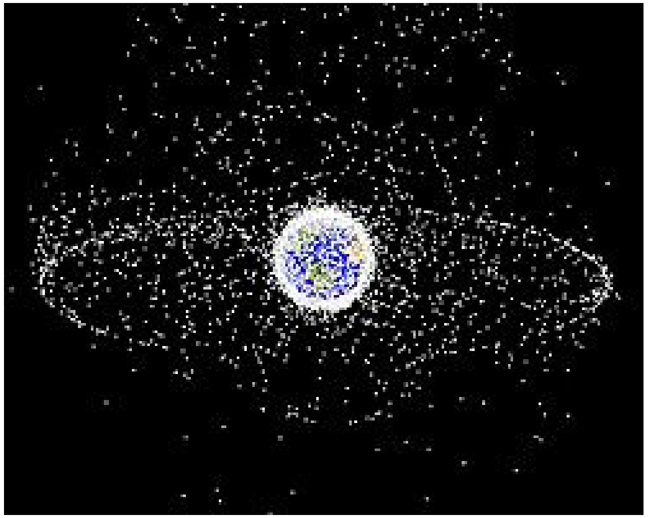
Some Applications of Nonnegative Tensor Factorizations (NTF) to Mining *Hyper*spectral & Related Tensor Data

Bob Plemmons
Wake Forest

Some Comments and Applications of NTF

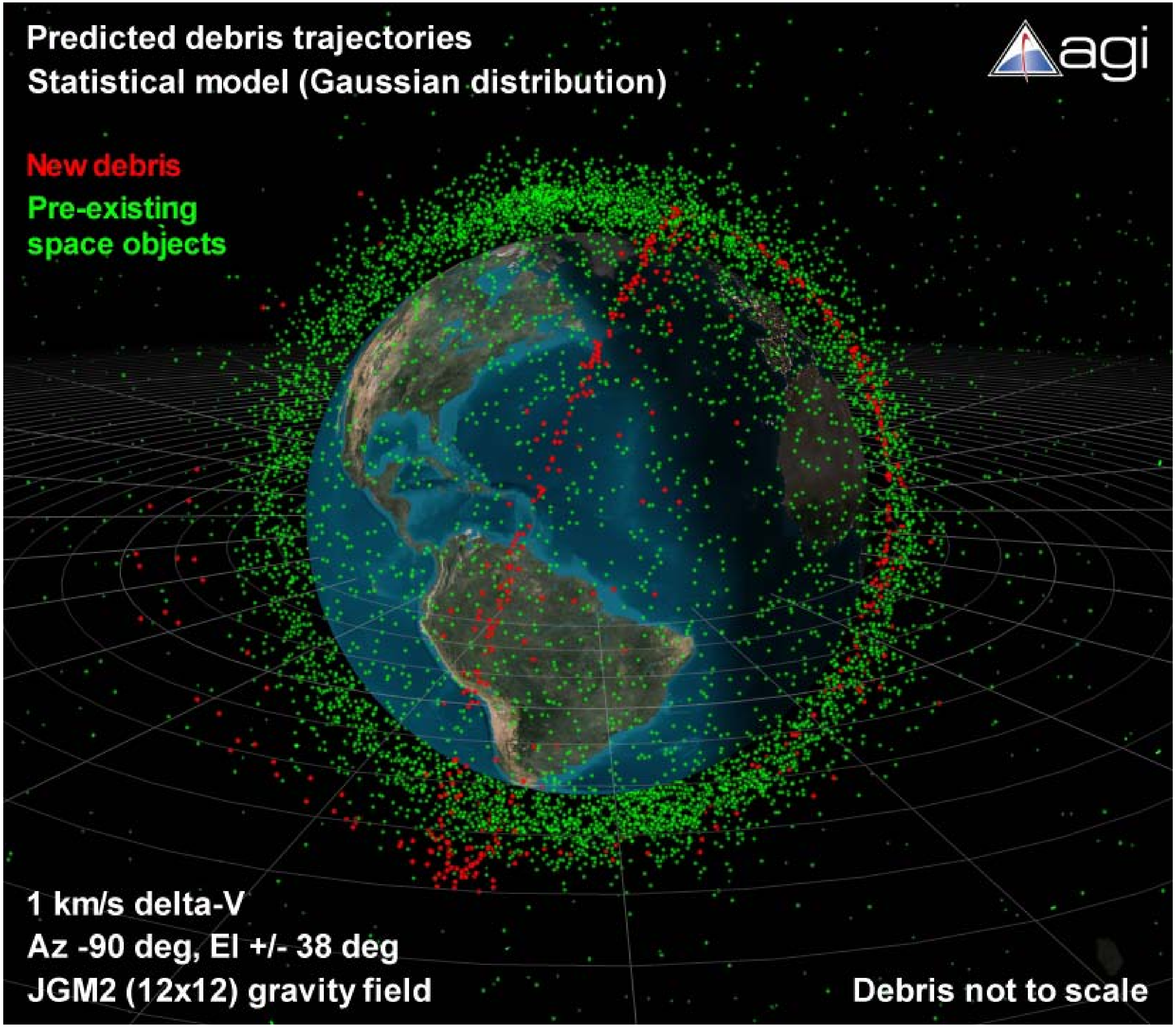
- Decomposition methods involve <u>nonlinear optimization</u> computations
- Spectral unmixing for (space) <u>object material identification with hyperspectral data</u>
  - Project for AFOSR involving UNM (Prasad), Duke (Brady), and WFU (Zhang, Pauca, Ple)
- Analysis of massive global multivariate <u>climate datasets</u> (very brief overview)
  - Project for NASA involving UTK (Berry) and WFU (Zhang, Pauca, Ple)
- Additional comments, problems, ideas

