Bank system is described regarding the organization, functions, and responsibilities of four subsystem elements: Input Processing, Information Storage, Inquiry Processing, and Reproduction and Dissemination. A 24-month implementation program is developed and costed which includes prototype and operational phases of 12 months each. Finally, a number of areas deserving additional study consideration are enumerated and discussed.

Section 4
LUNAR/EARTH DATA BANK SYSTEM – CONCEPTUAL DESIGN

4.1 INTRODUCTION

The utility of a data bank system depends equally on the volume and quality of data contained in the system, and on the capability of the system to retrieve particular data in response to the expressed needs of system users. Too frequently emphasis is directed at the collection of data with inadequate planning for the processing, storage, retrieval, and dissemination of such data. The most painstaking and costly effort can be expended in the collection and processing of data, but if equal consideration is not given to the indexing, storage, and retrieval functions, the data may lie dormant and unused.

This section derives from previously discussed user information system requirements (Appendix A) and describes the conceptual elements of a storage and retrieval system for Lunar/Earth data. No attempt is made to identify the functional elements presented with existing MSC organizational entities; rather, data bank functions are sequenced according to their precedence in the natural flow of data, and are grouped into subsystems on a basis of similarity of processing function. Which NASA organization would assume the responsibilities described within a particular subsystem is the proper consideration of a detailed systems specification.

Figure 4-1, which is the data organization and control section of Fig. 1-1, depicts the subsystem organization of the proposed Lunar/Earth data bank system as follows:

- Input Processing Subsystem
- Information Storage Subsystem
- Inquiry Subsystem
- Reproduction and Dissemination Subsystem

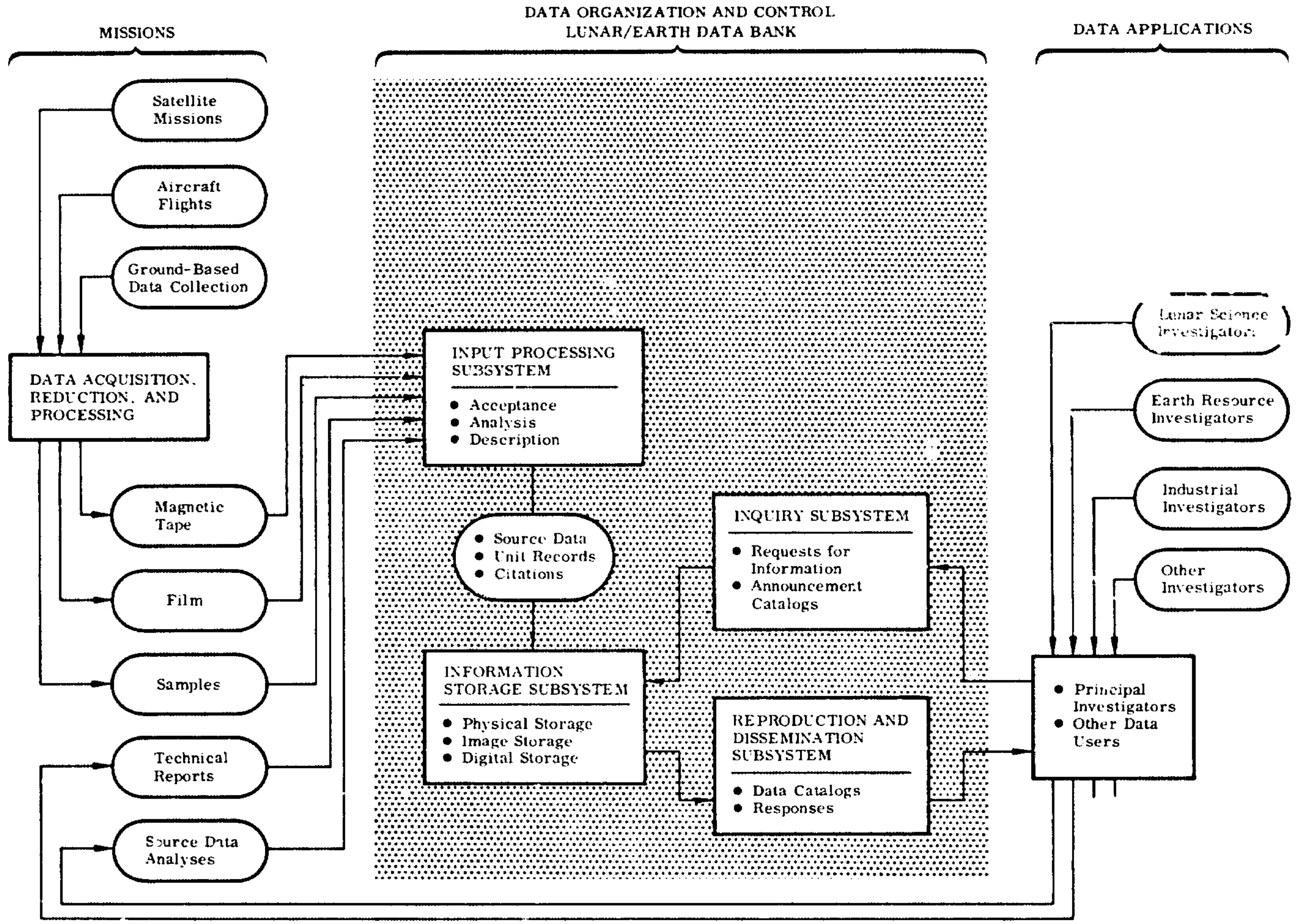


Fig. 4-1 Conceptual Data Flow, Lunar/Earth Data System

Inputs to the system originate from NASA programs, discipline-oriented investigators, industry research personnel, and other government programs, and take the form of magnetic tape, reels of film, maps, reports, data forms, and individual prints of film images. The suppliers of data to the system can be seen also to be users of data from the system. Just what data should be included in a data bank system and what the acceptance criteria for the data should be must be established by policy. The most expeditious way to render a data bank system ineffective is to inundate it with inadequately processed data of little general interest.

As the first point of contact for incoming data, the Input Processing Subsystem is responsible for review and acceptance of all data to be included in the Lunar/Earth Data Bank system, and for the initial processing of all data so accepted. Minimal standards for acceptance should require that the data include information relating to source, time of collection, geographical location represented (for images and maps), and collection media. The Input Processing Subsystem develops two products: citations and unit records. Citations are textual records describing physical units of source data, and unit records are microimage reproductions of source data. After this initial processing, source data, citations, and unit records are passed to the Information Storage Subsystem.

The Information Storage Subsystem provides physical storage for source data, image storage for unit records, and digital storage for citations and indexes. These indexes are produced by this subsystem from the citation records.

The Inquiry Subsystem represents the user/system interface functions. The objective of this subsystem is to provide the user means for identifying the subset of relevant source data items without having to exhaustively review the total collection of items. The inquiry subsystem will contain standing requests for information, but will also process special requests for data.

The Reproduction and Dissemination Subsystem is largely a service organization which supports the other subsystems. Responses to inquiries, as well as announcements of the availability of materials, are reproduced and disseminated by this subsystem.

The remainder of this section describes and illustrates each of the subsystems in greater detail.

4.2 INPUT PROCESSING SUBSYSTEM

The functional responsibility of the input processing subsystem is to review, accept, identify, analyze, describe, microimage, and submit for storage, selected data inputs from NASA programs and from extra-NASA data users.

4.2.1 Data Input, Review, and Acceptance

The types of data and conditions for acceptance of data for input to the Lunar/Earth data system should be rigorously defined by policy. Such policy should allow for a variety of data input, including:

Mission summary reports	Annotation data
Technical reports	Descriptive citations
Journal articles	Instrument traces
Mission data	Remote sensor imaging
Ground-truth data	Maps, charts

Physical forms of data are likely to include:

Magnetic tape	Film frames/prints
Rolls of film	Documents
Film strips/prints	Standard forms

Quality guidelines should spell out acceptance standards in terms of:

- Completeness of related information
- Image quality and sharpness
- General utility of material submitted
- Acceptable formats, sizes, length, etc.

It is expected that the acceptance policy will be modified as experience is gained in actual operation of the system. If such a data system is to be effective, it must be controlled, and must not be allowed to become a repository of data curiosities or an archival monument to ill-planned collection activity.

4.2.2 Data Identification, Analysis, and Description

Once accepted, the data must be divided into physical units and each such unit must be assigned an identification (or control) number. The identification number should indicate the year of acquisition and type of source data being identified, but beyond this should be a pure accession number. This number should be consistent in format to accession numbers already in use elsewhere in NASA. * An example of such a number is: 67T10015 where the first two digits indicate the year of acquisition, the letter is the source data type code (say, magnetic tape in this case), and the last five digits remain as a sequential serial within type and year (the 10,015th item of type T processed in year 1967). This identification or accession number serves to uniquely identify the defined physical unit of source data from that point forward.

Following identification, the unit of source data must be described (sometimes referred to as cataloging or indexing) in terms of form, content, format, originating environment, and related information. This descriptive information is developed in the form of a citation. (One example of a citation is a library catalog card.) Because the citation is a strictly formatted textual record, it lends itself to automated processing techniques, and serves as the basic record for all subsequent retrieval functions. The process of descriptive cataloging is discussed in detail in Appendix B and will only be summarily described here.

Information retrieval can be thought of as a series of screening levels, the first being directed at the total data store, and the final being the analysis of particular retrieved

---

*The accession number referred to is that used to identify reports announced in the STAR (Scientific and Technical Abstract Reports).

items of source data. The citation provides an intermediate product in the search, whose purpose is to assist the user in deciding the relevance of the source data it describes. This purpose defines the criteria for the information content of the citation. A citation should, as a consequence, contain any or all of the following information:

- Identification – accession number, mission, pass, item number
- Annotation – characteristics of the recording instrument, time of observation, sun angle, type of film, development characteristics, scale
- Geographical location – ground coverage of the image, political area
- Image content – cloud cover; major topographic features; identifiable geologic, oceanographic, hydrologic features; image clarity
- Related information – ordering procedure, price, availability, cross referenced data, etc.

Analysis and indexing can be done to any arbitrary degree of detail. In general, utility of indexing falls off (i.e., increases at a decreasing rate) and cost increases the more detailed and application-specific the analysis which is performed. It is assumed that information associated with the identification, annotation, and geographical location fields of the citation is supplied with the source data. As a result, the analysis question centers around the manner and degree of describing image content. It should be clear that gross features such as cloud cover, snow cover, image clarity, and surface type can be easily identified by unskilled, clerical personnel at low cost. If through such initial processing the investigator can screen out all source data images containing more than 70% snow or cloud cover, his time can be used to perform meaningful analysis of relevant material rather than to scan obscured images. Similar benefits accrue in the ordering and reproduction of image data. Initially the identification of image content should be limited to gross, application-independent features such as those indicated above. Such a proposal is not intended to limit the function of the input processing subsystem, but rather to initiate content on a solidly cost-effective basis. As requirements for additional content analysis are manifested, there is better justification for corresponding increases in and upgrading of professional staff.

System users can be utilized in two ways to increase the effectiveness of source data indexing. In the first case users should be requested to submit an indexing form along with any material provided by them for inclusion in the data bank. This technique has been used effectively by both NASA and Defense Documentation Center to obtain preliminary descriptive indexing of source material. Second, users can be requested to return an indexing form which would enumerate image content features identified during the course of their analysis of requested source data. In both cases, the submitted indexing would be reviewed by Input Processing personnel, and would be used as a supplement to other indexing. As a result the indexing quality and completeness would continually improve. Furthermore, the source data items in greatest use would receive the most thorough review and analysis.

It is anticipated that each type of source data may require a somewhat different citation format to effectively describe the contents of the material it represents, and the development of these citation formats is necessary prior to system implementation. To illustrate perhaps the most difficult type of citation, that describing imagery, a photograph of the San Francisco bay area was selected, and a representative citation was developed. These appear as Figs. 4-2 and 4-3, respectively. The entries in the citation record are discussed in detail in Appendix B, and its inclusion here is to illustrate general image indexing concepts.

By characterizing the important attributes of an image in a textual record (the citation), it is possible to apply a variety of proven data processing and text manipulation techniques to the retrieval of images. Computer intelligible indexes, for example, can be developed for each entry in the citation. Because these indexes can be searched rapidly and automatically by modern data processing equipment, the potential user has a powerful yet flexible means of specifying his requirements and identifying the corresponding set of relevant items from the data bank without having to exhaustively examine each source data item.

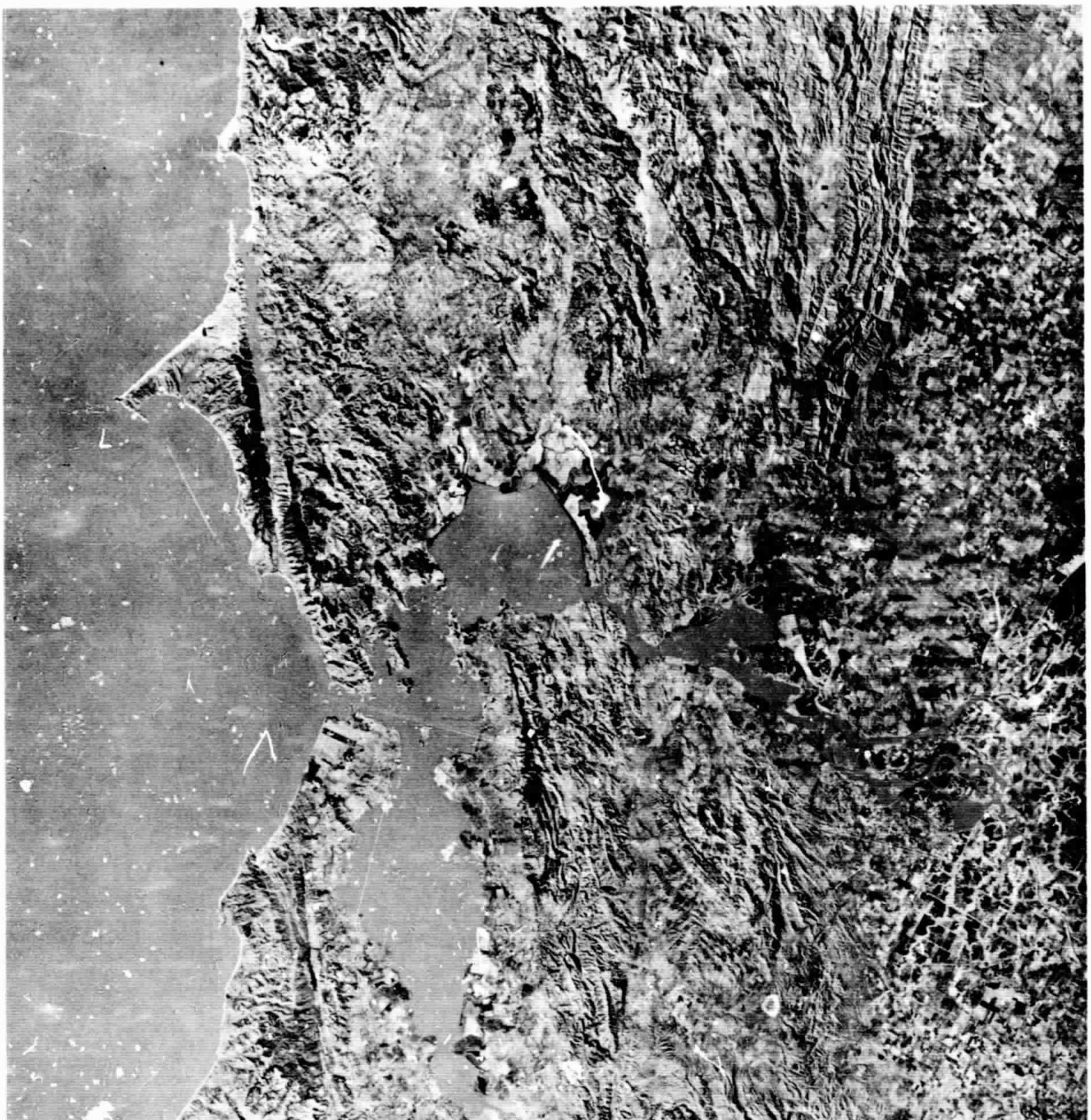


Fig. 4-2 Source Data Example

IDENTIFICATION DATA

Accession Number	Issue Date	Project	Date (UT)	Time (UT)	Mission Number	Pass (Orbit)	Frame	Recording Mode	No. of Times Requested
68 I 10063	1/09/68	A x z	7/07/67	2223	2188	513	1015	Photographic	413

GENERAL DESCRIPTION

Aerial photographs combined to simulate photograph of San Francisco Bay Area from a spacecraft at 300-mile altitude.

ANNOTATION DATA

Camera Type	Camera Number	Lens (efl) (in.)	Lens (efl) (in.)	Film Type	Spectral Range (mμ)	Filter Used (mμ)	Processing Method
KC-1	R-507	12	12	E/HSIR	540–900	A-25 (700–900)	D-19 (8 min)

Instrument Type	Instrument Model	Spectral Range (μA)	Tape Type Used	Storage Capacity (Char.)
N/A	N/A	N/A	N/A	N/A

Vehicle Track	Sun Alt.	Sun Azimuth	P. P. Pitch	P. P. Yaw	P. P. Roll	Vehicle Altitude	Scale
330°	045°	210°	30 min	1 deg	45 sec	30,000 ft	1/60,000

Remarks:

GEOGRAPHICAL LOCATION

WAC No.	Grid No.	Geopolitical Area	Area Cover (mi²)		N.W. Corner	S.W. Corner	N.E. Corner	S.E. Corner	Principal Point
364	R4, R5	Western USA (S. F. Bay Area)	12,500	Lat.	38° 25' N	36° 45' N	39° 05' N	37° 25' N	37° 53' 45" N
				Long.	123° 40' W	122° 45' W	122° 05' W	121° 00' W	122° 18' 18" W

Remarks: Positional accuracy given as: Corners to nearest 5 arc minutes
Principal Point to nearest 1 arc second

IMAGE CONTENT

Cloud Cover	Snow Cover	Water Cover	Land Cover	Image Quality
None	< 5%	30%	70%	Excellent

Remarks:

Other Features

1. San Francisco Bay
2. San Pablo Bay
3. Santa Clara Valley
4. San Joaquin Valley
5. Sacramento River Basin
6. San Andreas Fault
7. Pacific Ocean and coastline
8. Point Reyes
9. Coastal mountain range

PROCUREMENT INFORMATION

Order From	Data Forms Available	Source Data Location	Source Data Custody
MSC Houston, Tex.	See Remarks	MSC Houston, Tex.	Lunar/Earth Data Bank

Remarks:

Microfiche	$0.15 ea
35-mm neg	0.15 ea in aperture card
70-mm neg	0.20 ea in aperture card
105-mm neg	0.50 ea in envelope
9.5 × 9.5 in. neg	2.00 ea in envelope

REFERENCES

Accession Numbers	Issue Date
1. 68 I 10064	7/15/67
2. 68 I 10065	7/15/67
3.	
4.	
5.	
6.	
7.	
8.	
9.	
10.	

Fig. 4-3 Example of Image Citation

4.2.3 Unit Records

Unit records can serve as the final product for some users, and as an intermediate review product for other users. A most appealing unit record format is the COSATI standard 4- by 6-in. microfiche (Ref. 5). A single fiche can contain a nominal 60 microimages, allowing for the storage of entire technical reports or sequentially collected images on one piece of film.

The criteria for deciding which source data will be represented in microimage form, as well as distribution policies, should be determined in a system design phase.

4.2.4 Automatic Data Inputing

The input processing function discussed in this Section (4.2) is discussed in the context of a manual operation. However, the anticipated rate of acquisition of source data will be very high at times. To prevent the buildup of a massive backlog, an effort must be made to apply automatic processing wherever possible. For example, the one-to-one relationship between a unit record and its citation should be considered, where part of the citation's content is provided by the digital information acquired as magnetic tape source data.

4.3 INFORMATION STORAGE SUBSYSTEM

The functional responsibility of the information storage subsystem is to store source data, unit records, and citations in such a manner that they are quickly and easily retrievable given their accession or identification number. The storage subsystem will contain three types of storage:

- Physical storage for source data
- Image storage for unit records
- Digital storage for citations and indexes

4.3.1 Physical Storage of Source Data

The physical forms and information content of various classes of source data are summarized in the following tabulation.

Physical Form	Information Content
Magnetic Tape	Instrument traces, telemetry, vocal recording, digital data
Film	Images
Reports	Symbolic textual data
Standard Forms	Formatted symbolic data
Samples	Substantive quanta

There are three primary considerations in the physical storage of these types of source data:

- Storage environment
- Storage order
- Storage maintenance

Storage Environment. Environmental considerations include temperature and humidity requirements, shelf or bin construction, and handling equipment used in conjunction with the source data. Storage order refers to the discipline or rules which determine where a particular item of source data is stored. Because source data become obsolete, rules must be established for continually replacing source data of low utility with source data of higher utility.

The determination of specific environmental requirements for the various physical forms of source data is the proper subject of a systems design phase. Table 4-1 summarizes the general environmental conditions required for storage of photographic film and magnetic tape.

Storage Order. Objects are usually stored in one of two orders: serially on an as-received basis, or characteristically according to some attribute of the items being stored. The former storage order necessitates an index, whereas the latter allows browsing by attribute class. Where possible, a serial storage policy within a major source data category is desirable. That is to say, magnetic tape would be stored separately from film, but individual reels would be serially identified and sequentially

Table 4-1

SUMMARY OF ENVIRONMENTAL STORAGE REQUIREMENTS FOR FILM AND MAGNETIC TAPE

Unprocessed Film[a]	Temperature (°F)	Humidity (% RH)	Recommended (°F/% RH)
Spectroscopic films/plates	40 ± 10	30 to 40	40/40
Color emulsions (normal)	40 ± 10	30 to 40	40/40
Color emulsions (high control)	0 to -10	< 20	-10/20
Processed Film[a]			
Common B&W film	70	60	70/60
Motion picture film (B&W)	70 to 80	40 to 60	70/60
Color emulsions	40 to 50	25 to 40	40/30
Magnetic Tapes[b]			
Heavy duty type	40 to 90	20 to 80	Within limits specified
Mylar Tape[c]			
Heavy duty type	50 to 90	20 to 80	Within limits specified

(a) Telcon with representative of Eastman Kodak Company.
(b) Information from IBM Installation Planning Manual. No. C-226820-7. pages 8 and 9.
(c) Magnetic tape should be stored in a dust-proof container in a vertical position and should never come into contact with magnetic material at any time. Magnetic fields greater than 50 oersteds intensity can cause loss of information or introduction of noise.

stored. The overriding consideration is that a particular item of source data be readily locatable, given its serial number. It is the purpose of the inquiry subsystem, described below, to identify the particular items of source data which contain desired information.

