Multitemporal Region-Based Classification of High-Resolution Images by Markov Random Fields and Multiscale Segmentation

Gabriele Moser
Sebastiano B. Serpico
Outline

• Introduction
  – Multitemporal high-resolution image classification

• The proposed method
  – Overall architecture
  – Multitemporal multiscale region-based Markov random field
  – Parameter estimation and energy minimization

• Experimental results
  – Data sets and experimental set-up
  – Classification maps
  – Classification accuracies and comparisons

• Conclusion
• **Introduction**
  - Multitemporal high-resolution image classification

• The proposed method
  - Overall architecture
  - Multitemporal multiscale region-based Markov random field
  - Parameter estimation and energy minimization

• Experimental results
  - Data sets and experimental set-up
  - Classification maps
  - Classification accuracies and comparisons

• Conclusion
Introduction

Multitemporal high-resolution (HR) optical image classification:

- Important for environmental monitoring (e.g. in disaster prevention and management), urban area analysis, etc.
- 0.5 ÷ 10 m resolution by current (e.g., IKONOS, QuickBird, WorldView-2, GeoEye-1, SPOT-5 HRG) and forthcoming (e.g., Pleiades) missions.
- Need for modeling the temporal and spatial-geometrical information in multitemporal HR data.

A supervised multitemporal multiscale region-based HR image classifier is proposed, based on:

- Multiscale segmentation;
- Markov random field (MRF) models.
Outline

• Introduction
  – Multitemporal high-resolution image classification

• The proposed method
  – Overall architecture
  – Multitemporal multiscale region-based Markov random field
  – Parameter estimation and energy minimization

• Experimental results
  – Data sets and experimental set-up
  – Classification maps
  – Classification accuracies and comparisons

• Conclusion
The Proposed Method

• Motivation
  – Key role of spatial-contextual and multiscale information in HR image analysis.
  – Need for modeling the temporal context associated to multitemporal images.

• Key-ideas
  – Extracting multiscale information at each acquisition time by generating segmentation maps at different scales.
  – Fusing spatial, temporal, and multiscale information by a novel region-based MRF model
  – The model extends two previous models for single-time multiscale and single-scale multitemporal classification, respectively.
Outline

• Introduction
  – Multitemporal high-resolution image classification

• The proposed method
  – Overall architecture
  – Multitemporal multiscale region-based Markov random field
  – Parameter estimation and energy minimization

• Experimental results
  – Data sets and experimental set-up
  – Classification maps
  – Classification accuracies and comparisons

• Conclusion
Multitemporal Markov Random Fields

- **MRF model** for the spatio-temporal context
  - Representation of the statistical interactions between the pixel labels at each time $t_r \ (r = 0, 1)$ by using only local relationships in both the spatial and temporal domains:

$$ P(\ell_{ir}\mid\{\ell_{jr}\}_{j\neq i},\{\ell_{kr,1-r}\}) = P(\ell_{ir}\mid\{\ell_{jr}\}_{j\neq i},\{\ell_{k,1-r}\}_{k=i}) $$

Conditioning only to the labels of the (spatially or temporally) neighboring pixels, according to a given neighborhood system (here $3 \times 3$)

Spatial context of the $i$-th pixel in the image at each time $t_r$

Temporal context of the $i$-th pixel in the image at the other time $t_{1-r}$
Energy Function

- **MRF-based classification**
  - Minimization of a (posterior) energy function $U(\cdot)$, thanks to the Hammersley-Clifford theorem.
  - When using MRFs for data fusion, $U(\cdot)$ is a linear combination of energy contributions, each related to an information source.

- **Proposed region-based multitemporal MRF model**
  - Fusion of spatial, temporal, and multiscale information:

$$U(L_0, L_1 | S_0, S_1) = - \sum_{r=0}^{1} \left[ \sum_i \sum_{q=1}^{Q} \alpha_{qr} \ln P(s_{iqr} | \ell_{ir}) + \beta_r \sum_{i=j} P(\ell_{ir} | \ell_{jr}) + \gamma_r \sum_{i=k} P(\ell_{ir} | \ell_{k,1-r}) \right]$$

  - Transition probability from each class at a time to each class at the other time
  - Potts model for the spatial context

- Pixelwise probability mass function (PMF) of the segment labels in the segmentation map at each scale and each date, conditioned to each class

University of Genoa
Department of Biophysical and Electronic Engineering
Segmentation and PMF Estimation

- **Felzenszwalb & Huttenlocher** segmentation method [3]
  - Graph-based region-growing method depending on a **scale parameter**.
  - Segmentations at different scales by varying this parameter.

