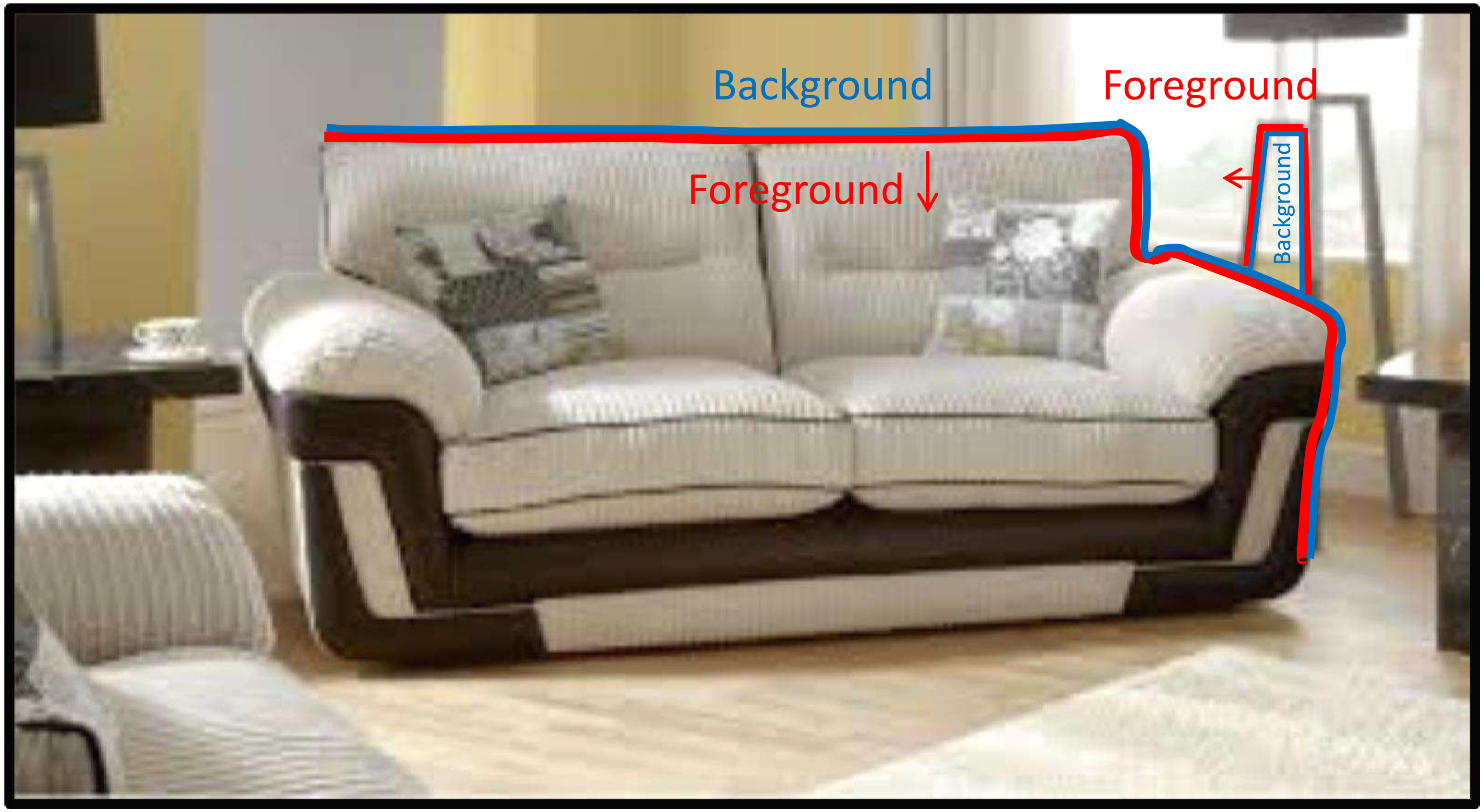
Fast Border Ownership Assignment with Bio-Inspired Features

Ching L. Teo, Cornelia Fermüller,
Yiannis Aloimonos
Computer Vision Lab and UMIACS
University of Maryland College Park

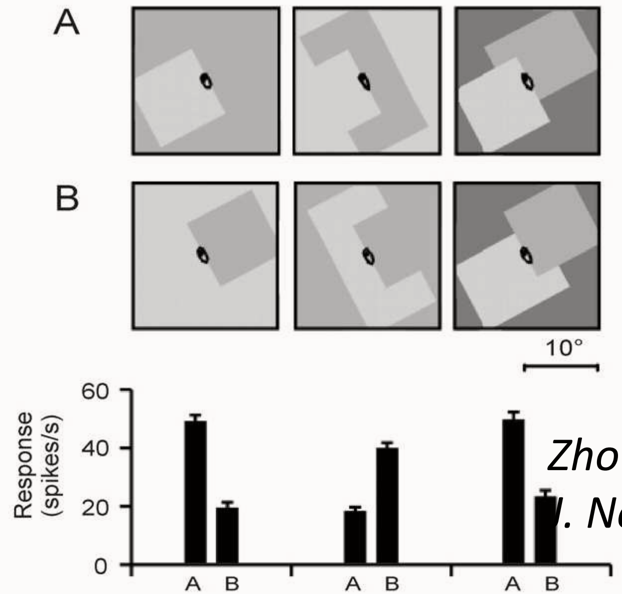
What is Border Ownership?



