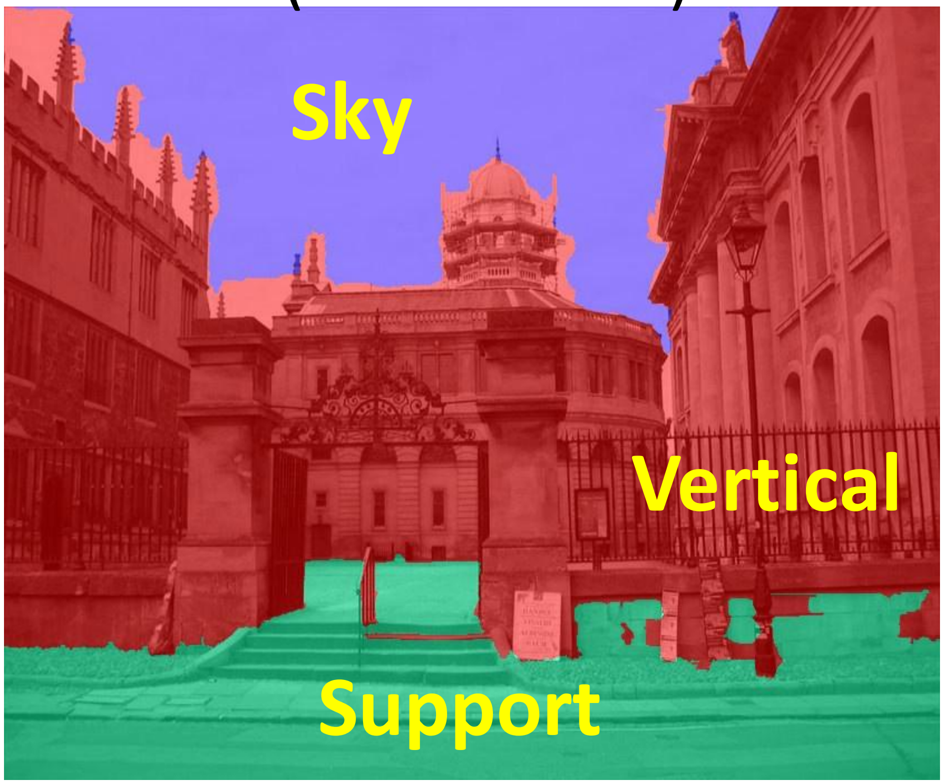
# Contextual Classification with Functional Max-Margin Markov Networks