- **Class-conditional PMF estimation**
  - Extension of a previous method proposed for a single-time region-based MRF.
  - **Relative-frequency estimate**, based on preliminary classification maps $M_0$ and $M_1$.
  - $M_0$ and $M_1$ generated here by a classical MRF classifier with $k$-NN estimates of the pixelwise class-conditional statistics.
Outline

• Introduction
  – Multitemporal high-resolution image classification

• The proposed method
  – Overall architecture
  – Multitemporal multiscale region-based Markov random field
  – Parameter estimation and energy minimization

• Experimental results
  – Data sets and experimental set-up
  – Classification maps
  – Classification accuracies and comparisons

• Conclusion
Parameter Estimation and Energy Minimization

- **Transition probabilities: expectation-maximization (EM)**
  - Converges (under mild assumptions) to ML estimates.
  - Uses the aforementioned $k$-NN pixelwise estimates.

- **Weight parameters ($\alpha$, $\beta$, $\gamma$) in the MRF**
  - Extension of a recent method based on the Ho-Kashyap algorithm.

- **Energy minimization: iterated conditional mode (ICM)**
  - Initialized with the preliminary maps $M_0$ and $M_1$.
  - Converges to a local energy minimum.
  - Usually good tradeoff between accuracy and time.
Outline

• Introduction
  – Multitemporal high-resolution image classification

• The proposed method
  – Overall architecture
  – Multitemporal multiscale region-based Markov random field
  – Parameter estimation and energy minimization

• Experimental results
  – Data sets and experimental set-up
  – Classification maps
  – Classification accuracies and comparisons

• Conclusion
Data Sets and Experimental Set-up

SPOT-5 HRG data set
- Peking (China), 4 bands, 1500 × 1160 pixels, 10-m resolution, acquisitions in 2003 and 2005.

QuickBird data set
- Phoenix (AZ), 4 bands, 300 × 300 pixels, 2.8-m resolution, acquisitions in 2003 and 2005.

Experimental comparisons
- With a previous multitemporal non-region-based single-scale classifier based on MRFs [6].
- With a previous multitemporal non-contextual single-scale classifier based on EM [8].
Outline

• Introduction
  – Multitemporal high-resolution image classification

• The proposed method
  – Overall architecture
  – Multitemporal multiscale region-based Markov random field
  – Parameter estimation and energy minimization

• Experimental results
  – Data sets and experimental set-up
  – Classification maps
  – Classification accuracies and comparisons

• Conclusion
"Phoenix": Classification Maps

- True color RGB
- EM method in [8]
- MRF method in [6]
- Proposed method

Legend:
- Urban
- Shrub-brush
- Herbaceous
- Barren
- Transitional
"Peking": Classification Maps

False color RGB

Proposed method

2003 2005

urban
agricultural
rangeland
barren
water

Department of Biophysical and Electronic Engineering
Outline

• Introduction
  – Multitemporal high-resolution image classification

• The proposed method
  – Overall architecture
  – Multitemporal multiscale region-based Markov random field
  – Parameter estimation and energy minimization

• Experimental results
  – Data sets and experimental set-up
  – Classification maps
  – Classification accuracies and comparisons

