Multi-polarized PALSAR and LandSat multi-modality data fusion for Crop Classification

Zhengwei Yang, Claire Boryan, Patrick Willis, Rick Mueller
USDA/NASS, Research & Development Division

Barry Haack
George Mason University
Outline

- Introduction
- Data
- Processing & Methods
- Classification Results and Discussion
- Conclusions
Introduction

- **USDA NASS’ Mission:**
  - Provides timely, accurate and unbiased agricultural statistics
- **NASS remote sensing based Crop Data Layer (CDL) program** provides acreage estimation. Its classification accuracy relies on sufficient quality image data. But the reality is:
  - Limited images available because of limited budget, cloud cover, time constraint.
- **Any type of image data available for improving crop identification should be studied**
  - Leads this radar data application study.
Why Studying Radar Image

- Test site Central Valley, California has many varieties of crops.
- Radar’s special characteristics may help to improve crop identification.
- Test images are available!
Objective of This Study

- Seeking answers to following questions:
  - Can fusion of Radar and LandSat data improve crop identification?
  - How big is the impact of the backscatter noise?
  - Can noise filtering help improving classification accuracy?
  - Can texture features help improving classification accuracy?
  - Can Radar images be helpful?
ALOS-PALSAR Radar Data

- L-band, polarimetric;
- Spatial resolution: 12.5m;
- Swath Width: 20-65km;
- Recorded on July 1, 2007
- Quad polarization: HH, HV, VH, VV
PALSAR Data – Central Valley, CA

Quad polarizations  HH polarization  HV polarization
LandSat TM

- Spatial resolution: 30m at nadir
- Quantization: 8 bits
- Spectral bands 7
  - Blue Band 1, 0.45-0.515 μm
  - Green Band 2, 0.525-0.605 μm
  - Red Band 3, 0.63-0.69 μm
  - NIR Band 4, 0.75-0.90 μm
  - SWIR (Band 5, 1.55-175 μm)
  - Band 6, 10.4 -12.5 μm
  - Band 7, 2.09-2.35 μm
- Repeat period: 16days.
- Swath width: 185km
- Path 43 Row 34 and Path 43 and Row 35, July 30, 2007
Data Processing

- LandSat mosaic (seamless and color balance)
- LandSat and PALSAR co-registration
- Radar noise filtering
- Image fusion
- Texture feature calculation
- Training & validation data preparation.
Image Registration

- LandSat and Palsar Images were registered.

- Unregistered Quad polys

- Registered Quad polys
Noise Filtering

- A few filters tested
- Median filter
  - Selects the value in the middle of the range of values within the moving window.
- Lee-Sigma filter
  - Uses the average of all pixel values within the moving window that fall within the designated range of standard deviations.
- Gamma-MAP filter
  - Maximizes the a posteriori probability density function. This filter attempts to derive the original pixel value which must lie between the local average and the degraded pixel value.
Filtered Images

Quad polarization (unfiltered)

Lee-Sigma 3x3

Median 3x3

Gamma-MAP 3x3

Median 5x5
Image Fusion Methods

- Many fusion methods: IHS, PCA, High pass filtering, Wavelet, Ehlers Fusion, Brovey, Difference & Ratio, Adding & Multiplication, etc.
- Image fusion can be performed at 3 fusion levels:
  - 1) Pixel; 2) Feature; 3) Decision level;
- The most popular pixel level methods:
  - Intensity-Hue-Saturation;
  - Principal Component Analysis;
- For classification, image bands from different sensors acquired on different dates can be stacked for input.
The approach for the computation of the principal components (PCs) comprises the calculation of:

1. Autocorrelation matrix;
2. Eigen-values, Eigenvectors;
3. Principal component;

PCA Fusion:
1. Replace the first principal component (Popular);
2. PCA of all multi-image data channels;
3. Reverse PCA.