Storage Maintenance. Assuming that storage space is a limited resource, provision must be made for purging the store of outdated and unlikely to be used material. The most effective purging method, of course, is the effective screening of system input. Only those items with a high utilization potential should be processed into the Lunar/Earth Data Bank system. The effectiveness of input screening policies should be evaluated by measuring the utilization of items stored. If overall utilization is low, either the material is inherently unuseful, the availability of the material is inadequately disseminated, or the system is unresponsive to the users' data needs. Measurement of utilization will provide guidelines for freeing storage space as well as planning future collection missions. Although measurement of stock turnover is essential to effective inventory control and ordering procedures in the business environment, it is infrequently employed within data storage systems. A field is provided in the citation panel (Fig. 4-2) which records the number of times the particular source data item was requested.

4.3.2 Unit Record Storage

There are many physical storage devices for unit records of various formats. The devices can be classified according to whether they provide:

- Random or serial access
- Automatic or manually controlled selection
- Limited or mass storage

File cabinets with 4 by 6 in. trays provide an inexpensive, random access, manual selection device of unlimited capacity to store standard microfiche. Reel microfilm provides a storage medium necessitating serial access which can be either automatically or manually controlled and has unlimited storage capacity (considering multiple reels). Video tape provides a high cost, serial access device of unlimited storage with an additional feature of easy transmission. At present there are no commercially marketed devices which provide users large capacity, automatically controlled, random access, microimage storage such as is required in a Lunar/Earth data bank. Such devices can be expected by the 1970 time frame. At present, Houston Fearless markets a device (CARD – Compact Automatic Retrieval Display) with automatically controllable, random access selection which stores up to 750 microfiche (Ref. 6).

Because of the current availability of storage devices of limited capability, and the anticipation of more fully adequate devices within the near term, development of a unit record base can begin at any time with a reasonable expectation of widespread device support.

4.3.3 Digital Storage for Citations and Indexes

The requirement of a digital storage system is to assist the user in identifying which physical items contain the information he requires. This identification is accomplished through the use of two types of files: citation files which contain descriptions and abstracts of source data items,* and index files which indicate the items containing particular attributes. The purpose of citation records and their associated indexes is to provide a descriptive medium which is intelligible to both man and computer. In this way, the data processing power of the computer can be coupled to the logical power of the man to create an effective information retrieval team.

As with physical storage of source data, there are three major considerations in the selection of storage media and design of the files:

- Physical storage media
- Storage order
- File maintenance

*There is no reason why the citation cannot provide substantive data to the user. In the case of a textual source data record such as an experiment summary form, the citation could include the record.

Physical Storage Media. There are several types of digital data storage devices currently marketed which differ according to cost, speed and means of access, and storage capacity. Magnetic tape is a low cost, low speed, serial access, high capacity medium, whereas magnetic cores are a high cost, high speed, random access, low capacity medium. Between these two extremes lie magnetic drums, disks, and several hybrid devices. In general, storage cost varies directly with capacity, randomness of access, and access speed, as indicated in Fig. 4-4 (Ref. 7). As a result, the design of any information system requires the careful configuration of a hierarchy of storage devices according to access requirements.

Because an interactive information retrieval system requires proportionately greater processing of index files than citation files. the indexes should be stored on relatively fast access media (such as disk, drum, or core), whereas the citation file can be stored on slower access media with higher capacity. (The serial access characteristic of magnetic tape discourages its use as a storage medium in an interactive environment.) In recent years, several of the manufacturers have announced devices which lend themselves to the requirements of citation storage as illustrated in Table 4-2.

Because several units of each of the devices described in Table 4-2 usually can be included in a specific system, the upper limit of random access digital storage can be in excess of a billion characters. It is also e... ..ed that storage capacity and storage cost will continue in apposition with further reductions in the cost of mass random access storage.

Storage Order. Because the index files can be processed independently of the citation file, the order of the latter can be largely a matter of convenience. and is usually determined on maintenance considerations. The index file for attributes should be organized by attribute value within attribute class. This index identifies the accession numbers of source data items containing particular attribute values. A citation location index provides the storage location for the citation corresponding to the identified accession numbers. Table 4-3 illustrates this or...nization.

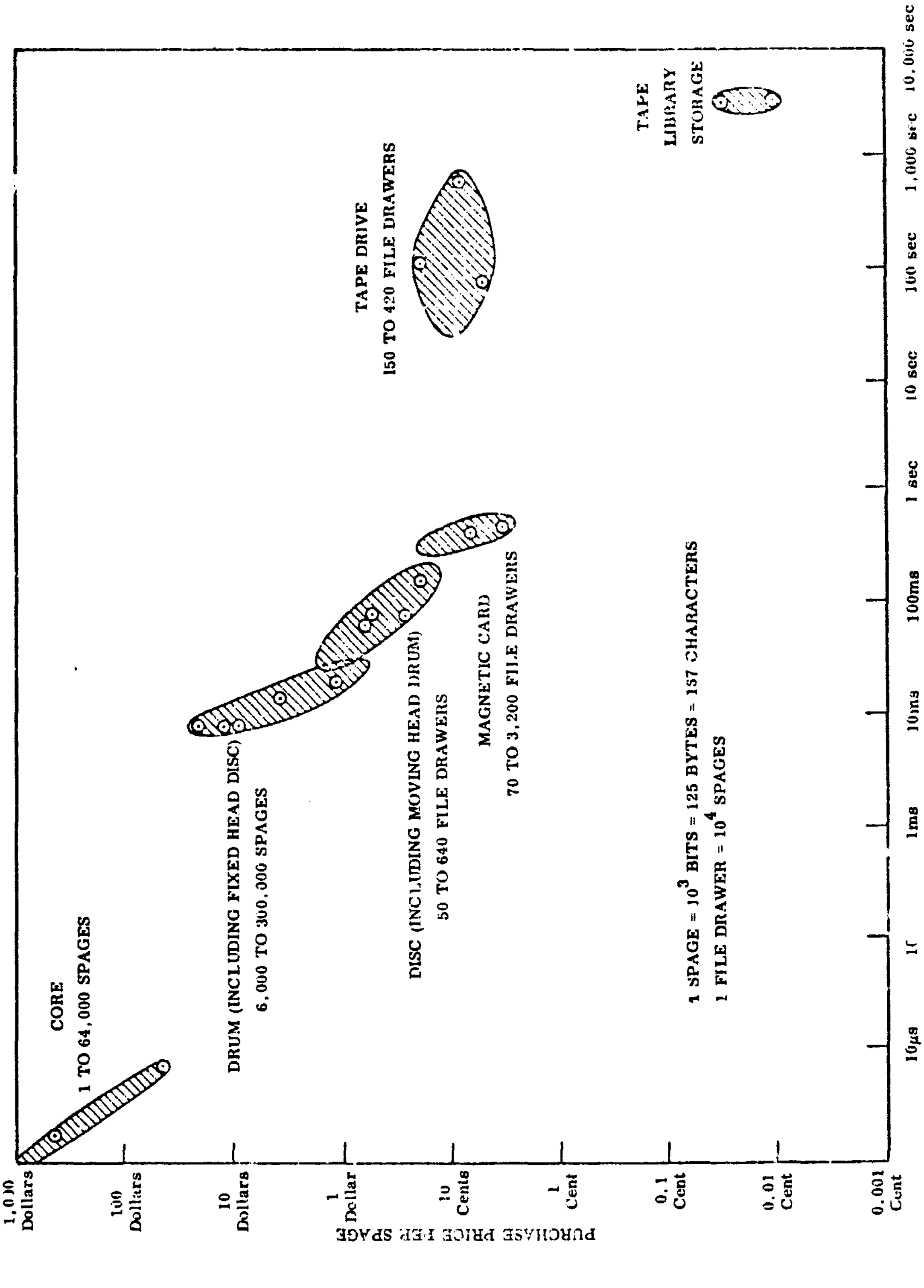


Fig. 4-4 Cost of Magnetic Media Storage Versus Access Time

Table 4-2

MASS DIGITAL STORAGE UNITS

Manufacturer		Nomenclature	Storage Technique	Data Transfer Rate (kc/sec)	Average Random Access Time (msec)	Capacity per Unit (characters)
Radio Corporation of America	70/568	RACE	Magnetic strip	70	375	560×10^6
National Cash Register	353-5	CRAM	Magnetic strip	51	125	82×10^6
UNIVAC		FASTRAND II	Drum	125	93	130×10^6
General Electric	DS 25	Mass random access file	Disk	300	116	201×10^6
Potter Instrument Company		RAM	Tape pack	600	90	50×10^6
International Business Machines	2321	Data cell	Magnetic strip	54	500	400×10^6
Honeywell	253	--	Magnetic strip	100	225	317×10^6

File Maintenance. The organization illustrated in Table 4-3 lends itself to file maintenance. Consider how the following maintenance problems would be handled:

- Delete citations
- Add citations
- Change existing citations
    (1) Result is no longer than previous citation
    (2) Result is longer than previous citation

Table 4-3

ILLUSTRATIVE INDEX FILE ORGANIZATION

Attribute Index			Citation Location Index	
Attribute Type	Attribute Value (deg sec)	Source Data Items	Accession Number	Storage Location
Topography	Mountain. (x)	68I10063	66B13459	N164
		68I10084	66B13460	N165
	⋮	⋮	66B13462	N166
	Plains (s)	67I10112	⋮	⋮
		67I10221	66I10025	1A23
		67I10314	66I10026	1A24
	⋮	⋮	⋮	⋮
Location Grid	0000	66I10122	67G10135	13B1
		66I10123	67G10136	13B2
	⋮	⋮	67G10137	S135
	OA13	66I10212	⋮	⋮
		66I11313	67I10011	1413
	⋮	⋮	67I10012	1414
Site Number	16	67I11342	67I10013	1A25
		67I13415		
	⋮	⋮	⋮	⋮

Given a list of accession numbers to be deleted, deletion is accomplished by substituting the location of a dummy citation in the citation location index. The citation itself is written on a transactions tape to provide an audit trail and to serve as a basis for updating the attribute index. Citations to be added are loaded after the previously highest storage location. The accession numbers and resulting storage locations are then merged with the previous citation location index.

In a separate pass, attributes keyed to accession numbers are sorted by attribute and merged with the existing attribute index. Existing citations can be changed by merely replacing the old record with the new if the resulting length is no greater than the previous length. If the change results in a longer record, it is treated as an addition and a deletion.

4.3.4 Summary of Information Storage Subsystem

The storage subsystem consists of three interrelated parts: physical storage for source data, image storage for unit records derived from source data, and digital storage for citations and associated indexes which describe source data. The purpose of the digital storage is to assist the user in identifying which source data items contain the information (i. e., the attributes) he requires. The unit records provide a convenient (in size and format) intermediate product to further assist the user in screening his response data. Source data items are the final product of a search. By furnishing several screening levels (index, citation, and unit record), only the relevant and needed source data will be requested. Because of the comparatively high cost of processing source data, these intermediate screening levels are required to assure that only pertinent source data are requested by the user.

4.4 INQUIRY SUBSYSTEM

The functional objective of the inquiry subsystem is to provide a means by which the potential user can identify a relevant subset of items within the total collection without having to exhaustively examine each item in the collection. This objective is

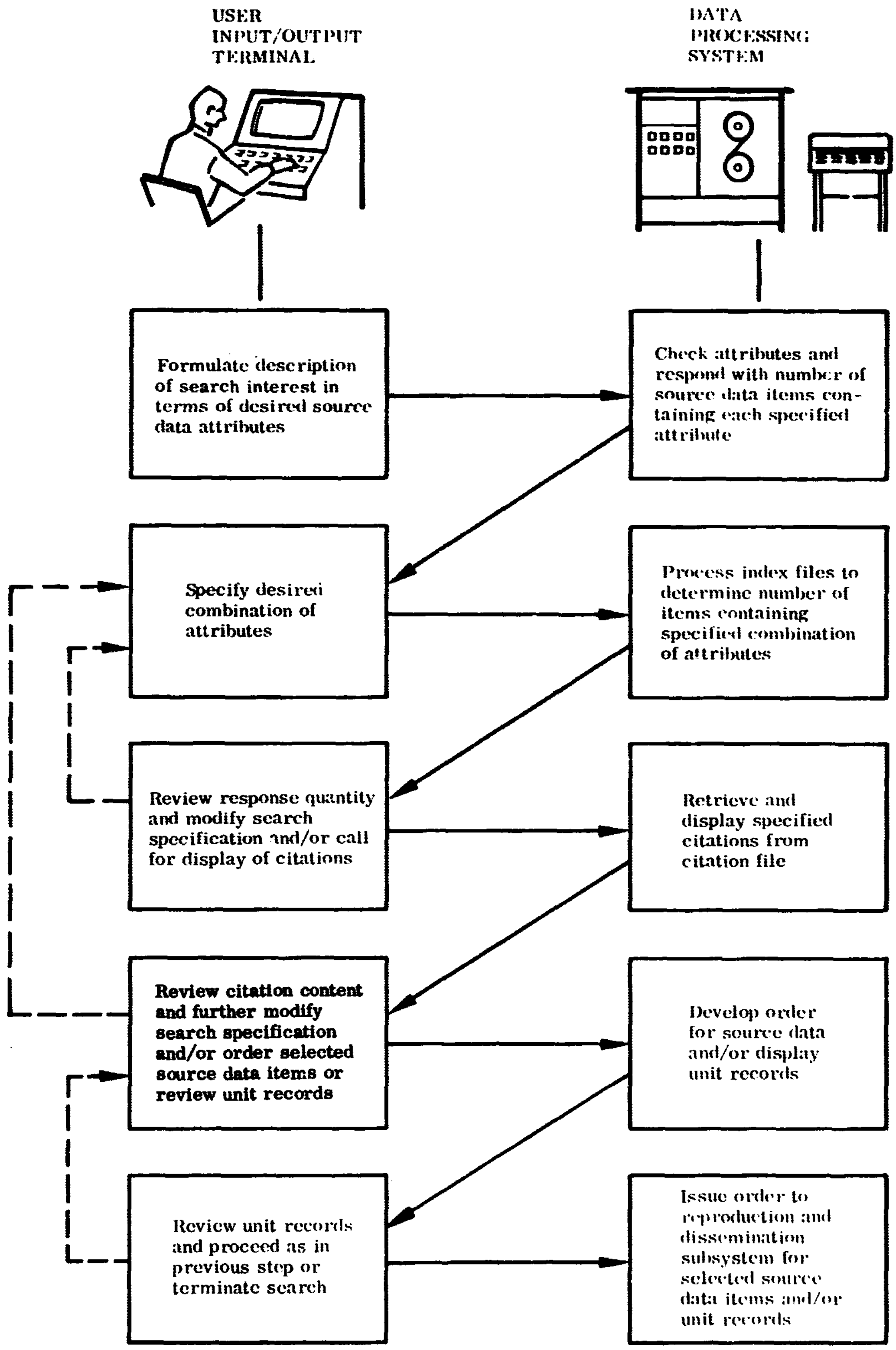


Fig. 4-5 Conceptual Information Flow of Inquiry Subsystem

4.5 REPRODUCTION AND DISSEMINATION SUBSYSTEM

The functional responsibility of the reproduction and dissemination subsystem is to disseminate announcement information regarding the availability of source data, to reproduce unit records, and to respond to requests for specific material by supplying original source data or reproduced copies of original source data as appropriate.

4.5.1 Announcement of Available Source Data

It is anticipated that the Earth resources data will be in wide demand among a variety of users (as opposed to data gathered for a single. unique purpose). If these data are to be utilized to the fullest extent, potential users must be made aware of the existence and availability of the data. The citation and unit record (described in subsection 4.2) can serve as the basis of an announcement catalog. Because the citation is in computer-readable form, it can be used to produce publishable text and provides a concise description of the available source data, as well as ordering information. The images of the unit records can be printed together with the citation text to give the potential user a more precise idea of source data applicability.

Such a system of announcement catalogs is fully compatible with and complementary to the current NASA announcement of technical reports in the STAR publications. As a result, cross referencing of technical reports and source data from Lunar and Earth programs is facilitated.

4.5.2 Distribution of Request Responses

Usually requestors should be provided with nonreturnable copies of requested source data. In cases where this is impossible, and in cases where the data are only of interest to a particular principal investigator, reproduction and dissemination subsystem personnel have a custodial responsibility for the source data. This responsibility can be exercised in conjunction with the citation, one field of which indicates the location of the source data, and when it was delivered to this location (Fig. 4-3). Changes in location can be documented using a citation change form.

4.5.3 Reproduction of Unit Records

The reproduction and dissemination subsystem provides reproduction support for the remainder of the data bank, and thus must include necessary photographic reproduction equipment. Substantially all such equipment is currently available within NASA.

4.6 REMOTE INFORMATION CENTERS

As use of the source data and unit records increases, it may become desirable to establish regional data centers at the sites of principal users. Such a center would have a complete collection of announcement catalogs and unit records, and could be linked to the central retrieval facility via a remote input/output terminal. Centralization of the citation storage and decentralization of the inquiry terminals provide a maximum of access convenience with a minimum of file maintenance problems. Any final network design should include a cost-effectiveness analysis of demand, telephone line costs, and satellite storage costs. The network configuration is an appropriate part of detailed system design.

Section 5
IMPLEMENTATION CONSIDERATIONS

5.1 INTRODUCTION

The previous section has outlined a storage and retrieval system for Lunar/Earth data. The success of such a system will depend largely on the implementation approach followed.

It is important that a retrieval system reflect the needs of its users. A system that is appropriate for professional searchers might be inappropriate for engineers: a system designed for a small, specially indexed collection can be unworkable with a large general collection of data such as that anticipated in a Lunar/Earth data bank.

It is difficult to evaluate the degree of usefulness of a particular retrieval system design without investing a large sum of money in the design and programming of sophisticated applications programs. Moreover, experience with the initial design frequently suggests additional desirable capabilities. What is needed is a powerful yet flexible information retrieval language which can be applied quickly and inexpensively to a large, complex data base. Such a language would allow MSC personnel to evaluate the strengths and limitations of such a system through first-hand experience.

Because of the state-of-the-art nature of the proposed system, it seems highly desirable to design an implementation program which allows the system to react to experience gained during initial operation of the system. Consequently it is recommended that implementation be initiated with a prototype phase followed by an operations phase. Knowledge and experience gained in the prototype environment can help to crystallize the requirements for a large-scale operational system.

5.2 PROTOTYPE PHASE

The objectives in developing a prototype Lunar/Earth Data Bank system are to accomplish the following:

- Demonstrate the feasibility and utility of the basic concept in an expeditious and cost-effective manner.
- Provide a means for learning by doing.
- Permit maximum experimental flexibility.
- Develop a basic system which can simply and logically be extended into an operations phase.

Where possible, a prototype should utilize existing hardware and software to avoid a costly development effort. Furthermore, if the prototype is to be truely representative of the final operating system, it should operate on a data base which includes a complete collection or representative data types. Data elements to be included in the prototype should be well organized and complete in terms of annotation. Finally, if the prototype is to attract a representative sample of users, it must offer substantive benefit to them and not provide mere academic exercise.

Ideally, a prototype system will be composed of breadboard components which can readily be reconfigured in response to experience gained during operation of the prototype, and the final prototype configuration can be utilized as the initial operational system.

Implementation of a prototype phase includes the following specific elements, each of which will be discussed:

- Retrieval system selection
- Data base generation
    - Data base selection
    - Citation format development
    - Indexing vocabulary development
    - Data base generation

- System demonstration
- System evaluation and operations phase implementation plan

5.2.1 Retrieval System Selection

Recently there have been developed several information retrieval computer programs. Although they are similar in objective – providing the user with desired information with a minimum of effort – they differ widely in capability. Prior to implementation of the prototype phase, either a particular existing system should be selected, or a system should be especially written for the prototype. Selection criteria for existing systems include:

- Demonstrated application in a variety of areas (library, personnel, etc.)
- Ability to operate with arbitrarily large data files (at least 1 to 2 million records)
- Ability to support multiple terminals
- Index based (as opposed to serial scan based) so that increasing file size does not materially increase retrieval time
- Random access storage (as opposed to tape or microfilm storage)
- Ability to modify easily search expression (without redoing the search) based on intermediate output from the file
- Immediate availability

There are relatively few systems which encompass more than one or two of these features. Companies having systems worthy of investigation in this regard include:

- DATA Corporation, Dayton, Ohio
- System Development Corporation, Santa Monica, Calif.
- Lockheed Palo Alto Research Laboratory (DIALOG System). Palo Alto, Calif.

The time and cost estimates in the following discussion are based on experience Lockheed has had in implementing the DIALOG system (described in Appendix C) on somewhat similar data bases, which may or may not be realistic in terms of the other cited software systems.

5.2.2 Data Base Generation

By far the most difficult problems in the implementation of the prototype Lunar/Earth Data Bank system are those relating to the selection and generation of the data base. It is suggested that this task be performed at .. ЗC with supported guidance from Lockheed personnel. The total task can be divided into several elements for discussion purposes as f illows:

- Data base selection
- Data base generation
- Unit record generation

Data Base Selection A major consideration in the valid demonstration of the capability of the system concept is the proper selection of the data base and file generation. It is obvious that the core problem in the concept presented is the handling of data derived from a variety of tape and imaging systems and its correlation with textual matter such as flight logs and ground-truth information. Consequently, the data base selected should be large enough and sufficiently diverse in content so that valid extrapolations may be made on the basis of results obtained by its manipulation.

It is recommended that 10,000 citations be generated from a variety of different remote sensors such as visible, infrared, and radar. It is conceivable that blocks of data can be extracted from those data obtained in the Earth Resources Aircraft Program, Gemini Program, or others possibly considered more appropriate. It would seem that various formats, scales, instrument types, resolution levels, and terrain types should be selected so that the sampling is indicative of an operational data base.

Data Base Generation Development of the citation format has been discussed in detail elsewhere (subsection 4.2.2 and Appendix B). It will remain to formulate the specific citation formats for each of the various data types, to test the formats on a sample of potential users, and to produce citation worksheets for use by Input Processing Subsystem personnel in cataloging the selected input data.

Following development of a preliminary indexing vocabulary, generation of citations could begin. A brief test using the citation format shown in Fig. 4-3 indicated that it takes approximately 6 to 10 min per citation for the cataloging function. Based on the above assumption of 10,000 citations, about 12 man months are required to catalog the initial data base.

