Fast 2D Border Ownership Assignment

Ching L. Teo, Cornelia Fermüller and Yiannis Aloimonos

Computer Vision Lab and UMIACS, University of Maryland, College Park, MD

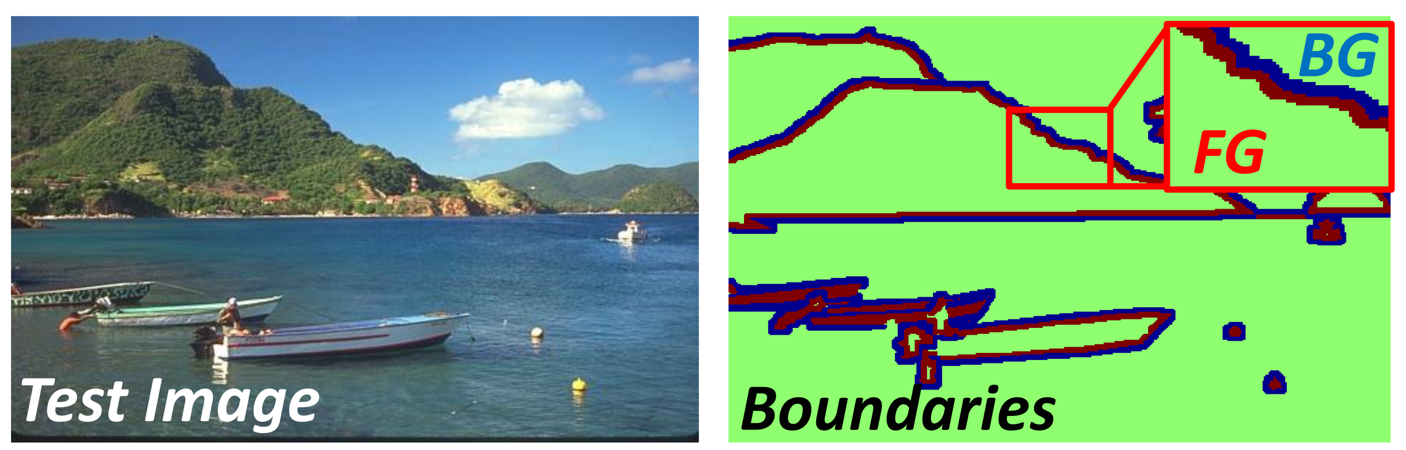
Code and data: http://www.umiacs.umd.edu/~cteo/BOWN_SRF

Abstract

We present a fast approach (~0.1s per 320x240 image) for detecting boundaries and **border ownership**, the relative ordinal depth along boundaries, using Structured Random Forests (SRF)s in real images. Key to the approach is the combination of local and global cues inspired from Gestalt psychology: **local shape**, **Extremal Edges** and **Gestalt-like** grouping patterns. Experimental evaluation over two diverse datasets of real images: a) The outdoor Berkeley Segmentation Dataset (BSDS) and b) The indoor NYU-Depth V2 highlights the speed, accuracy and generalizability of the approach compared to previous state-of-the-art multistage approaches.

What is Border Ownership?

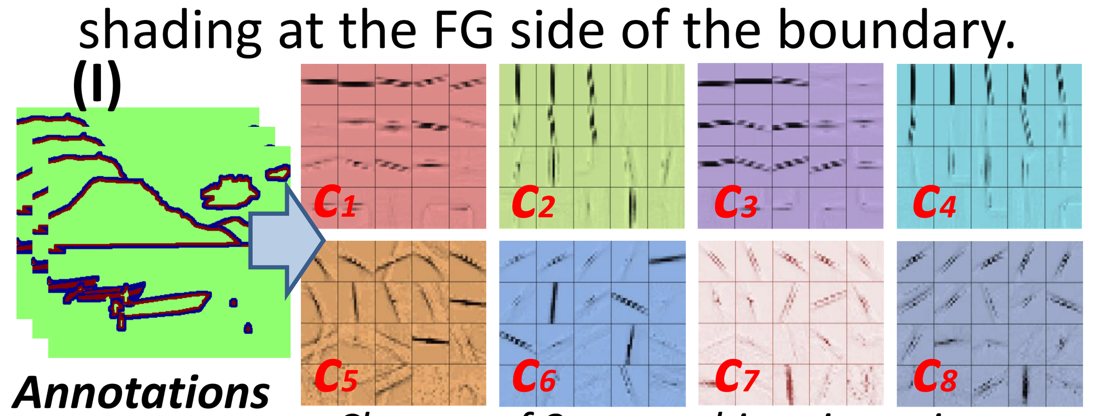
Given an image and its boundaries: regions where objects at different depth meet, the *border ownership assignment* problem is to determine which **side** of the boundary belongs to the object (foreground – FG) and which side is the background (BG).



