# Astronomy to medical imaging: Blackford Analysis

Alan Heavens

Institute for Astronomy, University of Edinburgh

with Ben Panter, Rob Tweedie, Mark Bastin, Will Hossack, Keith McKellar, Trevor Whittley





ROE Workshop 13th October 2010

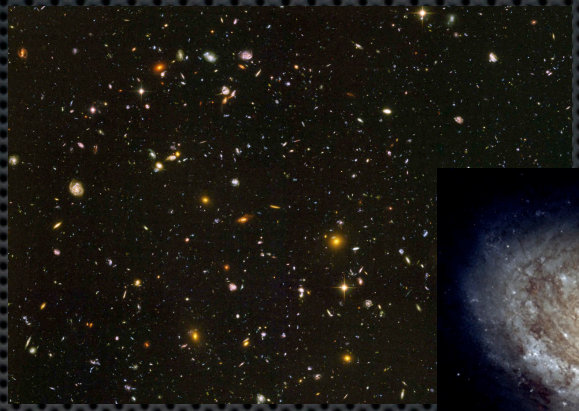
# Blackford Analysis

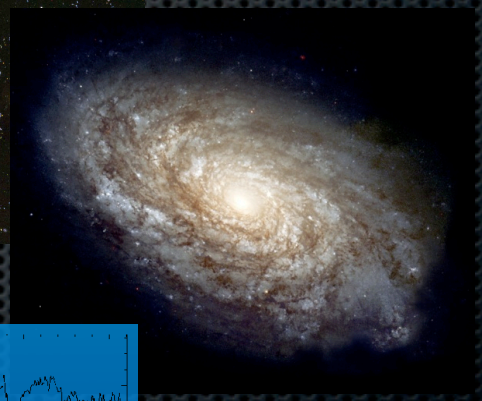
- Blackford Analysis Ltd. is a spinout company from the University of Edinburgh that builds processing solutions using MOPED and other technologies

- MOPED is a patented technology which speeds up computation involving large datasets

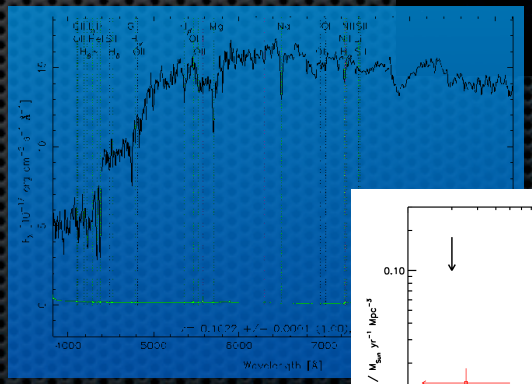

# Astrophysics Data Deluge

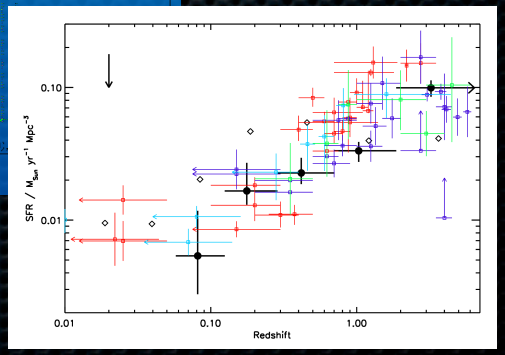
- Astronomy Datasets
- 1.4bn pixels every 30 sec