Motivations: Psychological & Biological

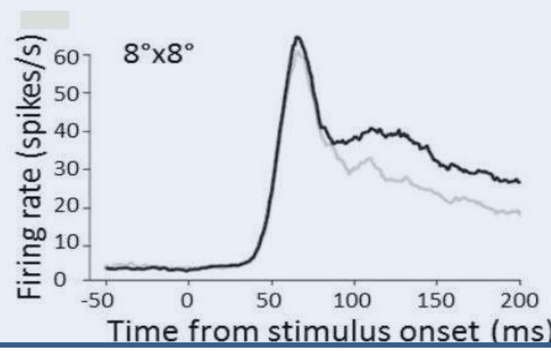
V2 cell that prefers bright Figure on the left

Selective nature of border ownership neurons (V2 & V4):



Zhou et al., J. Neuroscience 2000

Fast response time, <75ms from stimulus onset.

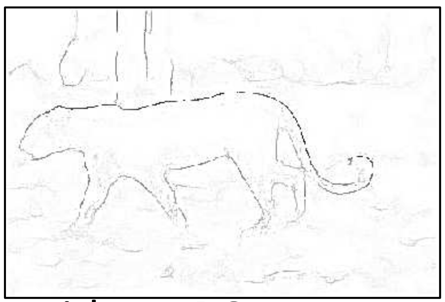


Sugihara et al., J. Neurophysiology 2011

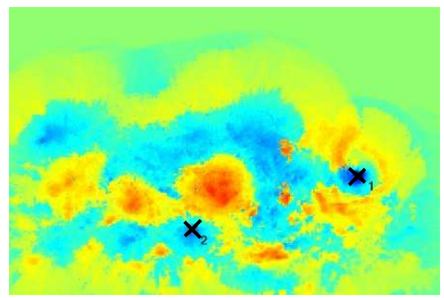
Figure-ground organization and attention are closely related

