Introduction to Remote Sensing Image Registration

Jacqueline Le Moigne,
NASA Goddard Space Flight Center

IGARSS 2017

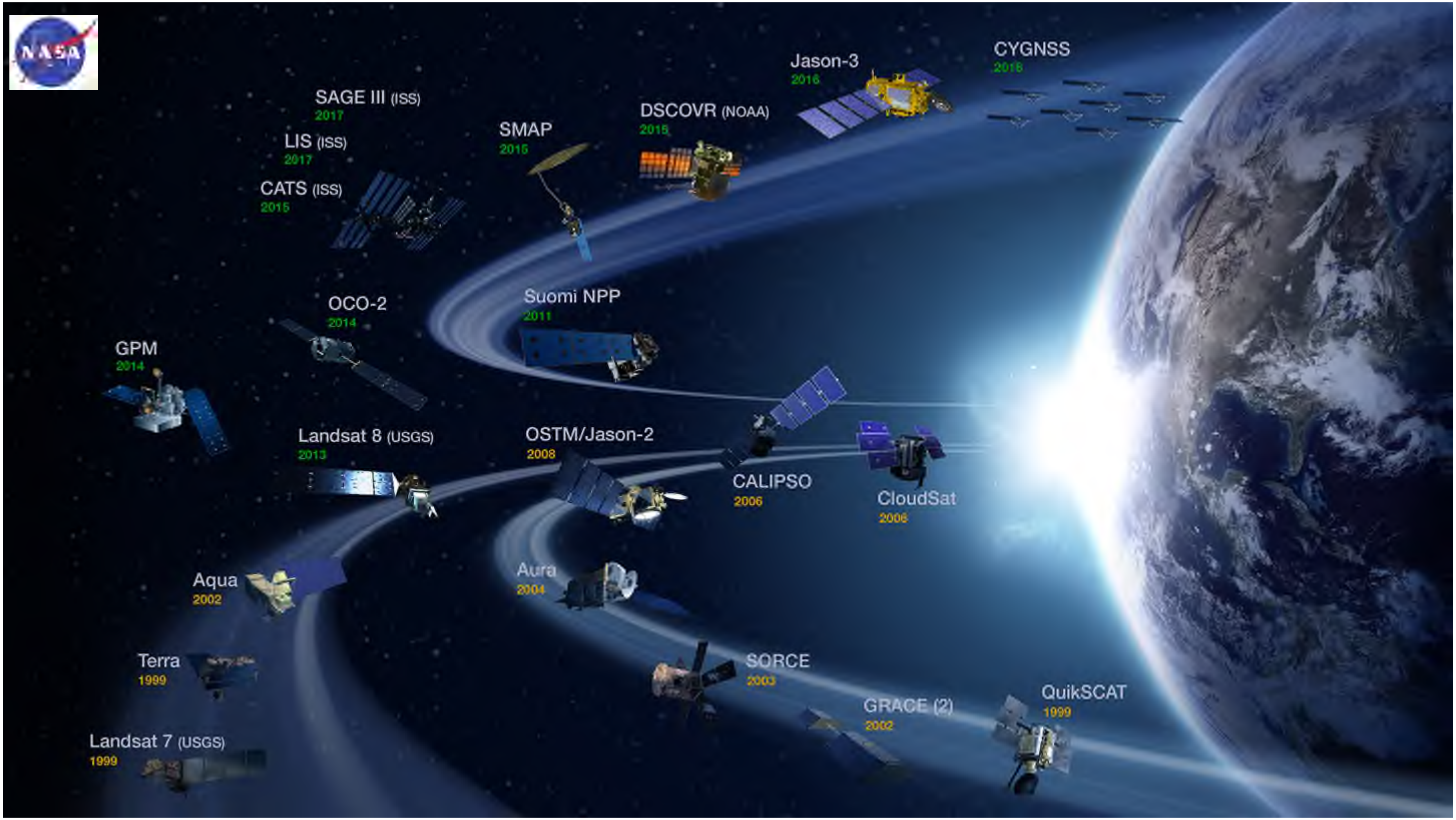
Problem Description

- Earth Science studies such as:
  - Climate change over multiple time scales
  - Predicting crop production
  - Monitoring land resources
  - Understanding the impact of human activity on major Earth ecosystems

- Addressed by using global and repetitive measurements provided by a wide variety of satellite remote sensing systems
  - Multiple-time or simultaneous observations of the same Earth features by different sensors
  - Global measurements with remote sensing systems
  - Complemented by regional and local measurements using ground and airborne sensors
  - Addressed by using global and repetitive measurements provided by a wide variety of satellite remote sensing systems

