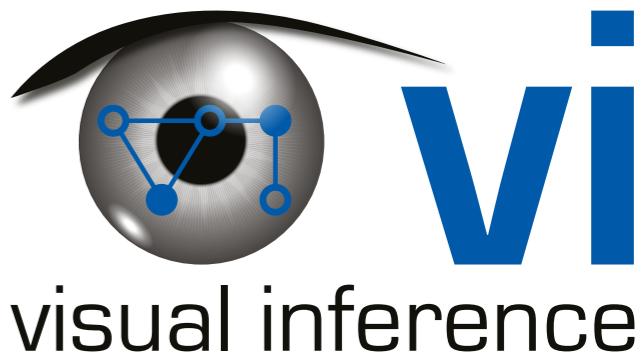
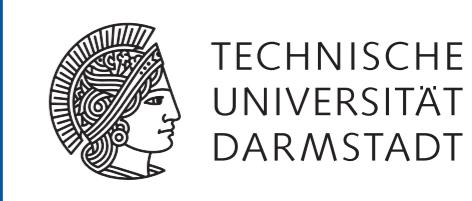


Outlook: Semantic Segmentation

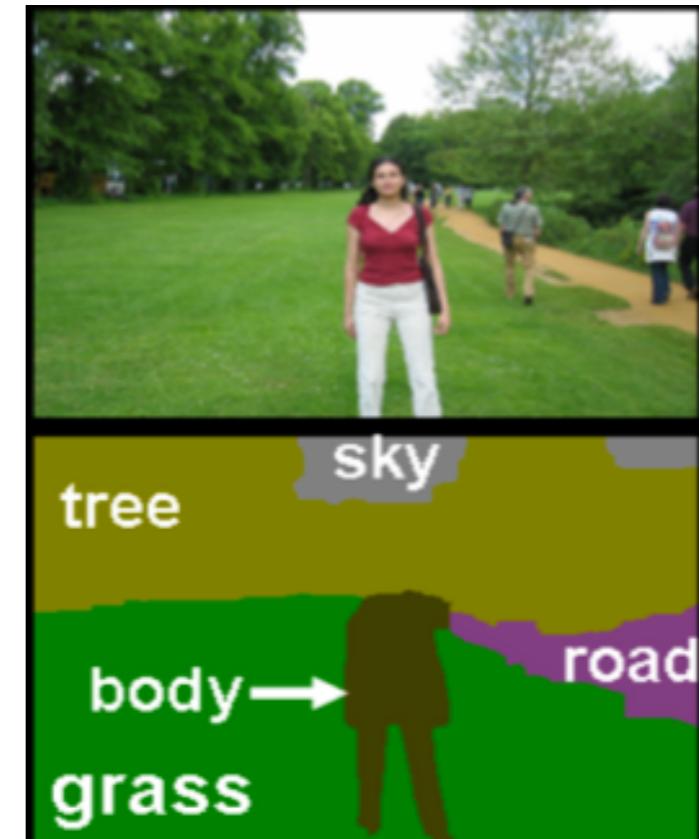
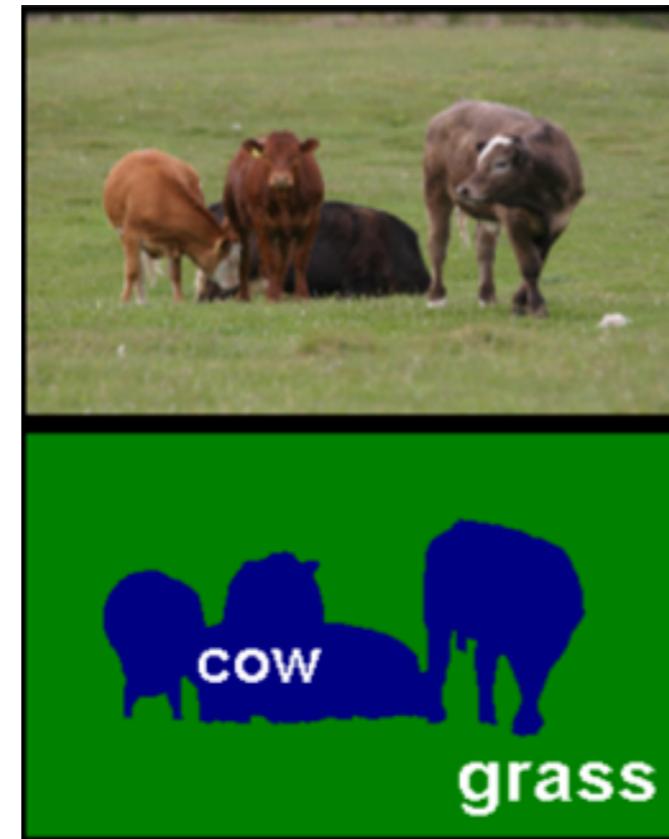
18.06.2014



