

Real-Time High-Performance Attention Focusing for Outdoors Mobile Beobots

Eric Pichon and Laurent Itti
Computer Science Department
University of Southern California

When confronted with cluttered natural environments, animals still perform orders of magnitude better than artificial vision systems in tasks such as orienting, target detection, navigation and scene understanding. The recent widespread availability of significant computational resources, however, in particular through the deployment of so-called “Beowulf” clusters of low-cost personal computers, leaves us little excuse for the enormous gap still separating biological from machine vision systems.

We describe a neuromorphic model of how our visual attention is attracted towards conspicuous locations in a visual scene. It replicates processing in posterior parietal cortex and other brain areas along the dorsal visual stream in the primate brain. The model includes a bottom-up (image-based) computation of low-level color, intensity, orientation and motion features, as well as a non-linear spatial competition which enhances salient locations in each of these feature channels. All feature channels feed into a unique scalar “saliency map” which controls where to next focus attention onto. Because it includes a detailed low-level vision front-end, the model has been applied not only to laboratory stimuli, but also to a wide variety of natural scenes. In addition to predicting a wealth of psychophysical experiments, the model demonstrated remarkable performance at detecting salient objects in outdoors imagery — sometimes exceeding human performance — despite wide variations in imaging conditions, targets to be detected, and environments.

The present paper focuses on a recently completed parallelization of the model, which runs at 30 frames/s on a 16-CPU Beowulf cluster, and on the enhancement of this real-time model to include motion cues in addition to the previously studied color, intensity and orientation cues. The parallel model architecture and its deployment onto Linux Beowulf clusters are described, as well as several examples of applications to real-time outdoors color video streams. Implementation on a 4-CPU rugged high-speed mobile robot, a “Beobot,” is also described. The model proves very robust at detecting salient targets from live video streams, despite large possible variations in illumination, rapid camera jitter, clutter, or omnipresent optical flow (e.g., when used on a moving vehicle). The success of this approach suggests that the neuromorphic architecture described may represent a robust and efficient real-time machine vision front-end, which can be used in conjunction with more detailed localized object recognition and identification algorithms to be applied at the selected salient locations.