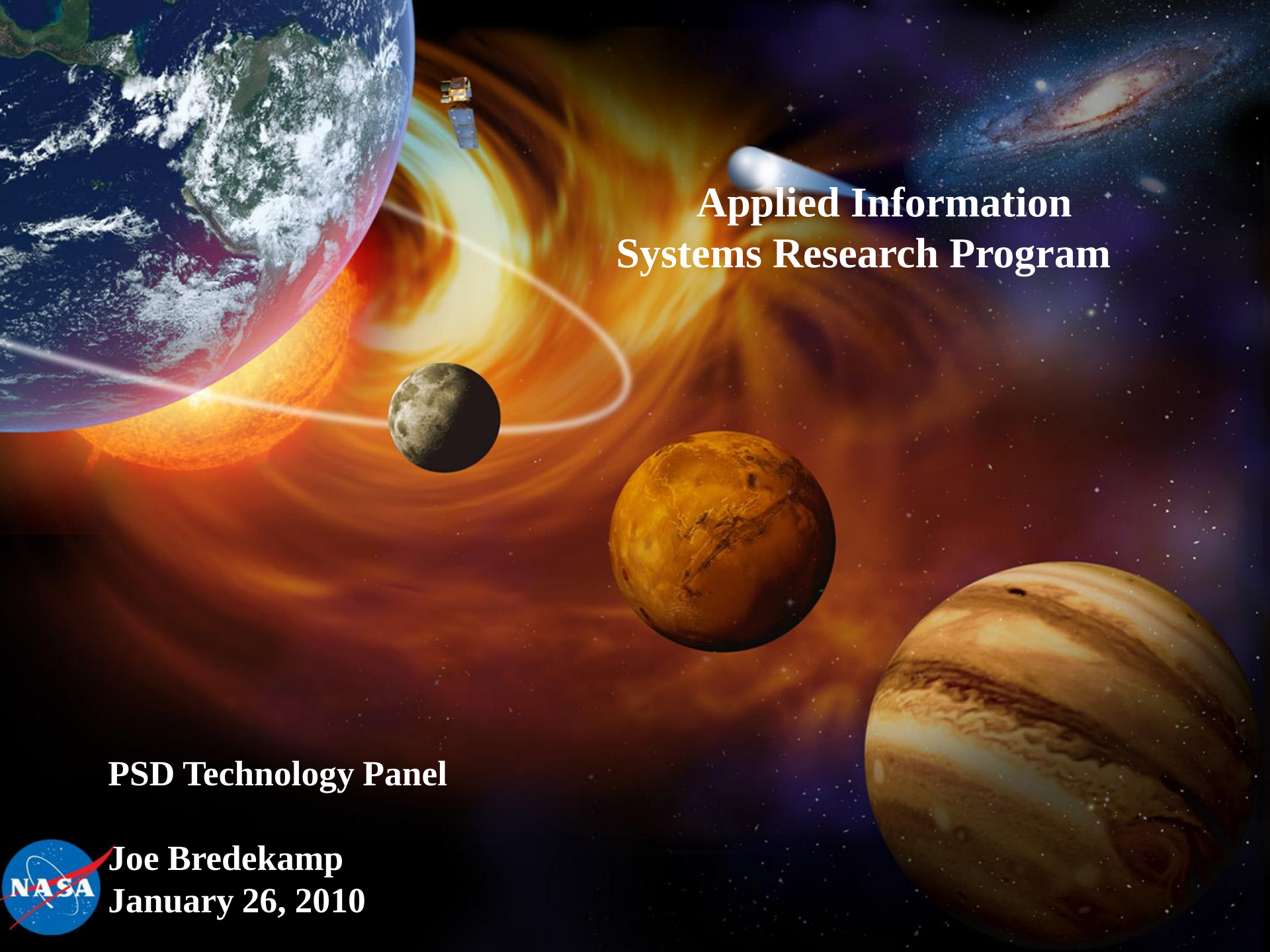


Applied Information Systems Research Program

PSD Technology Panel

Joe Bredekamp
January 26, 2010

Applied Information Systems Research Program

- ***Purpose*:** Exploit advances in computer science and technology to increase productivity of research endeavors sponsored by Science Mission Directorate