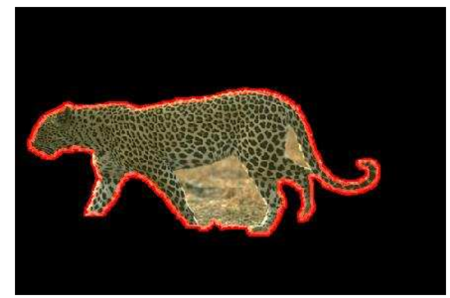
Craft et al, J. Neurophysiology 2007

Application: Figure-Ground Segmentation

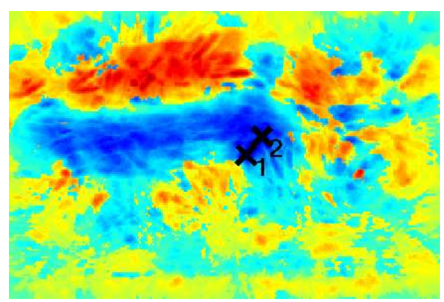


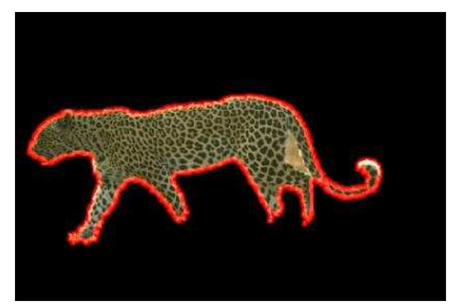
without BO



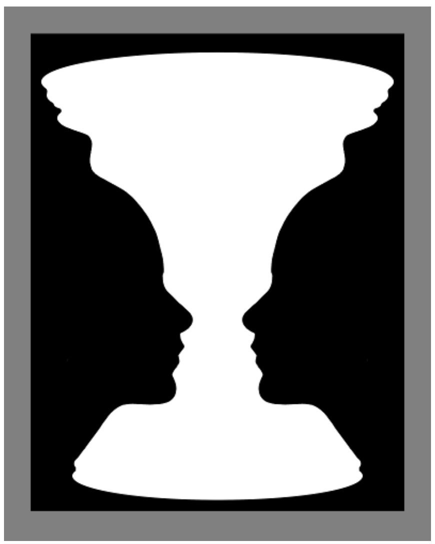


With BO

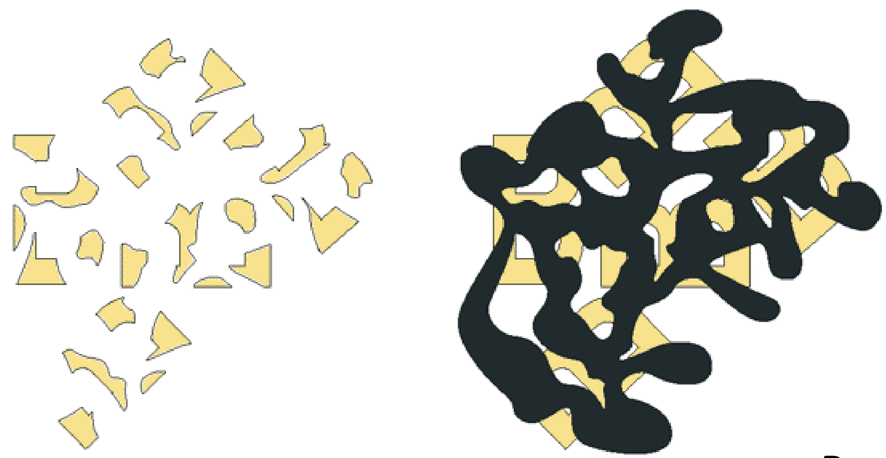




Shape perception and object recognition

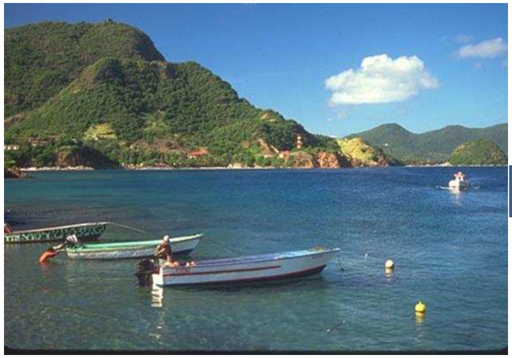


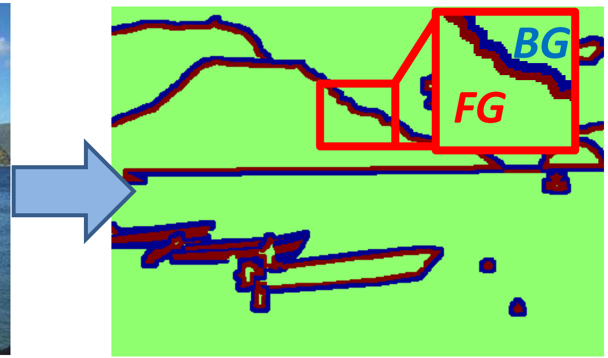
Rubin, 1915



Bregman, 1981

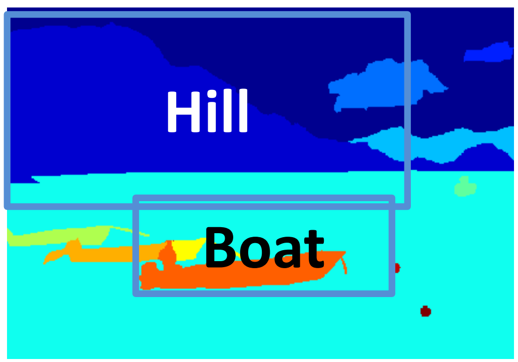
Related Work





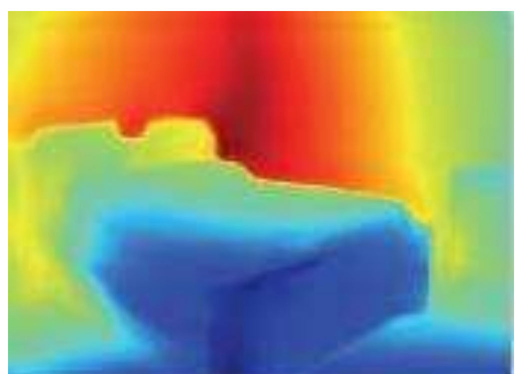
Border-ownership assignment to Edges using multiple cues + CRF

