From Contours to Regions: An Empirical Evaluation

presented by: Yingjie Tang
Outline:

- Motivation
- Contours VS Regions
- Contour Detection
- Oriented Watershed Transform
- Ultrametric Contour map
- Evaluation:
  - Benchmark & Results
- Conclusions & Contributions
Motivation of Segmentations:
1. Provide natural domains for computing features used in recognition
2. Visual tasks can benefit from the complexity reduction
Contours VS Segmentation:

Contours: Not closed
Segments: Closed and can provide partition

Hierarchical Segmentation from Contours

Contour Detector  OWT-UCM  Hierarchical Segmentation
gPb Detector: Oriented Gradient of Histograms

\[ \chi^2(g, h) = \frac{1}{2} \sum_i \frac{(g(i) - h(i))^2}{g(i) + h(i)} \]
gPb Detector: Multiscale/ Multi-channel Pb
gPb Detector:
weighted combination of multi scale local cues

\[ m_{Pb}(x, y, \theta) = \sum_s \sum_i \alpha_{i,s} G_{i,s}(x, y, \theta) \]

s: indexes scales
i: indexes feature channels (brightness, color a, color b, texture)

\( G_{i,s}(x, y, \theta) \) measures the histogram difference in channel i between two halves of disc centered at (x,y) at angle \( \theta \).
gPb Detector:
Globalization

\[ W_{ij} = \exp \left( -\max_{p \in \overline{ij}} \{ mPb(p) \} / \rho \right) \]

\( \overline{ij} \) is the line segment connecting \( i \) and \( j \) and \( \rho \) is a constant. We set \( r = 5 \) pixels and \( \rho = 0.1 \).

\[ sPb(x, y, \theta) = \sum_{k=1}^{n} \frac{1}{\sqrt{\lambda_k}} \cdot \nabla_{\theta} v_k(x, y) \]

the weighting by \( 1/\sqrt{\lambda_k} \) is motivated by the physical interpretation of the generalized eigenvalue problem as a mass-spring system
gPb Detector:  
Globalization

\[
gPb(x, y, \theta) = \sum_s \sum_i \beta_{i,s} G_{i,\sigma(i,s)}(x, y, \theta) + \gamma \cdot sPb(x, y, \theta)
\]

weights $\beta_{i,s}$ and $\gamma$ are learned by gradient ascent on the F-measure using the BSDS training images.
gPb Detector: Globalization

Benefit of Globalization
gPb Detector:
Globalization

Benefit of Globalization
Segmentation:

Contour Detector $E(x, y, \theta)$ → Oriented Watershed Transformation → Ultrametric Contour Map
Hierarchical Segmentation

- From contours to segmentation
  - Watershed Transform
    - Concept

—Hsin-Min Cheng
Hierarchical Segmentation

- From contours to segmentation
  - Watershed Transform
    - Example
Hierarchical Segmentation

- From contours to segmentation
  - Watershed Transform

\[ E(x, y) = \max_\theta E(x, y, \theta) \]

—Hsin-Min Cheng
Hierarchical Segmentation

- From contours to segmentation
  - Oriented Watershed Transform

\[ o(x,y) \approx 0 \quad \text{for each pixel on the arc} \]
\[ E(x,y) = E(x,y,0) \]
Segmentation: Watershed Transform.

Disadvantages: Since the contour detector produces a spatially extended response around strong boundaries, simply weighting each arc by mean value of $E(x,y)$ can introduce artifacts.
Segmentation: Watershed Transform  

\[ E(x, y) = \max_{\theta} E(x, y, \theta) \]

Watershed arcs computed from \( E(x,y) \).

\[ \text{Watershed arcs with an approximating straight line segment subdivision overlaid.} \]

\[ \text{Oriented boundary strength } E(x, y, \theta) \text{ for four orientations } \theta. \]

\[ \text{Oriented boundary strength } E(x, y, \theta) \text{ for four orientations } \theta. \]
Segmentation:
Contour Subdivision

Scale Invariant
Segmentation:
Ultrametric Contour Map

Base level: Over-segmentation

Moving between levels offers a continuous trade-off between these extremes

Upper level: Under-segmentation
Segmentation:
Ultrametric Contour Map

1. Construct graph \( G = (P_0, K_0, W(K_0)) \) given by OWT
2. Iteratively merge regions by removing min weight boundary.
3. Produces region tree where:
   - Root is entire image
   - Leaves are \( P_0 \)
   - \( \text{Height}(R) \) is boundary threshold at which \( R \) first appears
   - \( \text{Distance}(R_1, R_2) = \min \{ \text{Height}(R) : R_1, R_2 \subseteq R \} \)
Segmentation:

Ultrametric Contour Map

Merging Algorithm Description:

1) Select minimum weight contour:
   \[ C^* = \arg\min_{C \in \mathcal{K}_0} W(C). \]
2) Let \( R_1, R_2 \in \mathcal{P}_0 \) be the regions separated by \( C^* \).
3) Set \( R = R_1 \cup R_2 \), and update:
   \[ \mathcal{P}_0 \leftarrow \mathcal{P}_0 \setminus \{R_1, R_2\} \cup \{R\} \quad \text{and} \quad \mathcal{K}_0 \leftarrow \mathcal{K}_0 \setminus \{C^*\}. \]
4) Stop if \( \mathcal{K}_0 \) is empty.
   Otherwise, update weights \( W(\mathcal{K}_0) \) and repeat.
Evaluation:
On Berkeley Segmentation Dataset (BSDS)

- Precision-Recall on Boundaries
- Variation of Information
- Rand Index
- Segmentation Covering

Evaluation:
Precision-Recall on Boundaries

\[
\begin{array}{ll}
P' & P \\
A & B \\
C & D \\
\end{array}
\]

\[
R\text{(recall)} = \frac{A}{A+C} \\
P\text{(precision)} = \frac{A}{A+B} \\
F\text{-measure} = \frac{2PR}{P+R}
\]

- Optimal Dataset Scale (ODS)
- Optimal Image Scale (OIS)
- Average Precision (AP)
Evaluation:
Precision-Recall on Boundaries

Evaluating boundaries on the Berkeley dataset

The proposed approach outperformed the state of art
Evaluation:
Precision-Recall on Boundaries

Evaluating boundaries on the Berkeley dataset

<table>
<thead>
<tr>
<th>Method</th>
<th>ODS</th>
<th>OIS</th>
<th>AP</th>
</tr>
</thead>
<tbody>
<tr>
<td>human</td>
<td>0.79</td>
<td>0.79</td>
<td>—</td>
</tr>
<tr>
<td>gPb-owt-ucm</td>
<td>0.71</td>
<td>0.74</td>
<td>0.77</td>
</tr>
<tr>
<td>Mean Shift</td>
<td>0.63</td>
<td>0.66</td>
<td>0.62</td>
</tr>
<tr>
<td>NCuts</td>
<td>0.62</td>
<td>0.66</td>
<td>0.59</td>
</tr>
<tr>
<td>Canny-owt-ucm</td>
<td>0.58</td>
<td>0.63</td>
<td>0.59</td>
</tr>
<tr>
<td>Felz-Hutt</td>
<td>0.58</td>
<td>0.62</td>
<td>0.54</td>
</tr>
<tr>
<td>gPb</td>
<td>0.70</td>
<td>0.72</td>
<td>0.75</td>
</tr>
<tr>
<td>Canny</td>
<td>0.58</td>
<td>0.62</td>
<td>0.60</td>
</tr>
</tbody>
</table>

Boundary benchmarks on the BSDS
Evaluation:
Variation of Information & Rand Index & Segmentation Covering

\[ V I(C, C') = H(C) + H(C') - 2I(C, C') \]

Measures the distance between two segmentations in terms of their average conditional entropy

\[ PRI(S, \{G_k\}) = \frac{1}{T} \sum_{i<j} [c_{ij}p_{ij} + (1 - c_{ij})(1 - p_{ij})] \]

Rand Index between test and ground-truth segmentations S and G is given by the sum of the number of pairs of pixels that have the same label in S and G and those that have different labels in both segmentations, divided by the total number of pairs of pixels.

\[ \mathcal{O}(R, R') = \frac{|R \cap R'|}{|R \cup R'|} \]

\[ C(S' \rightarrow S) = \frac{1}{N} \sum_{R \in S} |R| \cdot \max_{R' \in S'} \mathcal{O}(R, R') \]
Evaluation:
Variation of Information & Rand Index & Segmentation Covering

<table>
<thead>
<tr>
<th>Method</th>
<th>ODS</th>
<th>OIS</th>
<th>Best</th>
<th>PRI</th>
<th>VI</th>
</tr>
</thead>
<tbody>
<tr>
<td>human</td>
<td>0.73</td>
<td>0.73</td>
<td></td>
<td>0.87</td>
<td>1.16</td>
</tr>
<tr>
<td>gPb-owt-ucm</td>
<td>0.58</td>
<td>0.64</td>
<td>0.74</td>
<td>0.81</td>
<td>1.68</td>
</tr>
<tr>
<td>Mean Shift</td>
<td>0.54</td>
<td>0.58</td>
<td>0.64</td>
<td>0.78</td>
<td>1.83</td>
</tr>
<tr>
<td>Felz-Hutt</td>
<td>0.51</td>
<td>0.58</td>
<td>0.68</td>
<td>0.77</td>
<td>2.15</td>
</tr>
<tr>
<td>Canny-owt-ucm</td>
<td>0.48</td>
<td>0.56</td>
<td>0.67</td>
<td>0.77</td>
<td>2.11</td>
</tr>
<tr>
<td>NCuts</td>
<td>0.44</td>
<td>0.53</td>
<td>0.66</td>
<td>0.75</td>
<td>2.18</td>
</tr>
</tbody>
</table>

