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

$$E(\mathbf{x}) = \boxed{\sum c_i x_i}$$

Pixel Colour

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

Unary Cost (c_i)

Dark (Bg)   Bright (Fg)

$x^* = \arg\min E(x)$

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

Binary Image Segmentation

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

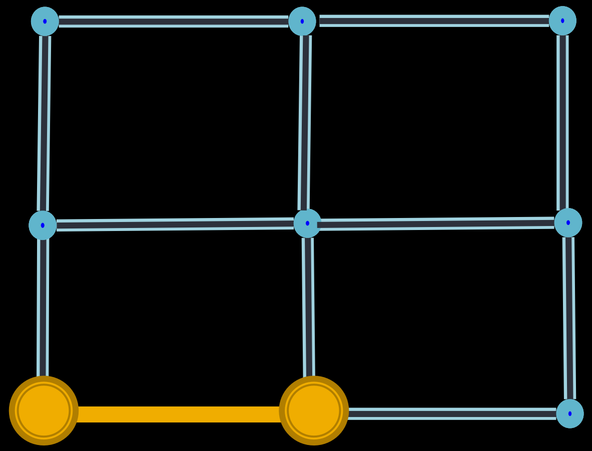
$\mathbf{x}$ in $\{0,1\}^n$

Pixel Colour

Smoothness Prior

Unary Cost (c_i)

Dark (Bg) Bright (Fg)



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

Binary Image Segmentation

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

Pixel Colour　　**Smoothness Prior**

x in {0,1}^{n}

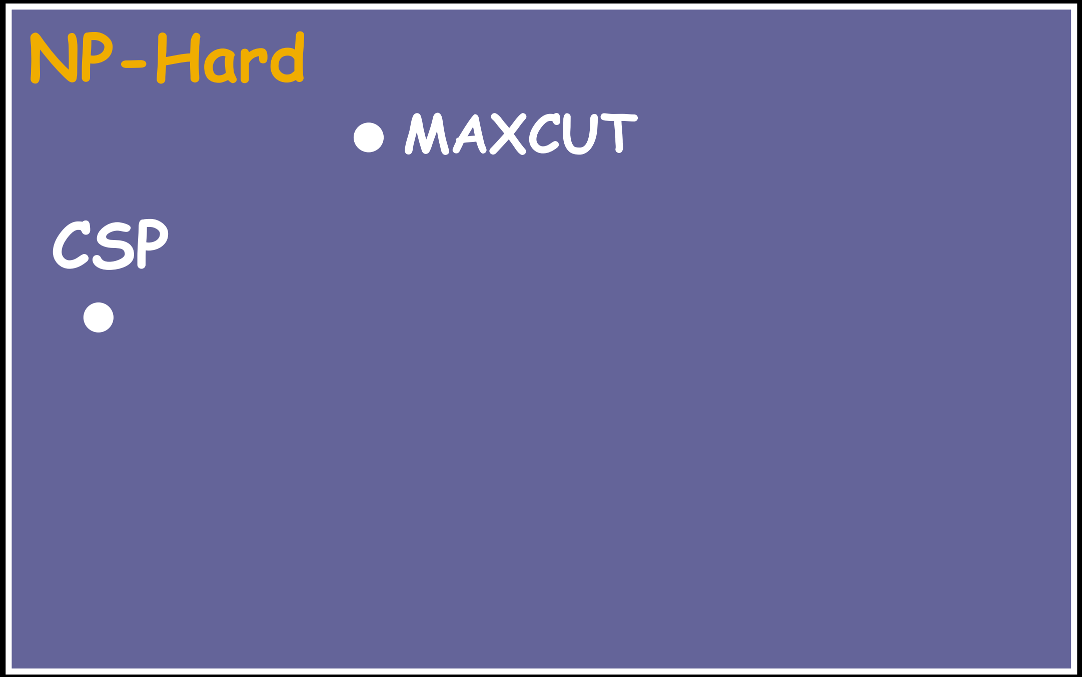
How to minimize E(x)?

x* = arg min E(x)

Old Solution

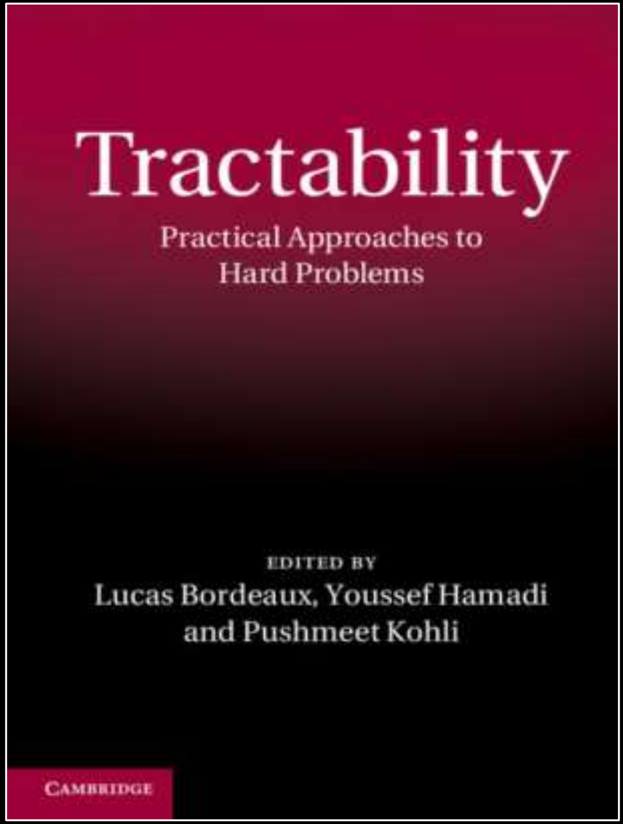
[Boykov and Jolly ' 01] [Blake et al. '04]

Energy Minimization Problems



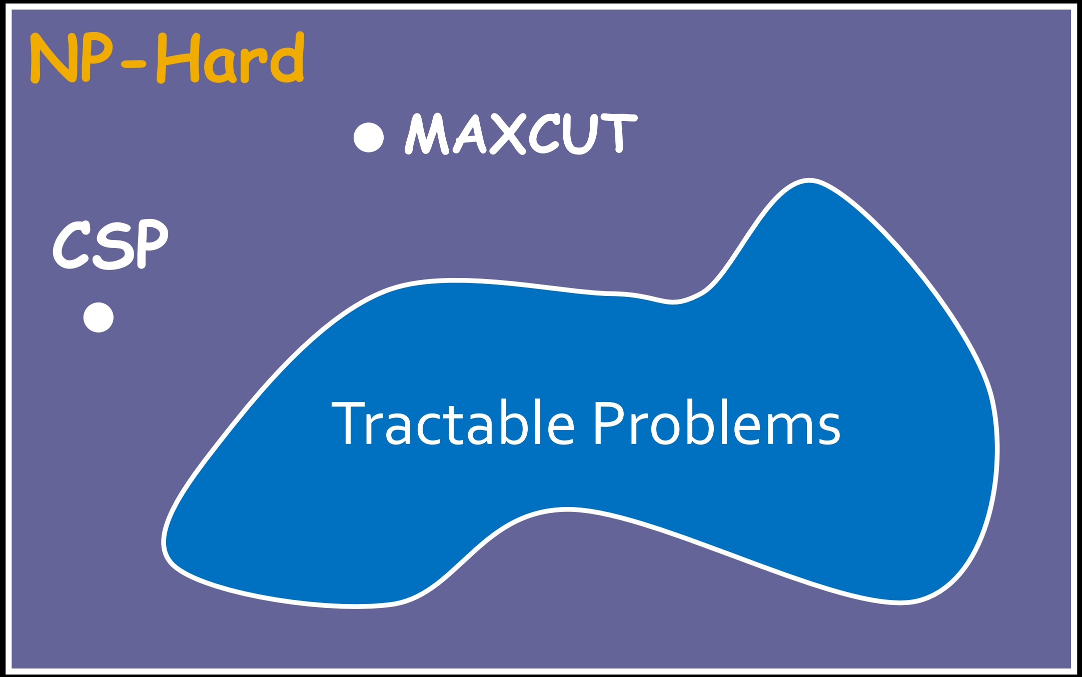
Space of Problems

Energy Minimization Problems



March 2014

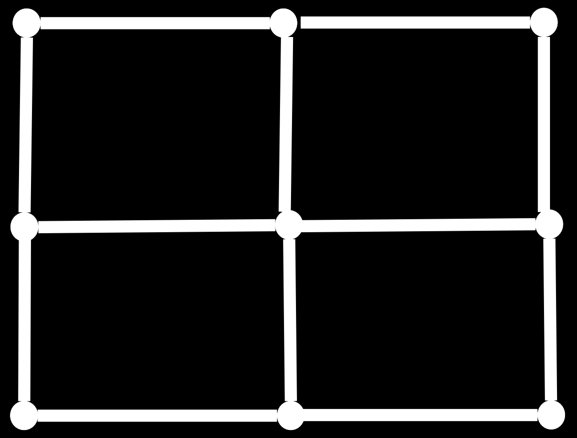
Perfect Graphs, Low-tree width, Submodular functions,
Structured decomposable functions …