• Conclusion
## Classification Accuracies

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Phoenix</td>
<td>urban</td>
<td>3545</td>
<td>4819</td>
<td>1329</td>
<td>1885</td>
<td>91.20%</td>
<td>100%</td>
<td>87.06%</td>
<td>98.36%</td>
<td>68.25%</td>
<td>89.60%</td>
</tr>
<tr>
<td></td>
<td>shrub and brush rangeland</td>
<td>2091</td>
<td>2091</td>
<td>1252</td>
<td>1252</td>
<td>100%</td>
<td>100%</td>
<td>96.49%</td>
<td>95.93%</td>
<td>88.34%</td>
<td>90.34%</td>
</tr>
<tr>
<td></td>
<td>herbaceous rangeland</td>
<td>283</td>
<td>528</td>
<td>119</td>
<td>284</td>
<td>100%</td>
<td>85.92%</td>
<td>99.16%</td>
<td>80.63%</td>
<td>99.16%</td>
<td>81.69%</td>
</tr>
<tr>
<td></td>
<td>mixed barren</td>
<td>1052</td>
<td>260</td>
<td>571</td>
<td>196</td>
<td>100%</td>
<td>83.16%</td>
<td>98.60%</td>
<td>32.14%</td>
<td>95.62%</td>
<td>47.45%</td>
</tr>
<tr>
<td></td>
<td>transitional areas</td>
<td>2109</td>
<td>—</td>
<td>1529</td>
<td>—</td>
<td>90.12%</td>
<td>—</td>
<td>84.3%</td>
<td>—</td>
<td>81.03%</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>overall accuracy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>94.42%</td>
<td>97.98%</td>
<td>90.31%</td>
<td>92.54%</td>
<td>81.58%</td>
<td>86.95%</td>
</tr>
<tr>
<td></td>
<td>average accuracy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>96.26%</td>
<td>92.27%</td>
<td>93.12%</td>
<td>76.76%</td>
<td>86.48%</td>
<td>77.27%</td>
</tr>
<tr>
<td>Peking</td>
<td>urban</td>
<td>2042</td>
<td>3365</td>
<td>2026</td>
<td>3342</td>
<td>50.25%</td>
<td>84.32%</td>
<td>46.54%</td>
<td>71.99%</td>
<td>41.41%</td>
<td>64.33%</td>
</tr>
<tr>
<td></td>
<td>agricultural</td>
<td>1697</td>
<td></td>
<td>926</td>
<td>1689</td>
<td>93.52%</td>
<td>90.59%</td>
<td>88.55%</td>
<td>84.84%</td>
<td>75.81%</td>
<td>77.26%</td>
</tr>
<tr>
<td></td>
<td>barren land</td>
<td>2460</td>
<td>2446</td>
<td>2594</td>
<td>2122</td>
<td>93.02%</td>
<td>65.93%</td>
<td>82.88%</td>
<td>54.38%</td>
<td>85.00%</td>
<td>76.63%</td>
</tr>
<tr>
<td></td>
<td>rangeland</td>
<td>5243</td>
<td>3202</td>
<td>3409</td>
<td>3905</td>
<td>90.64%</td>
<td>89.04%</td>
<td>93.49%</td>
<td>89.40%</td>
<td>90.17%</td>
<td>78.44%</td>
</tr>
<tr>
<td></td>
<td>water</td>
<td>843</td>
<td>653</td>
<td>885</td>
<td>624</td>
<td>100%</td>
<td>96.15%</td>
<td>90.62%</td>
<td>95.35%</td>
<td>93.90%</td>
<td>97.28%</td>
</tr>
<tr>
<td></td>
<td>overall accuracy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>84.07%</td>
<td>84.10%</td>
<td>80.30%</td>
<td>77.72%</td>
<td>77.75%</td>
<td>74.91%</td>
</tr>
<tr>
<td></td>
<td>average accuracy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>85.49%</td>
<td>85.21%</td>
<td>80.42%</td>
<td>79.19%</td>
<td>77.26%</td>
<td>78.79%</td>
</tr>
</tbody>
</table>

- Application with up to 5 scales.
- **Accurate classification maps** with “Phoenix” (overall accuracy, OA ≈ 94 ÷ 98%; average accuracy, AA ≈ 92 ÷ 96%).
- **Quite accurate maps** with “Peking” (OA ≈ 84%, AA ≈ 85%): strong spectral overlapping between the classes (especially “urban vs. barren land”).
### Experimental Comparisons