- Need to correlate and integrate all these complementary data

NASA Earth Global Measurements



Example of International Measurements

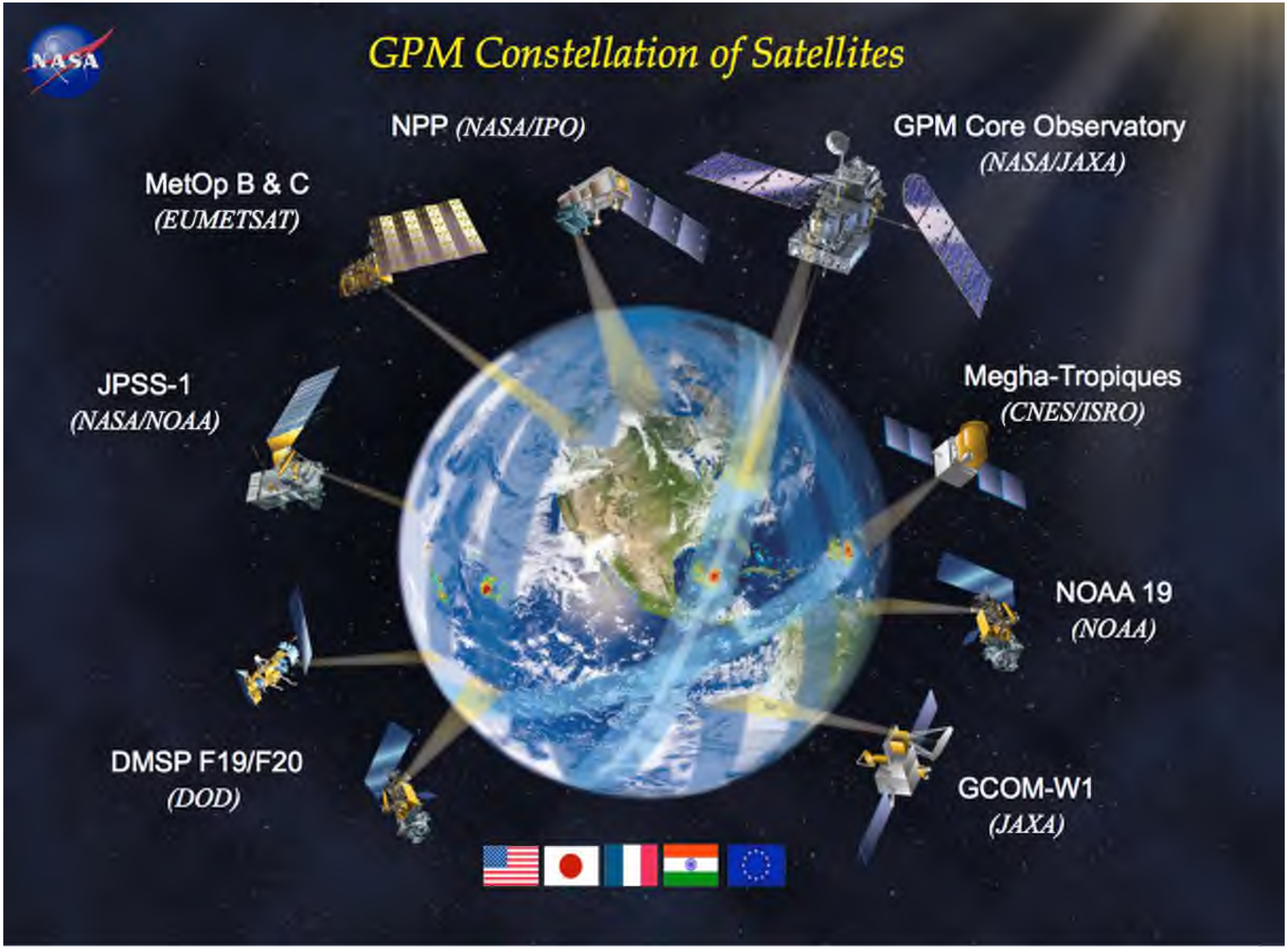
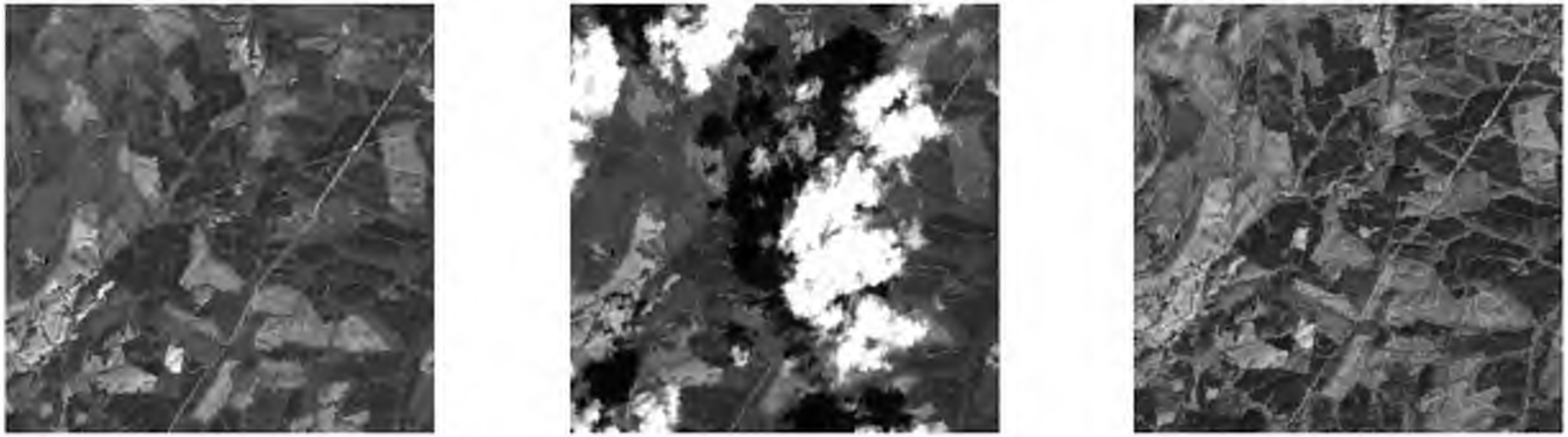


Image Registration Challenges

- Remote Sensing vs. Medical or Other Imagery
  - Variety in the types of sensor data and the conditions of data acquisition
  - Size of the data
  - Lack of a known image model
  - Lack of well-distributed "fiducial points" resulting in lack of algorithms validation
- Navigation Error
- Atmospheric and Cloud Interactions



Three Landsat images over Virginia acquired in August, October, and November 1999
(Courtesy: Jeffrey Masek, NASA Goddard Space Flight Center)

Image Registration Challenges

Atmospheric and Cloud Interactions

Atmospheric and Cloud Interactions

Baja Peninsula, California; 4 different times of the day (GOES-8)

(Reproduced from Le Moigne & Eastman, 2005)