**Dan Munoz** Drew Bagnell
Nicolas Vandapel Martial Hebert

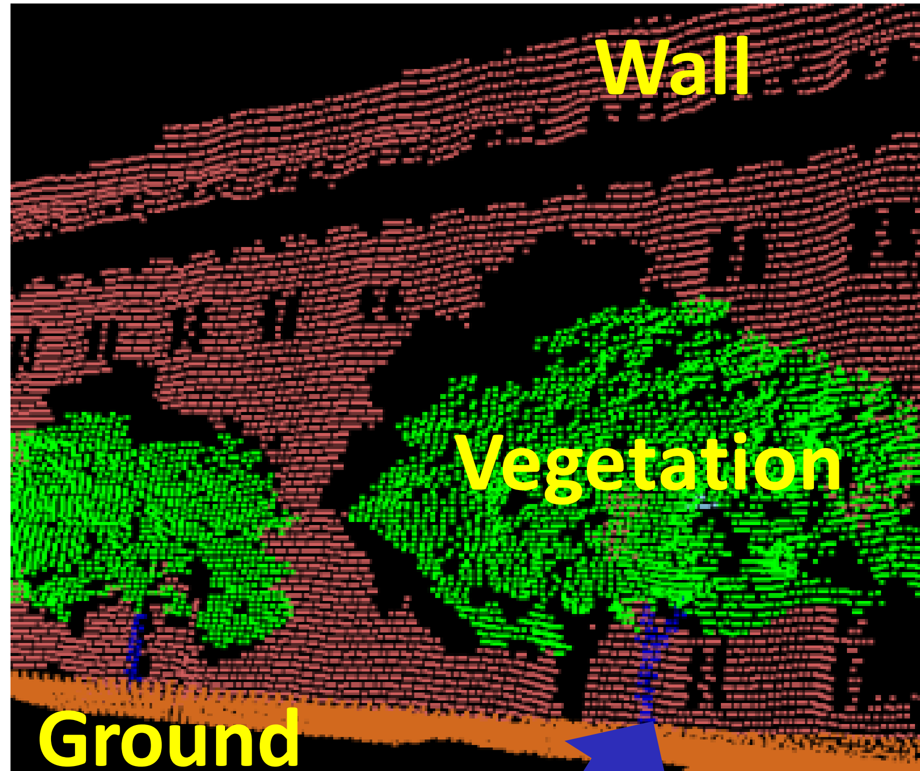
THE ROBOTICS INSTITUTE
Carnegie Mellon

# Problem

**Geometry Estimation**
(Hoiem *et al.*)



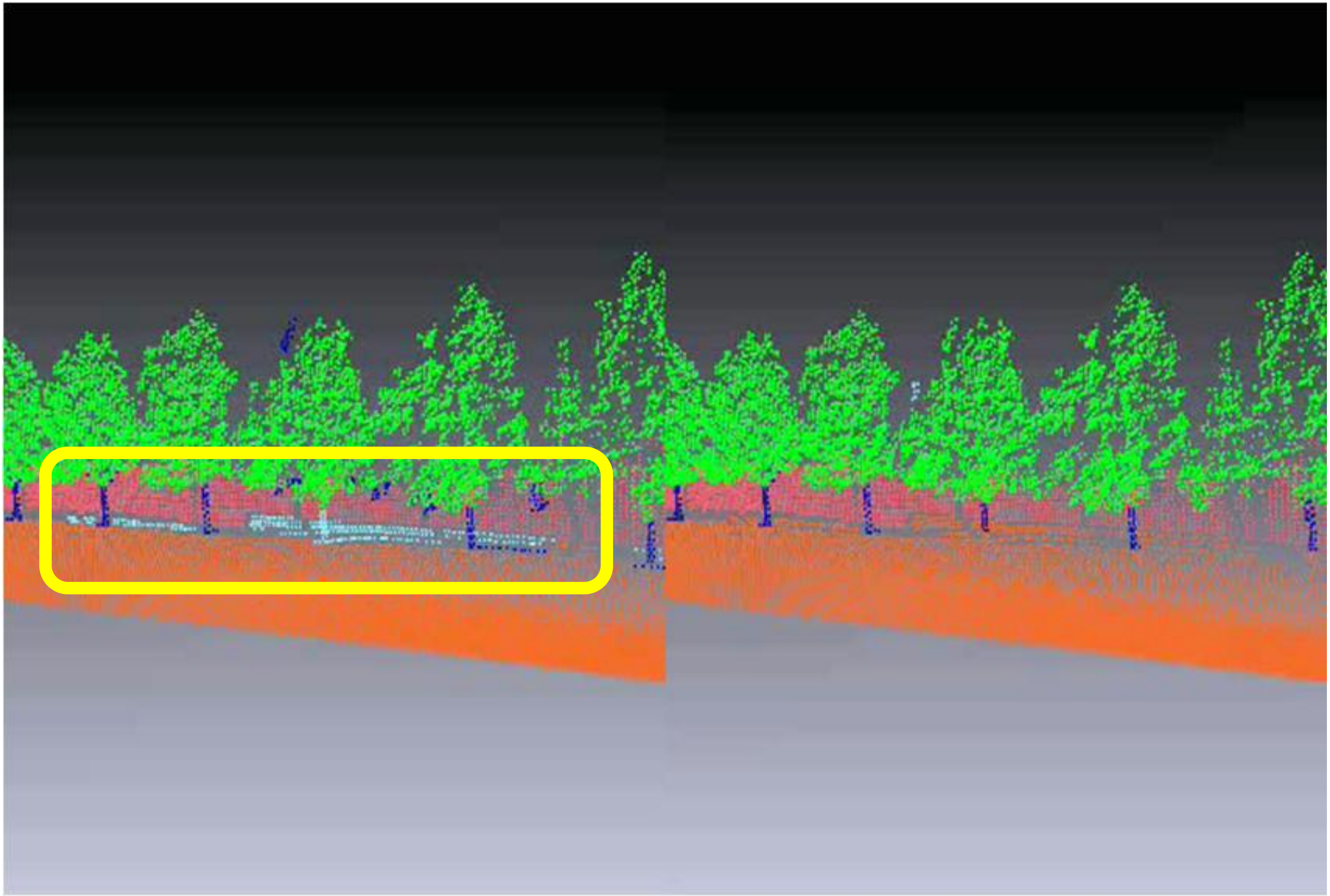

**3-D Point Cloud Classification**



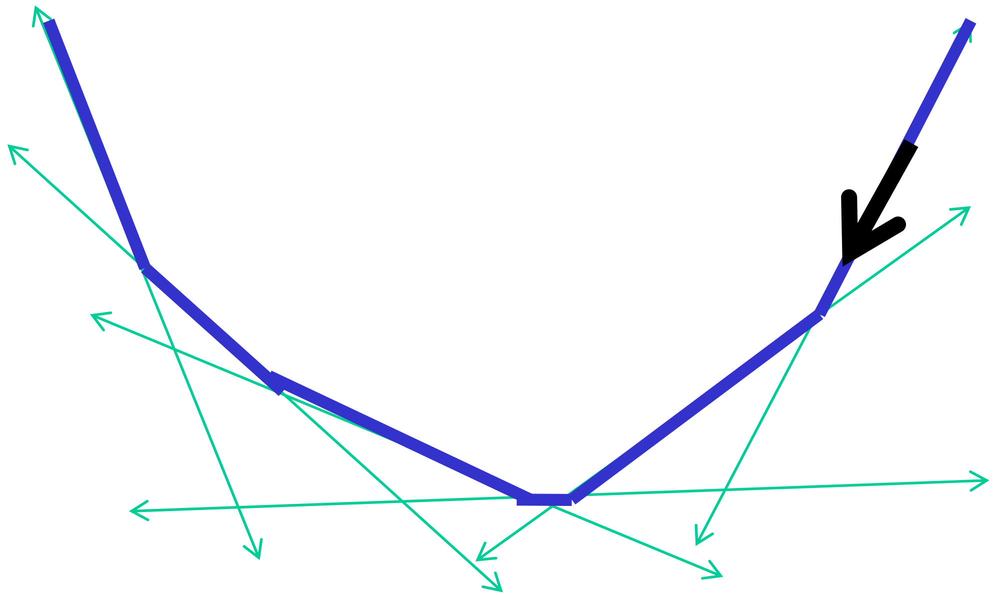

**Tree trunk**

Our classifications

# Room For Improvement



# Approach: Improving CRF Learning



**Gradient descent (w)** → **“Boosting” (h)**

- Friedman *et al.* 2001, Ratliff *et al.* 2007

\+ Better learn models with **high-order** interactions

\+ Efficiently handle **large** data & feature sets

\+ Enable **non-linear** clique potentials

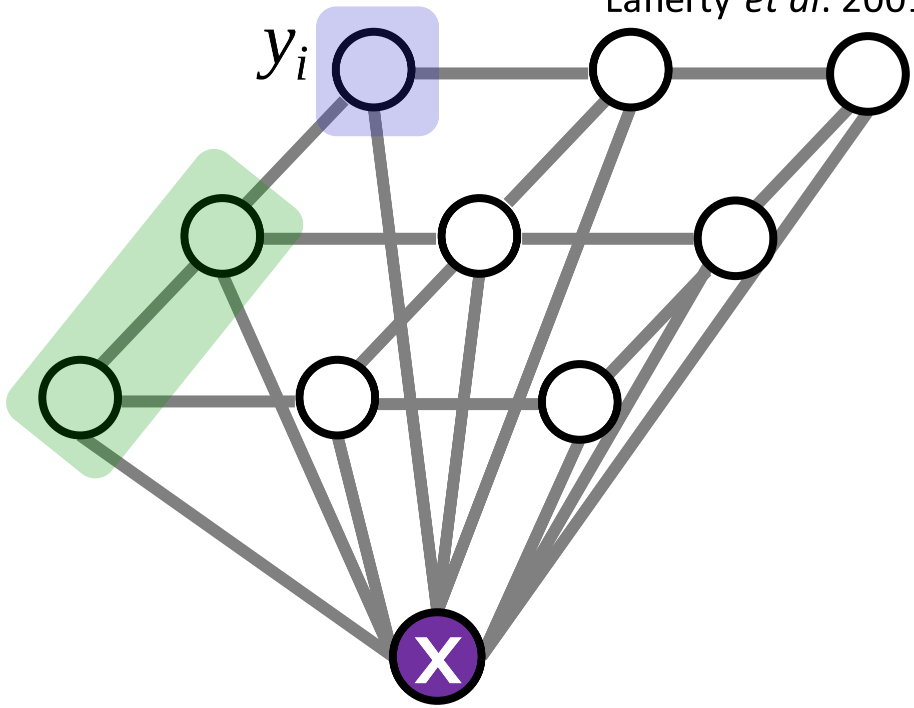
# Conditional Random Fields

Lafferty *et al.* 2001

- Pairwise model

$$\mathbf{Y} = \{Y_1, \ldots, Y_N\}$$

$$Y_i \in \underbrace{\{\ell_1, \ldots, \ell_K\}}_{\textit{Labels}}$$




$$P(\mathbf{y}|\mathbf{x}) = \frac{1}{Z} \exp\left[\sum_{i=1}^{N} \phi_i(y_i, \mathbf{x}) + \sum_{(ij)\in E} \phi_{ij}(y_i, y_j, \mathbf{x})\right]$$

- MAP Inference

$$\mathbf{y}^* = \arg\max_{\mathbf{y}} \sum_{i=1}^{N} \phi_i(y_i, \mathbf{x}) + \sum_{(ij)\in E} \phi_{ij}(y_i, y_j, \mathbf{x})$$

# Parametric Linear Model

**Weights**

$$\phi_i(y_i = l_k, \mathbf{x}) = \mathbf{w}^T \mathbf{f}_i(l_k, \mathbf{x})$$

Local **features** that describe **label**

# Associative/Potts Potentials

$$\phi_{ij}(y_i = l_k, y_j = l_k, \mathbf{x}) = \mathbf{w}^T \mathbf{f}_{ij}(l_k, l_k, \mathbf{x})$$

$$\phi_{ij}(y_i = l_k, y_j = l_l, \mathbf{x}) = 0$$

**Labels Disagree**

# Overall Score

$$\log P(\mathbf{y}|\mathbf{x}) = \sum_{i=1}^{N} \mathbf{w}^T \mathbf{f}(y_i, \mathbf{x}) + \sum_{(ij)\in E} \mathbf{w}^T \mathbf{f}(y_i, y_j, \mathbf{x}) - \log Z$$

$$= \underbrace{\mathbf{w}^T \mathbf{f}(\mathbf{y}, \mathbf{x})} - \log Z$$

**Overall Score** for a labeling **y** to all nodes

# Learning Intuition

- Iterate
  - Classify with current CRF model

$$\mathbf{y}^* = \arg\max_{\mathbf{y}} \sum_{i=1}^{N} \phi_i(y_i, \mathbf{x}) + \sum_{(ij)\in E} \phi_{ij}(y_i, y_j, \mathbf{x})$$

  - If $y_i^* \neq \hat{y}_i$ **(misclassified)**

$\varphi(\mathbf{f}_i(\hat{y}_i, \mathbf{x}))$ ➡ **increase** score

$\varphi(\mathbf{f}_i(y_i^*, \mathbf{x}))$ ➡ **decrease** score

  - *(Same update with edges)*

# Max-Margin Structured Prediction

Taskar *et al*. 2003

min **w** [ <u>**Best**</u> **score from** <u>**all**</u> **labelings (+*M*)** ] **–** [ **Score with** <u>**ground truth**</u> **labeling** ]

**Ground truth labels**

$$\min_{\mathbf{w}} \frac{\lambda}{2}||\mathbf{w}||^2 + \max_{\mathbf{y}}[\mathbf{w}^T\mathbf{f}(\mathbf{y},\mathbf{x}) + M] - \mathbf{w}^T\mathbf{f}(\hat{\mathbf{y}},\mathbf{x})$$

**Convex**

# Descending† Direction

(Objective)

$$\min_{\mathbf{w}} \frac{\lambda}{2}||\mathbf{w}||^2 + \max_{\mathbf{y}}[\mathbf{w}^T\mathbf{f}(\mathbf{y},\mathbf{x}) + M] - \mathbf{w}^T\mathbf{f}(\hat{\mathbf{y}},\mathbf{x})$$

**Labels from MAP inference**

$$\mathbf{g_w} = -\lambda\mathbf{w} - \mathbf{f}(\mathbf{y}^*, \mathbf{x}) + \mathbf{f}(\hat{\mathbf{y}}, \mathbf{x})$$

**Ground truth labels**

# Learned Model

$$\phi_i(y_i = l_k, \mathbf{x}) = \mathbf{w}^T \mathbf{f}_i(l_k, \mathbf{x}) = \sum_t \alpha_t \mathbf{g}_{\mathbf{w_t}}{}^T \mathbf{f}_i(l_k, \mathbf{x})$$

# Update Rule

- Unit step-size, and $\lambda = 0$

$$\mathbf{w}_{t+1} \leftarrow \mathbf{w}_t + \mathbf{g}_{\mathbf{w}}$$

$$\mathbf{w}_{t+1} \leftarrow \mathbf{w}_t + \mathbf{f}(\hat{\mathbf{y}}, \mathbf{x}) - \mathbf{f}(\mathbf{y}^*, \mathbf{x})$$

$$\mathbf{w}_{t+1} \mathrel{+}= \sum_{i=1}^{N} \mathbf{f}_i(\hat{y}_i, \mathbf{x}) - \mathbf{f}_i(y_i^*, \mathbf{x}) + \sum_{(ij)\in E} \mathbf{f}_{ij}(\hat{y}_i, \hat{y}_j, \mathbf{x}) - \mathbf{f}_{ij}(y_i^*, y_j^*, \mathbf{x})$$

**Ground truth** **Inferred**

# Verify Learning Intuition

- Iterate

$$\mathbf{w}_{t+1} \mathrel{+}= \sum_{i=1}^{N} \mathbf{f}_i(\hat{y}_i, \mathbf{x}) - \mathbf{f}_i(y_i^*, \mathbf{x}) + \sum_{(ij)\in E} \mathbf{f}_{ij}(\hat{y}_i, \hat{y}_j, \mathbf{x}) - \mathbf{f}_{ij}(y_i^*, y_j^*, \mathbf{x})$$

- If $y_i^* \neq \hat{y}_i$ **(misclassified)**

$\varphi(\mathbf{f}_i(\hat{y}_i, \mathbf{x}))$ ➡ **increase** score

$\varphi(\mathbf{f}_i(y_i^*, \mathbf{x}))$ ➡ **decrease** score

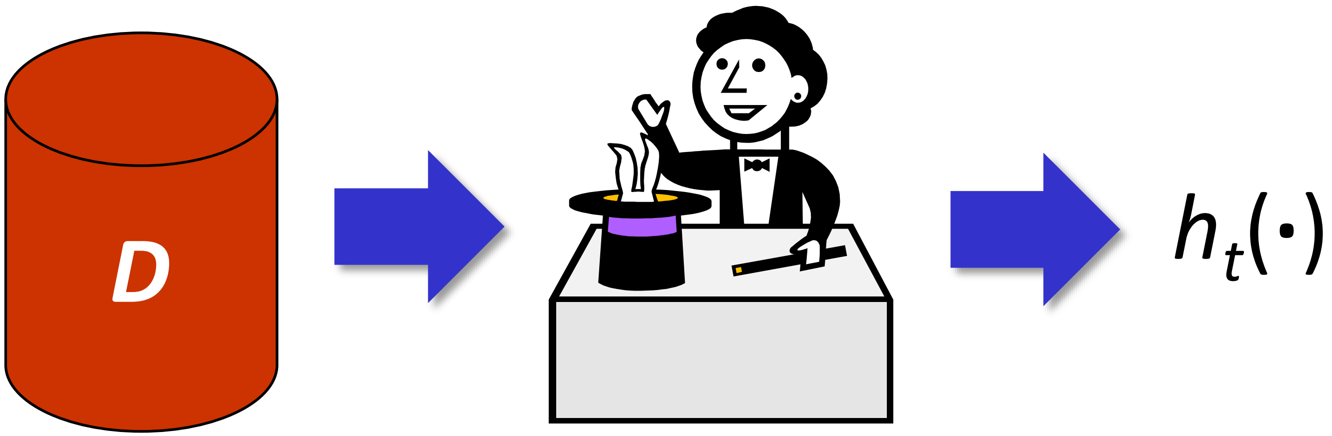
# Alternative Update

1. **Create training set:** *D*
   - From the **misclassified** nodes & edges

$$D = \left\{ \begin{array}{cc} \left( \boxed{\mathbf{f}_i(\hat{y}_i, \mathbf{x})}, +1 \right) & \left( \boxed{\mathbf{f}_{ij}(\hat{y}_i, \hat{y}_j, \mathbf{x})}, +1 \right) \\ \left( \boxed{\mathbf{f}_i(y_i^*, \mathbf{x})}, -1 \right) & \left( \boxed{\mathbf{f}_{ij}(y_i^*, y_j^*, \mathbf{x})}, -1 \right) \end{array} \right\}$$

# Alternative Update

1. Create training set: *D*
2. **Train regressor:** $h_t$



# Alternative Update

1. Create training set: *D*
2. Train regressor: *h*
3. **Augment model:**

$$\phi_i(y_i, \mathbf{x}) = \sum_t \alpha_t h_t(\mathbf{f}_i(y_i, \mathbf{x}))$$

(Before)

$$\phi_i(y_i, \mathbf{x}) = \sum_t \alpha_t \mathbf{g_{w_t}}^T \mathbf{f}_i(y_i, \mathbf{x})$$

# Functional M³N Summary

- Given features $\mathbf{x}$ and labels $\hat{\mathbf{y}}$
- for *T* iterations
  - Classification with current model

$$\mathbf{y}^* = \arg\max_{\mathbf{y}} \sum_{i}^{N} \phi_i(y_i, \mathbf{x}) + \sum_{ij \in E} \phi_{ij}(y_i, y_j, \mathbf{x}) + Margin$$

