Flysight

Research in computer vision may help computers of the future to see.

Consider the blowfly. NEC research scientist Rob de Ruyter van Steveninck certainly does. In one corner of his lab—located at NEC Research Institute in Princeton, New Jersey—a common blowfly is hanging out, sipping sugar water and watching the tube. A monitor is displaying a slightly changing pattern of lines.

Nearby, an oscilloscope displays the electrical pulses recorded by an electrode that taps into a single nerve cell in the fly's brain. As the pattern of lines moves across the fly's visual field, the on-screen waveforms fluctuate, synchronized to the audible crackle of radio static. If you pass your hand from left to right in front of the fly, the nerve cell's signal spikes, amplified visually and audibly by the equipment. Pass your hand right to left and the visual and audible signals are quiet—evidence of the fly's capacity to distinguish among different types of visual data.

What does a semiconductor company hope to learn from studying the sophisticated motion-detection and visual measurement systems of the blowfly? Understanding these systems may help scientists there figure out how computers may one day process visual information efficiently—a task for which the acrobatic blowfly is optimized. NEC is at the forefront of developing machine vision systems that may eventually enable computers to see.

Each fly eye is an array of 5,000 lenses, each with eight photoreceptors that are among the fastest known in the animal kingdom. Similar to a charge-coupled device (CCD) camera, the eye converts fluctuations in light intensity into electrical responses. The fly then processes the information, measuring velocity and sending instructions to its flight gear. The insect-vision team at NEC is trying to figure out how accurately the fly extracts visual information and encodes the information in its pixel array. The principles they're learning may one day be useful in developing artificial-vision applications.

DETECTING MOTION

NEC research scientist Jeffrey Mark Siskind is attacking the computing vision problem from another angle—motion detection. He has developed a computer system that identifies not just objects but also seven simple events, such as when he picks up a colored block, puts it down, and tips it over. It sounds rather mundane, but the potential applications are, in fact, extraordinary. One application Siskind envisions could help blind people navigate their environments.

In his lab, a camera hooked up to a PC records video for 3 seconds, during which time he tips over a block. The system has two trackers—one for color and one for motion—which record the relative and absolute positioning of the hand and the block. The system breaks down the momentous event into numerous steps and compares the positioning information with previously recorded events to identify the event accurately.

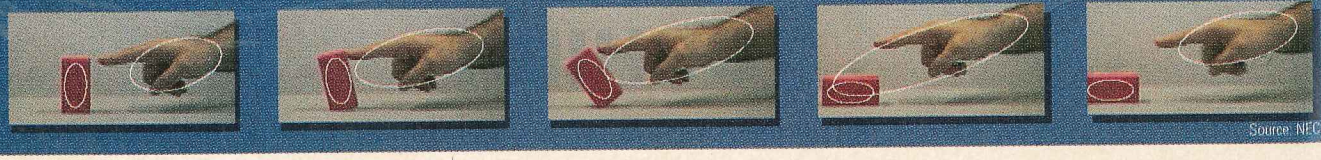
Ultimately, the technology may show up in smart glasses equipped with a camera that would feed image information to a processor, recognize events, and deliver an audible feed to a set of headphones that would describe the surrounding environment. The technology could also be used to train industrial robots to perform tasks by watching and then replicating tasks. And a smart conference room equipped with such technology would recognize a new speaker and automatically start the person's presentation.

The objective of Siskind's research is to endow computers with some of the expertise that humans have. "What allows us to describe the world in ordinary language?" he asks. His goal is to codify that understanding and enable computers to learn. "Kids watch and learn to correlate between language and what they see," he says. The ultimate goal, he explains, is to make the computer more reflective of the human.

Motion detection and computer vision are also subjects of research activities at IBM and Microsoft. Microsoft's research lab is working on incorporating gesture recognition into a future interface. And IBM has already demonstrated an interface that interprets gestures.—*Carol Levin*

Pushover

NEC's research in machine vision has produced a system that recognizes seven simple actions. In this *tip-over* event, a camera records 3 seconds of video, draws ellipses around the hand and block, and tracks their positioning. The computer then compares the values with models from previously recorded video, and it chooses the best match. Eventually, the technology could be used to train industrial robots and to make smart glasses that would verbally describe surroundings for blind people.



Source: NEC

ILLUSTRATION BY HAL MAYFORTH