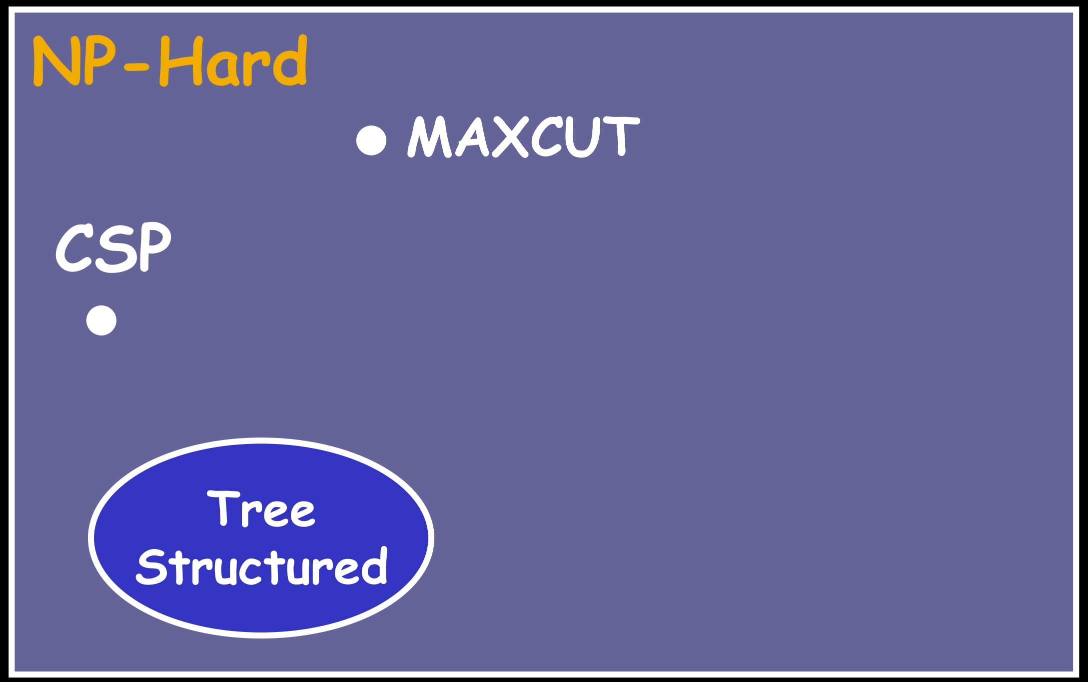
Space of Problems

Energy Minimization Problems

Tractability Properties



Structural
Tractability



Space of Problems

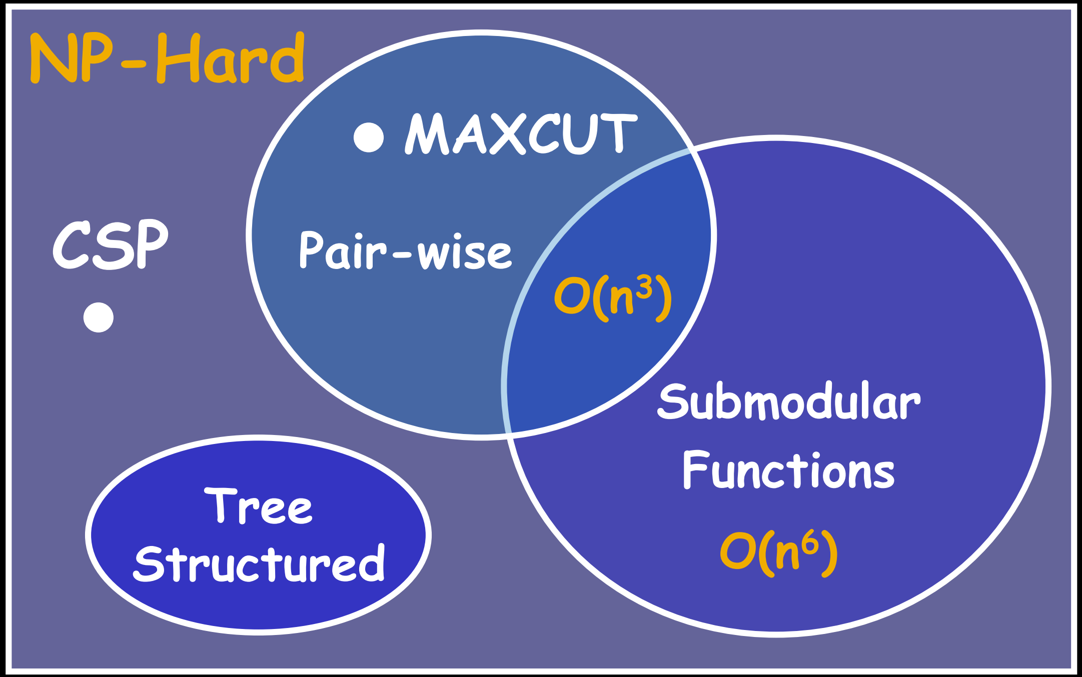
Energy Minimization Problems

Tractability Properties

Constraints on the terms of your energy functions

$d_{ij} \, |x_i - x_j|$

Language or Form Tractability

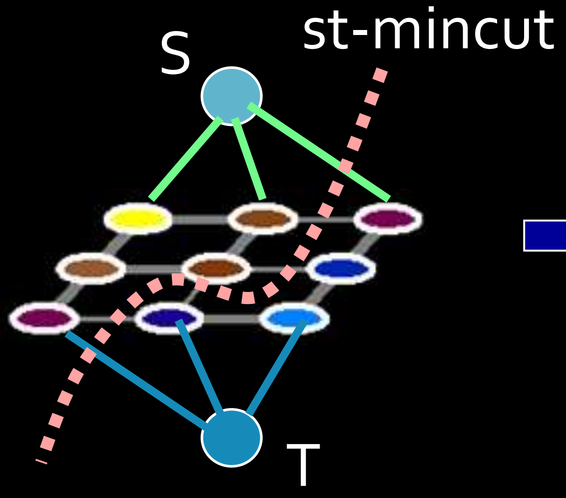


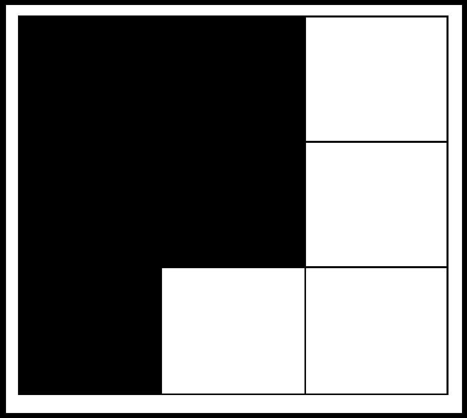
Space of Problems

So how does this work?

E(x)

Pairwise Submodular





Solution

[Hammer, 1965] [Kolmogorov and Zabih, 2002]

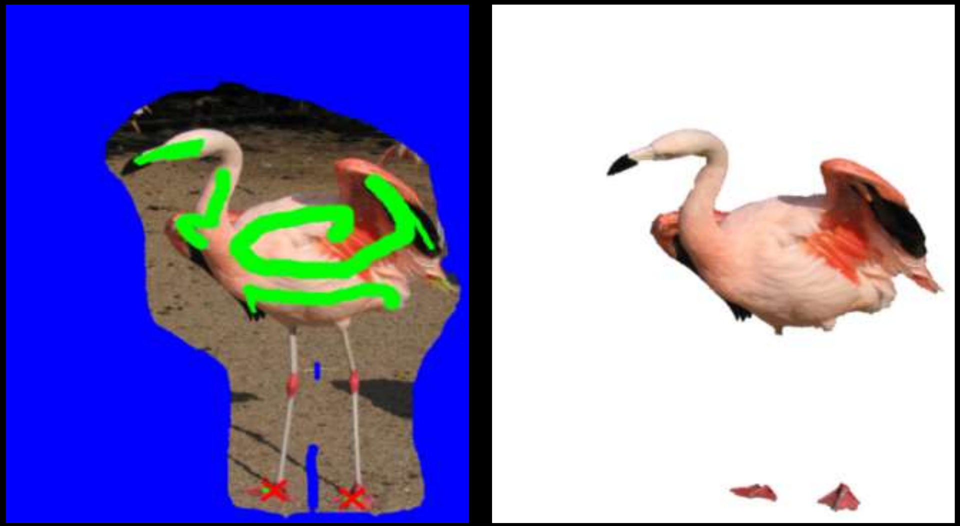
Demo

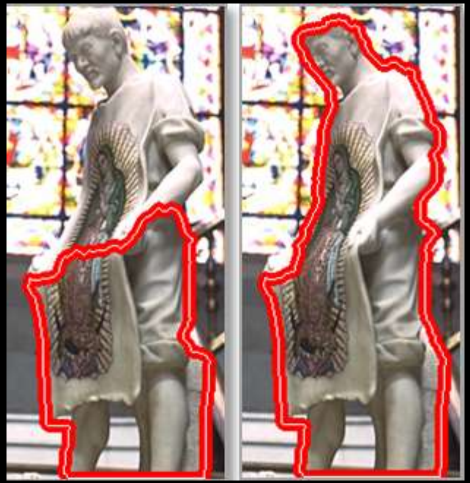
.. So what are the challenges?

