Local background enclosure for RGB-D salient object detection

David Feng, Nick Barnes, Shaoqi You, Chris McCarthy

1. Smart Vision Systems, Data61@CSIRO; 2. Research School of Engineering, ANU; 3. Department of Computer Science and Software Engineering, Swinburne University

Motivation

Prosthetic vision: low resolution and low dynamic range display, uniform sampling very poor

Depth saliency provides a way to identify interesting objects in the field of view, for tasks such as navigation with prosthetic vision.

1. Camera image capture
2. Image to signal
3. Signal to implant
4. Implant stimulation
5. Stimulation

Saliency System

Prosthetic vision devices

Depth

GRGD salient object detection - a challenging scene:

- Low colour contrast
- Low depth contrast

Previous work - depth contrast features:


Problem: saliency(P1) × saliency(P2)

Solution - look for scene structure "enclosed" by background:

P1: Local background enclosure saliency

Saliency system output for challenging images:

References


Measuring Enclosure:

The local background set of a patch P consists of all patches within a radius r with mean depth above a threshold from P:

\[ B(P) = \{ P' \in N_r | \bar{d}(P') > \bar{d}(P) + t \} \]

where \( \bar{d}(P) \) is the mean depth within a patch P.

The threshold T controls exclusivity of local background set:

\[ T = \sigma \bar{d}(P) \]

Here \( \sigma \) denotes neighbourhood depth standard deviation.

Density Functions

Angular Fill

Angular proportion of neighbourhood interacting with local background set:

\[ f(P) = \frac{1}{2\pi} \int \frac{1}{r^2} \pi \delta(r, \theta) \text{d}r \text{d}\theta \]

Angular Gap

Size of the largest angular region not intersecting the background set:

\[ g(P) = \frac{1}{2\pi} \int \frac{1}{r^2} \pi \delta(r, \theta) \text{d}r \text{d}\theta \]

Distribution Functions

Integrate the density functions f and g over the background threshold T in (0,\( \bar{d} \)) to improve robustness.

\[ P(P) = \int f(P') \cdot g(P') \text{d}P' \]

LBE value is given by:

\[ P(P) = \frac{1}{2\pi} \int \frac{1}{r^2} \pi \delta(r, \theta) \text{d}r \text{d}\theta \]

Results

LBE vs Depth Contrast

High contrast background regions are a common source of errors for depth contrast methods. LBE provides a better indication of foreground structure.

Variations in background structure particularly in indoor scenes result in false positives for contrast methods. LBE can reduce the error in these cases.

Angled planar surfaces such as the ground plane tend to be marked as salient by contrast methods, but have low background enclosure values.

Objects with high depth variance can lead to false negatives in contrast based methods. Enclosure values are more uniform across these objects.