Efficient Inference in Fully Connected CRFs with Gaussian Edge Potentials

Philipp Krähenbühl   Vladlen Koltun

philkr@stanford.edu   vladlen@stanford.edu

Department of Computer Science, Stanford University

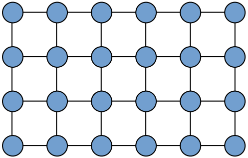
December 14, 2011

Multi-class image segmentation

Assign a class label to each pixel in the image

CRF models in multi-class image segmentation

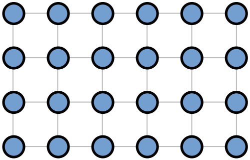
$$E(\mathbf{x}) = \sum_i \underbrace{\psi_u(x_i)}_{\text{unary term}} + \sum_i \sum_{j \in \mathcal{N}_i} \underbrace{\psi_p(x_i, x_j)}_{\text{pairwise term}}$$



- MAP inference in conditional random field
- Unary term
  - From classifier
  - TextonBoost [Shotton et al. 09]
- Pairwise term
  - Consistent labeling

CRF models in multi-class image segmentation

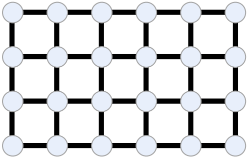
$$E(\mathbf{x}) = \sum_i \underbrace{\psi_u(x_i)}_{\text{unary term}} + \sum_i \sum_{j \in \mathcal{N}_i} \underbrace{\psi_p(x_i, x_j)}_{\text{pairwise term}}$$



- MAP inference in conditional random field
- Unary term
  - From classifier
  - TextonBoost [Shotton et al. 09]
- Pairwise term
  - Consistent labeling

CRF models in multi-class image segmentation

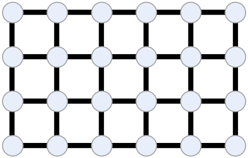
$$E(\mathbf{x}) = \sum_i \underbrace{\psi_u(x_i)}_{\text{unary term}} + \sum_i \sum_{j \in \mathcal{N}_i} \underbrace{\psi_p(x_i, x_j)}_{\text{pairwise term}}$$



- MAP inference in conditional random field
- Unary term
  - From classifier
  - TextonBoost [Shotton et al. 09]
- Pairwise term
  - Consistent labeling

Adjacency CRF models

$$E(\mathbf{x}) = \sum_i \underbrace{\psi_u(x_i)}_{\text{unary term}} + \sum_i \sum_{j \in \mathcal{N}_i} \underbrace{\psi_p(x_i, x_j)}_{\text{pairwise term}}$$

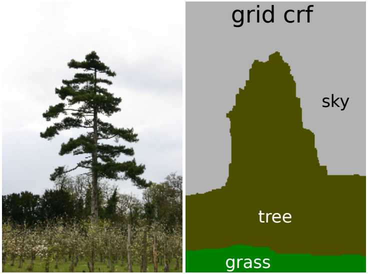


- Pairwise term
  - Neighboring pixels
  - Color-sensitive Potts model

$$\psi_p(x_i, x_j) = 1_{[x_i \neq x_j]} \left(w^{(1)} \exp\left(-\frac{|\mathbf{I}_i - \mathbf{I}_j|^2}{2\theta_\beta^2} \right) + w^{(2)} \right)$$

Adjacency CRF models

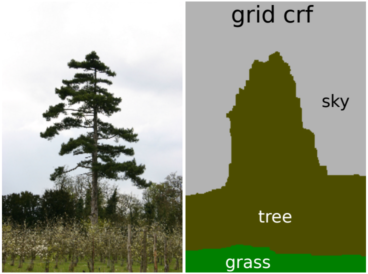
$$E(\mathbf{x}) = \sum_i \underbrace{\psi_u(x_i)}_{\text{unary term}} + \sum_i \sum_{j \in \mathcal{N}_i} \underbrace{\psi_p(x_i, x_j)}_{\text{pairwise term}}$$



- Efficient inference
  - 1 second for 50′000 variables
- Limited expressive power
- Only local interactions
- Excessive smoothing of object boundaries
  - Shrinking bias

Adjacency CRF models

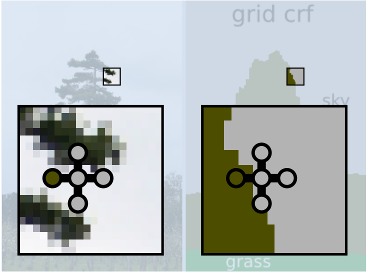
$$E(\mathbf{x}) = \sum_i \underbrace{\psi_u(x_i)}_{\text{unary term}} + \sum_i \sum_{j \in \mathcal{N}_i} \underbrace{\psi_p(x_i, x_j)}_{\text{pairwise term}}$$



- Efficient inference
  - 1 second for 50′000 variables
- Limited expressive power
- Only local interactions
- Excessive smoothing of object boundaries
  - Shrinking bias

Adjacency CRF models

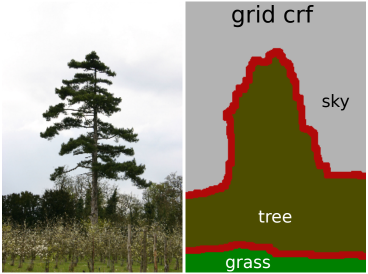
$$E(\mathbf{x}) = \sum_i \underbrace{\psi_u(x_i)}_{\text{unary term}} + \sum_i \sum_{j \in \mathcal{N}_i} \underbrace{\psi_p(x_i, x_j)}_{\text{pairwise term}}$$



- Efficient inference
  - 1 second for 50′000 variables
- Limited expressive power
- Only local interactions
- Excessive smoothing of object boundaries
  - Shrinking bias

Adjacency CRF models

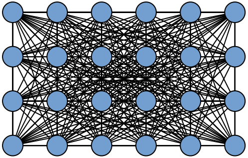
$$E(\mathbf{x}) = \sum_i \underbrace{\psi_u(x_i)}_{\text{unary term}} + \sum_i \sum_{j \in \mathcal{N}_i} \underbrace{\psi_p(x_i, x_j)}_{\text{pairwise term}}$$



- Efficient inference
  - 1 second for 50′000 variables
- Limited expressive power
- Only local interactions
- Excessive smoothing of object boundaries
  - Shrinking bias

Fully connected CRF

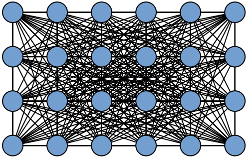
$$E(\mathbf{x}) = \sum_i \underbrace{\psi_u(x_i)}_{\text{unary term}} + \sum_i \sum_{j>i} \underbrace{\psi_p(x_i, x_j)}_{\text{pairwise term}}$$



- Every node is connected to every other node
  - Connections weighted differently

Fully connected CRF

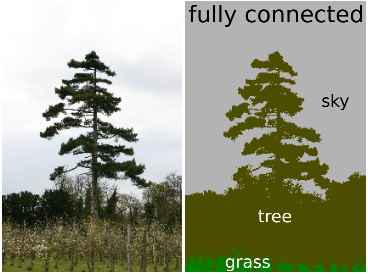
$$E(\mathbf{x}) = \sum_i \underbrace{\psi_u(x_i)}_{\text{unary term}} + \sum_i \sum_{j>i} \underbrace{\psi_p(x_i, x_j)}_{\text{pairwise term}}$$



- Every node is connected to every other node
  - Connections weighted differently