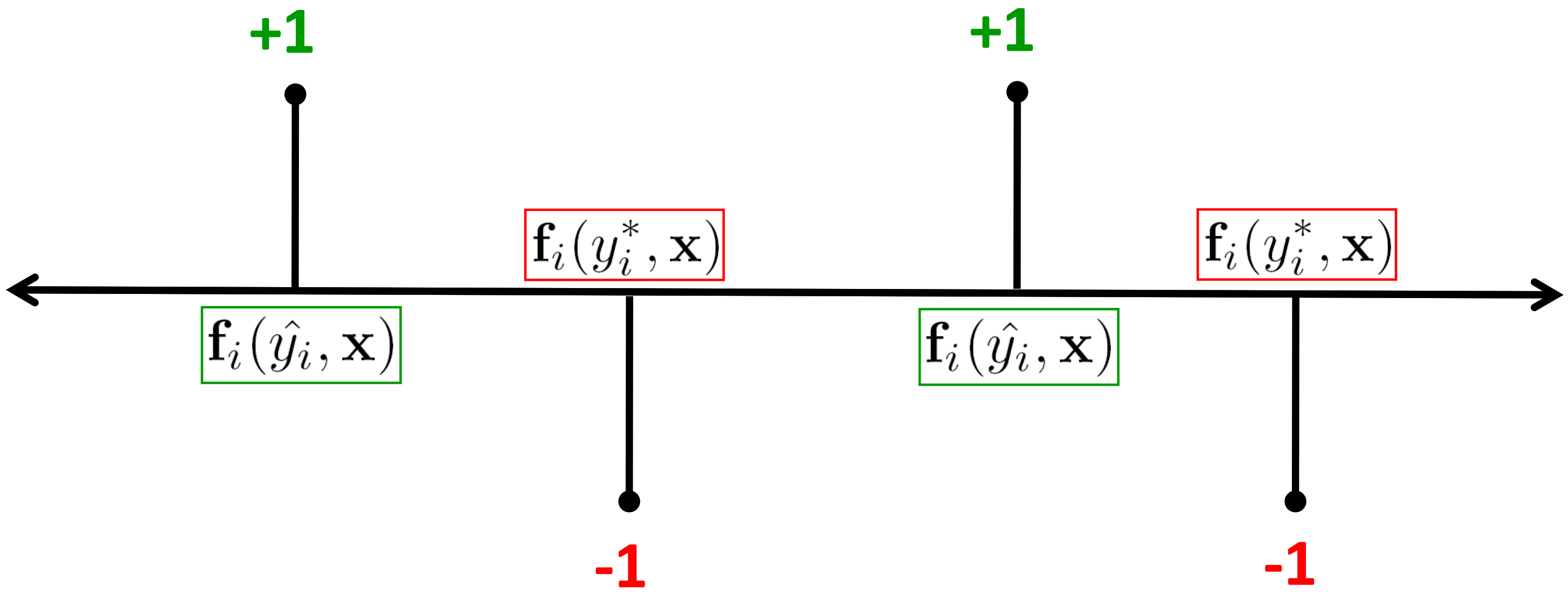
  - Create training set from **misclassified** cliques


  - Train regressor/classifier $h_t$

  - Augment model $\phi(\cdot) = \sum_t \alpha_t h_t(\cdot)$

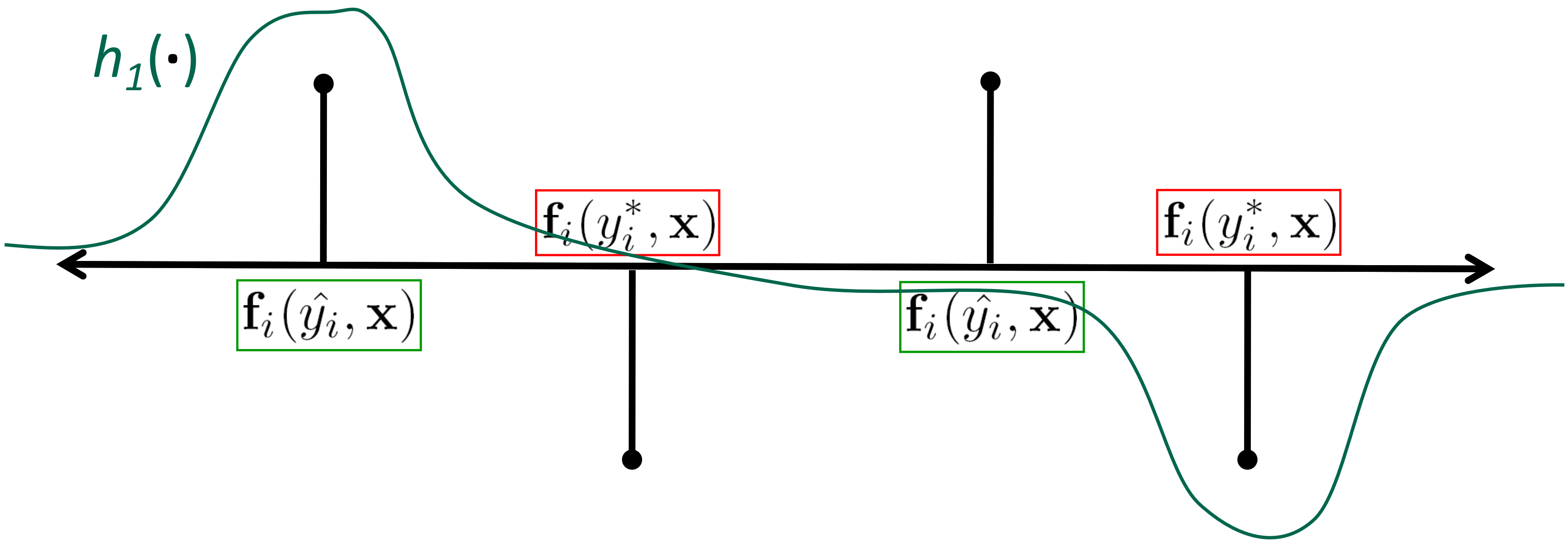
# Illustration

- Create training set

# Illustration

- Train regressor $h_t$

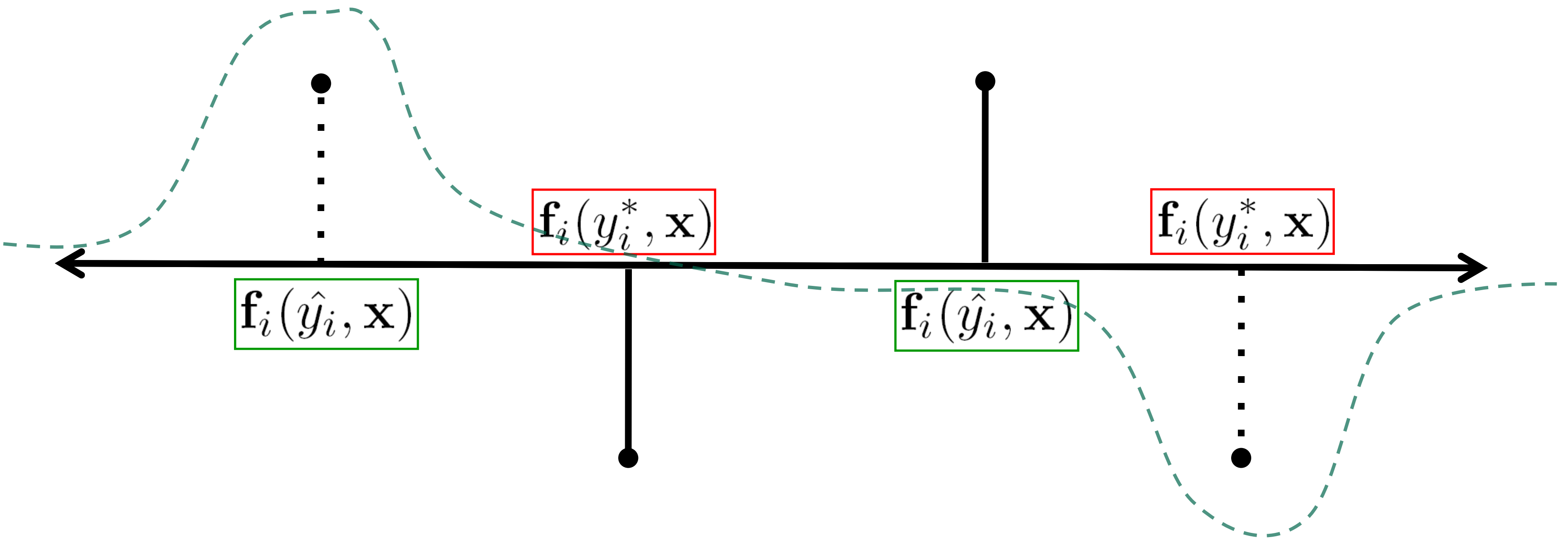



$$\phi(\cdot) = \alpha_1 h_1(\cdot)$$

# Illustration

- Classification with current **CRF** model

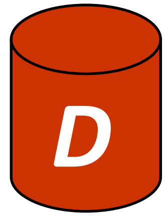
$$\mathbf{y}^* = \arg\max_{\mathbf{y}}$$

