The IUE Final Archive

J.D. Ponz

ESA-D/TOS, Villafranca, Apartado 50727, 28080 Madrid, Spain

1 Introduction

Archives are fundamental components in any space project because they contain information with very high value, based on observations that, in most of the cases, cannot be repeated. In the case of space astronomy, archives are unique tools that allow the study of variable phenomena, and provide reference information to calibrate instruments of new missions. In addition, new technologies in data storage and networking, together with algorithms developed in the last years, are opening new research areas. The correlation of the information in different archives and new statistical methods make it possible to analyze large, homogeneous data sets. Therefore, archives are essential sources of information that must be preserved for future generations of scientists.

In the context of the IUE project, these ideas were identified at an early stage, and a special effort has been dedicated to the definition of the mission archive and to the distribution of the spectra to the scientific community world-wide. IUE has led the way in promoting archive services.

The IUE Final Archive, was defined as a collaboration between NASA, ESA and PPARC, to generate a high quality and uniform spectral archive during the final phase of the mission – when specialized knowledge on the instrument and calibration procedures were still available –, and to maintain it accessible to the scientific community as the historical reference of the mission.

Within this project, a new data processing system was developed, NEWSIPS and was implemented at both observing stations so that each site would process its own data, creating in this form the master version of the archive.

The new processing system, together with the new calibrations, allowed to process all data in a uniform way. In addition, especial attention was dedicated to the compilation and verification of the image parameters (*Core Data Items*) required to process correctly the image, and essential for scientific analysis. Finally, given that each observing station process its own data, quality control procedures were defined so as to ensure full compatibility of output products.

2 The Production at Villafranca

The Final Archive production at Villafranca was implemented in two main phases: (1) Data preparation and (2) actual archive production.

During the data preparation phase, all raw images were collected and stored into optical disks and the Core Data Items, critical parameters for both scientific interpretation and reprocessing of an observation, where verified.

The archive production is based on NEWSIPS, developed and implemented at both observing stations. The main elements for the production system at Villafranca are:
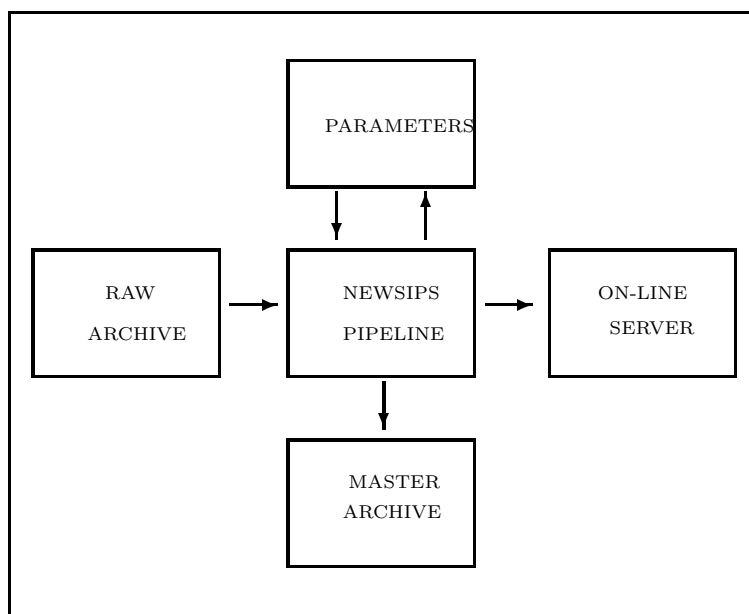
Raw archive stored on optical disks, provides the input images to the production pipeline.

Image parameters define the main object characteristics and instrumental setup. These parameters determine the calibration options, and provide quality control information.

Pipeline NEWSIPS, implemented under MIDAS. Output products are generated in FITS format.

Master archive with reduced spectra and intermediate data files stored on optical disks.

On-line data server extracted spectra – both two-dimensional and absolutely calibrated fluxes – are transferred to the distribution node after quality control, and are available on-line.

The complete system runs in a distributed environment, so that several nodes, structured in three main subsystems, are involved in the complete data cycle:

1. The optical disk subsystem, a VAX/VMS server to handle input and output staging areas and dedicated optical disk units.

2. The pipeline subsystem based on NEWSIPS, has been implemented in a set of distributed processing queues, running on several nodes: four DEC Ultrix workstations for low-dispersion and one DEC Unix workstation for high-dispersion processing.