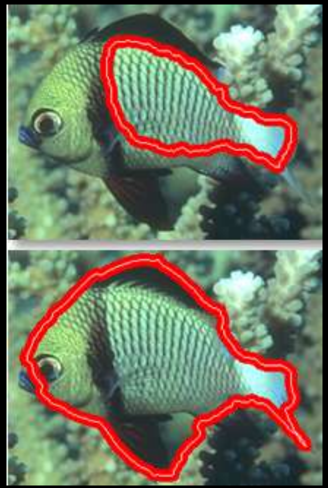
Modelling Challenges

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] ..

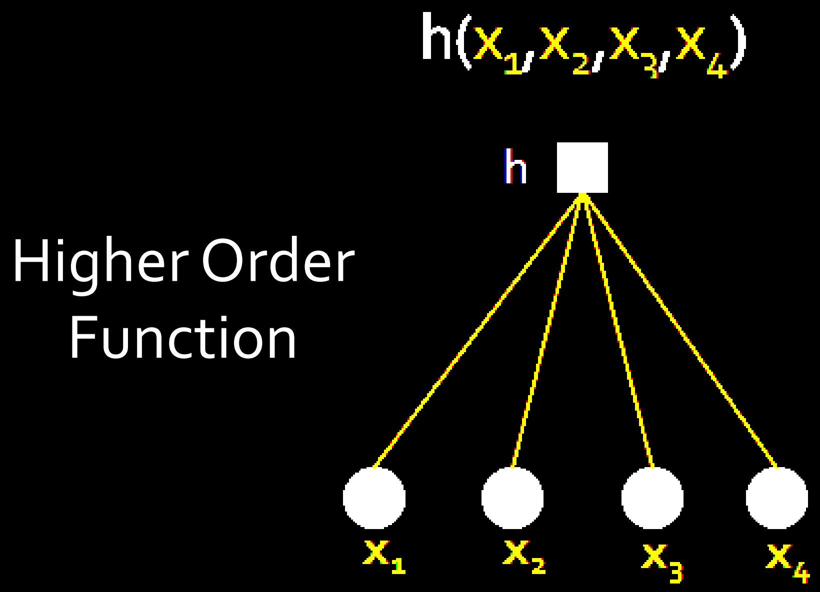
- Special purpose message computation

  [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

Example

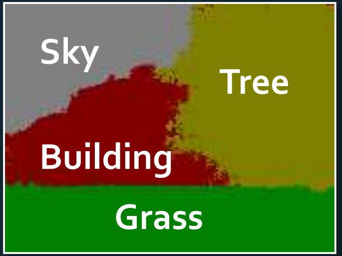
Labelling Consistency in Pixel Groups

Unary Potentials
[Shotton et al. ECCV 2006]

Colour, Location & Texture

Higher Order Potentials
(Defined using multiple Segmentations)

Higher Order Energy

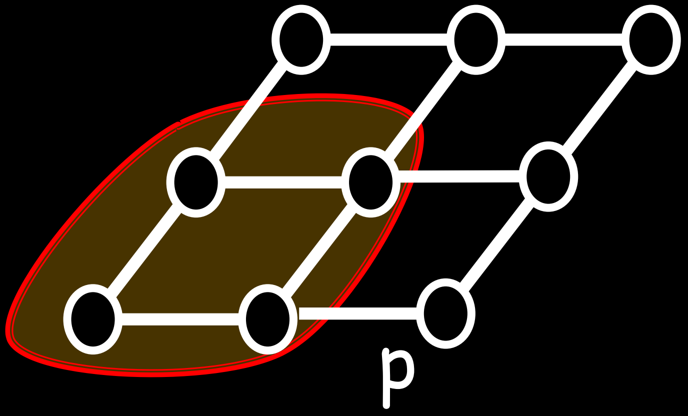


+

Pairwise Smoothness Potentials

+

Pixels belonging to a group should take the same label



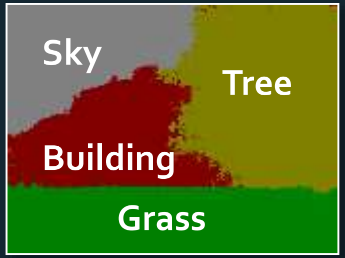
$$h(X_p) = \begin{cases} 0 & \text{if } x_i = L, I \in p \\ C & \text{otherwise} \end{cases}$$

Labelling Consistency in Pixel Groups

Unary Potentials
[Shotton et al. ECCV 2006]

Colour, Location & Texture

Higher Order Energy



+

Pairwise Smoothness Potentials

+

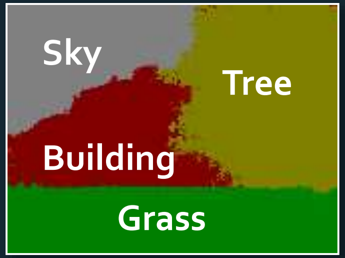
Higher Order Potentials
(Defined using multiple Segmentations)

Labelling Consistency in Pixel Groups

Unary Potentials
[Shotton et al. ECCV 2006]

Colour, Location & Texture

Higher Order Energy



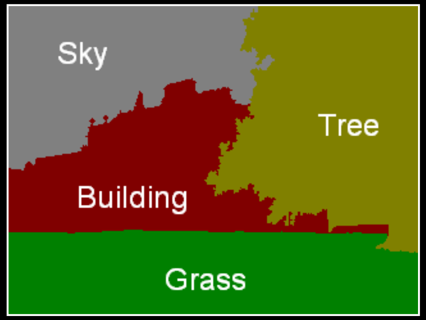
+

Pairwise Smoothness Potentials

+

Higher Order Potentials
(Defined using multiple Segmentations)

Energy Minimization

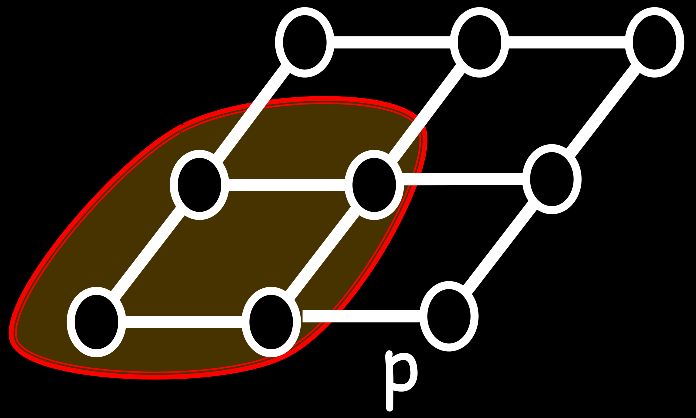


Segmentation Solution

Example

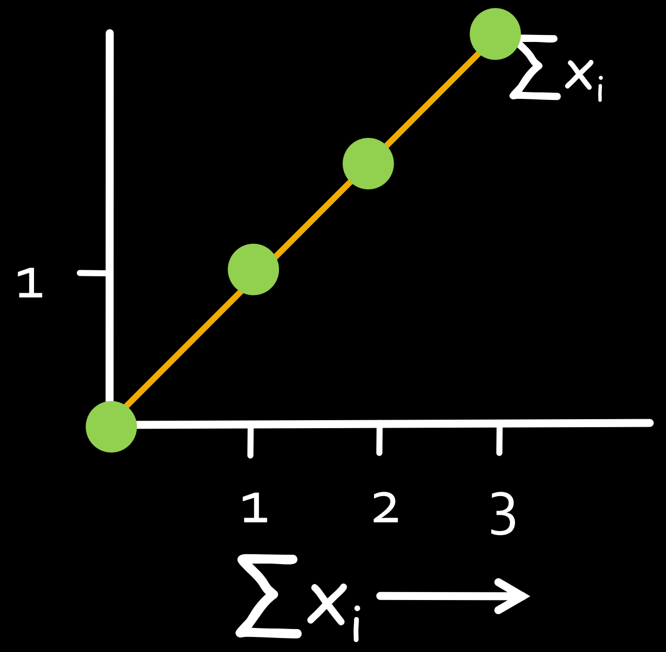
$$h(X_p) = \begin{cases} 0 & \text{if all } x_i = \text{"Tree"}(0),\ I \in p \\ 1 & \text{otherwise} \end{cases}$$

Pixels belonging to the group p should take the same label "tree"