Space Object Identification and Characterization from Spectral Reflectance Data Using NMF/NTF



- More than 30,000 known objects in orbit: various types of military and commercial satellites, rocket bodies, residual parts, and debris (many more objects there with 2007 Chinese and 2008 U.S. satellite kills)
- AFOSR project

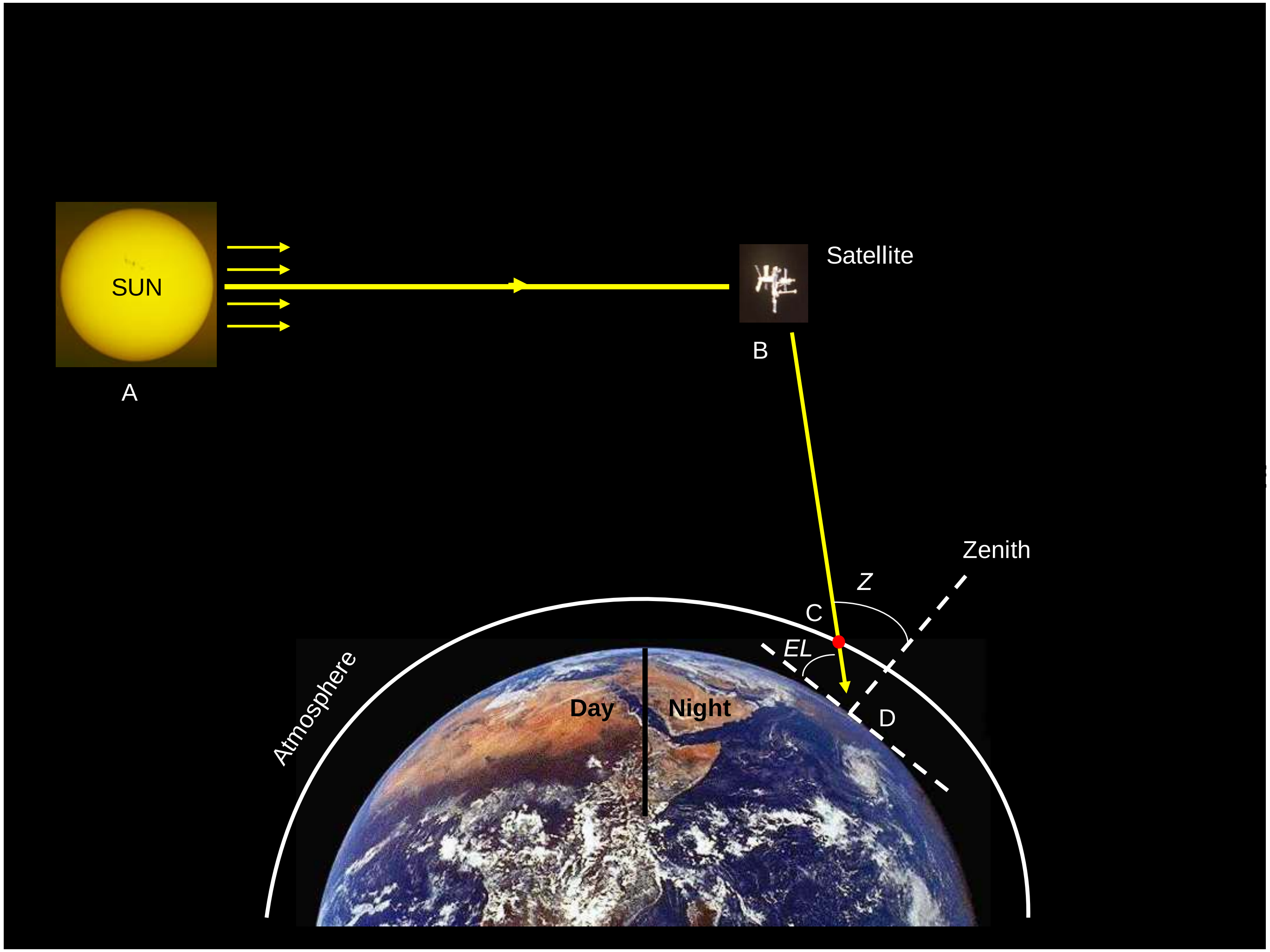
USA-Russian Satellite Collision – Feb 12