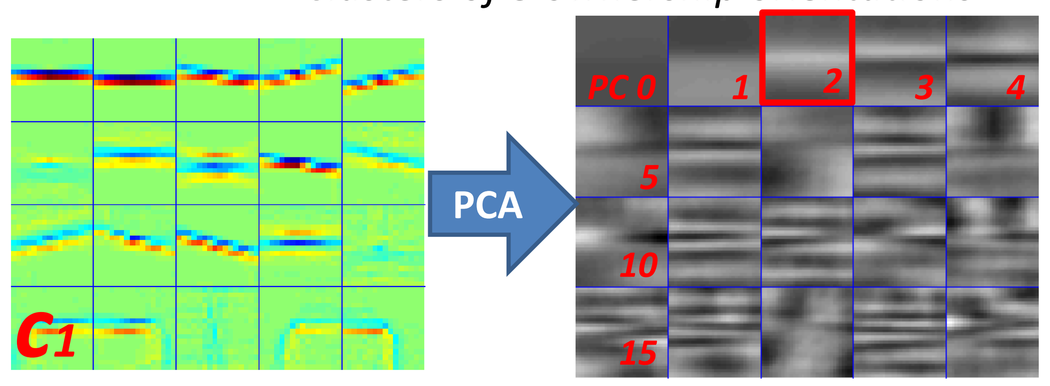
Border Ownership Cue 1: Extremal Edges (EE)

Extremal edges (EE), considered as one of the strongest cues for ownership[1], denote the specific change in grayscale intensities that occur along a true boundary of the object, with a distinctive shading at the FG side of the boundary.

(I)

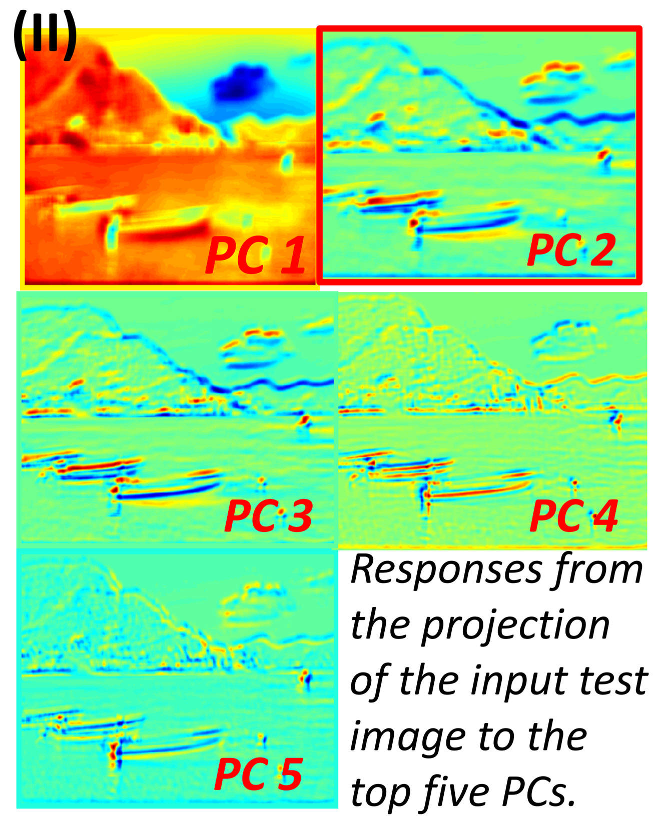


Clusters of 8 ownership orientations



PC2's grayscale variations are indicative of EE.

(II)



Responses from the projection of the input test image to the top five PCs.