Transformation settings:
Remap, Cubic convolution for resampling
PALSAR-LandSat Fused Images
PALSAR-LandSat Fused Images - Zoom-in View

Quad polarization image

HH Pol

HH Pol-LandSat fusion

LandSat image

HV Pol

HV Pol-LandSat fusion
Texture Features

- There are many texture features. Two were tested here:
- Mean Euclidean distance with 7x7 window
- Variance with 7x7 window
The existing USDA ground truth data (CLU & Admin 578 data) do not cover all crop types since the scene covered area is too small.

2007 California Cropland Data Layer (CDL) used for both training and validation.

CDL - a crop land cover product annually generated by USAD/NASS. It’s a raster data with crop identified.

Both training and validation data are stratified in sampling.
Training & Validation Data

- Yellow (corn)
- Green (soybean)
- Red (Cotton)
- Blue (Rice)
- Purple (Grape)

2007 CDL

Validation Sample
Classification Method

- **Classifier:**
  - Supervised decision tree classifier

- **Why - advantages:**
  - A white box model - easily explained by Boolean logic and easy to understand and interpret results;
  - Able to handle both numerical and categorical data;
  - Robust - tolerates training errors and cloud pixels;
  - Good computational performance.
  - No assumption of data distribution required;
  - Easy to validation;
  - Little data preparation needed;
  - Excellent scalability - no limit in number of data layers;

- **Inputs:**
  - Training data, various source radar/LandSat images
Classification with LandSat or PALSAR Pols

<table>
<thead>
<tr>
<th></th>
<th>Cotton</th>
<th>Rice</th>
<th>Alfalfa</th>
<th>Almonds</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Producer</td>
<td>User</td>
<td>Producer Kappa</td>
<td>Producer</td>
<td>User</td>
</tr>
<tr>
<td>LandSat</td>
<td>72.35</td>
<td>57.64</td>
<td>0.6332</td>
<td>56.27</td>
<td>67.55</td>
</tr>
<tr>
<td>hh</td>
<td>85.26</td>
<td>25.68</td>
<td>0.5752</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>hv</td>
<td>85.23</td>
<td>26.01</td>
<td>0.5845</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>vh</td>
<td>87.94</td>
<td>24.64</td>
<td>0.5954</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>vv</td>
<td>81.20</td>
<td>25.51</td>
<td>0.4973</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Quad Pols</td>
<td>76.80</td>
<td>28.20</td>
<td>0.5003</td>
<td>0.05</td>
<td>0.92</td>
</tr>
</tbody>
</table>

- LandSat performs better than Radar for all except for Almonds;
- Individual pols hh, hv, vh perform better than Quad pols for Cotton;
- Accuracies of Quad pol and individual pols are extremely low for Rice;
- For Almonds the producer accuracy of hv and vh are better than LandSat result;
- For cotton the producer accuracies of Quad pols and individual pols are significantly better than that of LandSat.
### Filtered Data Classification Result

<table>
<thead>
<tr>
<th></th>
<th>Cotton</th>
<th>Rice</th>
<th>Alfalfa</th>
<th>Almonds</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Producer</td>
<td>User</td>
<td>Producer Kappa</td>
<td>Producer</td>
<td>User</td>
</tr>
<tr>
<td>LandSat</td>
<td>72.35</td>
<td>57.64</td>
<td>0.6332</td>
<td>56.27</td>
<td>67.55</td>
</tr>
<tr>
<td>hh</td>
<td>85.26</td>
<td>25.68</td>
<td>0.5752</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>hv</td>
<td>85.23</td>
<td>26.01</td>
<td>0.5845</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>vh</td>
<td>87.94</td>
<td>24.64</td>
<td>0.5954</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>vv</td>
<td>81.20</td>
<td>25.51</td>
<td>0.4973</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>qp_gamma_map3</td>
<td>74.75</td>
<td>29.96</td>
<td>0.5043</td>
<td>0.19</td>
<td>6.56</td>
</tr>
<tr>
<td>qp_lee_sigma3</td>
<td>74.55</td>
<td>29.96</td>
<td>0.5015</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>qp_median3</td>
<td>74.37</td>
<td>29.72</td>
<td>0.4953</td>
<td>0.05</td>
<td>1.49</td>
</tr>
<tr>
<td>qp_median5</td>
<td>72.67</td>
<td>31.08</td>
<td>0.4940</td>
<td>0.38</td>
<td>14.29</td>
</tr>
<tr>
<td>Quad Pols</td>
<td>76.80</td>
<td>28.20</td>
<td>0.5003</td>
<td>0.05</td>
<td>0.92</td>
</tr>
</tbody>
</table>