Fully connected CRF

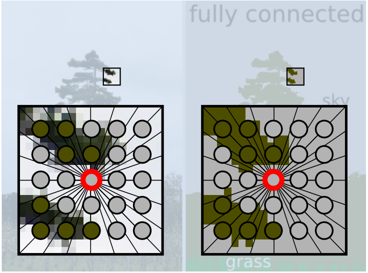
$$E(\mathbf{x}) = \sum_i \underbrace{\psi_u(x_i)}_{\text{unary term}} + \sum_i \sum_{j>i} \underbrace{\psi_p(x_i, x_j)}_{\text{pairwise term}}$$



- Long-range interactions
- No more shrinking bias

Fully connected CRF

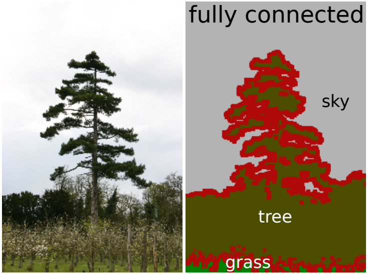
$$E(\mathbf{x}) = \sum_i \underbrace{\psi_u(x_i)}_{\text{unary term}} + \sum_i \sum_{j>i} \underbrace{\psi_p(x_i, x_j)}_{\text{pairwise term}}$$



- Long-range interactions
- No more shrinking bias

Fully connected CRF

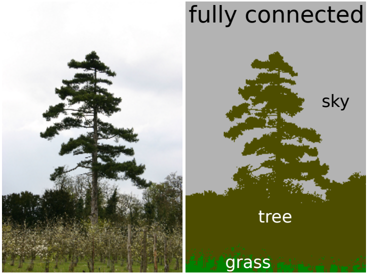
$$E(\mathbf{x}) = \sum_i \underbrace{\psi_u(x_i)}_{\text{unary term}} + \sum_i \sum_{j>i} \underbrace{\psi_p(x_i, x_j)}_{\text{pairwise term}}$$



- Long-range interactions
- No more shrinking bias

Fully connected CRF

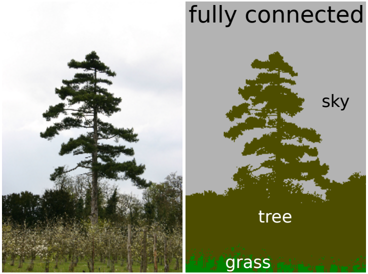
$$E(\mathbf{x}) = \sum_i \underbrace{\psi_u(x_i)}_{\text{unary term}} + \sum_i \sum_{j>i} \underbrace{\psi_p(x_i, x_j)}_{\text{pairwise term}}$$



- Region-based [Rabinovich et al. 07, Galleguillos et al. 08, Toyoda & Hasegawa 08, Payet & Todorovic 10]
  - Tractable up to hundreds of variables
- Pixel-based
  - Tens of thousands of variables
    - Billions of edges
  - Computationally expensive

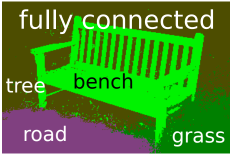
Fully connected CRF

$$E(\mathbf{x}) = \sum_i \underbrace{\psi_u(x_i)}_{\text{unary term}} + \sum_i \sum_{j>i} \underbrace{\psi_p(x_i, x_j)}_{\text{pairwise term}}$$



- Region-based [Rabinovich et al. 07, Galleguillos et al. 08, Toyoda & Hasegawa 08, Payet & Todorovic 10]
  - Tractable up to hundreds of variables
- Pixel-based
  - Tens of thousands of variables
    - Billions of edges
  - Computationally expensive

- Inference in 0.2 seconds
  - 50′000 variables
  - MCMC inference: 36 hrs
- Pairwise potentials: linear combinations of Gaussians



- Inference in 0.2 seconds
  - 50′000 variables
  - MCMC inference: 36 hrs
- Pairwise potentials: linear combinations of Gaussians

Model definition

$$E(\mathbf{x}) = \sum_i \underbrace{\psi_u(x_i)}_{\text{unary term}} + \sum_i \sum_{j>i} \underbrace{\psi_p(x_i, x_j)}_{\text{pairwise term}}$$

Gaussian edge potentials

$$\psi_p(x_i, x_j) = \mu(x_i, x_j) \sum_{m=1}^{K} w^{(m)} k^{(m)}(\mathbf{f}_i, \mathbf{f}_j)$$

- Label compatibility function μ
- Linear combination of Gaussian kernels

$$k^{(m)}(\mathbf{f}_i, \mathbf{f}_j) = \exp(-\frac{1}{2}(\mathbf{f}_i - \mathbf{f}_j)\Sigma^{(m)}(\mathbf{f}_i - \mathbf{f}_j))$$

- Arbitrary feature space $\mathbf{f}_i$

Model definition

$$E(\mathbf{x}) = \sum_i \underbrace{\psi_u(x_i)}_{\text{unary term}} + \sum_i \sum_{j>i} \underbrace{\psi_p(x_i, x_j)}_{\text{pairwise term}}$$

Gaussian edge potentials

$$\psi_p(x_i, x_j) = \mu(x_i, x_j) \sum_{m=1}^{K} w^{(m)} k^{(m)}(\mathbf{f}_i, \mathbf{f}_j)$$

- Label compatibility function μ
- Linear combination of Gaussian kernels

$$k^{(m)}(\mathbf{f}_i, \mathbf{f}_j) = \exp(-\frac{1}{2}(\mathbf{f}_i - \mathbf{f}_j)\Sigma^{(m)}(\mathbf{f}_i - \mathbf{f}_j))$$

- Arbitrary feature space $\mathbf{f}_i$

Model definition

$$E(\mathbf{x}) = \sum_i \underbrace{\psi_u(x_i)}_{\text{unary term}} + \sum_i \sum_{j>i} \underbrace{\psi_p(x_i, x_j)}_{\text{pairwise term}}$$

Gaussian edge potentials

$$\psi_p(x_i, x_j) = \boxed{\mu(x_i, x_j)} \sum_{m=1}^{K} w^{(m)} k^{(m)}(\mathbf{f}_i, \mathbf{f}_j)$$

- Label compatibility function μ
- Linear combination of Gaussian kernels

$$k^{(m)}(\mathbf{f}_i, \mathbf{f}_j) = \exp(-\frac{1}{2}(\mathbf{f}_i - \mathbf{f}_j)\Sigma^{(m)}(\mathbf{f}_i - \mathbf{f}_j))$$

- Arbitrary feature space $\mathbf{f}_i$

Model definition

$$E(\mathbf{x}) = \sum_i \underbrace{\psi_u(x_i)}_{\text{unary term}} + \sum_i \sum_{j>i} \underbrace{\psi_p(x_i, x_j)}_{\text{pairwise term}}$$

Gaussian edge potentials

$$\psi_p(x_i, x_j) = \mu(x_i, x_j) \boxed{\sum_{m=1}^{K} w^{(m)} k^{(m)}(\mathbf{f}_i, \mathbf{f}_j)}$$

- Label compatibility function μ
- Linear combination of Gaussian kernels

$$k^{(m)}(\mathbf{f}_i, \mathbf{f}_j) = \exp(-\frac{1}{2}(\mathbf{f}_i - \mathbf{f}_j)\Sigma^{(m)}(\mathbf{f}_i - \mathbf{f}_j))$$