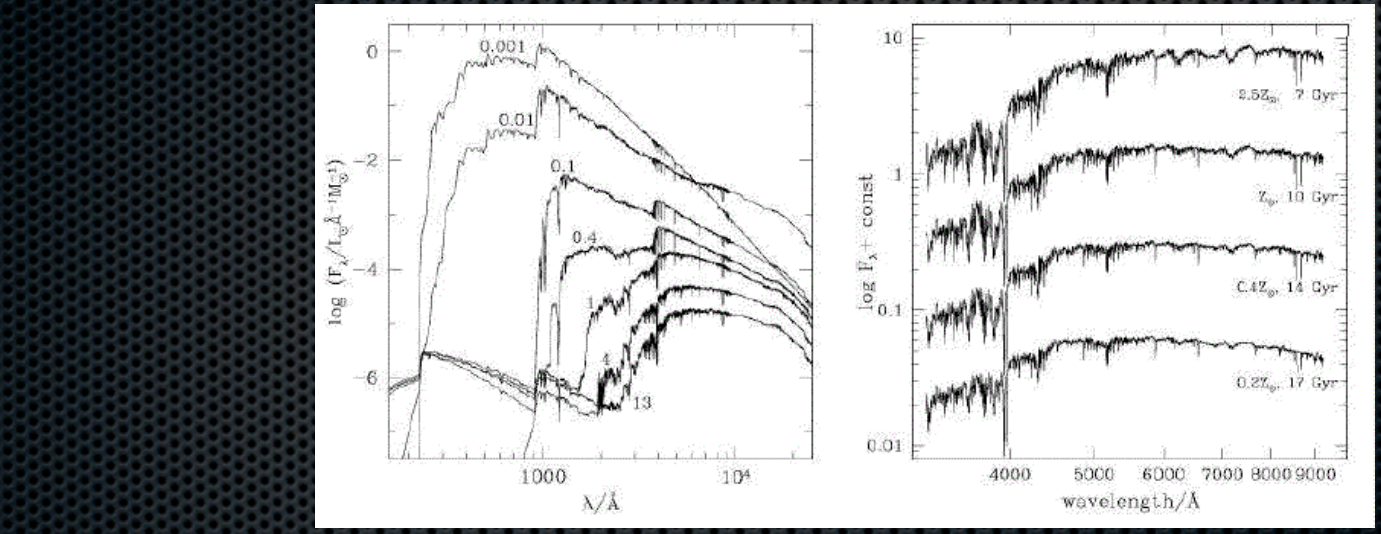
- > 1 TB / night image data
- Multi-object spectrographs
- Analysis Requires:
- Accurate answers to fitting of complex models





# Examples

- Galaxy spectra: flux measurements are sum total of starlight from stars of given **age** (simplest)