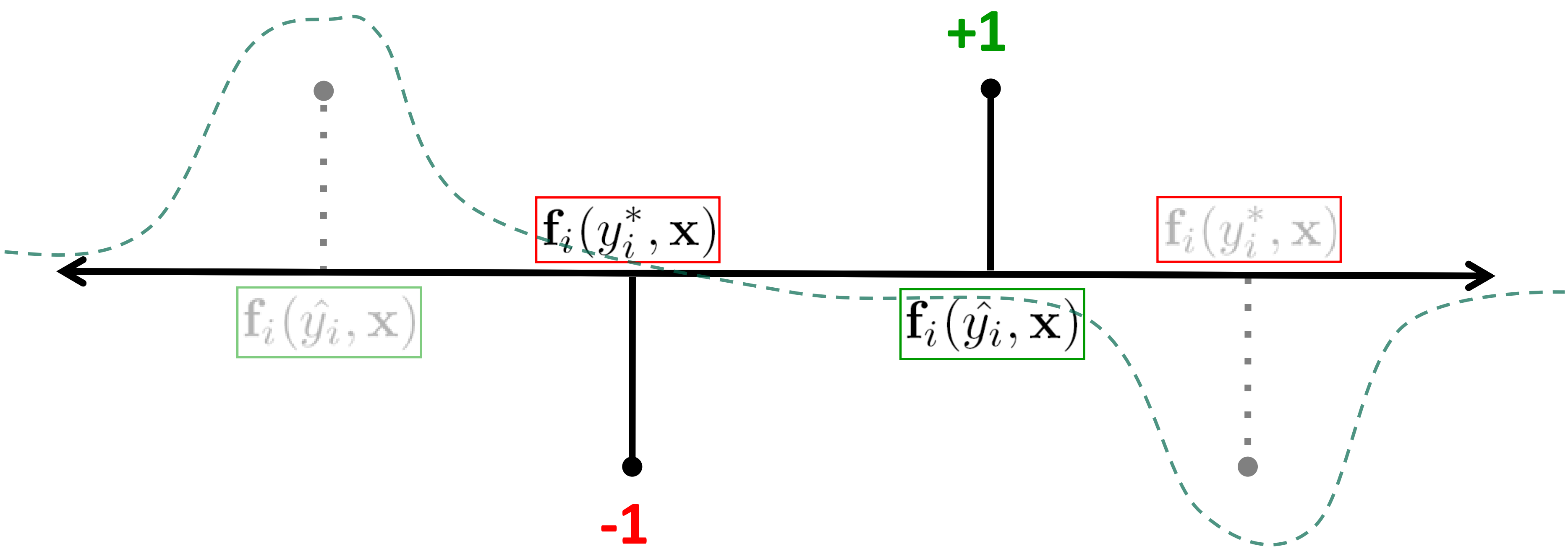



$$\phi(\cdot) = \alpha_1 h_1(\cdot)$$

# Illustration

- Create training set

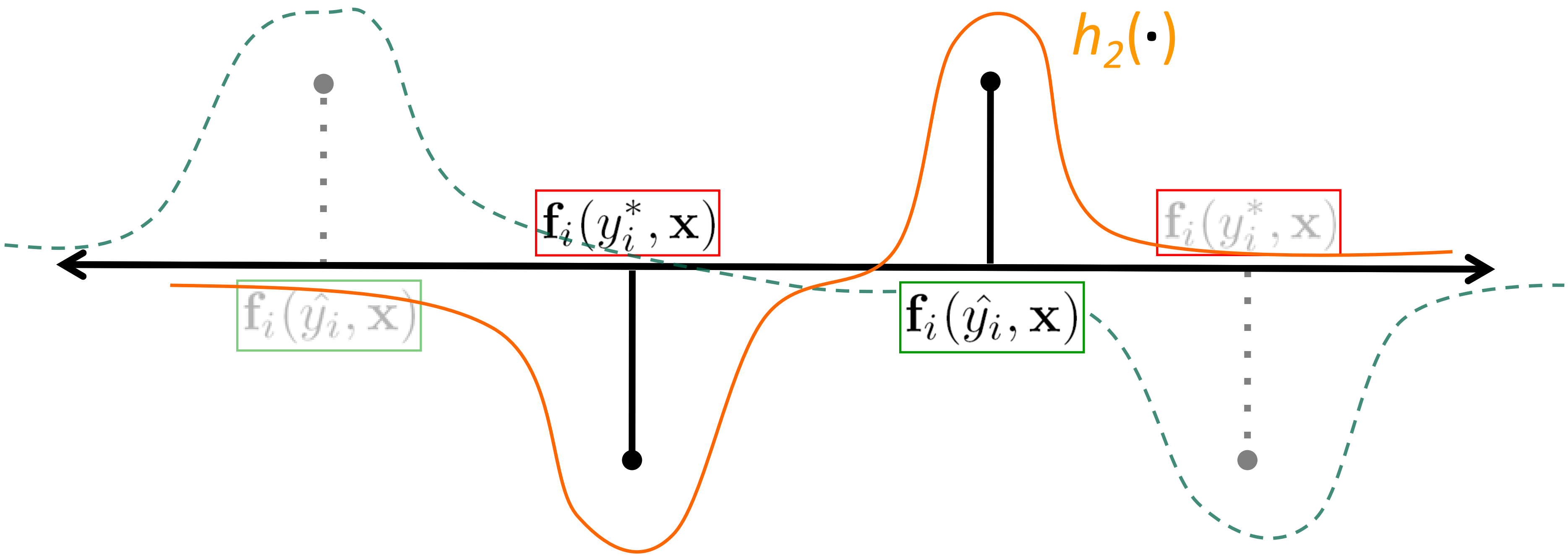



$$\phi(\cdot) = \alpha_1 h_1(\cdot)$$

# Illustration

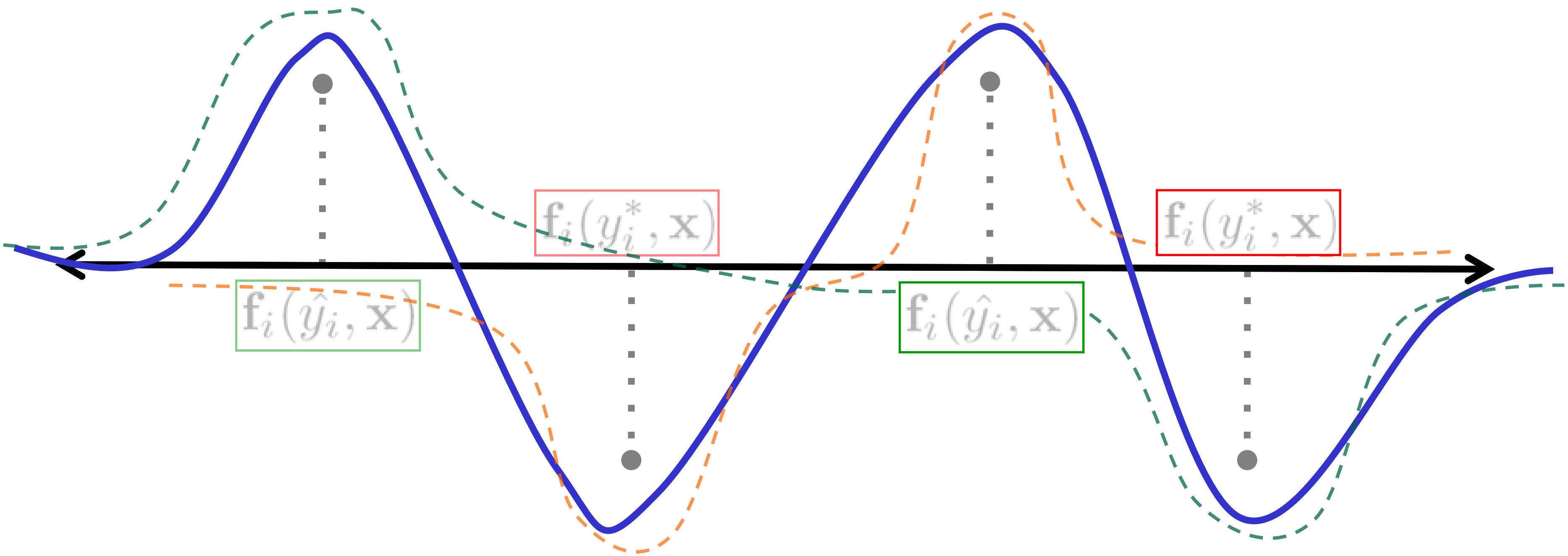
- Train regressor $h_t$




$$\phi(\cdot) = \alpha_1 h_1(\cdot) + \alpha_2 h_2(\cdot)$$

# Illustration

- Stop



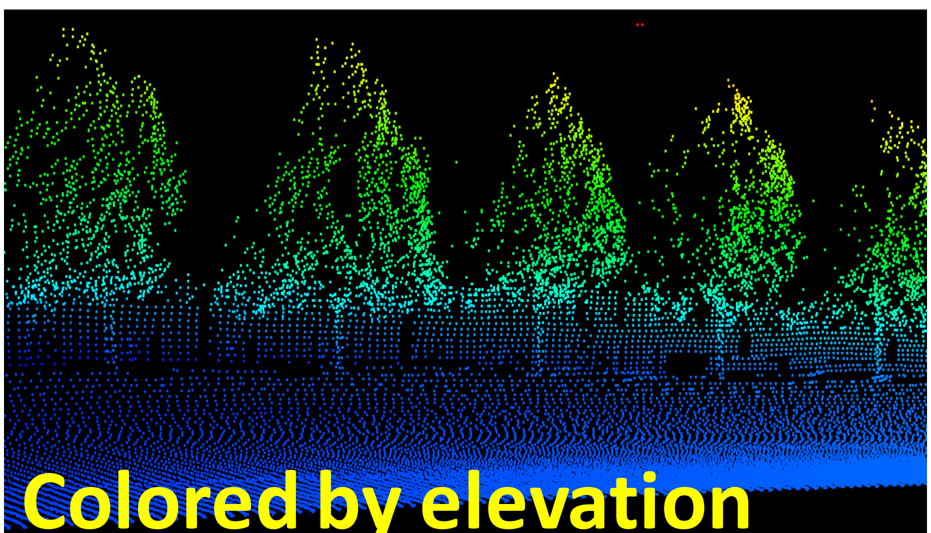

$$\phi(\cdot) = \alpha_1 h_1(\cdot) + \alpha_2 h_2(\cdot)$$

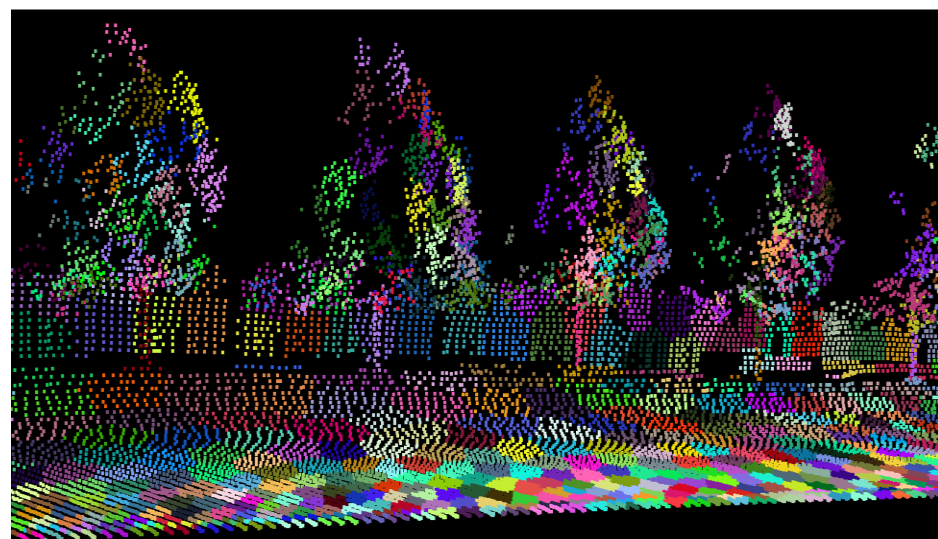
# Boosted CRF Related Work

- Gradient Tree Boosting for CRFs
  - Dietterich *et al.* 2004
- Boosted Random Fields
  - Torralba *et al.* 2004
- Virtual Evidence Boosting for CRFs
  - Liao *et al.* 2007

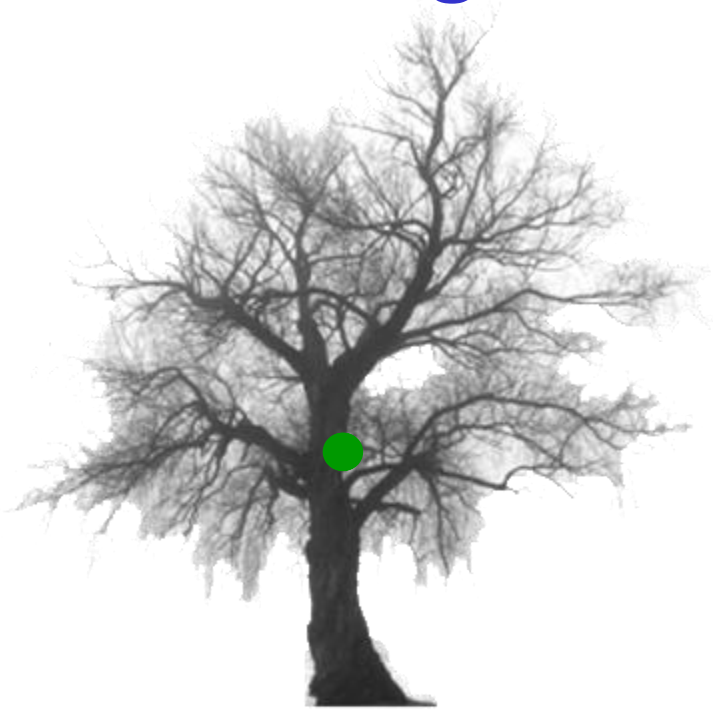
- Benefits of Max-Margin objective
  - **Do not need marginal probabilities**
  - (Robust) High-order interactions
    - ✓ Kohli *et al.* 2007, 2008

