Part III: Affinity Functions for Image Segmentation

Charless Fowlkes

joint work with David Martin and Jitendra Malik
at University of California at Berkeley
Q: What measurements should we use for constructing the affinities?

Similarity Cues

a) distance
b) region cues (patch similarity)
c) boundary cues (intervening contour)

What image measurements allow us to gauge the probability that pixels $i$ and $j$ belong to the same segment?
Boundary Processing

Region Processing

Texture

Brightness

Color

Original Image

\( W_{ij} \)

Proximity

\( \chi^2 \)
Learning Pairwise Affinities

$S_{ij}$ – indicator variable as to whether pixels $i$ and $j$ were marked as belonging to the same group by human subjects.

$W_{ij}$ – our estimate of the likelihood that pixel $i$ and $j$ belong to the same group conditioned on the image measurements.

- Use the ground truth given by human segmentations to calibrate cues.
- Learn a statistically optimal cue combination strategy in supervised learning framework
- Ecological Statistics: Measure the relative power of different cues for natural scenes
Part III: Affinity Functions for Image Segmentation
Individual Gradient Features

- 1976 CIE L*a*b* colorspace
- Brightness Gradient $BG(x,y,r,\theta)$
  - Difference of $L^*$ distributions
- Color Gradient $CG(x,y,r,\theta)$
  - Difference of $a^*b^*$ distributions
- Texture Gradient $TG(x,y,r,\theta)$
  - Difference of distributions of V1-like filter responses

\[
\chi^2(g,h) = \frac{1}{2} \sum_i \frac{(g_i - h_i)^2}{g_i + h_i}
\]
Texture Feature

- Texture Gradient $\text{TG}(x,y,r,\theta)$
  - $\chi^2$ difference of texton histograms
  - Textons are vector-quantized filter outputs
What about my favorite edge detector?

• Canny Detector
  – Canny 1986
  – MATLAB implementation
  – With and without hysteresis

• Second Moment Matrix
  – Nitzberg/Mumford/Shiota 1993
  – cf. Förstner and Harris corner detectors
  – Used by Konishi et al. 1999 in learning framework
  – Logistic model trained on full eigenspectrum
$P_b$ Images I

<table>
<thead>
<tr>
<th>Image</th>
<th>Canny</th>
<th>2MM</th>
<th>Us</th>
<th>Human</th>
</tr>
</thead>
<tbody>
<tr>
<td>![Image 1]</td>
<td>![Canny 1]</td>
<td>![2MM 1]</td>
<td>![Us 1]</td>
<td>![Human 1]</td>
</tr>
<tr>
<td>![Image 2]</td>
<td>![Canny 2]</td>
<td>![2MM 2]</td>
<td>![Us 2]</td>
<td>![Human 2]</td>
</tr>
</tbody>
</table>
### Pb Images III

<table>
<thead>
<tr>
<th>Image</th>
<th>Canny</th>
<th>2MM</th>
<th>Us</th>
<th>Human</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="canny1.png" alt="Canny" /></td>
<td><img src="2mm1.png" alt="2MM" /></td>
<td><img src="us1.png" alt="Us" /></td>
<td><img src="human1.png" alt="Human" /></td>
</tr>
<tr>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="canny2.png" alt="Canny" /></td>
<td><img src="2mm2.png" alt="2MM" /></td>
<td><img src="us2.png" alt="Us" /></td>
<td><img src="human2.png" alt="Human" /></td>
</tr>
</tbody>
</table>
Part III: Affinity Functions for Image Segmentation

Image + Contour Cues → Estimated Boundaries
Human Segmentations → Groundtruth Boundaries

Evaluate
Two Decades of Local Boundary Detection
How good are humans locally?

- Algorithm: $r = 9$, Humans: $r = \{5, 9, 18\}$
- Fixation(2s) -> Patch(200ms) -> Mask(1s)
Man versus Machine:
Intervening Contour

...turning a boundary map into Wij

1 - maximum $P_b$ along the line connecting i and j
Part III: Affinity Functions for Image Segmentation
Individual Patch Features

- Use same histogram based representation
- Brightness Similarity
  - Difference of L* distributions
- Color Similarity
  - Difference of a*b* distributions
- Texture Similarity
  - Difference of distributions of V1-like filter responses

\[
\chi^2(g, h) = \frac{1}{2} \sum_i \frac{(g_i - h_i)^2}{g_i + h_i}
\]
Detail: Clipping Patch Features

- Clip patch support using Pb in order to try and avoid “polluting” histograms.
Part III: Affinity Functions for Image Segmentation

- Image
- Region Cues
- Contour Cues
- Estimated Affinity (W)

- Human Segmentations
- Groundtruth Affinity (S)
- Evaluate

- Segment
Two Evaluation Measures

1. **Precision-Recall** of same-segment pairs
   
   - Precision is \( P(S_{ij} = 1 \mid W_{ij} > t) \)
   
   - Recall is \( P(W_{ij} > t \mid S_{ij} = 1) \)

2. **Mutual Information** between \( W \) and \( S \)

\[ p(s, w) \log \left( \frac{p(s)p(w)}{p(s,w)} \right) \]
Individual Features

Patches

Gradients
Clipping patch support improves $W_{ij}$ estimate
Cue Combination Models

- Classification Trees
  - Top-down splits to maximize entropy, error bounded
- Density Estimation
  - Adaptive bins using k-means
- Logistic Regression, 3 variants
  - Linear and quadratic terms
  - Confidence-rated generalization of AdaBoost (Schapire&Singer)
- Hierarchical Mixtures of Experts (Jordan&Jacobs)
  - Up to 8 experts, initialized top-down, fit with EM
- Support Vector Machines (libsvm, Chang&Lin)
  - Gaussian kernel, v-parameterization

- Logistic with quadratic terms is sufficient (performs as well as any classifier we tried)
Findings:

1. **Common Wisdom**: Use patches only / Use edges only  
   **Finding**: Use both in pairwise affinity framework.

2. **Common Wisdom**: Must use patches for texture  
   **Finding**: Not true. Possible to detect texture boundaries.

3. **Common Wisdom**: Color is a powerful grouping cue  
   **Finding**: True, but texture is better.

4. **Common Wisdom**: Brightness patches are a poor cue  
   **Finding**: True (shadows and shading).

5. **Common Wisdom**: Proximity is a (Gestalt) grouping cue  
   **Finding**: Proximity is a result, not a cause of grouping.
Affinity Model vs. Human Segmentation

![Precision-Recall Curve](image)

- **Patch**: F=0.60, MI=0.16
- **IC**: F=0.61, MI=0.14
- **Patch+IC**: F=0.65, MI=0.19
- **Humans**: F=0.77, MI=0.25
- **Iso**: F=0.77
Part III: Affinity Functions for Image Segmentation

- Image
- Region Cues
- Contour Cues
- Estimated Affinity (W)
- Eigenvectors
- Cluster
- Groundtruth Affinity (S)
- Human Segmentations

Evaluate
Extract Pb

Compute Eigenvectors

Gradient of eigenvectors
Evaluating the power of “globalization”