Region benchmarks on the BSDS

Evaluating regions on the BSDS
Conclusions & Contributions:

• Algorithm gPb-owt-ucm, offers the best performance on every dataset and for every benchmark criteria tested
• This algorithm is straightforward, fast, has no parameters to tune, and supports interactive user refinement
• The generic grouping machinery has found use in optical flow and object recognition applications.
Q & A
Constrained Parametric Min-Cuts for Automatic Object Segmentation
Outline
• Constrained Parametric Min-Cuts (CPMC)
• The Grid Geometry
• Fast Rejection
• Ranking Segments
• Experiments
Constrained Parametric Min-Cuts (CPMC)

Min-Cuts: a minimum cut of a graph is a cut (a partition of the vertices of a graph into two disjoint subsets that are joined by at least one edge) whose cut set has the smallest number of edges (unweighted case) or smallest sum of weights possible. —Wikipedia

Maximum flow: maximum flow problems involve finding a feasible flow through a single-source, single-sink flow network that is maximum. —Wikipedia
Constrained Parametric Min-Cuts (CPMC)

Object segmentation framework
Constrained Parametric Min-Cuts (CPMC)

Energy function in pixels:

\[ E^\lambda(X) = \sum_{u \in \mathcal{V}} D_\lambda(x_u) + \sum_{(u,v) \in \mathcal{E}} V_{uv}(x_u, x_v) \]

with \( \lambda \in \mathbb{R} \), and with the unary potential function being:

\[
D_\lambda(x_u) = \begin{cases} 
0 & \text{if } x_u = 1, u \notin \mathcal{V}_b \\
\infty & \text{if } x_u = 1, u \in \mathcal{V}_b \\
\infty & \text{if } x_u = 0, u \in \mathcal{V}_f \\
f(x_u) + \lambda & \text{if } x_u = 0, u \notin \mathcal{V}_f 
\end{cases}
\]

\[
V_{uv}(x_u, x_v) = \begin{cases} 
0 & \text{if } x_u = x_v \\
g(u, v) & \text{if } x_u \neq x_v 
\end{cases}
\]
Fast Rejection

Large set of initial segmentations (~5500)

High Energy

~2000 segments with the lowest energy

Low Energy

Cluster segments based on spatial overlap (at least 0.95)

Lowest energy member of each cluster (~154 in PASCAL VOC)

Credit: SasiKanth Bendapudi Yogeshwar Nagaraj
Ranking Segments:

3 sets of features:
• Graph partition properties (8 features)
• Region properties (18 features)
• Gestalt properties (8 features)
Experiments:

Three Datasets:
- Weizmann’s Segmentation Evaluation Database
  (with only one prominent foreground object in each and test F-measure criteria)
- MSRC dataset (nearly 11 objects in each image and is used to evaluate the quality of the pool of segments generated)
- VOC 2009 dataset (most complicated and used segmentation covering as an accuracy measure)
Experiments:
Segment Pool Quality

<table>
<thead>
<tr>
<th>Weizmann</th>
<th>F-measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPMC</td>
<td>0.93 ± 0.009</td>
</tr>
<tr>
<td>Bagon et al.</td>
<td>0.87 ± 0.010</td>
</tr>
<tr>
<td>Alpert et al.</td>
<td>0.86 ± 0.012</td>
</tr>
</tbody>
</table>

Average of best segment F-measure scores over the entire dataset.

<table>
<thead>
<tr>
<th>MSRC</th>
<th>Covering</th>
<th>N Segments</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPMC</td>
<td>0.85 ± 0.1</td>
<td>57</td>
</tr>
<tr>
<td>gPb-owt-ucm</td>
<td>0.78 ± 0.15</td>
<td>670</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>VOC2009</th>
<th>Covering</th>
<th>N Segments</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPMC</td>
<td>0.78 ± 0.18</td>
<td>154</td>
</tr>
<tr>
<td>gPb-owt-ucm</td>
<td>0.61 ± 0.20</td>
<td>1286</td>
</tr>
</tbody>
</table>

Average best image covering scores on MSRC and VOC2009 train+validation datasets
Experiments:
Segment Pool Quality

Quality of the segments in VOC2009 joint train and validation sets for the segmentation problem.
Evaluating the Ranking Object Hypothesis

Average best segment F-measure as the number of retained segments from the ranking is varied.
Evaluating the Ranking Object Hypothesis

Complementing the basic descriptor set with appearance and shape features improves the ranking slightly, but with a more expressive regressor, random forests, the basic set is still superior.
Conclusion & Contributions

- Presented an algorithm that casts the automatic image segmentation problem as one of finding a set of plausible figure-ground object hypotheses.

- It is shown that the proposed framework is able to generate small sets of segments that represent the objects in an image more accurately than existing state of the art segmentation methods.
Q&A