Semantic Segmentation

- ◆ Segmentation can be made more well-defined by coupling it with high-level reasoning
- ◆ Semantic segmentation = Segmentation + Recognition

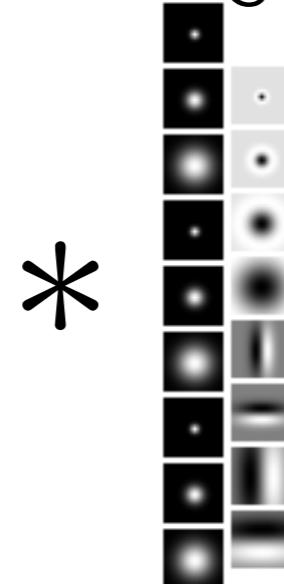
- ◆ segment the image
- ◆ determine the category at every pixel



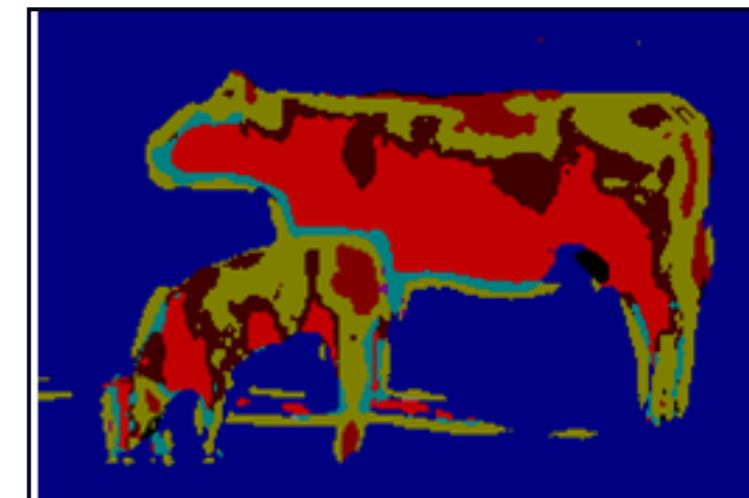
[Jamie Shotton]

First attempt: Pixel classification

- ◆ If we have a good classifier, this may work
- ◆ **TextonBoost** [Shotton et al., 2006]:
 - ◆ Texton features for image description



clustering →



- ◆ Dense visual words
- ◆ + shape filters (skipped here)

First attempt: Pixel classification

- ◆ If we have a good classifier, this may work
- ◆ **TextronBoost** [Shotton et al., 2006]:
 - ◆ Textron features for image description
 - ◆ Joint boosting for multi-class classification



| Object classes | Building | Grass | Tree | Cow | Sheep | Sky | Aeroplane | Water | Face | Car |
|----------------|----------|-------|------|------|-------|------|-----------|-------|------|------|
| Bike | Flower | Sign | Bird | Book | Chair | Road | Cat | Dog | Body | Boat |

[Shotton et al., 2006]

Training data (MSRC)

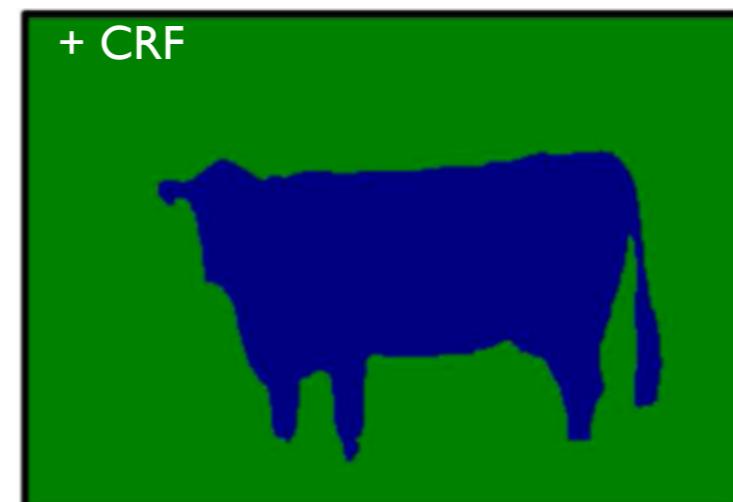
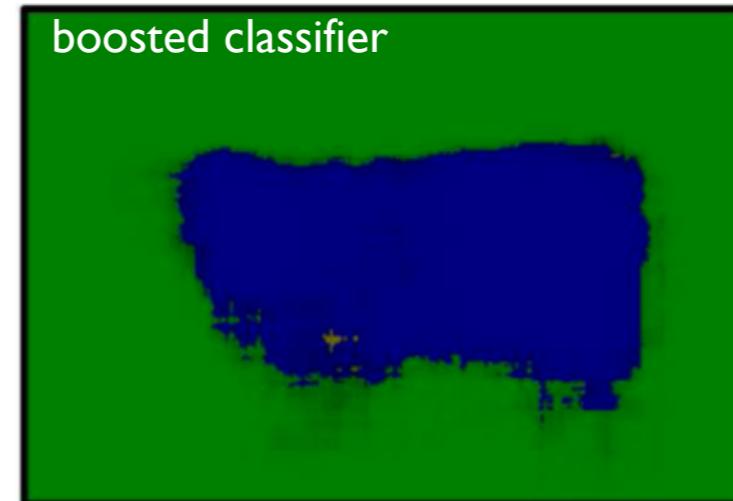


