

New algorithm improves robot vision



Except in fanciful movies like 2003's *The Matrix Revolutions*, where fearsome squid-like robots maneuvered with incredible ease, most robots are too clumsy to move around obstacles at high speeds. This is true in large part because they have trouble judging in the images they "see" just how far ahead obstacles are. This week, however, Stanford computer scientists will unveil a machine vision algorithm that gives robots the ability to approximate distances from single still images.

"Many people have said that depth estimation from a single monocular image is impossible," says computer science Assistant Professor Andrew Ng, who will present a paper on his research at the Neural Information Processing Systems Conference in Vancouver Dec. 5-8. "I think this work shows that in practical problems, monocular depth estimation not only works well, but can also be very useful."

With substantial sensor arrays and considerable investment, robots are gaining the ability to navigate adequately. Stanley, the Stanford robot car that drove a desert course in the DARPA Grand Challenge this past October, used lasers and radar as well as a video camera to scan the road ahead. Using the work of Ng and his students, robots that are too small to carry many sensors or that must be built cheaply could navigate with just one video camera. In fact, using a simplified version of the algorithm, Ng has enabled a radio-controlled car to drive autonomously for several minutes through a cluttered, wooded area before crashing.

Inferring depth

To give robots depth perception, Ng and graduate students Ashutosh Saxena and Sung H. Chung designed software capable of learning to spot certain depth cues in still images. The cues include variations in texture (surfaces that appear detailed are more likely to be close), edges (lines that appear to be converging, such as the sides of a path, indicate increasing distance) and haze (objects that appear hazy are likely farther).

To analyze such cues as thoroughly as possible, the software breaks images into sections and analyzes them both individually and in relationship to neighboring sections. This allows the software to infer how objects in the image appear relative to each other. The software also looks for cues in the image at varying levels of magnification to ensure that it doesn't miss details or prevailing trends—literally missing the forest for the trees.

Using the Stanford algorithm, robots were able to judge distances in indoor and outdoor locations with an average error of about 35 percent—in other words, a tree that is actually 30 feet away would be perceived as being between 20 and 40 feet away. A robot moving at 20 miles per hour and judging distances from video frames 10 times a second has ample time to adjust its path even with this uncertainty. Ng points out that compared to traditional stereo vision algorithms—ones that use two cameras and triangulation to infer depth—the new software was able to reliably detect obstacles five to 10 times farther away.

"The difficulty of getting visual depth perception to work at large distances has been a major barrier to getting robots to move and to navigate at high speeds," Ng says. "I'd like to build an aircraft that can fly through a forest, flying under the tree canopy and dodging around trees." Of course, that brings to mind another movie image: that of the airborne chase scene through the forest on the Ewok planet in *Return of the Jedi*. Ng wants to take that idea out of the realm of fiction and make it a reality.

Source: Stanford University

This document is subject to copyright. Apart from any fair dealing for the purpose of private study, research, no part may be reproduced without the written permission. The content is provided for information purposes only.