- Arbitrary feature space $\mathbf{f}_i$

Model definition

$$E(\mathbf{x}) = \sum_i \underbrace{\psi_u(x_i)}_{\text{unary term}} + \sum_i \sum_{j>i} \underbrace{\psi_p(x_i, x_j)}_{\text{pairwise term}}$$

Gaussian edge potentials

$$\psi_p(x_i, x_j) = \mu(x_i, x_j) \sum_{m=1}^{K} w^{(m)} k^{(m)}(\boxed{\mathbf{f}_i}, \boxed{\mathbf{f}_j})$$

- Label compatibility function μ
- Linear combination of Gaussian kernels

$$k^{(m)}(\mathbf{f}_i, \mathbf{f}_j) = \exp(-\frac{1}{2}(\mathbf{f}_i - \mathbf{f}_j)\Sigma^{(m)}(\mathbf{f}_i - \mathbf{f}_j))$$

- Arbitrary feature space $\mathbf{f}_i$

Detailed model definition

$$\psi_p(x_i, x_j) = \mu(x_i, x_j) \left(w^{(1)} \exp\left(-\frac{|\mathbf{p}_i - \mathbf{p}_j|}{2\theta_\alpha^2} - \frac{|I_i - I_j|}{2\theta_\beta^2}\right) + w^{(2)} \exp\left(-\frac{|\mathbf{p}_i - \mathbf{p}_j|}{2\theta_\gamma^2}\right) \right)$$

- Label compatibility
  - Potts model: $\mu(x_i, x_j) = 1_{[x_i \neq x_j]}$
  - Semi-metric model: $\mu(x_i, x_j)$ learned from data
- Appearance kernel
  - Color-sensitive model
- Local smoothness
  - Discourages pixel level noise

Detailed model definition

$$\psi_p(x_i, x_j) = \mu(x_i, x_j) \left(w^{(1)} \exp(-\frac{|\mathbf{p}_i - \mathbf{p}_j|}{2\theta_\alpha^2} - \frac{|I_i - I_j|}{2\theta_\beta^2}) + w^{(2)} \exp(-\frac{|\mathbf{p}_i - \mathbf{p}_j|}{2\theta_\gamma^2}) \right)$$

- Label compatibility
  - Potts model: $\mu(x_i, x_j) = 1_{[x_i \neq x_j]}$
  - Semi-metric model: $\mu(x_i, x_j)$ learned from data
- Appearance kernel
  - Color-sensitive model
- Local smoothness
  - Discourages pixel level noise

Detailed model definition

$$\psi_p(x_i, x_j) = \mu(x_i, x_j) \left(w^{(1)} \exp(-\frac{|\mathbf{p}_i - \mathbf{p}_j|}{2\theta_\alpha^2} - \frac{|I_i - I_j|}{2\theta_\beta^2}) + w^{(2)} \exp(-\frac{|\mathbf{p}_i - \mathbf{p}_j|}{2\theta_\gamma^2}) \right)$$

- Label compatibility
  - Potts model: $\mu(x_i, x_j) = 1_{[x_i \neq x_j]}$
  - Semi-metric model: $\mu(x_i, x_j)$ learned from data
- Appearance kernel
  - Color-sensitive model
- Local smoothness
  - Discourages pixel level noise

Detailed model definition

$$\psi_p(x_i, x_j) = \mu(x_i, x_j) \left(w^{(1)} \exp(-\frac{|\mathbf{p}_i - \mathbf{p}_j|}{2\theta_\alpha^2} - \frac{|I_i - I_j|}{2\theta_\beta^2}) + w^{(2)} \exp(-\frac{|\mathbf{p}_i - \mathbf{p}_j|}{2\theta_\gamma^2}) \right)$$

- Label compatibility
  - Potts model: $\mu(x_i, x_j) = 1_{[x_i \neq x_j]}$
  - Semi-metric model: $\mu(x_i, x_j)$ learned from data
- Appearance kernel
  - Color-sensitive model
- Local smoothness
  - Discourages pixel level noise

Detailed model definition

$$\psi_p(x_i, x_j) = \mu(x_i, x_j) \left(w^{(1)} \exp(-\frac{|\mathbf{p}_i - \mathbf{p}_j|}{2\theta_\alpha^2} - \frac{|I_i - I_j|}{2\theta_\beta^2}) + w^{(2)} \exp(-\frac{|\mathbf{p}_i - \mathbf{p}_j|}{2\theta_\gamma^2}) \right)$$

- Label compatibility
  - Potts model: $\mu(x_i, x_j) = 1_{[x_i \neq x_j]}$
  - Semi-metric model: $\mu(x_i, x_j)$ learned from data
- Appearance kernel
  - Color-sensitive model
- Local smoothness
  - Discourages pixel level noise

Inference

Find the most likely assignment (MAP)

$$\hat{x} = \underset{\mathbf{x}}{\operatorname{argmax}}\, P(\mathbf{x}) \qquad \text{where} \quad P(\mathbf{x}) = \exp(-E(\mathbf{x}))$$

Mean field approximation

- Find $Q(\mathbf{x}) = \prod_i Q(x_i)$ close to $P(\mathbf{x})$ in terms of KL-divergence $D(Q\|P)$
- $\hat{x}_i \approx \operatorname{argmax}_{x_i} Q(x_i)$

Inference

Find the most likely assignment (MAP)

$$\hat{x} = \operatorname*{argmax}_{\mathbf{x}} P(\mathbf{x}) \qquad \text{where} \quad P(\mathbf{x}) = \exp(-E(\mathbf{x}))$$

Mean field approximation

- Find $Q(\mathbf{x}) = \prod_i Q(x_i)$ close to $P(\mathbf{x})$ in terms of KL-divergence $D(Q\|P)$
- $\hat{x}_i \approx \operatorname{argmax}_{x_i} Q(x_i)$

Inference

Find the most likely assignment (MAP)

$$\hat{x} = \underset{\mathbf{x}}{\operatorname{argmax}}\, P(\mathbf{x}) \qquad \text{where} \quad P(\mathbf{x}) = \exp(-E(\mathbf{x}))$$

Mean field approximation

- Find $Q(\mathbf{x}) = \prod_i Q(x_i)$ close to $P(\mathbf{x})$ in terms of KL-divergence $D(Q \| P)$
- $\hat{x}_i \approx \operatorname{argmax}_{x_i} Q(x_i)$

Inference

Find the most likely assignment (MAP)

$$\hat{x} = \underset{\mathbf{x}}{\operatorname{argmax}} P(\mathbf{x}) \qquad \text{where} \quad P(\mathbf{x}) = \exp(-E(\mathbf{x}))$$

Mean field approximation

- Find $Q(\mathbf{x}) = \prod_i Q(x_i)$ close to $P(\mathbf{x})$ in terms of KL-divergence $D(Q \| P)$
- $\hat{x}_i \approx \operatorname{argmax}_{x_i} Q(x_i)$

Mean field approximation

$$Q_i(x_i = l) = \frac{1}{Z_i} \exp\left\{-\psi_u(x_i) - \sum_{l' \in \mathcal{L}} \mu(l, l') \sum_{m=1}^{K} w^{(m)} \sum_{j \neq i} k^{(m)}(\mathbf{f}_i, \mathbf{f}_j) Q_j(l')\right\}$$

