

A Multi-Cue Vision Person Tracking Module

MIT2000-05

Progress Report: July 1, 2000—December 31, 2000

Trevor Darrell

Project Overview

This project will develop a multi-cue person tracking system that will integrate stereo range processing with other visual processing modalities for robust performance in active environments.

Next generation intelligent environments and interfaces require low-cost, easily configurable person tracking systems to provide perceptual awareness of users. We will build a robust multi-cue vision module that will provide these services. By exploiting the near orthogonal error modes of different cues or sensing modalities, this system can be more robust and real-time than a system based on any single cue. We plan to implement this system on a single motherboard system. We hope to demonstrate a laptop-based system with (relatively) low-cost stereo camera heads.

Progress Through December 2000

We have focused so far on the stereo range modality, and have been developing methods for fast and robust range estimation. We have developed a module which combines fast, predictive range estimation with dense range background models. Dense range models are constructed using long-term observations with multi-aperture and variable illumination conditions. We have used this module as part of an integrated person tracking system, which integrates the information about tracked people in each view into a single 3-D frame before estimating distinct trajectories for each person. A technical report on the NTTMIT web site describes this system.

Research Plan for the Next Six Months

We plan to finish the real-time implementation of the stereo person tracking system described above, and add a flesh color tracking modality.