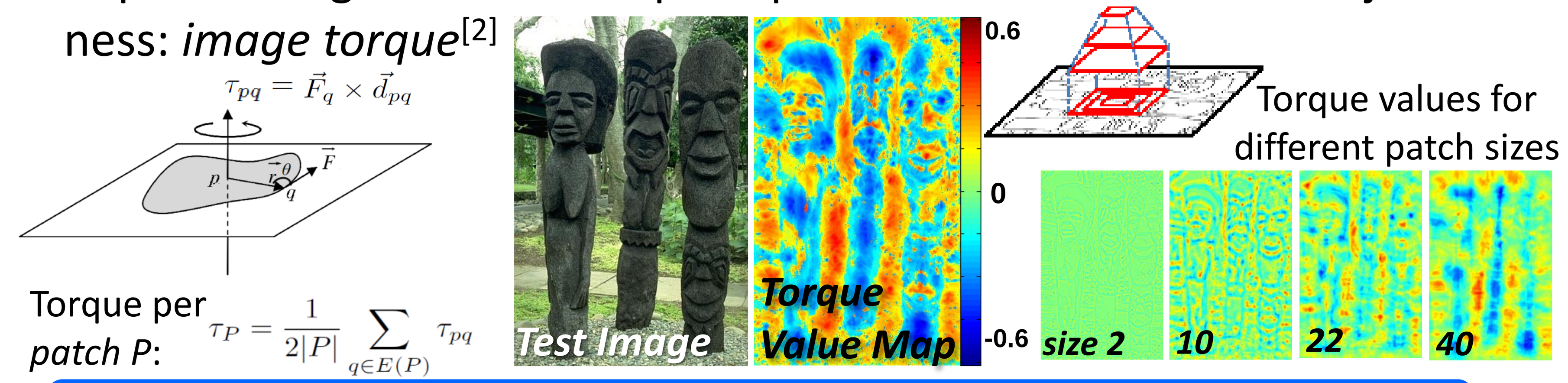
We analyze the intensity patterns within aligned patches over 8 ownership orientation clusters, using Principal Component Analysis (PCA) **(I)**. The top 5 principal components (PC) are then used as **spectral features (II)**, and the second PC encodes the EE feature.

A Mid-level Closure Operator: Image Torque

Implementing the "Gestalt" principle of **closure** to encode "object-ness: *image torque*[2]

$$\tau_{pq} = \vec{F}_q \times \vec{d}_{pq}$$

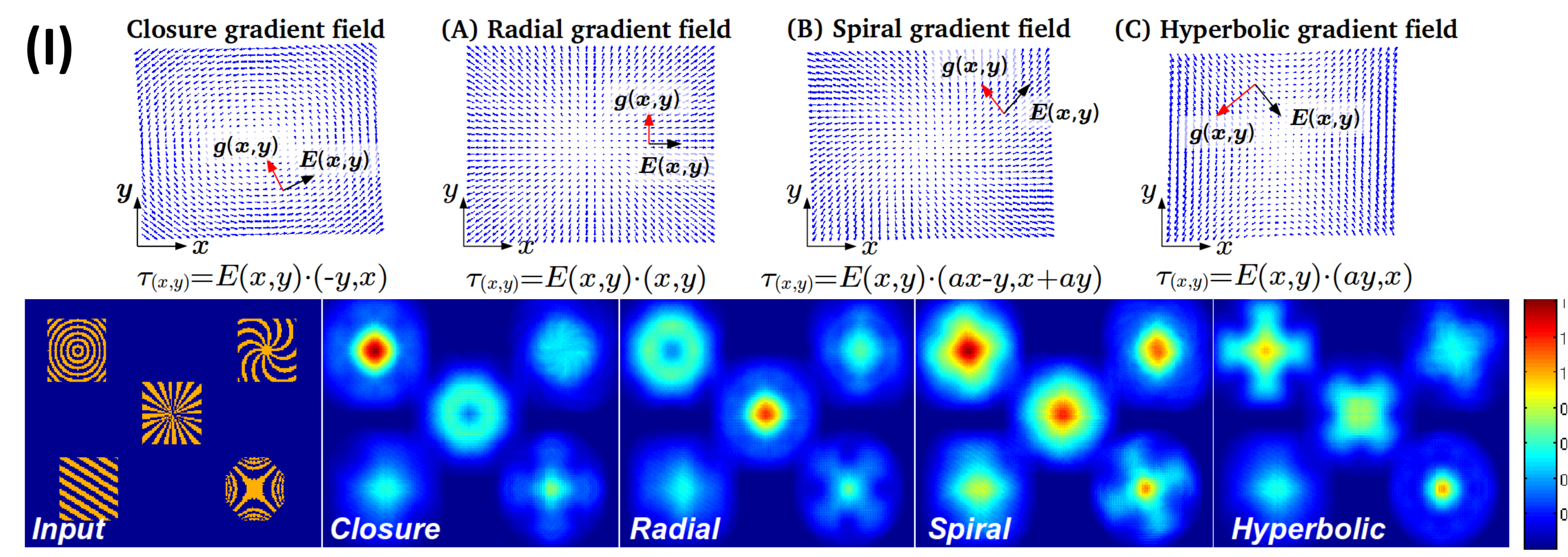
Torque per patch *P*: $\tau_P = \frac{1}{2|P|} \sum_{q \in E(P)} \tau_{pq}$