- ***Specific Goals:***
  - Investigate novel information technologies and computational methods that have the potential to increase productivity of the SMD research endeavors, and which would extend the state-of-the-practice in those endeavors;
  - Demonstrate the degree of relevance, applicability, and potential impact of emerging information technologies to SMD missions and programs; and
  - Foster interdisciplinary collaborations that span the space science, Earth science, and computer science disciplines.
- ***Notional Areas of Interest:***
  - Data/information synthesis and intelligent knowledge capture;
  - Computational methods and algorithms for scientific analysis and discovery;
  - Multidimensional data representation, visualization, and comparison;
  - Distributed collaborative frameworks; and/or
  - Autonomy and Systems Engineering

Applied Information Systems Research Program Heritage

- Virtual Observatory foundation activities
  - Szalay– Standardized Sky Partitioning– Java classes for Hierarchical Triangular Mesh algorithm following JSKY standards
  - Hanisch et al– Interoperable Sys for Archival Information Access
  - Roberts- Virtual Space Physics Observatory; VisBARD
  - McGlynn- SkyView Virtual Telescope
  - Crichton--Object-oriented middleware for PDS
- Tools and services
  - Borrill- Cosmic Microwave Background Analysis Tools (COMBAT)
  - Puetter- PIXON Image reconstruction technique
  - Loredo, Scargle, Jewell, Weinberg—Statistical tools, inference models, textural methods for analysis
  - Kinter- GrADS—Gridded Analysis and Display System
  - Cornillon- DODS– Distributed Oceanographic Data System, morphed to OpenDAP

Applied Information Systems Research Program

Funding: Approximately $6M /year

# Proposals	$ Requested	# Selected	$ Available
160	34	22	4.5
110	24	12	2

Issues:

- Technology infusion …..avoid "valley of death" syndrome
- Visibility and ownership within the divisions
- Need to reassess overall SMD strategy forward

AISR Portfolio
Computational Methods and Information Management

Bergeron, Daniel	U of New Hampshire	Interactive out-of-core visualization of very large multiresolution time series data
Borrill, Julian	Lawrence Berkeley Nat Lab	A High Performance Computing Framework for Cosmic Microwave Background Data Management
Broxton, Michael	Carnegie Mellon University	Automated Digital Terrain Model Generation from HiRISE and LROC Stereo Imagery
Brunner, Robert	U of Illinois	Advanced Astrophysical Algorithms to Novel Supercomputing Hardware
Connolly, Andrew	U Washington	Streaming the sky: high performance knowledge discovery
Gardner, Jeffrey	Carnegie Mellon	Enabling Massive Scientific Databases Through Automated Schema Design
Gardner, Jeffrey P	U Washington	Astronomy in the Cloud
Gogarten, Peter	University of Connecticut	Exploration of Novel Methods to Visualize Genome Evolution.
Gombosi, Tamas	University of Michigan	Development of an Adaptive Non-Ideal MHD Simulation Tool for Multiple Space Science Applications
Head, James	Brown University	Advanced Visualization in Solar System Exploration and Research (ADVISER)
Lo, Martin	JPL	MTool 1.0: Massive High Dimension Data Analysis and Scientific Discovery Tool
MacNeice, Peter	NASA Goddard	Magnetogram Synthesis - A Vital Data Analysis Component of A Space Weather Prediction Infrastructure