Ren et al., ECCV 2006
Leichter & Lindenbaum, ICCV 2009



Object proposals

Cheng et al. CVPR 2014
Yao et al., CVPR 2012
Gupta et al., CVPR 2013

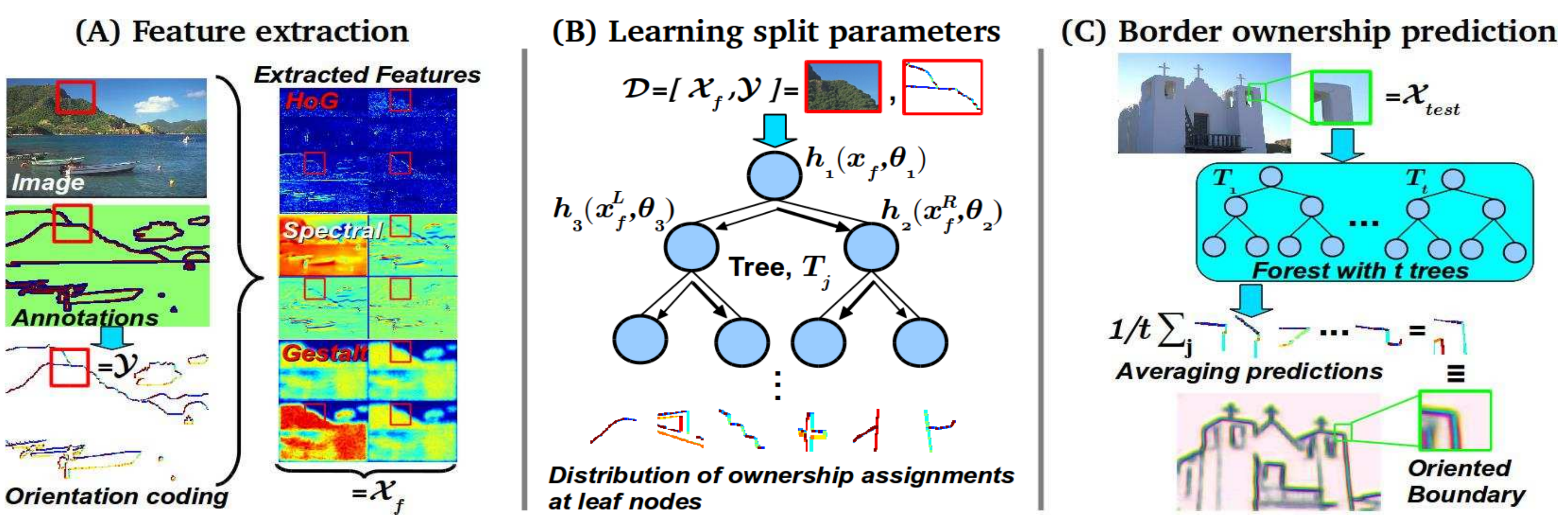


Depth + Normals from 2D

Saxena et al., PAMI 2009
Eigen et al, NIPS'14, ArXiv'15

Approach Overview

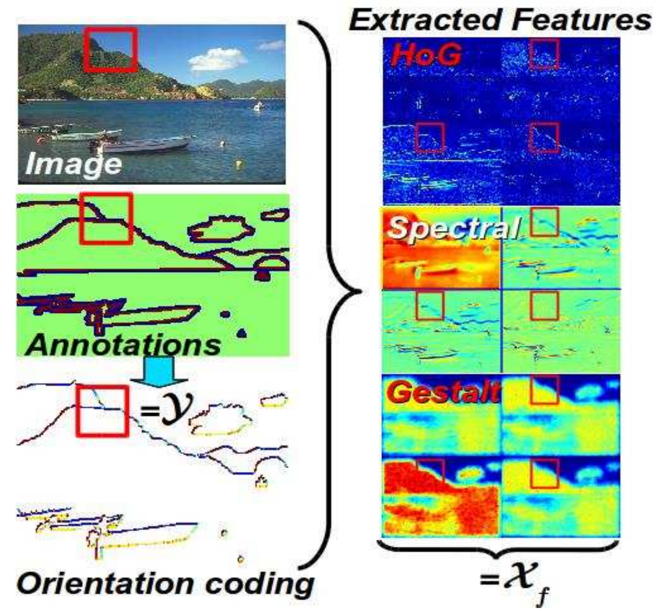
1. Extract patch-based features sensitive to ownership
2. Train a Structured Random Forest (SRF) that saves ownership structure at leaf nodes
3. Fast inference using SRF by averaging responses over all decision trees



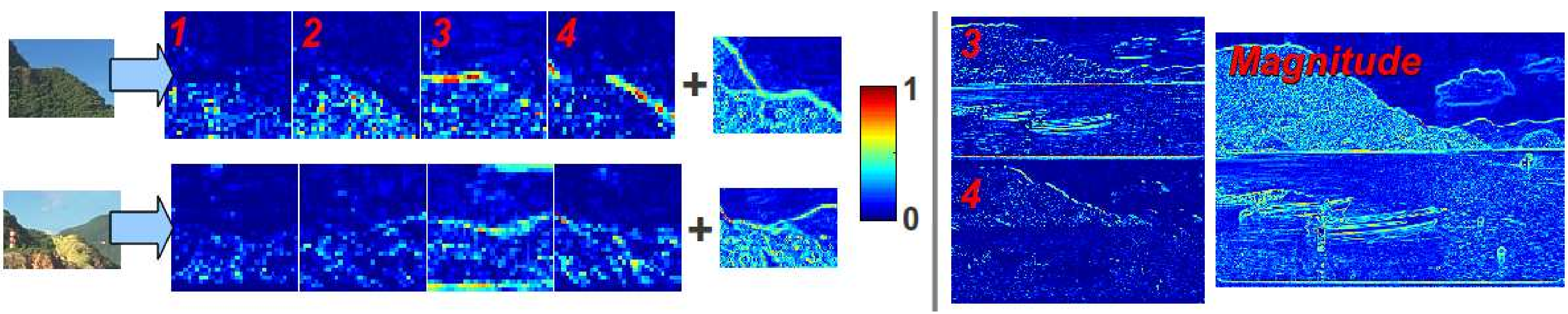
Feature Extraction

Three different features:

1. *Histogram of Gradients (HoG)*
2. *Spectral features of grayscale intensity*
3. *Gestalt-like grouping features*



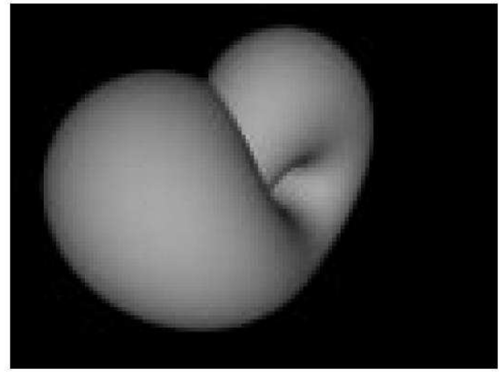
HoG orientations encodes local *shape*: **convexity and concavity**.
Magnitude localizes good boundary locations. *Palmer, Vision Science 1999*



Orientation (4 directions) + Magnitude

Feature Extraction: Local Ownership Cues

Extremal edges or ***image folds*** are characteristic changes in intensity along boundaries.

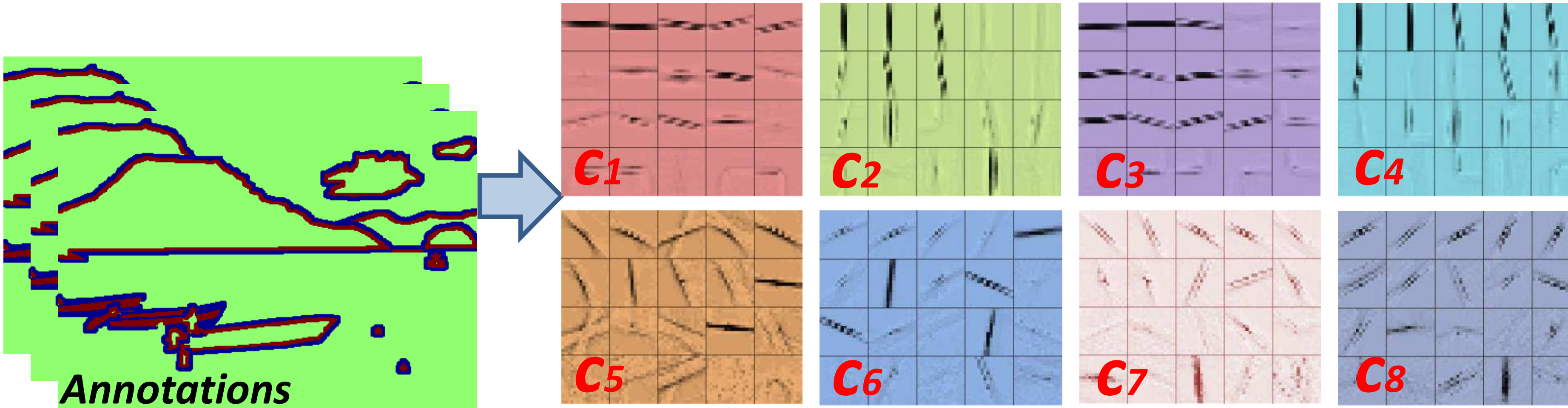



Huggins & Zucker, ICCV 2001

Psychophysical experiments have shown them to be one of the **strongest cues** for ownership.

