Robot See, Robot React

Scientists seek to provide computers with greater understanding of dynamic visual data.

Advances in visual processing may soon allow robot vehicles to travel autonomously across battlefields and city streets. Researchers are developing mathematical models that offer insight into how mechanical and biological systems interpret images for movement and navigation. The answers will provide a key to designing more sophisticated automated guidance systems for commercial and military use.

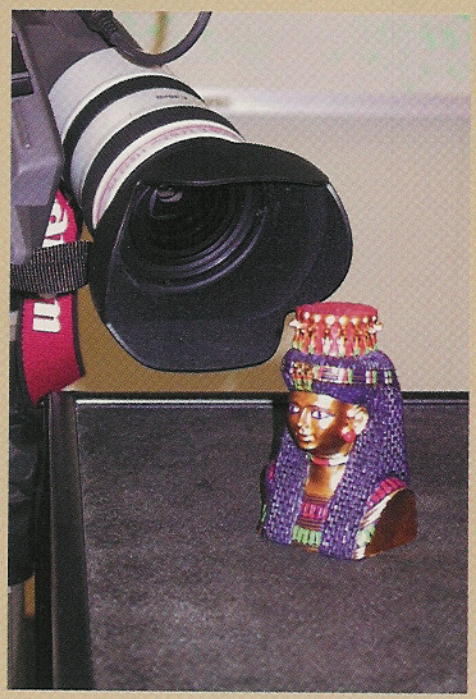
Permitting machines to navigate accurately using visual cues has been one major hurdle in robotics research. Although improvements in processing technology allow vehicles and aircraft to maneuver without human intervention, these systems continue to encounter difficulty in dynamic environments such as cities. While machines perform well in structured settings, the ability to perceive and move around in a totally unfamiliar setting would be a great step forward for robotics, experts say.

One of the institutions investigating the processes behind visual interpretation is the University of California at Los Angeles (UCLA). Researchers are studying dynamic vision, which is a computer's ability to use visual data about its surroundings and apply it in a changing environment to perform assigned tasks. According to Stefano Soatto, associate professor at UCLA's computer science department and head of the vision laboratory at the university's engineering school, potential applications for dynamic vision include basic robotic tasks, remote or autonomous movement and manipulation, tracking, docking, remote surgery, visual navigation and driving. But performing tasks like these is not simple. "It's not just a matter of connecting a camera to a computer. You need to interpret it. An image to a computer is just a bunch of zeros and ones. So how do you extract from these zeros and ones information about the shape of an object, its texture, position, orientation, motion and so forth?" he offers.

The crux of dynamic vision research is understanding how machines interpret the data they receive. For example, if a photograph is taken of an office that contains a desk, computer, books and a telephone, the human brain can immediately and effortlessly assess this information. But, if this image is scanned into a computer, it will be reduced to a series of numbers between zero and 255.

If the lighting in the room changes slightly and another picture is taken, a human will have no difficulty identifying the scene as being exactly the same as the original image. However, when this data is entered into a computer, all of the numbers will change. This process also will happen if the camera is moved slightly.

Much of how humans and machines interpret this information depends on a number of unknown factors, Soatto observes. "Your image depends on the shape of objects, on their material properties, light distribution, the position of the camera and its calibration—which is



Researchers at the University of California at Los Angeles are studying how machines interpret visual data. The results of this work will permit the creation of more sophisticated guidance and navigation systems for robot aircraft and ground vehicles.