Unit Record Generation Selected units of source data should be identified for the basis of unit record generation. The microfiche format should contain contiguous sequences of imagery (or pages of a report). Kodak has recently introduced Kodachrome microfiche which should be evaluated by inclusion in the microfiche store for those source data elements which are particularly benefited by color reproduction. Although it is not presumed that the microfiche will serve the needs of most system users, the extent of their usefulness should be evaluated as part of the prototype phase of operations.

Microfiche production costs are relatively low as can be seen in Table 5-1

Table 5-1

MICROFICHE PRODUCTION COSTS

Negative Size	Frames per Fiche	Type of Film	Cost of Master	Cost per Fiche	Cost per 1000 Frames	
					Masters	Fiche
8 × 10 in.	32 (16:1)	B & W	2.00	$0.15	68.00	$ 5.10
8 × 10 in.	32 (16:1)	Color	6.00	0.75	204.00	25.50
8 × 10 in.	98 (24:1)	B & W	2.00	0.15	22.00	1.65
8 × 10 in.	98 (24:1)	Color	6.00	0.75	66.00	8.25

The comparative cost of producing 1,000 frames of 35-mm positives (a format suggested for some unit record applications) is $160.00 for each thousand produced.

Microfiche viewing equipment is reasonably low cost varying between $100.00 and $400.00 for nonreproducing viewers, and between $1000.00 and $3000.00 for reproducing viewers.

5.2.3 System Demonstration

Principal investigators and other potential users should be introduced to the system prior to completion of the prototype phase and should be requested to evaluate system operation constructively. During this period it is anticipated that there will be several modifications to the indexing procedure and indexing vocabulary based on this initial experience.

While limited in operational scope by virtue of the size of the data bank, the demonstration of the prototype will simulate the input and output data flows. As additional subsystems are evolved, they can be demonstrated by means of the prototype system.

Subsequent to the issuance of a prototype system evaluation report, an operational system implementation plan should be prepared. This document will be based on the experience gained during the prototype phase and will contain the specifications and implementation requirements for the operations phase.

It is anticipated that the equipment for the operations phase of the implementation program can be installed at MSC 12 months after initiation of the program. During the period of prototype operation, detailed organizational responsibilities can be assigned, and physical facility plans can be effected. The careful monitoring of the prototype phase can provide the basis for valid planning and cost estimates for the following 5-year period. The success achieved in the development of data extraction and image analysis methodology will greatly affect not only the future costing of the operational system, but also the final design to be used in the operational system. Progress in all subsystem areas must be parallel since a lag in any area can affect the total system utilization potential.

5.2.4 Operational Phase

The major transition from prototype to operational system will be the finalizing of the system configuration, procurement of operational system hardware, and the phase-over of computer programs to the operational system. At this point, detailed data volumes can be estimated, and system parameters can be frozen. Based on the prototype system evaluation report, prototype programs should be modified and extended to be responsive with operational requirements.

The major tasks and functions associated with the implementation program are:

- Systems engineering – adaption and development of software for the Lunar/ Earth Data Bank system
- File generation – the production of descriptive citations; generation of machine-readable files
- Communications costs – cost of full duplex communication line between remote terminals and computers
- Operational system equipment – the lease/purchase cost of equipment for operational system
- Applications programming – the final Lunar/Earth Data Bank applications programs

Section 6
ADDITIONAL CONSIDERATIONS

6.1 INTRODUCTION

There is a variety of research currently underway related to improving the application of remote sensing systems and technology. Although not immediately concerned with the data bank problem, this research is directed at improving the quality of the data to be stored in a data bank, and at techniques which allow more meaningful analysis to be conducted on data to be retrieved from a data bank. Because of the importance of this research to the effective utilization of Lunar/Earth data, discussion of these related areas is included in this section. The study areas relate to functions shown in Fig. 1-1 which depicts an overall Lunar/Earth data flow as follows:

- Data acquisition – the acquisition of raw data from missions together with the necessary correlative data to make the raw data meaningful to the analyst
- Data processing – the manipulation of data including processes of developing, reformatting, enlarging, converting, etc.; i.e., processes necessary to produce an intelligible and efficiently organized record for storage and analysis
- Data analysis – the extraction of meaningful information from processed records

The present section recommends that additional study be directed in these related areas both within NASA by establishing a Remote Sensing Technology Laboratory, and elsewhere by continuing and extending existing university and industry Lunar/Earth sciences programs.

In addition to development in these technical areas, it is felt that a study should be performed to determine how best to disseminate Lunar/Earth data, and to educate potential users in techniques which can be used in its exploitation. *

6.2 DATA ACQUISITION

Data acquisition developments are closely tied to instrumentation development and should be closely coordinated with development of the data handling system. To arrive at an effective information storage and retrieval system design, it is necessary to consider the form and annotative content of the data acquisition system. In systems that are in the process of development, the degree and type of annotation included on the imagery should be carefully analyzed so that identification is foolproof and data processing and analysis are facilitated by having time and geometry references available on each frame of imagery. Simultaneous firing of shutters and the recording of exposure and calibration references are examples of areas to be considered. In essence, an overview of the entire data acquisition and information handling systems should be taken to evaluate the alternative design tradeoffs.

Because many data are now available without the necessary accompanying annotation data, it is also necessary to study current systems with a view toward adapting them most efficiently to the data handling systems being currently developed.

6.3 DATA PROCESSING

The area of data processing can be divided into (1) <u>image processing</u> which includes film-related processes of development, enlargement, enhancement, and duplication, and (2) <u>digital data processing</u> which involves processes such as reduction, reformatting, converting, sorting, merging, editing, and enhancement. It is important that additional studies be performed in the data processing area to better delineate requirements and their relationships to information handling concepts now being evolved.

---

*Much of the material contained in this section has been derived from internal MSC studies, particularly from the Mapping Sciences Laboratory.

6.3.1 Image Processing Techniques

Image data are obtained directly by photographic techniques, indirectly through electronic imaging systems, or by combinations of photographic and electronic scanning processes. The following are specific areas of image processing which require further study:

- Image development – preparation of imagery for scanning, interpretation, and compilation
- Image enhancement – increasing the legibility of the image components to aid in interpretation and screening
- False color techniques – increasing the discrimination between similar images by means of composite color presentations
- Image storage techniques – development of storage media, retrieval devices using such media, and display devices suitable for remote recall and display of pictorial data
- Correlation of imagery – provision of multiple display of imagery which allows comparison of like data points, change detection, signature determination, and preparation of jointless mosaics

To optimize the extraction of data from the image, it is vital that consideration be given to precise control of the film processing. The production of high quality monochromatic and polychromatic reproductions, enlargements, and overlays, and the precise rectifications used in precise interpretation and mensuration, require controls and techniques beyond the normal processing laboratory. Closely related to this area is that of edge and area image enhancement which enables more accurate analysis as well as analysis not otherwise possible without enhancement. Although techniques for performing enhancement now exist, these techniques must be integrated into the overall processing system, and new techniques must be devised.

The use of color to supply an additional dimension in the data extraction and interpretation process has been well established. Continued development is proceeding in the application of false color techniques for discriminating similar images which appear in sets of multispectral negatives. By projecting positives of each spectral negative onto a screen in a registered fashion using a different primary color for each, a composite color rendition of the scene is formed. This additive color presentation will produce every hue; the saturation of the color will depend on the relative density differences between the spectral positives. The brightness of the image color will be a function of the minimum density level of the image on one of the spectral positives used in the projection. Adjustment of the three primary colors will achieve the desired color, and the composite presentation will show all density differences between the individual positives as colors, with a shade of grey appearing where the same density appears on all positives. Thus, detection of images where the density differences are slight is possible by noting differences in hue, brightness, and saturation in relation to the background. Accurate superpositioning of images is, of course, basic to the process.

6.3.2 Digital Data Processing Techniques

Digital data processing can consist of conversion of analog signals to digital form, with subsequent storage or processing of these signals, or can involve manipulation and storage of alphanumeric material. The following areas of digital data processing require further study:

- Automatic indexing – the processing of alphanumeric data to extract significant words, phrases, and values which characterize the contents of the record being indexed. *
- Data correlation – the automatic association of time and position related data.
- Data reduction – the reduction of raw data to standardized record formats which are more suitable for retrieval or for further processing.

*For a review of the present state-of-the-art, see the chapter "Indexing and Classification" by H. Borko in the book Automated Language Processing, edited by H. Borko, Wiley, New York, 1967.

The data bank concepts discussed in the present report require that large quantities of data be processed so that the time, position, and textual information are efficiently extracted and correlated. Nonimaging sensor data can be correlated with aerial photographs of the terrain by recording the time of shutter opening together with the sensor output signal. It is possible to derive the path of the sensor boresight on the photograph by using flight parameters obtained from onboard navigation equipment (or derived from the photographs themselves) and boresight calibration tests which relate the position of the camera and sensor. The association of sensor position on the ground with the instantaneous sensor return can be accomplished by a digital computer using techniques described in Ref. 43.

The systems aspect of the indexing, correlation, and reduction procedures thus becomes evident: the techniques must be tailor-made to allow for the characteristics of the incoming data, and the final output must be compatible with the data bank and the retrieval techniques used in the data bank. Progress in indexing, correlation, and reduction will allow more economic preparation of relevant data bank material.

6.4 DATA ANALYSIS

The extent toward which data analysis must be conducted as a manual rather than an automatic process will limit the frequency and completeness of analysis. Additional studies which should be performed toward increasing automation include:

- Pattern recognition – the automatic determination of image content through an analysis of the digitized representation of the image
- Change detection – detection of changes through repetitive coverage of the same area, followed by automatic correlation of image content based on some criteria of change
- Signature correlation – the analysis of image patterns of materials to determine invariant patterns which identify the materials under various conditions
- Automated photogrammetric analysis – the automation of the mathematics of perspective and projective geometry to enable mapping not only from frame photography but also from strip, panoramic scan, and line-scan images

Pattern recognition and change detection techniques have received much attention over the past several years. If perfected, these techniques could be used to automatically screen and, to some extent, index incoming image data. The procedures are currently somewhat unsophisticated and very specialized as indicated in Appendix B, but further research appears to hold promise.

Because different types of rocks, vegetation, and other material have different spectral distribution of radiation, it should be possible to discriminate surface materials using remote sensing equipment. Due to the importance of its ultimate application to remote sensing, work in the signature correlation area should be encouraged and extended.

Automated photogrammetry is currently in its infancy with some work taking place in the automated readout of map position and elevation information from frame photography. Automatic plotters can profile and contour three to ten times faster than human operators with accuracy and repeatability of orientation better than that achieved through manual operation. (Some plotters automatically print orthophotographs using height information from a stereopair.)

Much work remains to be done in developing not only the mathematical relationships relative to the geometry of strip and scan imagery, but also in reducing these mathematical formulations to operating systems utilizing computer techniques.

The need for these developments has arisen because of the limitations of frame photography in meeting the sophisticated needs of world-wide and planetary mapping. The unique geometry of these systems, the need for remote communication links, and the tremendous amount of data to be analyzed make it mandatory that research in the field of computer automation of photogrammetry be broadened.

6.5 REMOTE SENSING TECHNOLOGY LABORATORY

Effective utilization of data from a state-of-the-art sensor complex can only be attained through the application of the newest and most powerful techniques. Although it would be possible to fund each of several principal investigators sufficiently to maintain currency with most remote sensing techniques, the cost of duplication of effort would be high and coordination of the investigations would be difficult. What might be more desirable would be the establishment of a Remote Sensing Technology Laboratory to perform experimental research in the areas enumerated in this section.

The principal functions of this laboratory would include:

- Development and pilot application of advance image enhancement and data reduction techniques
- Calibration and establishment of field standards
- Determination of the reliability of calibration data for each mission
- Application of optimum image enhancement and radiometric data reduction techniques to special user problems
- Simulation of anticipated sensor products and outputs by the use of laboratory and computer models
- Optimization of sensor parameters for specific applications
- Analysis and simulation of unusual anomalies detected by the sensor on specific missions
- Experimental evaluation of state-of-the-art processing and analysis techniques developed elsewhere

6.6 DISSEMINATION AND DATA EXPLOITATION

It is not at all unusual to find that a large amount of effort has gone into the technical aspects of an information systems design, but that correspondingly little effort has gone into such considerations as dissemination of information and user training.* This appears to be the case to date in the Lunar/Earth sciences field: an area largely overlooked has been the consideration of methodology by which potential users of

*For a review of the present state and problems of user studies see the chapter "Information Needs and Uses in Science and Technology," by S. and M. Herner, in Annual Review of Information Science and Technology, Vol. 2, 1967, C. A. Cuadra, ed., New York, Wiley, 1967.

Lunar/Earth resources data may be informed of the existence of such data and educated as to its proper use and application. Study is therefore required to develop methods of stimulating a dynamic interchange between the sources of data and the prospective user community. The following are some of the potential aids to data dissemination and exploitation:

- Educational programs – organized educational seminars which travel from region to region, designed to acquaint potential users with the Lunar/Earth resources material. The seminars could consist of lectures as well as demonstrations of actual use of the system by means of remote terminals.
- Distribution of informational material – general informational material and lists of currently available detailed material could be circulated to selected potential users. The users could indicate their interests, and appropriate follow-ups then made.
- Selective dissemination of information – Selective Dissemination of Information (SDI) systems bring published papers or their titles to the user on the basis of comparison of their subject content with the "profile" of the user interests. *
- Publication of abstracts – it should be possible to incorporate selected abstracts of Lunar/Earth sciences material into the NASA Scientific & Technical Aerospace Reports (STAR) system.

---

*For a discussion of distribution of information by means of published literature, see the chapter "Techniques for Publication and Distribution of Information," by the American Institute of Physics staff, in the volume cited in the previous footnote. Abstracting and indexing services are discussed, as is communication outside of conventional publication channels.

6.7 SUMMARY

The components of data acquisition, processing, and analysis have been described, and the need for research on the automation of many of the operations has been stressed. Two specific recommendations were made: one for the establishment of a Remote Sensing Technology Laboratory which would develop integrated, highly automated acquisition, processing, and analysis systems; and another for the development of techniques for dissemination of Lunar/Earth data, and the education of potential users.

Section 7

REFERENCES

7.1 CITED REFERENCES

1. "Future Nations Space Objectives," and "Hearings Before the Subcommittee on Space Science and Applications," of the Committee on Science and Astronautics, U.S. House of Representatives, 89th and 90th Congress, respectively.

2. DOD Manual for Building a Technical Thesaurus, prepared by Projext LEX of the Office of Naval Research, Apr 1966

3. Space Data Thesaurus, NASA-TM-X-59379, NASA Goddard Space Flight Center. Nov 1966

4. "Remote Sensor Aircraft Data Gathering System, Data Processing and Distributing Unit," Space Applications Programs Office, MSC, NASA, Mar 1966

5. COSATI Microfiche Structural Specifications, Federal Council for Science and Technology, Office of Science and Technology, Jul 1965

6. Andries van Dam, A Survey of Pictorial Data Processing Techniques and Equipments, (AD 626-155), University of Pennsylvania, Aug 1965

7. California Statewide Information System Study, Interim Study, Y-82-65-2, Lockheed Missiles & Space Company, 14 May 1965

8. Peaceful Uses of Earth Observation Spacecraft, Survey of Applications and Benefits, Vol, I, II, III, University of Michigan for NASA, Sep 1966

9. R. N. Colwell, "Some Practical Examples of Multiband Spectral Reconnaissance." American Scientist, Vol. 49

10. D. F. Marble and E. N. Thomas, "Some Observations on the Utility of Multispectral Photography for Urban Research," Proceedings of the Fourth Symposium of Remote Sensing of Environment, Apr 1966

11. R. H. Alexander, "Geographic Research Potential of Earth Satellites," Proceedings of the Third Symposium of Remote Sensing of Environment, Oct 1964

12. F. E. Kinsman, "Some Fundamentals in Non-Contact Electromagnetic Sensing for Geoscience Purposes," Proceedings of the Third Symposium on Remote Sensing, Oct 1964

13. W. Fischer, "Color Aerial Photography in Photogeologic Interpretation," Photogrammetric Engineering, Vol. 24, Sep 1963

14. H. O. Rydstrom, "Interpreting Local Geology From Radar Imagery," Proceedings of the Fourth Symposium on Remote Sensing of Environment

15. C. J. Robinove, "Preliminary Evaluation of Airborne and Spaceborne Remote Sensing Data for Hydrologic Uses," Technical Letter NASA-50, U. S. Geological Survey, Jul 1966

16. G. C. Ewing, "The Outlook for Oceanographic Observations From Satellites," Woods Hole Conference, Aug 1964

17. Tables, NASA Headquarters, SA 67-17011, Jun 1967

18. Woods Hole Conference, Space Science Board, Summer 1965

19. Proposal for a Group of Photographic Experiments for Manned Lunar – Orbital Mission (Apollo Application Flights), submission by U. S. Govt. Agencies Coast and Geodetic Survey et al. to NASA, Washington, D. C., 15 Oct 1965

20. "Interpretation of Lunar Probe Data," ed. J. Green, A. A. S. Science and Technology Series, Washington, D. C., 1967

21. North American Aviation Co., List of Lunar Exploration Experiments, NASA Contract, Downey, Calif., 1966

22. M. F. Meir, "Multispectral Sensing Tests at South Cascade Glacier, Washington," Proceedings of the Fourth Symposium on Remote Sensing, Apr 1966

23. C. Molineux, "Aerial Reconnaissance of Surface Features With the Multiband Spectral System," Proceedings of the Third Symposium on Remote Sensing, Oct 1964

24. E. F. Yost and S. Wenderoth, "Multispectral Color Aerial Photography," Photogrammetric Engineering, Vol. 33, No. 9, Sep 1967

25. R. G. Tarkington, "Color and False-Color Films for Aerial Photography," Photogrammetric Engineering, Vol. 26, Sep 1960

Manual of Photointerpretation – American Society of Photogrammetry, 1960

26. J. G. Vidder, "Evaluation of Ektachrome and Multiband Photography in California Range," Technical Letter 17, U.S. Geological Survey, Aug 1966

27. P. D. Lowman and J. A. McDivitt, Terrain Photography on the Gemini IV Mission, Preliminary Report No. G-779, Goddard Space Flight Center, NASA

28. P. D. Lowman, Synoptic Terrain Photography (Experiment S-5) During Gemini 6 and Gemini 7, Goddard Space Flight Center, NASA

29. F. F. Sabins, "Geologic Aspects of Infrared Imagery," Indio Hills, Calif., Mar 1967

30. R. S. Vickers and R. J. P. Lyon, "Infrared Sensing From Spacecraft – A Geological Interpretation," AIAA, Thermo-Physics Conference Paper 67-284, New Orleans

31. R. E. Wallace, "Use of Infrared Imagery in Study of the San Adreas Fault System," Technical Letter 42, U.S. Geological Survey, Jun 1966

32. D. E. Harris, "Terrain Mapping by Use of Infrared Radiation," Photogrammetric Engineering, Vol. 30, No. 1, Jan 1964

33. R. K. Moore, Radar Scatterometry – An Active Remote Sensing Tool, CRES Report No. 61-11, University of Kansas

34. J. H. Simons, "Some Applications of Side-Looking Radar," Proceedings of the Third Symposium of Remote Sensing of Environment, Oct 1964

35. L. F. Dellwig and R. K. Moore, "The Geological Value of Simultaneously Produced Like and Cross-Polarized Radar Imagery," J. of Geophysical Research, Vol. 71

36. R. D. Ellermeir and D. S. Simonett, Imaging Radars on Spacecraft as a Tool for Studying the Earth, CRES Report No. 61-6, University of Kansas, Nov 1965

37. E. H. Whitten, W. A. Beckin, and R. D. Hobson, Aspects of Geological Sampling at Test Sites, Northwestern University, Report No. 4, Jul 1966

38. J. Lintz and J. Quade, "Supporting Measurements of Passive Microwave Radiometer Sites," Technical Letter No. 1, Mackay School of Mines, University of Nevada

39. J. Lintz and J. Quade, "Conference on Ground Measurements for the Instrument and Geologic Teams," Technical Letter No. 3, Mackay School of Mines, University of Nevada

40. W. T. Brandhorst and P. F. Eckert, Guide to the Processing, Storage, and Retrieval of Bibliographic Information at the NASA Scientific and Technical Information Facility," NASA CR-62033, National Aeronautics and Space Administration, Jun 1966

41. J. G. Santoro, Techniques for Data Handling in a Unit Record Photographic Data Reduction System – Project SCRAM, AD-804843, Rome Air Development Center, Nov 1966

42. ORL Experiment Program, International Business Machines Corp., Contractor Report, Feb 1966

43. W. G. Eppler and R. D. Merrill, Semiautomatic Preparation of Radiometric Synthetic Reference Maps, 5-26-68-3, Lockheed Palo Alto Research Laboratory, Palo Alto, Calif., May 1968

7.2 UNCITED REFERENCES

Accession List for the Earth Resources, Aircraft Program Data, NASA MSC, Sep 1967

Analysis of Remote Sensing Data Requirements by Experiment, Remote Sensing Data Facility MSC, Houston, Texas

P. C. Badgley, L. Childs, and W. L. Vest, "The Application of Remote Sensing Instruments in Earth Resource Surveys," Geophysics, Vol. XXXII, No. 4, Aug 1967

W. D. Carter, "Summary of Significant Results of Remote Sensing Studies in 1965," Technical Letter NASA, U. S. Geological Survey, 14 May 1966

Detailed Plan and Status Report of U.S. Geological Survey, Research in Remote Sensing Under the Natural Resources Space Applications Program, Prepared by the U.S. Geological Survey for NASA

Earth Photographs From Gemini III, IV, and V. National Aeronautics and Space Administration, NASA SP-129, 1967

Proposal for a Multiband Synoptic Photographic Experiment for Manned Earth Orbital Missions NASA/OSSA Photographic Team, Jun 1966

Report to the Space Science Board on the Space Science and Applications Programs, NASA Headquarters, Washington, D. C., Nov 1966

Sample Submittal Manual – Instructions for Submittal of Samples for Chemical Spectrographic or Physical Properties Analysis, U. S. Geological Survey, Jun 1967

Appendix A
LUNAR/EARTH DATA BANK USER REQUIREMENTS

Based on a review of remote sensing literature and discussions held with principal investigators and NASA MSC personnel, the potential user communities were identified as including personnel in the following scientific and application areas:

- Earth Resource Disciplines – agriculture and forestry, geodesy, cartography and cultural resources, geology, hydrology and water resources, and oceanography
- Lunar Science Disciplines – cartography, geodesy, geology, geomchemistry, geophysics, biology, physics, astronomy, and engineering
- Technology Development – instrumentation development, ground truth correlation

This appendix defines the objectives of these discipline and application groups and derives the information requirements which are most amenable to collection by remote sensor.

