Motivation

- Saliency, segmentation, and object detection are related concepts in human vision.
- We present a joint framework for object proposal generation that is based on saliency and segmentation.

Saliency: from iNVT to VOCUS2

Saliency computation can be simple and fast. The seminal saliency system iNVT by Itti et al. [1] contained all necessary concepts of saliency computation. We show how to adapt Itti’s approach to obtain state-of-the-art performance on current benchmarks [2]:

- Fusion: Equal treatment of 3 channels I, RG, BY
- Different opponent color space
- Main change:
  - **Twin Pyramids** for Difference-of-Gaussians (one center, one surround pyramid per channel)
  - Enables flexible center-surround ratio
- Scale-space with several scales per layer
- Optional: Location prior (e.g. center bias)
- Optional: Segment-based saliency based on object proposal detection method

Results

- Average on datasets: Kitchen D (superpixels) (superpixels) (superpixels)
- C++ Code at: http://www.iai.uni-bonn.de/~frintrop/vocus2.html

Object Proposal Generation [3,5]

Extensions to Video [3] and RGB-D Data [4]