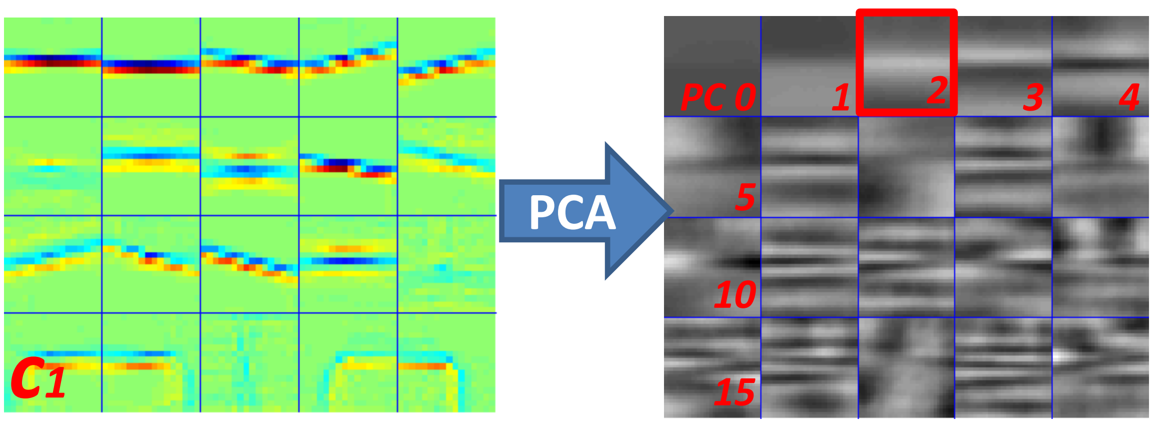
Ghose & Palmer, J. Vision 2010

Feature Extraction: Local Ownership Cues



“Sketch token” clusters of 8 ownership directions

Lim et al., CVPR’13



PC2 displays grayscale variations indicative of extremal edges.

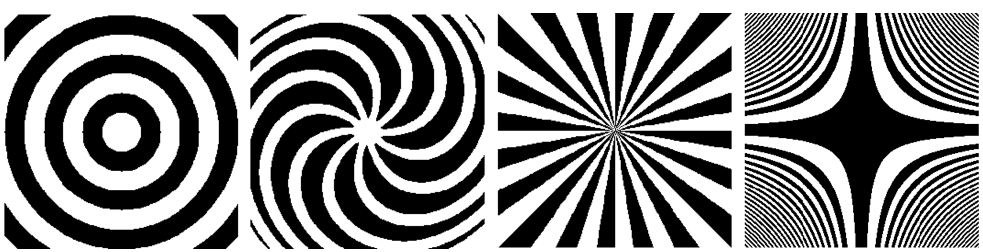
Ramenahalli et al., CISS’11

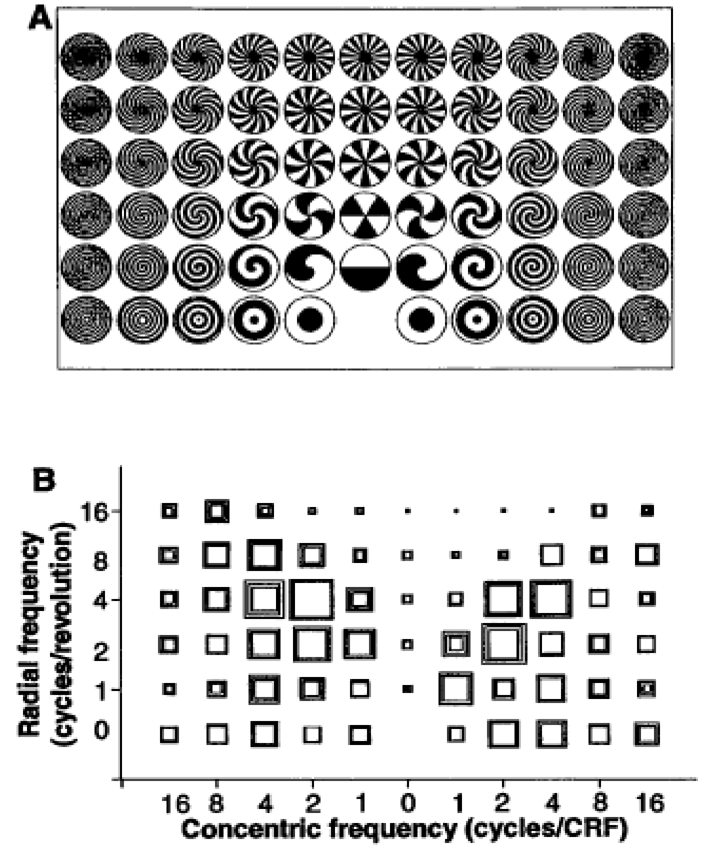
Feature Extraction: Global Ownership Cues

Border ownership is also determined by longer range (global) contextual cues.

