ARCHIVING AND RETRIEVAL OF DATA FROM THE *INTERNATIONAL ULTRAVIOLET EXPLORER (IUE)* MISSION

W. H. Warren Jr.
National Space Science Data Center
NASA/Goddard Space Flight Center

D. F. Alderman
Astronomy Department
Computer Sciences Corporation

ABSTRACT

The data base of IUE spectra available for dissemination to the scientific community presently contains approximately 22,000 separate images. Due to storage space and computer processing time limitations, it was decided to store the data on high density (6250 bpi) specially formatted tapes and to retrieve and reconvert them to the originally supplied format upon request. The archival/retrieval system for the data and an automatic data entry system for creation and maintenance of an indexing data base are described. Suggestions are made for the use of standard formats for future image data storage, retrieval and dissemination.

I. INTRODUCTION

Ultraviolet spectral data returned by the IUE spacecraft are processed by the IUE observatories at the Villafranca Satellite Tracking Station (VILSPA) near Madrid, Spain and at NASA's Goddard Space Flight Center and are subsequently archived at the National Space Science Data Center (NSSDC)/World Data Center A for Rockets and Satellites, the VILSPA data center, and at the World Data Center C at Appleton Laboratory in the U.K. The three agencies which operate the spacecraft maintain independent archives both for safety purposes and so that data can be disseminated by each agency to its respective user community. The NSSDC receives data tapes from both the NASA and VILSPA IUE observatories on medium density (1600 bpi) tapes and must index and store the data for distribution to the scientific community. Although the NASA observatory now provides the images on blocked tapes, the tapes provided by both agencies were unblocked until about a year ago, and the VILSPA tapes are still not blocked. Since a rather large compression factor is possible in going from 1600 bpi unblocked to 6250 bpi blocked format, it was decided to process at NSSDC all data from both observatories to

C. Jaschek and W. Heintz (eds.), Automated Data Retrieval in Astronomy, 115–117.

the higher density with the longest possible physical records (blocks), thereby greatly reducing the number of tapes while simultaneously decreasing computer processing time. All tapes received are therefore processed by a special program to reblock them to 6250 bpi with variable block size and spanned logical records (VBS format).

Table 1 demonstrates the advantages of processing the data for storage and retrieval purposes. The VBS record format results in a decrease of about a factor of 10 in the input/output computer time when data are retrieved to prepare requester tapes. The significantly fewer

Table 1. Comparison of Original and Reblocked Tapes

	Original Tapes	Archive Tapes
Density	1600 bpi	6250 bpi
Blocking	Unblocked	Blocked 32,000 bytes/record
Record Format	Variable	Variable, Spanned
Number of Tapes	3084	310

number of archive tapes also greatly simplifies the retrieval of images since fewer tapes need be read for recovery of a given number of files. More importantly, due to limited storage space for tapes at our computer facility, archive tapes must be stored at another location and sent to the computer area to be processed; hence, the physical handling of tapes and requirements for storage at the computer center are markedly decreased.

II. THE DATA MANAGEMENT SYSTEM

Even with the fewer number of tapes resulting from reblocking of the data, it is obviously of great importance to use a computerized list of images to locate data to be retrieved from the archive tapes. Maintenance of an image indexing data base has proved to be a significant part of the IUE task, since we require the index to contain enough information for cross checking to locate mis-specified and incomplete identifications on requests (several errors have been discovered in this way). Until recently, the indexing data base was updated manually, with the usual error and identification problems, especially for images which had been reprocessed and supplied as replacements. Also, the archiving and project facilities maintained independent data bases containing considerable amounts of redundant information and software for data base management. With the development by the IUE project and observatory of more sophisticated management and housekeeping routines, it was decided that the data bases should be merged, thus allowing the use of previously developed software for editing and automatic entry of new records to the unified data base.

Upon completion of each high-density archive tape, a program is run which determines the locations of all images on the tape. This information is written to a disk data set in the form of commands to be used by a subsequent background editing program which updates the data base. The disk data set is quality checked at an interactive computer terminal and added to another disk data set, which is actually used as the input to the editing program. After the editing run has been made, the input data set is reinitialized and ready for additional entries for newly completed archive tapes.

A special accounting program, previously used to monitor the IUE project data base, has been expanded to analyze and compute statistics for NSSDC information in the merged data base. Weekly reports are now produced which list images delivered to NSSDC but overdue for processing or shipment to the European data centers.

III. SUMMARY

The archiving and indexing of IUE spectroscopic data at NASA has progressed from a largely manual operation to almost an entirely automatic procedure through the development of processing and data management software. At present the automated system is used only for data taken at the NASA IUE Observatory, but efforts are now underway to create an automated data base system for VILSPA data. The new system has proven very effective in eliminating most errors and increasing the efficiency of our retrieval and dissemination efforts. The improved indexing and decreased turn-around time for data availability will provide faster and easier IUE data accessibility for research and analysis purposes.

An additional possible improvement for data processing of observations taken in a guest-observer environment such as IUE, where processed data are transmitted to each observer for further processing and analysis at one's home institution, would be to adopt a standard format for data transmission. The current IUE image-type format, although well suited to the information content, consists of mixed mode IBM binary and EBDIC data which are sometimes difficult to read and process by other computer systems. The use of a flexible format such as FITS (Wells and Greisen 1979), which has been designed specifically for the transport of image data to other installations, would greatly facilitate data handling by the astronomers themselves, an obviously very important aspect of a guest-observer structured program.

REFERENCE

Wells, D.C., and Greisen, E.W. 1979, in *Image Processing in Astronomy*, Ed. G. Sedmak, M. Capaccioli, R.J. Allen, Osservatorio Astronomico di Trieste, p. 445.

https://doi.org/10.1017/S0252921100082786 Published online by Cambridge University Press