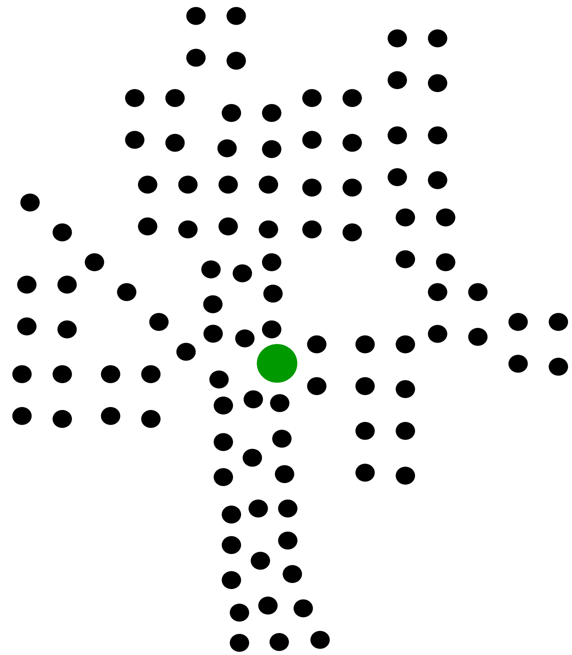
# Using Higher Order Information

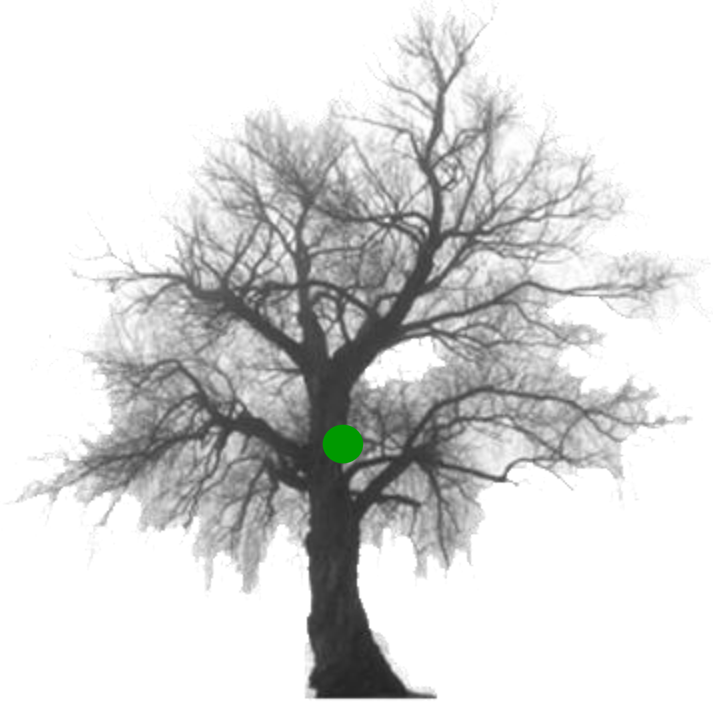
# Region Based Model

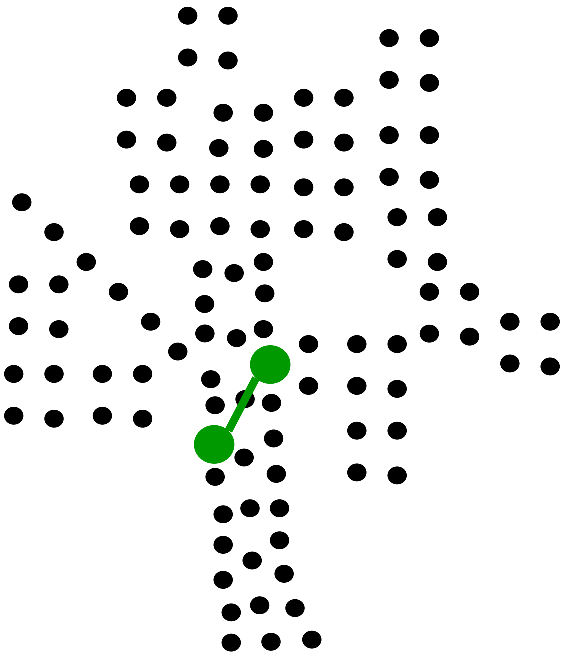




$$\mathbf{y}^* = \arg\max_{\mathbf{y}} \sum_{i=1}^{N} \phi_i(y_i, \mathbf{x}) + \sum_{(ij)\in E} \phi_{ij}(y_i, y_j, \mathbf{x}) + \sum_{c \in S} \phi_c(\mathbf{y}_c, \mathbf{x})$$

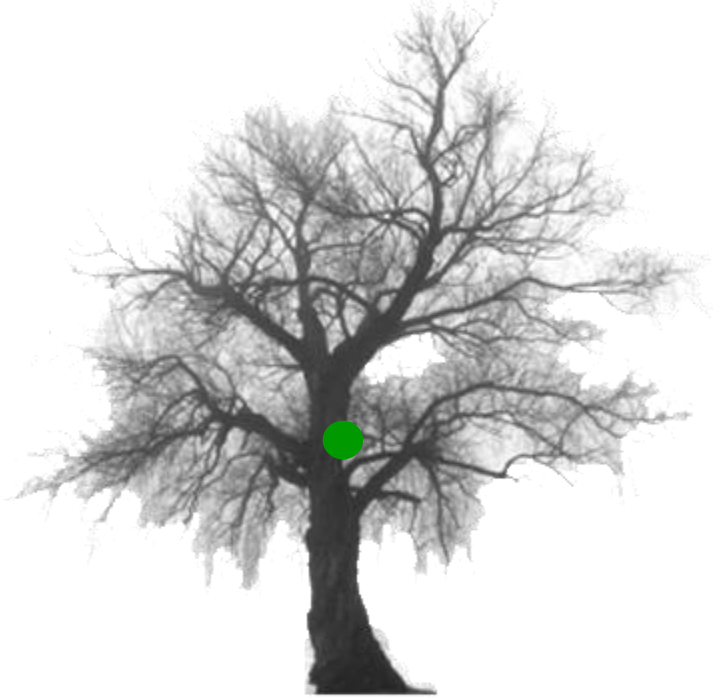
# Region Based Model

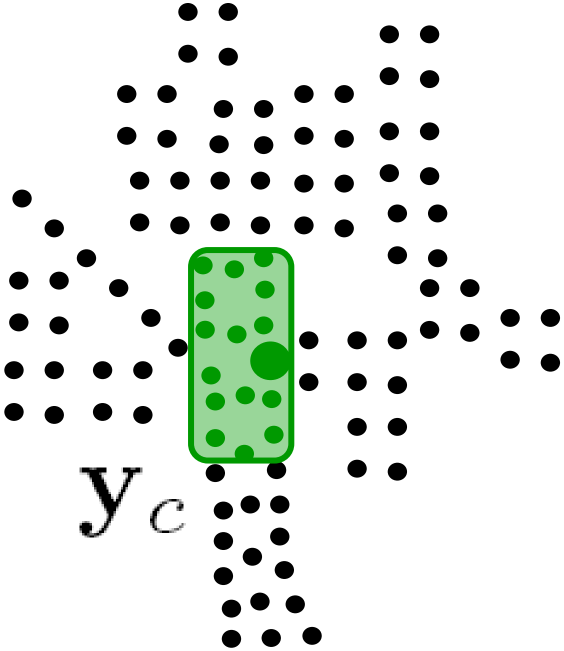




$$\mathbf{y}^* = \arg\max_{\mathbf{y}} \sum_{i=1}^{N} \phi_i(y_i, \mathbf{x}) + \sum_{(ij)\in E} \phi_{ij}(y_i, y_j, \mathbf{x}) + \sum_{c\in S} \phi_c(\mathbf{y}_c, \mathbf{x})$$

# Region Based Model

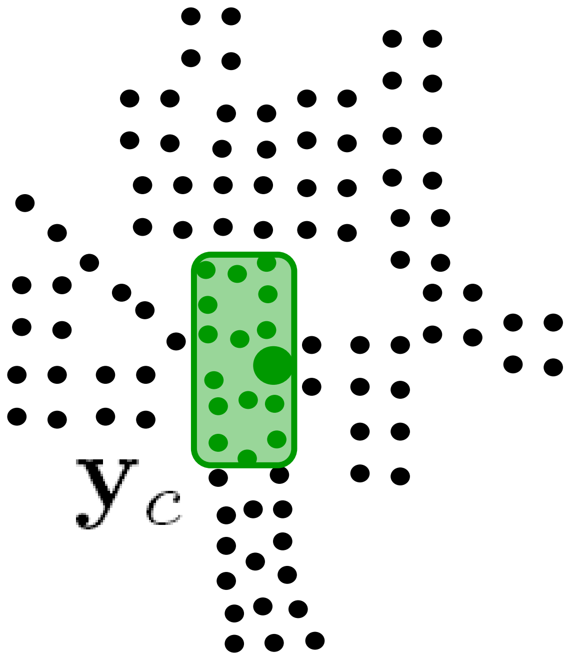




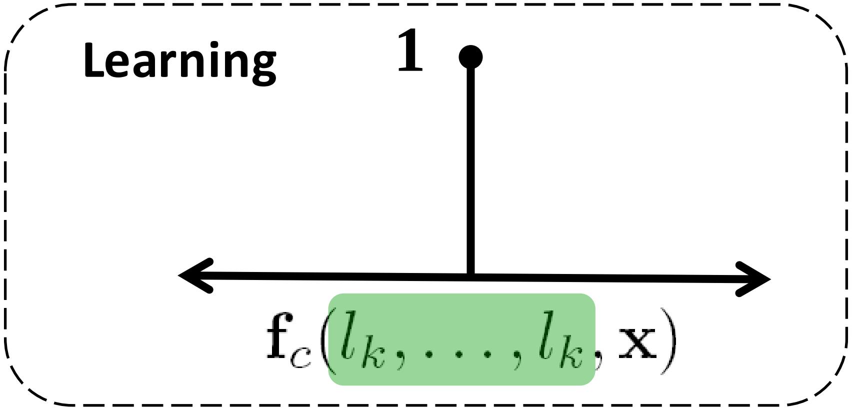

$$\mathbf{y}^* = \arg\max_{\mathbf{y}} \sum_{i=1}^{N} \phi_i(y_i, \mathbf{x}) + \sum_{(ij)\in E} \phi_{ij}(y_i, y_j, \mathbf{x}) + \sum_{c \in S} \phi_c(\mathbf{y}_c, \mathbf{x})$$

- Inference: graph-cut procedure
  - $P^n$ Potts model (Kohli *et al.* 2007)

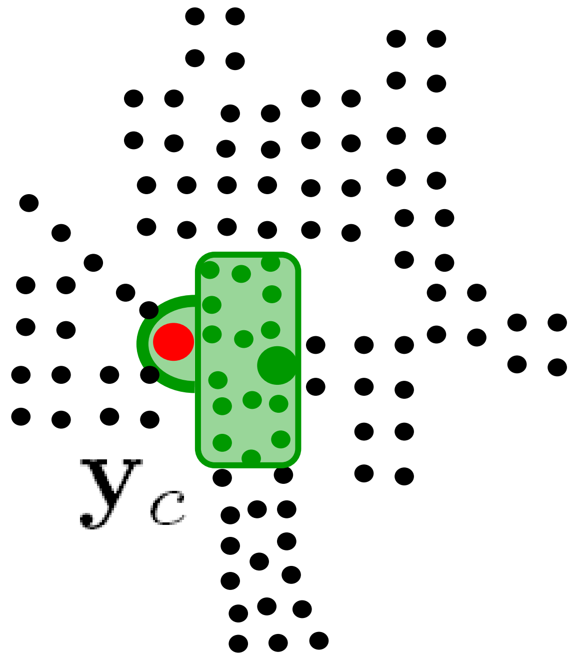
# How To Train The Model




$$\phi_c(l_k, \ldots, l_k, \mathbf{x}) = \sum_t \alpha_t h_t(\mathbf{f}_c(l_k, \ldots, l_k, \mathbf{x}))$$