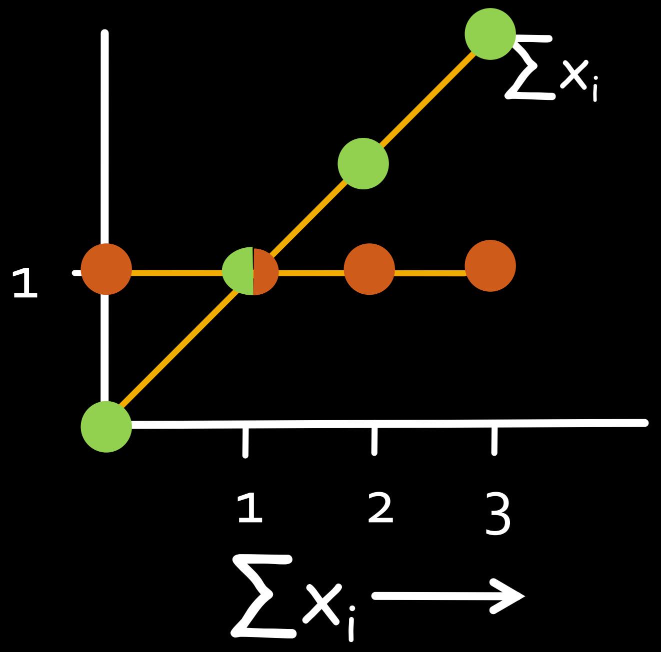
Transforming higher order potentials

$$f(x) = \begin{cases} 0 & \text{if all } x_i = 0 \\ 1 & \text{otherwise} \end{cases}$$



Transforming higher order potentials

$$f(x) = \begin{cases} 0 & \text{if all } x_i = 0 \\ 1 & \text{otherwise} \end{cases}$$



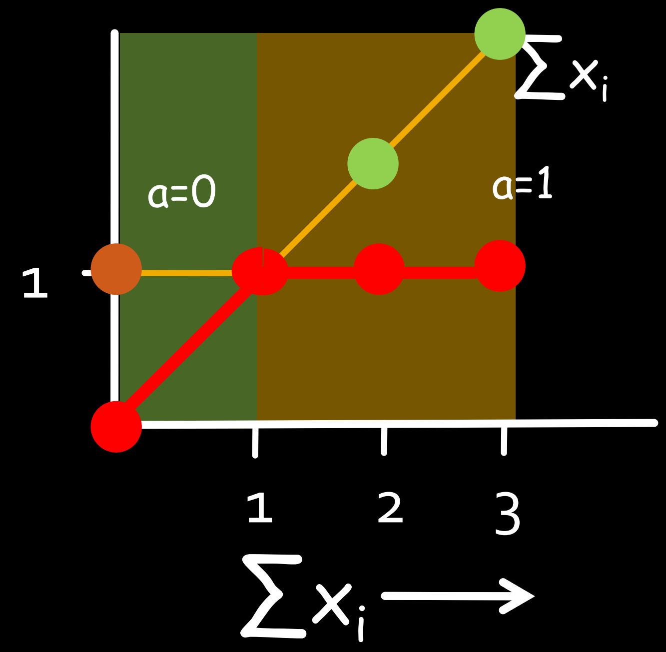
Transforming higher order potentials

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

Higher Order Submodular Function

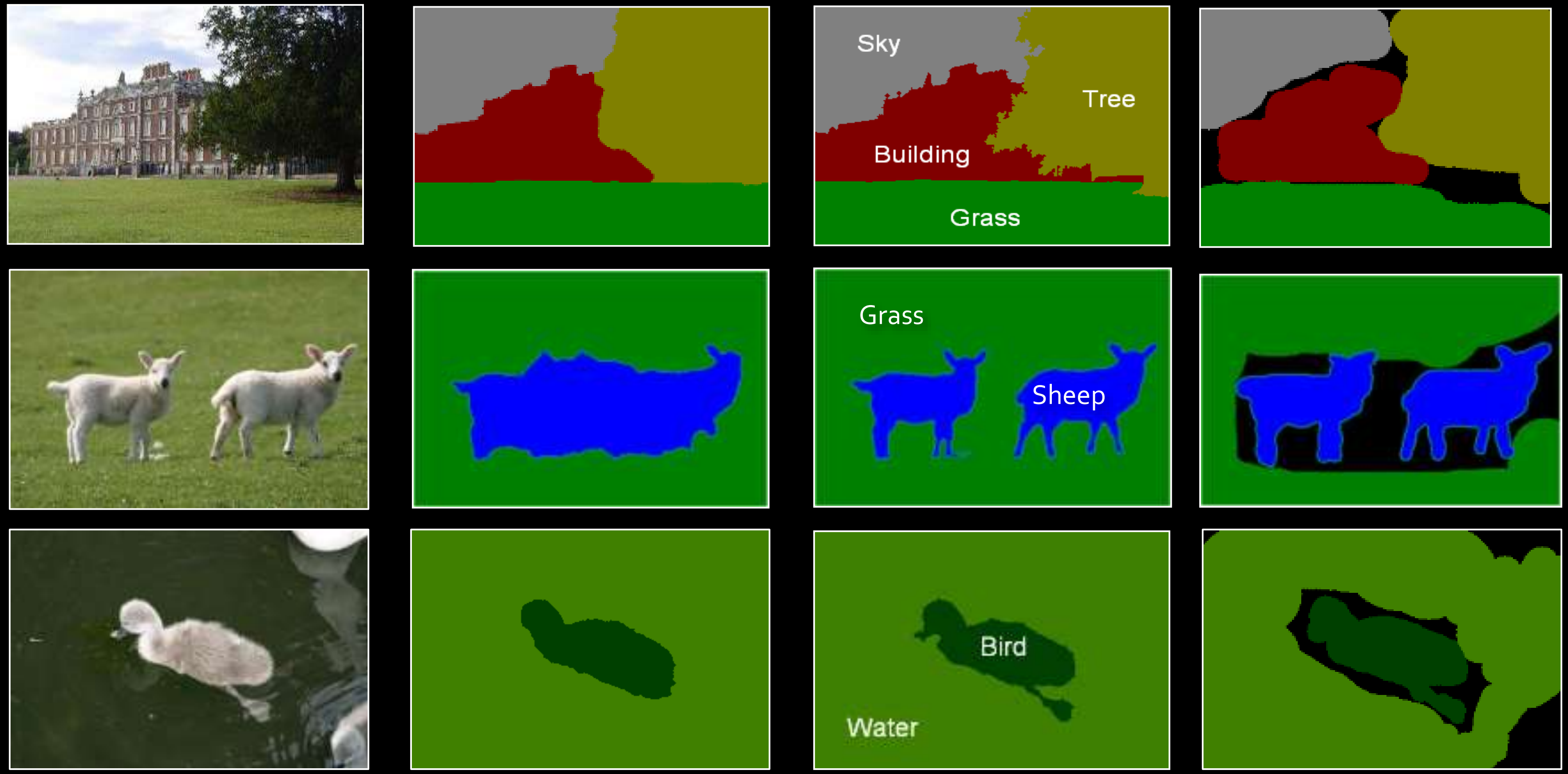
Quadratic Submodular Function

$$f(x) = \begin{cases} 0 & \text{if all } x_i = 0 \\ 1 & \text{otherwise} \end{cases}$$



More Results

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



[Runner-Up, PASCAL VOC 2008]

Another Example

Overcoming short-boundary bias

$$E(X) = \sum c_i x_i + \boxed{\sum d_{ij}|x_i - x_j|}$$

Encourages short boundaries

Image

Segmentation

Penalize types of boundaries not the actual number of boundaries!

Cooperative Cuts [Jegelka and Bilmes, CVPR 2011]

Overcoming short-boundary bias

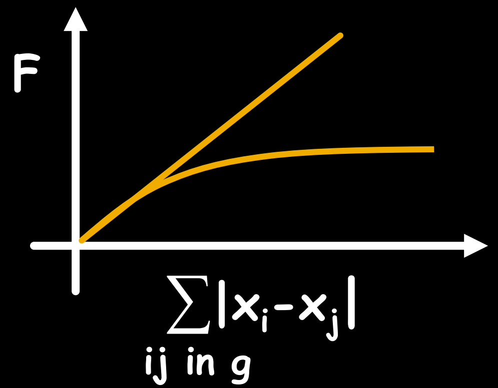
$$E(X) = \sum c_i x_i + \cancel{\sum d_{ij} |x_i - x_j|}$$