AISR Portfolio
Computational Methods and Information Management

McGlynn, Thomas	NASA Goddard	Presence, Personalization and Persistence: A New Model for Doing Science in a Collaborative Archive Environment
Mighell, Kenneth	NOAO	Parallel-Processing Astrophysical Image-Analysis Tools
Olson, Kevin	Drexel University	PARAMESH: A parallel, adaptive, grid tool for the Space Sciences.
Ptak, Andrew	Johns Hopkins U	On-the-fly and Grid Analysis of Astronomical Images for the Virtual Observatory
Quinn, Thomas	University of Washington	Visualization of Terascale Datasets with Impostors
Richardson, Mark	California Institute of Technology	Planetary Atmospheric Data Assimilation System
Schombert, James	University of Oregon	ExGKS: Extragalactic Knowledge System
Sussman, Alan	U of Maryland	Robust Grid Computing using Peer-to-Peer Services
Szabo, Adam	NASA Goddard	Extending Virtual Observatory Capabilities Through Grid Enabled Resources
Szalay, Alexander	Johns Hopkins University	Framework for Fast Spatial Searches Using the Hierachical Triangular Mesh

AISR Portfolio
Computational Methods and Information Management

Ford, Eric B.	U of Florida	Parallel Integration of ODEs with GPUs with Application to N-Body Integrators and Planetary Systems
Szapudi, Istvan	U of Hawaii	Scalable Algorithms for Analysis of Megapixel CMB Maps and Large Databases
Thakar, Aniruddha	Johns Hopkins U	Large Scale on Demand Cross-Matching with Open SkyQuery
Tricarico, Pasquale	Planetary Science Institute	A Distributed Computing System Supporting Near Earth Asteroids Research
Young, Eliot	Southwest Research Institute	Neural Net Spectral Analysis of Icy Volatiles in the Solar System
Young, Eliot	Southwest Research Institute	Application of Iterative Blind Deconvolution

AISR Portfolio
Machine Learning and Statistical Techniques

Burl, Michael	Jet Propulsion Laboratory	Directed Exploration of Complex Systems
Gardner, Jeffrey	Carnegie Mellon University	Knowledge Discovery in a Virtual Universe
Johnson, Jay	Princeton U	Novel Higher-Order Statistical Method for Extracting Dependencies in Multivariate Geospace Data Sets
Kargupta, Hillol	University of Maryland Baltimore County	Distributed and Peer-to-Peer Data Mining for Scalable Analysis of Data from Virtual Observatories
Kashyap, Vinay	Harvard/SAO	Developing Methods to Incorporate Calibration Uncertainties in Data Analysis
Knuth, Kevin	University at Albany	Bayesian Source Separation for Astrophysical Spectra: Application to PAHs
Loredo, Thomas	Cornell Un	Quantified Uncertainty: Flexible Probabilistic Modeling of Dynamic Spectra and Other Astronomical Data
Merenyi, Erzsebet	Rice University	A neural map view of planetary spectral images for precision data mining and rapid resource identifi
Miller, Chris	NOAO/CTIO	Facilitated Access and Application of Computational Astrostatistics Algorithms
Morris, Robin	Ames/USRA	Statistical Inference for Scientific Instruments: Event Analysis for the Gamma-Ray Large Area Space
Morris, Robin	Universities Space Research Association	Improving Remote Sensed Data Products Using Bayesian Methodology for the Analysis of Computer Model Output

AISR Portfolio
Machine Learning and Statistical Techniques

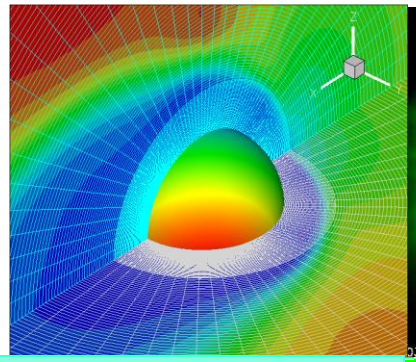
Nasraoui, Olfa	University of Louisville	Solar Loop Mining to Support Studies of the Coronal Heating Problem
Newman, Timothy	UA Huntsville	Auroral Phenomenon Localization, Classification, and Temporal Evolution Tracking
Scargle, Jeffrey	NASA Ames	Novel Methods for the Analysis of Photon-Limited Data
Schafer, Chad	Carnegie MU	Stochastic Models for High-Dimensional, Nonstandard Data
Stepinski, Tomasz	USRI/LPI	Automated Identification and Characterization of Landforms on Mars
Strachan, Leonard	Harvard/SAO	Innovative Techniques for Producing Corrected Synoptic Maps for a UVCS Outflow Velocity Database
Szalay, Alexander	JHopkins U	Application of Advanced Statistical Crossmatching Technologies for the Hubble Legacy Archive
Talukder, Ashit	NASA JPL	Global cyclone detection and tracking (GLYDER) for climate variability studies: A multisensor data processing approach
Turmon, Michael	Jet Propulsion Laboratory	Predicting the Perfect Storm: Feature Identification and State Estimation for Nonlinear Systems
Wagstaff, Kiri	Jet Propulsion Laboratory	Detecting Transient Surface Features with Dynamic Landmarking
Weinberg, Martin	U Massachusetts	Enabling Bayesian Inference for the Astronomy Masses
Wettergreen, David	Carnegie Mellon University	Automated Orbital Mapping: Statistical Data Mining of Orbital Imagery to Analyze Terrain, Summarize its Characteristics and Draft Geologic Map