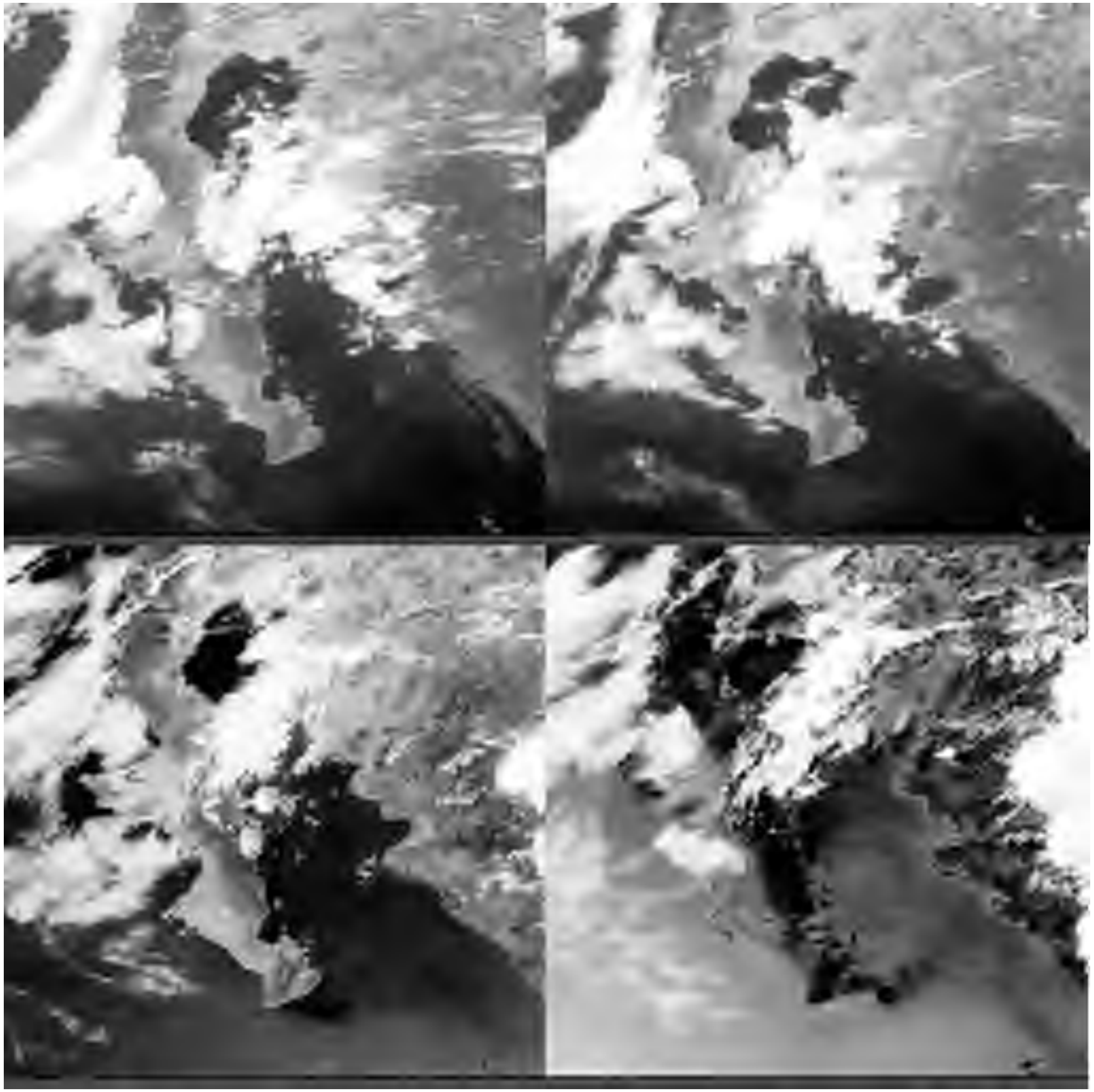
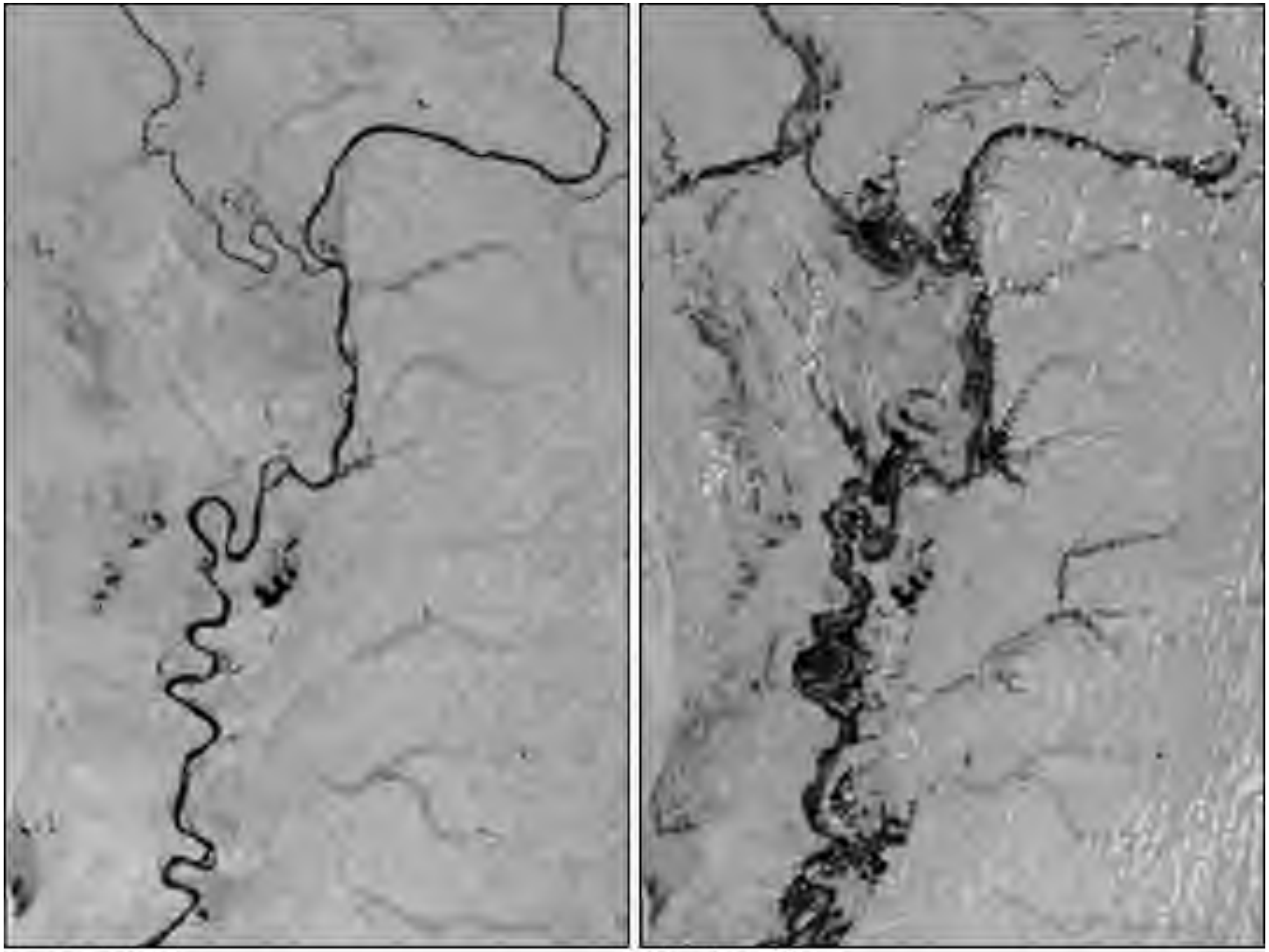


Image Registration Challenges

Multi-Temporal

Mississippi and Ohio Rivers before & after Flood of Spring 2002 (Terra/MODIS)



April 25, 2002 | May 18, 2002

What is Image Registration?

