Inference with Higher-order Models

Pushmeet Kohli
Microsoft Research
Binary Image Segmentation

$E(x)$

$x \in \{0, 1\}^n$

Image (D)

[Boykov and Jolly '01] [Blake et al. '04]
Binary Image Segmentation

\[ E(x) = \sum c_i x_i \]

Pixel Colour

Unary Cost \( (c_i) \)
Dark \( (Bg) \)    Bright \( (Fg) \)

\[ x^* = \arg \min E(x) \]

\[ x \in \{0,1\}^n \]

[Boykov and Jolly '01] [Blake et al. '04]
Binary Image Segmentation

\[ E(x) = \sum c_i x_i + \sum d_{ij} |x_i - x_j| \]

Pixel Colour  Smoothness Prior

Unary Cost \((c_i)\)

Dark (Bg)  Bright (Fg)

x in \(\{0,1\}^n\)

[Boykov and Jolly '01] [Blake et al. '04]
Binary Image Segmentation

\[ E(x) = \sum c_i x_i + \sum d_{ij} |x_i - x_j| \]

Pixel Colour + Smoothness Prior

How to minimize \( E(x) \)?

\[ x^* = \text{arg min } E(x) \]

Old Solution

[Boykov and Jolly '01] [Blake et al. '04]
Energy Minimization Problems

Space of Problems

NP-Hard

• MAXCUT

CSP

•
Energy Minimization Problems

Perfect Graphs, Low-tree width, Submodular functions, Structured decomposable functions ...

Space of Problems

Tractable Problems

CSP

MAXCUT

NP-Hard

March 2014

Tractability: Practical Approaches to Hard Problems
Edited by Lucas Bordeaux, Youssef Hamadi and Pushmeet Kohli

Cambridge
Energy Minimization Problems

Tractability Properties

Structural Tractability

NP-Hard
- MAXCUT

CSP

Tree Structured

Space of Problems
Energy Minimization Problems

Tractability Properties

Constraints on the terms of your energy functions
\[ d_{ij} |x_i - x_j| \]

Language or Form Tractability

Space of Problems
So how does this work?

Pairwise Submodular

E(x)

S

T

st-mincut

Solution

[Hammer, 1965] [Kolmogorov and Zabih, 2002]
Demo
.. So what are the challenges?
Modelling Challenges
Does not enforce connectivity

Short boundary bias

Cannot enforce priors on label counts
Examples of Higher order Models

- Taskar et al. 02 – associative potentials
- Roth & Black 05 – field of experts
- Kohli Kumar Torr. 07 – segment consistency
- Kohli Ladicky Torr 08 – segment consistency
- Woodford et al. 08 – planarity constraint
- Vicente et al. 08 – connectivity constraint
- Nowozin & Lampert 09 – connectivity constraint
- Ladický et al. 09 – consistency over several scales
- Woodford et al. 09 – marginal probability
- Delong et al. 10 – label occurrence costs
- Ladicky et al. 10 – label set co-occurrence costs
- Jegelka and Bilmes 11 – label set co-occurrence costs
- ... many others
Different Approaches

- Transformation schemes
  [Kolmogorov’02] [Kohli et al.07, 08,09]

- Constrained Inference using Parametric Mincuts
  [Kolmogorov’07] [Lim et al.08, 14]

- Decomposition Techniques
  [Woodford et al. ’09] [Komodakis’ 09] ..

- Iterative refinement of constraints (Connectivity and Bounding Box Potentials)
  [Nowozin and Lampert ‘08, Lempitsky et al.’09] ..

- Special purpose message computation
  [Gupta and Sarawagi ’07, 08] [Tarlow et al. 09] ..

- Learning to preserve higher-order statistics
  [Pletscher and Kohli ‘12]
Different Approaches

- Transformation schemes
  
  [Kolmogorov'02] [Kohli et al.07, 08,09]

- Constrained Inference using Parametric Mincuts
  
  [Kolmogorov'07] [Lim et al.08, 14]

- Decomposition Techniques
  
  [Woodford et al. ‘09] [Komodakis‘ 09].

- Iterative refinement of constraints (Connectivity and Bounding Box Potentials)
  
  [Nowozin and Lampert ‘08, Lempitsky et al. ’09].

- Special purpose message computation
  
  [Gupta and Sarawagi ‘07, 08] [Tarlow et al. 09].

