CS 395T Visual Recognition
Learning to Detect a Salient Object

Chao Jia
2012/10/26
Goal of this paper

- Detect the (unique) salient object in an image

Applications
  - Image resizing
  - Object recognition
Overview

• General steps
• MSRA dataset
• Different salient feature maps
• CRF inference & learning
• Results
• Multiple salient object detection
• Conclusions
MSRA dataset

- 20000 images with user labeled bounding boxes of salient objects

MSRA dataset

- Ground truth: labeled by 3-9 users
General steps

saliency feature maps

binary saliency map

search for bounding box
Salient feature maps

- Multi-scale contrast
  - 6-level Gaussian pyramid
Salient feature maps

• Multi-scale contrast
Salient feature maps

- Center-surround histogram

works better for homogeneous background
Salient feature maps

• Color spatial distribution
  – 6-component GMM modeling of color

  skin of people (partially colluded by clothes

  image borders

  special object in background

  car

  image borders
CRF inference & learning

- Binary labeling
- Weighting parameters learned from user labeling (bounding boxes, not binary salient maps)
- Brute force search of bounding box
  - 7 aspect ratios (1:2 to 2:1)
  - 11 window sizes (2% to 60% of the image area)
Results

• Success for easy examples
Results

• Success for harder examples
Results

• Success for harder examples
Results

• Three cues perfectly complement each other
Results

- Failure: very complicated scene
Results

- Failure: similar to background
Results

• Failure: spatially apart components with similar color

Not always a problem unless the image is very colorful
(More components in GMM needed)
Results

• How much better than just putting a fixed size window in the center?

\( \frac{1}{4} \text{ size; same aspect ratio} \)

\[
\begin{align*}
\text{Precision} & = \frac{\sum_x g_x a_x}{\sum_x a_x} \\
\text{Recall} & = \frac{\sum_x g_x a_x}{\sum_x g_x}
\end{align*}
\]
Multiple salient object detection

• PASCAL VOC 2007 dataset
  – Multiple salient objects
  – More complicated scenes
Multiple salient object detection

generate the binary mask

score each possible window

Non-maximum suppression: remove a window if it overlaps with another window with a higher score

score = # saliency pixels × (# saliency pixels / window area)
Multiple salient object detection
Multiple salient object detection
Multiple salient object detection

• Different overlap threshold in NMS
Multiple salient object detection

- Precision-Recall Curve on PASCAL VOC 2007 – 1000 images; 3004 objects

This paper

Algorithm that especially designed to detect multiple salient objects

ICCV 2011 “Salient Object Detection by Composition”
Multiple salient object detection

- Still acceptable if there’s only one salient object
Multiple salient object detection

- A few successful results on multi-object detection

The binary masks are correct and separable.
Multiple salient object detection

- Failure examples

- Too far away

- Binary mask completely wrong
- Hard to separate
- Is the ground truth really salient?
Conclusions

• Works well for single salient object detection.
  – Three cues complement each other
  – Bounding box annotation is very robust to errors in binary saliency labeling.
  – MSRA dataset is relatively simple: Central fixation bias naïve baseline works well too.

• The algorithm is not suitable for multi-object detection.
  – Only one connected components most of the time
  – Binary salient mask fails
    • Spatially wide spread of objects will make the color spatial distribution cue less accurate.