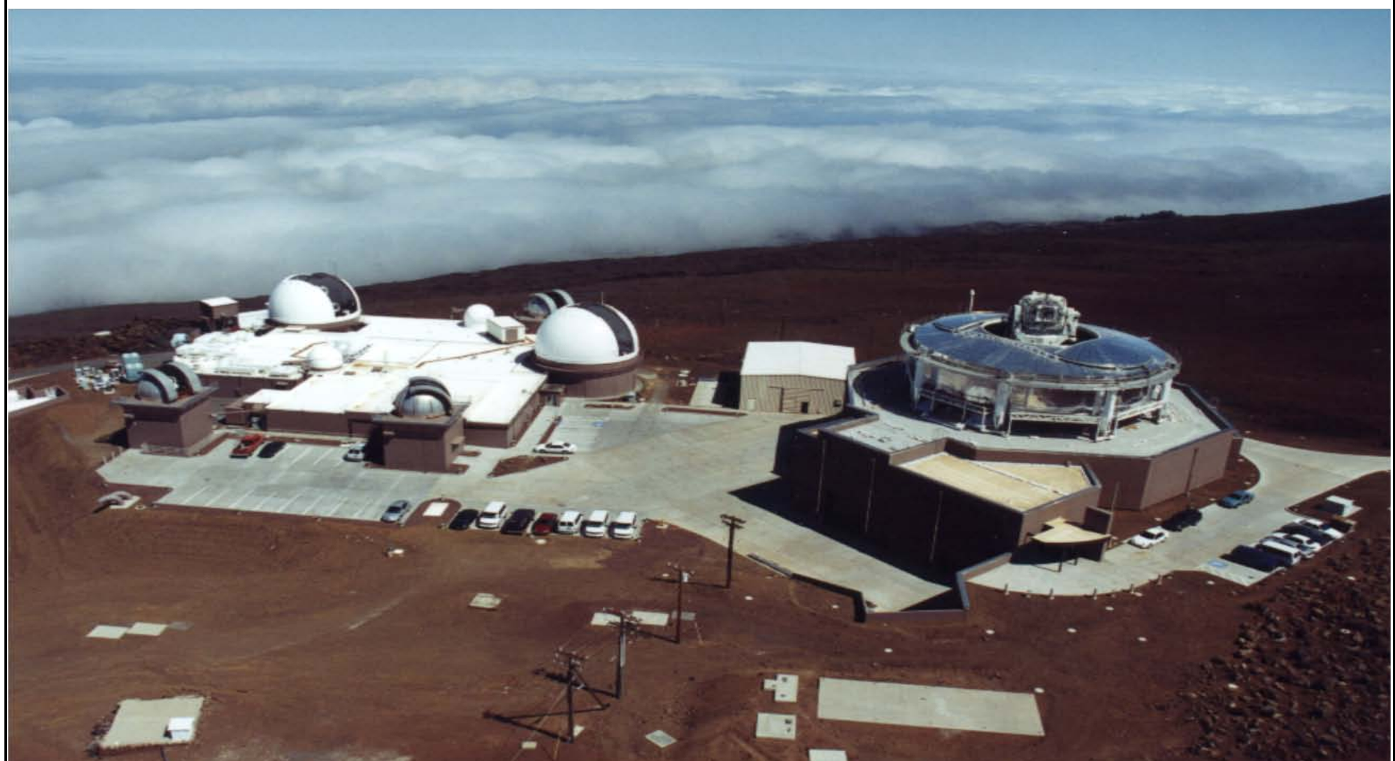
Overview of the SSA Problem

- Space activities require accurate information about orbiting objects for space situational awareness (SSA)
- Many objects are either in:
  - Geosynchronous orbits (about 40,000 KM from earth), or
  - Near-Earth orbits, but too small (e.g., space mines, debris) to be resolved by optical imaging systems
- Objectives: data compression, identification of materials and fractional abundances