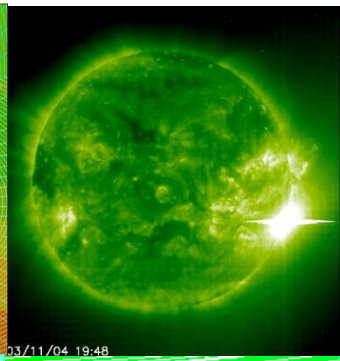
AISR Portfolio
Autonomy and Systems Engineering

Belbruno,Edward	Princeton U	Mission Extension Using Sensitive Trajectories and Autonomous Control (contd)
Bose, Tamal	Utah State U	Adaptive Algorithms For Optimal Classification and Compression of Hyperspectral Images
Crumpler, Larry	New Mexico Museum of Natural History and Science	Tactical Rover-based Martian Geologic Mapping
Dasu, Aravind	Utah State U	An Integrated Software Environment to Design Polymorphic Fault Tolerant Processors for Command and Control Functions on RadHard FPGAs
Dolan, John	Carnegie Mellon University	An Analytical Tool for Robot Mission Reliability Prediction
Giuliano, Mark	Space Telescope Science Institute	Multi-Objective Optimal Scheduling for Space Science Applications
Gruenwald, Le	University of Oklahoma	Estimating Missing Data in Sensor Network Databases Using Data Mining to Support Space Data Analysis
Jewell, Jeffrey	NASA JPL	EPISODE - Software for Trajectory Generation and Nonlinear Continuous Control in the Presence of Uncertainty
Li, Rongxing	Ohio State U	Integration of Orbital, Descent and Ground Imagery for Topographic Capability Analysis in Mars Landed Missions
Lorenz, Ralph	JHU Applied Physics Lab	Intelligent Sensor Network Study of Dust Devils
Williams, Brian	MIT	Monitoring and Diagnosis of Complex Software and Hardware for Earth Observing Missions

BATS-R-US is a Multi-Physics Code

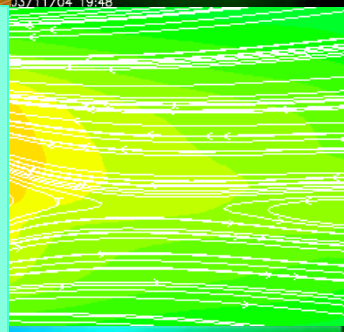
- Compressible fluid dynamics
- Ideal MHD with mass addition
- Physics-based energy transport
  - Heat conduction
  - Wave energy transport