3. The on-line archive subsystem provides the required input parameters, and, after processing and quality control, stores output parameters and extracted spectra. This subsystem is implemented in a DEC Unix server. Its central components are the WWW server and the relational database, Oracle, that stores input and output parameters as required by the pipeline, and provides access to the catalogue to retrieve extracted spectra.

In addition, output products from GSFC are periodically ingested in the archive on optical disk and extracted spectra are catalogued and stored in the on-line archive for data distribution.

Quality control is the paramount requirement in a reference archive of this nature. Quality control procedures are applied both to data and software.

Automatic and interactive procedures are implemented to insure the completeness and accuracy of the information at three levels: (1) input parameters and raw images, (2) pipeline and (3) master archive.

Input parameters are interactively verified using as reference the hand-written observing logs. In addition, automatic procedures are used to check the validity of the values and the correct size and format of the input raw images.

The pipeline includes a quality control mechanism based on thresholds for certain output parameters; using this mechanism, alarms are triggered to identify images requiring a special action, to correct the input information or a specific process.

Quality control on the master archive is based on automatic procedures to check that files have the correct size and format, and on interactive examination of a selected set of files per optical disk.

The complete software system is under configuration control. In addition, special effort has been dedicated to its quality at three levels: (1) science verification, (2) compatibility tests and (3) system stability.

Science verification is done by analyzing a set of images selected so that they would cover an ample range of observing modes. In order to enforce compatibility of products, a set of images were processed at both sites and compared for each NEWSIPS release. Finally, system stability is monitored by processing the same set of data at periodic intervals.

Finally, operations have been setup to minimize manual intervention in the complete archive production cycle, with interactive procedures centered into quality control tasks. This was an essential requirement since most of the archive production was done while the satellite was still in operation.

3 Project Documents

The project was implemented following ESA Software Engineering Standards. A selected list of project documents is included below for reference.

1. "ESA Software Engineering Standards". ESA PSS-05-0, Issue 2, 1991.

2. "IUE Final Archive Reprocessing Reference Document" version 3.0, July 1992.

3. "IUE Final Archive Reprocessing at VILSPA. Users Requirements Document", version 1.4. SD/8002-03/MB-RG-AT/920718.

4. "IUE Final Archive Reprocessing at VILSPA. Software Requirements Document", version 2.1. SD/8003-03/JDP-TM/920720.

5. "IUE Final Archive Reprocessing at VILSPA. Architectural Design Document. Part I: Data Preparation", version 1.2. SD/8005-02/JDP-IY-MB-JJAG/920721.

6. "IUE Final Archive Reprocessing at VILSPA. Architectural Design Document. Part II: Production System", version 1.1. SD/8006-01/JDP/920731.

7. "CDI Verification Users Manual" version 1.4. SD/8007-04/MB-JDP/920120.

8. "NEWSIPS Operations Manual", version 1.2. SD/8012-02/ESL-IY/940915.

9. "IUEFA Production Database Description", SD/8014-00/IY/950329.

10. "Core Data Items Definition". TN/8002-03/AT/920120.

11. "FITS Format for Final Archive", version 1.7. TN/8005-07/JDP/940709.

12. "Calibration Plan". TN/8006-00/AC/900320.

13. "SOAR Software Intallation and Testing". TN/8008-00/IY/900604.

14. "CDI Extraction from Headers. Algorithms and Data flow". TN/8009-01/IY/910118.

15. "Setup for CDI Extraction within the ISAS VAX Cluster". TN/8010-00/EOP/910129.

16. "Statistics on VILSPA Raw Images". TN/8016-01/IY-JDP/911219.

17. "The FITS IMAGE extension. A proposal". TN/8017-01/JDP/911218.

18. " Verification of the IUEFA Production System". TN/8020-01/JDP-IY-MB/930414

19. "IUEFA Quality Control Procedures". TN/8022-04/JDP/931014

20. "NewSIPS Operations Plan at VILSPA". TN/8023-01/TM/930423

21. "IUEFA Production DB Operations Manual". TN/8027-00/IY/9310130

22. "IUEFA Output Products Retrieval". TN/8034-02/IY/940830

23. "CDI Population Guide". TN/8035-01/IY/940715

24. " Optical disk generation". TN/8038-00/IY-TM-JDP/950315