A.1 EARTH RESOURCES REQUIREMENTS

A.1.1 Introduction

The use of Earth-orbiting satellites as platforms for sensing of the Earth's surface has unique advantages over the more conventional ground-based or airborne methods of observation. Because of the great altitude and rapid movement of the satellite, it is possible to observe large areas of the Earth's surface in a relatively short time and at moderate expense. Detailed studies are currently being made by public and private organizations under the auspices of the National Aeronautics and Space Administration to consider the potential applications of observation spacecraft for both scientific and economic purposes (Ref. 8).

To determine the data needs of potential users, the information requirements of several disciplines within the area of Earth Sciences are described, as well as the broad instrumentation characteristics necessary to collect the information. Disciplines discussed are:

- Agriculture/forestry
- Geodesy
- Geography, cartography, and cultural resources
- Geology
- Hydrology and water resources
- Oceanology

Although the disciplines are discussed separately, the data needs will not be oriented by discipline; rather, the separate coverages will be integrated into a set of overall information requirements.

A.1.2 Agriculture and Forestry

The expanding population and economic growth of the United States and the world require the continual development of the world agricultural capacity and forest resources. An accurate global inventory and continuing assessment of food and fiber resources is necessary for agricultural expansion. This is also generally applicable to forest resources where reliable information is needed in advance of forest exploitation as well as for forest management and protection.

Some specific applications of remote sensing techniques to agriculture and forestry are (Ref. 9):

- Gathering of data on plant vigor and disease to aid in the increase of agriculture and forest production
- Accurate, timely, and broad-scale surveys combined with automatic pattern recognition methods to yield information necessary for improved productivity, development, and utilization of these resources. Such surveys would yield

information useful in soil classification, land use capability, crop identification, forest species identification, crop acreage control programs, and forest fire detection

- Production of world land-use maps as well as synoptic photographs for planning and operational interpretation

The following technical developments have been indicated as important within agriculture and forestry:

- Determination of spectral signatures of agricultural and forestry phenomena
- Automated pattern recognition procedures as well as enhancement and screening techniques
- Evaluation of a global land-use satellite system which would provide continuous coverage of the Earth at constant sun angle using three-band photography and possibly radar

A. 1. 3 Geodesy

While the relationship of geodesy to Earth resources is rather indirect, its contributions to almost all of the applications discussed in this report will become more evident as the objectives of this discipline are covered. Geodesy is fundamental to such sciences and disciplines as cartography, astronomy, navigation, and Earth-orbiting space applications. The broad geodetic objectives are:

- Geometric geodesy – To determine the size and shape of the Earth and locate with high precision a number of reference points in a unified coordinate system.
- Physical geodesy – To describe the Earth's fixed and time-variable planetary gravity field and to define a unified mass-centered coordinate system.

The application of space technology will enable the achievement of these geodetic objectives with a precision essentially unachievable by the use of other techniques.

There are a number of applications of remote-sensing techniques related to these geodetic objectives:

- Establishment of properly scaled geodetic control essential to topographic mapping (one of the basic requirements for economic development of natural resources)
- Horizontal positioning of oceanographic measurements in the accurate location of ships and the establishment of precisely located permanent control points on the ocean bottom (which will provide the control necessary in the practical utilization of the marine environment)
- Monitoring of horizontal and vertical land motions along fault zones with respect to a fixed coordinate system and comparing absolute motions of widely separated points
- Accurate monitoring of the ocean surface for rapid detection of tsunami waves
- Determination of tides
- More accurate determination of the wandering terrestrial pole
- Investigation of time-variable processes on the deep ocean bottom such as deposition and erosion
- Measurement of continental drift
- Measurement of glacier motion with respect to absolute datum

Accomplishment of these applications is dependent on the continuation of the National Geodetic Satellite Program. In geometric geodesy, use can be made of the already programmed GEOS-B and a replacement for the passive ECHO-1 (designated PAGEOS-B). A ground network of primary and secondary stations with ties between the network and the "absolute" stations can provide data required to transform the coordinates of the network into a mass-centered system. Also necessary are scientific reference stations requiring higher precision for basic geophysical investigations.

In physical geodesy, the launching of four satellites (designated GEDY) at 1-year intervals will be required in addition to GEOS-B. These satellites will carry laser reflectors and Doppler devices and will enable passive camera tracking.

A.1.4 Geography, Cartography, and Cultural Resources

Geography is the science of the Earth's surface, form, physical features, natural and political divisions, climate, productions, and population. Information pertaining to this science is concerned with the interaction and flow of natural phenomena and human activities. Geography is closely related to cartography which is the depiction of the physical surface of the Earth in the form of maps useful for scientific, engineering, and commercial purposes.

The urgent need in geography is for uniform data collected over short periods of time on a worldwide basis to calibrate scientific models such as those dealing with urban growth (Ref. 10) and expansion. Information currently available varies not only in quantity and quality but also in space and time.

Synoptic coverage and ground truth combined in an articulated information system can provide more accurate and more efficient delimitation of the broad physical and cultural patterns on the Earth's surface (Ref. 11). Specific applications of space technology are:

- Plotting of gross surface patterns to reflect man's activities on a yearly, seasonal, or shorter time basis to enable inventory of cultural and natural resources and thus evaluate the potential for further land development
- Rapid determination of areas of poverty or of distress resulting from natural disasters
- Determination of the dynamics of change, expansion of cities, conversion of agricultural land to other uses, and changes in snow and ice coverages, through repetitive coverage
- Monitoring of transportation on sea, air, continental waterways, railroads, and highways
- Estimations of census through synoptic coverage coupled with ground truth
- Global map making

These applications, which are by no means complete, emphasize the fact that satellite global data collection will meet many of the information needs of geographers and cartographers. A central data bank may be the best way to store, process, and disseminate these data.

A. 1. 5 Geology

The basic objective in the application of geology to the study of the Earth's crust is for the location of minerals, oil, and gas. Exploration theory leads to the recognition and localization of structural and/or lithologic conditions which are favorable for mineral occurrence. The use of satellites will result in broad synoptic surveys with multispectral photography and radar (and possibly magnetic and gravity field mapping) permitting the trained observer to locate geologic structure and features not evident in small-scale observations (Refs. 12, 13, and 14).

A basic problem exists in applying remote sensor data, i.e., the electromagnetic and force-field data, to the geological process used in exploration. Extensive effort is required in the interpretation of the new types of data proposed. This effort is undeniably more essential for the geologist than the R&D effort on the sensors themselves. Without an increase in the understanding of geologic remote sensing, more refined sensor systems would be superfluous. At present a quantum jump in understanding sensor data is required.

Among the applications that exist in geology are the following:

- A geologic resource study of North and South America by satellite using synoptic space photography and radar imagery
- Development of line-edge enhancement techniques to improve structural analysis
- Acceleration of the development of radar imagery interpretation techniques to optimize the information extraction

After these applications have been made, a more sophisticated program can be undertaken which will emphasize extensive study of remotely sensed data at other wavelengths. This program would seek to expand beyond the simple structural linement maps to a compositional display based on recognition of electromagnetic signature and the various spatial and temporal distributions of these signatures. Such a program should include the following aspects:

- Research on information available for mineral identification at various wavelengths
- Development of the theory of surficial geology to understand how useful geologic information for exploration can best be obtained from remote sensing
- Study of the various physical aspects of the coupling between geological conclusions and the sensing process such as sun angle, polarization, spectral properties of minerals, vegetation, and soil structure in the surface layer
- Sensor development in longer wavelength radar (providing deeper penetration), and multichannel single aperture scanning systems in the infrared, visible, and ultraviolet wavelengths
- Research in the data processing and data interpretational aspects of future systems

A. 1. 6 Hydrology and Water Resources

Hydrology, the science of the properties and laws of water, is concerned with the study of water in the atmosphere and on the earth, and the application of water resources to the pursuits of man. The objectives of hydrology fall into two categories: those objectives required immediately for better utilization of water resources and those required for a better understanding of the hydrologic cycle.

Applications relating to these objectives are (Ref. 15):

- The ability to forecast precipitation, temperature, and stream flow (short and long term)

- Reliable information on evaporation from water and land, transpiration, and consumptive water uses
- Reliable information on rainfall distribution and drainage basins
- Accurate soil moisture data over large areas
- Accurate data on snow cover, snow water content, and the rate of snow melt
- Better rainfall information for the inaccessible headwaters of drainage basins
- Better information on currents and tidal effects in harbors and estuaries, the extent of salt water intrusion, and the effects of stream flow and wastes on coastal waters

Accomplishment of many of these applications is anticipated by a hydrology communications satellite, HCS, a hydrology sensing satellite, HSS-1, and the HSS-2, an advanced sensing satellite.

The HSS-1 would utilize:

- Panchromatic and color photography
- Multispectral infrared photography
- Infrared and radar imagery

These instruments would provide data on:

- Snow cover and ice occurrence in rivers and glaciers
- Near-shore underwater detail in coastal waters, estuaries, large lakes, and major reservoirs
- Details of saline intrusion, circulation patterns, and visible pollutant distribution in coastal waters, estuaries, and lakes
- Geomorphology of river basins and changes in shoreline

The HSS-2 may provide the following:

- Improved resolution of HSS-1 items
- Areal extent, intensity, and total amount of rainfall

- Horizontal movement of atmospheric water vapor
- Salinity and biochemical characteristics of water
- Rates of stream flow
- Elevation of reservoirs, rivers, and lakes

Other developments of importance are:

- Remote sensors to determine snow cover, density, depth, and temperature
- Side-looking radar systems for ice surveillance
- Detailed study of existing imagery to determine its applicability

Prompt processing and dissemination of such photographs to interested users is required.

A.1.7 Oceanography

The primary objective of oceanography is the study of the physical and biological character of the ocean at all depths, its behavior as a dynamic global system, and its interaction with the coasts and inhabitants of the coast. An aircraft or satellite remote-sensing program will sample the top few hundred feet of the ocean which is that portion of greatest importance to man in his fishing, shipping, and coastal activities; it includes the photosynthesis zone providing the entire biological resonance of the sea and the air-sea interface across which the major energy inputs from sun and wind occur.

Among the specific applications of space technology, areas with particular promise are (Ref. 16):

- Prediction of fish location by measurement of sea temperature and chlorophyll content
- Surveillance of coastal waters by monitoring temperature changes, storm surges, wave and current conditions, pollution patterns, river run-off, sedimentation, and shore erosion

- Surveillance of the oceans on a global scale as a major service to shipping; monitoring of sea state, ocean currents, floating ice, and ice cover in northern latitudes
- Regular recording of the height of the ocean over the globe to contribute to calculations of tidal current and circulatory dynamics of the ocean

The most significant parameters and associated equipment required for these applications are:

- Sea surface temperature (IR, microwave, and telemetry)
- Imagery (photography, imaging radar, and IR)
- Drift rate of floating objects (IR, buoys, etc.)
- Sea ice and bergs (imagery by radar)
- Spectrograms (chlorophyll detection, sea color, bioluminescence, and fluorescence)
- Sea state (radar roughness)
- Dynamic topography of sea surface (radar or laser altimeter)

This will require vigorous R&D programs closely coupled to ground-truth testing programs. Perhaps the greatest benefits can be accrued by evolvement of a combined system utilizing (1) aircraft and satellites with various remote sensors, (2) collection and relay of data from buoys and ships, and (3) central oceanographic data handling banks.

A. 1. 8 Summary of Earth Resource Information Requirements

The objectives of the various disciplines and major applications of remote sensing have been indicated to establish the envelope of user requirements upon which the indexing of data will be accomplished. These applications must then be analyzed as to the potential phenomena which will make up the required observations. This, in turn, requires the distinction of specific characteristics or observables. These characteristics are not all observable at the same resolution level or wavelength.

Therefore, a combination of remote sensing instrumentation of varying resolution levels and wavelength will be flown simultaneously to obtain the required coverage.

An organized presentation relating spatial resolution to earth resource discipline application areas is hown in Table A-1. A ready reference comparison of discipline area, application, phenomena and characteristic, resolution required, operating wavelength, and equipment is presented in Tables A-3 through A-6 (Ref. 17). Appendix A contains an indicative correlated summary of observables across a variety of discipline lines which is useful in establishing the envelope of data elements required by potential users.

A.2 LUNAR SCIENCES REQUIREMENTS

A.2.1 Introduction

Data received from the Moon will contribute to the answering of questions in eight different sciences and technologies (Ref. 18). The specific breakdown of lunar information into these eight disciplines is not a rigid allocation nor is it essential to an indexing system, but is rather an effort to cover the types of information to be handled within the system. These disciplines are:

Cartography
Geology
Geochemistry
Geophysics
Biology
Physics
Astronomy
Engineering

Collection of information about the Moon will be accomplished by surface exploration, first at points and then later by traverses. Knowledge gained in this manner will then be extrapolated to cover other areas of the Moon by use of sensors carried in lunar-orbiting vehicles which can survey vast areas within short time periods. The over

Table A-1

NATURAL RESOURCE APPLICATIONS GROUPED BY RESOLUTION REQUIREMENTS[a]

Spatial Resolution[b] (meters)	Agriculture/Forestry	Geography[c]	Geology[c]	Hydrology	Oceanography[c]
< 20	Timberline, waterline and snowline studies Grass, brush, and timberland interfaces Vegetation density Tree count Tree crown diameter Crop species Crop acreage Irrigation studies Fields of smaller sizes, 10 acres or less Livestock census	Population and cultural studies Fishing boat activities Land use studies Topo-mapping 1:250,000 and larger scales Plant cover and soils Forest types Thematic mapping	Delineation of small small folds Delineation of small linear elements Delineation of stratigraphic sequences Lithologic units Soil compaction Slope stability Permeability studies Ore deposits Local geothermal anomalies Tectonic studies Glaciological studies (local)	Groundwater discharge Subaqueous features of lakes Detection of water pollution, inland areas (rivers, lakes, bays) Effluents of major rivers Monitoring lake and reservoir levels Evapotranspiration Water surface roughness Rainfall Salt content Drainage basins Water regimens of valley glaciers Snow surveying Reservoir sedimentation	Ice surveillance Snow/ice and ice/water interface studies Wave profile Shoals and coastal mapping (bottom topography) Currents (long shore) Coastal marine processes (tidal variations) Estuarine and shoreline morphology Sea level and sea slope Sea mammals
20 – 100	Timberline and snowline studies Fields of larger sizes, 10 acres or more Soil temperature Detection of forest fires	Water resources Gross cultural studies Geomorphology studies Gross land-use studies Topo mapping, scales smaller than 1:250,000 Pollution (air, land, water) Thematic mapping	Delineation of folds and linear elements Soil compaction Slope stability Gross geothermal studies Geomorphic studies Glaciological studies Mineral belts Permafrost	Evapotranspiration Water surface roughness Rainfall Salt content Drainage basins Water regimens of valley glaciers Snow surveying Reservoir sedimentation	Sea survace thermal mapping Cold region thermal structure Fresh/salt water interface Water pollution, large areas, oceanic, harbor areas Ocean waves Currents (offshore) Biological studies (fish & other populations) Wave refraction studies Volcanic activity
100 – 300	Timberline, snowline and desertline studies Fields of gross sizes (rangelands, etc.)	Land-use studies Thematic mapping	Delineation of large folds Delineation of linear elements Lithologic units Geothermal studies Volcanic studies Metallogenic provinces Inventory of ice features	Evapotranspiration Water surface roughness Rainfall Monitoring lake and reservoir levels	Currents (offshore) Water masses Upwelling areas
300	Soil moisture	Cloud studies Land-use studies Thematic mapping	Delineation of large folds and faults Slope stability	Evapotranspiration Rainfall Snow surveying	Sea state Delineation of pack and cap ice margins Sea water color analysis

(a) Requirements submitted by agencies and representatives of the various geoscience disciplines involved.
(b) Resolution (side dimension in meters of resolvable objects) is here defined as the ability to resolve Earth-surface distances per line pair at a target contrast of 2:1 under actual flight conditions. Resolutions less than 30 meters will not necessarily be attained on early missions.
(c) Cartography, geodesy, and geophysics applications integrated in geography, geology, and oceanography.

Table A-2

AGRICULTURE/FORESTRY -- PARTIAL SUMMARY OF OBJECTIVES AND REQUIREMENTS

Application	Phenomenon	Characteristic	Detector	Resolution (meters)	Spectrum
Lumber and Conservation	Inventory and distribution	Boundaries/topographic and texture	Metric camera	6–35	0.4–0.9μ
			Multispectral camera	6–35	0.4–0.9μ
			IR imager	6–35	3–14μ
			Microwave imager	30–300	3–35 GC
			Radar imager	6–35	0.5–10 GC
		Yield	Multispectral camera	3–20	0.4–0.9μ
			IR imager	20–60	3–14μ
			Spectrometer	20–200	3–14μ
			Radiometer	20–200	3–14μ
			Radar imager	6–35	0.5–10 GC
		Type and density	Multispectral camera	3–20	0.4–0.9μ
			Metric camera	3–20	0.4–0.9μ
			IR imager	6–35	0.4–3μ
			Spectrometer	6–35	0.4–14μ
			Radiometer	6–35	0.4–14μ
			Radar imager	3–20	0.5–10 GC
			Multispectral camera	3–20	0.4–0.9μ
			IR imager	3–20	0.4–3μ
	Fire and reclamation	Damage	IR imager	30–200	3–14μ
			Metric camera	30–200	0.4–0.9μ
		Pattern and temperature discontinuity	IR imager	30–300	3–14μ
			Metric camera	30–200	0.4–0.9μ
		Soil moisture/texture	IR spectrometer	6–60	3–14μ
			IR radiometer	6–60	3–14μ
			Radar scatterometer	30–300	0.3–0.6 GC
			Microwave imager	30–300	3–35 GC
Farming	Inventory and distribution	Boundaries/topographic	Metric camera	3–20	0.4–0.9μ
			Multispectral camera	3–20	0.4–0.9μ
			IR imager	6–35	3–14μ
			Microwave imager	30–300	0.3–35 GC
			Radar imager	6–35	0.5–10 GC
		Crop statistics type/density and yield	IR imager	20–35	3–14μ
			IR spectrometer	20–35	3–14μ
			IR radiometer	20–35	3–14μ
			Metric camera	20–35	0.4–0.9μ
			Multispectral camera	20–35	0.4–0.9μ
	Infestation	Pattern/damage	Multispectral	2	0.4–1.0μ
			tracking tele	6–35	
			IR imager	6–35	3–14μ
			Spectrometer	6–35	3–14μ
			Radiometer	6–35	3–14μ
			Multispectral camera	6–35	0.4–0.9μ
	Usable land	Soil texture/moisture and vegetation	Multispectral camera	30–200	0.4–0.9μ
			IR imager	30–200	3–14μ
			Spectrometer	30–200	3–14μ
			Radiometer	30–200	3–14μ
			Metric camera	30–200	0.4–0.9M
			Microwave radiometer	30–300	0.3–35 jc
			Radar scatterometer	30–300	0.3–0.6 GC

Table A-3

GEOGRAPHY – PARTIAL SUMMARY OF OBJECTIVES AND REQUIREMENTS

Application	Phenomenon	Characteristics	Detector	Resolution (meters · F)	Spectrum
Transportation	Identify facilities	Terminals, buildings	Pan photo	3 – 15	0.4 – 0.7μ
	Roads, RR tracks	Roads, RR tracks	Metric photo	5 – 20	0.4 – 0.9μ
		Traffic count	Pan/UHR photo	2 – 5	0.4 – 0.7μ
	Location of new facilities	Maps – 1:50,000 scale	Pan photo	5	0.4 – 0.7μ
			Metric photo	10 – 20	0.4 – 0.9μ
		Maps – 1:250,000 scale	Metric photo	10 – 30	0.4 – 0.9μ
		Cultural/economic factors	Multi photo	10 – 100	0.3 – 1.0μ
			Radar imager	10 – 100	5 – 10 GC
Navigation	Topography	Maps – 1:100,000 scale	Metric photo	10 – 20	0.4 – 0.9μ
		Maps – 1:250,000 scale	Metric photo	10 – 30	0.4 – 0.9μ
Urban Planning	Locate settlements	Boundary and topography	Metric photo	10 – 100	0.4 – 0.9μ
			Microwave 1	10 – 100	0.4 – 35GC
	Type of settlements	Color, texture, color contrast	Multi photo	15 – 30	0.3 – 1.0μ
			IR imager	30 – 100 1°	1 – 3μ
	Distribution in settlement	Pattern density of dwelling	Metric photo	3 – 15	0.4 – 0.9μ
			Multi photo	10 – 100	0.3 – 1.0μ
			IR imager	10 – 100 1°	0.3 – 14μ
			Radar imager	10 – 100	5 – 10GC
	Occurrence of recreation areas	Color, texture, shape	Multi photo	10 – 100	0.3 – 1.0
			Radar imager	10 – 100	5 – 10GC
	Population dist.	Population count	UHR photo	1	0.4 – 0.7μ
	New sites for settlements	Map – 1:50,000 scale	Pan photo	5	0.4 – 0.7μ
			Metric photo	10 – 20	0.4 – 0.9
		Map – 1:250,000 scale	Metric photo	10 – 30	0.4 – 0.9
		Cultural/economic factors	Multi-photo	10 – 100	0.3 – 1.0μ
			Radar imager	10 – 100	5 – 10GC
		Land use intensity	Multi-photo	10 – 100	0.3 – 1.0μ
			UHR photo	2 – 5	0.3 – 1.0μ
			IR imager	100 1°	3 – 14μ
		Spectral signature	IR spec/rad	200	0.4 – 3μ
Industrial Planning	Classification of facilities	Topographic features	Metric photo	10 – 20	0.4 – 0.9μ
		Color, texture	Multi-photo	10 – 100	0.3 – 1.0μ
		Heat budget	IR radiometer	200	8 – 14μ
		Detail features	UHR photo	2	0.3 – 1.0μ
		Spectral signature	IR spec/rad	100	0.4 – 3μ

LOCKHEED PALO ALTO RESEARCH LABORATORY
LOCKHEED MISSILES & SPACE COMPANY
A GROUP DIVISION OF LOCKHEED AIRCRAFT CORPORATION