- Image Registration/Feature-Based Precision Correction vs. Navigation or Model-Based Systematic Correction
  1. Orbital, Attitude, Platform/Sensor Geometric Relationship, Sensor Characteristics, Earth Model, etc.
  2. Navigation within a Few Pixels Accuracy
  3. Image Registration Using Selected Features (or Control Points) to Refine Geo-Location Accuracy
- Mathematical Framework
  - I1(x,y) and I2(x,y): images or image/map
    - ✓ find the mapping (f,g) which transforms I1 into I2: **I2(x,y) = g(I1(fx(x,y),fy(x,y))**
      a. f : spatial mapping
      b. g: radiometric mapping
  - Spatial Transformations "f"
    - ✓ Translation, Rigid, Affine, Projective, Perspective, Polynomial, …
  - Radiometric Transformations "g" (Resampling)
    - ✓ Nearest Neighbor, Bilinear, Cubic Convolution, ...
- Algorithmic Framework (Brown, 1992)
  1. **Search Space** of potential transformations
  2. **Feature Space** of information extracted from the 2 datasets
  3. **Similarity Metric** used to match the 2 sets of features
  4. **Search Strategy** to find the optimal transformation
  5. **Resampling Method** to create the corrected image
  6. **Validation Method** to evaluate the accuracy of the registration

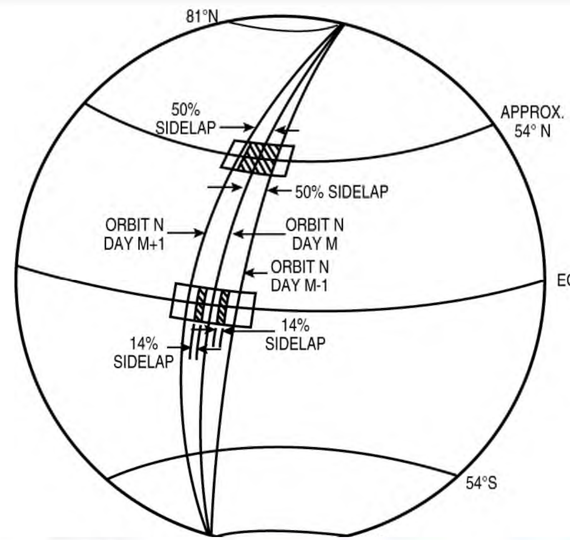
What is Image Registration?

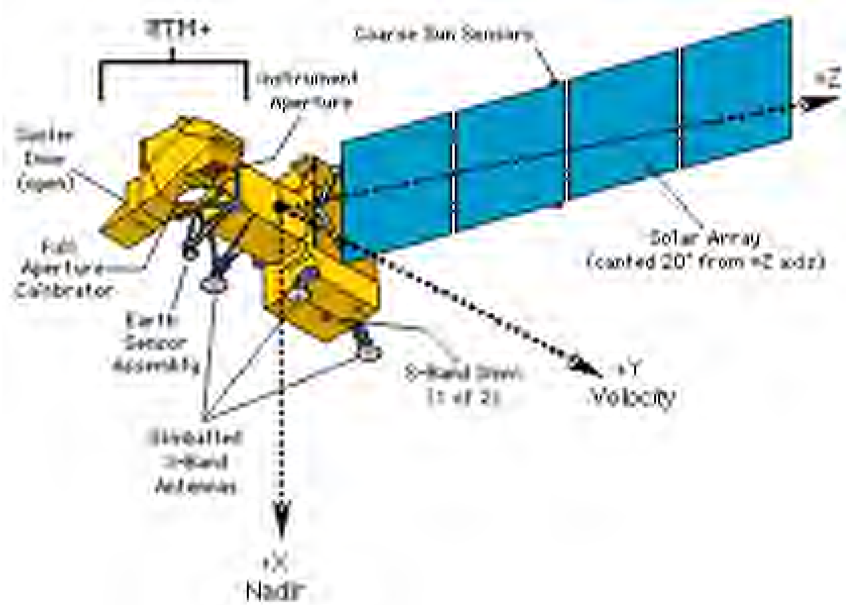
- Image Registration/Feature-Based Precision Correction vs. Navigation or Model-Based Systematic Correction
  1. Orbital, Attitude, Platform/Sensor Geometric Relationship, Sensor Characteristics, Earth Model, etc.
  2. Navigation within a Few Pixels Accuracy
  3. Image Registration Using Selected Features (or Control Points) to Refine Geo-Location Accuracy
- Mathematical Framework
  - I1(x,y) and I2(x,y): images or image/map
    - ✓ find the mapping (f,g) which transforms I1 into I2: **I2(x,y) = g(I1(fx(x,y),fy(x,y))**
      a. f : spatial mapping
      b. g: radiometric mapping
  - Spatial Transformations "f"
    - ✓ Translation, Rigid, Affine, Projective, Perspective, Polynomial, …
  - Radiometric Transformations "g" (Resampling)
    - ✓ Nearest Neighbor, Bilinear, Cubic Convolution, ...
- Algorithmic Framework (Brown, 1992)
  1. **Search Space** of potential transformations
  2. **Feature Space** of information extracted from the 2 datasets
  3. **Similarity Metric** used to match the 2 sets of features
  4. **Search Strategy** to find the optimal transformation
  5. **Resampling Method** to create the corrected image
  6. **Validation Method** to evaluate the accuracy of the registration

Operational Solutions

The Landsat ETM+ Example

- Sensor Knowledge
  - Sensor geometry
  - Sensor to platform
  - Orbit
  - Terrain data (DEM)
  - Radiometric model
- Geodetic accuracy
  - Database of GCPs derived from USGS data
  - Normalized correlation
  - Updates navigation models
  - Results: RMSE ~54m
- Band-to-band registration
  - Selected tie-points in high-freq. arid regions
  - Normalized correlation
  - Subpixel by second order fit to 3x3 neighborhood
  - Result: 0.1 to 0.2 subpixel





