Robots With a Lot of Nerve

Eye-hand coordination may help a robot learn to see

By IVARS PETERSON

Robots have a tough time coping with an unpredictable environment. It's easy enough to program a robot to pick up an object if the object is exactly where it's supposed to be and is facing the right way. Similarly, a programmed robot doesn't have much difficulty moving from one place to another over a smooth, flat, unencumbered floor. The trouble comes when a robot stumbles into unfamiliar territory. Faced with something not covered by its instructions, the robot may suddenly stop as if in a trance, or it may crash ahead in a drunken fashion, oblivious to its surroundings.

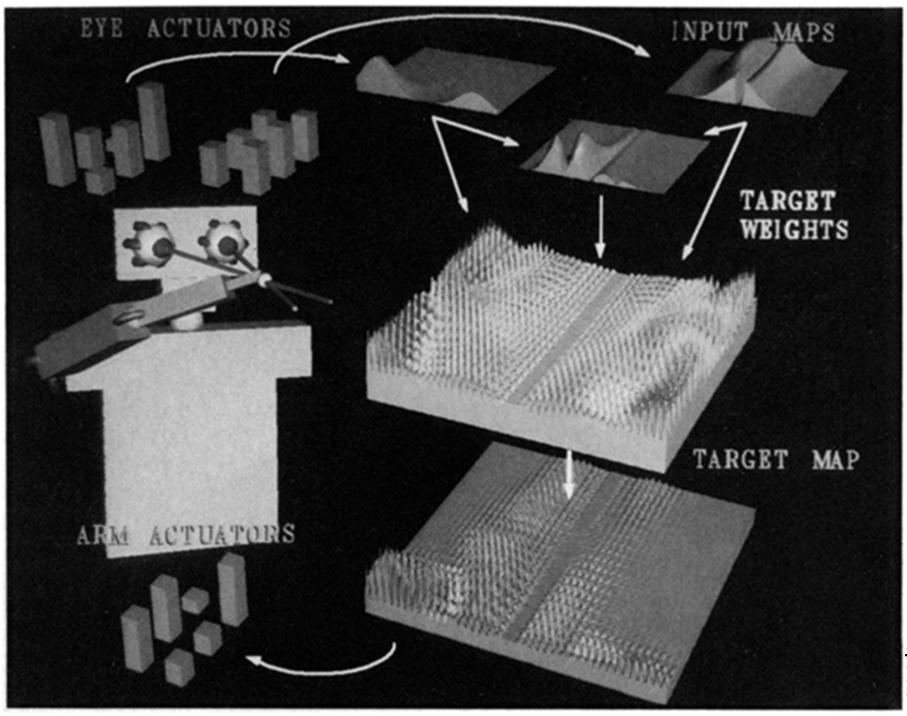
A human being, on the other hand, has the ability to adapt to new situations and to learn from them. This suggests that the principles of organization built into the human brain may also be useful for controlling robots and as models for "intelligent" computers.

Consequently, a large number of researchers, in both industry and universities, have been trying to simulate how the brain's nerve cells may be interconnected. In their research, these scientists construct and test mathematical models of neural networks. They look for arrangements that show a capacity for storing and retrieving information, recognizing patterns, understanding everyday speech and performing other complex tasks (SN: 1/24/87, p.60).

Michael Kuperstein of Wellesley (Mass.) College has focused on the connection between eye movement and muscle action, say, in an arm. His studies have led to the development of a new computer architecture that permits robots to learn about their surroundings.

"The architecture is composed of a number of interacting networks that enable robots to develop coordination by themselves," says Kuperstein. "These robots adapt to unforeseen changes in their mechanics by achieving their own sense of space. The brain-like architecture will even compensate for partial damage to itself."

His two-eyed, one-armed model robot, consisting of a pair of cameras mounted on one block and a two-jointed arm attached to an adjacent block, mimics in a cartoonish way a human's physiognomy. Kuperstein presented his results at the 1987 IEEE International Conference on Robotics and Automation, held recently in Raleigh, N.C.



Illustrations: Kuperstein

Kuperstein's robot learns to see by relating eye orientation to arm position. Signals from six actuators controlling each camera are transformed into an input map. The input maps from both cameras, together with a map conveying stereo information, are then multiplied by target weights, represented as an array of spines. The result is a new target map, which defines six arm actuator values used to position the robot's arm.

Kuperstein's work is based on several important findings from studies of how the brain processes visual information (SN:6/28/86,p.408). "I wanted to study neural network problems from as much of a biological basis as possible," he says. "My approach has developed by analyzing how biological systems determine movement control."

In the 1960s and '70s, Alan Hein and Richard Held of the Massachusetts Institute of Technology argued that movement is coordinated with vision. They showed that kittens that couldn't see their limbs and were not allowed to walk but were carried from place to place failed to develop the capacity to guide their bodies through space. In contrast, kittens that could see the results of their own leg movements were able to navigate successfully.

Kuperstein extends this result. "My hypothesis," he says, "is that we recognize, say, a cup not by its visual registration on our senses but rather by the combination of what we see and what we manipulate. As we move that cup through space, and we hold it as we move it, so will our muscles move."

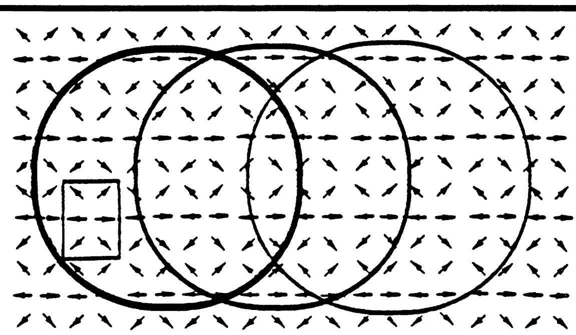
Moved closer, the cup looks bigger. Its shape and angles change. But the elbow joint also moves. The changing visual image registered on the eye's retina combines with changing muscle positions and tensions to produce a stable image in the mind. Even though the actual image is changing, a person's ability to recognize a cup remains unaffected.

