KEYNOTE TALK
Tuesday, Nov 7, 2006
9AM–10AM

ISVC 2006: International Symposium on Visual Computing
Lake Tahoe, November 6-8, 2006

How to Win the Semantic Robot Vision Challenge

Daniel DeMenthon
National Science Foundation

Abstract

Our Robust Intelligence group at NSF is sponsoring a challenge competition organized by Paul Rybski and Alexei Efros (CMU) called the *Semantic Robot Vision Challenge*. It is a treasure hunt for mobile robots. Target objects will be spread around a conference room along with decoy objects. One hour before the competition, mobile robots will receive a text list of the target objects from a memory stick. The robots will be allowed during that hour to access the internet and autonomously learn about the visual appearance of the requested objects. Then the robots will race against the clock, roaming around the room and capturing snapshots of the requested objects into their memory sticks. The jury will consider completion times and collected snapshots to select a winner. The first competition will take place in 2007 at the AAAI Conference in Vancouver under partial AAAI sponsorship. Our goal with this challenge is to accelerate the pace of discovery and code sharing in object category recognition, scene understanding, text-based query of videos, active vision and visual navigation. In my talk I will give details about the challenge, and I will describe how close we are in the computer vision field to being able to succeed with this type of task, and what breakthroughs may still be needed for a competent performance in the competition.

Speaker Bio-Sketch: Daniel DeMenthon is Program Director for Computer Vision Research at the National Science Foundation since 2005. He is also an Associate Research Professor with the Laboratory for Language and Media Processing (LAMP) at University of Maryland. He graduated from Ecole Centrale de Lyon in France in 1972. After additional graduate degrees in Applied Mathematics from France, and Offshore Engineering & Naval Architecture from U.C. Berkeley, he received a Ph.D. in Computer Science from Université Joseph Fourier in Grenoble, France, in 1993. He developed fast algorithms for pose (POSIT, softPOSIT) and segmentation (Hierarchical Mean Shift). POSIT is now part of the OpenCV open source computer vision library. He was the winner of an Invention of the Year Award at University of Maryland for his 3D mouse development work using POSIT. He is the author of five U.S. patents and three European patents, and edited the book "Video Mining" with Azriel Rosenfeld and David Doermann. His present research focuses on document image processing and classification, object recognition, and fast alignment between video streams and 3D models of cities. For details, see http://www.cfar.umd.edu/~daniel