Online, self-supervised terrain classification via discriminatively trained submodular Markov random fields

Paul Vernaza, Ben Taskar, Daniel D. Lee

GRASP Lab
University of Pennsylvania

May 23, 2008
1. Discriminative MRFs for image classification

2. Efficient inference and learning algorithms

3. Experimental results
Semi-supervised terrain classification

- Goal: Learn appearance of obstacles, ground from stereo classification
- Real-time implementation on a mobile robot for long-range planning

Paul Vernaza, Ben Taskar, Daniel D. Lee
Stereo classification

- Obstacles identified by disparity deviation from ground plane
- Ground plane estimated via iteratively reweighted least squares (EM-like algorithm)
Independent-pixel classification

- Independent classification is easy, but suffers from tunnel vision
Relaxing spatial independence via MRF modeling

\[
\begin{align*}
\phi_1(1) &= 1 \\
\phi_1(0) &= 100 \\
\phi_2(1) &= 50 \\
\phi_2(0) &= 2 \\
\phi_{1,2}(0,0) &= 10 \\
\phi_{1,2}(1,1) &= 20 \\
\phi_{1,2}(0,1) &= 0 \\
\phi_{1,2}(1,0) &= 0
\end{align*}
\]

\[
\text{graph of pixels (2-pixel example)}
\]

\[
\text{Probability}(x) \propto \phi(x) = \prod_i \phi_i(x_i) \prod_{<i,j>} \phi_{ij}(x_i, x_j), \ x_i \in \{0, 1\}
\]

- Submodular MRF assigns high probability to labelings \( x \) that maximize product of affinities (global satisfaction measure)
Discriminatively-trained MRFs

\[
\begin{align*}
\log \phi_i(0) &= -w_0^T f_i \\
\log \phi_i(1) &= -w_1^T f_i \\
\log \phi_{ij}(0,0) &= -w_{00} f_{ij} \\
\log \phi_{ij}(1,1) &= -w_{11} f_{ij}
\end{align*}
\]

(nonnegative features \( f \))

\[
\begin{align*}
\min_{x, w} \quad & \|w\|^2 \\
\text{subject to} \quad & \text{Energy}(x, w, f) - \text{Energy}(\hat{x}, w, f) \geq \text{Hamming}(x, \hat{x}) \\
\end{align*}
\]

\((\text{Energy} = -\log(\phi(x)), \hat{x} = \text{desired labeling})\)

- (Taskar 2005) Learn \( w \) to make \( \hat{x} \) mode of MRF (Max Margin MRF)
- Discriminative training sidesteps issue of intractable inference; admits convex formulations, optimization methods
MPE inference problem

\[
\begin{align*}
\min_{p,x} & \quad \sum_{ij \neq ts} x_{ij} \phi_{ij} \\
\text{subject to} & \quad p_i - p_j \leq x_{ij}, \forall ij \in \mathcal{E} \\
& \quad p_t - p_s \leq -1 \\
& \quad 0 \leq x_{ij} \leq 1
\end{align*}
\]

\[
\begin{align*}
\max_f & \quad f_{ts} \\
\text{subject to} & \quad \sum_{ij \in \mathcal{E}} f_{ij} = \sum_{ji \in \mathcal{E}} f_{ji}, \forall j \in \mathcal{N} \\
& \quad 0 \leq f_{ij} \leq \phi_{ij}, \forall ij \in \mathcal{E} \setminus ts
\end{align*}
\]

- Given MRF weights, find the most probable binary MRF assignment
- Can be formulated as linear program with *totally unimodular* constraints
MPE inference via graph cuts

(Kolmogorov and Zabih 2004) Graph cut allows solution of combinatorial inference problem in polynomial time
Learning: intuition

- Learning possible via large quadratic program
- Alternative: iteratively adjust weights to create/suppress bottlenecks where necessary
Learning as subgradient optimization

\[
\begin{align*}
\min_{w \geq 0} & \quad ||w||^2 + C \left( N_n + \sum_{ij \in \mathcal{E}} (w^T f_{ij}) \hat{x}_{ij} \right) - \\
& \min_{x \in Q} \sum_{ij \in \mathcal{E}} (w^T f_{ij} + \hat{x}_{ij} (\delta_{is} + \delta_{jt})) x_{ij}
\end{align*}
\]

- Intuitive procedure is actually (almost) a projected subgradient method on the MMM objective
- Iteration (feature matrix F, desired cut \(\hat{x}\)):
  \[
  w \leftarrow (w - \eta (2w + F^T (\hat{x} - \tilde{x})))_+
  \]
  \(\tilde{x}\) is cut with respect to “adversarially tweaked” weights to ensure margin (robustness criterion)
- Similar to Max Margin Planning (Ratliff et. al 2006)
Feature selection

- Node features: “bin indicators”; equivalent to histogram classifier when edge features are ignored
- Edge features: concatenation of node features
Summary of method

- For each image...
  1. Calculate near-field obstacle/ground segmentation from stereo
  2. Calculate image features for classification
  3. Use stereo segmentation to incrementally train MRF-based classifier on given features
  4. Evaluate classifier on the current image
Experimental details

- MRF classifier compared against histogram (independent) classifier on stereo log data
- Graph cut implemented with Boykov and Kolmogorov’s publicly available code
- Runs on LAGR robot
- Source code available
**Batch training experiments**

- **Easy dataset**
- **Hard dataset**

<table>
<thead>
<tr>
<th></th>
<th>% Correct Histogram</th>
<th>% Correct MRF</th>
<th>% Correct MRF - Histogram</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Easy” training set</td>
<td>88.40</td>
<td>94.85</td>
<td>6.46</td>
</tr>
<tr>
<td>“Hard” training set</td>
<td>72.47</td>
<td>84.99</td>
<td>12.52</td>
</tr>
<tr>
<td>“Easy” held-out set</td>
<td>83.50</td>
<td>92.21</td>
<td>8.71</td>
</tr>
<tr>
<td>“Hard” held-out set</td>
<td>68.26</td>
<td>79.76</td>
<td>11.50</td>
</tr>
</tbody>
</table>

- MRF classifier significantly outperforms independent classifier

Paul Vernaza, Ben Taskar, Daniel D. Lee
Online experiments

Images randomly permuted, one subgradient step performed per image
As in perceptron, error rate converges to a low value
## Computational efficiency

<table>
<thead>
<tr>
<th></th>
<th>Subgradient calculation (ms.)</th>
<th>Graph cut inference (ms.)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean</strong></td>
<td>37.2</td>
<td>7.6</td>
</tr>
<tr>
<td><strong>Standard deviation</strong></td>
<td>5.1</td>
<td>2.2</td>
</tr>
</tbody>
</table>

- Statistics for 160x120 images, 2.4 GHz P4 Xeon
- Computation time scales roughly linearly with number of pixels
Histogram vs. MRF

Video
Conclusions

- Context significantly boosts image classification performance
- Learning and prediction possible in real-time despite combinatorial nature of problem
- Future work: non-binary classification, learning initial oversegmentation