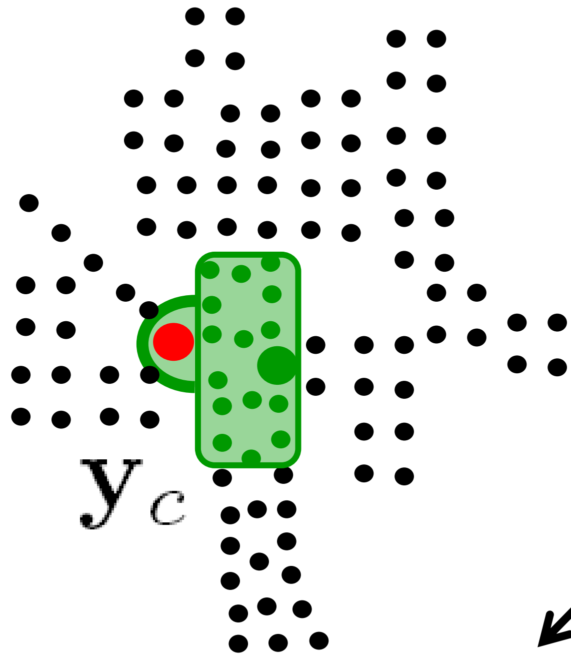
# How To Train The Model




$$\phi_c(l_l, l_k, \ldots, l_k, \mathbf{x}) = 0$$

**Learning**
(ignores features from clique *c*)

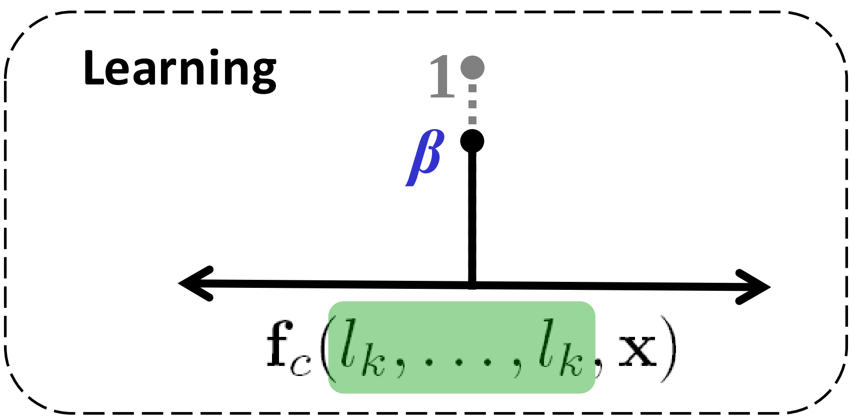
# How To Train The Model




**Robust** $P^n$ Potts

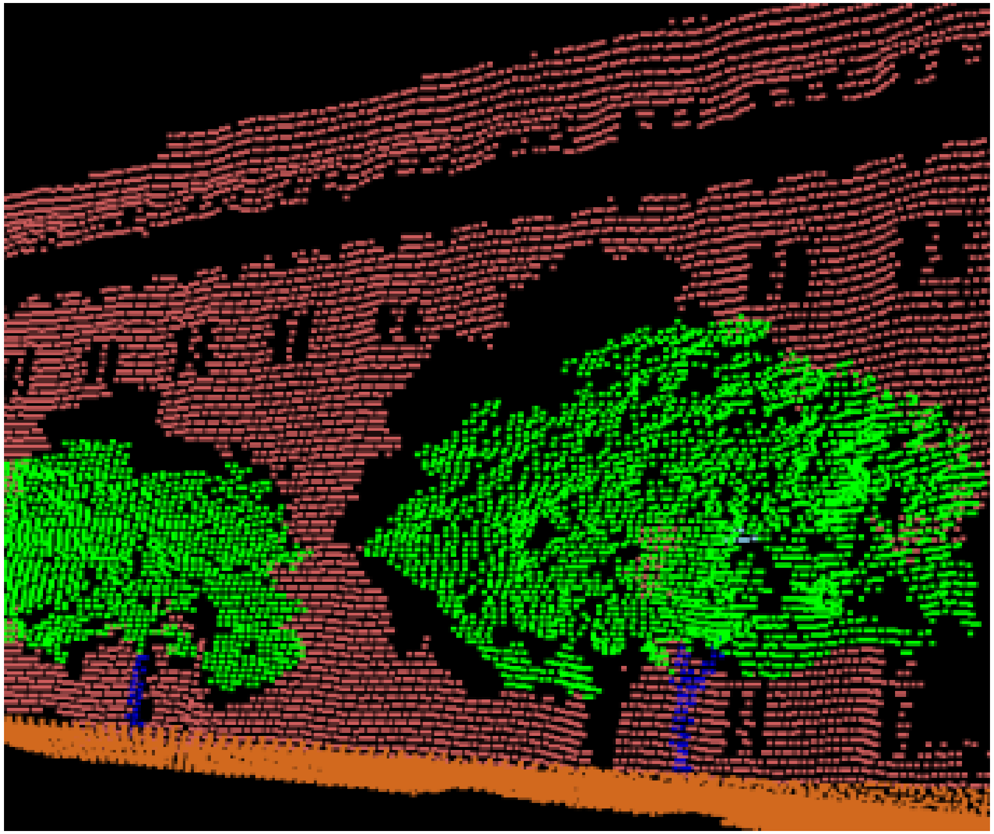
Kohli *et al.* 2008

$$\phi_c(l_l, l_k, \ldots, l_k, \mathbf{x}) = \boldsymbol{\beta}\, \phi_c(l_k, \ldots, l_k, \mathbf{x})$$



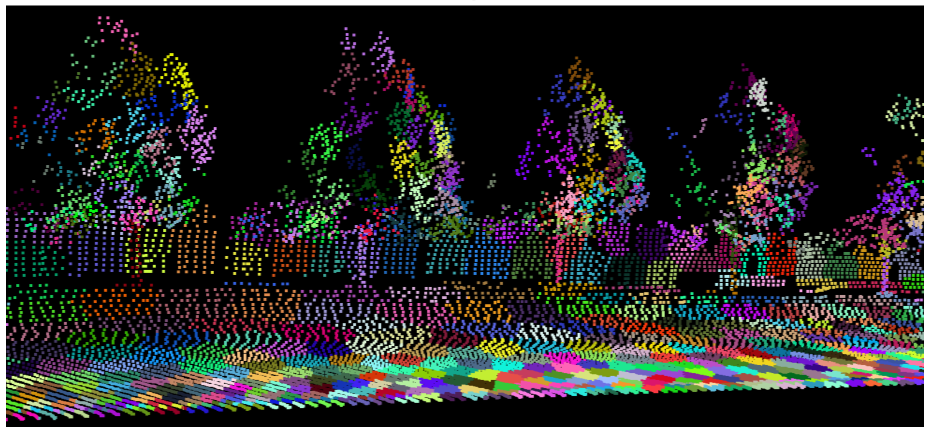
# Experimental Analysis

- **3-D Point Cloud Classification**
- Geometry Surface Estimation

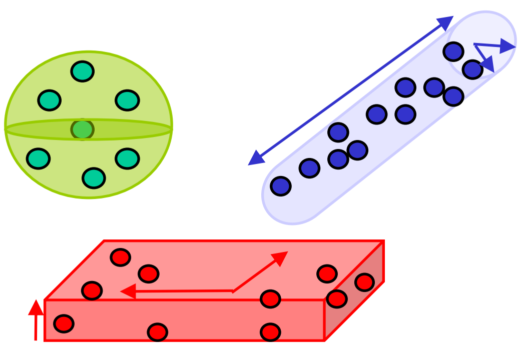


# Random Field Description

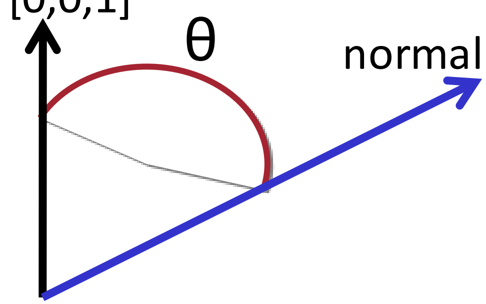
- **Nodes:** 3-D points
- **Edges:** 5-Nearest Neighbors
- **Cliques:** Two K-means segmentations