Operational Solutions

Normalized Cross-Correlation (NCC) Often Used

Instrument	Satellite	Resolution	Similarity	Subpixel
ASTER	Terra	15m-90m	NCC w/ DEM	Fit to surface
GOES	GOES I-M	1km-8km	NCC w/ vector coastlines	Bi-section search
MISR	Terra	275m	NCC w/ DEM	Least squares
MODIS	Terra	250m-1km	NCC w/ DEM	Fixed grid
HRS	SPOT	2.5m	NCC w/ DEM	Not described
ETM+	Landsat-7	15m-60m	NCC to arid region CPs	Fit to surface
VEGETATION	SPOT	1km	NCC w/ DEM	Not described

Image Registration Algorithm Classifications

- Area-Based vs. Feature-Based
  - Often Combination of Area- and Feature-Based

- Alternate Classification:
  - Manual Registration
  - Correlation-Based Methods
  - Fourier-Domain and Other Transform-Based Methods
  - Mutual Information and Distribution-Based Approaches
  - Feature-Point Methods
  - Contour- and Region-Based Approaches

Feature Extraction

- Features:
  - Gray levels
  - Salient points - *Matched point-to-point or globally*
    - Edge or edge-like, e.g., Sobel, Canny
    - Fourier coefficients
    - Gabor, Wavelets, Directional Gabor or Wavelets, Shearlets, etc.
    - Corners, e.g., Kearny, Harris and Stephens, Shi and Tomasi
  - Lines (Hough and Generalized), Contours (Govindu et al), Regions (Region Segmentation, e.g., Tilton)
    - Marked Point Processes (MPP): probabilistic framework with configuration space consisting of an unknown number of parametric objects
  - Scale invariant feature transform (SIFT-Lowe) and variants, e.g., Speeded Up Robust Features (SURF)
  - More recently, Neural Networks (NN) have been used for registration

Similarity Metrics

- **Cross-correlation**
  - Maximize *cross-correlation* over image overlap

$$I_1(x,y) \circ I_2(x,y) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} I_1(m,n) I_2(x+m, y+n)$$

- **Normalized cross-correlation (NCC)**
  - Maximize normalized cross-correlation

$$NCC_{I_1,I_2}(x,y) = \frac{\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}\left[I_1(m,n)-\bar{I}_1\right]\left[I_2(x+m,y+n)-\bar{I}_2\right]}{\sqrt{\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}\left[I_1(m,n)-\bar{I}_1\right]^2 \cdot \sum_{m=0}^{M-1}\sum_{n=0}^{N-1}\left[I_2(x+m,y+n)-\bar{I}_2\right]^2}}$$

- **Mutual information (MI):**

Maximizes the degree of statistical dependence between the images

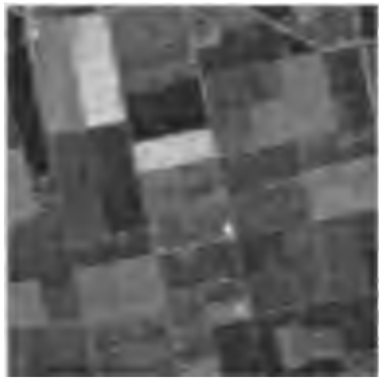
$$MI\left(I_1,I_2\right) = \sum_{g_1}\sum_{g_2} p_{I_1,I_2}\left(g_1,g_2\right)\cdot \log\left(\frac{p_{I_1,I_2}\left(g_1,g_2\right)}{p_{I_1}\left(g_1\right)\cdot p_{I_2}\left(g_2\right)}\right),$$

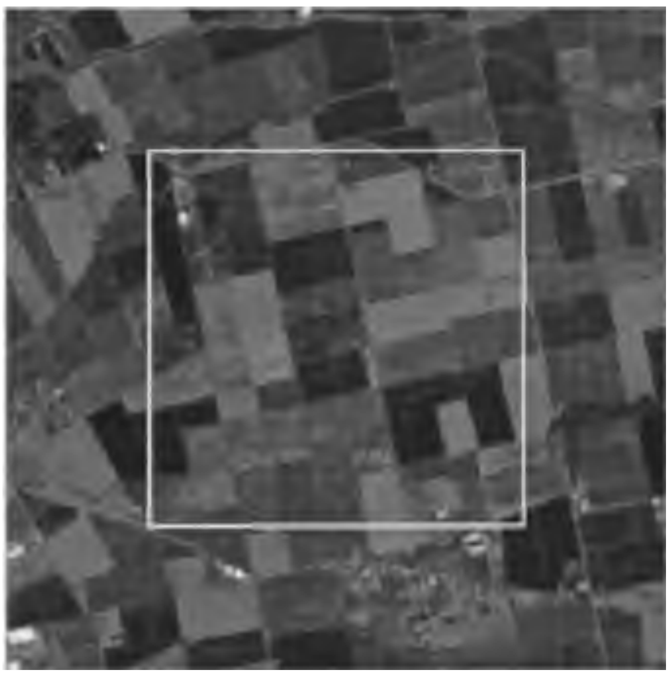
or using histograms, maximizes

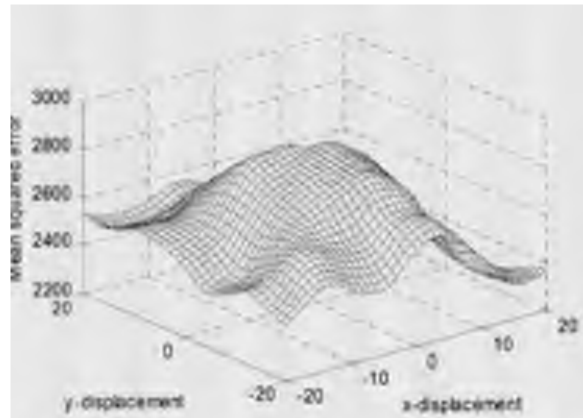
$$MI\left(I_1,I_2\right) = \frac{1}{M}\sum_{g_1}\sum_{g_2} h_{I_1,I_2}\left(g_1,g_2\right)\cdot \log\left(\frac{Mh_{I_1,I_2}\left(g_1,g_2\right)}{h_{I_1}\left(g_1\right)\cdot h_{I_2}\left(g_2\right)}\right)$$