Overcoming short-boundary bias

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

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

$$\boxed{h_g(X_p) = F\Big(\sum_{ij \text{ in } g} |x_i - x_j|\Big)}$$

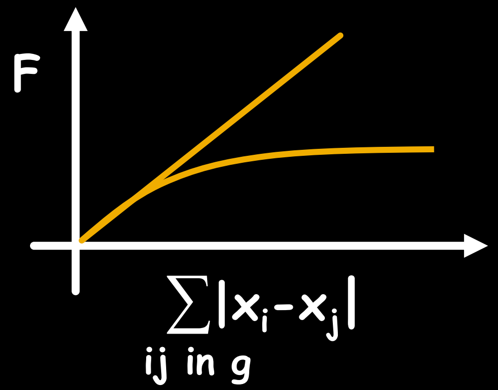


Overcoming short-boundary bias

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

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

$$\boxed{h_g(X_p) = F\left(\sum_{ij \text{ in } g} |x_i - x_j| \right)}$$

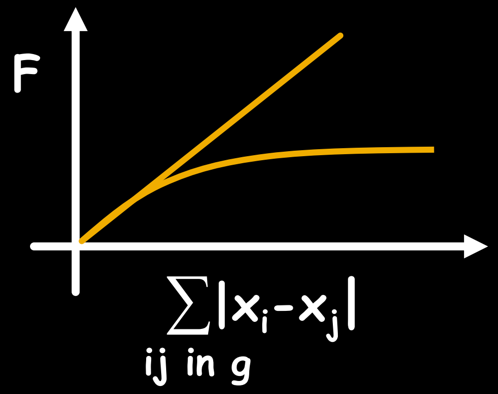


Overcoming short-boundary bias

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

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

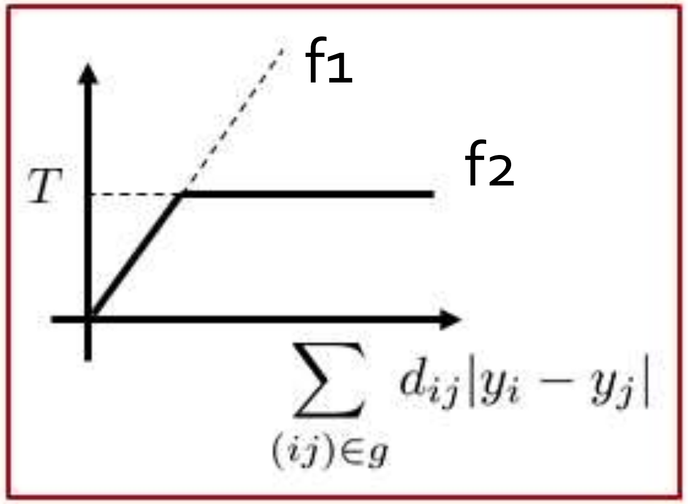
$$\boxed{h_g(X_p) = F\left(\sum_{ij \text{ in } g} |x_i - x_j|\right)}$$



Needs Special Purpose Minimization

Transform Higher order model

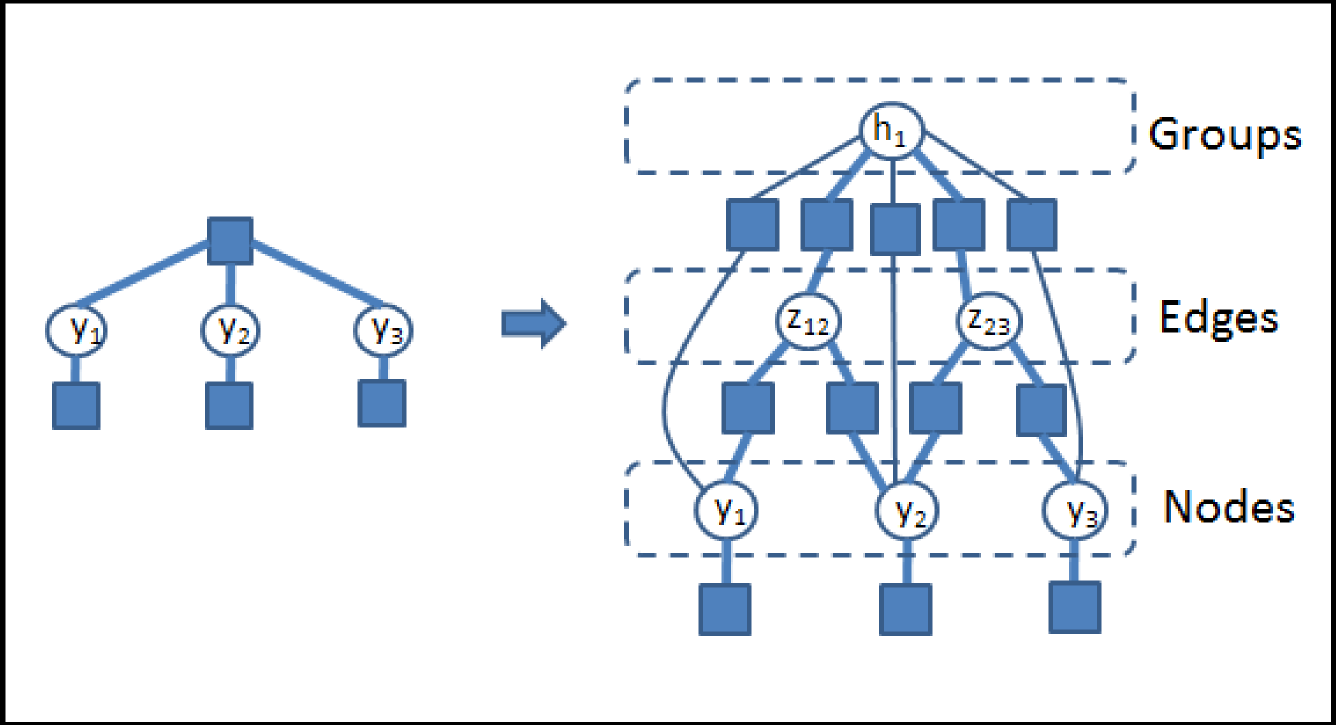
$$H_g(Y) = \min\left\{ \overset{\boxed{\text{f}_1}}{\sum_{(ij)\in g} d_{ij}|y_i - y_j|}, \quad \overset{\boxed{\text{f}_2}}{T} \right\}$$



Transformation

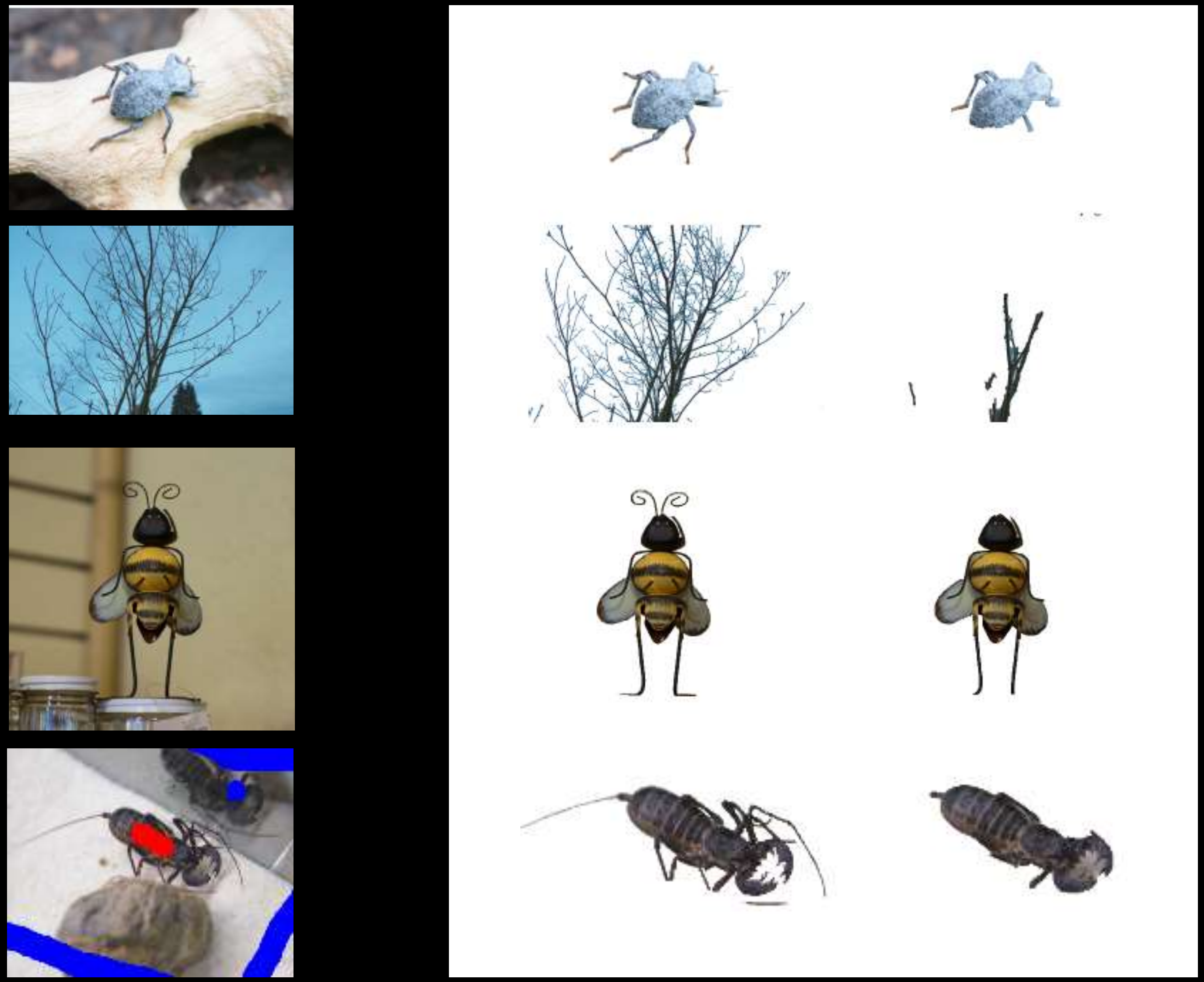
$$H_g(Y) = \min_{h_g \in \{0,1\}} h_g \underbrace{\sum d_{ij}(y_i + y_j - 2y_i y_j)}_{\boxed{\text{f}_1}} + \underbrace{T(1 - h_g)}_{\boxed{\text{f}_2}}$$

Transformation



So what happens to the results?