The creation and observation of a reflectance spectrum



Maui Space Surveillance Site



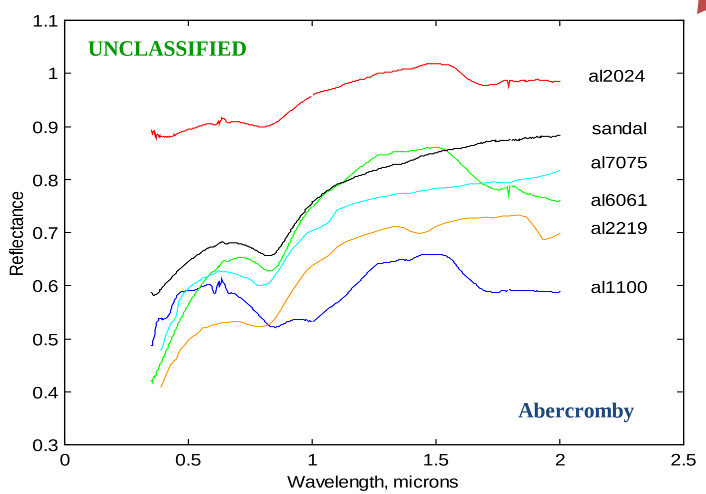
Space Situational Awareness (SSA) by Monitoring Space Objects

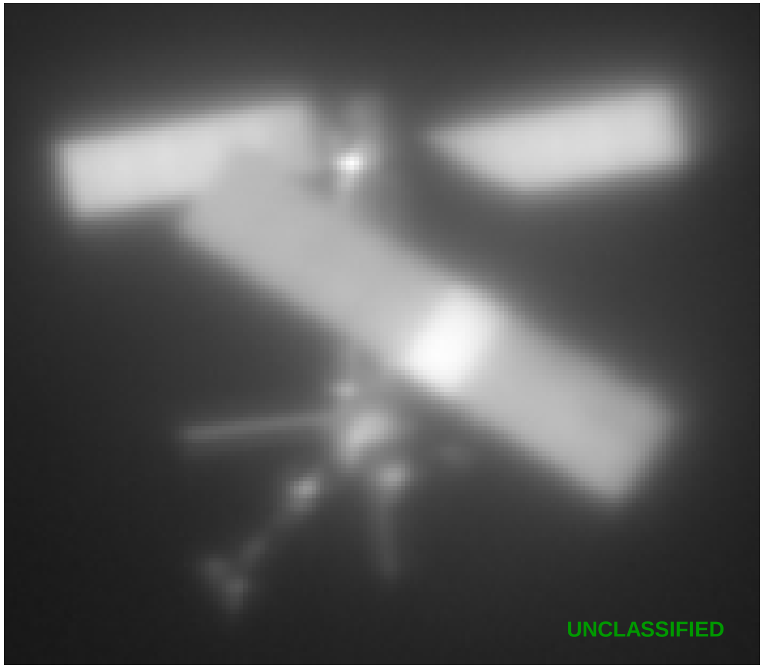
- ‘Listen’ (laser enabled vibrometry)
- ‘Smell’ (chemical sensing with spectrometers)
- ‘Touch’ (scatterometry/polarimetry for surface texture information)
- ‘See’ (by sequential speckle <video> imaging)
- ‘Characterize Materials’ for SOI (spectral imaging) (hyperspectral data mining)

All can involve processing tensor data.