Table A-4

GEOLOGY – PARTIAL SUMMARY OF REQUIREMENTS AND OBJECTIVES

Application	Phenomenon	Characteristic	Detector	Resolution (meters C°)	Wavelength
Petroleum, coal and ore detection	Surface patterns	Lithology	Pan camera	6	0.4–0.7 μ
			IR spec	300 1.0	8–15 μ
			Radar imager	30	~10 GC
		Outcrops	UV imag/spec	30	0.3–0.6 μ
		Magnetic field variation	Magnetometer	300–1000	N/A
		Small folds	Metric camera	10	0.4–0.7 μ
		Small drainage patterns	Radar imager	20	~10 GC
			IR imag	300 1.0	8–15 μ
			UV imag/spec	60	0.3–0.6 μ
		Large folds	Radar imager	300	~10 GC
		Large drainage patterns	UV imag/spec	150	0.3–0.6 μ
			Metric camera	6	0.4–0.7 μ
			IR imager	600 0.1	8–15 μ
Volcano prediction	Surface feature changes	Temperature variation	IR imager	300 1.0	8–15 μ
			Pass micro imag	150 1.0	
			Multispec cam	30	0.3–1.0 μ
		Lithologic identification	IR spec	300	8–15 μ
			IR imag	300	8–15 μ
		Spatial relations	Multispec cam	50	0.3–1.0 μ
			Radar imager	200	~10 GC
Earthquake prediction	Surface stress and discontinuities	Linear microtemperature anomalies	Multispec cam	30	0.3–1.0 μ
			IR imager	40 1.0	8–15 μ
		Slope distribution	Metric camera	30	0.4–0.7 μ
		Crust anomalies	Gravity meter	10^{-11} g's	N/A
		Soil moisture	Multispec cam	50	0.3–1.0 μ
			IR imager	200	8–15 μ
Engineering geology	Geothermal power sources	Temperature anomalies	IR imager	300	8–15 μ
		Surface gas	Absorption spec	200	0.3–0.7 μ
	Landslide prediction	Soil moisture	IR imager	300	8–15 μ
		Slope distribution	Metric camera	30	0.4–0.7 μ
		Crust anomalies	Gravity meter	10^{-11} g's	N/A

Table A-5

HYDROLOGY – PARTIAL SUMMARY OF REQUIREMENTS AND OBJECTIVES

Application	Phenomenon	Characteristic	Detector	Resolution (meters C°)	Wavelength
Water inventory	Water inflow into basins, rivers, and streams	Effluents of major rivers	Pan camera	10	0.5–0.7μ
			IR spec	100 1.0	8–15μ
			UV imag/spec	100	0.3–0.6μ
		Lake and reservoir levels	Pan camera	300	0.5–0.7μ
			Multi teles	3	0.5–0.7μ
			Laser alt	1/3 vertical	~1μ
		Drainage basin features	Multispec cam	100	0.2–1.0μ
			IR spec	100 1.0	8–15μ
Flood control	Excess surficial water	Location and areal coverage of floods	Metric cam	100–300	0.5–0.7μ
			Radar imager	10–50	~10 GC
		Rainfall monitor	Radar imager	30	~10 GC
			IR imager	300	8–15μ
			Pan camera	500	0.5–0.7μ
Water pollution	Natural and industrial pollution	Color	Pan camera	10	0.5–0.7μ
			IR spec	100 1.0	8–15μ
		Spectral signature	UV imag/spec	100	0.3–0.6μ
		Salt content	Multispec cam	100	0.3–1.0μ
Water conservation	Evaporation and transpiration	Evapotranspiration	IR imager	300 1.0	8–15μ
			IR spec	300 0.5μ	8–15μ
			Pass micr spec	1000 0.1	
			Multispec cam	30	0.3–1.0μ
			Radar imag	30	~10 GC
Water resources	Seeps and springs	Temperature variation	IR imager	10–50	8–15μ
			Multispec cam	10–100	0.3–1.0μ
		Quality of water	Multispec cam	10–100	0.3–1.0μ
	Glaciology	Frozen water inventory	IR spec	100 1.0	8–15μ
			Metric camera	30	0.5–0.7μ
		Snow survey	Metric cam	30	0.5–0.7μ
			Pass micro rad	3000 0.5	

Table A-6

OCEANOGRAPHY -- PARTIAL SUMMARY OF OBJECTIVES AND REQUIREMENTS

Application	Phenomenon	Characteristics	Detector	Resolution (meters °C)	Spectrum
Shipping	Sea state	Wave height	Radar scatt	1000	0.4 – 3 GC
	Currents	Surface temp. gradients	IR imager	30 – 300	8 – 13μ
			IR radiometer	30 – 300 0.1 – 0.2	10 – 12μ
			Microwave imag	30 – 300	1 – 6 GC
			Microwave rad	20 – 150 0.5 – 2.0	1 – 6 GC
		Water color tones	Multiband photo	30	0.3 – 1.0μ
	Hazards Icebergs Ice masses	Temperature anomalies	IR imager	20	8 – 13μ
			IR rad	20 150 1.0 – 2.0	10 – 12μ
		Water/ice interface	Radar imager	20	1.5 – 10 GC
			Microwave rad	30 – 300 1.0 – 2.0	1 – 6 GC
			Microwave imag	30 – 300	1 – 6 GC
Sea food production	Upwelling	Surface temperature gradient	IR imager	150	8 – 13μ
			IR radiometer	150 0.5 – 1.5	10 – 12μ
			Microwave imag	30 – 300	1 – 10 GC
			Microwave rad	30 – 300 0.5 – 1.5	1 – 10 GC
	Currents and eddies	Water color tones	Multiband photo	30	0.3 – 1.0μ
		Surface temperature gradient	IR imager	30 – 300	8 – 13μ
			IR radiometer	30 – 300 0.1 – 2.0	10 – 12μ
	Bottom topography	Wave refraction and color tones	Multiband photo	25	0.3 – 1.0μ
	Bioluminescence	Color Tones	Multiband photo	60	0.3 – 1.0μ
	Oil slicks	Vapor signature	Absorption spec	60	0.3 – 0.6μ
Coastal Geography	Shoreline topography	Land/water interface	Metric photo	20	0.4 – 0.9μ
			Radar imager	20	1 – 30 GC
		Color tones and contrast	Multiband photo	20	0.3 – 1.0μ
	Water effluents and sediment transport	Water color tones	Multiband photo	20	0.3 – 1.0μ
	Sea levels and slopes	Surface Elevation	Radar alt	0.2 – 1.0 vert.	10 – 60 GC
			Laser alt	0.2 – 1.0 vert.	0.3 – 0.9μ
Marine Biology	Bioluminescence	Color tones	Multiband photo	60	0.3 – 1.0μ
	Red tides	Color tones	Multiband	60	0.3 – 1.0μ
	Plankton	Color tones	Multiband photo	30	0.3 – 1.0μ
	Schools of fish and algae	Color tones	Multiband photo	30	0.3 – 1.0μ

flying sensors will scan and photograph parameters which have previously been scanned and photographed on very limited areas of the lunar surface.

A.2.2 Cartography/Geodesy (Selenography/Selenodesy)

Selenography/selenodesy has as its general objective the making of maps of the Moon's surface. The maps would be of various scales and contour intervals, and would depict and name most lunar surface features. The maps would have grids established on a longitude-latitude grid which could, in turn, develop other grid or coordinate systems.

Basic to the making of accurate lunar maps is the establishment of the precise figure of the Moon (i.e., its departure from a sphere) and a system or net of control points of the required accuracy. Accurate lunar surface maps are needed to pinpoint the location of data collection points and to permit precise navigation.

To fulfill the demands of map making, the required information elements are (Ref. 19):

- Precise metric photography of the entire lunar surface with end, sidelap, and stereo coverage
- Control network on the lunar surface consisting of a series of accurately located points
- Precise knowledge of conditions of photography
- Complete orbital data of cameras to give the figure of the Moon
- Knowledge of lunar tides
- Slope studies
- Recognition of lunar surface features
- Measurements of lunar features
- Measures of roughness

A summary of the information channeled back from the Moon and related to cartography yields:

- Slope data – percentage breakdown
- Grid and coordinate systems – latitude and longitude, grid areas, military grids

- Planimetric and contour maps at large and small scales, large and small contour intervals
- Lunar feature maps
- Lunar surface roughness analysis (photogrammetric and photometric) by area
- Lunar shape or figure
- Lunar tide data (solid-body tide)

Cartography requires specific instrumentation so that data returned to Earth can be used to produce maps with the help of plotting instruments. The following are needed:

- Metric cameras (orbit equipment) consisting of multiband photography, UV IR visual, 70-mm and 9-1/2-in. width film, with Earth return by telemetry
- Radar (orbit instrument) for imaging – requires electrical measurements on lunar surface for calibration
- Radio transponders (landed on Moon) to receive and transmit signals used for precise distance measurement
- Radar altimeter (in orbiter) to measure altitude of cameras precisely
- Chronometer and star camera (landed on Moon) to measure location and give fix on control points
- Corner reflectors to pinpoint surface locations on radar imagery, that is, control points or ground truth locations

A.2.3 Geology (Selenology)

Geology of the Moon is concerned with the geologic mapping of the lunar surface and subsurface, the study of the processes that affect the lunar surface, and the determination of the age, origin, and history of the lunar features. The geologic mapping entails putting boundaries around the areal limits of mappable rock units and describing their structural attitudes. The subsurface mapping will be made possible by exposures in drill holes and in crater walls. The study of the processes responsible for present and past lunar surface morphology entail investigating erosion, transportation, and deposition as well as internal movement or tectonism (Ref. 20). The age, origin, and history of the Moon and of lunar surface and subsurface features requires investigation based on data in part given by geochemical and geophysical studies.

To achieve the geological objectives, the following space technology applications are required:

- Base maps must be created (see cartography)
- Manned or unmanned landings must be made to achieve lunar exploration of surface geology and to collect samples
- Surface and orbital imagery must be done in selected wavelengths with long-continued monitoring of processes required in some instances
- Ground-truth sites must be mapped and studied covering crucial areas (Ref. 21)
- Drilling and crater wall studies should be carried on to get information about the lunar subsurface
- Specific study will be made of the predominant processes of impact cratering and volcanism

Among the parameters of interest are:

- Surface roughness – this parameter relates to astronaut landing, mobility of wheeled vehicles, geologic origin of rock, and impact cratering forces
- Surface features – origin, nature, and significance of rills, ridges, and crater pits
- Rock type of a surface area relating to origin, process, and history
- Structural features – folds, faults, and joints
- Evidence of processes – landslides, rolling stones, and secondary cratering by flying ejecta
- Evidence of external effects – results of meteorite impact, radiation darkening, and blanketing by ejecta
- Formation boundaries – as distinguished by changes in albedo, roughness, color, and reflectivity
- Age relations of superposition or transection

Instruments required are both orbiter and surface type.

Orbiter instruments are:

(1) Cameras of metric mapping precision type using multiband applications of film, 70-mm and 9-1/2 in. width, for detection of hotspots and for volcanism
(2) Radar IR and UV scanner imagers for detection of rock composition and for subsurface structure depiction

Orbit image and scan records will be telemetered to Earth or returned intact in a capsule.

Surface instruments are:

(1) Sample collecting tools, containers
(2) Geologic tools – clinometers for measuring dips of rocks, hammers, maps, notebooks, tape recorders for notes, TV cameras, and multiband hand cameras for recording what is seen
(3) Gas chromatographs – for chemical analysis of gas at suspected active volcanic sites
(4) Mass spectrometers – for chemical analysis for elemental composition of rocks or gases
(5) X-ray spectrometers – for rock composition analysis of samples
(6) Seismic equipment – for measurement of volcanic and meteor impact activity

A. 2.4 Geochemistry (Selenochemistry)

Geochemistry is supportive to geology but is itself an important area of lunar science. It is concerned with the chemistry of lunar soil, rocks, and atmospheres, and their origin, age, and history. The study of volcanic phenomena on the Moon, of radioactivity, and of age dating of rocks will be very important. The analysis for age, elemental, and mineral composition will be done by sampling gases and rock for in situ, lunar laboratory, or Earth-returned study.

Samples collected on the surface, from drilling, and from sources of volcanic gas will be analyzed for type, origin, age, history, fabric, structure, and value for engineering use. Knowledge of elemental and mineral composition is a requirement for geologic mapping, establishment of ground truth, and the performance of lunar surface exploration from an engineering or practical standpoint. Once surface determinations are made in significant areas, overflights by orbiting instruments can be keyed into these results and the knowledge extended to other parts of the lunar surface. Orbiting sensor acquisition requirements are:

- Sensing thermal anomalies (volcanic activity)
- Sensing rock types (for mapping)
- Sensing gas emissions (volcanoes)
- Sensing roughness (cratering studies, and trafficability and landing hazards)
- Sensing radioactivity (for study of heat regime)

Mineral names, rock names, and bulk chemical analyses are descriptors for the soil and rock samples of the lunar surface. In addition, there are the following parameters applicable to samples of soil and gas:

- Age of the sample – as obtained from radioactive age dating
- Origin assigned to sample – volcanic (or igneous), sedimentary, meta-type
- Fabric, structure, changes, and texture of rock sample
- History deduced from origin, such as indication of burial depth of cooling as inferred from crystal size
- Likely origin of gas sample, such as from a volcanic vent

Surface exploration instruments consist of collecting apparatus and containers plus the following analytical and sensing instruments:

- Gas chromatograph – for gas analysis of small quantities
- Mass spectrometer – for analysis of minute quantities of gas or solids
- X-ray spectrometer – for mineral composition analyses of rocks
- Gieger counter – for radioactivity detection
- Wet chemical analytical equipment for standard chemical studies in a lunar laboratory for elemental analysis

- Radio and TV links to base vehicle or Earth
- Voice tapes of observations
- Maps for location of samples

Overflight instrumentation in orbiters will be the following:

- IR-visible-UV sensors for locating thermal anomalies and mineral signatures
- Radioactivity sensors

A.2.5 Geophysics (Selenophysics)

As applied to the Moon, geophysics studies the nature, origin, history, and mapping of parameters such as gravity, magnetism, and solid body waves. This science includes tectonism, differentiation, moonquakes, and active and passive seismic exploration.

The specific measurements needed to provide an understanding of the various lunar body geophysical parameters are the following:

- Magnetic field recordings taken at frequent intervals in time and over the lunar surface
- Gravity field recordings taken in a similar manner of both absolute readings (a few) and relative readings (a dense pattern)
- Passive seismic recordings using seismographs of both long and short periods to record waves produced by lunar solid body vibration
- Monitoring equipment (12 or more instruments) to establish a net to record natural quakes of internal and external origin (heavy meteorites, impact of heavy space vehicles)
- Samples taken to establish palomagnetic fields

Data to be expected from lunar geophysical research include passive and active seismology data and maps. Passive seismology data involve recordings of seismographs on multiple-track tapes; as reduced data this yields depth, location, and strength of quakes or impacts. Active seismology data are recovered on tapes from planned and sited explosions set off to delineate subsurface density interfaces and the structure of folds and faults, volcanic features, and stratigraphic relations such as nonconformities and thinning or thickening beds. The maps derived will have gravity or magnetic readings added to them so that anomalies can be located to indicate areas of high or low magnetism or gravity leading to location of iron concentrations, isostatic adjustment, or presence of unusually dense or light rock masses.

Orbiting measurements include gravity-gradient determinations depicting the rate of change of the gravity field. Also, the magnetic field can be sampled in orbit. Since lunar field strength is expected to be low, sensitive magnetometers will be needed. Surface or ground-truth instrumentation will consist of gravimeters (pendulum) for a few absolute gravity determinations plus gravimeters of a mobile type for numerous rapid gravity determinations. Seismographs will be of two types, mobile for active seismic experiments and essentially stationary for passive monitoring of quakes and high-energy impacts.

A.2.6 Biology

The primary objective of the biological sciences in lunar exploration is the finding of evidence of organic life. If life is found, its extent and nature, history, origin, and processes would be studied. Uncontaminated samples from the richest organic sources would be sought. Prevention of contamination by astronauts and Earth laboratories must also be attempted. Ices and gases, if found, would be sampled for organic evidences. Test sites to study the growth of Earth life-forms on the Moon would be provided.

Data sought would be based on:

- Samples – uncontaminated Earth-returned samples would afford the best test for evidence of life
- Analyses – chemical analyses, gas chromatography, mass spectrometry, all would provide in situ or lunar laboratory life presence tests, to prevent Earth return of many unproductive samples

- Observation of macroscopic and microscopic samples to determine if they contain or support life forms, in situ or in the laboratory
- Observation of environmental conditions such as temperature, water or ice presence, and cyclical rejuvenation of spores in any aqueous environment
- Subsurface conditions will be described by means of samples taken by drilling
- Evidence of life would be mapped in terms of its density
- All electromagnetic spectrum sensors would be calibrated using an area where life has been detected to determine ground-to-overflight correspondence of sensitivity of instrumentation
- Surface and orbiter exploration would have gas "sniffer" instrumentation

To summarize the biological data we would expect to note:

- Listing of species of life forms
- Lists of suggestive organic compounds
- Notes on processes exhibited by any life forms discovered
- TV monitoring of such forms over long periods – also for Earth forms undergoing lunar tests
- Characterization of vigor of all reported life forms, including stages of encystment or encapsulation
- Ground-truth data using sensors for later tests of correspondence to remote sensing signatures
- Voice tapes of descriptions of sterile techniques

Instrumentation necessary for the biological lunar study includes:

- Gas chromatographs sensitive to key organic chemicals and employing systems requiring very small gas volume
- Mass spectrometer able to operate at low concentrations
- Wet chemical analysis in missions where a laboratory is feasible, manned or unmanned
- Multiband photography and wide range spectral sensors to catch signatures of any organic discovery site

- Uncontaminated sample collection required to ensure that organic chemicals are not speciously counted
- Long-time monitoring of test and discovery areas via TV link and sensor link to environment
- Radiation and particle influx monitoring on a long-term basis

A. 2. 7 Physics

Within the compass of physics pertaining to the lunar surface and surrounding volume are the phenomena relating to fields and particles and other forces and effects not covered under geophysics. Several effects will be induced by solar phenomena and several by the Earth's influence. Important are solar particles and radiation, lunar magnetic and gravitational fields, thermal gradients, atmospheres and atmospheric pressures, temperatures, and densities. Electrical phenomena such as conductivities and resistivities of atmosphere and surface materials will be studied.

Most of the data will be read as numeric readings from instruments. Pressure gauges, geiger counters, gravity gradient instrument[illegible]meters, magnetometers, and electrical parameter instruments will be employed, many of which will be continuous long-term monitors.

A. 2. 8 Astronomy

Astronomical research concerning the Moon will be of two categories: (1) astronomy of the Moon and (2) astronomy of other astronomical objects with the Moon as a "platform."

The objectives of the first category will be the study of the astronomy of the Moon as a member of the solar system. This would be concerned with refinements of the lunar ephemeris, yielding more exact Earth-Moon distance calculations, more exact libration predictions, and more precise orbital data.

The second category would be that of lunar-based astronomy directed toward study of the sun, the balance of the solar system, and the stars. The steady atmosphereless conditions for observation would be a great advantage to astronomical observers. Observations from lunar orbiters would be somewhat less advantageous but would yield data pertaining to lunar occultations of stars and planets as well as observations at times of solar exposure and times of screening from solar effects.

The use of instruments such as radar altimeters, star field cameras using multiband film, corner reflectors for pinpointing lunar surface locations, and telescopes for visual, x-ray, and radio observations will greatly enhance astronomical knowledge.

A.2.9 Engineering Properties

The purpose of engineering studies is the establishment of information about the operation of instruments and equipment on the lunar surface. Considerable scientific information is necessary on bearing strength of soils, slumping and settling of the surface, safety factors for human living and exploring, and human factors such as trafficability, vision, movement, and shelter. All the instrumentation will undergo "shakedown" during actual operation and will transmit operational data to a central laboratory or back to Earth. Construction requirements will receive "feedback" from geology, physics, and other sciences to establish a practicable, safe, and inexpensive shelter. As an example, caves left by old emptied lava flow tubes would furnish protection against micrometeorites and radiation, large heat cycles, and possibly might yield ices of useful varieties.

A very important aspect of engineering studies would be the soil mechanics investigation. From this activity would come knowledge of bearing strength in pounds per square inch, slump resistance, chemical relevance to construction problems, and physical properties affecting digging and cementing. The heat-transfer characteristics of soil will be very important to the requirements of thermal shielding.

Engineering knowledge related to the mechanical properties of rocks will determine the feasibility of their use in building shelters. Drilling procedures and safety of lunar surface and subsurface living will also be influenced by these properties. As a very minimum, the conditions of long-term protection for monitoring instruments will be of importance. Freedom from complete breakdown due to radiation or particle influx must be ensured early on. Exposure panel and test panel monitoring will be crucial to finding which lubricants, component parts, or protective covers work well in various lunar environments.

Many of the data from these tests will be obtained from monitors recording and transmitting performance data There will also be voice reports on tapes and direct telemetry. Some instruments will be monitored on a "change only" basis. Some data will be in the form of imagery from photography and TV.

Instrumentation required will include thermometers, speedometers, voltmeters, and operational recording instruments. Astronaut-operated equipment will be of the soil- and rock-testing category. Results will be recorded on tapes directly or by voice records and descriptions.

TV cameras and visual still and movie cameras would be used with direct Earth return of film or by telemetry link.

A.3 TECHNOLOGY DEVELOPMENT REQUIREMENTS

A.3.1 Introduction

The determination of which natural and cultural resource phenomena are to be acquired from space is intimately related to the development of space flight instruments, subsystems observational procedures, and interpretational techniques. It is necessary to correlate lunar and Earth resource information requirements in the data storage and retrieval system and also to correlate the equipment designers' ground-truth and interpretational requirements with the resulting data bank.

An extensive array of remote sensing equipment is currently being flight-tested and ground-truth studies are being made to gain a better understanding of the capabilities of this equipment so that the data obtained may be fully utilized. The Earth Resources Aircraft contains equipment covering the electromagnetic spectrum. Test flights over selected sites, with carefully annotated flight logs and ground-truth data, are obtaining data in various forms which, with effective storage and retrieval, will help in perfecting both the equipment flown and interpretational techniques.

The comprehensive annotation of these flight data and acquisition of carefully controlled ground truth data will not only enable rapid and comprehensive retrieval, but will also enable complete and rapid assessment of this information. In this way modifications in design and technique are efficiently implemented.

A.3.2 Remote Sensing Instrumentation

Perhaps the most organized way of presenting the interrelationship of remote sensing systems for Earth resource surveys is by reference to Fig. A-1. This figure presents a breakdown of the sensing bands related to atmospheric transmission, energy source, interactions, and detectors.