Results – Interactive Segmentation



Image

Ground Truth

Pairwise Model

Results – Semantic Segmentation

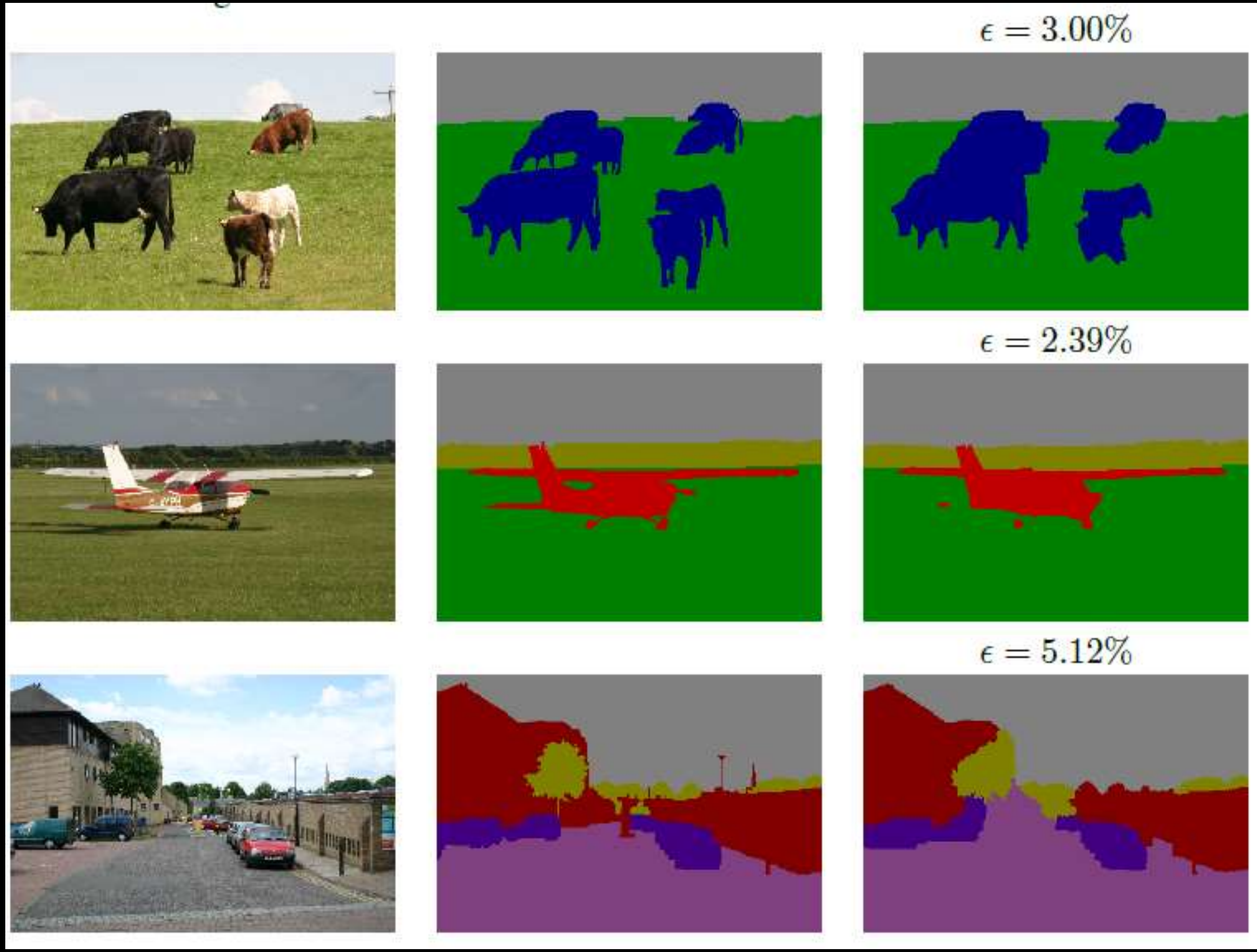


Image | Ground Truth | Pairwise Model

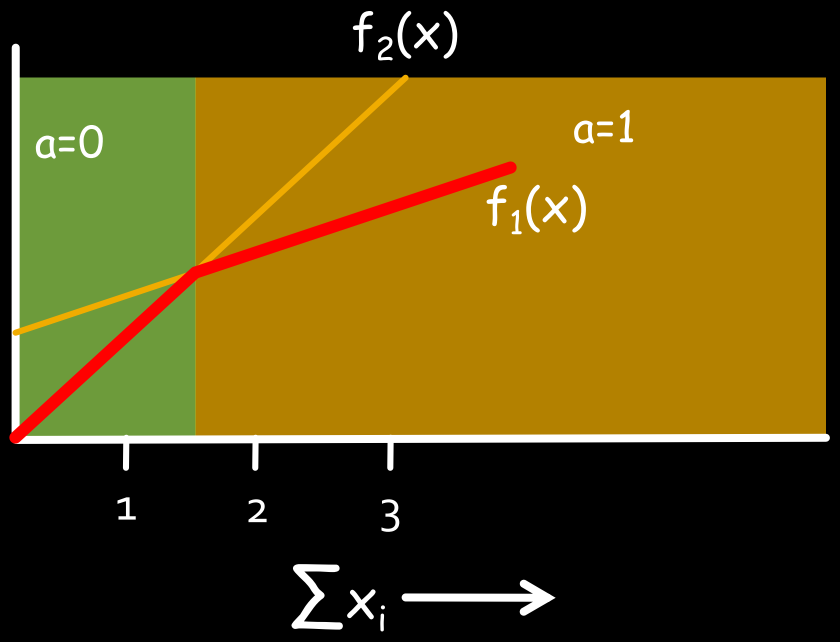
Approach 2: Constrained Inference

Lower Envelope Representation

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

Higher Order Submodular Function

Quadratic Submodular Function



Lower envelop of concave functions is concave

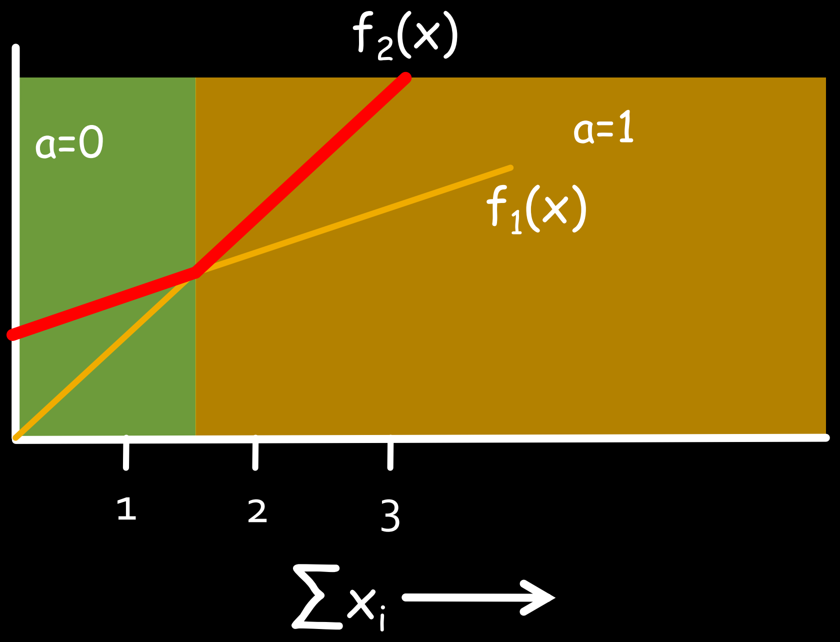
[Kohli et al. '08]

Upper Envelopes

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

Higher Order Submodular Function

Quadratic Submodular Function



Upper envelope of linear functions is convex

Very Hard Problem!!!!

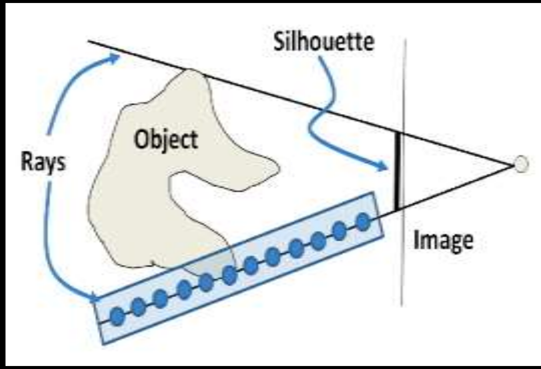
[Kohli and Kumar, CVPR 2010]

Why Upper Envelopes?