Craft et al., J. Neurophysiology 2007

Implementation through visual operators that capture four grouping or "Gestalt" patterns:



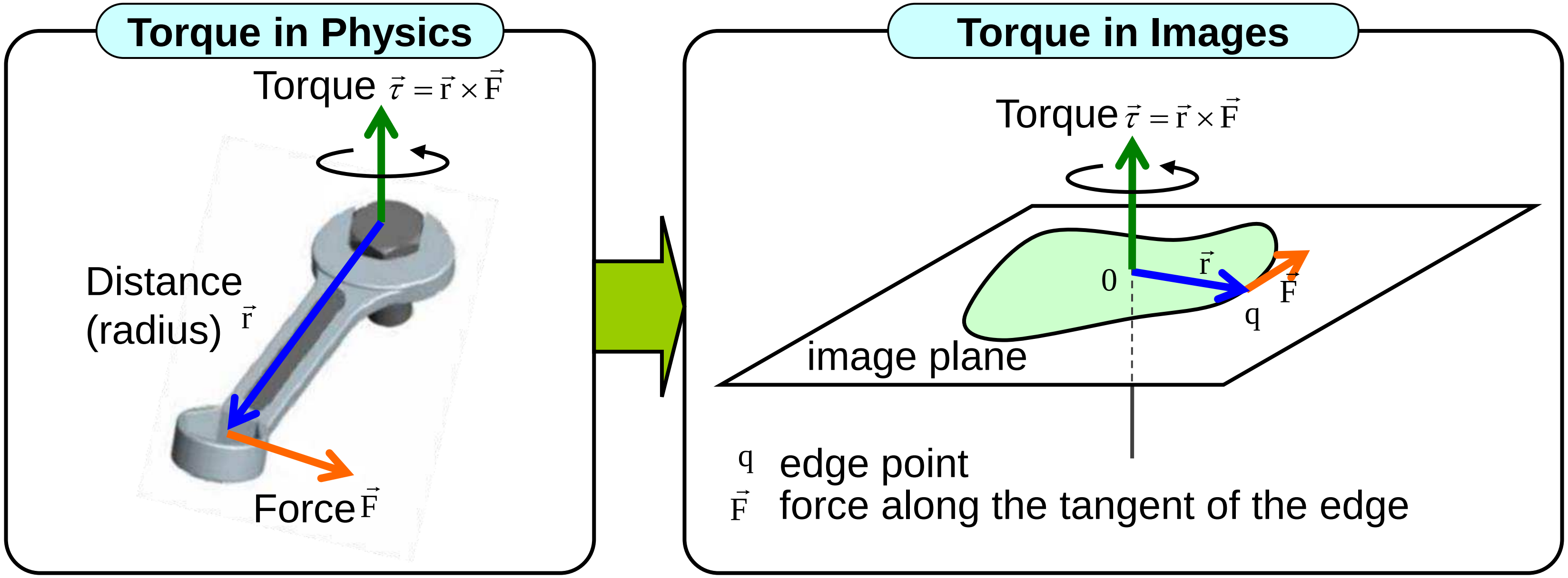


A: Gratings. B. Responses of a V4 cell

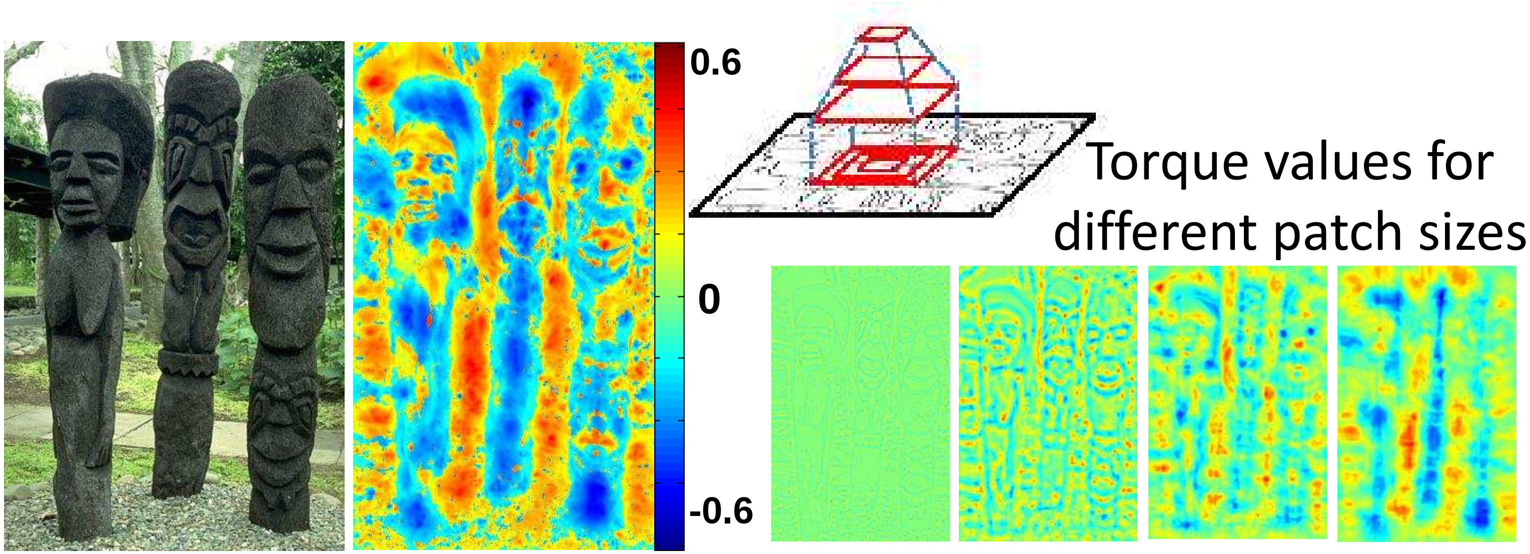
Cells tuned to these patterns have been observed area V4 of macaques :