| <i>Object classes</i> | Building | Grass | Tree | Cow | Sheep | Sky | Aeroplane | Water | Face | Car |
|-----------------------|----------|-------|------|------|-------|------|-----------|-------|------|------|
| Bike | Flower | Sign | Bird | Book | Chair | Road | Cat | Dog | Body | Boat |

[Shotton et al., 2006]

First attempt: Pixel classification

- ◆ Results so far:
 - ◆ Good at recognizing objects
 - ◆ Pixel-wise segmentation accuracy rather low

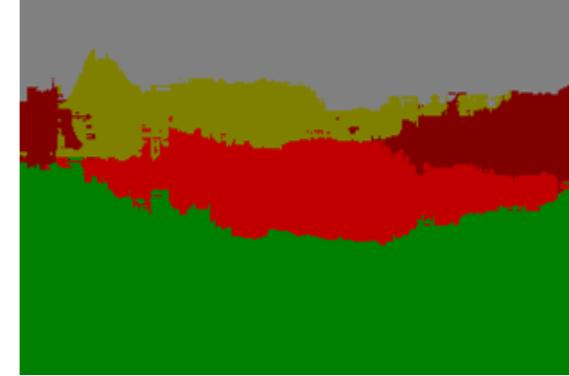


[Shotton et
al., 2006]

TextonBoost -- CRF

- ◆ CRF model:

structure texture
potentials



$$\begin{aligned} \log P(\mathbf{c}|\mathbf{x}, \boldsymbol{\theta}) = & \sum_i \psi_i(c_i, \mathbf{x}; \boldsymbol{\theta}_\psi) + \pi(c_i, \mathbf{x}_i; \boldsymbol{\theta}_\pi) + \lambda(c_i, i; \boldsymbol{\theta}_\lambda) \\ & + \sum_{(i,j) \in \mathcal{E}} \phi(c_i, c_j, \mathbf{g}_{ij}(\mathbf{x}); \boldsymbol{\theta}_\phi) - \log Z(\boldsymbol{\theta}, \mathbf{x}) \end{aligned}$$

[Shotton et al., 2006]

TextonBoost -- CRF

- ◆ CRF model:

$$\begin{aligned}\log P(\mathbf{c}|\mathbf{x}, \boldsymbol{\theta}) = & \sum_i \psi_i(c_i, \mathbf{x}; \boldsymbol{\theta}_\psi) + \pi(c_i, \mathbf{x}_i; \boldsymbol{\theta}_\pi) + \lambda(c_i, i; \boldsymbol{\theta}_\lambda) \\ & + \sum_{(i,j) \in \mathcal{E}} \phi(c_i, c_j, \mathbf{g}_{ij}(\mathbf{x}); \boldsymbol{\theta}_\phi) - \log Z(\boldsymbol{\theta}, \mathbf{x})\end{aligned}$$

color
potential



[Shotton et al., 2006]

TextonBoost -- CRF

- ◆ CRF model:

location
potential



$$\begin{aligned} \log P(\mathbf{c}|\mathbf{x}, \boldsymbol{\theta}) = & \sum_i \psi_i(c_i, \mathbf{x}; \boldsymbol{\theta}_\psi) + \pi(c_i, \mathbf{x}_i; \boldsymbol{\theta}_\pi) + \lambda(c_i, i; \boldsymbol{\theta}_\lambda) \\ & + \sum_{(i,j) \in \mathcal{E}} \phi(c_i, c_j, \mathbf{g}_{ij}(\mathbf{x}); \boldsymbol{\theta}_\phi) - \log Z(\boldsymbol{\theta}, \mathbf{x}) \end{aligned}$$

[Shotton et al., 2006]

TextonBoost -- CRF

- ◆ CRF model:
 - ◆ inference using graph cuts + alpha expansion

$$\log P(\mathbf{c}|\mathbf{x}, \boldsymbol{\theta}) = \sum_i \psi_i(c_i, \mathbf{x}; \boldsymbol{\theta}_\psi) + \pi(c_i, \mathbf{x}_i; \boldsymbol{\theta}_\pi) + \lambda(c_i, i; \boldsymbol{\theta}_\lambda) \\ + \sum_{(i,j) \in \mathcal{E}} \phi(c_i, c_j, \mathbf{g}_{ij}(\mathbf{x}); \boldsymbol{\theta}_\phi) - \log Z(\boldsymbol{\theta}, \mathbf{x})$$



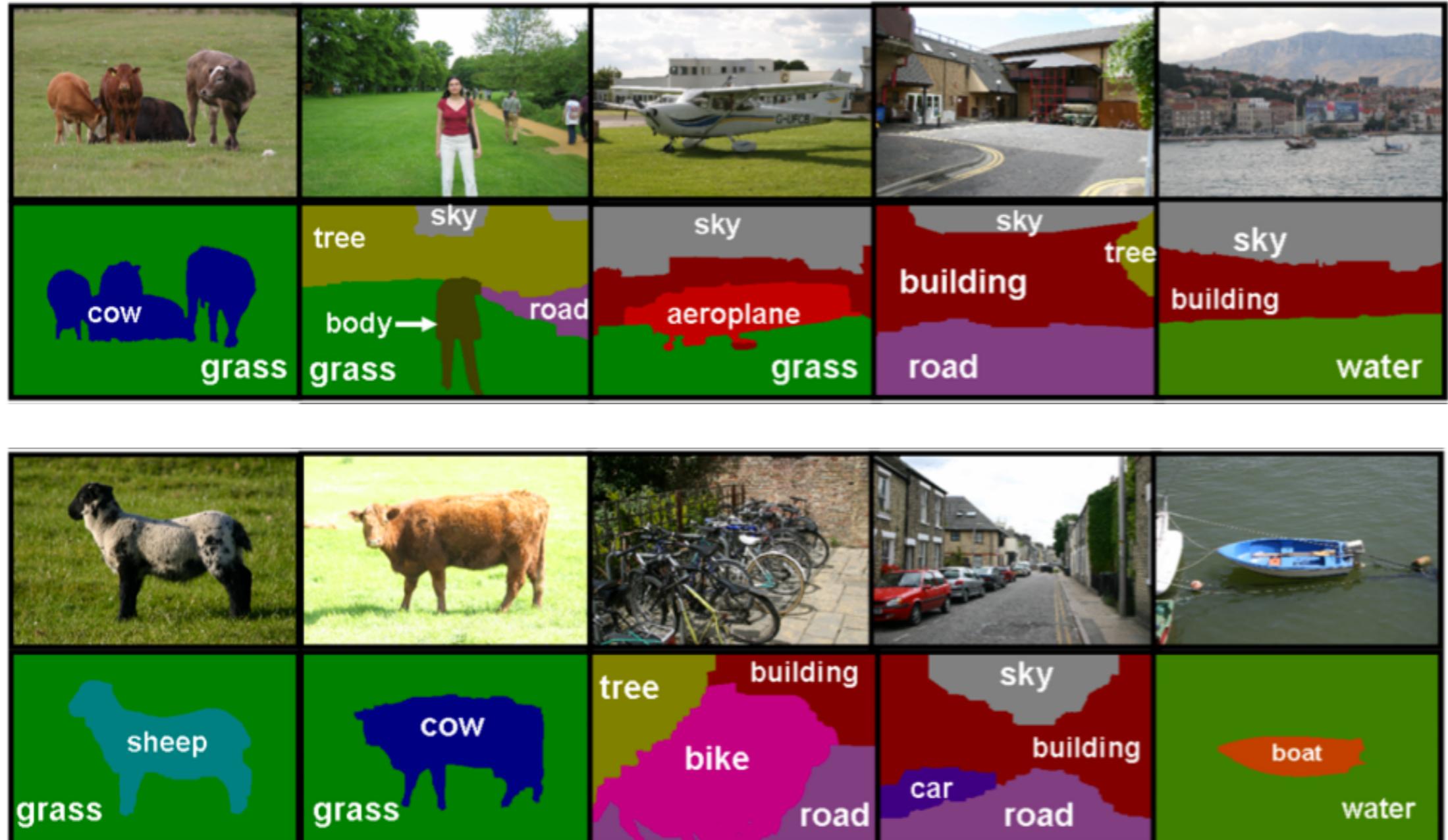
edge map

contrast-
sensitive
edge potential
(Potts)

partition
function

[Shotton et al., 2006]

Example results



| Object classes | Building | Grass | Tree | Cow | Sheep | Sky | Aeroplane | Water | Face | Car |
|----------------|----------|-------|------|------|-------|------|-----------|-------|------|------|
| Bike | Flower | Sign | Bird | Book | Chair | Road | Cat | Dog | Body | Boat |

[Shotton et al., 2006]