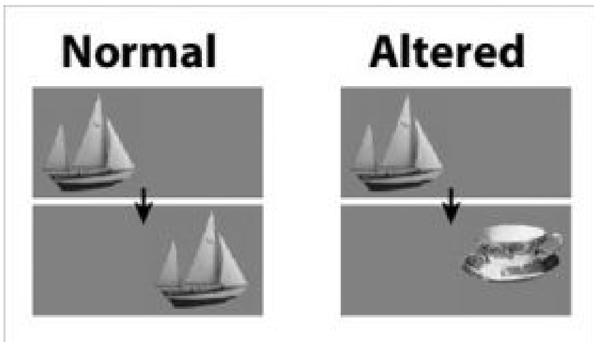
Watch and learn: Time teaches us how to recognize visual objects

September 11 2008



Example of stimuli used by Li and DiCarlo. An image of a sailboat is presented at different locations in the visual field. In most cases, it remains unchanged as the eyes move toward it, but at one test location, the stimulus is swapped for a teacup during the eye movement. As monkeys are exposed to this altered visual world, neurons in their brains begin to confuse the two objects when shown at the test location. The confusion is exactly that which is expected if the brain uses temporal contiguity to teach it how to recognize objects. Source / Nuo Li and James DiCarlo, McGovern Institute for Brain Research at MIT

(PhysOrg.com)—In work that could aid efforts to develop more brain-like computer vision systems, MIT neuroscientists have tricked the visual brain into confusing one object with another, thereby demonstrating that time teaches us how to recognize objects.

It may sound strange, but human eyes never see the same image twice. An object such as a cat can produce innumerable impressions on the retina, depending on the direction of gaze, angle of view, distance and so forth. Every time our eyes move, the pattern of neural activity changes, yet our perception of the cat remains stable.

"This stability, which is called 'invariance,' is fundamental to our ability to recognize objects—it feels effortless, but it is a central challenge for computational neuroscience," explained James DiCarlo of the McGovern Institute for Brain Research at MIT, the senior author of the new study appearing in the Sept. 12 issue of *Science*. "We want to understand how our brains acquire invariance and how we might incorporate it into computer vision systems."

A possible explanation is suggested by the fact that our eyes tend to move rapidly (about three times per second), whereas physical objects usually change more slowly. Therefore, differing patterns of activity in rapid succession often reflect different images of the same object. Could the brain take advantage of this simple rule of thumb to learn object invariance?

In previous work, DiCarlo and colleagues tested this "temporal contiguity" idea in humans by creating an altered visual world in which the normal rule did not apply. An object would appear in peripheral vision, but as the eyes moved to examine it, the object would be swapped for a different object. Although the subjects did not perceive the change, they soon began to confuse the two objects, consistent with the temporal contiguity hypothesis.

In the new study, DiCarlo and graduate student Nuo Li sought to understand the brain mechanisms behind this effect. They had monkeys watch a similarly altered world while recording from neurons in the inferior temporal (IT) cortex—a high-level visual brain area where object invariance is thought to arise. IT neurons "prefer" certain objects and respond to them regardless of where they appear within the visual field.

"We first identified an object that an IT neuron preferred, such as a sailboat, and another, less preferred object, maybe a teacup," Li said. "When we presented objects at different locations in the monkey's peripheral vision, they would naturally move their eyes there. One location was a swap location. If a sailboat appeared there, it suddenly became a teacup by the time the eyes moved there. But a sailboat appearing in other locations remained unchanged."

After the monkeys spent time in this altered world, their IT neurons became confused, just like the previous human subjects. The sailboat neuron, for example, still preferred sailboats at all locations—except at the swap location, where it learned to prefer teacups. The longer the manipulation, the greater the confusion, exactly as predicted by the temporal contiguity hypothesis.

Importantly, just as human infants can learn to see without adult supervision, the monkeys received no feedback from the researchers. Instead, the changes in their brain occurred spontaneously as the monkeys looked freely around the computer screen.

"We were surprised by the strength of this neuronal learning, especially after only one or two hours of exposure," DiCarlo said. "Even in adulthood, it seems that the object-recognition system is constantly being retrained by natural experience. Considering that a person makes about 100 million eye movements per year, this mechanism could be fundamental to how we recognize objects so easily."

The team is now testing this idea further using computer vision systems viewing real-world videos.

Source Massachusetts Institute of Technology