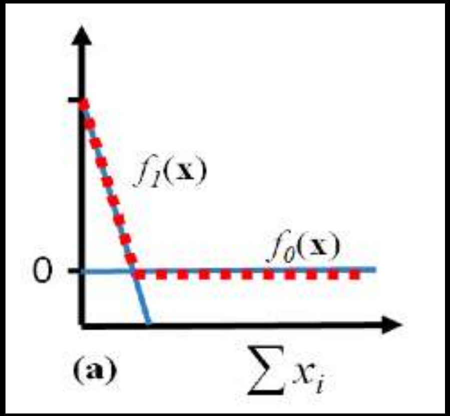
SILHOUETTE CONSTRAINTS

[Sinha et al. '05, Cremers et al. '08]

3D RECONSTRUCTION



Rays must touch silhouette at least once

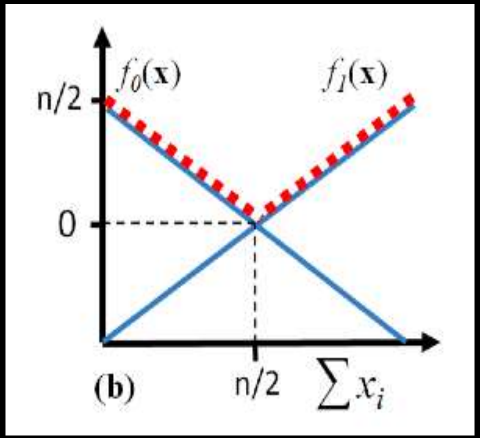


SIZE/VOLUME PRIORS

[Woodford et al. 2009]

BINARY SEGMENTATION

Prior on size of object (say n/2)



[Kohli and Kumar, CVPR 2010]

Approach 2: Constrained Inference

$$x^* = \arg\min_{x} E(x)$$

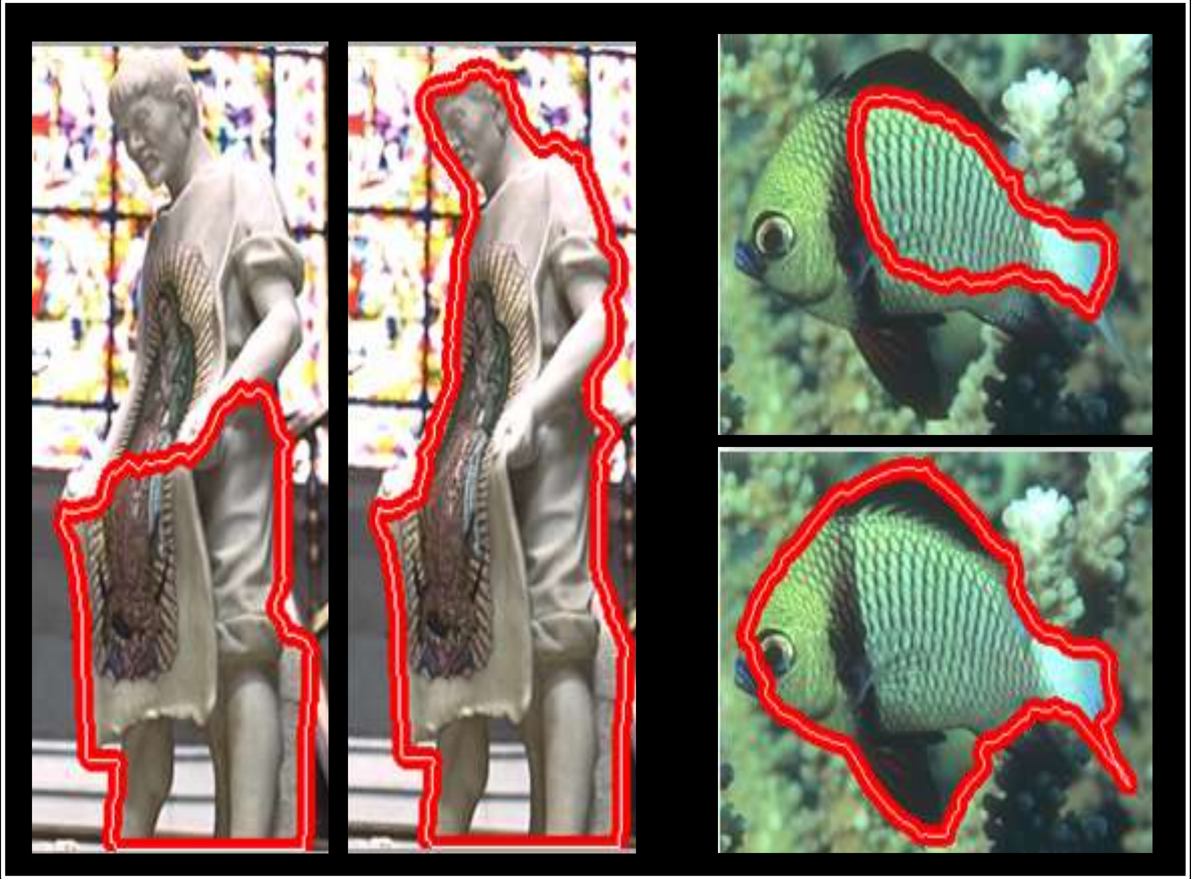
Such that:

x has area A

$$\sum_{i \in V} x_i = A$$

x has boundary length B

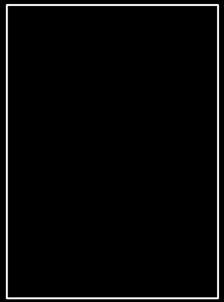
$$\sum_{(i,j) \in E} |x_i - x_j| = B$$



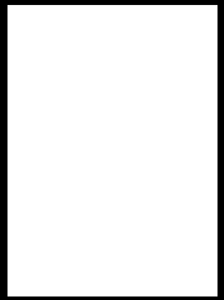
Approach 2: Constrained Inference

Parametric Maxflow/st-Mincut $H^{\lambda}(\mathbf{x}) = E(\mathbf{x}) + \lambda \sum_{v \in V} x_v$

$\lambda = +\infty$ $\quad \lambda \quad$ $\lambda = -\infty$



$\sum x_v = 0$



$\sum x_v = n$

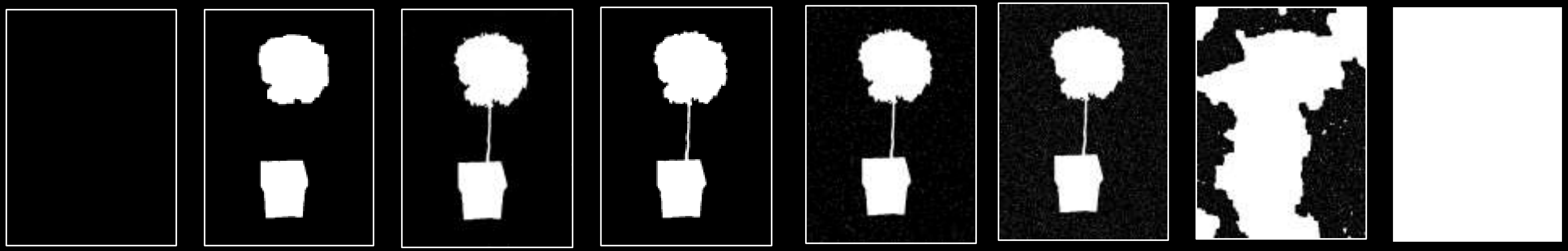
$$\underbrace{\qquad\qquad\qquad\qquad}_{x^* = \arg\min H^{\lambda}(\mathbf{x})}$$