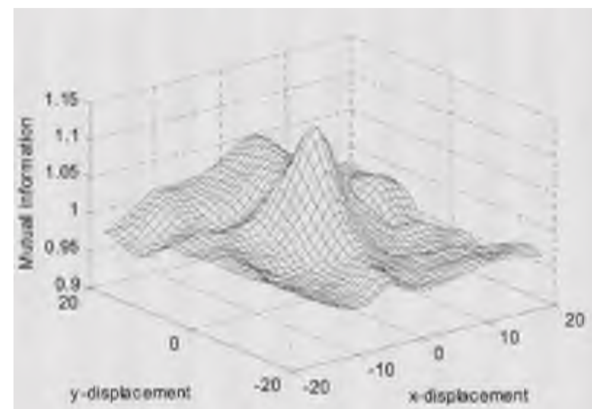
where M is the sum of all histogram entries, i.e., number of pixels (in overlapping subimage)

Similarity Metrics (cont.)









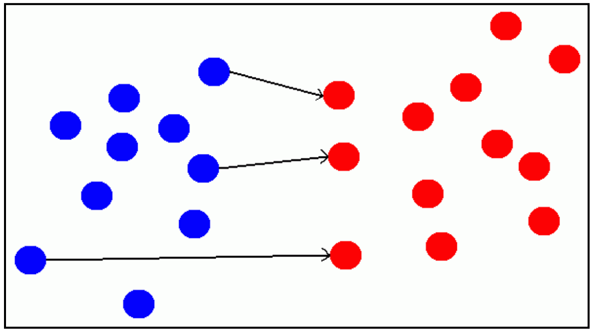
MI vs. L_2-norm and NCC applied to Landsat-5 images

Other Similarity Metrics

- **Partial Hausdorff distance (PHD):**

$$H_K(I_1, I_2) = K^{th}_{p_1 \in I_1} \min_{p_2 \in I_2} \operatorname{dist}(p_1, p_2),$$

where $1 \le K \le |I_1|$ (Huttenlocher *et al*, Mount *et al*)



- **Discrete Gaussian mismatch (DGM):**

$$\mathrm{DGM}_\sigma(I_1, I_2) = 1 - \frac{\sum_{a \in I_1} w_\sigma(a)}{|I_1|}$$

where $w_\sigma(a)$ denotes the *weight* of point *a*, and $w_\sigma(a) = \exp\left(-\frac{\operatorname{dist}(a, I_2)^2}{2\sigma^2}\right)$

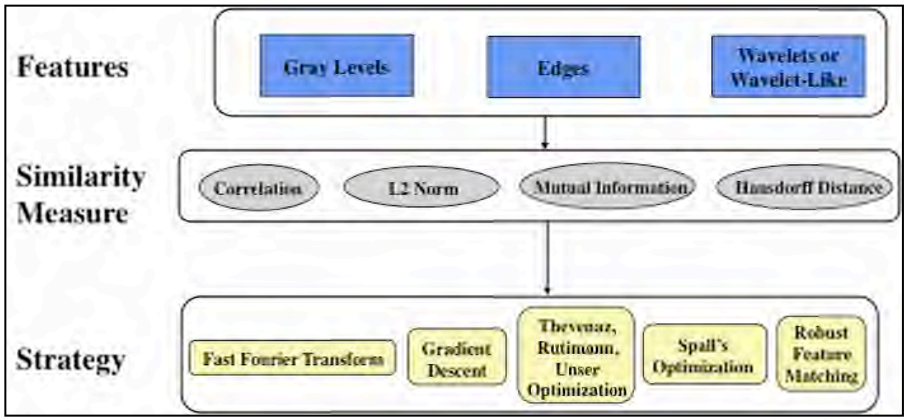
Image Matching Strategies

- Matching strategies matched with feature extraction techniques
- Some methods:
  - Exhaustive Search
  - FFT/Phase Correlation – Fourier Mellin Transform
  - Optimization:
    - Steepest Gradient Descent
    - Levenberg-Marquart
    - Stocchastic Gradient
  - Robust Feature Matching (RFM)
  - Genetic algorithms (including binary shapes)
  - Neural Networks (esp. for quantum & cognitive computing)
- Global or local registration
- Various image representations, e.g., Multi-resolution and quadtrees