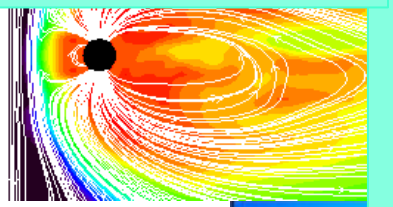


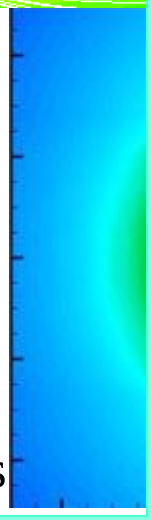


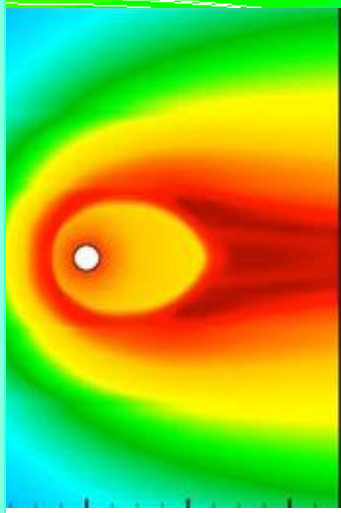
Supported by AISRP

- Resistive MHD
  - Current driven resistivity models
  - Non-gyrotropic resistivity model
- Hall MHD
  - Keeps the Hall term in Ohm's law
  - More realistic reconnection rate
- Semi-relativistic MHD
  - Displacement current is kept in Ampère's law
  - Limits all wave speeds by c
- Multi-species/Multi-fluid MHD
  - Separate continuity, momentum and energy equations

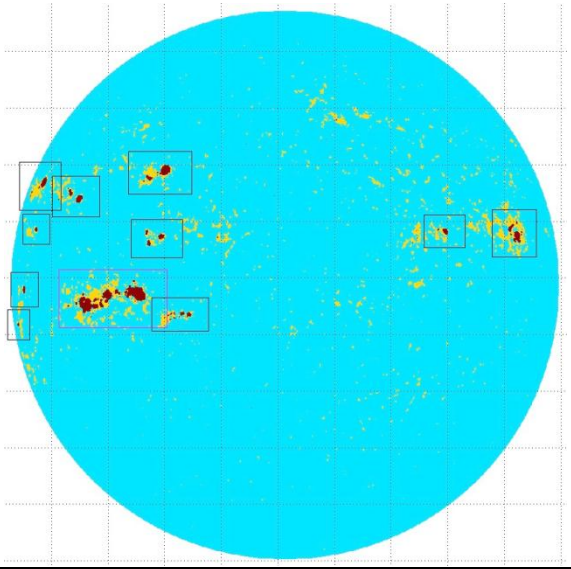








Classification of Solar Imagery



Objective:

Develop and validate algorithms for object identification, tracking, and track analysis

Enabling Technical Innovations:

- Markov Random Field (MRF) classifiers
- Single-link most-likely tracker
- Kalman filter object-behavior modeling

	Milestones
FY07	Integrated and tested improved MRF optimizer
	Calibrated tracker to agree with NOAA catalog
FY08	Port to HMI pipeline, develop models for proxy HMI data
	Greater tracking automation and coverage

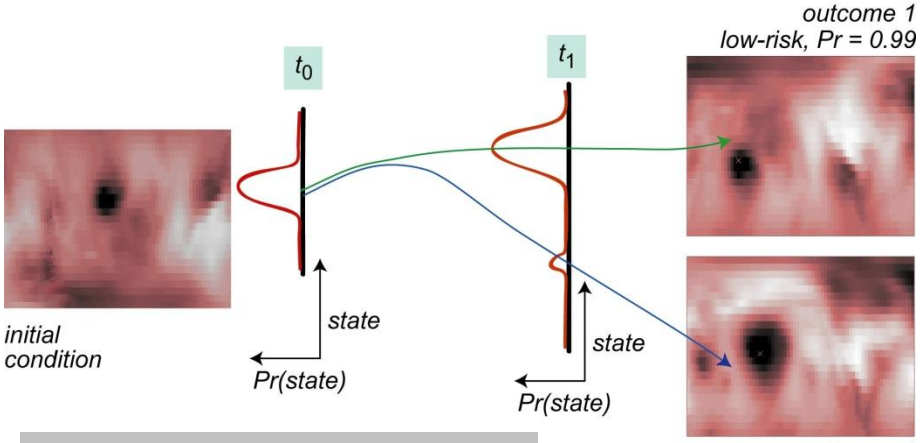
Mission Customers: Solar Dynamics Observatory (HMI), SoHO/MDI

Expected Impact:

- Address data overload problem by concentrating attention on objects of interest
- Develop models for objects via their intrinsic behavior

PI: Michael Turmon

Predicting the Perfect Storm: Project Summary



Measure the Disparate Outcomes Due to Nonlinear State Evolution

Objective:

Provide a conditional forecasting capability, enabled by posing scenarios using human-interpretable features, not at the level of grid cells.