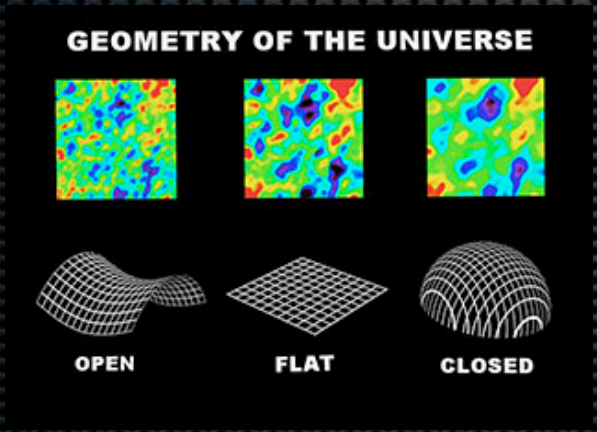

- 2 parameters = age, mass

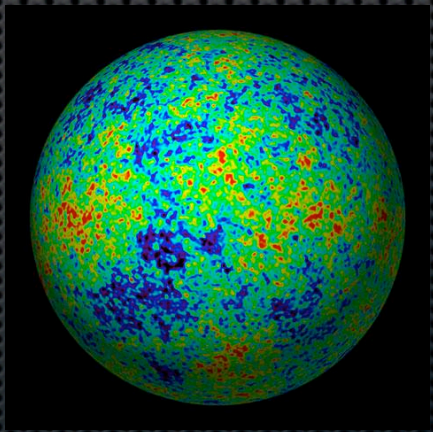
# Inverse problems

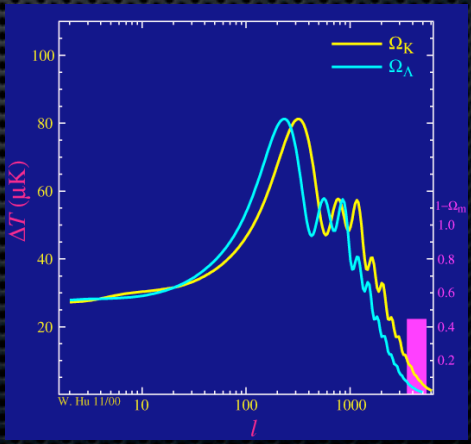
- **Parameter Estimation:** given some data, and a model, what are the most likely values of the parameters, and what are there errors?
- Astronomical problems often have
  - LARGE DATASETS and
  - (fairly) LARGE PARAMETER SPACES

# Example: cosmology

- Model:  Big Bang theory
- Parameters: Expansion rate, density of ordinary matter, density of dark matter, dark energy content… (around 15 parameters)



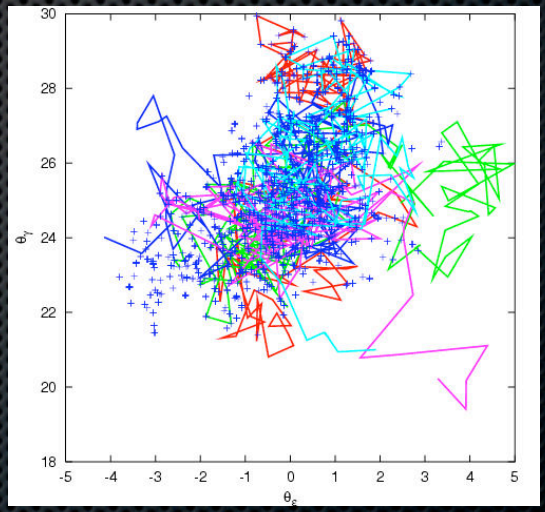






Scale

# Dealing with large parameter spaces

- Don’t explore it all - generate a chain of random points in parameter space
- Most common technique is MCMC (Markov Chain Monte Carlo)




Variants: Hamiltonian Monte Carlo, Nested Sampling,Gibbs Sampling...

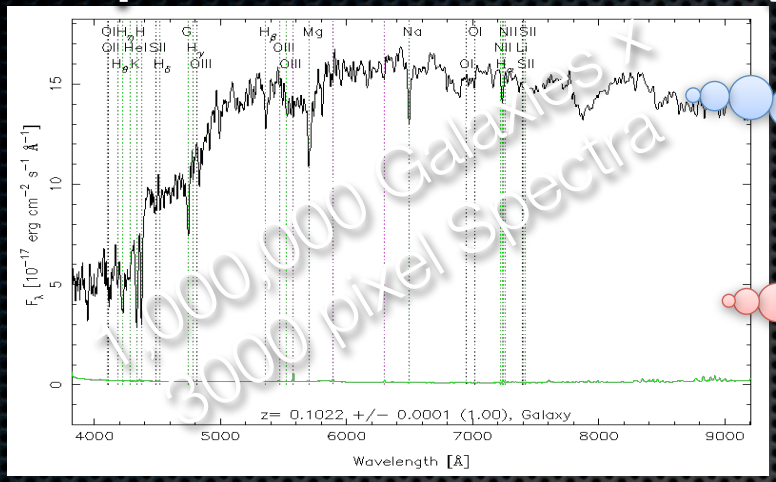
# Large data sets

- What scope is there to reduce the size of the dataset?
- Why?  Faster analysis
- Can we do this without losing accuracy?
- Depending on where the information is coming from, often the answer is yes.