Current DOD/NASA Imaging of Space Objects

- Current “operational” capability for spectral imaging of space objects – imaging and non-imaging
- Panchromatic images, AEOS
- Non-imaging spectra, SPICA



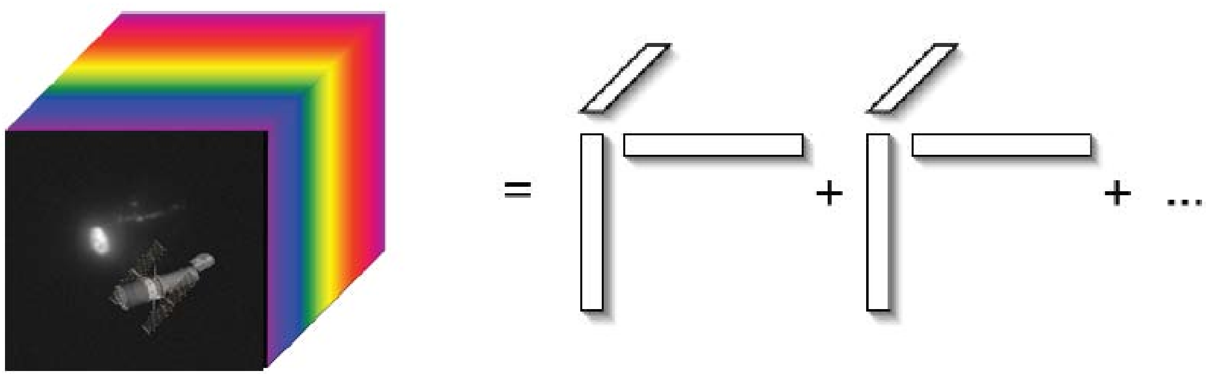


Hyperpectral Imaging – ASIS System on Maui

Nonnegative Tensor Factorization (NTF)

Definition 1. *Let* $\mathcal{T} \in \mathbb{R}^{D_1 \times D_2 \times D_3}$ *be a nonnegative tensor and* $\hat{\mathcal{T}} = \sum_{i=1}^{k} \boldsymbol{x}^{(i)} \circ \boldsymbol{y}^{(i)} \circ \boldsymbol{z}^{(i)}$ *a tensor in CP factored form, where* $\boldsymbol{x}^{(i)} \in \mathbb{R}^{D_1}, \boldsymbol{y}^{(i)} \in \mathbb{R}^{D_2}, \boldsymbol{z}^{(i)} \in \mathbb{R}^{D_3}$. *Then a rank-k nonnegative approximate tensor factorization problem is defined as:*

$$\min_{\tilde{\mathcal{T}}} ||\mathcal{T} - \tilde{\mathcal{T}}||_F^2, \text{ subject to } \tilde{\mathcal{T}} \geq 0. \quad (1)$$



NTF Methods We Used

- ANLS for PARAFAC model
- Projected gradient block coordinate descent method (Lin) with an improved Amijo's rule
- Preprocessing by adaptive re-sampling using total variation minimization criteria (works better than using wavelet basis, in our case)
- Nonlinear optimization methods
- Reference: Journal of Opt. Soc. Amer., Vol. 25, pp. 3001-30012, Dec. 2008. http://www.opticsinfobase.org/josaa/Issue.cfm

Experiments with Hyperspectral Data

- 177 x 193 x 100   3-D model of Hubble satellite
- Assign each pixel a certain spectral signature from lab data supplied by NASA. 8 materials used
- Bands of spectra ranging from .4 μm to 2.5 μm, with 100 evenly distributed spectral values. Re-sampling based on total variation minimization
- Spatial blurring followed by Gaussian and Poisson noise and applied over the spectral bands

Materials Assigned to Pixels

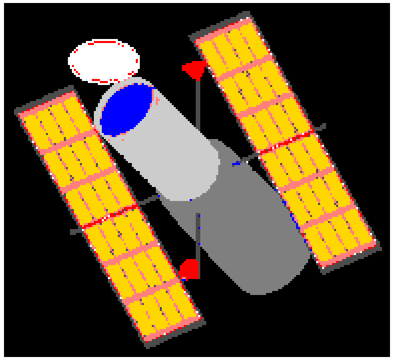
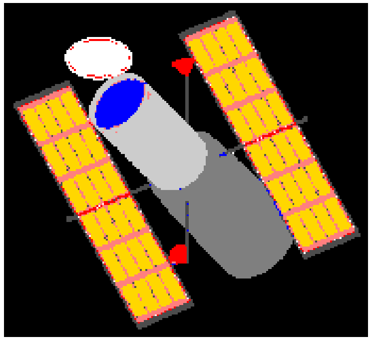


