

- [Home](#)
- [About the CVR](#)
- [News](#)
- [Members](#)
- [Seminar Series](#)
- [Conference](#)
- [Resources](#)
- [CVR Summer School](#)
- [Research Labs](#)
- [Training at the CVR](#)
- [Partnering with the CVR](#)
- [Contact Us](#)

- Friday, September 29, 2006

The perceptual grouping of local features for image annotation and vision-based navigation

Local features, such as interest points, have become increasingly popular in computer vision due to their distinctiveness as well as their invariance. However, their limited spatial scope as well as their focus on image texture (as opposed to world structure) means that they cannot be readily grouped in a bottom-up fashion. Instead, their grouping into semantically meaningful collections must be driven by higher-level knowledge. In this talk, I will explore two such perceptual grouping frameworks. In the first half, we draw on the figure captions of cluttered images to learn, in a machine translation framework, associations between nouns and frequently occurring subsets of local features. These collections, in turn, can be used to annotate uncaptioned, cluttered scenes. In the second half, we turn to the problem of learning, in a graph partitioning framework, an optimal set of visual landmarks, each made up of a collection of interest points. These collections, in turn, allow a mobile robot to navigate its environment. In each case, the issues of feature stability and spatial coherence play a critical role in the grouping of local features. We demonstrate both frameworks on sets of real images. The first half of the talk is joint work with Mike Jamieson, Sven Wachsmuth, and Suzanne Stevenson, while the second half is joint work with Pablo Sala, Robert Sim, and Ali Shokoufandeh.

Sven Dickinson
U of Toronto