Introduction to Remote Sensing

Image Registration

Jacqueline Le Moigne,
NASA Goddard Space Flight Center

IGARSS 2017
Problem Description

• Earth Science studies such as:
  o Climate change over multiple time scales
  o Predicting crop production
  o Monitoring land resources
  o 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
  o Multiple-time or simultaneous observations of the same Earth features by different sensors
  o Global measurements with remote sensing systems
  o Complemented by regional and local measurements using ground and airborne sensors
  o 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

GPM Constellation of Satellites

- NPP (NASA/IPO)
- GPM Core Observatory (NASA/JAXA)
- MetOp B & C (EUMETSAT)
- JPSS-1 (NASA/NOAA)
- NOAA 19 (NOAA)
- DMSP F19/F20 (DOD)
- GCOM-W1 (JAXA)
- Megha-Tropiques (CNES/ISRO)
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

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)
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**
  o $I_1(x,y)$ and $I_2(x,y)$: images or image/map
    ✓ find the mapping $(f,g)$ which transforms $I_1$ into $I_2$: $I_2(x,y) = g(I_1(f(x,y),f(y,y)))$
      a. $f$: spatial mapping
      b. $g$: radiometric mapping
  o **Spatial Transformations** “$f$”
    ✓ Translation, Rigid, Affine, Projective, Perspective, Polynomial, …
  o **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**
  - $I_1(x,y)$ and $I_2(x,y)$: images or image/map
    - find the mapping $(f,g)$ which transforms $I_1$ into $I_2$: $I_2(x,y) = g(I_1(fx(x,y),fy(x,y)))$
      - $f$: spatial mapping
      - $g$: 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**

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

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

• Alternate Classification:
  o Manual Registration
  o Correlation-Based Methods
  o Fourier-Domain and Other Transform-Based Methods
  o Mutual Information and Distribution-Based Approaches
  o Feature-Point Methods
  o 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
  \[
  I_1(x, y) \ast I_2(x, y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I_1(m, n) I_2(x + m, y + n)
  \]

- **Normalized cross-correlation (NCC)**
  - Maximize normalized cross-correlation
  \[
  NCC_{I_1, I_2}(x, y) = \frac{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} [I_1(m, n) - \bar{I}_1][I_2(x + m, y + n) - \bar{I}_2]}{\sqrt{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} [I_1(m, n) - \bar{I}_1]^2} \cdot \sqrt{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} [I_2(x + m, y + n) - \bar{I}_2]^2}}
  \]

- **Mutual information (MI):**
  - Maximizes the degree of statistical dependence between the images
  \[
  MI(I_1, I_2) = \sum_{g_1} \sum_{g_2} p_{I_1, I_2}(g_1, g_2) \cdot \log \left( \frac{p_{I_1, I_2}(g_1, g_2)}{p_{I_1}(g_1) \cdot p_{I_2}(g_2)} \right),
  \]
  or using histograms, maximizes
  \[
  MI(I_1, I_2) = \frac{1}{M} \sum_{g_1} \sum_{g_2} h_{I_1, I_2}(g_1, g_2) \cdot \log \left( \frac{Mh_{I_1, I_2}(g_1, g_2)}{h_{I_1}(g_1) \cdot h_{I_2}(g_2)} \right)
  \]
  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):**

\[
H_K(I_1, I_2) = K^\text{th} \min_{p_2 \in I_2} \text{dist}(p_1, p_2),
\]

where \(1 \leq K \leq |I_1|\) (Huttenlocher *et al*., Mount *et al*)

\[
\text{dist}(p_1, p_2) = \sqrt{(p_{1x} - p_{2x})^2 + (p_{1y} - p_{2y})^2}
\]

\[
1 \leq K \leq |I_1|
\]

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

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

where \(w_\sigma(a)\) denotes the weight of point \(a\), and \(w_\sigma(a) = \exp\left(-\frac{\text{dist}(a, I_2)^2}{2\sigma^2}\right)\)
Image Matching Strategies

• Matching strategies matched with feature extraction techniques

• Some methods:
  o Exhaustive Search
  o FFT/Phase Correlation – Fourier Mellin Transform
  o Optimization:
    ▪ Steepest Gradient Descent
    ▪ Levenberg-Marquart
    ▪ Stochastic Gradient
  o Robust Feature Matching (RFM)
  o Genetic algorithms (including binary shapes)
  o 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

- 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*

**Edge, Wavelet and Wavelet-Like Based Registration Framework**

- **Features**
  - Gray Levels
  - Edges
  - Wavelets or Wavelet-Like

- **Similarity Measure**
  - Correlation
  - L2 Norm
  - Mutual Information
  - Hausdorff Distance

- **Strategy**
  - Fast Fourier Transform
  - Gradient Descent
  - Thrun, Rutimann, Unser Optimization
  - Spall’s Optimization
  - Robust Feature Matching

**Wavelet Features**

**SAR image (1024x1024)**

**Shearlet Features**
Some Recent Image Registration Results
Landsat Warped and Noise Experiments

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 $\mu = 0$ and variance $\sigma^2 = 0.05$