- For Cotton, filtered results are inferior to original polarization results
- For Rice, Alfalfa and Almonds, accuracies of filtered quad pols are better than unfiltered quad pols
- Among filtered data, different methods with smaller window 3x3 yield close results. The larger window 5x5 gives better result than 3x3 window for Median filter.
### Classification Results using PALSAR-LandSat Fusion/Texture

<table>
<thead>
<tr>
<th></th>
<th>Cotton</th>
<th>Rice</th>
<th>Alfalfa</th>
<th>Almonds</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Producer</td>
<td>User</td>
<td>Producer Kappa</td>
<td>Producer</td>
<td>User</td>
</tr>
<tr>
<td>LandSat</td>
<td>72.35</td>
<td>57.64</td>
<td>0.6332</td>
<td>56.27</td>
<td>67.55</td>
</tr>
<tr>
<td>hh_landsat</td>
<td>73.46</td>
<td>59.75</td>
<td>0.6500</td>
<td>48.92</td>
<td>56.43</td>
</tr>
<tr>
<td>hv_landsat</td>
<td>74.70</td>
<td>60.52</td>
<td>0.6659</td>
<td>48.17</td>
<td>55.11</td>
</tr>
<tr>
<td>vh_landsat</td>
<td>74.41</td>
<td>59.68</td>
<td>0.6609</td>
<td>49.44</td>
<td>55.64</td>
</tr>
<tr>
<td>vv_landsat</td>
<td>72.51</td>
<td>59.14</td>
<td>0.6377</td>
<td>48.21</td>
<td>55.44</td>
</tr>
<tr>
<td>Quad Pols</td>
<td>76.80</td>
<td>28.20</td>
<td>0.5003</td>
<td>0.05</td>
<td>0.92</td>
</tr>
<tr>
<td>qp_med7_var7</td>
<td>52.70</td>
<td>33.76</td>
<td>0.3021</td>
<td>1.69</td>
<td>7.64</td>
</tr>
</tbody>
</table>

- Palsar-LandSat fusion performs better than both Quad pols and LandSat for Almonds and Cotton.
- Palsar-LandSat fusion results are worse than LandSat for Rice and similar to LandSat for Alfalfa;
- **Overall, the fused results perform better than radar data alone**
- Including texture features don’t improve classification accuracy. They even deteriorate the results. This means that texture features computed with less than or equal to 7x7 window does not help classification accuracy improvement at all!
Classification Result - CDL

2007 CDL (LandSat)  
VH_LandSat fusion result
Conclusions

- Texture features calculated with less than or equal to 7x7 windows don’t help classification accuracy improvement at all. They even reduced classification accuracy. However, it’s not clear that if using larger window size will be helpful or not.

- The backscatter noise significantly affects the classification accuracy as evidenced by the facts that the filtered data with large window size performed better than the unfiltered Radar data.

- The fusion of Radar and LandSat data can improve crop identification as evidenced by the facts that the fused results performs better than radar data alone and Palsar-LandSat fusion performs better than both Quad pols and LandSat for Almonds and Cotton. However, it does not help for every crop type. In general, Radar images can be helpful if properly used.
THANK YOU

QUESTION?

Contact Information:
Email: zhengwei_yang@nass.usda.gov
Tel: 703-877-8000