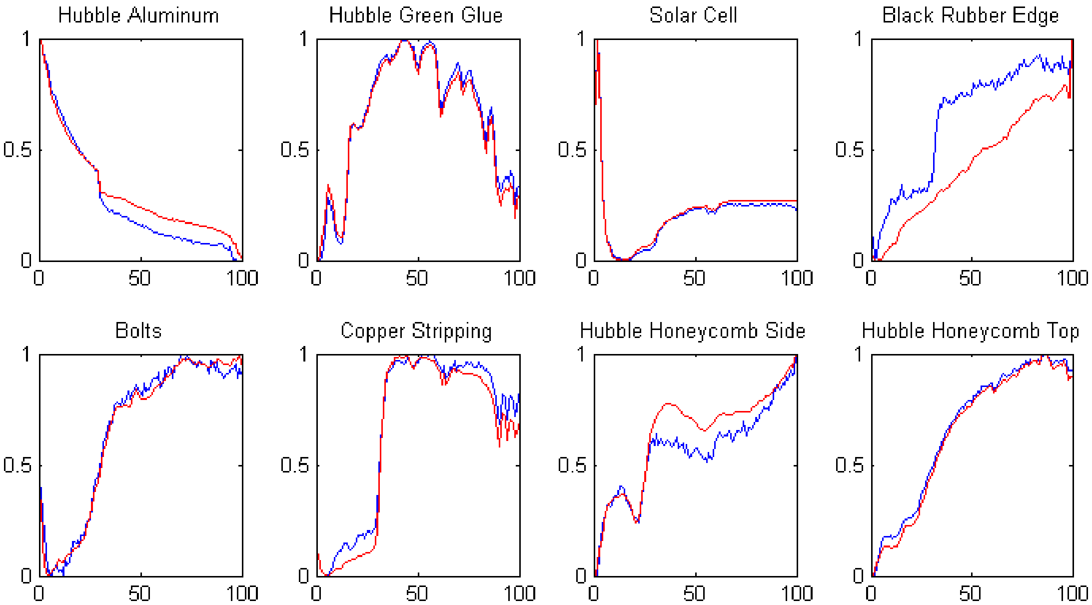
Table 1: Materials, colors and fractional abundances

Material	Color	Fractional Abundance (%)
Hubble Aluminum	light gray	19
Hubble Green Glue	dark gray	12
Hubble Honeycomb Top	white	4
Hubble Honeycomb Side	blue	3
Solar Cell	gold	37
Bolts	red	3
Black Rubber Edge	dark gray	8
Copper Stripping	cyan	13

Material Identification using NTF

- Factors from NTF compared with a material spectral signature library from AFRL/NASA for identification purposes.
- The following graphs show individual material signature comparisons and identified materials spatial support coded in different colors, using four datasets.





Global Climate Changes – NASA Data

- Data mining techniques are commonly used for the discovery of interesting patterns in earth science data.
- Such patterns can help to both understand and predict changes in climate and the global carbon cycle.
- Regions of earth partitioned into sub-regions described land- or sea-based parameters. Patterns within these subregions mined to reveal both spatial and temporal autocorrelation.
- We identify regions (or clusters) of the earth which have similar short- or long-term characteristics.
- Earth scientists are interested in patterns that reflect deviations from normal seasonal variations (e.g., El Niño and La Niña).
- Interpreting these patterns can facilitate a better understanding of biosphere processes. Can effect policy decisions at a global scale.

A Parallel Algorithm for Approximate Nonnegative Tensor Factorization

Peter Zhang[1], Michael W. Berry[2], Todd Torgersen[3] and Bob Plemmons[3,4]

Department of Biostatistical Sciences, Wake Forest University Health Sciences

Department of Electrical Engineering and Computer Science, University of Tennessee

