Science Data Processing for the Multi-angle Imaging SpectroRadiometer

Graham Bothwell, Earl G. Hansen, Robert E. Vargo
Jet Propulsion Laboratory, California Institute of technology
4800 Oak Grove Drive
Pasadena, CA 91109-8099

Abstract - **Ground processing of data from the Multi-angle Imaging SpectroRadiometer (MISR) instrument, part of NASA's Earth Observing System (EOS), uses new and unique science algorithms because MISR's multi-camera observing method has not previously been used anywhere. Extensive prototyping was required from a relatively primitive status. The data volume is large, necessitating an innovative software design approach that maximizes throughput. While the routine processing software was developed at the Jet Propulsion Laboratory (JPL), data processing occurs at the NASA Langley Research Center using the EOS Core System (ECS), a collaborative arrangement that works well. Since the launch of MISR on the Terra Spacecraft in December 1999, with the availability of actual mission data, MISR's actual computational needs have become better known, and increased efficiencies are being implemented to make best use of available computing resources. This science software is a sound foundation for processing the data from potential future multi-angle instruments.**

I. INTRODUCTION

The Multi-angle Imaging SpectroRadiometer (MISR) [1] is one of the instruments aboard the Terra spacecraft, which is part of NASA's Earth Observing System (EOS), and was launched in December 1999. Data from MISR are processed into standard data products by the Atmospheric Sciences Data Center at NASA Langley Research Center (LaRC), also known as the LaRC Distributed Active Archive Center (DAAC). This paper addresses development of this data processing capability, in particular the science software that was implemented by the MISR Project at the Jet Propulsion Laboratory (JPL).

Fig. 1 illustrates the processing stream for standard MISR data products at the DAAC. Processing takes place within separately executable entities called Product Generation Executables (PGEs). The PGE is the task entity for production scheduling, and each PGE can contain one or more executable processes. MISR has standard products at Levels 1, 2, and 3. (Product levels are defined in the EOS Reference Handbook. [2])

II. UNIQUE ALGORITHMS

MISR uses a multi-angle multi-camera concept not previously flown in space, requiring new and unique approaches to science algorithms that have no inheritance from earlier missions or instruments. For example, there are new and extensive geometric calibration techniques, unique methods for retrieval of atmospheric aerosol information, and stereo techniques that have not previously been implemented for extensive, on-going systematic operation.

Key to the development of the processing capability was assembly of a team that included the requisite discipline specialists, and the co-location of the primary system developers within that team. In effect, a unified team developed algorithms and implemented the operational software system. Extensive prototyping occurred at JPL, such as for the Level 1 photogrammetric algorithms and the Level 2 aerosol/surface products. The main area where the developers relied on science team members resident outside of JPL for initial prototyping and ongoing collaboration is the cloud algorithms, which originate from the University of Arizona, University College London, and the University of Illinois. Other prototyping was done at Los Alamos National Laboratory, Boston University, and the University of Miami. In all these cases, the MISR developers implemented new processing codes based on specifications and prototypes by the outside team members and their associates.

Substantial development at the MISR Science Computing Facility (SCF) was necessary to prepare ancillary tables and databases to be referenced during standard product generation. For example, Level 1B2, which is the geometric rectification and coregistration of the nine sets of MISR imagery [3], required several workyears of effort to prepare a global Digital Elevation Model (DEM) based on DTED-1 and other sources; a Projection Parameter (PP) data set used in translating navigation and pointing information to ground location; and Reference Orbit Imagery (ROI) that will allow improved ground location through image matching. These tasks are computationally intensive, to the extent that care was needed to ensure the work needed to be done once only. For example, one production run for Projection Parameters required approximately four months of continuous processing using 16 CPUs of a Silicon Graphics Origin-2000 computer.