Gamma and X-Ray Region

Scintillometers and high-energy particle detectors are useful at the short wavelength end of the electromagnetic system. These sensors were utilized extensively in uranium exploration. Since the cessation of these activities, the gamma-ray spectrometers have been further developed in the search for potassium, uranium, and thorium deposits and for low-level radiation detection from normal rocks and soils. Experiments indicate that airborne instruments may be used for gamma-ray pulse-height analysis. X-ray emission (fluorescence) is a highly developed laboratory technique but has not been used extensively for airborne studies due to the attenuation properties of our atmosphere and the unfeasibility, at present, of making active x-ray systems airborne because of their power requirements.

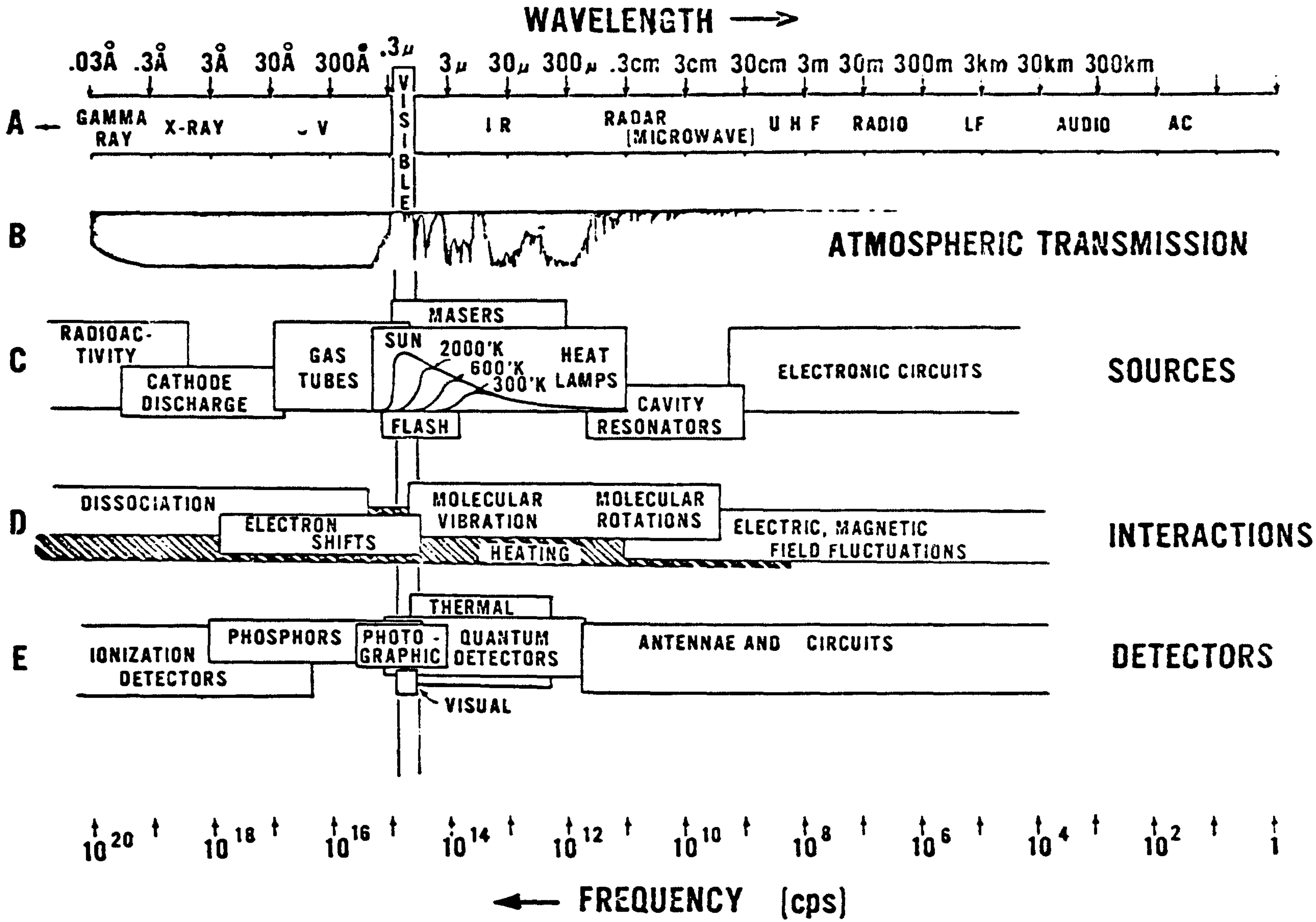


Fig. A-1 Characteristics of the Electromagnetic Spectrum Which are of Significance in Remote Reconnaissance (Redrafted From Article by R. N. Colwell, Photogrammetric Eng., Vol. XXIX, 1963)

Ultraviolet Region

The next useful imaging band in the spectrum is the ultraviolet. As in the visible and infrared bands, ultraviolet sensing is done at the skin of the radiating surface. The ultraviolet radiation (UV), defined as the area less than 4000 Å, is heavily absorbed by Earth's atmosphere and therefore has its greatest potential in the remote sensing of planets having little or no atmosphere.

Passive experiments utilizing UV radiation are restricted to the near UV (4000 Å and longer) and are concerned with luminescence phenomena where solar UV excites certain terrain materials (phosphates and calcites); the resultant fluorescent radiation occurs in the visible spectrum. Consequently, most UV photographs are taken with appropriate filtering between 4200 and 3900 Å. As indicated in Fig. A-1, at the shorter wavelengths the contribution from atmospheric scattering becomes sufficiently large to obscure the surface radiation. In examining the UV spectra of the possible lunar surface rocks, it is found that the greatest differences in reflectivity occur at wavelengths shorter than 2200 Å. Conceivably, high contrast imagery may be obtained in this area. UV spectrometric methods are also employed and center around UV stimulated fluorescence and luminescence.

In fluorescence, the emitted radiation is simultaneous with the exciting radiation; in luminescence, emitted radiation takes place after the excitation has been removed. Since the stimulated radiation takes place usually in the visible, the possibility of detecting fluorescence from air or spaceborne platforms does exist, but it is not yet a proven technique.

Visible Region

Next in the spectrum is the visible portion, which is by far the most thoroughly developed and understood of all the remote sensing systems. The photograph still remains the most compact and precise information storage medium so that in spite of the development of infrared and radar systems, camera systems continue to play a major role in remote sensing. Selection of the proper type of camera format size and focal length requires a detailed tradeoff analysis which is dependent on the particular application. There are many lenses, films, filters, and speed combinations which make it possible to sample several spectral bands at once. The conventional wavelength range is considered to be from 4000 to 7000 Å whereas the photographic or near infrared is from 7000 Å to 1.0 μ (10,000 Å). The photographic system operates in a passive mode detecting the reflectance of "visible light" by means of photographic emulsions [panchromatic, ektachrome (true color), or ektachrome infrared (false color)]. Conventional black and white photography requires little, if any, explanation because of its acceptance and widespread use in the Earth Sciences.

More recently, a system has been developed which obtains photographic coverage from 0.4 to 0.9 μ (Refs. 22, 23, and 24). Coverage is subdivided by the use of nine matched lenses with different filters to take nine simultaneous exposures of the same ground area. Each of the resultant nine photographs taken during one exposure is of identical image size. Six of the exposures are on two rolls of 70-mm panchromatic film and three exposures are on a roll of 70-mm infrared film. Eight of the exposures represent spectral signatures in the eight different narrowband portions of the spectrum, the ninth has full IR sensitivity. A color-separation technique is used to enhance the tonal differences imaged in the multiband photography. Selected frames from the eight bands provide the tonal values necessary to employ a color separation system. This technique enables operation at a bandwidth which will give maximum tonal contrast. What finally evolves after photographic processing is a composite photograph representing differences of reflectances between bands. Where no color differences are evident, cancellation occurs. This approach is considered to have some advantages over conventional color photography largely because the sensitivity to temperature, humidity, exposure, and production variables of conventional color photography makes it a difficult medium to work with. The high resolution of wide-bandwidth reconnaissance film is also maintained.

The use of color photography for remote sensing in the visible spectrum supplies an additional dimension not evident in black and white photography (Refs. 25 and 26).

It is evident that the gradation of colors supplies subtleties contained in the scene which would not be evident otherwise. Recent Gemini photography contains excellent examples of this process; color variation with depth is very evident showing the contours of the coastal areas to about 60 feet in depth (Refs. 27 and 28).

When using "false color" techniques, the colors in the scene are approximated but colorimetric match does not imply spectral match. Any portion of the spectrum to which photographic materials are sensitive can be recorded and the colorant developed by this record can be any that are available without regard to the spectral region initially chosen. These are known as "false color." Spectral regions can be chosen to emphasize objects which visually are quite similar. Ektachrome Infrared is based on this principle and emphasizes the difference in infrared reflection between green vegetation and most visually similar green objects. When a yellow filter is used to absorb the blue light, the chlorophyll of the vegetation will appear red while other green objects will appear a dull blue. In a similar manner, any two (or more) different portions of the spectrum discussed can be selected to emphasize spectral differences occurring in the objects of interest. Several areas within the Earth sciences disciplines are making use of this extra dimension to record and extract additional data. Notable examples are the airborne detection of crop diseases and loss of vigor in trees prior to ground evidence.

Infrared Region

Beyond the visible exists the infrared region within whose boundaries two basic types of remote sensing systems are available. The first is a system using infrared sensitive film and filters for sensing the very shortest of infrared wavelengths (from 0.72 to 1.0 μ). This is the system which has been summarized in the previous discussion on multiband imagery. Equipment to operate in the near infrared (1 to 4 μ), medium infrared (4 to 15 μ), and far infrared (15 to 1000 μ) has been developed to extend the limits of photographic films. Such equipment operates within the windows available in the Earth's atmosphere and records the radiance received from the source after it has been emitted by (or reflected from) the Earth's surface. The atmosphere's composition allows relatively undisturbed transmission of infrared radiations in the 2- to 5- and 7- to 15-μ ranges. It is important to remember that the similarity between the infrared imagery and conventional photography ends with the fact that they are both pictorial representations of the scene. The distribution of contrast on the thermal infrared imagery represents the pattern of energy radiated from the scene due to its temperature and emissivity distribution (Refs. 29 and 30).

Since the thermal infrared spectrum falls between those of visible light and radar waves, its radiation exhibits some of the characteristics of both visible light and microwave radio waves. It can be optically focused and directed by lenses and mirrors or dispersed by prisms. It can also propagate through some materials that are opaque to visible light.

The basic elements of the airborne thermal infrared mapping system consist of a plane mirror scanner whose axis of rotation is parallel to the longitudinal axis of the aircraft. This mirror reflects radiation to a parabolic mirror that is used to focus the radiation on the infrared detector. The detector senses the incoming radiation and converts radiation changes into an electrical signal. This signal is amplified and used to modulate the resulting current through a lamp whose output is an intensity-modulated spot of light. The image is focused on the surface of a photographic film and is scanned across the film in synchronism with the rotating scanner. The film moves across the exposure station at a speed proportional to the aircraft speed-to-altitude ratio and the result after processing is a strip thermal map of the area flown over, in which the film density represents relative effective infrared radiation temperature.

Infrared systems produce good daylight imagery, but, since they respond to energy radiated from beyond the visible spectrum, night infrared with sensitivity in the middle and far regions yields excellent results. For many purposes, infrared data are best collected after dark since there is no interference from solar insulation. Dust and haze may be penetrated depending on the size of the aerosol particles although clouds, high surface winds, and rain reduce image quality. It is not difficult to envision the applicability of such a system to the Earth sciences. Mapping of hot springs, glaciers, and volcano areas have been accomplished successfully. Alternating bands of warmer and cooler imagery are easily correlated to sandstone and siltstone beds. The sharp differentiation seen between moist and dry areas is also outstanding (Refs. 31 and 32). This system has been used for plotting forest fire perimeters normally obscured by smoke to assist fire fighters.

Radar Region

In the wavelengths longer than infrared, there is a variety of radar systems used for both passive and active sensing. The physics of the passive microwave systems is strictly in the domain of infrared physics (wavelength range 0.1 to 10 cm). Some advantages that they have are greater penetration into the terrain than infrared, lower attenuation in the passage of the rays through the atmosphere, and, provided that the emittance of the surface layers is known, they may determine subsurface temperatures. Each radiance measurement combines three semi-independent variables – reflectance, absorption, and temperature – of the surface and also of each of the layers below the surface down to the effective attenuation depth of the wavelength in this medium. Differences in the emittance of objects at the same (ambient) temperatures can be mapped. Poor reflectors (high emittance) look "hot" while good reflectors look "cold" even if both are at the same temperature. For this reason, icebergs (with emittance almost unity) look hotter than the surrounding seawater (with a much lower emittance and thereby reflecting the truly colder sky). Melt water can change the emittance of ice markedly if it collects into pools. Temperature detection between bodies measured to tenths of a degree is only possible if it is presumed that the emittance of the two bodies are the same (Ref. 33).

Up to the present time, emphasis has focused on active radar imaging systems such as side-looking radar (SLAR). The wavelength range involved is from 1 to 300 cm with only certain bands being used in practice (K, X, S, L, and P). Radar produces imagery which is largely independent of the climate and time of day because the longer wavelengths penetrate fog, haze, and clouds with minimum signal loss. (Rains attenuate the signals but the extent of loss depends on the system wavelength and rainfall rate.) System components such as cathode ray tube and film cannot record all signal levels received at the antenna with equal discrimination. Whether the film records maximum differences between high or low intensity signals depends on the requirements of the particular mission. The radar image is affected by the manner in which the object is illuminated, its surface roughness, and the geometry.

The basic advantages of radar imagery are day-night all-weather overview with the ability to penetrate forest cover. Specific advantages exist in its use for geological studies. Overview and penetration make it possible to depict distinctly geological features such as linear scarps. Structural features such as faults, attitude of strata, and drainage patterns are very obvious. Fine and coarse surface materials can be discriminated. Radar has been used in discriminating old and new ice and in the study of glaciers in which moraines, crevasses, and texture patterns are made clear (Refs. 34, 35, and 36).

A.3.3 Ground-Truth Correlation

The remote-sensing instrumentation discussed in the previous section has indicated parameters relating to several systems, particularly emphasizing those operating in the far infrared and radar bands. To fully utilize the instrumentation records or imagery, they must be correlated with calibrated test site information so that interpretational techniques may be perfected. This requires a meticulously executed program of gathering ground-truth data during the course of flights (Refs. 37, 38, and 39). Appendix C is a technical letter report of a "Conference on Ground Measurements for the Instrument and Geologic Teams" held at the Mackay School of Mines, University of Nevada. The report outlines the basic ground-truth requirements and parameters and is an excellent coverage of the data required to enable valid correlations for the development of interpretational techniques.

A.3.4 Data Annotation

To complete the system requirements for the input data to the data bank, the proper annotation of collected data must be considered. Annotation data refers to that data describing the equipment and conditions under which particular data is collected. Proper annotation methods enable all relevant equipment data, conditions of use, and related ground truth data to be correlated. The extraction and meaningful interpretation of collected data is severely hampered without such correlation.

Appendix B
LUNAR/EARTH DATA INDEXING CONCEPTS

B.1 INTRODUCTION

Lunar/Earth data indexing concepts are described with specific emphasis on image indexing. The indexing of report data and other textual material has evolved over many years with several systems emerging as the accepted standards (Ref. 40). Present developments in the indexing of textual data have evolved systems such as permuted subject indices, inverted indices as represented by the Lockheed DIALOG system, and other techniques adaptable to ADP operations. Much work remains to be done in the area of image indexing. The indexing concepts developed here for image indexing follow closely other ADP systems (Ref. 41). The file organization is patterned after that used by the NASA Scientific & Technical Information Facility.

B.2 GENERAL CONSIDERATIONS IN IMAGE INDEXING

B.2.1 Differences Between Textual and Photographic Data

With a textual object, an indexer can extract verbal data from the document itself and produce a usable index image of the document without any real understanding of what the document actually contains. Unless we can conceive of an index composed of actual picture segments, the indexer must be able to obtain some true comprehension of the picture contents.

A textual document is usually created in accord with some specific objective of the author. While a potential user may be more interested in a tangential facet of the document, the use cannot be too far removed from the intended theme of the document. Many graphical objects have no central or organizing theme (e.g., a terrain photograph), and a description of the same object from two different points of view may be completely disjointed.

The implication of these comments is that indexing pictorial data must be more application oriented than the indexing of textual data, and for equivalent performance the indexer of pictorial data must have a higher degree of familiarity with the subject matter (or be better supported with indexing aids).

B.2.2 Types of Photographic Representations

A photographic representation of a scene can assume many forms, thus posing a variety of problems for an indexer of such data. The simplest representation is a conventional colored or black and white print, in focus and noise free, obtained by an optical system sensitive only to the visual spectrum. A more complex representation is again a single image; but this time obtained either outside of the visual spectrum (e.g., by infrared photography), or by the detection of coherent radiation (side-looking radar or laser photography), or even a multispectral representation combining both visible and nonvisible radiation. Additional complexity is encountered in photographic representations which include stereo pairs (in which the scene is represented by two images which must be superimposed, usually by special equipment, for viewing) and mosaic or strip photography (where the scene was partitioned during the photographic process, and may or may not be reassembled of indexing time).

The main point to be noted here is that the problem of "readability" (and perhaps even that of physical access), which is almost never even considered in discussions of textual indexing, is a serious consideration when dealing with photographic representations.

B.2.3 The Descriptive Process

A reasonably comprehensive retrieval system for pictorial (as well as textual) data will offer an abstract as well as a collection of descriptors for each item in the file. Even assuming the indexer can interpret what is present in a photograph, the preparation of a verbal description is a complex art form offering a large variety of choices.

First we note that the description of a pictorial object can be oriented toward forming a basis for reconstructing (at least in wide form) the object. In this case, the description would employ a relatively simple vocabulary, and would be largely concerned with the location, size, and shape (color, texture, etc.) of simple structures which can be identified in the photograph. Such a description would have to be extensive.

A second type of description would be oriented toward distinguishing the pictorial scene from other similar scenes, or placing it in one of a set of distinct categories. In this case, the indexer would have to have considerable knowledge of what is typical and what is not typical for the subject matter of the photograph. He would also have to know what the possible classification categories might be. The resulting description could be relatively short, but would necessarily resort to a much more complex technical vocabulary.

A third type of description is one oriented toward a specific process, such as information retrieval. In this case, the description could be closer to an enumeration of relevant entities (major structures appearing in the photograph might be completely ignored). The indexer could be extremely specialized and therefore quickly trained. If desirable, a collection of descriptive forms and names or indexing terms could be compiled to simplify the writing of such a description.

B.2.4 The Descriptive Citation

A formal descriptive record which identifies and describes a unit of source data such as an image is called a citation. Because of its textual format, it lends itself to computer processing, and indexes can be generated from the contents of various fields in the citation. The citation is a convenient means for providing:

- A formal record which consolidates information from a variety of sources pertaining to a particular image
- An indexing base for automatic data processing equipment
- A medium for announcement of image availability as well as an intermediate search product in information retrieval

The different classes of information to be included in the citation emerge from these objectives. The first class reflects the need for a definitive compilation of the physical characteristics describing the image acquisition and the necessary geometric relationships required for image rectification. The second class relates to the geographic location of the image. The third class of indexing information describes the image content, and those parameters of general or specific interest to principal investigators and other users.

The first two classes of indexing represent the constants supporting the interpretation of the image content. The data fields for identification data, annotation data, related information, and geographical location are included. The third class is represented by the general textual description of the image content. Other ancillary data may be included on the citation form such as procurement information. A continuing problem is deciding an efficient tradeoff between the amount of detail to be included and the cost of its inclusion.

B.3 AN IMAGE INDEXING EXAMPLE

Image indexing involves a process and a product he citation. In this example the citation was produced from a photograph of the S... Francisco Bay area. Following a description of this format, the process which would be used in an operational environment at MSC will be described. This format is used as an example to illustrate a concept rather than the ultimate format which would be required. The citation must be used to describe not only imagery but also the other forms of source data.

B.3.1 The Phot graph

Figure B-1 represents an unusual photograph. It consists of a mosaic formed from 591 photographs taken from an altitude of 30,000 feet. The camera lens focal length was 6 inches, giving a scale of 1:60,000. The mosaic photograph was printed to simulate a picture taken with a 12-inch lens from a spacecraft at an altitude of 300 miles. The

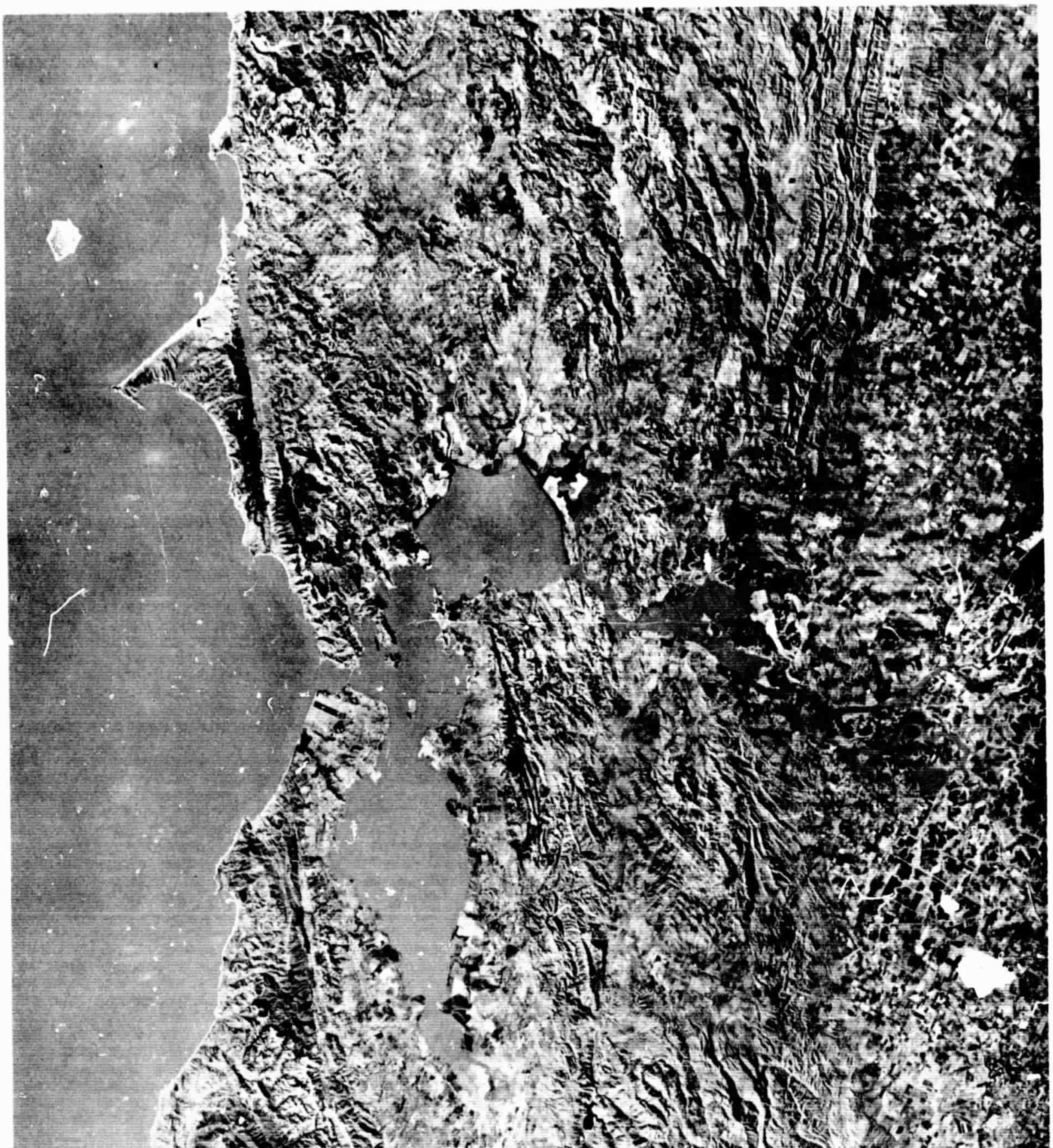


Fig. B-1 Source Data Sample

various photographs were taken during the first part of 1956. Due to the absence of other specific information, geographic coordinates and scale are close approximations. Annotation data are fictitious, with entries chosen to simulate the type of information required. This photograph was selected as an example due to its clarity, knowledge of its geopolitical area coverage, and the fact that various disciplines can derive information from its content.