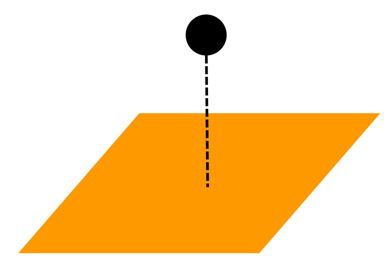
- **Features**



Local shape

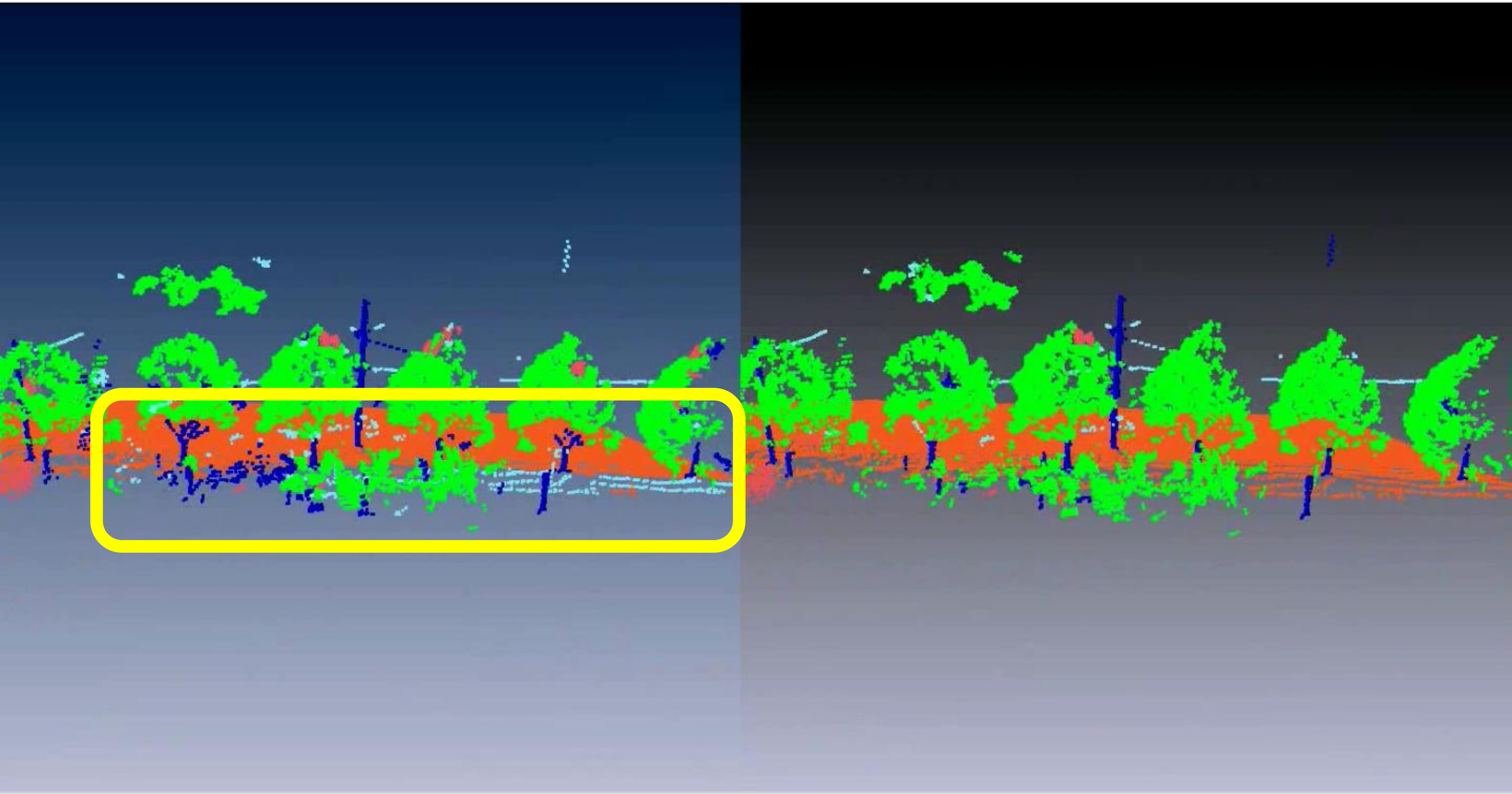



Orientation



Elevation

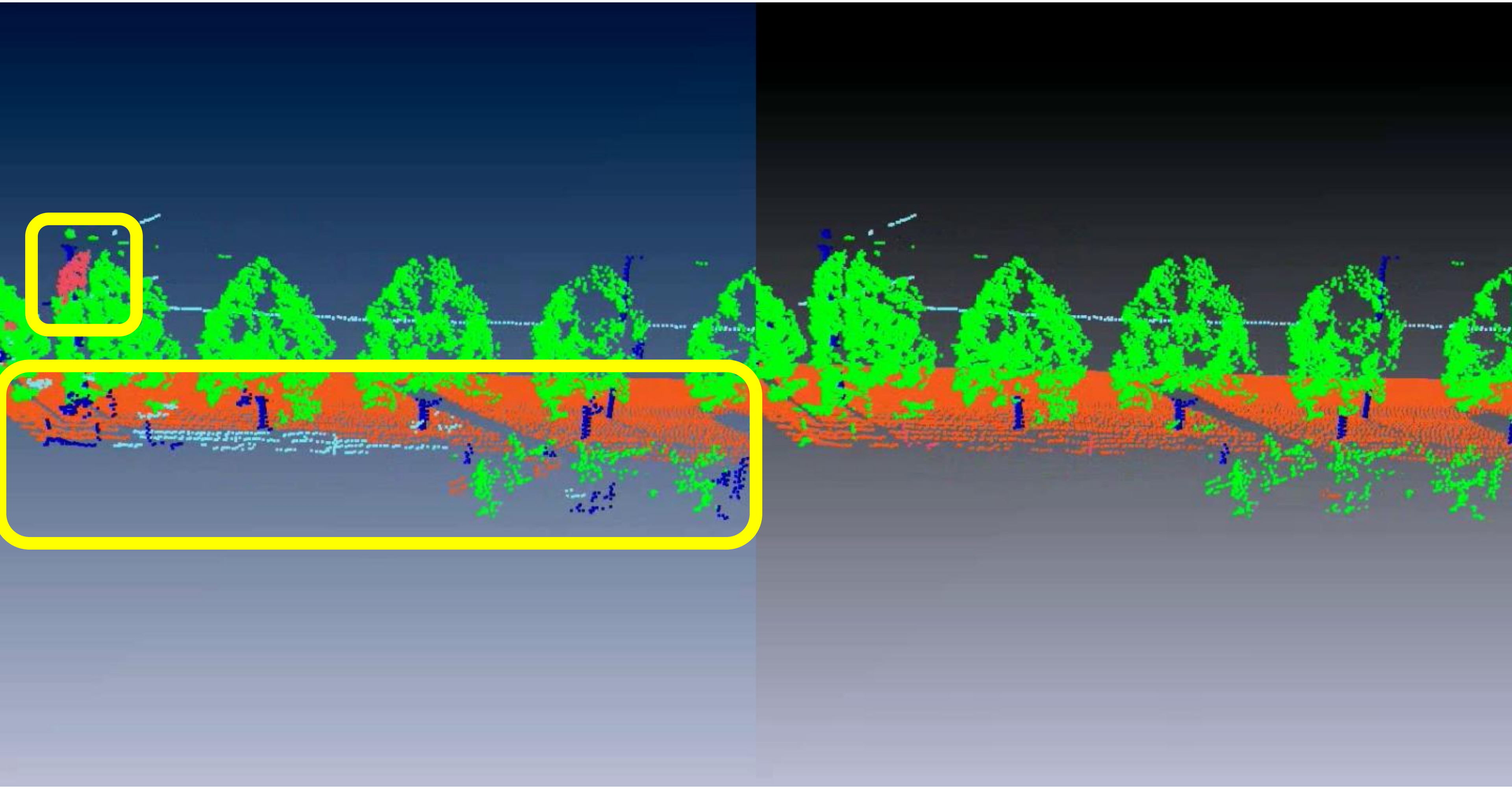
# Qualitative Comparisons



Parametric

**Functional** (this work)

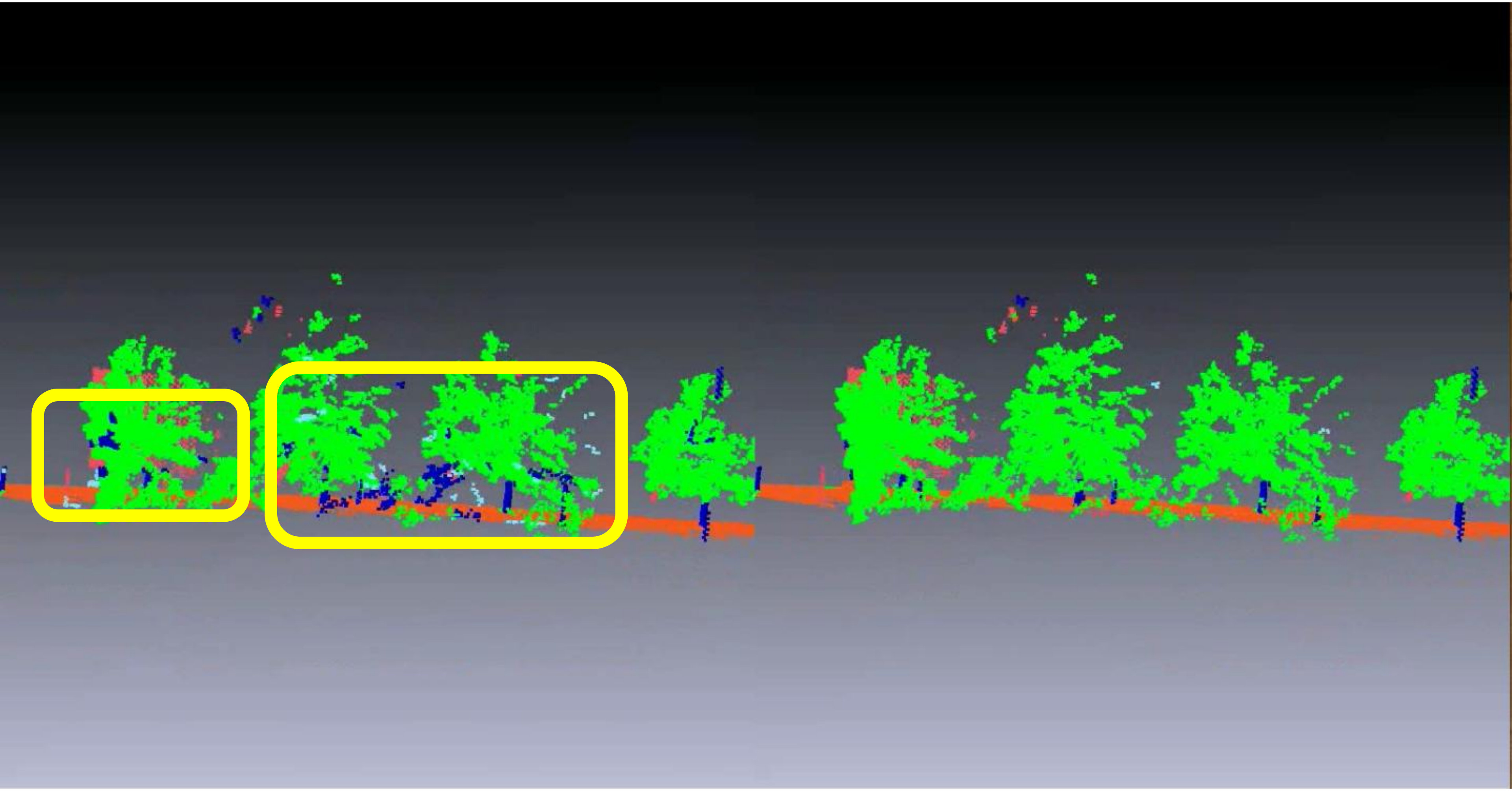
# Qualitative Comparisons



Parametric

**Functional** (this work)

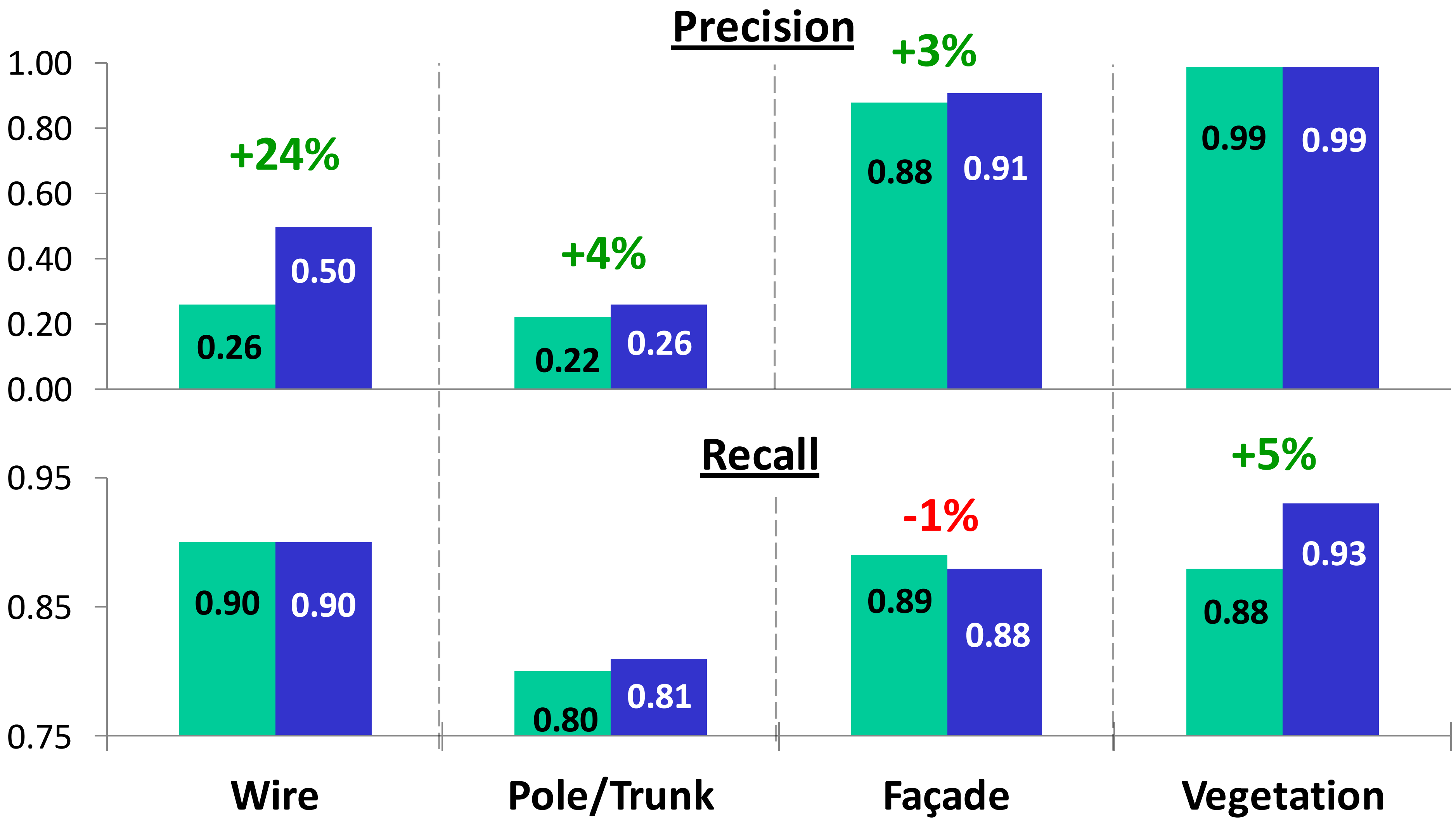
# Qualitative Comparisons



Parametric **Functional** (this work)