Gallant et al., Science 1993

The Image Torque: Global Closure



The Image Torque: Global Closure

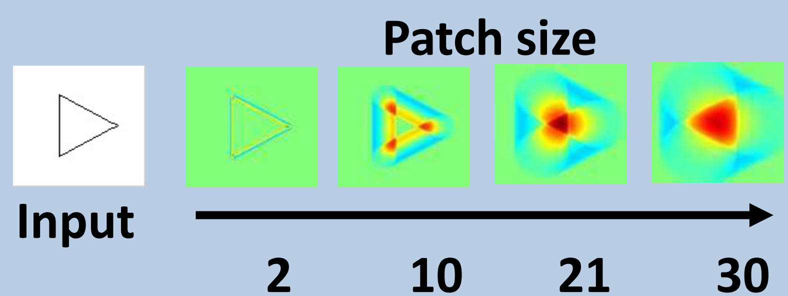


Torque value map

size 2 10 22 40

Key Properties:

1. *Largest response for* ***closed contours*** *at scale of patch.*
2. *Useful for attention and proto-segmentation.*

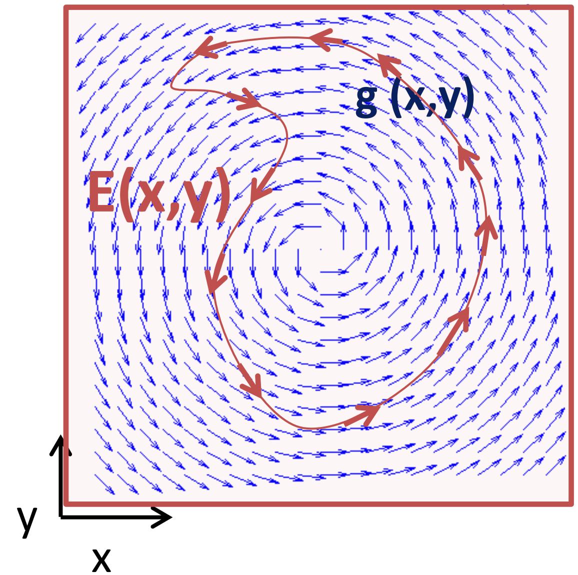


Nishigaki et al., CVPR 2012

Feature Extraction: Global Ownership Cues

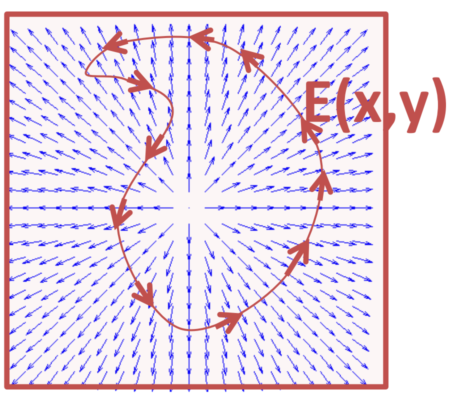
Rewriting the image torque as a *scalar product of* the edges (tangent vectors) *E(x,y)*

and a circular gradient field g(x,y).



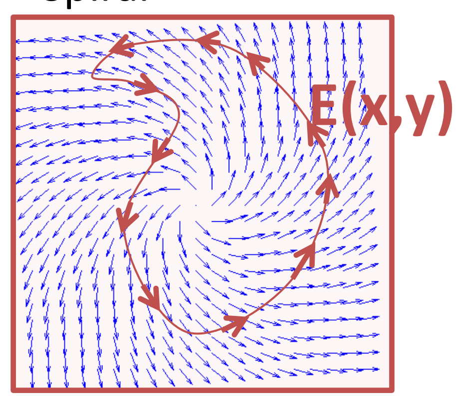
$$\tau_{pq} = \vec{F}_q \times \vec{d}_{pq} \quad \equiv \quad \tau(x,y) = E(x,y) \cdot (-y, x)$$