# MOPED weighting vectors

- MOPED automatically weights each data point in an optimum way, creating a vastly smaller dataset.
- In many cases, *the errors from the compressed dataset are no larger than those from the entire dataset*
- It is NOT obvious that this is possible

# Spectral Analysis




12 populations, each with variable metal content = 24 parameters

Dust Screen affects all components = 25th parameter

Spectrum of 3000 points is reduced to 25 numbers

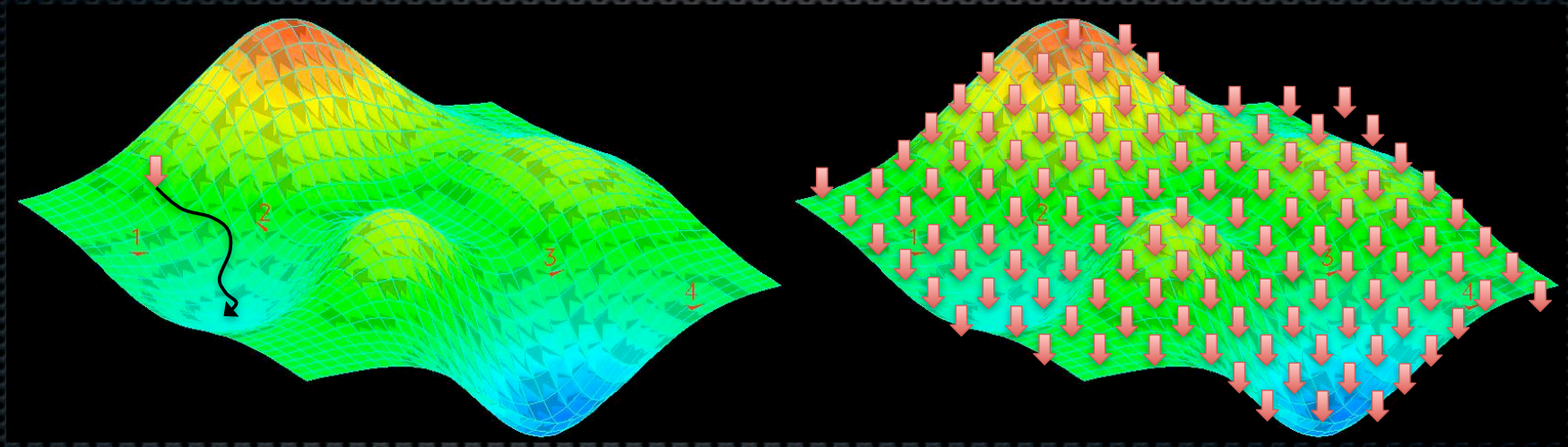
Time: 10 hours -> 1 minute per spectrum – **600x** faster with MOPED

Allowed 1M spectra to be investigated

Nature 428, 625-627 (8 April 2004)

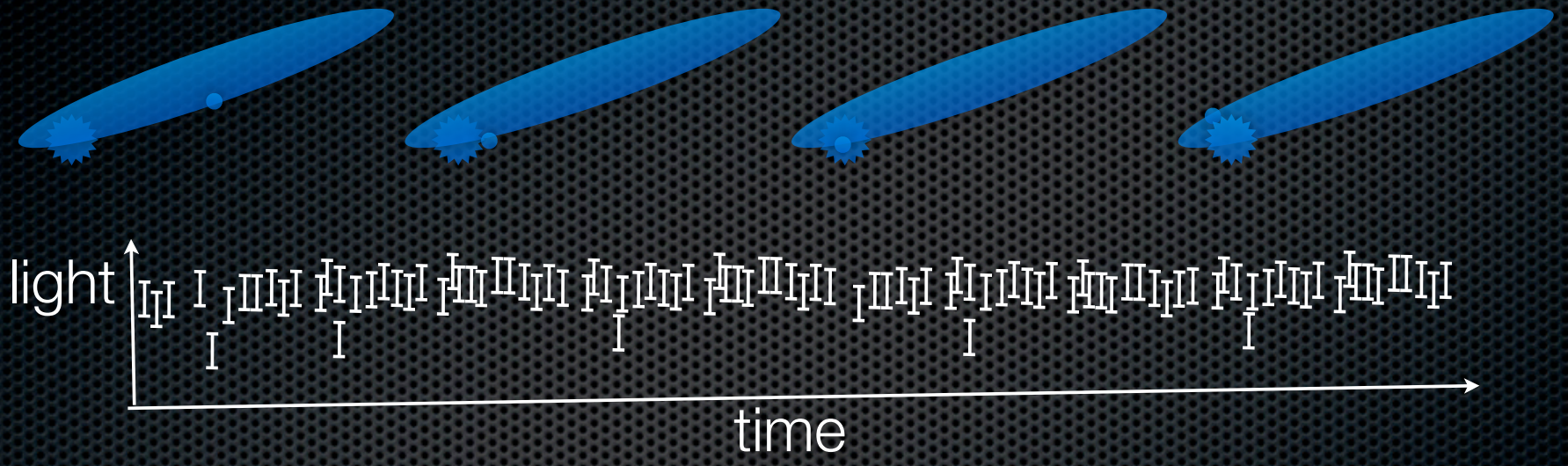
# How does MOPED work?