Department of Computer Science, Wake Forest University

Department of Mathematics, Wake Forest University

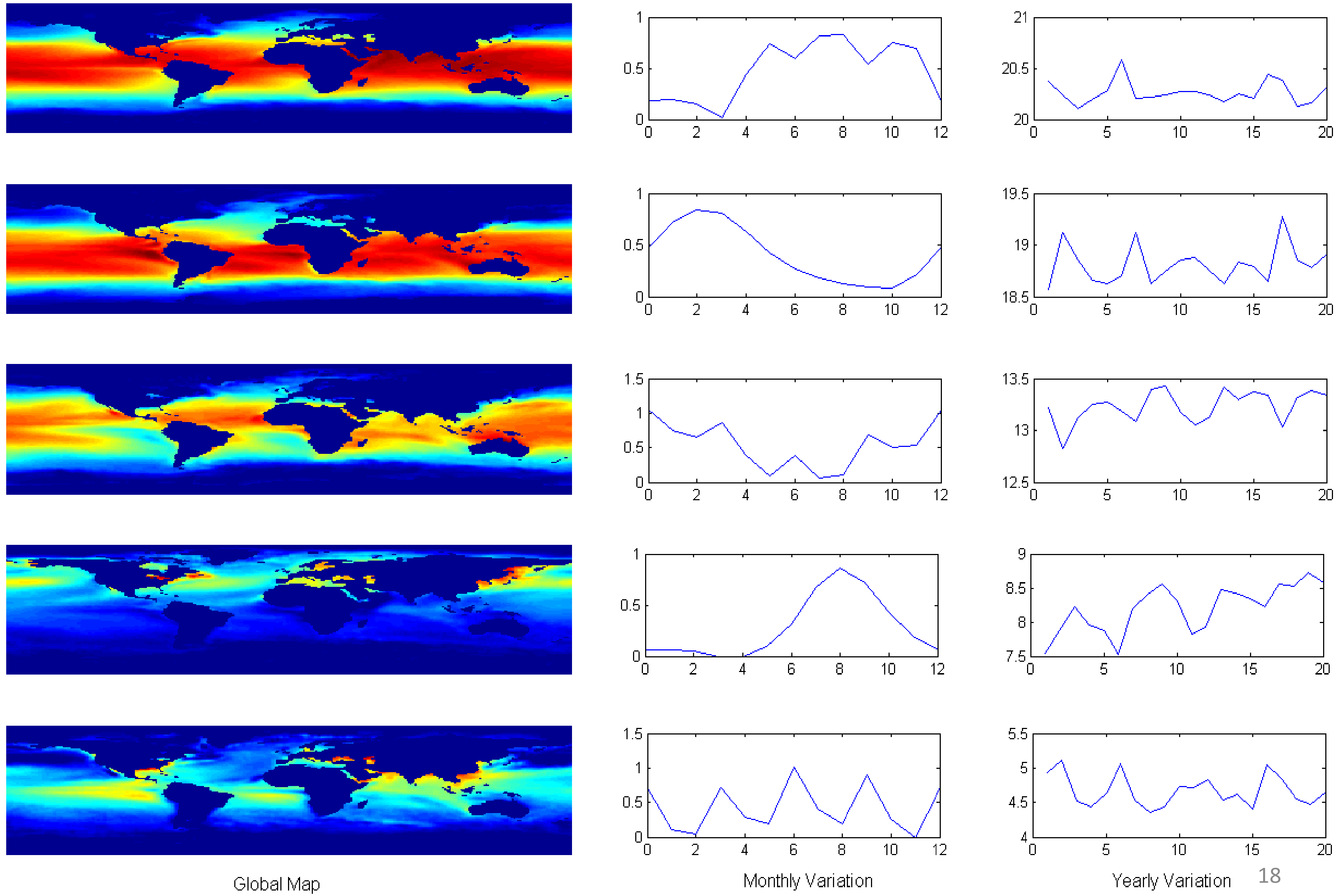
March 4, 2009

Variables Being Considered in Study

- Sea surface temperature ←
- Land surface temperature
- precipitation
- Normalized difference vegetation index
- Geopotential elevation for 500 mb pressure
- Geopotential elevation for 1000 mb pressure

Study spatial patterns and associated time indices

Sea Surface Temperature Change Patterns Obtained using NTF
(Sample slide from Zhang's Talk in Carla and Misha's Mini at SIAM-CSE)



Array Imaging Application

(Practical Enhanced Resolution Integrated Optical Digital Imaging Camera)

PERIODIC Project

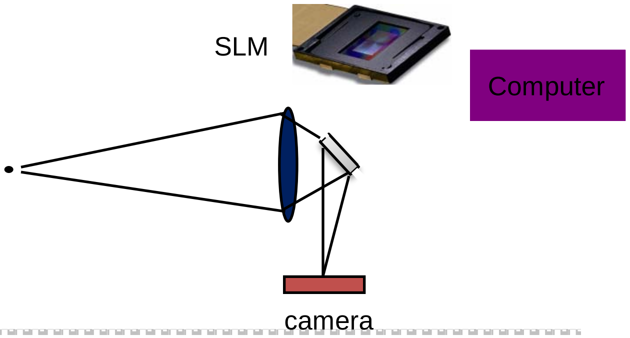
Demonstration at IARPA
20 February 2009

Prototype Camera Systems

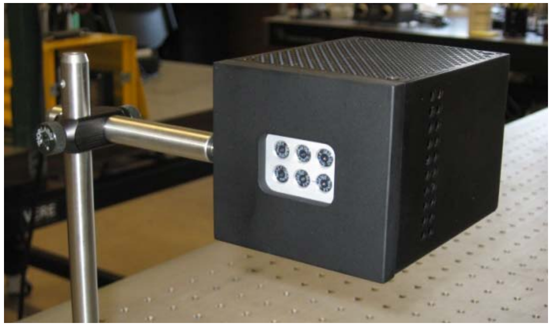
Spectral Diversity
multi-spectral prototype