- Learning to preserve higher-order statistics
  
  [Pletscher and Kohli ‘12]
Approach 1: Transforming higher order potentials

Ensure tractability of the transformed problem
Labelling Consistency in Pixel Groups

Unary Potentials

[Shotton et al. ECCV 2006]

Colour, Location & Texture

Higher Order Energy

- Sky
- Tree
- Building
- Grass

Pairwise Smoothness Potentials

Higher Order Potentials
(Defined using multiple Segmentations)

 Pixels belonging to a group should take the same label

$$h(X_p) = \begin{cases} 
0 & \text{if } x_i = L, I \in p \\
C & \text{otherwise}
\end{cases}$$
Unary Potentials [Shotton et al. ECCV 2006]
Colour, Location & Texture

Higher Order Potentials (Defined using multiple Segmentations)

Labelling Consistency in Pixel Groups

Higher Order Energy

Sky + Tree
Building + Grass

Pairwise Smoothness Potentials
Unary Potentials
[Shotton et al. ECCV 2006]

Colour, Location & Texture

Higher Order Energy

Sky

Tree

Building

Grass

Unary Potential

Pairwise Smoothness Potentials

Higher Order Potentials
(Defined using multiple Segmentations)

Energy Minimization

Segmentation Solution
Example

$$h(X_p) = \begin{cases} 0 & \text{if all } x_i = \text{"Tree"}(0), I \in p \\ 1 & \text{otherwise} \end{cases}$$

Pixels belonging to the group $p$ should take the same label “tree”
Transforming higher order potentials

\[ f(x) = \begin{cases} 
0 & \text{if all } x_i = 0 \\
1 & \text{otherwise}
\end{cases} \]
Transforming higher order potentials

\[
f(x) = \begin{cases} 
0 & \text{if all } x_i = 0 \\
1 & \text{otherwise}
\end{cases}
\]
Transforming higher order potentials

\[ \min_x f(x) = \min_{x,a \in \{0,1\}} a + \bar{a} \sum x_i \]

Higher Order Submodular Function

Quadratic Submodular Function

\[ f(x) = \begin{cases} 
0 & \text{if all } x_i = 0 \\
1 & \text{otherwise} 
\end{cases} \]
More Results

Image (MSRC-21)  
Pairwise CRF  
Higher order CRF  
Ground Truth

Sheep
Grass

[Runner-Up, PASCAL VOC 2008]
Another Example
Overcoming short-boundary bias

\[ E(\mathbf{x}) = \sum c_i x_i + \sum d_{ij} |x_i - x_j| \]

Encourages short boundaries

Penalize types of boundaries not the actual number of boundaries!

Cooperative Cuts [Jegelka and Bilmes, CVPR 2011]
Overcoming short-boundary bias

$$E(X) = \sum c_i x_i + \sum d_i |x_i - x_j|$$
Overcoming short-boundary bias

\[ E(X) = \sum c_i x_i + \sum d_{ij} |x_i - x_j| + \sum_{g \in G} h_g(X) \]

- Divide edges into different types
- Incorporate a higher order consistency potential over the edges

\[ h_g(X) = F \left( \sum_{ij \in g} |x_i - x_j| \right) \]
Overcoming short-boundary bias

\[ E(X) = \sum c_i x_i + \sum d_{ij} |x_i - x_j| + \sum h_g(X_p) \]

- Divide edges into different types
- Incorporate a higher order consistency potential over the edges

\[ h_g(X_p) = F \left( \sum_{ij \in g} |x_i - x_j| \right) \]
Overcoming short-boundary bias

\[ E(X) = \sum c_i x_i + \sum d_{ij} |x_i - x_j| + \sum_{g \in G} h_g(X_p) \]

- Divide edges into different types
- Incorporate a higher order consistency potential over the edges

\[ h_g(X_p) = F (\sum_{ij \in g} |x_i - x_j|) \]
Transform Higher order model

\[
H_g(Y) = \min \left\{ \sum_{(ij) \in g} d_{ij} |y_i - y_j|, \quad T \right\}
\]

\[
H_g(Y) = \min_{h_g \in \{0, 1\}} h_g \sum d_{ij} (y_i + y_j - 2y_i y_j) + T(1 - h_g)
\]
Transformation
So what happens to the results?
Results – Interactive Segmentation

Image

Ground Truth

Pairwise Model
Results – Semantic Segmentation

<table>
<thead>
<tr>
<th>Image</th>
<th>Ground Truth</th>
<th>Pairwise Model</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image.jpg" alt="Image" /></td>
<td><img src="ground_truth.png" alt="Ground Truth" /></td>
<td><img src="pairwise_model.png" alt="Pairwise Model" /></td>
</tr>
<tr>
<td><img src="image2.jpg" alt="Image" /></td>
<td><img src="ground_truth2.png" alt="Ground Truth" /></td>
<td><img src="pairwise_model2.png" alt="Pairwise Model" /></td>
</tr>
<tr>
<td><img src="image3.jpg" alt="Image" /></td>
<td><img src="ground_truth3.png" alt="Ground Truth" /></td>
<td><img src="pairwise_model3.png" alt="Pairwise Model" /></td>
</tr>
</tbody>
</table>

$\epsilon = 3.00\%$

$\epsilon = 2.39\%$

$\epsilon = 5.12\%$
Approach 2: Constrained Inference
Lower Envelope Representation

\[
\min_{x} f(x) = \min_{x,a \in \{0,1\}} f_1(x) a + f_2(x) (1-a)
\]

Lower envelope of concave functions is concave

Higher Order Submodular Function

Quadratic Submodular Function

[ Kohli et al. '08 ]
Upper Envelopes

\[
\min_x f(x) = \min_{x \in \{0,1\}} \max_{a \in \{0,1\}} f_1(x)a + f_2(x)(1-a)
\]

Upper envelope of linear functions is convex

[Very Hard Problem!!!!] [Kohli and Kumar, CVPR 2010]
Why Upper Envelopes?