Moving the eye instead of the cup also changes the image. "In some way," says

Kuperstein, "the brain is able to take into account voluntary movement of our sensory apparatus relative to the world and to maintain a stable perception of the world." That may happen because the brain correlates image changes with the action of certain sets of muscles, allowing it to interpret the effect of eye movements correctly.

Hence, Kuperstein's chief assumption is that all objects can be represented by how they're manipulated. This provides a finite vocabulary — essentially muscle signals — for describing objects.

Kuperstein also uses the experimental finding that much of the information entering by way of the eye, ear or some other sensory organ is generally registered in particular areas of the brain. Moreover, that information appears to be distributed locally across a set of brain cells in a way that closely resembles the way in which the information was gathered. For example, as the eye moves from side to side, signals go successively to groups of adjacent or nearby brain cells.



In Kuperstein's architecture, each one of the six limb actuators is represented by an arrow pointing in a particular direction. These arrows are distributed evenly across the target map. An input signal activates all the arrows within, say, a circle and modifies the weights of those elements. Slightly different camera positions send signals to different parts of the target map. Moving the cameras from left to right, for example, shifts the input circles from left to right across the target map. The collective output from all active arrows of a specific type determines the activity of each limb actuator. The rectangular area shown in the diagram represents one complete set of limb actuators.

Kuperstein's model robot is deliberately designed to look and act, in a limited sense, like an animal or a human being. It consists of two movable cameras attached to one block and two coupled limbs, representing an arm with joints at the shoulder and elbow, attached to another block. The shoulder joint has two pairs of "actuators," representing shoulder muscles, and the elbow joint has one pair. These allow the upper limb to rotate in two directions, and the lower limb to rotate in one. The stereo camera "eyes" can move back and forth or up and down. Each camera is controlled by six actuators, which shift the camera's position.

This model robot is a descendant of a single-jointed, one-eyed version that Kuperstein and mathematician Stephen Grossberg of Boston University developed a few years earlier to test the basic concepts involved in controlling robot movements. "We felt that if the brain could solve a behavior in a certain way," says Kuperstein, "then probably it would want to use that solution many times over, if it could, on other problems."

At first, the robot knows practically nothing about the space around it. Its control system has no *a priori* information about how sight is related to arm movement, how limb lengths and joint angles are related to endpoint position and how actuator signals affect joint angles. Learning begins when a random activity generator activates what Kuperstein calls a "target map," which puts the robot's arm in various positions. The robot's eyes sense where the arm's endpoint rests, generating an "input map" that's related to the angle of gaze of the eyes. The trick is to correlate what the robot sees (the input map) with what it does (the target map).

Once the relationship between camera angles used to point to a target and the arm joint angles needed to get to the target is established, the robot at any future time knows how to reach any object that happens to be within its field of view and causes the same convergence of its eyes. In other words, says Kuperstein, "when the system knows how to see where it moves, then it can move where it sees."

In Kuperstein's model, each camera generates an input map — a hilly, graphical landscape not unlike a topographic map — that corresponds to the position of the camera and the state of its actuators. Different camera positions generate different input maps. In addition, the input maps from the left and right cameras are combined into a single "disparity" map that provides stereo information.

All three camera input maps send signals to a target map, which is made up of interleaved elements representing the six actuators in the robot's two limbs. The input signals are modified by numerical weights associated with each element in the target map. As learning occurs, the weights change. The output from the target map generates six numbers that activate the three sets of "muscles" in the robot's arm. These numbers position the end of the two-jointed arm to the target.

Initially, the target is the endpoint of the robot's arm itself. Step by step, the robot learns to correlate its camera positions with a randomly specified arm placement. "It's actually making mistakes," says Kuperstein, "and learning from its own mistakes." Discrepancies between input and output maps guide this learning. Later, when the cameras focus on a particular spot of light located somewhere in space, the robot should be able to move its arm to the right location.

"How accurately the arm's position reaches the target," says Kuperstein, "will depend on how well the target map weights allow the input map signals to be correlated with the limb-actuator representations." Computer simulations show that the average error, when target map ouputs are compared with randomly generated limb activations, is about 4.3 percent after several thousand trials.

"From my point of view," says Kuperstein, "the beauty of the model is that it is adaptable. You can change the muscle strength or the actuator strength. You can change the joint length. You can change the optics of the eye, the distance between the eyes. The system will still be able to reorganize itself and maintain its self-consistency."

At the same time, Kuperstein's neural network model interacts with the world and responds to changes and uncertainty in the environment. It organizes its own perception of the world and modifies that perception as needed.

His model is also fault-tolerant. Because of the way the target map is organized as a set of interleaved elements and the way information is distributed across these elements, the breakdown of a few processors causes only a minor change in the whole network's output. Moreover, the system corrects itself — in effect, circumventing the problem processors.

Kuperstein believes it would be relatively easy to build a robot-control system based on his proposed neural architecture. He has already applied for patents to cover the basic concepts involved in his network design.

"As I see it," he says, "the model can be generalized to other behaviors." That includes robot functions such as walking over uneven terrain or recognizing and grasping unfamiliar objects.

"The vision I work toward," he says, "is that this architecture will be usable in many different sensory-motor combinations." The same architecture, in different contexts or settings, may someday allow such "neural" robots to learn to see, touch and hear. □