Class-Specific, Top-Down Segmentation

Eran Borenstein & Shimon Ullman

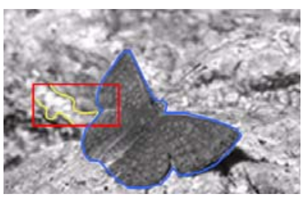
Presented by Chia-Chih Chen

Overview (1)

- The major goal of image segmentation is to identify structures in the image that are likely to correspond to scene objects
- Classic image-based segmentation methods use continuity of grey-level, texture, and bounding contours
- Where is the object boundary? 

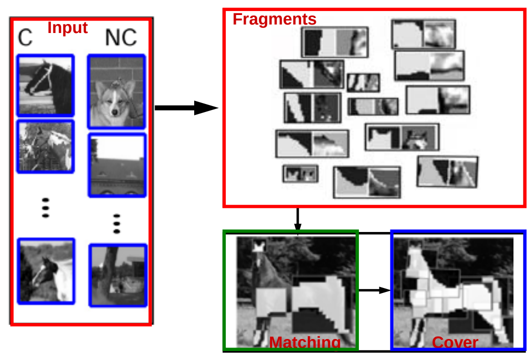
Reference slides: www.frc.ri.cmu.edu/users/josephad/TopDownBottomUpSeg.ppt

Overview (2)



- The class can help resolve ambiguities!
- Segmentation is guided by a stored representation of the shape of objects within a general class

Method Overview



Method Outline

- **Fragment Extraction**
  - **Figure Ground Label**
  - **Reliability Value**
- Fragment Matching
  - Individual Correspondences
  - Consistency
  - Reliability
- Segmentation
  - Optimal Cover

Fragment Extraction

- Calculate the strength of responses *Si* of *Fi* in C and NC
- Decide *θi* according to Neyman-Pearson decision theory $p(S_i > \theta_i | NC) \leq \alpha$
- Select top K fragments according to $p(S_i > \theta_i | C)$ (hit rate), K decide size of fragment set
- Two more factors are added to each fragment:

Figure-ground label

Reliability

$$r_i = \frac{p(S_i > \theta_i | C)}{p(S_i > \theta_i | NC)} = \frac{\text{detection rate}}{\alpha}$$

Method Outline

- Fragment Extraction
  - Figure Ground Label
  - Reliability Value
- **Fragment Matching**
  - **Individual Correspondences**
  - **Consistency**
  - **Reliability**
- Segmentation
  - Optimal Cover

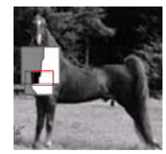
Fragment Matching

- Individual Correspondence

$$s_i(p,I) = w \cdot \underbrace{\boxed{N_{cor}(p,I)|_{\substack{\text{Object}\\ \text{Pixels}}}}}_{\text{Region Correlation}} + (1-w)\cdot \underbrace{\boxed{E_{dge}(p,I)}}_{\text{Edge Detector}}$$

- Consistency

$$c_{ij} = \frac{\#\ \text{Consistent Overlapping Pixels}}{\max(\text{Total Overlap}, \mu_{ij})}$$



- Reliability

$$r_i = \frac{p(S_i > \theta_i|C)}{p(S_i > \theta_i|NC)} = \frac{\text{detection rate}}{\alpha}$$

Method Outline

- Fragment Extraction
  - Figure Ground Label
  - Reliability Value
- Fragment Matching
  - Individual Correspondences
  - Consistency
  - Reliability
- **Segmentation**
  - **Optimal Cover**

Segmentation – Optimal Cover (1)

- The best cover should maximize individual match quality, consistency and reliability
- Thus the cover score is written:

$$cs = \underbrace{\sum_i r_i \cdot s_i \cdot f_i}_{\substack{\text{Individual match}\\ \text{and reliability}}} + \underbrace{\frac{1}{\lambda}\sum_{i,j} \beta_{ij} \cdot f_i \cdot f_j}_{\text{Consistency}}$$

Rewards for match quality and reliability

Penalizes for inconsistent overlapping fragments

$$\beta_{ij} = (c_{ij} - \beta) \cdot (r_i s_i + r_j s_j)$$

Interaction is zero for non-overlapping pairs

Constant that determines the magnitude of the penalty for insufficient consistency

Segmentation – Optimal Cover (2)

- Initialize with a sub-window that has the maximal concentration of reliable fragments
  - Similarity of all the reliable fragments is examined at 5 scales at all possible locations
- Iterative Algorithm:
  - Select a small number (M=15) of good candidate fragments
  - Add to cover a subset of the M fragments that maximally improve the score
  - Remove existing fragments inconsistent with new cover (fragments with cumulative negative score)
  - Guaranteed to converge to a local maximum

Results

$$r = \frac{|S \cap F|}{|S \cup F|}$$

- Paper algorithm: 0.71
- Normalized-cuts: 0.31
- Random segmentation: 0.23

Conclusions

- Demonstrate the feasibility of using class-based criteria to generate segmentation corresponds to visual objects
- The cover algorithm resembles solving jigsaw puzzle
- (# of reliable fragments)(# of pixels in each scale)(# of scales)
- Future work: 1) using pyramid of image segments
  2) boundaries can be refined by image-based methods