The Role of Image Understanding in Contour Detection

Larry Zitnick (Microsoft Research) and Devi Parikh (TTIC)

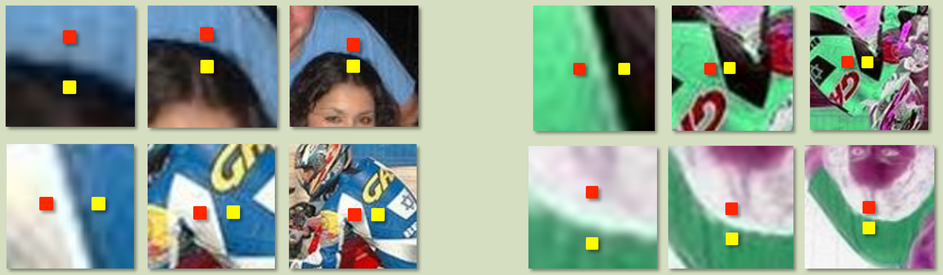
Overview

We examine the relative importance of low-, mid- and high-level cues to gain a better understanding of their role in detecting object contours in an image. To accomplish this task, we conduct numerous human studies and compare their performance to several popular machine algorithms for segmentation and contour detection.

Dataset

- 185 images in the SUN dataset (Choi et al., CVPR 2010).
- Extracted patches from 240 locations
- Half on object boundaries as per SUN ground truth annotations and half off
- A third of locations with low, medium and high gradients each
- 96,000 responses to each of 4 questions on Mechanical Turk

Do the red and yellow squares lie on the same object, or different objects?

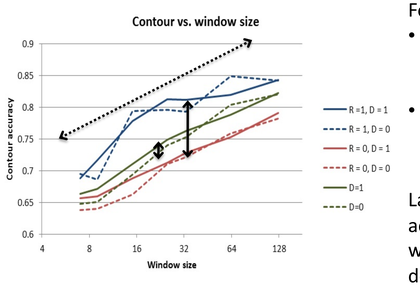


Human studies

We asked the subjects a series of four questions:

- Do the red and yellow squares lie on the same object, or different objects? The possible answers were: "Same object" or "Different object".
- Is the object under the red square in front of or behind the object under the yellow square? The possible answers were: "Red in front of Yellow", "Yellow in front of Red" and "Neither".
- Which object does the red square belong to? Subjects were to provide a one word freeform answer.
- Which object does the yellow square belong to? Subjects were to provide a one word freeform answer.

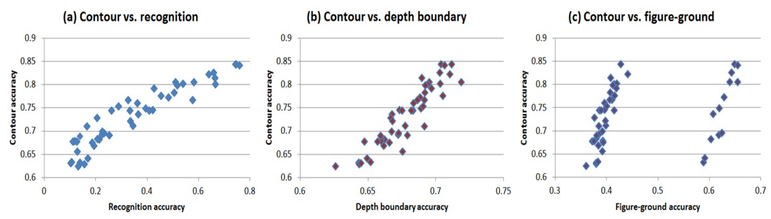
Mid- vs. High-Level Cues



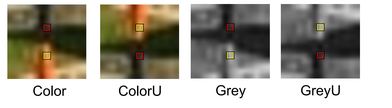
For patches of the same size:

- If at least one object is recognized, the contour is more likely to be correctly classified.
- However, the correct labeling of a depth boundary has a minimal effect on contour classification.

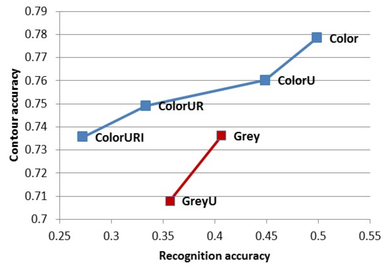
Large improvement in contour detection accuracy as patch size increases even when conditioned on recognition and depth boundary detection



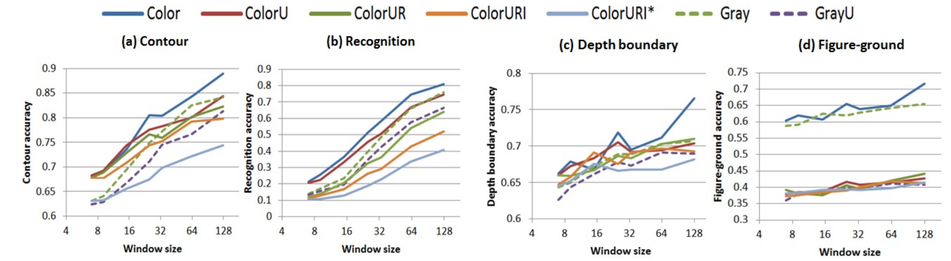
Contour detection vs. recognition



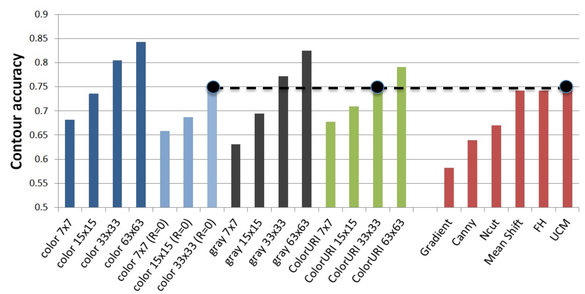
U = Upside down
R = Rotated colors (RGB = GBR)
I = Inverted colors (R = 255 – R)
* = Only one color inverted

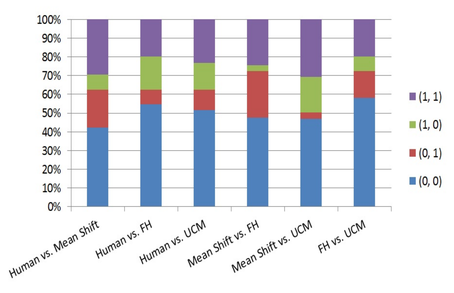


- Contour accuracy **improves** as the patches become easier to recognize.
- Mid-level information from non-local gradients, textures and edges is constant.
- Provides strong evidence of a causal relationship between recognition and contour detection



Human vs. machine





Gradient: naïve local gradient method
Canny: Classical canny edge detector
Ncut: Normalized cuts segmentation algorithm, Shi & Malik, 2000
Mean Shift: Mean-shift clustering algorithm for image segmentation, Comaniciu & Meer 2002
FH: Graph-based segmentation approach by Felzenszwalb & Huttenlocher, 2004
UCM: Ultrametric Contour Maps segmentation approach, Arbelaez et al., 2011

Conclusions

- Our findings suggest that the current state-of-the-art contour detection algorithms perform as well as humans using low-level cues.
- We find evidence that the recognition of objects, but not occlusion information, leads to improved human performance.
- When at least one object is recognized by humans, their contour detection performance increases over current machine algorithms.
- Mid-level cues appear to offer a larger performance boost than high-level cues such as recognition.