Methods:

Extensions of ensembles for representing competing hypotheses about system state. Learning dynamical rules for features from data.

Customers: Ocean/atmosphere forecasting and NWP, mission simulation

	$K	Milestones
FY08	200	Identify & track objects (sim. data)
		Identify & track objects (obs. data)
FY09	200	Learn forward object dynamics
		Identify dynamic path to end state
FY10	200	Uncover low-probability storm path

Sponsor: NASA SMD (AISR)

Unique Innovations:

Rather than repeated simulation ("dart throwing") to uncover high-risk scenarios, our system allows experts to pose them via high-level features and directly compute their relative odds.

Maturity:

Beginning (2007) TRL is 2-3; exit (2010) TRL is 4.

PI: M. Turmon (JPL)
Team: Jewell, Chin, Ghil (UCLA/ENS)

A Neural Map View of Planetary Spectral Images

for Precision Data Mining and Rapid Resource Identification

Project Summary

We develop neural machine learning algorithms based on advanced Self-Organizing Maps for learning, visualizing, identifying the detailed structure of high-dimensional manifolds. This facilitates precise capture of data clusters of vastly varying statistical properties, including discovery of interesting, rare ones; and supports subsequent sophisticated supervised classification. We also use neural learning to identify the importance of data dimensions in a novel way. We investigate the composition of planetary surfaces from spectral and hyper-spectral imagery with these tools.

600,000 pixels

Impact on Science

-Precise exploitation (deep information mining) of large, complex scientific spectral data sets collected in space and Earth science missions.

- Novel knowledge discovery, data mining and prediction capabilities.

- Envisioned as intelligent, real-time on-board clustering/classification tool: algorithms lend themselves for implementation in massively parallel hardware, to continuous learning on chip and to reconfiguring as needed.

- Similar potential benefits to bioinformatics (microspectral imagery), and other areas with complex high-D data sets.

Last Year's Highlights

Invented a new representation & visualization for learned SOM know-ledge, more complete than other exis-ting SOM visualzations; and new quality measures of SOM learning

Adapted (mathematically) neural relevance learning (invented in 2001) for hyperspectral dimensions; selected "important" dimensions are different from PCA or wavelet selected dimensions and retain more information

Completed comprehensive surface mapping from Mars Pathfinder images

2 journal papers; 4 refereed conference papers; 7 presentations; 3 journal papers in preparation

Plans for Next Year

Improve cluster capture based on our new SOM knowledge representation & visualization

Develop novel cluster validity indices that incorporate the new knowledge representation

Cluster/classify spectral images from the Mars Exploration Rovers, and from the Cassini VIMS; predict surface properties of icy satellites

Erzsébet Merényi, PI
Rice University

Robert H. Brown, Co-I
University of Arizona

William H. Farrand, Co-I
Space Science Institute

Tomas Villmann, Collaborator
University of Leipzig;
Colin Fyfe, Collaborator
University of Paisley

URL: http://www.ece.rice.edu/~erzsebet

NASA AISR 2004-07

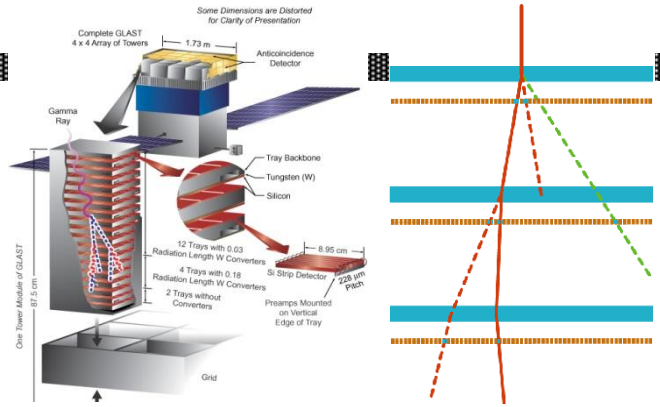
Event Analysis for the Gamma-ray Large Area Space Telescope