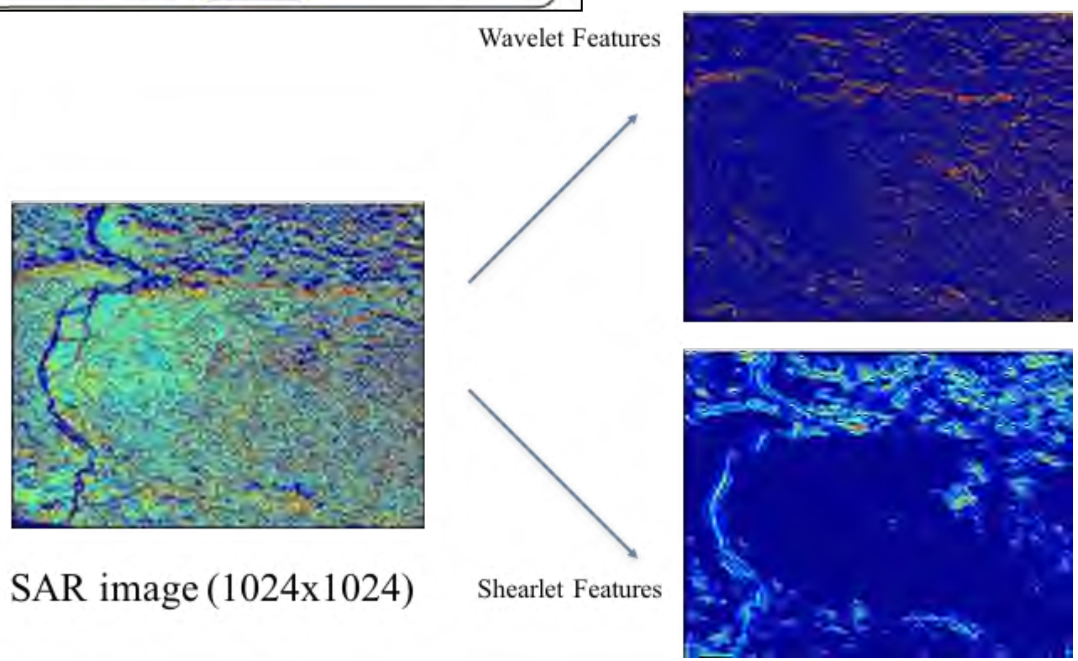
Some Recent Image Registration Results

Wavelet and Wavelet-Like Based Algorithms

Edge, Wavelet and Wavelet-Like Based Registration Framework

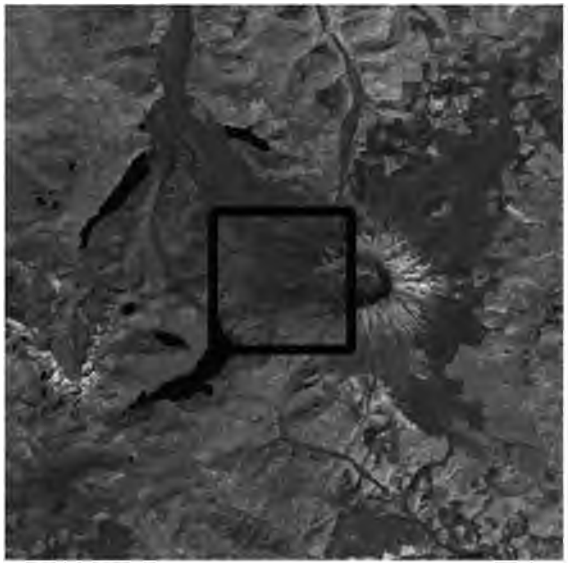


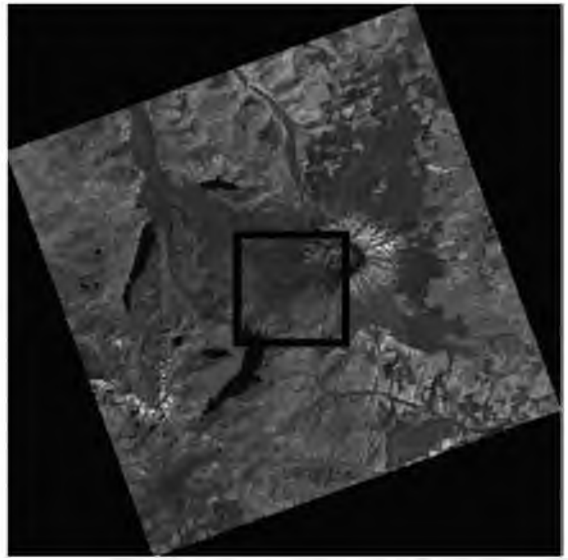
- Wavelets are fundamentally *isotropic*, i.e., no directional sensitivity
- Generalization of wavelets to be *anisotropic* => *Shearlets*, which ***refine the wavelet construction by including a directional component***



Some Recent Image Registration Results

Landsat Warped and Noise Experiments





Geometrically warped synthetic input images. The full source image is 1024 x 1024 Landsat-5 TM image from the Mount Hood are. The extracted images are 256 x 256.

256 x 256 Landsat-7 ETM+ images of Washington, DC, (left) without and (right) with Gaussian noise added. The parameters for the noise are mean μ = 0 and variance σ2 = 0.05

Shearlet-Based Registration Results

As a Function of Warp

Comparison of Registration Algorithms for Landsat-TM Geometrically Warped Synthetic Experiments

Registration Technique	Number of Converged Experiments (out of 200)	Percentage of Converged Experiments	Mean RMSE	Standard Deviation RMSE	Relative Improvement
Spline Wavelets	108	54.00%	.0019	.0017	-
Simoncelli Band-Pass	21	10.50%	.0045	.0014	-
Simoncelli Low-Pass	113	55.50%	.0040	.0036	-
Shearlets	154	77.00%	3.9513	1.5506	-
Shearlet+ Spline Wavelets	154	77.00%	.0058	.0062	42.59%
Shearlet+ Simoncelli Band-Pass	154	77.00%	.0080	.0050	633.33%
Shearlet + Simoncelli Low-Pass	154	77.00%	.0081	.0081	36.28%

Shearlet-Based Registration Results

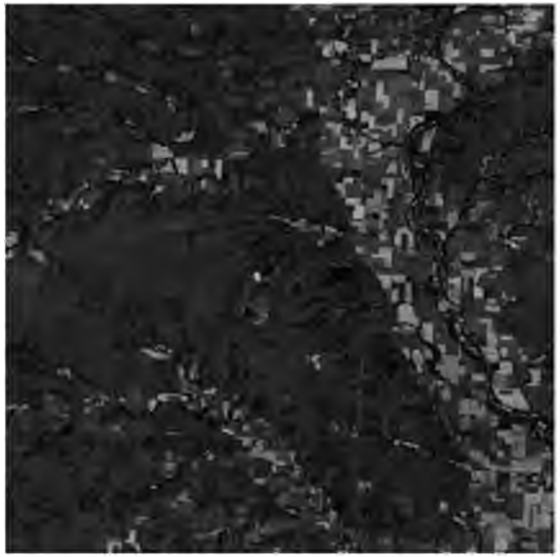
As a Function of Noise

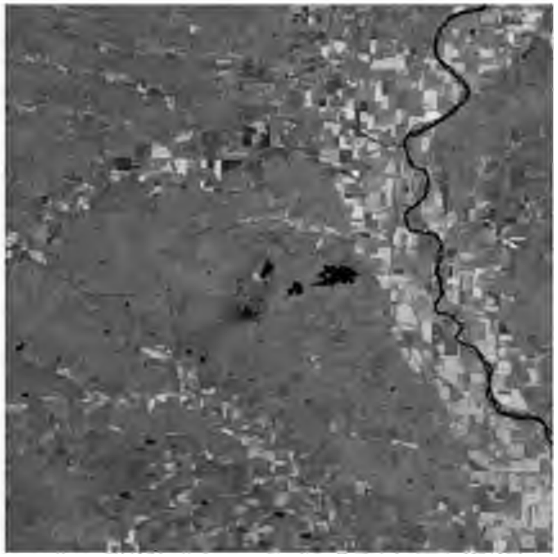
Comparison of Registration Algorithms for Noisy Landsat-ETM+ Synthetic Experiments (Variance = 0.05)

