# GET SMART

## ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

## Pave Robotics' Future

By Nick Adde

A robot cannot write this article. Not yet, anyway.

The day when a cyborg journalist can be assigned a task — in this case, writing a story about artificial intelligence — interview experts, respond to their answers, assess the totality of the material and craft an article that accurately digests it is a long way off.

Sooner, realistically, an intelligent robot could assist the skin-and-bones journalist who is writing the story. It could respond to specific commands from the ever-industrious writer, enabling him or her to stay focused on the computer keyboard and the task at hand. Go into the break room and find a misplaced cell phone. Search the Internet for specific work in artificial intelligence, at a given university, company or government agency. Bring it here.

Or get coffee. A robot with artificial intelligence capability that is new to the office even could respond, "Cream or sugar? Where are the cups?"

Once it learns where those items are, it would never have to ask again. The diligent scribe would have more time to stare at the computer screen and hope he or she is coming up with the right words.

But while comparing the present cutting edge of artificial intelligence technology with a fully functional humanoid robot would be akin to the difference between a Viking long boat and the space shuttle, those who could benefit from it would have a much shorter wait — perhaps within the next decade, rather than a millennium.

### Sensing a Purpose

Indeed, researchers are hard at work building machines that can be given a job, or sequence of jobs, and perform them without assistance from a human operator. At the University of Washington in Seattle, a team is trying to develop robots that have a richer understanding of their environment, with the ability to recognize objects, pictures, and people and interact with them. A key component for these robots — its 3-D camera system — comes straight from the video-game world.

At Carnegie Mellon University in Pittsburgh, another group is taking advantage of the same 3-D camera to teach robots three key components essential to artificial intelligence — the abilities to assess the world they are in, make decisions and act upon those decisions.

And at the Massachusetts Institute of Technology, another group is building a system that would enable the U.S. Navy to land unmanned aerial vehicles onto the decks of aircraft carriers. The UAS, also equipped with 3-D cameras, would be able to follow and obey the hand and arm gestures that landing signal officers use to guide manned aircraft in.

"Our research has been compelling," says Manuela Veloso, a computer science and robotics professor at Carnegie Mellon. "We have robots that move completely and autonomously in the ninth floor of our building."

It is possible she said, because of something called symbiotic autonomy — the ability to rely upon other machines, the Web, human beings and their own databases to learn to perform tasks.

"They are capable of asking for help," she says.

A major breakthrough came when Microsoft Corp. introduced the Kinect system for its Xbox video game platforms. With Kinect, Veloso said, 3-D cameras that once cost thousands of dollars could be had for under $200 apiece.

"They [Kinect cameras] are remarkable. Robots can now get depth easily in images and identify objects with much more reliability and accuracy. We envision many more robots now being made available because of these particular sensors," Veloso says.

Veloso admits that the project has caused her to alter the way she approaches her job.

"Somehow, I, as a researcher, thought intelligence was about doing it all yourself," Veloso said.

Thus, an office or home someday would rely upon multiple robots — which she calls "cobots," or companion robots — that can share information, rather than a single entity, Veloso believes.

"That's the goal we're going for in unmanned systems. Program them, set them up so they are outside [the] boundaries of what we think they can do alone," Veloso says. "It's too complex a world. You will never be able to tell a machine what to do in every conceivable case. Humans are much better at handling novelty and uncertainty," she says.

Dieter Fox, a computer science and engineering associate professor at the University of Washington, is taking advantage of the inexpensive 3-D capability Kinect affords, to design and develop robots with precise object-recognition capabilities.

Fox, who made inroads in developing robots that functioned as tour guides in museums more than a decade ago, intends to use the Kinect camera system to enhance their object-recognition and manipulation capabilities.

In the foreseeable future, Fox says, a consumer could buy a robot and bring it home. On the first day, the owner would show the robot a very specific object, such as a particular coffee mug. The robot in the future would recognize that mug among others.