While the example is interesting, its primary purpose is to illustrate the preparation of a citation. The format for the citation has been developed to standardize the entry of those parameters which will identify the time period, instrumentation details, location, general image content, and applicable references for each image acquired. The format contains a well balanced choice of basic data which, upon demand from the user, will service multiple disciplines.

B.3.2 The Citation

The citation shown in Fig. B-2 contains fields for identification data, annotation data, geographical location data, image content data, and procurement information. The accession number, the first entry in the identification data field, indicates that the image shown in Fig. B-1 was cataloged in 1968 (the first two digits), that the citation describes an image (I), and that this is the 63rd image cataloged in 1968 in the 10,000 series. The remainder of the identification data field is self explanatory.

The annotation data section has subsections to indicate the collection instrument characteristics, and the vehicle positional and attitude information. Employment of the Principal Point (PP) pitch, yaw, and roll information in conjunction with the vehicle track will help define displacement of the PP and the nadir point.

The Geographical Location section has been designed to permit a rapid assessment of the general area contained in the imagery. In the proposed system, the use of World Aeronautical Charts (WAC's) is employed. The use of grid numbers allows a further refinement of a localized coverage within the general area covered by the

IDENTIFICATION DATA

Accession Number	Issue Date	Project	Date (UT)	Time (UT)	Mission Number	Pass (Orbit)	Frame	Recording Mode	No. of Times Requested
68 1 10063	1/09/68	A x z	7/07/67	2223	2188	513	1015	Photographic	413

GENERAL DESCRIPTION

Aerial photographs combined to simulate photograpn of San Francisco Bay Area from a spacecraft at 300-mile altitude.

ANNOTATION DATA

Camera Type	Camera Number	Lens (efl) (in.)	Lens (efl) (in.)	Film Type	Spectral Range (mμ)	Filter Used (mμ)	Processing Method
KC-1	R-507	12	12	E/HSIR	540 900	A-25 (700 900)	D 19 (8 min)

Instrument Type	Instrument Model	Spectral Range (μÅ)	Tape Type Used	Storage Capacity (Char.)
N/A	N/A	N/A	N/A	N/A

Vehicle Track	Sun Alt.	Sun Azimuth	P.P. Pitch	P.P. Yaw	P.P. Roll	Vehicle Altitude	Scale
330°	045°	210°	30 min	1 deg	45 sec	30,000 ft	1/60,000

Remarks:

GEOGRAPHICAL LOCATION

WAC No.	Grid No.	Geopolitical Area	Area Cover (mi²)		N.W. Corner	S.W. Corner	N.E. Corner	S.E. Corner	Principal Point
364	R4, R5	Western USA (S. F. Bay Area)	12,500	Lat.	38° 25' N	36° 45' N	39° 05' N	37° 25' N	37° 53' 45" N
				Long.	123° 40' W	122° 45' W	122° 05' W	121° 00' W	122° 18' 18" W

Remarks: Positional accuracy given as: Corners to nearest 5 arc minutes
Principal Point to nearest 1 arc second

IMAGE CONTENT

Cloud Cover	Snow Cover	Water Cover	Land Cover	Image Quality
None	< 5%	30%	70%	Excellent

Remarks:

Other Features

1. San Francisco Bay
2. San Pablo Bay
3. Santa Clara Valley
4. San Joaquin Valley
5. Sacramento River Basin
6. San Andreas Fault
7. Pacific Ocean and coastline
8. Point Reyes
9. Coastal mountain range

PROCUREMENT INFORMATION

Order From	Data Forms Available	Source Data Location	Source Data Custody
MSC Houston, Tex.	See Remarks	MSC Houston, Tex.	Lunar/Earth Data Bank

Remarks:

Microfiche	$0.15 ea	
35-mm neg	0.15 ea in aperture card	
70-mm neg	0.20 ea in aperture card	
105-mm neg	0.50 ea in envelope	
9.5 × 9.5 in. neg	2.60 ea in envelope	

REFERENCES

Accession Numbers	Issue Date
1. 68 1 10064	7/15/67
2. 68 1 10065	7/15/67
3.	
4.	
5.	
6.	
7.	
8.	
9.	
10.	

Fig. B-2 Image Citation Example

WAC number if useful. Geopolitical area nomenclature and the total area contained within the imagery can also be employed for a rapid perusal of potential information to the user. The geographical coordinates of the four corners of the area covered provide a third means of rapid assessment.

Image content quite generally gives the percentage of cloud, snow, water, and land coverage. Image quality is expressed as excellent, good, fair, and poor. Specific resolution values will result from a detailed analysis by the user. Specific image content will evolve from the same analysis. Under "other features," prominent, well known, or easily recognized surface features are identified which help to define specific areas of image content.

Procurement information is predicated on the assumption that all data will be stored at MSC, Houston, Texas. Cost estimates are approximate and intended to depict that type of information which is available to the user and its approximate cost. References will be added as deemed advisable by the Lunar/Earth Data Bank or at the request of a user.

B.4 THE IMAGE INDEXING PROCESS

Figure B-3 shows the information flow in an image indexing process at MSC for Lunar/Earth data. The accession number represents the only arbitrary entry on the citation (except for year and data form code); all other data are derived from mission planning, program office, principal investigator, and instrumentation/flight log documents. At the moment of image capture, these data become the constants for image interpretation. The preparation of the citation may even precede the availability of the actual image from the photo processing. The assignment of the accession number to a set of "constant data" requires the matching of the image to its unique citation. At this time the content analysis could be performed. The citation and film processing are prepared in parallel. During Step 1 the mission data, program office, and principal investigator directives are abstracted. The flight log and recorded data (magnetic tape) are processed (abstracted) and are posted to the

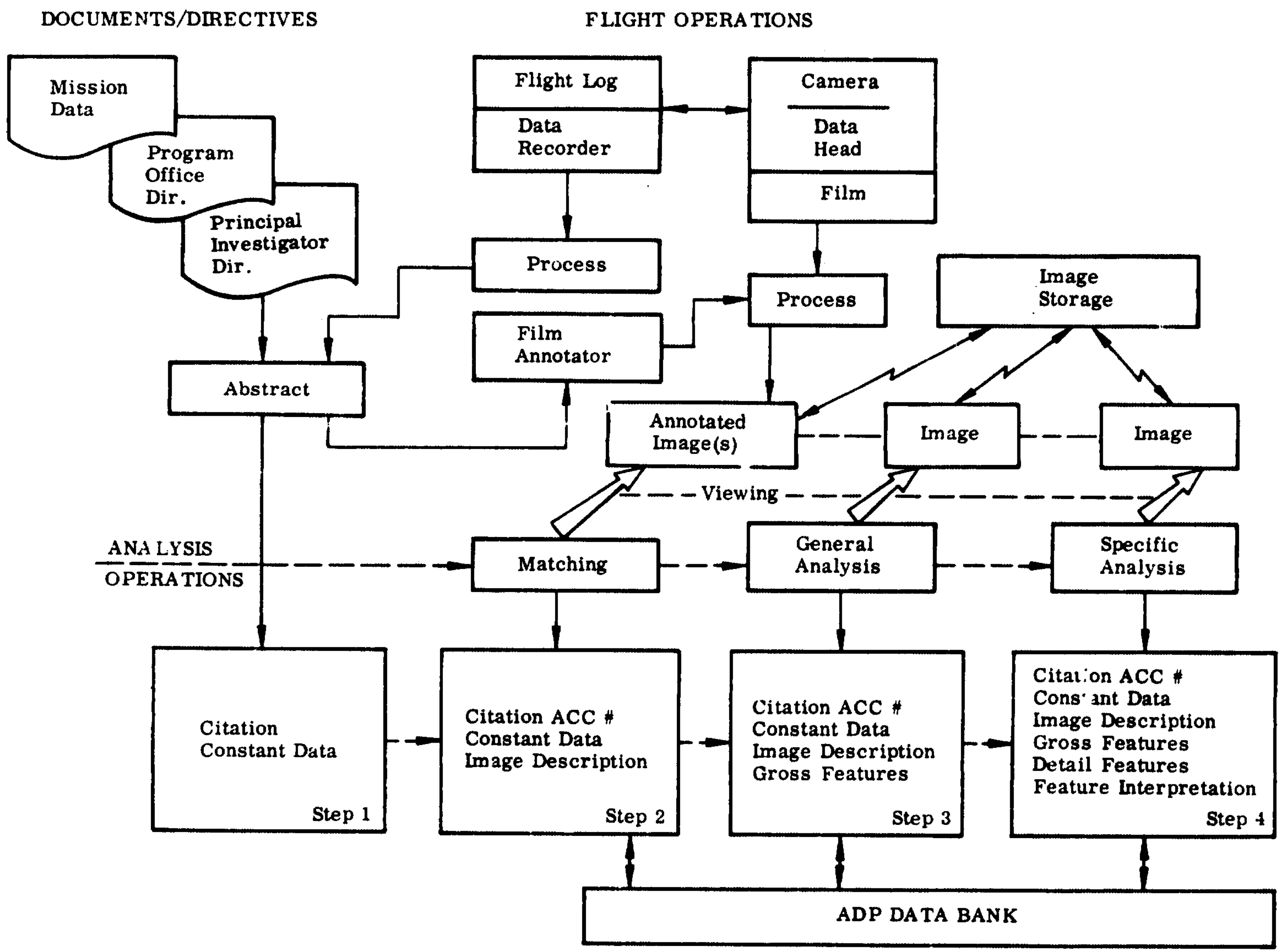


Fig. B-3 Citation Processing, Example

citation. In parallel with the citation preparation, relevant data are annotated on the image for correlation and matching in Step 2. During Step 2 the accession number is entered on both the image and the citation. At this point in the sequence, a one-to-one relationship is established between the image and the citation, and an image search is possible based on these limited data acquisition objectives and instrumentation constants. This capability is indicated on the diagram by flow lines to the ADP for the citation and image storage (by accession number) for the image data. These steps conclude the entry of the first class of descriptive data.

The second data class relates the image to its geographical location. These data may be derived from a time-spatial relationship for each exposed image. These data are obtained from a correlation of the flight log, and recorder outputs to ground check points identified both during the data acquisition and later during film processing. The geographical location of the image is identified by the principal point and the four corners. When ground check points appear on the image they must be identified for ease of location. The geographical location, principal point plus corner coordinates, is completed during Step 2.

The image content analysis function produces the third class of indexing data. This class includes identification of gross image features (Step 3) and the identification of detail features and feature interpretations (Step 4).

The general analysis function requires a minimum of skill relating to the image interpretation. This analysis makes an assessment of the image quality, identifying cloud cover, and other visibility limiting factors (reflections, etc.). Further assessment is made on the basic coverage identifying land/water ratio and the discernible gross features. Some of these gross features are as follows:

- Land/water, lakes/oceans
- Geographic area identification
- Land features, mountains, plains, rivers, deserts
- Water features, islands, underwater feature visibility

The identification of these gross features in conjunction with the constant data affords a significant advancement in source data retrieval with a high probability for obtaining the desired specific features.

The assessment of the detailed features would be accomplished primarily by personnel trained in the specific field of investigation, i.e., forestry, hydrology, oceanography, etc. Such feature interpretation is the responsibility of the principal or associate investigators and their staff. The detail to which the feature analysis is performed is based on several factors. First, if the principal investigators are directly associated with the program office and data facilities at MSC, the feature extraction could be performed by both MSC and principal investigator team members providing gross and detail feature identification; if the principal investigator is located remote to the data processing activity, the level of analysis would depend on the available analysis services at MSC. Second, the analysis team personnel capability and number of members (excluding principal investigators) would depend on the available project funds, the total volume of data to be screened, the objectives of the analyses, and other factors relating to data interpretation. The level of analysis must be continually weighed against the data contribution. The techniques of data acquisition may produce a high level of redundant data or multiple frames of imagery containing similar gross features, i.e., as expected for high resolution images in a continuous strip. For this situation, a single analysis would include n frames in a single citation. Under this concept the citation describes an indexable unit which may include any number of images or any length of film.

The "specific analysis" refers to the identification of detail features on the photographs. Another aspect of the detail analysis includes feature interpretation. It should be acknowledged that a wide overlap exists between feature identification and interpretation. This overlap results from the different experience levels of the analysts.

The task of photo interpretation is treated here in relation to the Earth sciences. The lunar sciences also require extensive photointerpretation; however, the techniques utilized for the Earth sciences are directly applicable to the lunar tasks, particularly in the areas of cartography, geography, geodesy, and geology. The Earth sciences tasks also include natural resources in agriculture, forestry, hydrology, and oceanography.

The objectives for each data acquisition mission are defined by the principal investigator and his associates. The instrumentation is designed to meet these objectives and similarly the analyses level performed is expected to favor those areas supporting these objectives. The nature of data is universal to many different objectives. The identification of the basic coverage is useful to investigators in other fields. There exists a need to develop an indexing and classification system for the Earth science data enabling an improved, efficient utilization of the total data bank. The capability of the general-purpose computer can provide an effective cross correlation between the fields in Earth sciences technology. A thesaurus should be evolved relating specific terms and related items. A hierarchy of classifications would provide the basic headings for indexing and cross referencing. This thesaurus would be developed to aid the general analyst in photointerpretation, enabling the identification of features in acceptable terminology.

The hierarchy of classification should be relatively independent of the specific science. The first order of classification might identify the continental platforms, ocean basins, and gross details at the interface. The second order of classifications could describe the land areas, detailing features such as mountains, plateaus, plains, and deserts; for the water areas, feature identification would include lakes, oceans, island populations, and underwater features when visible. The third order of classification would identify more detailed features such as hills, valleys, buttes, mesas, rivers, vegetation, and other features. More detailed levels in the hierarchy may be desirable; however, as detail increases, indexing and retrieval generality decreases. For example, the distinctions between forestry and agriculture or industrial development and natural phenomena takes into account the man-made patterns which are detectable. The thesaurus of terms is a function of the detectable features which are different for aircraft and satellite data. The photographic scale and image resolution provide different sets of resolvable phenomena requiring new descriptors for features and attributes.

The development of this thesaurus will enable many researchers to share the information in the data bank. The use to which these data may be put depends on the need and resourcefulness of the investigators. Table B-1 (derived from Ref. 42) identifies some features and attributes which would be considered for inclusion in this thesaurus. The thesaurus would be developed primarily from most frequently used terms.

The basic objective for data classification is to provide a data bank supporting research and study of Earth features. The capability for interpretation of these features from instruments must be determined. As these technologies are developed, the ability to retrieve the original data is necessary to substantiate new concepts for feature extraction or data interpretation.

B.5 MACHINE AIDS TO IMAGE INDEXING

B.5.1 Automatic Classification of Pictorial Data (Pattern Recognition)

Almost all present pattern recognition schemes partition the recognition problem (after data acquisition) in two parts. The first part is involved with the problem of feature extraction, operations on a pattern which determine identifying features or characteristics of the pattern. The second part involves the decision-making procedure which classifies the pattern by comparing pattern characteristics with those of a reference set of patterns. The decision-making procedure might be arrived at through statistical-decision-theory considerations, or perhaps through some informal intuitive approach; but in any case, the performance potential of the decision-making element is highly dependent on the output of the feature extraction.

While the decision-making element can be designed so as to be relatively general in its ability to accept inputs from a wide variety of feature extractors, the feature

Table B-1

DISCIPLINE-INDEPENDENT, FEATURE-DESCRIPTIVE INDEX TERMS

<u>Entities (Features)</u>

Acoustic noise
Air
Airports
Albedo
Ash
Barometric pressure
Beaches
Biology
Biological pollutants
Bioluminescence
Breezing
Chemical pollutants
Cloud cover
Crustal features
Cultural features
Dwellings
Earth tilt
Ecological features
Faults
Floods
Folds
Harbors
Ice
Industry
Joints
Lakes
Lake shore
Land
Lava
Lineaments
Meteorological features
Mountains
Ocean
Ocean currents
Pingos
Population
Radiation
Railways
Rainfall
Raw material
Rivers
River banks
Roads
Roof density
Schools of fish
Sediment
Shock swarms
Silting
Smoke and fire
Snow
Soil
Surface energy
Surface strain
Swarms
Tectronic features
Thermal energy
Traffic type
Transportation
Turbulence
Upwelling
Vegetation
Water
Water vapor
Weather
Wind

<u>Attributes</u>

Area
Boundaries
Chemistry
Circulation
Color
Coverage
Cycle
Damage
Decay
Deformation
Density
Destruction
Development
Dielectric constant
Diffusion
Distribution
Drainage
Drift
Dynamics
Elevation
Energy
Erosion
Extent
Flow direction
Flow rate
Frequency
Heat
Height
Infestation
Location
Migration
Moisture
Movement
Pattern
Periodicity
Planting pattern
Power
Proximity
Reflectivity
Salinity
Sea state
Shade
Silting
Slope
Spectral signature
Temperature
Texture
Thickness
Turbidity
Variation
Velocity
Vigor
Volume
Water loss
Wave characteristic

extractor itself must be very closely matched to the process on which it is to act. The feature extractor can be looked at as a collection of complete pattern recognition devices, each such device designed to detect one specific feature. Thus, the design of the feature extractor poses two problems to the designer; he must first decide which features are to be extracted for input to the decision-making element, then he must use his knowledge of the process producing the patterns to design the sensors and logic which will detect the features of interest.

As can be seen from these comments, the designer must know rather precisely what he is looking for if he is expected to do a good job on the design of the essential feature extractor. It should also be realized that most of the theoretical work done to date in the area of pattern recognition has been concerned with the decision-making element. There is almost no theoretical basis for the design of the feature extractor. Operating pattern recognition systems which perform successfully owe their success in large part either to the simplicity of the problem to which they are applied, or to the intuitive cleverness of the designer in constructing the feature extractor. Thus, the near-term potential for a useful general-purpose system for fully automatic pattern classification of pictorial data is rather small.

B. 5. 2 Semiautomatic Classification of Pictorial Data (Screening of Photographs, Detection of Changes in Photographic Images)

Machine screening (crude classification) of photographic data according to criteria such as the occurrence of gross change (for which comparative cover exists), or the percentage of cloud cover, or the "texture" and structural properties of the photographic content, offers reasonable promise for near-term accomplishment. None of the machine screening techniques is yet available for practical application and indeed many problems still remain to be solved.

Machine image detection procedures, which have received the most attention so far, are still overly sensitive to such irrelevant factors as camera position and orientation, sun location and resulting shadows, atmospheric conditions, etc. Percentage cloud cover assessment and screening can probably be accomplished with today's techniques, but the utility and cost-effectiveness of this type of screening by machine is open to serious question. The machine partitioning and screening of photographs based on their texture and structural properties is a relatively new area of investigation (some work in this area is being carried out at Cornell Aeronautical Laboratories), and its potential is still unresolved.

It is not unreasonable to consider human screening of pictorial data in this section on semiautomatic classification of data. Here we would have relatively unskilled personnel (perhaps using machine aids or data already filtered by machine) performing rough classification of pictorial data, either according to the criteria given or with respect to somewhat more sophisticated criteria requiring the ability to distinguish between land, sea, mountains, desert, etc.

B. 5. 3 Machine Aids for Human Classifiers and Indexers of Pictorial Data

Machine aids to the human processing of pictorial data fall into many categories. These can include:

- Display equipment – hardware which permits viewing the pictorial data. Such hardware includes optical projectors (including stereo receivers) with a capability for magnifying, focusing, and accessing imagery locally available in hard copy form; facsimile and long distance xerography for obtaining hard copy imagery from a remote source (some ability to adjust magnification and contrast); electronic displays (e.g., TV monitors, cathode ray tubes, electroluminescent and mosaic indicators) with a capability for magnifying, changing aspect ratio, and adjusting contrast of locally or remotely available imagery.
- Rectification equipment – hardware which removes distortions introduced into the imagery by the acquisition process, including nonlinear geometric distortions, bandpass distortions, intensity distortions, and various types of random noise. Processing techniques for rectification of these distortions is still highly dependent on the skill of operator, and on the particular imagery and acquisition system involved. The hardware employed is typically

a system involving a digital computer and a flying spot scanner. Thus, rectification equipment is obviously not suited to on-line use by personnel indexing or abstracting pictorial data.

- Mensuration aids – measurements derived from photographic data can have two applications. The measurements can result in annotation on the photographic data itself, or they can be used as part of a report (index, or abstract) about the photographic data. In the first case, an example of annotations which can later be used by an indexer are contour markings. Such annotations can be produced automatically (some operator intervention required) by existing hardware which uses stereo pairs as input. In the second case, we are typically integrated in area and distance measurements. A wide range of such instruments currently exists, from inexpensive hand-held rulers and planimeters, to computer-implemented routines for making and adjusting the measurements according to scale, projection angle, and other parameters associated with the photographs.
- Clerical and information retrieval aids – devices, forms, and formal procedures which can be used as tools by a human indexer of pictorial data in the performance of his task. These can include specially designed forms, vocabulary and phrase lists, transcribing aids (including a terminal to a computer-controlled report writing system), and a retrieval system for obtaining photographic keys and comparative cover to be used as a pictorial guide for the human indexer. A wide variety of such aids currently exists.