Carefully constructed compression step allows rapid exploration of problem.




12D registration example: **1,000,000** combinations / second

# Planets Around Distant Stars

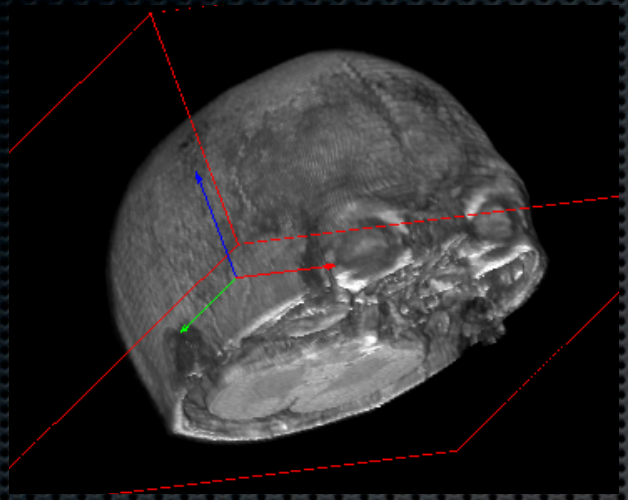


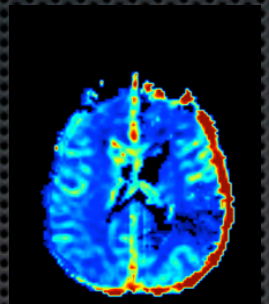

Brute force: 3 years CPU

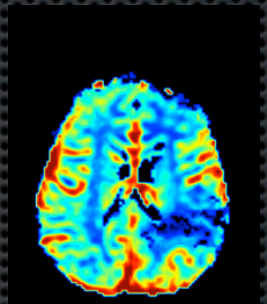
MOPED: 1 day CPU (1000x speed up)