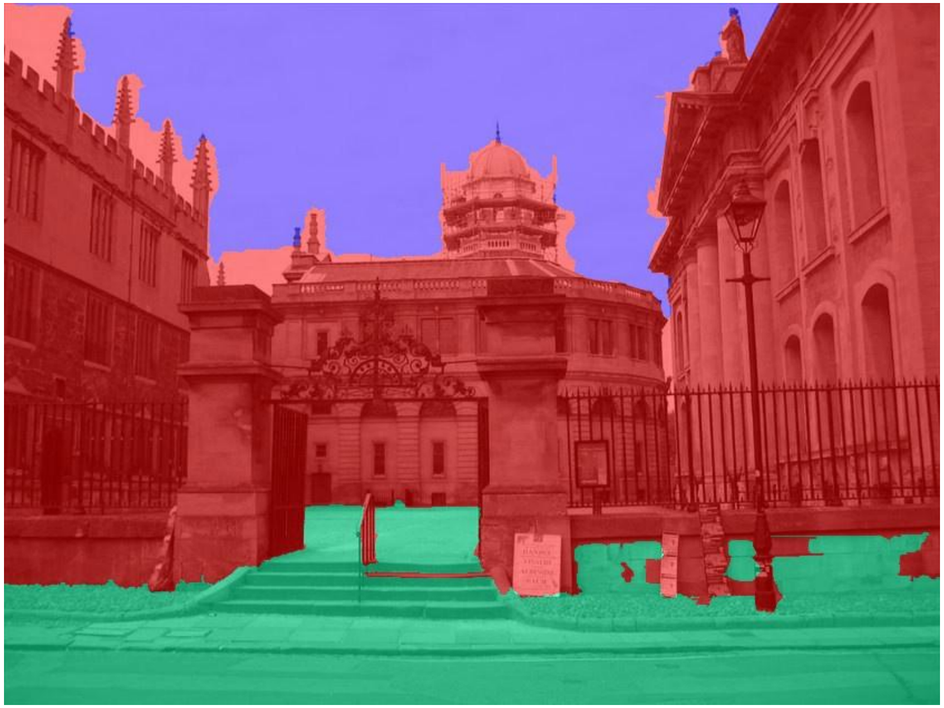
# Quantitative Results (1.2 M pts)

- Macro* AP: Parametric 64.3% **Functional 71.5%**



# Experimental Analysis

- 3-D Point Cloud Classification
- **Geometry Surface Estimation**



# Random Field Description

- **Nodes**: Superpixels (Hoiem *et al.* 2007)
- **Edges**: (none)
- **Cliques**: 15 segmentations (Hoiem *et al.* 2007)


More Robust

- **Features** (Hoiem *et al.* 2007)
  - Perspective, color, texture, etc.
- **1,000** dimensional space

# Quantitative Comparisons

| | Ground | Vertical | Sky |
|---|---|---|---|
| Ground | 0.74 | 0.24 | 0.02 |
| Vertical | 0.24 | 0.70 | 0.07 |
| Sky | 0.03 | 0.20 | 0.78 |
| | Accuracy: 72.8% | | |

Parametric (Potts)

| | Ground | Vertical | Sky |
|---|---|---|---|
| Ground | 0.83 | 0.16 | 0.00 |
| Vertical | 0.09 | **0.89** | 0.02 |
| Sky | 0.00 | 0.10 | 0.89 |
| | Accuracy: **87.1**% | | |

Hoiem *et al.* 2007

| | Ground | Vertical | Sky |
|---|---|---|---|
| Ground | 0.84 | 0.15 | 0.01 |
| Vertical | 0.13 | 0.83 | 0.04 |
| Sky | 0.02 | 0.07 | 0.91 |
| | Accuracy: 84.9% | | |

**Functional** (Potts)

| | Ground | Vertical | Sky |
|---|---|---|---|
| Ground | **0.85** | 0.14 | 0.01 |
| Vertical | 0.13 | 0.84 | 0.04 |
| Sky | 0.01 | 0.05 | **0.94** |
| | Accuracy: 86.0% | | |

**Functional** (Robust Potts)

# Qualitative Comparisons

Parametric (Potts)

**Functional** (Potts)

# Qualitative Comparisons

Parametric (Potts)

**Functional** (Potts)

**Functional** (Robust Potts)

# Qualitative Comparisons

Parametric (Potts)

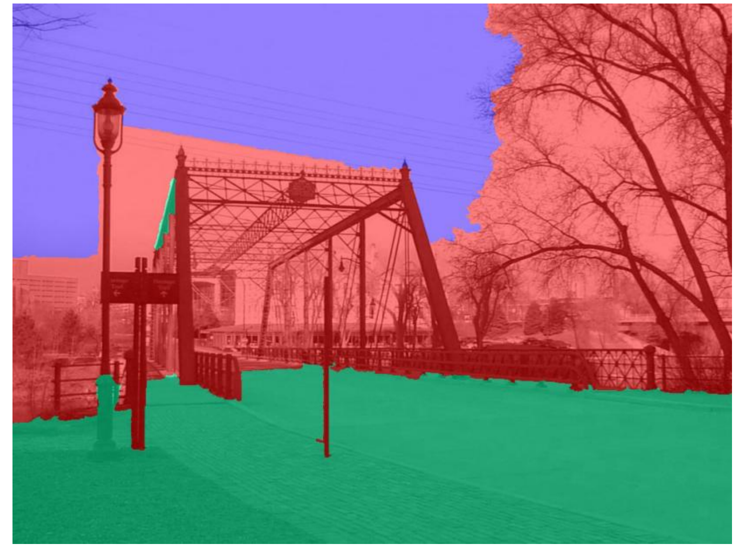
Hoiem *et al.* 2007

**Functional** (Potts)

**Functional** (Robust Potts)

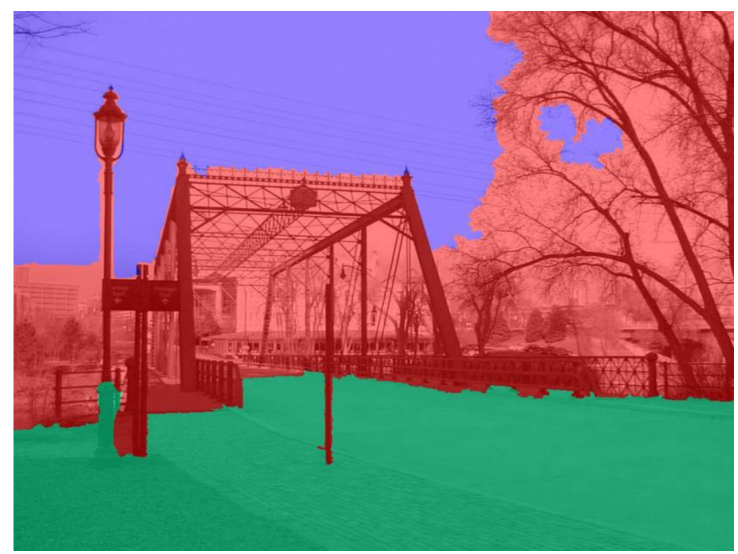
# Qualitative Comparisons

Parametric (Potts)



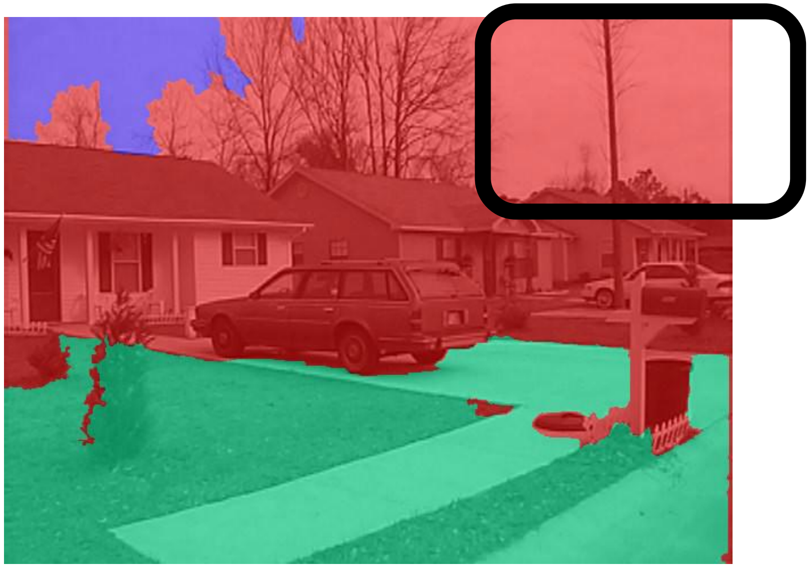
Hoiem *et al.* 2007

**Functional** (Potts)

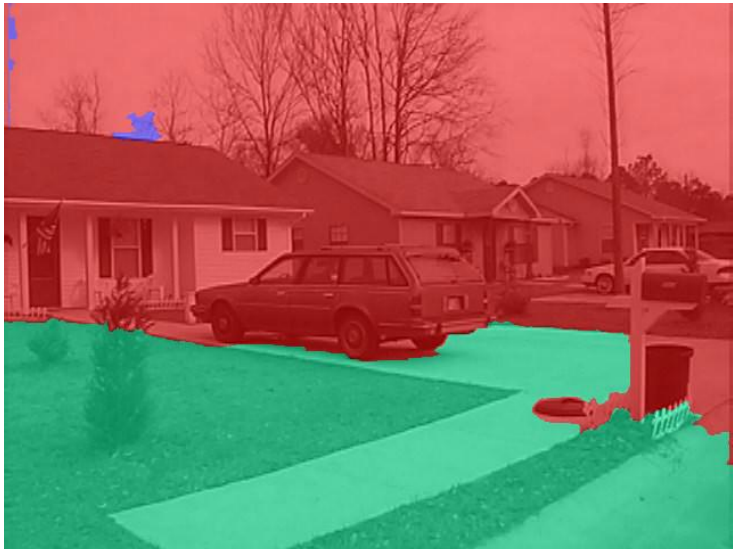


**Functional** (Robust Potts)

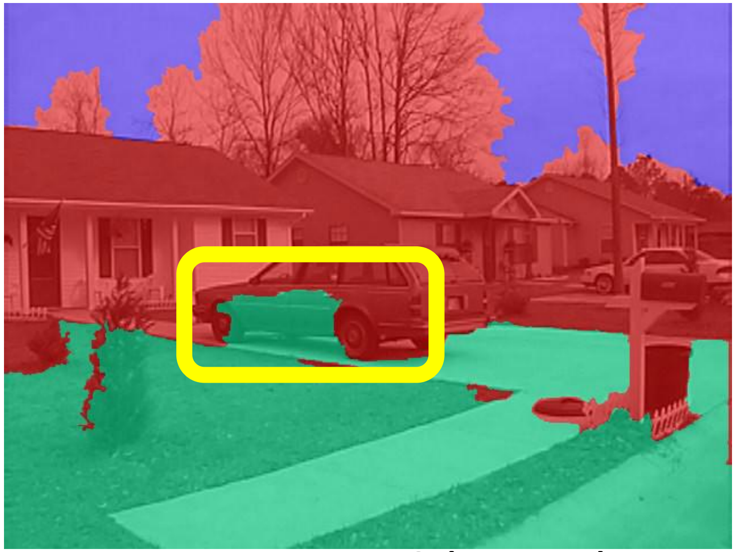
# Qualitative Comparisons



Parametric (Potts)



Hoiem *et al.* 2007



**Functional** (Potts)

**Functional** (Robust Potts)

# Conclusion

- **Effective** max-margin learning of high-order CRFs
  - Especially for large dimensional spaces
  - Robust Potts interactions
  - Easy to implement

- Future work
  - Non-linear potentials (decision tree/random forest)
  - New inference procedures:
    - ✓Komodakisand Paragios 2009
    - ✓Ishikawa 2009
    - ✓Gould *et al.* 2009
    - ✓Rother *et al.* 2009

# Thank you

- Acknowledgements
  - U. S. Army Research Laboratory
  - Siebel Scholars Foundation
  - S. K. Divalla, N. Ratliff, B. Becker

- Questions?