"It would interact with humans in a more natural way," Fox says. "You could tell the robot what to do with gestures."

Such devices would have obvious practical applications in the eldercare domain, helping them "live a more independent life, rather than in a nursing home," Fox says.

While Fox and his colleagues continue to make inroads, he cautions that many obstacles need to be overcome before such devices achieve ubiquity. Manipulator arms, for example, are still very expensive.

And robots still have trouble "dealing with the huge complexity of everyday environments, recognizing 20 different objects [among] the hundreds of thousands of different objects," Fox says.

"It's not going to happen within the next five years," Fox says.

## UAS With Machine Learning?

Meanwhile, Randall Davis, a professor at MIT's Computer Science and Artificial Intelligence Lab, is working with doctoral candidate Yale Song in collaboration with the Office of Naval Research that would make unmanned aircraft as commonplace on aircraft carrier decks as manned aircraft.

Before that happens, UAS must be able to recognize the sequences of hand and arm signals that tell pilots to turn brakes on or off, get ready for positioning on catapults, turn left or right, stop or go. Audible communications are useless in the noisy, busy and potentially dangerous environment.

Davis and Song also are incorporating 3-D cameras in their prototype devices, which would be able to recognize body and hand poses and combine them to recognize and distinguish among gestures. The system would have to see the motions, and determine what they mean, by assessing statistical data about factors such as the percentage of time one arm might be stuck straight out from the shoulder while the other arm is pointed up in the air, for instance.

"These gestures are sometimes in a static position and sometimes moving," Davis says.

The AI system would "need to recognize not just body position at this instance, but then be able to go from a sequence of detected body positions over time to determine that this is a particular kind of gesture," Davis says.

If machine learning gets integrated into unmanned systems, as is the goal of an ONR program, UAS landing on aircraft carriers, like the X-47B UCAS aims to in 2013, could follow the gestures of those on the carrier signaling the aircraft. Photo courtesy Northrop Grumman Corp.

Davis compares the activity to a person watching a movie, without consciously noticing how body positions and gestures change with each frame.

"Nobody is telling you this," Davis says. "It's easy for people, but a whole lot harder to teach a computer to do it."

It would seem simple enough merely to design a program that recognizes gestures, but that will not necessarily be good enough because of the variation in human gestures, Song says. If everyone issued a given command, for instance, by putting one arm up at a 45-degree angle, such a plan would work. But human beings do not act that way; one landing signal officer may put an arm up at a 30-degree angle, another at 50 degrees. Also, the same landing officer may execute such gestures energetically at the beginning of a watch and less so at the end as fatigue becomes a factor.

"That's why we need a machine-learning approach, which is more robust to unexpected events," Song says.

"What makes our gesture recognizer interesting is that it learns statistical relationship between body and hand poses using graphical models, a popular machine-learning technique," Song says. "This significantly reduces the amount of training data and computational resources."

Song claims that their method was able to perform on par with conventional approaches using only half the training data.

One day, Davis says, he hopes their research would expand further into the realm of gesture recognition and interaction.

"You could wave your hand at a TV set, or a media center, and control it. After all, why not? Why should I have to keep looking for the damned remote control, which is stuck under the couch?" Davis says.

And if Davis had is way, the writer staring at the computer screen would be able to use gestures to communicate with sources and editors.

"You and I aren't having this conversation with a keyboard and mouse," Davis says. "Human interaction gets left behind when interacting with a keyboard and a mouse."

While Davis and Song have forged significant inroads in gesture recognition — with 87 percent accuracy to date, Davis considers their project as basic research.

"While 87 percent is promising, it needs to be better than that to be used in a high-impact application. I'm not sure how much harder the last 13 percent will be than the first 87 percent. We anticipate it will be harder, but we're not sure."

*Freelance writer Nick Adde has covered technology and personnel matters, primarily relating to the U.S. armed forces, since 1983.*