Project Summary

Develop statistical methodology to analyze the response of the Large Area Telescope (LAT) instrument on the GLAST satellite. Account for the uncertainty in the structure of the event in the analysis. Use the known statistics of the physical processes to build up a statistical analysis methodology that accurately captures and uses the physics of the detector to give an accurate estimation of the azimuth, elevation and energy of the incident photons.

Robin D Morris
USRA-RIACS
rdm@riacs.edu

Johann Cohen-Tanugi
Stanford Linear Accelerator Center



Applications

The projects primary application is to event analysis for GLAST, however the methods developed will be suitable for other instruments. The event analysis methodology is being implemented within the GLEAM (GLAST Event Analysis Machine) software framework, (GLEAM is built on the GAUDI framework for high-energy physics instrument analysis.) The code developed can therefore be adopted directly by the GLAST mission when it reaches a suitable level of maturity and has been demonstrated as effective.

Last Year's Highlights

- Code development to analyze events from electrons incident on the LAT
- Identification of production of photons in the conversion layers
- Invited presentation at "Interface 2006"
- Presentations at Valencia 8, SCMA IV and univ. seminars

Plans for Next Year

- Develop the analysis methodology for all the physics processes
- Extensive testing on simulated gamma-ray events
- Code release to GLAST project

NASA AISR 2004-07

Scalable Algorithms for Fast Analysis of Megapixel CMB Maps and Large Astronomical Databases

Istvan Szapudi

szapudi@ifa.hawaii.edu

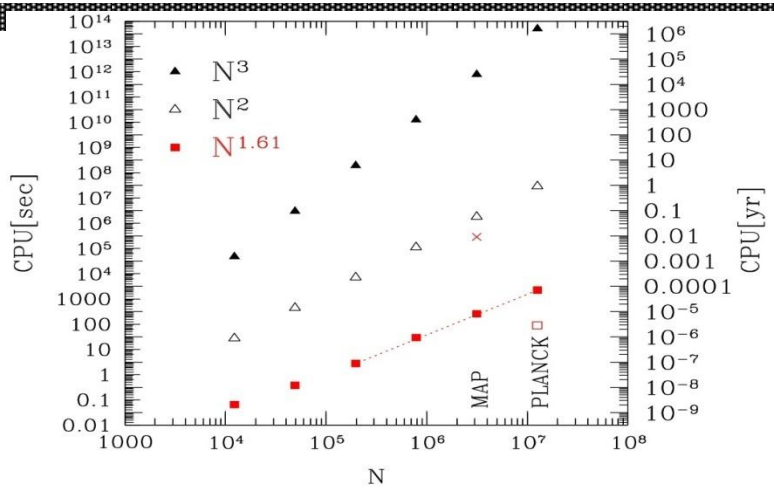
Project Summary

Develop fast scalable methods to measure the angular spatial statistics of the Cosmic Microwave Background.

Plans for Upcoming Year

We plan to generalize the our methods

- higher order statistics
- angular three-point functions
- bi-spectra
- Euclidean versions



Applications

Development of SpICE (Spatially Homogeneous Correlation Estimator) in the previous cycle. WMAP1 angular power spectrum and WMAP2 cumulant correlator power spectra have been analysed with SpICE. WMAP + SDSS cross correlations have been calculated with SpICE.

Highlights

- Spice has been downoaded over 260 times
- Official Level 2 Planck software
- Constrained Dark Energy from measurements of galaxy-CMB correlations
- Constrained primordial non-Gaussianity

CMB, Statistics

http://www.ifa.hawaii.edu/~szapudi

AISRP NNG06GE71G