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

As a Function of Warp

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

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

As a Function of Noise

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

<table>
<thead>
<tr>
<th>Registration Technique</th>
<th>Number of Converged Experiments (out of 201)</th>
<th>Percentage of Converged Experiments</th>
<th>Mean RMSE</th>
<th>Standard Deviation RMSE</th>
<th>Relative Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spline Wavelets</td>
<td>31</td>
<td>15.42%</td>
<td>0.0579</td>
<td>0.0001</td>
<td>-</td>
</tr>
<tr>
<td>Simoncelli Band-Pass</td>
<td>42</td>
<td>20.90%</td>
<td>0.0805</td>
<td>~ 0</td>
<td>-</td>
</tr>
<tr>
<td>Simoncelli Low-Pass</td>
<td>67</td>
<td>33.33%</td>
<td>0.0560</td>
<td>~ 0</td>
<td>-</td>
</tr>
<tr>
<td>Shearlets</td>
<td>98</td>
<td>48.76%</td>
<td>1.8486</td>
<td>1.1933</td>
<td>-</td>
</tr>
<tr>
<td>Shearlet + Spline Wavelets</td>
<td>98</td>
<td>48.76%</td>
<td>0.0468</td>
<td>~ 0</td>
<td>216.13%</td>
</tr>
<tr>
<td>Shearlet + Simoncelli Band-Pass</td>
<td>98</td>
<td>48.76%</td>
<td>0.0805</td>
<td>~ 0</td>
<td>133.33%</td>
</tr>
<tr>
<td>Shearlet + Simoncelli Low-Pass</td>
<td>99</td>
<td>48.76%</td>
<td>0.0560</td>
<td>~ 0</td>
<td>46.27%</td>
</tr>
</tbody>
</table>
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

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

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

<table>
<thead>
<tr>
<th>Registration Technique</th>
<th>Number of Converged Experiments (out of 41)</th>
<th>Percentage of Converged Experiments</th>
<th>Mean RMSE</th>
<th>Standard Deviation RMSE</th>
<th>Relative Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spline Wavelets</td>
<td>25</td>
<td>60.98%</td>
<td>.2389</td>
<td>.0137</td>
<td>-</td>
</tr>
<tr>
<td>Simoncelli Band-Pass</td>
<td>18</td>
<td>43.90%</td>
<td>.2492</td>
<td>~ 0</td>
<td>-</td>
</tr>
<tr>
<td>Simoncelli Low-Pass</td>
<td>34</td>
<td>82.93%</td>
<td>.2100</td>
<td>~ 0</td>
<td>-</td>
</tr>
<tr>
<td>Shearlets</td>
<td>38</td>
<td>92.68%</td>
<td>.6678</td>
<td>.3917</td>
<td>-</td>
</tr>
<tr>
<td>Shearlet+Spline Wavelets</td>
<td>38</td>
<td>92.68%</td>
<td>.2465</td>
<td>.0336</td>
<td>52.00%</td>
</tr>
<tr>
<td>Shearlet+Simoncelli Band-Pass</td>
<td>38</td>
<td>92.68%</td>
<td>.2492</td>
<td>~ 0</td>
<td>111.11%</td>
</tr>
<tr>
<td>Shearlet+Simoncelli Low-Pass</td>
<td>38</td>
<td>92.68%</td>
<td>.2100</td>
<td>~ 0</td>
<td>11.76%</td>
</tr>
</tbody>
</table>
### Shearlet-Based Registration Results
#### Multimodal Experiments (cont.)

Comparison of Registration Algorithms for LIDAR to Optical Multimodal Experiments

<table>
<thead>
<tr>
<th>Registration Technique</th>
<th>Number of Converged Experiments (out of 101)</th>
<th>Percentage of Converged Experiments</th>
<th>Mean RMSE</th>
<th>Standard Deviation RMSE</th>
<th>Relative Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spline Wavelets</td>
<td>55</td>
<td>54.46%</td>
<td>3.4499</td>
<td>.0012</td>
<td>-</td>
</tr>
<tr>
<td>Simoncelli Band-Pass</td>
<td>61</td>
<td>60.40%</td>
<td>3.6542</td>
<td>.0174</td>
<td>-</td>
</tr>
<tr>
<td>Simoncelli Low-Pass</td>
<td>86</td>
<td>85.15%</td>
<td>3.5918</td>
<td>.0066</td>
<td>-</td>
</tr>
<tr>
<td>Shearlets</td>
<td>44</td>
<td>87.13%</td>
<td>15.6428</td>
<td>6.1668</td>
<td>-</td>
</tr>
<tr>
<td>Shearlet + Spline Wavelets</td>
<td>60</td>
<td>59.41%</td>
<td>3.4222</td>
<td>~ 0</td>
<td>9.09%</td>
</tr>
<tr>
<td>Shearlet + Simoncelli Band-Pass</td>
<td>65</td>
<td>64.36%</td>
<td>3.6518</td>
<td>.0174</td>
<td>6.56%</td>
</tr>
<tr>
<td>Shearlet + Simoncelli Low-Pass</td>
<td>88</td>
<td>87.13%</td>
<td>3.5912</td>
<td>.0083</td>
<td>2.33%</td>
</tr>
</tbody>
</table>
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:
    +/− 45º 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:
  o Feature Extraction
  o Similarity Metrics
  o Search Strategies

• Components combined appropriately and adapted to:
  o Type of data (e.g., edge- vs. texture-rich)
  o Size of data and computational resource needed
  o Required accuracy
  o Initial conditions

• Future Work:
  o Systematic assessment of various algorithms
  o Creating benchmark datasets