Appendix C

DIALOG: An operational on-line reference retrieval system

by ROGER K. SUMMIT
Lockheed Palo Alto Research Laboratory
Palo Alto, California

INTRODUCTION

Classification systems in the sciences usually provide an unambiguous structure of mutually exclusive, collectively exhaustive categories. The same formal structuralization, when strictly applied to the classification of technical literature for retrieval purposes, has proved inadequate. At another extreme, approaches to indexing which preclude any hierarchical association are similarly disappointing. The dual dilemma is illustrated in the following quotation:*

> The English language is so rich that [even] many of the most explicit technical and scientific concepts may be represented by several different word symbols (or combinations of word symbols). This apparent literary advantage can become a formidable retrieval disadvantage unless [some means] is developed to enable the user to express his information needs in his vocabulary, and to retrieve relevant information expressed by an originator in his own entirely different vocabulary. Similarly, a user must be able to express his information needs on a generic, or "inclusive class," level with reasonable expectation that the documents retrieved will include not only those which discussed information on the generic class level, but also those which discussed information on the level of the specific members of that generic class.

The differences among information retrieval systems today relate to the manner in which the two problems of ambiguity and specificity are treated. Off-line, batch-processed retrieval suffers an inherent disadvantage of providing no intermediate results for user evaluation and subsequent search redefinition. For these and other reasons, it is felt by many that an on-line computer system, which allows a user to converse directly with the computer in his quest for relevant document citations, can provide a more effective environment for information retrieval than is possible with off-line systems. The on-line system permits information retrieval to be a highly individualized process with respect to time of occurrence, question at hand, and characteristics of user.

Computer technology currently allows the configuration of a real-time, user-directed information storage and retrieval system. Less understood, however, is the problem of directing and controlling such a hardware configuration so as to allow a user who is neither knowledgeable about nor interested in computers to obtain useful results from a large file of document descriptions (citations) in a rapid, convenient, and effective manner.

Related to a previous experiment, CONVERSE, the DIALOG system was developed to investigate the effectiveness of a flexible, user-directed language in accomplishing reference retrieval.

Dialog development

The effectiveness of an on-line, user-directed retrieval system lies in the degree to which it can accomplish the following:

- Provide a variety of "command" functions for communication, search, and display of information from which the user can select those most appropriate to his particular problem.
- Provide the flexibility to include additional commands or other operational modes as new search techniques are developed.
- Assist the user in search definition and in full employment of system capabilities.
- Allow intermediate user evaluation of search results with subsequent request refinement.
- Require a minimum of bookkeeping or remembering on the part of the user in the association of retrieved references with request expressions.
- Minimize elapsed time between query and response.

*"The Engineers Joint Council Action Plan," Appendix II, *Thesaurus of Engineering Terms*, New York, Engineers Joint Council, May 1964.

- Eliminate need for "middle-man" request interpretation by system specialists.
- Allow real-time interaction between user and system for search guidance.

Although a "free-form" language was considered for communication between the user and the computer, it was decided that a better balance between man inconvenience and machine inconvenience was attained through the use of several predefined commands which could be modified by the user according to his own needs. Such a structure allows modular development of the system and also permits the easy incorporation of additional commands if or as the need for them arises.

It was felt that although most users would not be familiar with Boolean algebra, some method of coordinate searching should be allowed. The conclusion was the development of the COMBINE command. If A and B are sets of documents the first of which contains descriptor A and the second of which contains descriptor B, COMBINE A + B results in a set of documents each of which contains either index term A or index term B; COMBINE A*B results in a set of documents each of which contains both index terms A and B; COMBINE A – B results in a set containing term A but not term B. Any set can be used in subsequent COMBINE commands to recursively partition the reference set into successively more relevant subsets. In this manner, a Boolean search strategy evolves in stepwise fashion, and the user is provided information at the conclusion of each step to assist him in defining the next step.

The dialog language

The current operating environment of the DIALOG system in an IBM 360/30 (32 thousand bytes of core) together with two 2311 disk packs (7.5 million bytes each) for programs and intermediate storage, a 2321 Data Cell (415 million bytes) mass storage device for the reference corpus, a 1443 off-line printer, and a 2260 display/1053 printer input/output terminal (Figure 1). The reference file consists of some 300,000 NASA announced citations.

The DIALOG system provides a number of commands which appear as the upper case or shift values of the top row of keys on the display keyboard (Figure 2). The depression of these command keys, together with entry of associated operands, enable the user to instruct the computer in a desired sequence of operations. A search consists of (1) identifying and selecting descriptors (subject of index terms) which reflect the user's interest, (2) combining descriptors into search expressions, and (3) examining retrieved citations and modifying search expressions.

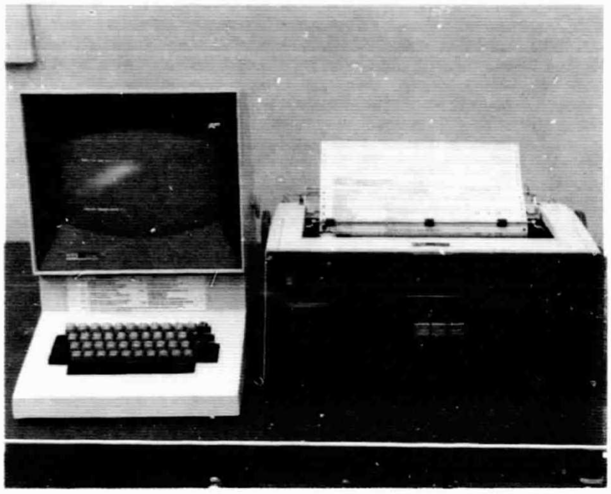


Figure 1. – Remote access terminal

Figure 2. – Remote terminal keyboard

Identification and selection of descriptors

To determine whether a certain term has been used to index documents under consideration, the command EXPAND, entered together with the term, causes a display of a list of actual descriptors alphabetically close to the term entered. Each descriptor is shown with a temporary identification number by which it may be referenced as long as it appears on the display. For each descriptor so displayed, the number of citations to which that descriptor was assigned, as well as the number of terms conceptually related to that descriptor, are also displayed. A display of the conceptually related terms for any displayed descriptor may be obtained by depressing the EXPAND key and entering the descriptor identification number appearing on the display. Any displayed term can be selected by depressing the SELECT command and then entering the descriptor identification number which appears on the display. SELECT causes the citations containing the selected descriptor to be collected for further processing. Each selected descriptor is assigned an identification number and is typed out on the console printer together with the number of citations to which it has been assigned as an index term.

Combination of descriptors

The number of citations associated with any given descriptor is likely to be large (500 to 10,000). Although it is possible to display citations from a single-descriptor set, it is probably more efficient to specify a combination of descriptors which must be present in a citation before it is retrieved. The COMBINE command is used for this purpose. Assume a person is interested in documents pertaining to welding defects in aluminum, and has selected the terms: (1) WELDING (used in 2239 citations), (2) DEFECT (used in 1206 citations), and (3) ALUMINUM (used in 7137 citations). By combining these three terms (i.e., COMBINE 1*2*3 where * stands for "and"), a fourth set of 13 citations results, each of which contains all three terms.

By allowing the repeated use of sets generated by one COMBINE command in the definition of other COMBINE commands, the user can converge in step-wise fashion on citations of interest. At each step he is provided the size of the resultant set and can either examine individual citations in that set, or modify the set by combining it with other sets (with the COMBINE command).

Examination of retrieved citations and search expression modification

Citations can be displayed wherever desired with the DISPLAY command. This operation will frequently allow the user to discover new descriptors and add them to his search list to further specify his interest.

The KEEP command allows the user to set aside or save selected citations for later printout either on the console printer (TYPE command) or on the off-line printer (PRINT command).

It is possible to further restrict a retrieved set of citations by year of announcement, announcement media (IAA, STAR, CSTAR), or announcement series number. The command LIMIT is provided for this purpose. Although primarily of interest to library personnel, this command allows limiting on any or all of the categories just described.

A search example using dialog

Assume that the user is interested in reports dealing with the transfer of aerospace technology to industry. On the bottom line of the display screen (Figure 2) he will see the message:

ENTER NEXT COMMAND ►

This is an indication that searching can begin. The user initiates his search by depressing the BEGIN SEARCH command key. This results in an interview display which, when completed, appears as in Display 1.

```
PLEASE ENTER THE INFORMATION REQUESTED BELOW.  ENTER EACH LINE AS COMPLETED.
-----------------------------------------------------------------------------
1. SEARCH TITLE: TECHNOLOGY TRANSFEN TO INDUSTRY
2. NAME: E.L.BRADBURY
3. MAIL STOP: 301-3
4. BRANCH: ELECTRONICS
```

Entry of the final item (the word "electromagnetics" in this case) causes the console printer to print out the search heading at the top of Figure 3. Each subsequent step of the search is typed out on the console printer and appears as Figure 3. (The reader should refer to this figure to follow the search procedure.)

```
SEARCH TITLE:  TECHNOLOGY TRANSFER TO INDUSTRY
DATE:          05/23/67
REQUESTOR:     E.L.BRADBURY,301-3,ELECTRONICS

                    SET  NO.IN   DESCRIPTION OF SET
COMMAND-OPERAND(S) NO.   SET    (+=OR,*=AND,-=NOT)
------------------ --- ------ ---------------------------------------------
E-TECHNOLOGY
E-E5
S-E5                 1   2666 TECHNOLOGY
S-E10                2    364 AEROSPACE TECHNOLOGY
C-1+2                3   2745 1+2
S-TRANSFER           4   7891 TRANSFER
S-INDUSTRY           5   1126 INDUSTRY
C-3*4*5              6     23 (1+2)*4*5
D-6
S-UTILIZATION        7    394 UTILIZATION
C-4+7                8   8266 4+7
C-3*8*5              9     40 (1+2)*5*(4+7)
C-9-6               10     17 ((1+2)*5*(4+7))-((1+2)*4*5)
D-10
P-9
                          1-40        ITEMS HAVE BEEN PRINTED.
OUR DIVISION, A HEADQUARTERS COMPONENT, OFTEN HAS TO RESPOND IN A SHORT TIME TO
REQUESTS FOR INFORMATION FROM OUR MANAGEMENT   THUS FAR THIS SYSTEM FOR QUICK  A
CCESS TO INFORMATION IS THE ONLY ONE I HAVE SEEN THAT CAN MEET OUR NEEDS.

TOTAL TIME ELAPSED FOR THIS SEARCH IS   8.49 MINUTES.
```

Figure 3. – Search example (console printer output)

To begin his search, the user wishes to see if "technology" has been used as an index term, and, if so, how many citations contain it as an index item. The user depresses the EXPAND key on the display keyboard and types in "technology." The console printer prints out a command echo (Figure 3) consisting of the first letter of the command (E) together with its operand ("technology"). This provides the user a visual check on what the computer received. The command response then appears on the display screen as shown in Display 2.

```
                    EXPAND-TECHNOLOGY
REF        DESCRIPTOR                      CITATIONS REL. TERMS REF
E1    TECHNICAL                               347               E1
E2    TECHNICAL DRAWING                                   1     E2
E3    TECHNICAL WRITING                         4               E3
E4    TECHNIQUE                              4696        16     E4
E5   *TECHNOLOGY                             2666         4     E5
E6    TECHNOLOGY /GEN/                          2               E6
E7    TECTONIC MOVEMENT                        33               E7
E8    TECTONICS                                68               E8
E9    TEE                                      17               E9
ENTER NEXT COMMAND▶
```

Notice that the display shows the terms alphabetically near to "technology" (which itself is indicated with an *). By displaying the alphabetically near terms, the user is able to see not only if and to what extent the term he entered has been used as an index term, but also any spelling or ending variations on the term which have been used (e.g., weld versus welding). The user need not even spell the term correctly. "E" numbers are assigned to the displayed index terms for reference purposes.

It can be seen that "technology" is used in an index term in 2,666 citations and has 4 related terms entries in the thesaurus. (Related terms refer to conceptually or hierarchically related terms which are usually associated with a particular term.) To examine the related terms, the user depresses EXPAND and types in E5 which results in Display 3.

```
                    EXPAND-E5
REF        DESCRIPTOR                      CITATIONS REL. TERMS REF
E5   *TECHNOLOGY                             2666         4     E5
E10   AEROSPACE TECHNOLOGY                    364               E10
E11   BIOTECHNOLOGY                            68               E11
E12   MILITARY TECHNOLOGY                     189         1     E12
E13   REACTOR TECHNOLOGY                      161               E13


ENTER NEXT COMMAND▶
```

The user notices the descriptor "aerospace technology" and reasons that for his purposes "technology" and "aerospace technology" are equivalent. He thus selects the two terms "technology" and "aerospace technology," and combines the corresponding sets into a third set containing 2,745 citations (in which each citation contains either "technology" *or* "aerospace technology" as index terms). The console typewriter response for these commands is shown in lines 3 through 5 in Figure 3. (Note that OR is coded as "+," whereas AND is coded "*.")

The user now continues his search by selecting "transfer" and "industry" (shown as lines 6 and 7 of Figure 3). He is now ready to combine his selected terms to define his search topic. He wants each citation retrieved to contain either "technology" or "aerospace technology" and "transfer" and "industry." He can effect such a set by depressing the COMBINE key and typing in "3*4*5." This command results in set 6 which contains 23 citations (shown as line 8 of Figure 3).

To display these citations, the user depresses the DISPLAY key and types in 6 (the set number), which results in a display of the first citation in the set as shown in Display 4.

```
                         DISPLAY     6/2/1
65A31673       00/07/65      UNCLASSIFIED
  SPIN-OFF FROM SPACE.   (N;ASA INFORMATION SYSTEM TO ASSIST TRANSFER OF TECHNOL
OGICAL DATA FROM SPACE PROGRAMS TO POTENTIAL BENEFICIARIES)
  15KERR, B. M.    /NASA, SCIENTIFIC AND TECHNICAL 16*INFORMATION DIV., WASHINGT
ON, D.C./. 203049 40SCIENCE JOURNAL, VOL. 1, JUL. 1965, P. 85-90.
  KERR, B. M.
  / AEROSPACE/*AEROSPACE TECHNOLOGY/ DATA/ INDUSTRY/ INFORMATION/*INFORMATION RE
TRIEVAL/*NASA PROGRAM/ PROGRAM/ RETRIEVAL/ SPACE/ TECHNOLOGY/ TITANIUM/ TRANSFER


ENTER NEXT COMMAND▶
```

Note that the response contains the three specified terms: "transfer," "aerospace technology," and "industry." Successive items can be displayed by depressing ENTER.

Assume the user continues stepping through set 6 to item 3 (Displays 5 and 6). Individual citations can be printed by depressing the PRINT key.

```
                         DISPLAY     6/2/2
65N16989*   NASA-CR-51214   NASR-162   00/06/63     UNCLASSIFIED
  AEROSPACE RESEARCH APPLICATIONS CENTER SUMMARY REPORT, 1 APRIL TO 30 JUNE 1963
   (AEROSPACE RESEARCH APPLICATIONS - CONFERENCE)
  WEIMER, A. M.
  1E576200INDIANA UNIV. FOUNDATION, BLOOMINGTON.
  / AEROSPACE/*AEROSPACE TECHNOLOGY/ APPLICATION/ COMMERCIAL/*CONFERENCE/ INDUST
RY/ NASA PROGRAM/ RESEARCH/ TRANSFER


ENTER NEXT COMMAND▶
```

```
                         DISPLAY     6/2/3
66N13375#   NASA-CR-68620 ER-S8-1844   NASW-1139   00/04/65     UNCLASSIFIED
  SPACE "TECHNOLOGY "APPLIED TO "MAN*S "EARTHLY "NEEDS - "A FEASIBILITY STUDY ON
 THE TRANSFER OF AEROSPACE TECHNOLOGY TO INDUSTRY USE   (FEASIBILITY STUDY ON AC
CELERATING TRANSFER OF AEROSPACE TECHNOLOGY TO COMMERCIAL INDUSTRY - AEROSPACE L
ITERATURE APPLICABILITY TO INDUSTRY)
  BROCK, A. W. DEMBICZAK, W. J. NAGY, A.
  AS334557AMERICAN MACHINE AND FOUNDRY CO., SANTA AS334557BARBARA, CALIF.
  / AEROSPACE/*AEROSPACE TECHNOLOGY/ APPLICATION/ COMMERCIAL/ EVALUATION/*INDUST
RY/ INFORMATION/*INFORMATION RETRIEVAL/ LITERATURE/ QUALITY/ RETRIEVAL/ SURVEY/
TECHNICAL/ TECHNOLOGY/ TRANSFER/ UTILIZATION
ENTER NEXT COMMAND▶
```

The user continues stepping through the set of citations and notices in item 3 that the term "utilization" is used in the same sense as "transfer" (i.e., "technology transfer to industry" versus "technology utilization by industry"). He thus decides to broaden his search expression to include "transfer" or "utilization." This is accomplished with the following commands:

Command	*Operand(s)*
SELECT	UTILIZATION
COMBINE	4 + 7
COMBINE	3*8*5

(The results of these commands can be followed in Figure 3.) The final COMBINE command results in set 9 containing 40 citations. To see if the broadened definition returned relevant citations, the user wishes to examine the items in set 9 which do not appear in set 6 (the first search expression). This is accomplished by COMBINE 9-6 (resulting in set 10 containing 17 citations) and DISPLAY 10. A few of the results are shown in Displays 7, 8, and 9.

```
                         DISPLAY   10/2/1
65N18316#   NASA-CR-50668   NASR-63/03/   00/00/63    UNCLASSIFIED
 (UTILIZATION OF NASA SPACE TECHNOLOGY BY MIDWESTERN INDUSTRY)
  64N63-18316  MIDWEST RESEARCH INST., KANSAS CITY, 61MO. UTILIZATION OF NASA-GE
NERATED SPACE TECHNOLOGY 428Y MIDWESTERN INDUSTRY  QUARTERLY PROGRESS REPORT 43N
O. 3, 5 MAY - 5 AUG. 1962 H. M. GADBERRY  <1965< 6630P /NASA CONTRACT NASR-63/03
// /NASA CR-50668/ 65OTS-  $2.60 PH, $1.10 MF
  GADBERRY, H. M.
  MZ513670MIDWEST RESEARCH INST., KANSAS CITY, MO.
  / CONCEPT/*INDUSTRY/*NASA PROGRAM/ SPACE/ TECHNOLOGY/ UTILIZATION

ENTER NEXT COMMAND▶
```

```
                         DISPLAY   10/2/2
65N83633    NASA-TM-X-51711   27/04/64     UNCLASSIFIED
  THE NASA PROGRAM FOR STIMULATING INDUSTRIAL UTILIZATION OF GOVERNMENT-SPONSORE
D TECHNOLOGY
  DENNISON, J. T.
  NE368373NATIONAL AERONAUTICS AND SPACE ADMINISTRATION, NE368373WASHINGTON D.
C.
  / CONFERENCE/ INDUSTRY/ NASA PROGRAM/ PROGRAM/ SIMULATION/ TECHNOLOGY/ UTILIZA
TION

ENTER NEXT COMMAND▶
```

```
                         DISPLAY   10/2/3
66N12622      00/00/65     UNCLASSIFIED
  THE UNIVERSITY AND TECHNOLOGY UTILIZATION   (UNIVERSITY PROGRAMS AND TECHNOLOG
Y UTILIZATION - EDUCATION AND INDUSTRY)
  TERMAN, F. E.
  S0380476STANFORD UNIV., CALIF.
  / CONFERENCE/ DEVELOPMENT/*EDUCATION/*INDUSTRY/ NASA PROGRAM/ PROGRAM/ RESEARC
H/ SCIENCE/ SPACE/ TECHNOLOGY/ TRAINING/ UNIVERSITY/*UNIVERSITY PROGRAM/ UTILIZA
TION

ENTER NEXT COMMAND▶
```

Any or all of these citations or their accession numbers (identification numbers) can be printed out for future reference (in this example the user prints set 9). The search expression could be further broadened or narrowed by including additional terms. The retrieved sets could be limited by various parameters such as date and publication type. The search is completed when the user depresses END SEARCH. This command results in Display 10 which the user completes. Entry of this information causes elapsed search time and user comments to be printed on the console typewriter, and clears the computer for the next search.

```
PLEASE ENTER COMMENTS, SUGGESTIONS AND CRITICISMS IN THE SPACE BELOW. DEPRESS
ENTER UPON COMPLETION.
-------------------------------------------------------------------------------
 OUR DIVISION, A HEADQUARTERS COMPONENT, OFTEN HAS TO RESPOND IN A SHORT TIME TO
 REQUESTS FOR INFORMATION FROM OUR MANAGEMENT.  THUS FAR THIS SYSTEM FOR QUICK
ACCESS TO INFORMATION IS THE ONLY ONE I HAVE SEEN THAT CAN MEET OUR NEEDS!



TOTAL TIME ELAPSED FOR THIS SEARCH IS   8.49 MINUTES.
```

CONCLUSION

The DIALOG language was developed as a proprietary product by the Information Sciences group at the Lockheed Palo Alto Research Laboratory. It has been implemented with the NASA collection of over 300,000 citations, and is currently being applied to a collection of personnel summaries.

Search topics which have been executed using DIALOG include:

- Interaction of magnetosphere and solar wind
- Vibrational excitation of carbon dioxide
- Molybdenum disulfide as a solid lubricant in spacecraft
- Gas phase reactions of fluorocarbons with oxygen and nitrogen

Search times vary considerably among users, depending on the user's experience and on the complexity of the search. A reasonable average, however, is about 30 minutes of elapsed time and 5 to 10 minutes of computer time per search. At $125 per hour average machine charge, this represents a cost of $10 to $20 per search.

In summary, there are five important characteristics of the DIALOG language:

- The search question is constructed at search time (rather than at index time as is the case with a manual system).
- DIALOG is designed for nonspecialists; i.e., the users themselves, and thus avoids one communication barrier.
- The command language is independent of the particular data it searches.
- As an on-line system, it allows continual redefinition of the search question, based on examination of intermediate results.
- Control of the process lies with the user; the computer merely serves as a data-processing extention of the user.

BIBLIOGRAPHY

D L DREW R K SUMMIT R I TANAKA and R B WHITELEY

An on-line technical library reference retrieval system
American Documentation 17 No 1 3 1966
E HERBERT
Information transfer
International Science and Technology No 51 26 1966
N S PRYWES
Brousing in an automated library through remote access
Computer Augmentation of Human Reasoning Spartan Books Inc Washington D. C. 1965
Information
Scientific American Book W. H. Freeman & Comapny San Francisco Calif. 1966