Registration Technique	Number of Converged Experiments (out of 201)	Percentage of Converged Experiments	Mean RMSE	Standard Deviation RMSE	Relative Improvement
Spline Wavelets	31	15.42%	.0579	.0001	-
Simoncelli Band-Pass	42	20.90%	.0805	~ 0	-
Simoncelli Low-Pass	67	33.33%	.0560	~ 0	-
Shearlets	98	48.76%	1.8486	1.1933	-
Shearlet+ Spline Wavelets	98	48.76%	.0468	~ 0	216.13%
Shearlet+ Simoncelli Band-Pass	98	48.76%	.0805	~ 0	133.33%
Shearlet + Simoncelli Low-Pass	99	48.76%	.0560	~ 0	46.27%

Shearlet-Based Registration Results

Multimodal Experiments





1024 x 1024 images of (left) ETM+ Infared/Red band and (right) Near-Infared/NIR band of the Konza Prairie

Pixels computed by SIFT in the LIDAR shaded-relief (left) and optical (right) images of Washington State, connected by line segments. Note the lack of correspondence; such points are unsuitable for a registration algorithm.

Shearlet-Based Registration Results

For LIDAR Data

Comparison of Registration Algorithms for LIDAR Warped Synthetic Experiments

Registration Technique	Number of Converged Experiments (out of 201)	Percentage of Converged Experiments	Mean RMSE	Standard Deviation RMSE	Relative Improvement
Spline Wavelets	74	36.82%	.3552	.0256	-
Simoncelli Band-Pass	42	20.90%	.0074	~ 0	-
Simoncelli Low-Pass	72	35.82%	.2412	.0166	-
Shearlets	108	53.73%	.0304	.0012	-
Shearlet+ Spline Wavelets	111	55.22%	.3222	.0143	50.00%
Shearlet+ Simoncelli Band-Pass	108	53.73%	.0075	~ 0	157.14%
Shearlet + Simoncelli Low-Pass	111	55.22%	.2432	~ 0	54.71%

Shearlet-Based Registration Results

Multimodal Experiments

Comparison of Registration Algorithms for ETM+ Infrared to NIR Multimodal Experiments

Registration Technique	Number of Converged Experiments (out of 41)	Percentage of Converged Experiments	Mean RMSE	Standard Deviation RMSE	Relative Improvement
Spline Wavelets	25	60.98%	.2389	.0137	-
Simoncelli Band-Pass	18	43.90%	.2492	~ 0	-
Simoncelli Low-Pass	34	82.93%	.2100	~ 0	-
Shearlets	38	92.68%	.6678	.3917	-
Shearlet+ Spline Wavelets	38	92.68%	.2465	.0336	52.00%
Shearlet+ Simoncelli Band-Pass	38	92.68%	.2492	~ 0	111.11%
Shearlet + Simoncelli Low-Pass	38	92.68%	.2100	~ 0	11.76%

Shearlet-Based Registration Results

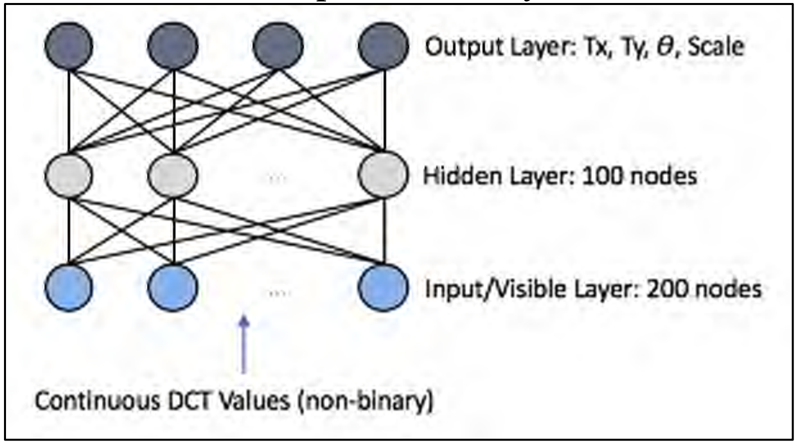
Multimodal Experiments (cont.)

Comparison of Registration Algorithms for LIDAR to Optical Multimodal Experiments

Registration Technique	Number of Converged Experiments (out of 101)	Percentage of Converged Experiments	Mean RMSE	Standard Deviation RMSE	Relative Improvement
Spline Wavelets	55	54.46%	3.4499	.0012	-
Simoncelli Band-Pass	61	60.40%	3.6542	.0174	-
Simoncelli Low-Pass	86	85.15%	3.5918	.0066	-
Shearlets	44	87.13%	15.6428	6.1668	-
Shearlet + Spline Wavelets	60	59.41%	3.4222	~ 0	9.09%
Shearlet + Simoncelli Band-Pass	65	64.36%	3.6518	.0174	6.56%
Shearlet + Simoncelli Low-Pass	88	87.13%	3.5912	.0083	2.33%

Preliminary Image Registration Results
Using Artificial Neural Networks

Feed-forward neural network (FF-NN) for subpixel accuracy



- Using **Discrete Cosine Transform (DCT)** coefficients as input provides subpixel accuracy
- **Input:** 100 DCT coefficients from reference image + 100 DCT coefficients from test image
- **Output:** Transformation Variables (Tx, Ty, Θ, s)
- **Score:** Subpixel registration accuracy if mean RMS error < 1.0 per pixel

- Subpixel accuracy on 50% of the test images in < 500 training epochs. Running for longer increases accuracy
- Training set must be large enough to capture the range of values for rotation/translation in the test set
  - Training set of 100 images randomly rotated/translated from a source image is enough to learn:
    +/- 45° rotation coupled with +/- 10 pixels translation
  - Training set of 300 images is enough to learn:
    +/- 120º rotation, no translation
    +/- 80 pixels translation, no rotation
- Current experiments using Deep Belief Networks and Restricted Boltzman Machines

Conclusions

- Brief introduction to remote sensing image registration and its main components:
  - Feature Extraction
  - Similarity Metrics
  - Search Strategies
- Components combined appropriately and adapted to:
  - Type of data (e.g., edge- vs. texture-rich)
  - Size of data and computational resource needed
  - Required accuracy
  - Initial conditions
- Future Work:
  - Systematic assessment of various algorithms
  - Creating benchmark datasets