$10^{10}$ combinations

# Medical imaging: registration

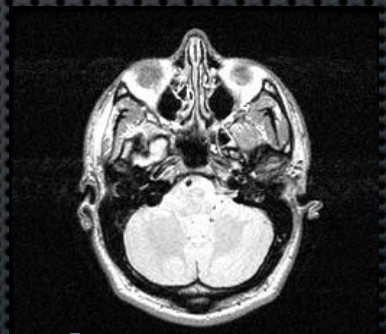


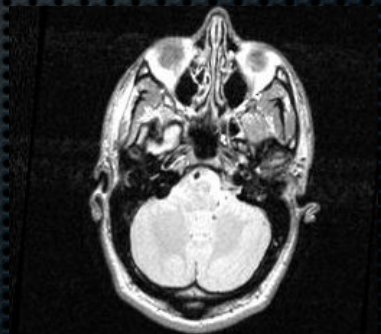


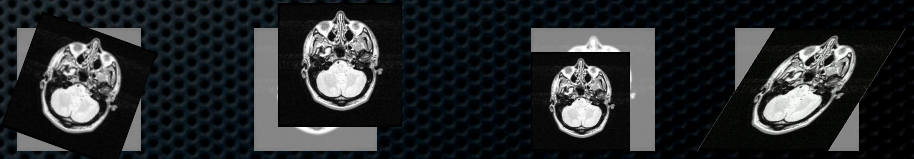

Stroke lesion

Image Distortions

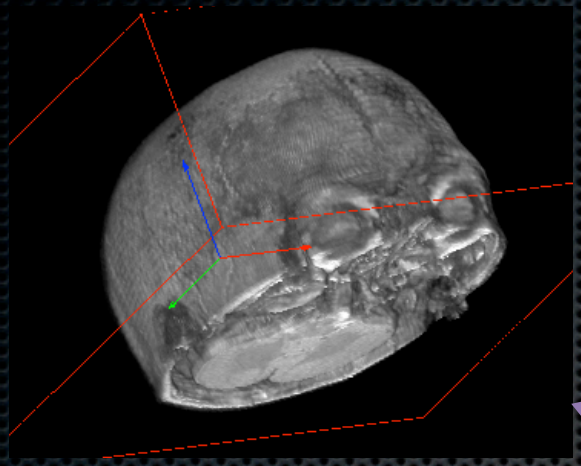




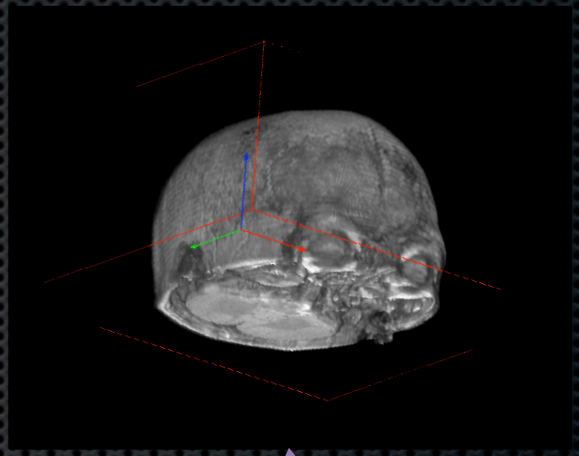
- MRI scans:
  512x512x100 voxels
  **$2.6 \times 10^7$ data**
- Affine distortions: **12 parameters**



# Registration in MR / CT (implemented)

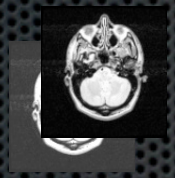


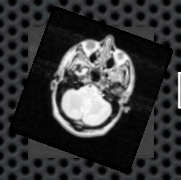
**Registration takes 1 second, down from 10 minutes without MOPED**