Approach 2: Constrained Inference

Parametric Maxflow/st-Mincut

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

$\lambda = +\infty$ λ $\lambda = -\infty$



$\sum x_v = 0$ $\sum x_v = n$

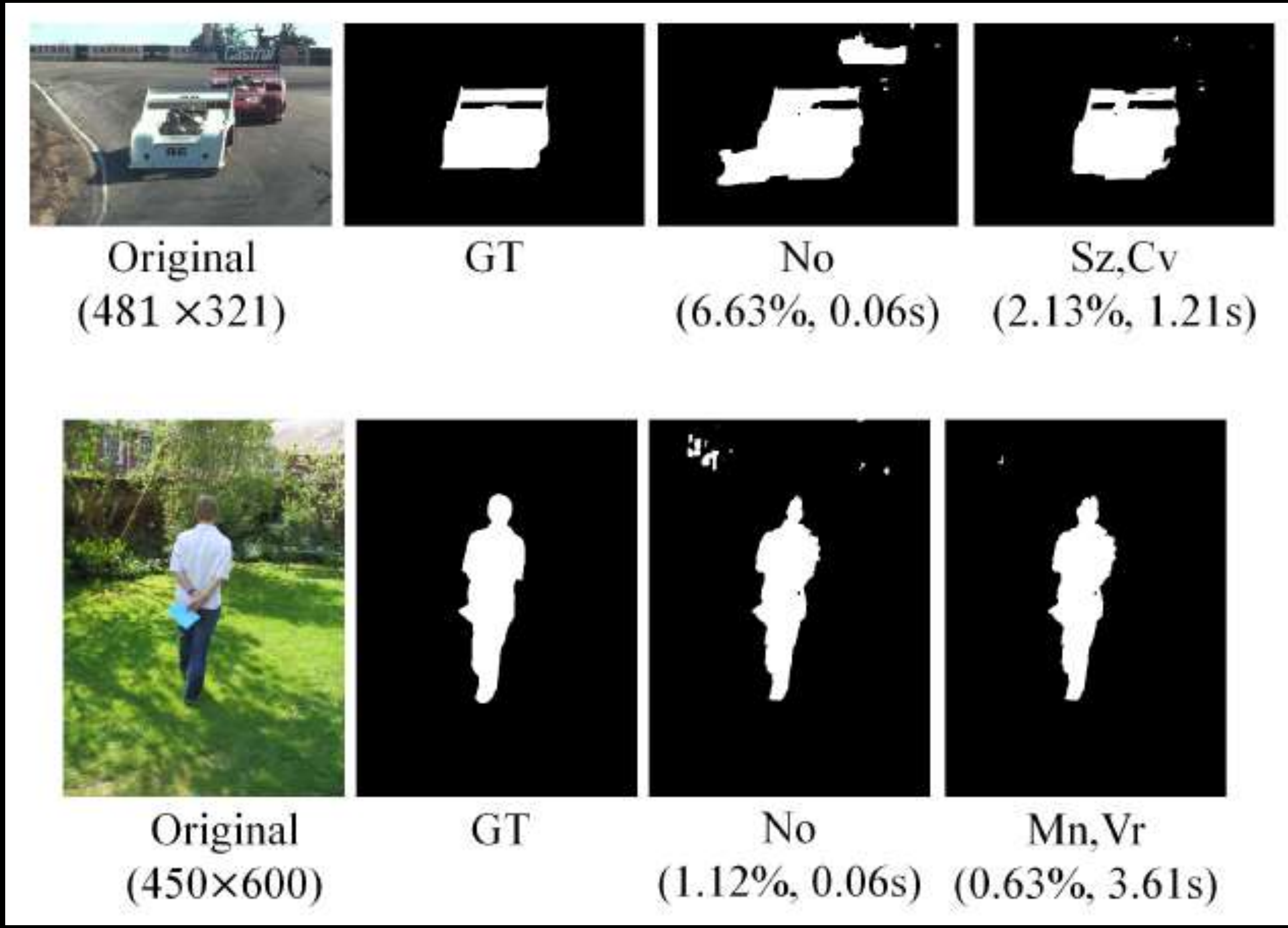
$$x^* = \arg\min H^{\lambda}(\mathbf{x})$$

Lemma 1. *If an assignment* $\mathbf{x}$ *minimizes the energy function* H^{λ} *for some* λ, $E(\mathbf{x})$ *is minimum under the same label count as* $\mathbf{x}$.

[Gallo et al. 1986] [Kolmogorov et al. ICCV 2007]

Approach 2: Constrained Inference

Generalized
(Multi-dimensional)
Parametric Maxflow



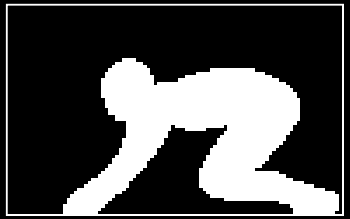
Approach 3: Learning

Low-order Models for Enforcing Higher-order Statistics

data: $\{\mathbf{x}^k, \mathbf{y}^k\},\ k = 1..K$

$$E(\mathbf{x}, \mathbf{y}, \mathbf{w}) = \mathbf{w}^\top \psi(\mathbf{y}, \mathbf{x})$$

x



y

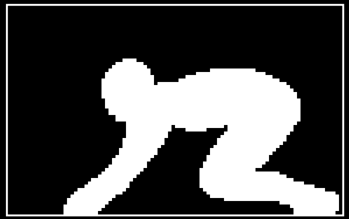
Can we find parameters that lead to solutions consistent with higher-order statistics?

[Pletscher and Kohli, AISTATS 2012]

Low-order Models for Enforcing Higher-order Statistics

data: $\{\mathbf{x}^k, \mathbf{y}^k\},\ k = 1..K$

$\mathbf{x}$



$\mathbf{y}$

$$E(\mathsf{x}, \mathsf{y}, \mathsf{w}) = \mathbf{w}^\top \psi(\mathsf{y}, \mathsf{x})$$

$$\min_{\xi \geq 0, \mathbf{w}, \mathsf{U}} \quad o(\mathbf{w}, \mathsf{U}) := \frac{1}{2} \|\mathbf{w}\|^2$$

$$\text{sb.t.} \quad \mathbf{w}^\top \left[\psi(\mathbf{x}^k, \mathbf{y}) - \psi(\mathbf{x}^k, \mathbf{y}^k)\right] \geq \sum_{i \in \mathcal{V}} y_i \neq y_i^*$$

[AISTATS 2012]

Low-order Models for Enforcing Higher-order Statistics

data: $\{\mathbf{x}^k, \mathbf{y}^k\},\ k = 1..K$

$$E(\mathbf{x}, \mathbf{y}, \mathbf{w}) = \mathbf{w}^\top \psi(\mathbf{y}, \mathbf{x})$$

x y

$$\min_{\xi \geq 0, \mathbf{w}, \mathbf{U}} \quad o(\mathbf{w}, \mathbf{U}) := \frac{1}{2}\|\mathbf{w}\|^2$$

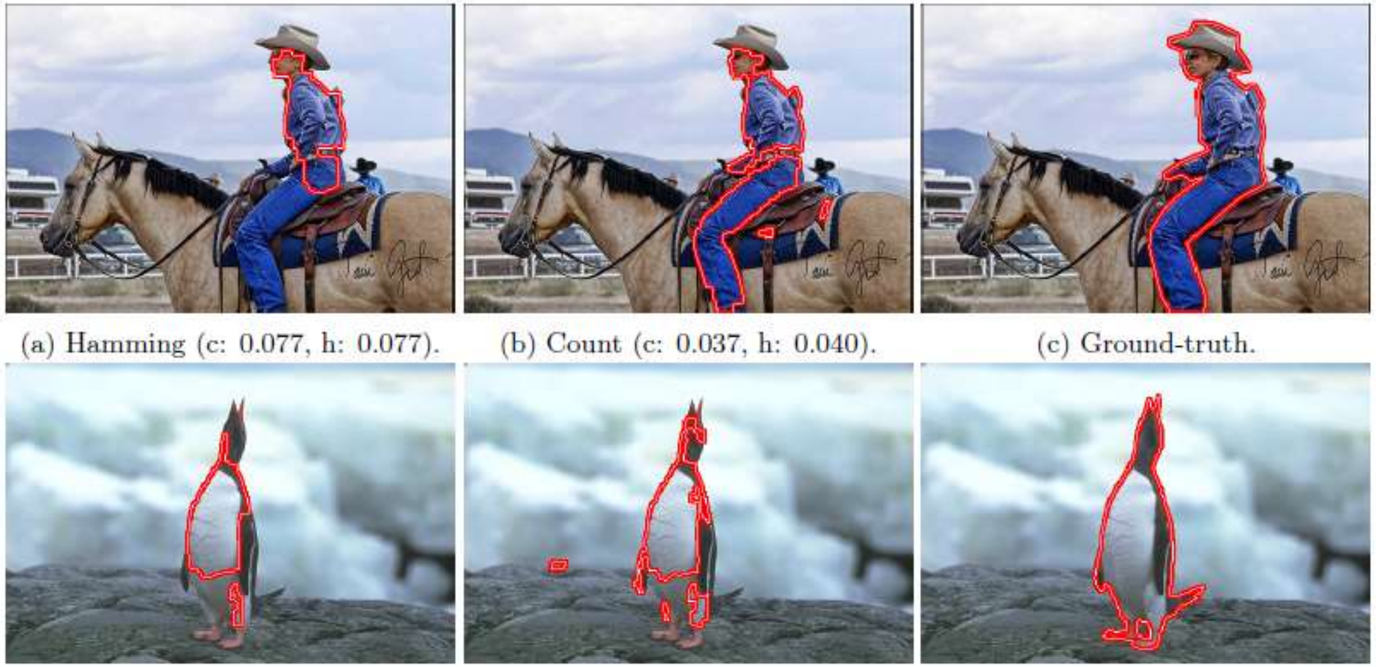
$$\text{sb.t.} \quad \mathbf{w}^\top \left[\psi(\mathbf{x}^k, \mathbf{y}) - \psi(\mathbf{x}^k, \mathbf{y}^k)\right] \geq \left|\sum_{i \in \mathcal{V}} y_i - \sum_{i \in \mathcal{V}} y_i^*\right|$$

Requires solution of

$$\min_{y} E(y, x, w) - \left|\sum_{i \in \mathcal{V}} y_i - \sum_{i \in \mathcal{V}} y_i^*\right|$$

[AISTATS 2012]

Learning Higher-order Model for Enforcing Low-order statistics



(a) Hamming (c: 0.077, h: 0.077). (b) Count (c: 0.037, h: 0.040). (c) Ground-truth.

(d) Hamming (c: 0.047, h: 0.047). (e) Count (c: 0.040, h: 0.043). (f) Ground-truth.

Eval \ Train		Hamming better (%)	Count better (%)
4/S	Hamming	52.1 ± 7.0	47.9 ± 7.0
	Count	33.8 ± 8.3	66.2 ± 8.3
4/D	Hamming	39.4 ± 6.1	60.6 ± 6.1
	Count	29.6 ± 8.3	70.4 ± 8.3

[AISTATS 2012]

Challenges and Opportunities

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

Thanks for listening. Questions?