SILHOUETTE CONSTRAINTS
[Sinha et al. ‘05, Cremers et al. ‘08]

3D RECONSTRUCTION
Rays must touch silhouette at least once

SIZE/VOLUME PRIORS
[Woodford et al. 2009]

BINARY SEGMENTATION
Prior on size of object (say n/2)

[Kohli and Kumar, CVPR 2010]
\[ x^* = \underset{x}{\arg \min} \ E(x) \]

Such that:

- \( x \) has area \( A \)
  \[ \sum_{i \in V} x_i = A \]
- \( x \) has boundary length \( B \)
  \[ \sum_{(i,j) \in E} |x_i - x_j| = B \]
Approach 2: Constrained Inference

Parametric Maxflow/st-Mincut

\[ H^\lambda(x) = E(x) + \lambda \sum_{v \in V} x_v \]

\[ x^* = \arg \min H^\lambda(x) \]

\[ \sum x_v = 0 \quad \lambda = +\infty \]

\[ \sum x_v = n \quad \lambda = -\infty \]

\[ \lambda = +\infty \]
Approach 2: Constrained Inference

Parametric Maxflow/st-Mincut

\[ H^\lambda(x) = E(x) + \lambda \sum_{v \in V} x_v \]

Lemma 1. If an assignment \( x \) minimizes the energy function \( H^\lambda \) for some \( \lambda \), \( E(x) \) is minimum under the same label count as \( x \).

[Gallo et al. 1986] [Kolmogorov et al. ICCV 2007]
Approach 2: Constrained Inference

Generalized (Multi-dimensional) Parametric Maxflow
Approach 3: Learning
Low-order Models for Enforcing Higher-order Statistics

Can we find parameters that lead to solutions consistent with higher-order statistics?

\[ E(x, y, w) = w^T \psi(y, x) \]

[Data: \{x^k, y^k\}, \ k = 1..K]

[Pletscher and Kohli, AISTATS 2012]
Low-order Models for Enforcing Higher-order Statistics

\[ E(x, y, w) = w^\top \psi(y, x) \]

Data: \( \{x^k, y^k\}, \ k = 1..K \)

\[
\begin{align*}
\min_{\xi \geq 0, w, U} & \quad \sigma(w, U) := \frac{1}{2} \|w\|^2 \\
\text{s.t.} & \quad w^\top [\psi(x^k, y) - \psi(x^k, y^k)] \geq \sum_{i \in \mathcal{V}} y_i - y_i^* 
\end{align*}
\]

[AISTATS 2012]
Low-order Models for Enforcing Higher-order Statistics

\[ E(x, y, w) = w^T \psi(y, x) \]

\[
\min_{\xi \geq 0, w, U} \quad \sigma(w, U) := \frac{1}{2} \|w\|^2 \\
\text{s.t.} \quad w^T [\psi(x^k, y) - \psi(x^k, y^k)] \geq \left| \sum_{i \in V} y_i - \sum_{i \in V^*} y_i^* \right|
\]

Requires solution of

\[
\min_y E(y, x, w) - \left| \sum_{i \in V} y_i - \sum_{i \in V^*} y_i^* \right|
\]

[AISTATS 2012]
Learning Higher-order Model for Enforcing Low-order statistics

<table>
<thead>
<tr>
<th>Train</th>
<th>Eval</th>
<th>Hamming better (%)</th>
<th>Count better (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4/S</td>
<td>Hamming</td>
<td>52.1 ± 7.0</td>
<td>47.9 ± 7.0</td>
</tr>
<tr>
<td></td>
<td>Count</td>
<td>33.8 ± 8.3</td>
<td>66.2 ± 8.3</td>
</tr>
<tr>
<td>4/D</td>
<td>Hamming</td>
<td>39.4 ± 6.1</td>
<td>60.6 ± 6.1</td>
</tr>
<tr>
<td></td>
<td>Count</td>
<td>29.6 ± 8.3</td>
<td>70.4 ± 8.3</td>
</tr>
</tbody>
</table>
Challenges and Opportunities

- Adaptive data-driven representations for Higher order Potentials
- Global potentials that encode topology constraints
- Efficiency

Thanks for listening. Questions?