Radial



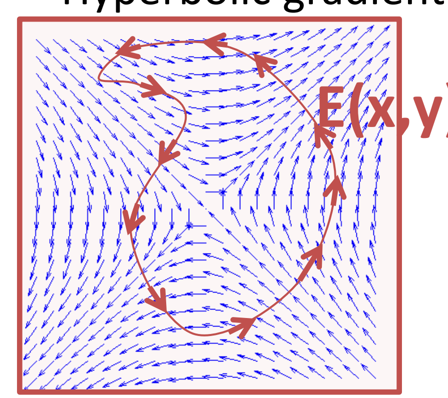
$g(x,y) = (x,y)$

Spiral



$g(x,y) = (ax - y, y + ax)$

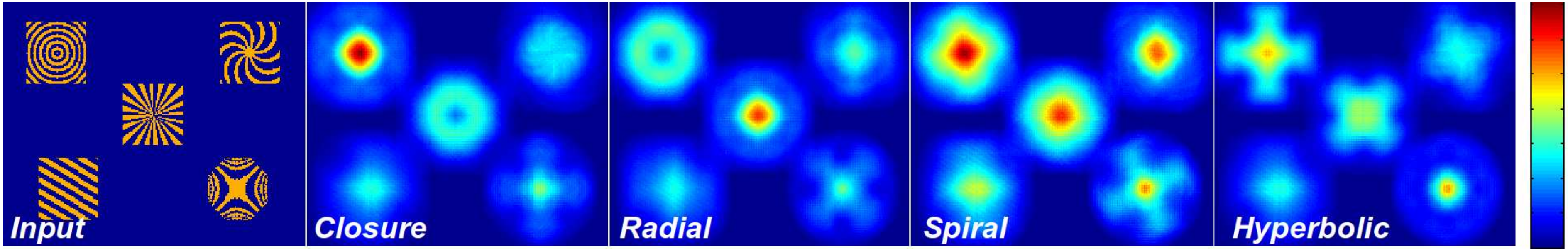
Hyperbolic gradient field



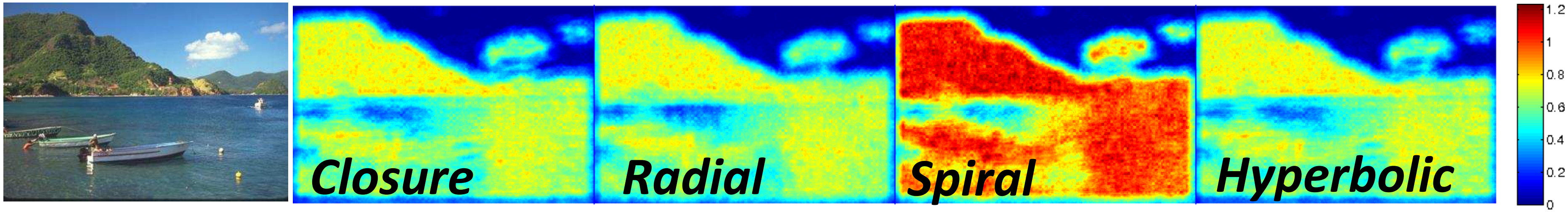
$g(x,y) = (ay, x)$

Feature Extraction: Global Ownership Cues

Responses over geometric patterns
(max response over multiple scales):



Responses over a real image are the **Gestalt-like features**
(over multiple scales):



Results

Predicted boundaries (red) and ownership (FG: green, BG:blue)

BSDS (100 training/100 testing)

Martin et al., PAMI 2004

NYU-Depth (795 training/ 654 testing)

Silberman et al., ECCV 2012

Ownership prediction accuracy:

Ren et al., ECCV 2006
Leichter & Lindenbaum, ICCV 2009

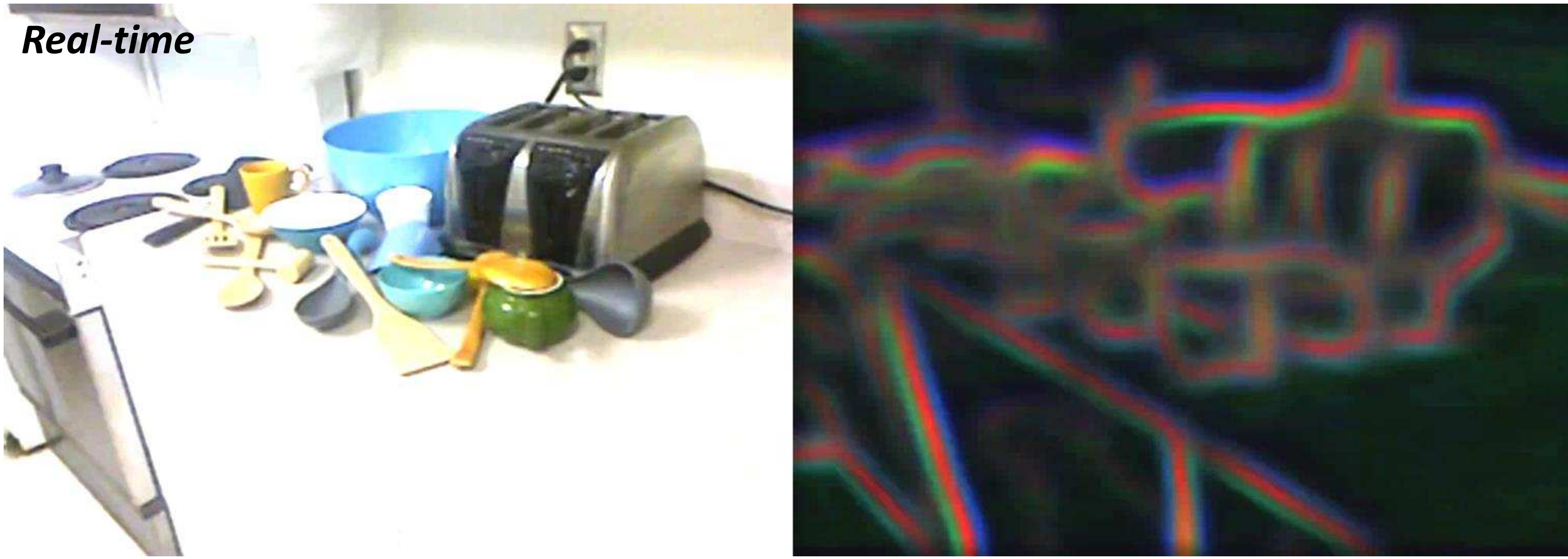
Feature set	BSDS	NYU-Depth
HoG	72.0%	66.0%
+ Spectral (no contour tokens)	73.1% (72.0%)	67.0% (65.6%)
+ Spectral (contour tokens)	74.0% (72.3%)	68.1% (66.7%)
+ Gestalt patterns	74.4% (72.7%)	**68.4**% (66.7%)
All features + Spectral (NYU)	**74.7**% (72.8%)	-
Global-CRF	68.9%	-
2.1D-CRF	69.1%	-

Boundary prediction accuracy:

Arbelaez et al., PAMI 2011
Dollar et al., PAMI 2015

Method	BSDS-500	NYU-Depth
Our approach	0.73,0.74,0.76	0.63,0.64,0.60
gPb-owt-ucm	0.73,**0.76**,0.73	0.63,0.66,0.56
SE	0.73,0.75,**0.77** (SE-SS)	**0.65**,**0.67**,**0.65** (SE-RGB)

Results



Red: Boundaries, Green: Foreground, Blue: Background

Outlook & Summary

We have presented an approach for **border ownership assignment**:

1. *Fast inference with state-of-the-art results.*
2. *Simultaneous boundary detection.*
3. *Validates the usefulness of shape, Gestalt and Extremal edge cues for this task.*

Closely related to:

- *Saliency and attention*
- *Object proposals*
- *Application: Layered segmentation*
- *Application: Scene understanding*

Acknowledgments

Code, Data and Full Results

www.umiacs.umd.edu/~cteo/BOWN_SRF/