<table>
<thead>
<tr>
<th>data set</th>
<th>class</th>
<th>training samples</th>
<th>test samples</th>
<th>proposed method</th>
<th>method in [6]</th>
<th>method in [8]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phoenix</td>
<td>urban</td>
<td>3545 4819</td>
<td>1329 1885</td>
<td>91.20% 100%</td>
<td>87.06% 98.36%</td>
<td>68.25% 89.60%</td>
</tr>
<tr>
<td></td>
<td>shrub and brush rangeland</td>
<td>2091 2091</td>
<td>1252 1252</td>
<td>100% 100%</td>
<td>96.49% 95.93%</td>
<td>88.34% 90.34%</td>
</tr>
<tr>
<td></td>
<td>herbaceous rangeland</td>
<td>283 528</td>
<td>119 284</td>
<td>100% 85.92%</td>
<td>99.16% 80.63%</td>
<td>99.16% 81.69%</td>
</tr>
<tr>
<td></td>
<td>mixed barren</td>
<td>1052 260</td>
<td>571 196</td>
<td>100% 83.16%</td>
<td>98.60% 32.14%</td>
<td>95.62% 47.45%</td>
</tr>
<tr>
<td></td>
<td>transitional areas</td>
<td>2109 —</td>
<td>1529 —</td>
<td>90.12% —</td>
<td>84.30% —</td>
<td>81.03% —</td>
</tr>
<tr>
<td></td>
<td>overall accuracy</td>
<td></td>
<td></td>
<td>94.42% 97.98%</td>
<td>90.31% 92.54%</td>
<td>81.58% 86.95%</td>
</tr>
<tr>
<td></td>
<td>average accuracy</td>
<td></td>
<td></td>
<td>96.26% 92.27%</td>
<td>93.12% 76.76%</td>
<td>86.48% 77.27%</td>
</tr>
<tr>
<td>Peking</td>
<td>urban</td>
<td>2042 3365</td>
<td>2026 3342</td>
<td>50.25% 84.32%</td>
<td>46.54% 71.99%</td>
<td>41.41% 64.33%</td>
</tr>
<tr>
<td></td>
<td>agricultural</td>
<td>2460 1697</td>
<td>926 1689</td>
<td>93.52% 90.59%</td>
<td>88.55% 84.84%</td>
<td>75.81% 77.26%</td>
</tr>
<tr>
<td></td>
<td>barren land</td>
<td>2460 2446</td>
<td>2594 2122</td>
<td>93.02% 65.93%</td>
<td>82.88% 54.38%</td>
<td>85.00% 76.63%</td>
</tr>
<tr>
<td></td>
<td>rangeland</td>
<td>5243 3202</td>
<td>3409 3905</td>
<td>90.64% 89.04%</td>
<td>93.49% 89.40%</td>
<td>90.17% 78.44%</td>
</tr>
<tr>
<td></td>
<td>water</td>
<td>843 653</td>
<td>885 624</td>
<td>100% 96.15%</td>
<td>90.62% 95.35%</td>
<td>93.90% 97.28%</td>
</tr>
<tr>
<td></td>
<td>overall accuracy</td>
<td></td>
<td></td>
<td>84.07% 84.10%</td>
<td>80.30% 77.72%</td>
<td>77.75% 74.91%</td>
</tr>
<tr>
<td></td>
<td>average accuracy</td>
<td></td>
<td></td>
<td>85.49% 85.21%</td>
<td>80.42% 79.19%</td>
<td>77.26% 78.79%</td>
</tr>
</tbody>
</table>

- More accurate results by the method in [8] (MRF fusion of spatio-temporal context) than by the one in [6] (modeling only temporal correlation).
- More accurate results by the proposed method than by both previous techniques. This suggests the effectiveness of the method in fusing spatial, temporal, and multiscale information for region-based multitemporal classification.
Outline

• Introduction
  – Multitemporal high-resolution image classification

• The proposed method
  – Overall architecture
  – Multitemporal multiscale region-based Markov random field
  – Parameter estimation and energy minimization

• Experimental results
  – With QuickBird images
  – With SPOT-5 HRG images

• Conclusion
Conclusion

• A novel multitemporal classifier has been proposed for HR images, based on a region-based multiscale MRF.
  - Capability to fuse spatio-temporal context and multiscale information associated to a multitemporal HR data set.
  - Accurate results with different sensors (QuickBird, SPOT-5) and resolutions (2.8 m, 10 m).
  - Accuracy improvements, compared to previous multitemporal methods based on temporal or spatio-temporal models.
  - Confirmation of the relevance of multiscale information in HR image analysis.

• Possible future generalizations
  - Integrating edge and/or texture information.
  - Approaching global energy minimization (e.g., graph-cuts).
  - Comparisons with other techniques for multitemporal VHR classification
References

1. S. Li, Markov random field modeling in image analysis, Springer, 2009