The amount of code in MISR's standard science processing is approximately 0.5 million lines of code (LOC) excluding comments, while the code at the SCF for production of ancil-

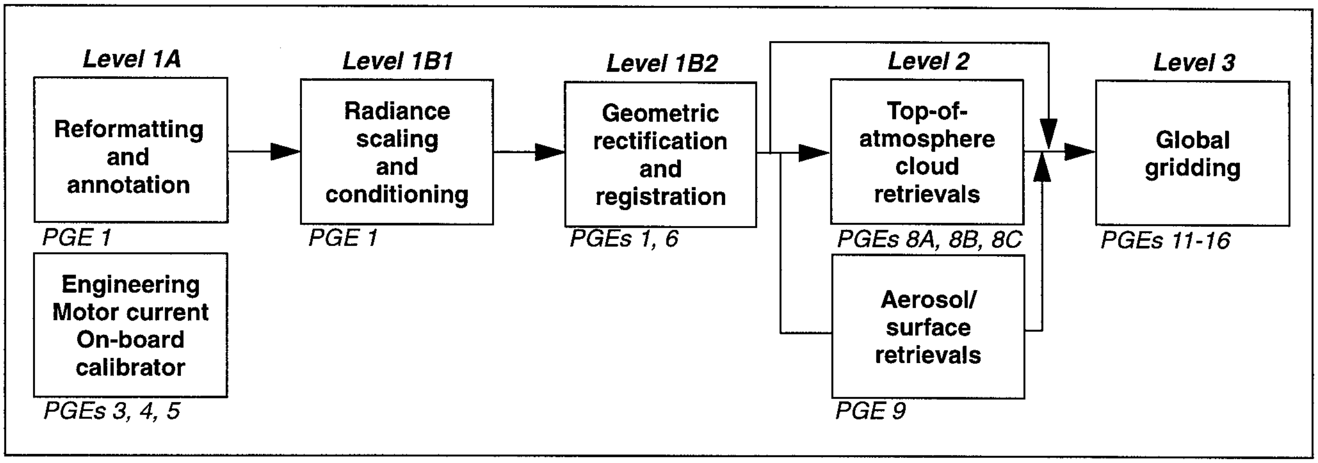

Fig. 1: MISR standard processing

lary data sets, quality assessment, and calibration is approximately 0.4 million LOC excluding comments.

III. THROUGHPUT AND EFFICIENCY

Because of the large volume of MISR data, over 35 Gigabytes of raw data from the instrument each day, and the computationally intensive nature of the processing, the MISR science software was designed from the start to maximize throughput. Thus it takes advantage of opportunities for data parallelism and concurrent processing, minimizes disk accesses, and attempts to optimize use of the EOS Data and Information System (EOSDIS) capabilities. It was also necessary to have processing that runs entirely unattended. Here are some examples.

A. Data access and data sharing

A multi-process architecture was developed by which data are passed from process to process in memory, avoiding the overhead of writing and reading data from storage between processing steps. All Level 1 and Level 2 processing utilizes this approach. It was implemented at a significant development and integration expense to the MISR team, since this is much more complicated than a simple one based on single-processes that have disk input/output (I/O) between the processes. This architecture has achieved its goal very well.

B. Reworking of design

Initial measurements of aerosol/surface processing speed using actual code in mid-1999 suggested a deviation by approximately two orders of magnitude from the estimates that were used to specify the installed processing hardware. This being an obvious threat to the success of the MISR experiment, it was decided to completely rework the software design. Over fifty workmonths were expended in this effort, and the results were dramatic, reducing the 100-fold increase to within specification.

Although further, minor software-engineering-based improvements to the MISR software will continue to be made, the future evolution of the software will be driven primarily by evolution of the algorithms. The software will also continue to be upgraded to ensure that it can be readily adapted at short notice for future missions. As a new experiment, MISR has benefited from the long lead time of the Terra mission, and will be able to benefit future missions where the development time may be shorter, and where the software will need to be readily portable to the most cost-effective hardware platforms of that time.

IV. OPERATING ENVIRONMENT

MISR science data processing is designed to operate within the EOS Core System (ECS) computer environment, which uses primarily Silicon Graphics, Inc., (SGI) mini-supercomputers, and which includes system-wide functions embracing data ingest, data archiving, data staging, job control, science processing, product distribution, and user services. Unique characteristics of the MISR data necessitated MISR-unique aspects of the ground processing system implemented for the Terra missions. For example, the data are downlinked and front-end handled in the traditional fashion of time-aggregation, but must transition into an orbit-by-orbit handling because this is the way that the ground processing is done. Another example is that, because MISR data products are produced on a per orbit basis, they require subsetting for users who are unable to accept very large products. Also under development is a MISR-unique data base of Quality Assessment (QA) data, required because the product parameters within whole-orbit granules vary along the swath.