**Transform:**

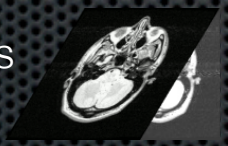
[x, y, z] translations

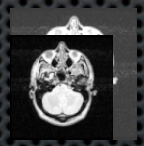




[xy, yz, zx] rotations

[x, y, z] shears





[x, y, z] scales

2 x [512 x 512 x 100] volumes

**12D** => massive number of possible solutions

**52 Million** Voxels

# What kind of problem is best?

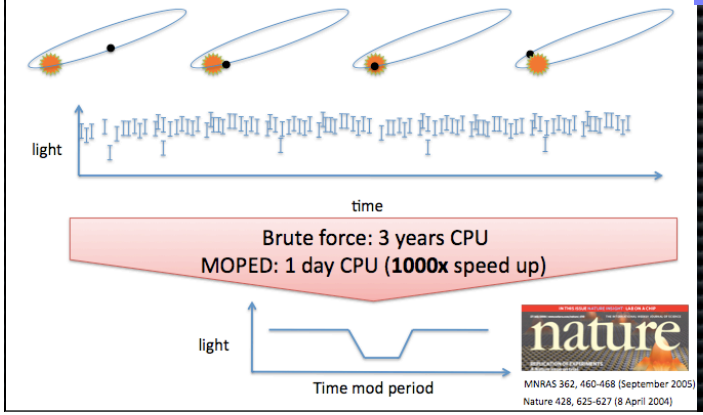
**Data**

**MOPED**

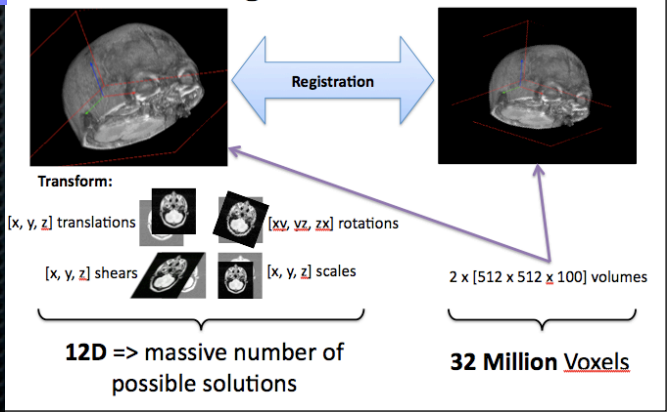
**Model**

Large dataset
Currently too slow
Fitting a model

## Planets Around Distant Stars

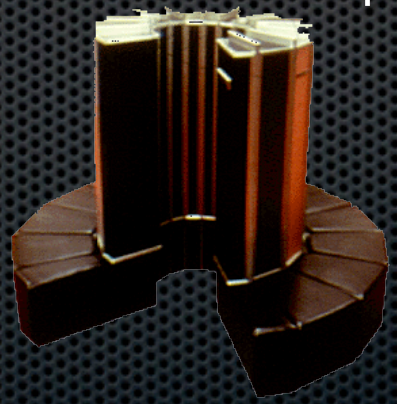



## Linear Registration in MR / CT



# Where could MOPED be applied?

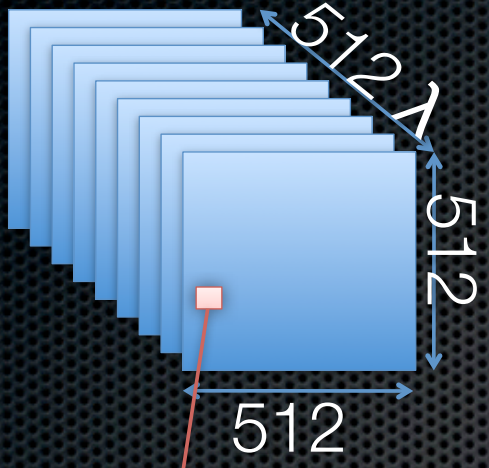
Faster Processing?

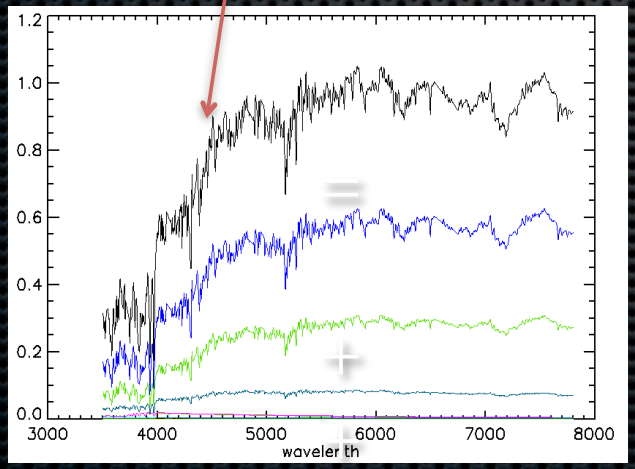


Smaller Hardware?

Bigger Datasets?

