Fast 2D Border Ownership Assignment

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

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 (SRFs) 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.

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).

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, $\theta$, by computing the Gini impurity measure over 8 class labels of ownership orientations, used previously for spectral features.

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 work 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.

References