

CIS Research Seminar - Thursday 11/15, 3PM, Jianbo Shi, Wu & Chen

Date: Thursday, November 15

Time: 3PM

Location: Wu & Chen

Speaker: Jianbo Shi

Title: Recognition and Segmentation

Abstract:

Our goal is to achieve large-scale object recognition, with learning, but with very few training examples. My main belief is visual intelligence occurs at multiple interconnected levels of perception, and they should be coupled tightly. I will present our recent works on integrating recognition with segmentation.

Bottom-up semantic image parsing. In many recognition tasks, one needs not only to detect an object, but also parse it into semantically meaningful parts. Borrowing concepts from NLP, we propose a bottom-up parsing of increasingly more complete partial object shapes guided by a composition tree. We demonstrate quantitative results from this challenging task on a dataset of baseball players with wide pose variation. There are two key innovations of our algorithm. First, at each level of parsing, we evaluate shape as a whole, rather than the sum of its parts, unlike previous approaches. This allows us to model nonlinear contextual effects on parts combination. Second, the parsing hypothesis is generated by bottom-up segmentation and grouping, while verification is achieved by top-down shape matching. By forcing the hypothesis and verification steps to be mutually independent, we reduce enormous false alarms (hallucinations) often occurring in background clutter.

Image matching. Image matching is a key building block for image search, visual navigation and long range motion correspondence. Our matching algorithm combines the discriminative power of feature correspondences with the descriptive power of matching segments. We introduce the notion of co-saliency for image matching. Co-saliency matching score favors correspondences that are consistent with "soft" image segmentation as well as with local point feature matching. We express the matching algorithm via a joint image graph whose edge weights represent intra- as well as interimage relations. We have demonstrated its application in the context of visual place recognition.

I will also briefly present our results on mid-level vision, shape

from shading and contour grouping using graph formulation.

This is joint work with Praveen Srinivasan, Alexander Toshev, Qihui Zhu, and Kostas Daniilidis.

-----