# Hyperspectral Imaging




Acquire 512x512 spectra from hyper-spectral camera, each with 512 wavelengths. (Data alone is **134Mpix**)

Use MOPED **to fit spectral features on camera**, returning parametric fit for each pixel, vastly reducing data size and allowing further analysis

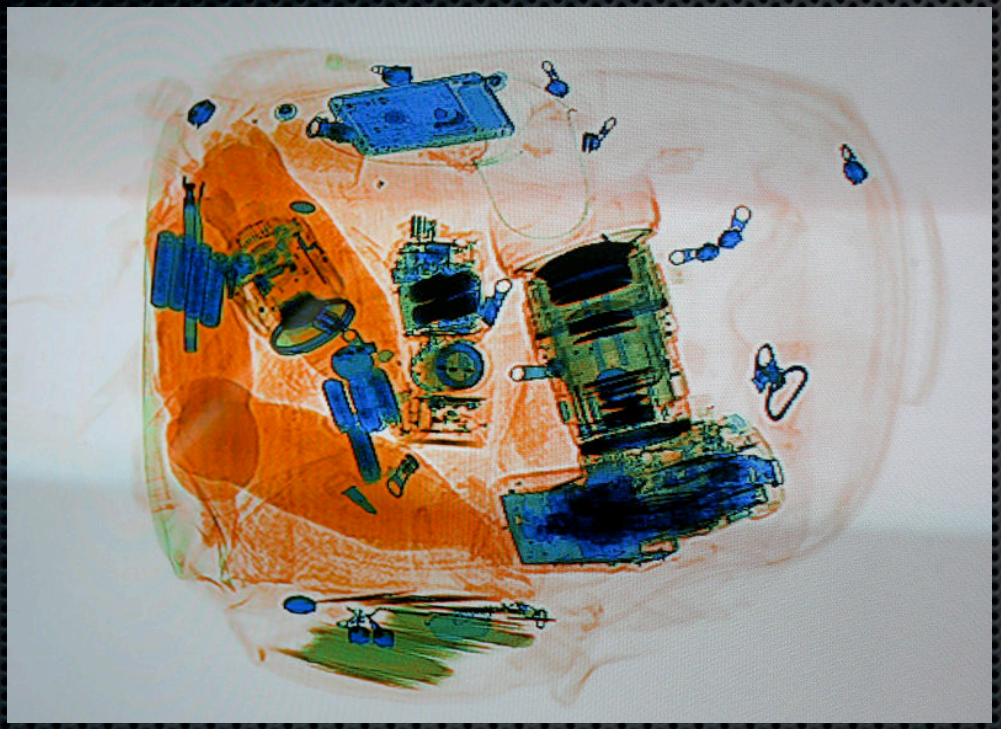
Problem: Spectral lines alone cannot distinguish
Answer: Fit whole spectrum?

-> Distinguish between chemical explosives and fuel

Determine spectral elements of interest – foliage, Semtex, chemical explosive, fuel, earth, rock & brick – MOPED filters to combine slices to reveal presence of spectra. Potentially eliminate real world lighting variations (reflections etc.) by incorporating as nuisance parameters

Original Datacube**: 134Mpix**
Processed Data**: 0.26Mpix** / parameter
**10** 'filters' – **50x** fewer data to transmit
Processing Speedup: **$512/N_{parameters}$**

# Security Scanning



X-ray / THz scan identifies 30 objects in luggage

MOPED used to immediately compare each object to library of known threats in every orientation

Go/no-go advice given, or alerts operator for follow up physical examination

Multi-frequency scans allow spectral analysis of liquids
MOPED enables fast spectral fitting (see prev. slide)

# Summary


- Blackford Analysis is using solutions designed for dealing with large Astronomical Datasets for rapid and accurate analysis of data
- Applications in medical imaging, security etc. - anywhere where speed and accuracy are important
- Arose from blue-skies research

Contact: Dr. Ben Panter

ben@blackfordanalysis.com www.BlackfordAnalysis.com