Physics Mining of Multi-Source Data Sets
Goddard Space Flight Center, Greenbelt, Maryland

Powerful new parallel data mining algorithms can produce diagnostic and prognostic numerical models and analyses from observational data. These techniques yield higher-resolution measures than ever before of environmental parameters by fusing synoptic imagery and time-series measurements. These techniques are general and relevant to observational data, including raster, vector, and scalar, and can be applied in all Earth- and environmental-science domains. Because they can be highly automated and are parallel, they scale to large spatial domains and are well suited to change and gap detection. This makes it possible to analyze spatial and temporal gaps in information, and facilitates within-mission re-planning to optimize the allocation of observational resources.

The basis of the innovation is the extension of a recently developed set of algorithms packaged into MineTool to multi-variate time-series data. MineTool is unique in that it automates the various steps of the data mining process, thus making it amenable to autonomous analysis of large data sets. Unlike techniques such as Artificial Neural Nets, which yield a blackbox solution, MineTool’s outcome is always an analytical model in parametric form that expresses the output in terms of the input variables. This has the advantage that the derived equation can then be used to gain insight into the physical relevance and relative importance of the parameters and coefficients in the model. This is referred to as “physical mining of data.” The capabilities of MineTool are extended to include both supervised and unsupervised algorithms, handle multi-type data sets, and parallelize it.

The innovations include: (1) Physics mining algorithms, enabling derivation of analytical relations and physical models from observational data; (2) Automated, parallel algorithms, enabling a high degree of automation and parallelization, scaling to large spatial domains well-suited to change and gap detection; (3) Local versus global modeling, to generate locally optimal models appropriate to a specific geospatial region accounting for the unique setting and conditions; (4) Fusion of multi-source, multi-type data that yield higher-resolution measures than ever before by fusing synoptic imagery and independent time-series measurements; and (5) Calculation of Palmer’s Drought Severity Index Analogue.

Successful completion of this project will lead to a major breakthrough in the climate study in particular, and to analysis of multi-source data as applied to the hydrologic cycle affecting climate change impacts and resource management.

This work was done by John Helly, Homa Karimabadi, and Tamara Sipes of Caltech for NASA’s Jet Propulsion Laboratory. Further information is contained in a TSP (see page 1), GSC-15802-1

Photogrammetry Tool for Forensic Analysis
This acquires visual crime and accident scene data for later processing.
John F. Kennedy Space Center, Florida

A system allows crime scene and accident scene investigators the ability to acquire visual scene data using cameras for processing at a later time. This system uses a COTS digital camera, a photogrammetry calibration cube, and 3D photogrammetry processing software.

In a previous instrument developed by NASA, the laser scaling device made use of parallel laser beams to provide a photogrammetry solution in 2D. This device and associated software work well under certain conditions. In order to make use of a full 3D photogrammetry system, a different approach was needed.

When using multiple cubes, whose locations relative to each other are unknown, a procedure that would merge the
data from each cube would be as follows:
1. One marks a reference point on cube 1, then marks points on cube 2 as unknowns. This locates cube 2 in cube 1's coordinate system.
2. One marks reference points on cube 2, then marks points on cube 1 as unknowns. This locates cube 1 in cube 2's coordinate system.
3. This procedure is continued for all combinations of cubes.
4. The coordinate of all of the found coordinate systems is then merged into a single global coordinate system.

In order to achieve maximum accuracy, measurements are done in one of two ways, depending on scale: when measuring the size of objects, the coordinate system corresponding to the nearest cube is used, or when measuring the location of objects relative to a global coordinate system, a merged coordinate system is used.

Presently, traffic accident analysis is time-consuming and not very accurate. Using cubes with differential GPS would give absolute positions of cubes in the accident area, so that individual cubes would provide local photogrammetry calibration to objects near a cube.

This work was done by John Lane of ASRC Aerospace for Kennedy Space Center. Further information is contained in a TSP (see page 1). KSC-12975