Crowdsourcing Image Segmentation using SVG
Edward Kim and Xiaolei Huang

Department of Computer Science and Engineering,
Lehigh University, Bethlehem, PA, USA

October 17, 2011
VERY BIG Problem

- Understanding images for a machine
VERY BIG Problem

- Understanding images for a machine
Motivate the Problem

- Imagine you want to train a computer to recognize object
  - teach a computer what a bottle looks like, or what a tree looks like
Motivate the Problem

• Imagine you want to train a computer to recognize object
  • teach a computer what a bottle looks like, or what a tree looks like
• Step further, search for image or movie based upon content
  • images of me and my family next to Mickey Mouse
BIG Problem

- Point & Speak
BIG Problem

- Point & Speak
- Build a large dataset of annotated images
Obtain a large database of annotate images

- How do we get this dataset?
Multiple ways to get the job done

Bad - “Me”

- 30 sec - 1 minute to annotate an object in an image
- 1,000 image database could take up to 17 hours
- 12 million images [deng2009], which could take a user 22.83 years
Multiple ways to get the job done

• Bad - “Me”
  • 30 sec - 1 minute to annotate an object in an image
  • 1,000 image database could take up to 17 hours
  • 12 million images [deng2009], which could take a user 22.83 years

• Good - “You”
Multiple ways to get the job done

- **Bad - “Me”**
  - 30 sec - 1 minute to annotate an object in an image
  - 1,000 image database could take up to 17 hours
  - 12 million images [deng2009], which could take a user 22.83 years

- **Good - “You”**
- **Better - “Computer & You”**
- **Best - “Computer”**
Outline

- Motivate the Problem
- Introduction
- Related Work
- Methodology & Framework
  - Segmenting objects in images
  - Crowdsourcing via Amazon Mechanical Turk
- Experiments and Results
- Conclusion and Past & Future Work
Related Work

- [Yao 2007] - Lotus Hill Dataset
- [Russell 2008] - LabelMe
  - Online annotation system
  - Also utilizes Mechanical Turk (just “you”)
  - Non-standard storage and visualization (XML)

- Our system
  - uses standardization, visualization of XML using SVG
  - Computer & You
Methodology & Framework

- Segmentation Methods
  - Manual segmentation
Manual Segmentation

- `<image href>` to display a raster image
- `onclick` events, draw `<circle>`s
- clicking on first circle closes the annotation and appends a `<polygon>` with the clicked points

```xml
<?xml version="1.0" ?>
<svg baseProfile="full" xmlns="http://www.w3.org/2000/svg" xmlns:xlink="http://www.w3.org/1999/xlink" width="200" height="300" viewBox="0 0 200 300" version="1.1">
  <title>Strawberry example</title>
  <desc>Manual annotation. Strawberry example</desc>
  <g xmlns="http://www.w3.org/2000/svg" id="area">
    <image xmlns:xlink="http://www.w3.org/1999/xlink" id="rawimage" xlink:href="strawberry.jpg" x="0" y="0" width="200" height="300"/>
    <g id="manual" transform="translate(0,0)">
      <polygon points="123.57 132.76 143.90 158.110 154.148 138.200 115.229 95.230 74.208 47.163 31.113 52.91 46.74 43.64 49.51 55.57 69.64 86.62 83.54 106.50" stroke="rgb(0,255,0)" stroke-width="2" fill="none"/>
    </g>
  </g>
</svg>
```
Semi-Automatic Segmentation

- Use interactive computer segmentation methods to provide more accurate object segmentations.
Semi-Automatic Image Segmentation

- We use active contours [Kass 1988]
- Send “rough” polygon to the server via AJAX.
- Active contours evolves the “rough” polygon towards the edges of the image.
- Returns the SVG result

```xml
<svg baseProfile="full" xmlns="http://www.w3.org/2000/svg" xmlns:xlink="http://www.w3.org/1999/xlink"
   width="200" height="300" viewBox="0 0 200 300" version="1.1">
  <title>Strawberry example</title>
  <desc>Semi-automatic annotation strawberry example</desc>
  <g xmlns="http://www.w3.org/2000/svg" id="area">
    <image xmlns:xlink="http://www.w3.org/1999/xlink" id="rawimage" xlink:href="strawberry.jpg" x="0" y="0"
           width="200" height="300"/>
    <g id="manual" transform="translate(0,0)">
      <polygon fill-opacity="0.4" fill="green" stroke="rgb(0,255,0)" stroke-width="1" points="41,129 41,126 41,127 41,126 41,125 41,124 42,123 42,122 43,120 43,119 43,118 44,117 44,116 45,115 45,114 45,113 46,112 46,111 46,110 46,109 47,108 47,107 48,106 48,105 48,104 49,103 49,102 50,101 51,100 52,99 52,98 53,97 53,96 53,95 53,94 53,93 53,92 53,91 52,90 52,89 52,88 52,87 52,86 52,85 53,84 52,83 52,82 51,81 51,80 50,79 50,78 49,77 48,76 48,75 47,74 46,73 46,72...cut several lines of points...
      </polygon>
    </g>
  </g>
</svg>
```
Amazon Mechanical Turk

- create Human Intelligence Tasks (HITs)
- Given instructions on how to annotate and what to annotate
- provide a manual segmentation and present both manual and semi-automatic results
Amazon Mechanical Turk

- Ask which result they prefer
- Store result in MySQL database
- After completing the HIT, the worker is paid $0.05.
Results

- 242 images from the ETHZ dataset which consists of 5 classes
Results

- Publish 5 HITs per image for a total of 1,210 HITs
- Within 3 hours and 11 minutes, all 1,210 HITs were completed by 71 unique workers
- Average time per HIT, 1 minute 39 seconds
- Worker cost = $0.05 per hit = $60.50
- Amazon Fees = $6.05
- Total cost = $66.05
Results, “Computer & You”

- In 93 out of 242 images at least one worker said that the semi-automatic segmentation result was better than their manual annotation.
Visualization of results
Visualization of results
Results - Comments

- 12 comments related to “nice”, “nice work”, or “good job”
- “I am very lucky to work this job”
- “your computer work not better then me”
- “a cup has a topological gap, hard to handle”
Conclusion

- We proposed a system for facilitating image annotation
- Crowdsourced the image segmentation task online using SVG
Additional Applications
Additional applications
Thank you!

- Questions?