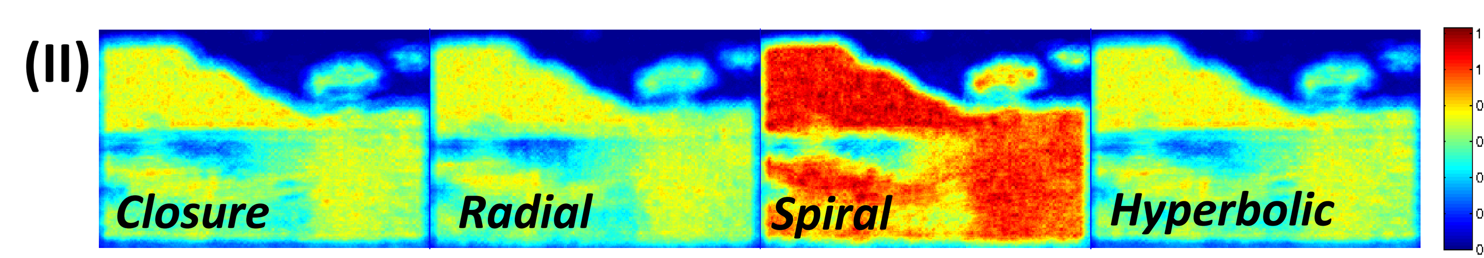
Border Ownership Cue 2: Gestalt-like Patterns

Additional patterns beyond closure have been observed in area V4 of macaques[3]. Besides closure, we extend image torque to 3 more Gestalt patterns: *radial, spiral* and *hyperbolic* **(I)**. The responses of the operator are then used as **"Gestalt"-like features (II)**.

(I)