Five
prototypes

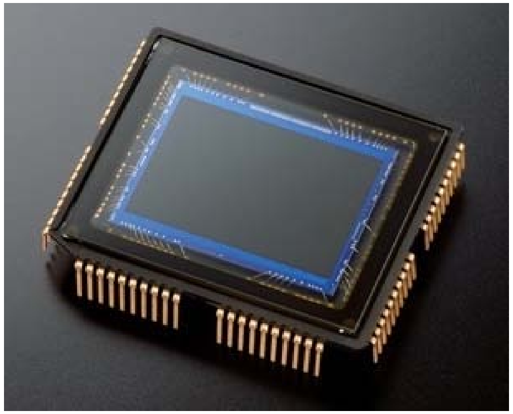
PSF engineering
"reconfigurable phase diversity"



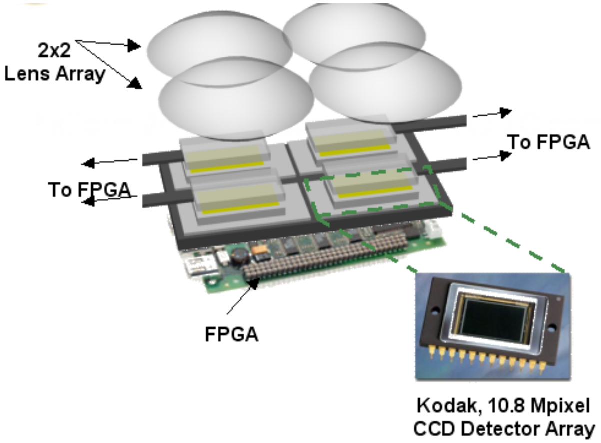
Polarization diversity
full stokes polarimetric imager



Temporal diversity
Short range "lock in" imager



"Brains on Board" imager



PERIODIC Array Imaging Objectives

- Balance processing capabilities imaging systems through concurrent design and joint optimization of all elements
- Achieve a particular imaging objective with minimal resources
- Seamless integration of sensing and processing algorithms using **multi-way arrays (tensors)**
- Our approach: design multi-aperture multi-diversity compact imaging systems

22

Sensing/Reconstruction Approaches

- Analyze lock-in sensing with modulated/gated illumination – "temporal diversity"
- Use of reconfigurable high-res SLM testbed to implement multiple diversities – how to optimize them for different classes of scenes?
- Explore theoretically a number of applications – fingerprint/hand/skin-based biometrics, IED detection
- Nonnegative Tensor Factorization (NTF) vs. physically motivated compressive reconstruction approaches, e.g., those based on non-separable geometric primitives, wavelets, etc.
- Novel data-fusion strategies for multi-diversity data

Array Based Digital Super Resolution Hardware Implementation

	Estimated Performance	Development Cost	Power (FLOPS/watt)	Flexibility
GPUs	Very High	Very High (**GPU development systems for embedded applications are not yet available)	Very Low	Very High
FPGAs	High	Medium/Low	High	High
ASICs	Very High	Very High	High	Low
DSP	High	Low	Medium	High
Multicore MPs	Medium	Very Low	Low	Very High
CELL	High	Medium	Very Low	Medium

Comments from Jack Dongarra (HPC WS at WFU, Feb 12-13)

- For the last few decades or more, the **research investment** strategy has been overwhelmingly **biased in favor of hardware**.
- This **strategy needs to be rebalanced**
  - The return on investment is more favorable to software.
  - Hardware has a half-life measured in years, while software has a half-life measured in decades.
- **No Moore's Law for software, algorithms and applications**

Final Items

- Andrzej Cichocki, et al. (Tokyo): **Book** (2009) – "Nonnegative Matrix and Tensor Factorizations, With Applications to Exploratory Multiway Data Analysis and Blind Source Separation"

- Problems: Re-sampling, deblurring and/or denoising tensor arrays of scientific data before analysis with NTF
  - Compressed sensing, coded apertures, massive multi-dimensional image-related datasets (Workshop 02/25-26/2009 at Duke)