- Initialize $Q_i(x_i) \leftarrow \frac{1}{Z_i} \exp\{-\phi_u(x_i)\}$
- **while** not converged
  - Message passing: $\tilde{Q}_i^{(m)}(l) \leftarrow \sum_{j \neq i} k^{(m)}(\mathbf{f}_i, \mathbf{f}_j) Q_j(l)$
  - Compatibility transform: $\hat{Q}_i(x_i) \leftarrow \sum_{l \in \mathcal{L}} \mu^{(m)}(x_i, l) \sum_m w^{(m)} \tilde{Q}_i^{(m)}(l)$
  - Local update: $Q_i(x_i) \leftarrow \exp\{-\psi_u(x_i) - \hat{Q}_i(x_i)\}$
  - Normalize: $Q_i(x_i)$

Mean field approximation

$$Q_i(x_i = l) = \frac{1}{Z_i} \exp\left\{-\psi_u(x_i) - \sum_{l' \in \mathcal{L}} \mu(l, l') \sum_{m=1}^{K} w^{(m)} \sum_{j \neq i} k^{(m)}(\mathbf{f}_i, \mathbf{f}_j) Q_j(l')\right\}$$

- Initialize $Q_i(x_i) \leftarrow \frac{1}{Z_i} \exp\{-\phi_u(x_i)\}$
- **while** not converged
  - Message passing: $\tilde{Q}_i^{(m)}(l) \leftarrow \sum_{j \neq i} k^{(m)}(\mathbf{f}_i, \mathbf{f}_j) Q_j(l)$
  - Compatibility transform: $\hat{Q}_i(x_i) \leftarrow \sum_{l \in \mathcal{L}} \mu^{(m)}(x_i, l) \sum_m w^{(m)} \tilde{Q}_i^{(m)}(l)$
  - Local update: $Q_i(x_i) \leftarrow \exp\{-\psi_u(x_i) - \hat{Q}_i(x_i)\}$
  - Normalize: $Q_i(x_i)$

Mean field approximation

$$Q_i(x_i = l) = \frac{1}{Z_i} \exp\left\{ -\psi_u(x_i) - \sum_{l' \in \mathcal{L}} \mu(l, l') \sum_{m=1}^{K} w^{(m)} \sum_{j \neq i} k^{(m)}(\mathbf{f}_i, \mathbf{f}_j) Q_j(l') \right\}$$

- Initialize $Q_i(x_i) \leftarrow \frac{1}{Z_i} \exp\{-\phi_u(x_i)\}$
- **while** not converged
  - Message passing: $\tilde{Q}_i^{(m)}(l) \leftarrow \sum_{j \neq i} k^{(m)}(\mathbf{f}_i, \mathbf{f}_j) Q_j(l)$
  - Compatibility transform: $\hat{Q}_i(x_i) \leftarrow \sum_{l \in \mathcal{L}} \mu^{(m)}(x_i, l) \sum_m w^{(m)} \tilde{Q}_i^{(m)}(l)$
  - Local update: $Q_i(x_i) \leftarrow \exp\{-\psi_u(x_i) - \hat{Q}_i(x_i)\}$
  - Normalize: $Q_i(x_i)$

Mean field approximation

$$Q_i(x_i = l) = \frac{1}{Z_i} \exp\left\{-\psi_u(x_i) - \sum_{l' \in \mathcal{L}} \mu(l, l') \sum_{m=1}^{K} w^{(m)} \sum_{j \neq i} k^{(m)}(\mathbf{f}_i, \mathbf{f}_j) Q_j(l')\right\}$$

- Initialize $Q_i(x_i) \leftarrow \frac{1}{Z_i} \exp\{-\phi_u(x_i)\}$
- **while** not converged
  - Message passing: $\tilde{Q}_i^{(m)}(l) \leftarrow \sum_{j \neq i} k^{(m)}(\mathbf{f}_i, \mathbf{f}_j) Q_j(l)$
  - Compatibility transform: $\hat{Q}_i(x_i) \leftarrow \sum_{l \in \mathcal{L}} \mu^{(m)}(x_i, l) \sum_m w^{(m)} \tilde{Q}_i^{(m)}(l)$
  - Local update: $Q_i(x_i) \leftarrow \exp\{-\psi_u(x_i) - \hat{Q}_i(x_i)\}$
  - Normalize: $Q_i(x_i)$

Mean field approximation

$$Q_i(x_i = l) = \frac{1}{Z_i} \exp\left\{-\psi_u(x_i) - \sum_{l' \in \mathcal{L}} \mu(l, l') \sum_{m=1}^{K} w^{(m)} \sum_{j \neq i} k^{(m)}(\mathbf{f}_i, \mathbf{f}_j) Q_j(l')\right\}$$

- Initialize $Q_i(x_i) \leftarrow \frac{1}{Z_i} \exp\{-\phi_u(x_i)\}$
- **while** not converged
  - Message passing: $\tilde{Q}_i^{(m)}(l) \leftarrow \sum_{j \neq i} k^{(m)}(\mathbf{f}_i, \mathbf{f}_j) Q_j(l)$
  - Compatibility transform: $\hat{Q}_i(x_i) \leftarrow \sum_{l \in \mathcal{L}} \mu^{(m)}(x_i, l) \sum_m w^{(m)} \tilde{Q}_i^{(m)}(l)$
  - Local update: $Q_i(x_i) \leftarrow \exp\{-\psi_u(x_i) - \hat{Q}_i(x_i)\}$
  - Normalize: $Q_i(x_i)$

Mean field approximation

$$Q_i(x_i = l) = \frac{1}{Z_i} \exp\left\{-\psi_u(x_i) - \sum_{l' \in \mathcal{L}} \mu(l, l') \sum_{m=1}^{K} w^{(m)} \sum_{j \neq i} k^{(m)}(\mathbf{f}_i, \mathbf{f}_j) Q_j(l')\right\}$$

- Initialize $Q_i(x_i) \leftarrow \frac{1}{Z_i} \exp\{-\phi_u(x_i)\}$
- **while** not converged
  - Message passing: $\tilde{Q}_i^{(m)}(l) \leftarrow \sum_{j \neq i} k^{(m)}(\mathbf{f}_i, \mathbf{f}_j) Q_j(l)$
  - Compatibility transform: $\hat{Q}_i(x_i) \leftarrow \sum_{l \in \mathcal{L}} \mu^{(m)}(x_i, l) \sum_m w^{(m)} \tilde{Q}_i^{(m)}(l)$
  - Local update: $Q_i(x_i) \leftarrow \exp\{-\psi_u(x_i) - \hat{Q}_i(x_i)\}$
  - Normalize: $Q_i(x_i)$

Mean field approximation

$$Q_i(x_i = l) = \frac{1}{Z_i} \exp\left\{ -\psi_u(x_i) - \sum_{l' \in \mathcal{L}} \mu(l, l') \sum_{m=1}^{K} w^{(m)} \sum_{j \neq i} k^{(m)}(\mathbf{f}_i, \mathbf{f}_j) Q_j(l') \right\}$$

- Initialize $Q_i(x_i) \leftarrow \frac{1}{Z_i} \exp\{-\phi_u(x_i)\}$
- **while** not converged
  - Message passing: $\tilde{Q}_i^{(m)}(l) \leftarrow \sum_{j \neq i} k^{(m)}(\mathbf{f}_i, \mathbf{f}_j) Q_j(l)$
  - Compatibility transform: $\hat{Q}_i(x_i) \leftarrow \sum_{l \in \mathcal{L}} \mu^{(m)}(x_i, l) \sum_m w^{(m)} \tilde{Q}_i^{(m)}(l)$
  - Local update: $Q_i(x_i) \leftarrow \exp\{-\psi_u(x_i) - \hat{Q}_i(x_i)\}$
  - Normalize: $Q_i(x_i)$

Mean field approximation

Runtime analysis for N variables

$O(N)$ Initialize $Q_i(x_i) \leftarrow \frac{1}{Z_i}\exp\{-\phi_u(x_i)\}$

~ 10 **while** not converged

$O(N^2)$ Message passing: $\tilde{Q}_i^{(m)}(l) \leftarrow \sum_{j\neq i} k^{(m)}(\mathbf{f}_i, \mathbf{f}_j) Q_j(l)$

$O(N)$ Compatibility transform: $\hat{Q}_i(x_i) \leftarrow \sum_{l\in\mathcal{L}} \mu^{(m)}(x_i, l) \sum_m w^{(m)} \tilde{Q}_i^{(m)}(l)$

$O(N)$ Local update: $Q_i(x_i) \leftarrow \exp\{-\psi_u(x_i) - \hat{Q}_i(x_i)\}$

$O(N)$ Normalize: $Q_i(x_i)$

Mean field approximation

Runtime analysis for N variables

$O(N)$ Initialize $Q_i(x_i) \leftarrow \frac{1}{Z_i} \exp\{-\phi_u(x_i)\}$

~ 10 **while** not converged

$O(N^2)$ Message passing: $\tilde{Q}_i^{(m)}(l) \leftarrow \sum_{j \neq i} k^{(m)}(\mathbf{f}_i, \mathbf{f}_j) Q_j(l)$

$O(N)$ Compatibility transform: $\hat{Q}_i(x_i) \leftarrow \sum_{l \in \mathcal{L}} \mu^{(m)}(x_i, l) \sum_m w^{(m)} \tilde{Q}_i^{(m)}(l)$

$O(N)$ Local update: $Q_i(x_i) \leftarrow \exp\{-\psi_u(x_i) - \hat{Q}_i(x_i)\}$

$O(N)$ Normalize: $Q_i(x_i)$

Efficient message passing using high-dimensional filtering

- Update all $\tilde{Q}_i^{(m)}(l)$ simultaneously

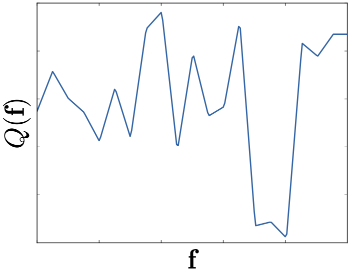
$$\tilde{Q}_i^{(m)}(l) = \sum_{j \neq i} k^{(m)}(\mathbf{f}_i, \mathbf{f}_j) Q_j(l)$$

- Efficiently computed using a cross-bilateral filter [Paris & Durand 09, Adams et al. 09, Adams et al. 10]
  - Permutohedral lattice [Adams et al. 10]

High-dimensional filtering [Paris & Durand 09]

$$\overline{Q}_i^{(m)}(l) = \sum_{j \in \mathcal{V}} \exp\left(\frac{1}{2}(\mathbf{f}_i^{(m)} - \mathbf{f}_j^{(m)})^2\right) Q_j(l)$$

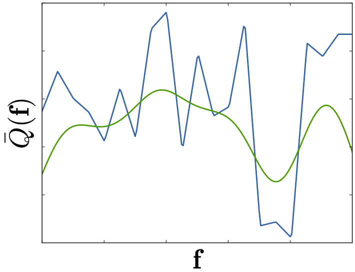
- High-dimensional input signal $Q_j(l)$
- Gaussian convolution $\overline{Q}_i^{(m)}(l) = \mathcal{G} \otimes Q_j(l)$
  - Band-limited, smooth function
- Can be reconstructed from a number of samples
  - Nyquist theorem



High-dimensional filtering [Paris & Durand 09]

$$\overline{Q}_i^{(m)}(l) = \sum_{j \in \mathcal{V}} \exp\left(\frac{1}{2}(\mathbf{f}_i^{(m)} - \mathbf{f}_j^{(m)})^2\right) Q_j(l)$$

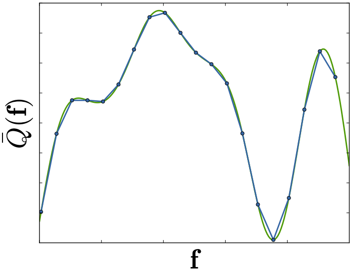
- High-dimensional input signal $Q_j(l)$
- Gaussian convolution $\overline{Q}_i^{(m)}(l) = \mathcal{G} \otimes Q_j(l)$
  - Band-limited, smooth function
- Can be reconstructed from a number of samples
  - Nyquist theorem



High-dimensional filtering [Paris & Durand 09]

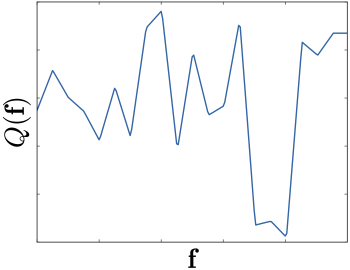
$$\overline{Q}_i^{(m)}(l) = \sum_{j \in \mathcal{V}} \exp\left(\frac{1}{2}(\mathbf{f}_i^{(m)} - \mathbf{f}_j^{(m)})^2\right) Q_j(l)$$

- High-dimensional input signal $Q_j(l)$
- Gaussian convolution $\overline{Q}_i^{(m)}(l) = \mathcal{G} \otimes Q_j(l)$
  - Band-limited, smooth function
- Can be reconstructed from a number of samples
  - Nyquist theorem



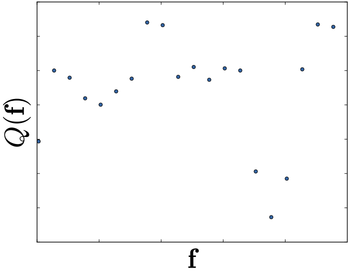
High-dimensional filtering [Paris & Durand 09]

- Downsample input signal $Q_j(I)$
- Blur the sampled signal
- Upsample to reconstruct the filtered signal $\overline{Q}_j(I)$



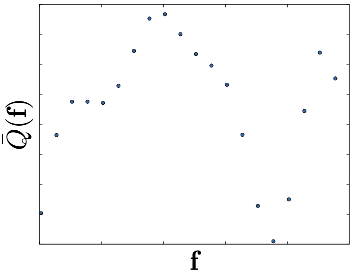
High-dimensional filtering [Paris & Durand 09]

- Downsample input signal $Q_j(l)$
- Blur the sampled signal
- Upsample to reconstruct the filtered signal $\overline{Q}_j(l)$



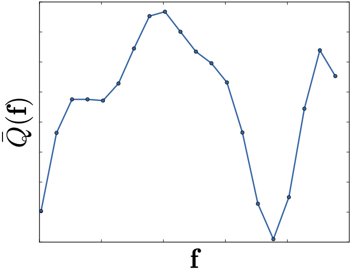
High-dimensional filtering [Paris & Durand 09]

- Downsample input signal $Q_j(l)$
- Blur the sampled signal
- Upsample to reconstruct the filtered signal $\overline{Q}_j(l)$



High-dimensional filtering [Paris & Durand 09]

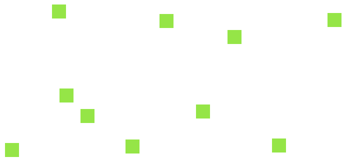
- Downsample input signal $Q_j(l)$
- Blur the sampled signal
- Upsample to reconstruct the filtered signal $\overline{Q}_j(l)$



High-dimensional Filtering [Adams et al. 10]

Permutohedral lattice

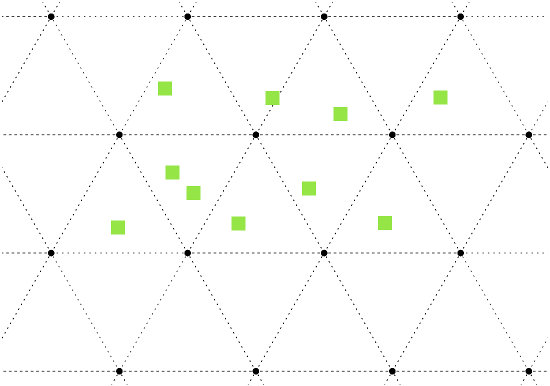
- High-dimensional signal $Q_j(l)$



High-dimensional Filtering [Adams et al. 10]

Permutohedral lattice

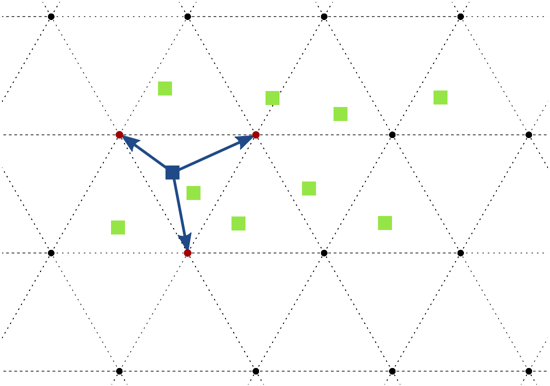
- Sample high-dimensional space



High-dimensional Filtering [Adams et al. 10]

Permutohedral lattice

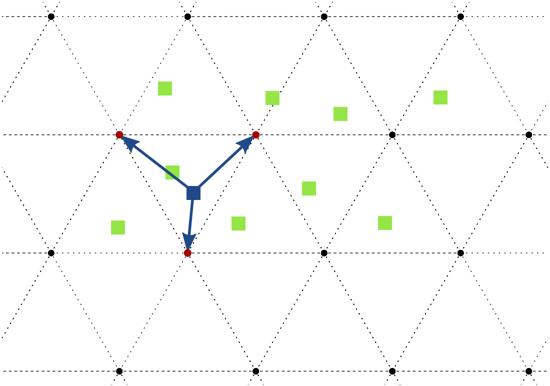
- Downsampling



High-dimensional Filtering [Adams et al. 10]

Permutohedral lattice

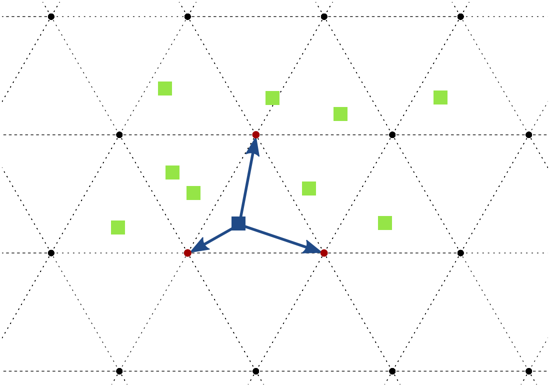
- Downsampling



High-dimensional Filtering [Adams et al. 10]

Permutohedral lattice

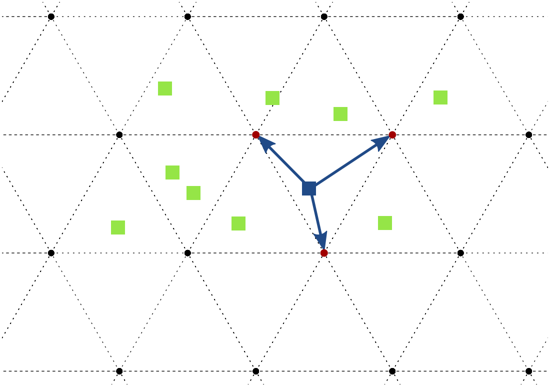
- Downsampling



High-dimensional Filtering [Adams et al. 10]

Permutohedral lattice

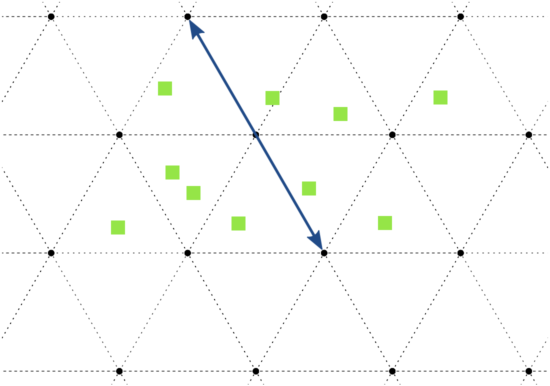
- Downsampling



High-dimensional Filtering [Adams et al. 10]

Permutohedral lattice

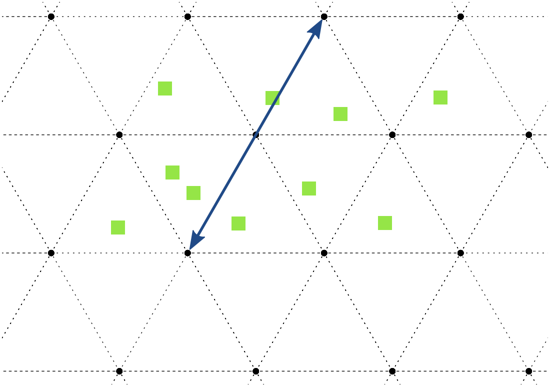
- Blurring



High-dimensional Filtering [Adams et al. 10]

Permutohedral lattice

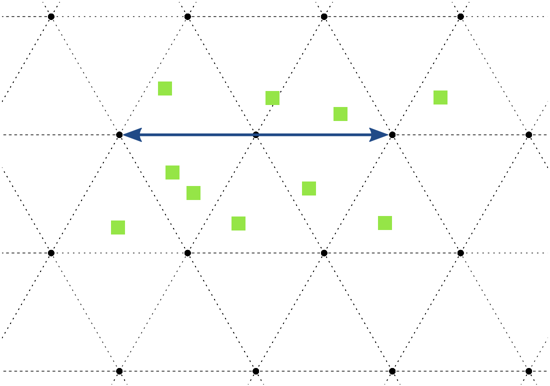
- Blurring



High-dimensional Filtering [Adams et al. 10]

Permutohedral lattice

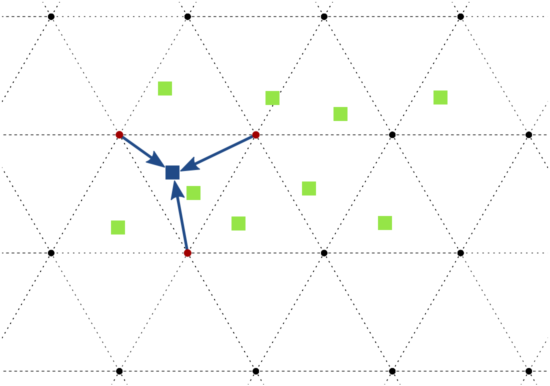
- Blurring



High-dimensional Filtering [Adams et al. 10]

Permutohedral lattice

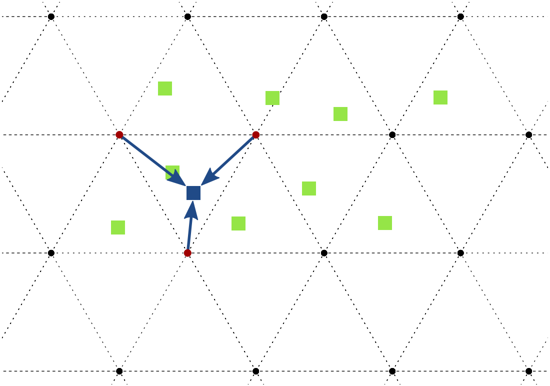
- Upsampling



High-dimensional Filtering [Adams et al. 10]

Permutohedral lattice

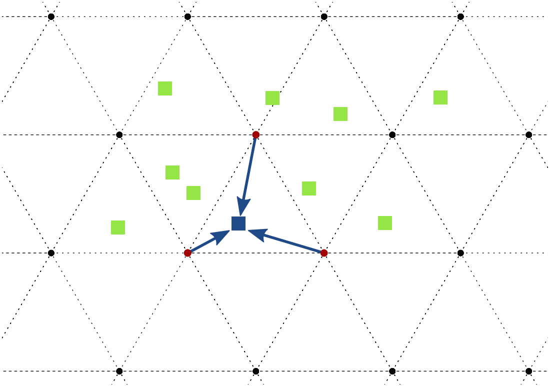
- Upsampling



High-dimensional Filtering [Adams et al. 10]

Permutohedral lattice

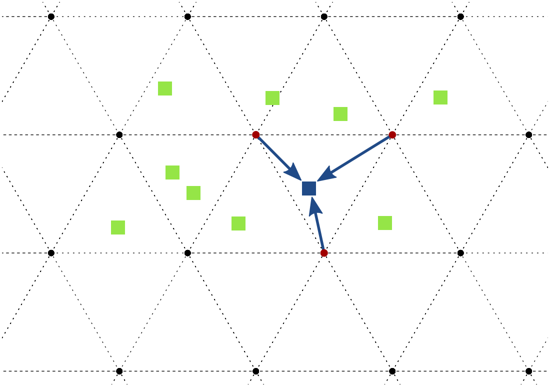
- Upsampling



High-dimensional Filtering [Adams et al. 10]

Permutohedral lattice

- Upsampling



Mean field approximation

Runtime analysis for N variables

$O(N)$ Initialize $Q_i(x_i) \leftarrow \frac{1}{Z_i} \exp\{-\phi_u(x_i)\}$

~ 10 **while** not converged

$O(N^2)$ Message passing: $\tilde{Q}_i^{(m)}(l) \leftarrow \sum_{j \neq i} k^{(m)}(\mathbf{f}_i, \mathbf{f}_j) Q_j(l)$

$O(N)$ Compatibility transform: $\hat{Q}_i(x_i) \leftarrow \sum_{l \in \mathcal{L}} \mu^{(m)}(x_i, l) \sum_m w^{(m)} \tilde{Q}_i^{(m)}(l)$

$O(N)$ Local update: $Q_i(x_i) \leftarrow \exp\{-\psi_u(x_i) - \hat{Q}_i(x_i)\}$

$O(N)$ Normalize: $Q_i(x_i)$

Mean field approximation

Runtime analysis for N variables

$O(N)$ Initialize $Q_i(x_i) \leftarrow \frac{1}{Z_i} \exp\{-\phi_u(x_i)\}$

~ 10 **while** not converged

$O(N)$ Message passing: High-dimensional filtering

$O(N)$ Compatibility transform: $\hat{Q}_i(x_i) \leftarrow \sum_{l \in \mathcal{L}} \mu^{(m)}(x_i, l) \sum_m w^{(m)} \tilde{Q}_i^{(m)}(l)$

$O(N)$ Local update: $Q_i(x_i) \leftarrow \exp\{-\psi_u(x_i) - \hat{Q}_i(x_i)\}$

$O(N)$ Normalize: $Q_i(x_i)$

Mean field approximation

Runtime analysis for N variables

$O(N)$ Initialize $Q_i(x_i) \leftarrow \frac{1}{Z_i} \exp\{-\phi_u(x_i)\}$

~ 10 **while** not converged

- $O(N)$ Message passing: High-dimensional filtering
- $O(N)$ Compatibility transform: $\hat{Q}_i(x_i) \leftarrow \sum_{l \in \mathcal{L}} \mu^{(m)}(x_i, l) \sum_m w^{(m)} \tilde{Q}_i^{(m)}(l)$
- $O(N)$ Local update: $Q_i(x_i) \leftarrow \exp\{-\psi_u(x_i) - \hat{Q}_i(x_i)\}$
- $O(N)$ Normalize: $Q_i(x_i)$

Learning

$$\psi_p(x_i, x_j) = \boxed{\mu(x_i, x_j)} \sum_{m=1}^{K} \boxed{w^{(m)}} \exp(-\frac{1}{2}(\mathbf{f}_i - \mathbf{f}_j)\Sigma^{(m)}(\mathbf{f}_i - \mathbf{f}_j))$$

- Efficient learning using high-dimensional filtering for μ and $w^{(m)}$
- Grid search for $\Sigma^{(m)}$
  - Non-Gaussian convolution

Learning

$$\psi_p(x_i, x_j) = \mu(x_i, x_j) \sum_{m=1}^{K} w^{(m)} \exp(-\frac{1}{2}(\mathbf{f}_i - \mathbf{f}_j) \boxed{\Sigma^{(m)}} (\mathbf{f}_i - \mathbf{f}_j))$$

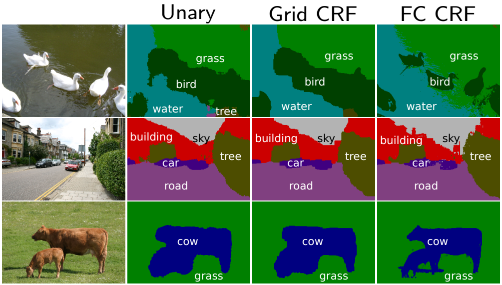
- Efficient learning using high-dimensional filtering for μ and $w^{(m)}$
- Grid search for $\Sigma^{(m)}$
  - Non-Gaussian convolution

Results: MSRC

MSRC dataset

- 591 images
- 21 classes

	Time	Global	Avg
Unary	-	84.0	76.6
Grid CRF	1s	84.6	77.2
FC CRF	**0.2s**	**86.0**	**78.3**

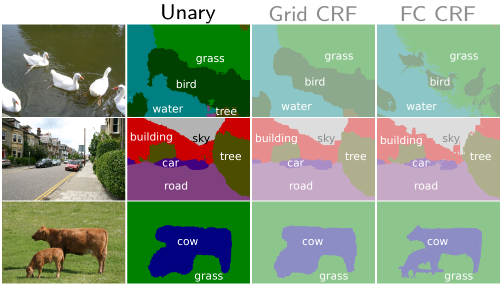


Results: MSRC

MSRC dataset

- 591 images
- 21 classes

	Time	Global	Avg
Unary	-	84.0	76.6
Grid CRF	1s	84.6	77.2
FC CRF	**0.2s**	**86.0**	**78.3**

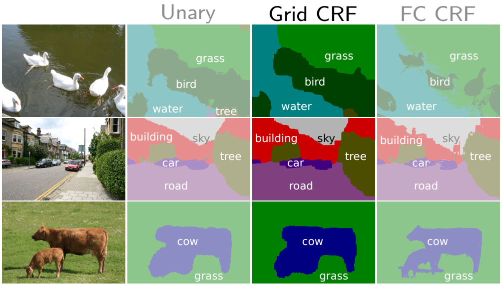


Results: MSRC

MSRC dataset

- 591 images
- 21 classes

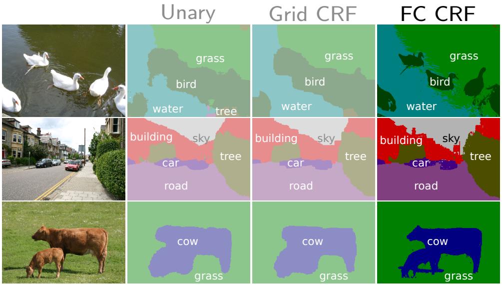
	Time	Global	Avg
Unary	-	84.0	76.6
Grid CRF	1s	84.6	77.2
FC CRF	**0.2s**	**86.0**	**78.3**



MSRC dataset

- 591 images
- 21 classes

	Time	Global	Avg
Unary	-	84.0	76.6
Grid CRF	1s	84.6	77.2
FC CRF	**0.2s**	**86.0**	**78.3**



Results: MSRC - Trimap

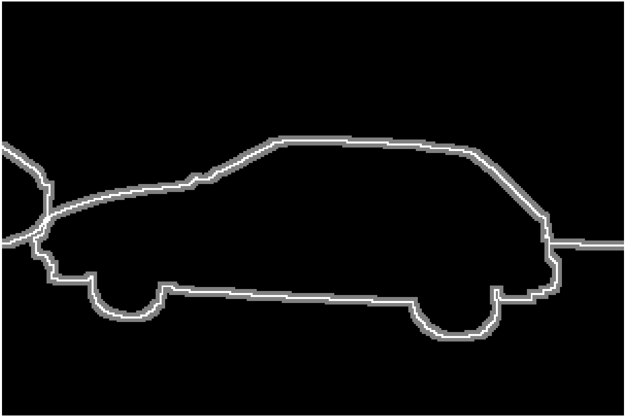
- 94 images
- hand annotated pixel accurately (30 min each)
- Trimap [Kohli et al. 2009]
  - Percentage of misclassified pixel around object boundaries

Results: MSRC - Trimap

Results: MSRC - Trimap

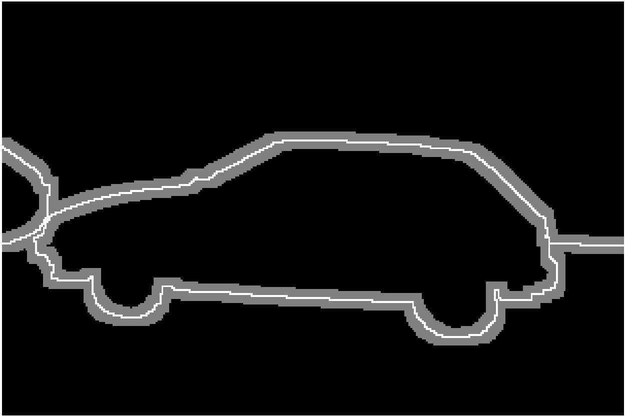
Results: MSRC - Trimap

4 pixel trimap

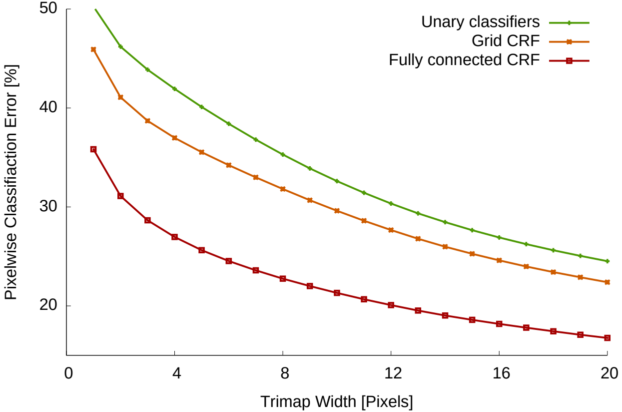


Results: MSRC - Trimap

8 pixel trimap



Results: MSRC - Trimap

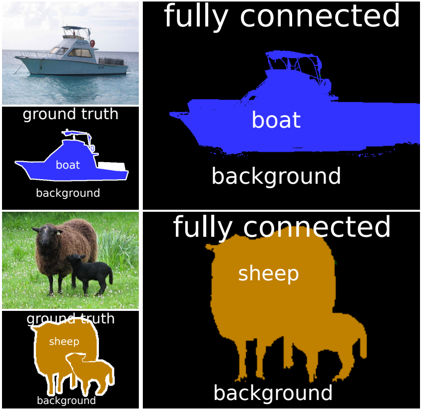


	Trimap width $= \infty$
Unary	16.8 ± 1.5
Grid CRF	15.2 ± 1.5
FC CRF	$\mathbf{11.8} \pm \mathbf{0.7}$

PASCAL VOC 2010 dataset

- 1928 images
- 20 classes + background

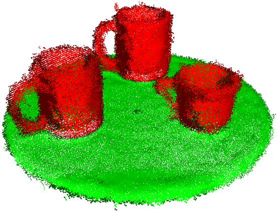
	Time	Acc
Unary	-	27.6
Grid CRF	2.5s	28.3
FC Potts	0.5s	29.1
FC label comp	**0.5s**	**30.2**



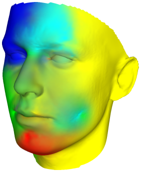
Other domains

Fully connected CRFs in other domains (ongoing work)

- Point clouds (XYZ + normal + color)

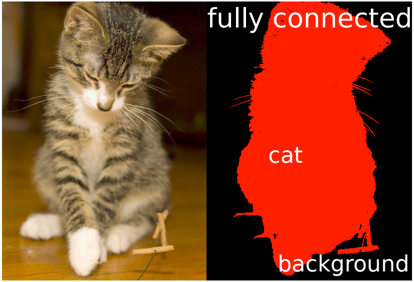


- Meshes (XYZ + normal)



Summary

- Fully connected CRF model
  - Pairwise terms: linear combination of Gaussians
- Efficient inference
  - Linear in number of variables
  - Independent of number of pairwise terms



Future work

- Better inference than mean field
  - Serial filtering
- Continuous variables
  - Depth reconstruction
  - Optical flow
- Non-Euclidean spaces
  - Geodesic or diffusion distance
  - Meshes
  - General graphs
- Beyond simple label compatibility
  - Feature dependent label compatibility

Future work

- Better inference than mean field
  - Serial filtering
- Continuous variables
  - Depth reconstruction
  - Optical flow
- Non-Euclidean spaces
  - Geodesic or diffusion distance
  - Meshes
  - General graphs
- Beyond simple label compatibility
  - Feature dependent label compatibility

Future work

- Better inference than mean field
  - Serial filtering
- Continuous variables
  - Depth reconstruction
  - Optical flow
- Non-Euclidean spaces
  - Geodesic or diffusion distance
  - Meshes
  - General graphs
- Beyond simple label compatibility
  - Feature dependent label compatibility

Future work

- Better inference than mean field
  - Serial filtering
- Continuous variables
  - Depth reconstruction
  - Optical flow
- Non-Euclidean spaces
  - Geodesic or diffusion distance
  - Meshes
  - General graphs
- Beyond simple label compatibility
  - Feature dependent label compatibility

Questions

Webpage and Code:

`http://graphics.stanford.edu/projects/densecrf/`

Poster W14