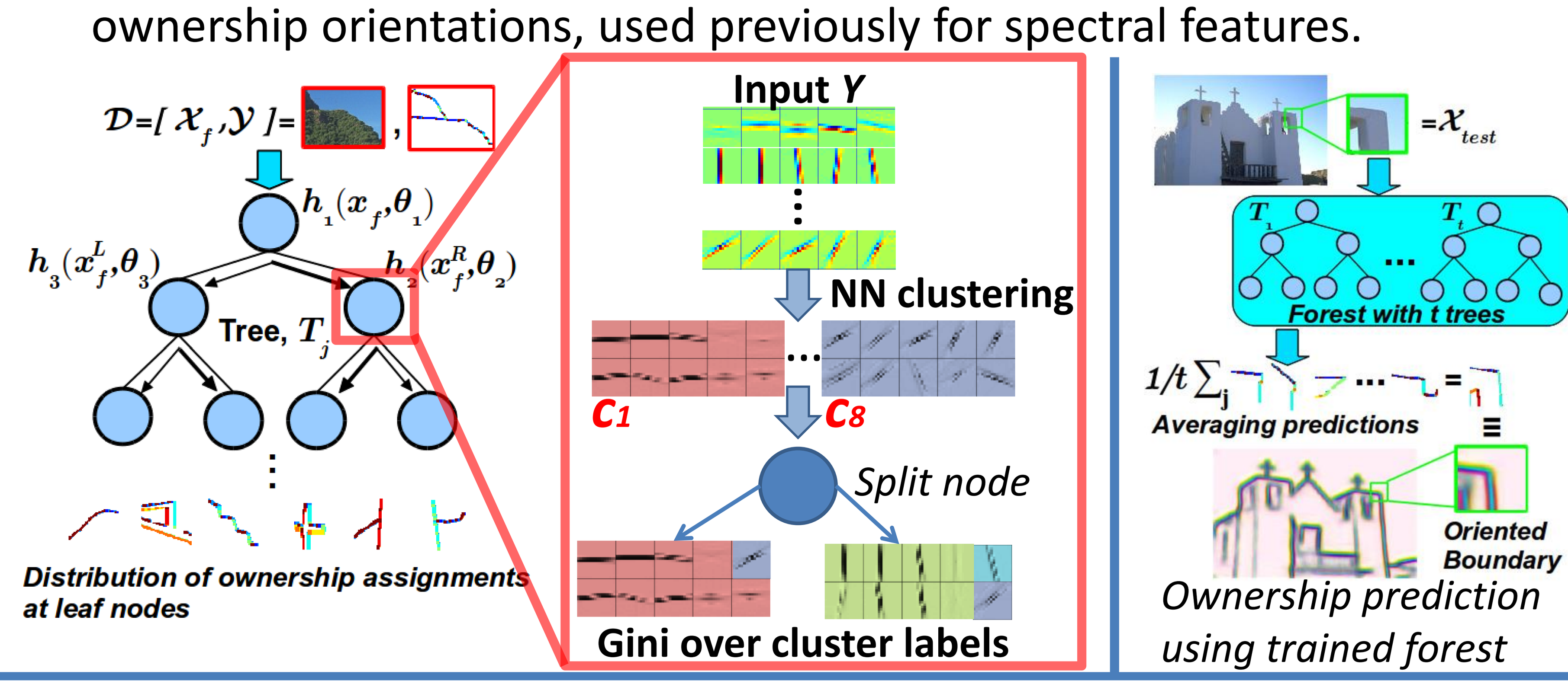
(II)



Responses for different Gestalt patterns over the input test image

SRF for Border Ownership Assignment

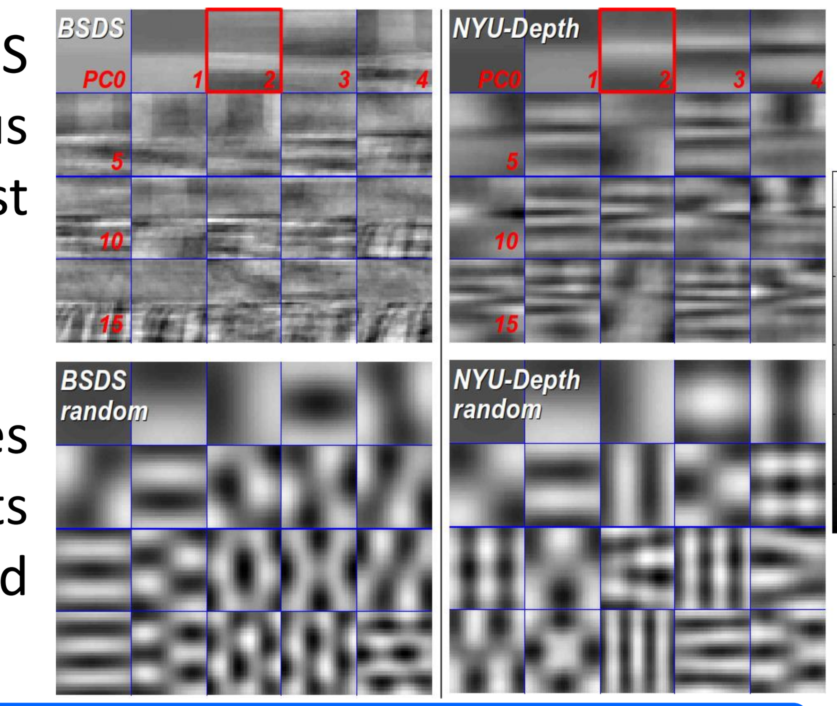
We train a SRF that associates these features with ownership annotations. The goal is to find the optimal splitting parameter, Θ_i, by computing the Gini impurity measure over 8 **class labels** of ownership orientations, used previously for spectral features.