A continuing major challenge is that the processing hardware capacity is based on estimates made several years ago, when many of the algorithms were still at the prototype stage. While hardware augmentation has been speeded up, there has also been substantial MISR software development effort to fit within the installed capacity. This has been highly successful with respect to product volume, and much has also been done to improve the processing performance.

Having no heritage, MISR was dependent upon early versions of production code to make any performance measurements. These codes were not viable until mid-1999. Besides indicating the science software performance, such measurements have also assisted greatly in improving the understanding of system behavior. For example, early estimates of processing needs were based mainly on hand calculations, and included an application environment overhead of 100% to allow for metadata generation, product structure constructions related to the Hierarchical Data Formal (HDF) and HDF-EOF formats used for the EOS products, input staging, job control, and any other production environment requirements. Experience with the system since Terra launch has shown that because of system downtime for maintenance, Level 0 data delays, inefficiencies in data staging, limitations on maximal CPU utilization, and general system instability, a more realistic overhead allowance would be 175% to 200% with respect to science software execution times.

V. DEVELOPMENT EXPERIENCE

The MISR science developers and the NASA Langley DAAC are on opposite sides of the North America continent, but this has presented no real problems. During the on-orbit phase of the mission, activities have been greatly facilitated by a dedicated high-speed link between the two sites. MISR staff visited Langley for extended periods during the early software deliveries. The delivery process has now settled down into a satisfactory routine, and is done entirely remotely.

One of the larger lessons emerging from MISR's experience is that the development of this kind of science software is unavoidably difficult, encompassing far more than a straightforward "bullet-proofing" of science prototype code. Instead, it is a major design and implementation challenge. Some of the issues include: error handling; unit testing; working through the data; using the system environment dependencies; autonomous mass production; continually changing requirements; broad dependencies; pressure to write code before requirements are complete; lack of prototypes for many of the algorithms.

Associated with this is the challenge of developing simulated data sets for testing the algorithms and software. Without a predecessor instrument, simulated data were constructed by aggregating Landsat scenes and applying a transformation to simulate multi-angle views. This was a complex job, and the extent of the simulated data that MISR developed was only a single swath without atmospheric effects. This was used extensively for end-to-end testing of both Level 1 and Level 2 processing.

VI. CONCLUSION

The first generation of MISR Level 1 data products became available to the science community in July 2000, and Level 2 products became available in February 2001. This is a major accomplishment for such a new instrument concept launched little more than a year ago. Validation of the algorithms is in progress.

Although the MISR science software will be subject to ongoing revision and improvement during the Terra mission, the foundations of the design and the implementation are well established, and highly successful. Many innovations and specialized approaches were necessary because of the newness of the MISR multi-angle observing concept. These are groundbreaking achievements that will no doubt be the foundation of data processing for future on-orbit, multi-angle instruments.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the dedication and inspiration of the MISR science data system team. This work was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. For further information, see the MISR web site at *http://www-misr.jpl.nasa.gov.*

REFERENCES

[1] D.J. Diner, J. Beckert, T. Reilly, C. Bruegge, J. Conel, R. Kahn, J. Martonchik, T. Ackerman, R. Davies, S. Gerstl, H. Gordon, J-P. Muller, R. Myneni, P. Sellers, B. Pinty, and M. Verstraete, "Multi-angle Imaging SpectroRadiometer (MISR) instrument description and experiment overview," IEEE Trans. Geosci. Rem. Sens. **36**, pp. 1072-1087, 1998.

[2] "EOS Data and Information System (EOSDIS)," EOS Reference Handbook, pp 27-35, 1999

[3] V.M. Jovanovic, M. M. Smyth, J. Zong, R. Ando, G.W. Bothwell, "MISR photogrammetric data reduction for geophysical retrievals," IEEE Trans. Goesci. Rem. Sens., **36**, No. 4, 1998