Experiment 1: Comparing Spectral Features

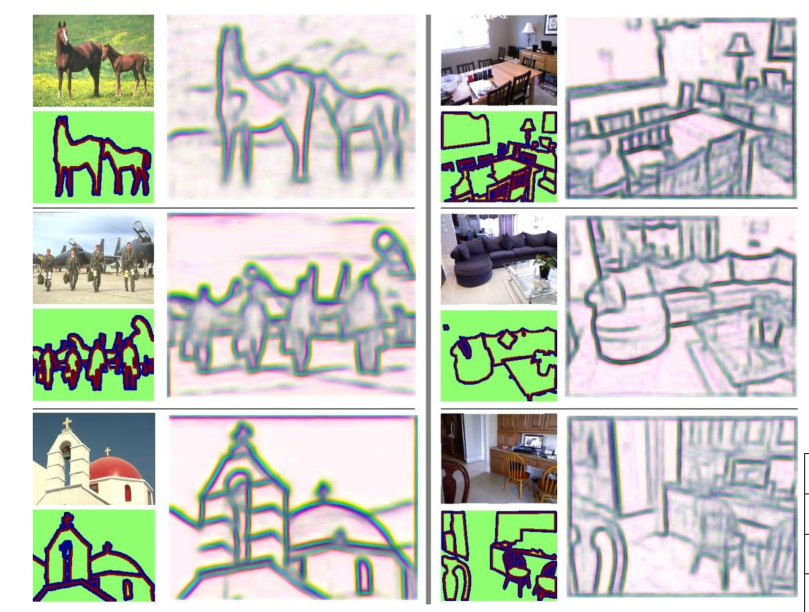
Spectral features from PCA of BSDS and NYU-Depth datasets versus random patches (below). The first 20 PCs are shown.

Notice the significant differences between the principal components obtained from boundary and random patches.



Experiment 2: Ownership and Boundary Accuracy

Feature ablation experiments over two datasets and comparison with two CRF-based border-ownership assignment approaches[4,5] are reported. Furthermore, evaluation of the *boundary prediction* accuracy using the BSDS-500 benchmark[6] yields comparable accuracies with state-of-the-art boundary detectors[6,7].



Example results: (L) BSDS dataset and (R) NYU-Depth. Blue: boundary, red: FG, yellow: BG

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 [4]	68.9%	-
2.1D-CRF [5]	69.1%	-

Feature ablations and comparisons with baselines

Method	BSDS-500	NYU-Depth
Our approach	0.73,0.74,0.76	0.63,0.64,0.60
gPb-owt-ucm [6]	0.73,**0.76**,0.73	0.63,0.66,0.56
SE [7]	0.73,0.75,**0.77** (SE-SS)	**0.65**,**0.67**,**0.65** (SE-RGB)

Boundary prediction accuracy [ODS, OIS, AP]

Conclusions

A real-time, state-of-the-art approach for border ownership assignment that combines perceptually plausible features with the Structured Random Forest classifier is described. Future works will focus on adding new features (motion and other Gestalt cues) and explore how ownership information can be exploited to improve segmentation and scene understanding.

Acknowledgments

This research was funded in part by the support of the European Union under the Cognitive Systems program (project POETICON++), the National Science Foundation under INSPIRE grant SMA 1248056 and DARPA Grant (W911NF-14-1-0384) under the project: Shared Perception, Cognition and Reasoning for Autonomy.

[1] Ghose and Palmer. "Extremal edges versus other principals of figure-ground organization". *J. Vision*, 10(8):3, 2010.
[2] Nishigaki et al. "The image torque: a new tool for mid-level vision". In *Proc. IEEE CVPR*, 502–509, 2012.
[3] Gallant et al. "Neural responses to polar, hyperbolic and Cartesian gratings". *J. Neurophys.*, 76(4), 2718–39, 1996.
[4] Ren et al. "Figure/ground assignment in natural images". In *Proc. ECCV*, 614–627, 2006.
[5] Leichter et al. "Boundary ownership by lifting to 2.1d". In *Proc. IEEE ICCV*, 9–16, 2009.
[6] Arbelaez et al. "Contour detection and hierarchical image segmentation". *IEEE T-PAMI*, 33(5):898–916, 2011.
[7] Dollár and Zitnick. "Fast edge detection using structured forests". In *IEEE T-PAMI, 2015*.

Scan for more Info: