

[sciencedaily.com](https://www.sciencedaily.com)

## New AI sees like a human, filling in the blanks

4-5 minutes

---

Computer scientists at The University of Texas at Austin have taught an artificial intelligence agent how to do something that usually only humans can do -- take a few quick glimpses around and infer its whole environment, a skill necessary for the development of effective search-and-rescue robots that one day can improve the effectiveness of dangerous missions. The team, led by professor Kristen Grauman, Ph.D. candidate Santhosh Ramakrishnan and former Ph.D. candidate Dinesh Jayaraman (now at the University of California, Berkeley) published their results today in the journal *Science Robotics*.

Most AI agents -- computer systems that could endow robots or other machines with intelligence -- are trained for very specific tasks -- such as to recognize an object or estimate its volume -- in an environment they have experienced before, like a factory. But the agent developed by Grauman and Ramakrishnan is general purpose, gathering visual information that can then be used for a wide range of tasks.

"We want an agent that's generally equipped to enter environments and be ready for new perception tasks as they arise," Grauman said. "It behaves in a way that's versatile and able to succeed at

different tasks because it has learned useful patterns about the visual world."

The scientists used deep learning, a type of machine learning inspired by the brain's neural networks, to train their agent on thousands of 360-degree images of different environments.

Now, when presented with a scene it has never seen before, the agent uses its experience to choose a few glimpses -- like a tourist standing in the middle of a cathedral taking a few snapshots in different directions -- that together add up to less than 20 percent of the full scene. What makes this system so effective is that it's not just taking pictures in random directions but, after each glimpse, choosing the next shot that it predicts will add the most new information about the whole scene. This is much like if you were in a grocery store you had never visited before, and you saw apples, you would expect to find oranges nearby, but to locate the milk, you might glance the other way. Based on glimpses, the agent infers what it would have seen if it had looked in all the other directions, reconstructing a full 360-degree image of its surroundings.

"Just as you bring in prior information about the regularities that exist in previously experienced environments -- like all the grocery stores you have ever been to -- this agent searches in a nonexhaustive way," Grauman said. "It learns to make intelligent guesses about where to gather visual information to succeed in perception tasks."

One of the main challenges the scientists set for themselves was to design an agent that can work under tight time constraints. This would be critical in a search-and-rescue application. For example, in a burning building a robot would be called upon to quickly locate

people, flames and hazardous materials and relay that information to firefighters.

For now, the new agent operates like a person standing in one spot, with the ability to point a camera in any direction but not able to move to a new position. Or, equivalently, the agent could gaze upon an object it is holding and decide how to turn the object to inspect another side of it. Next, the researchers are developing the system further to work in a fully mobile robot.

Using the supercomputers at UT Austin's Texas Advanced Computing Center and Department of Computer Science, it took about a day to train their agent using an artificial intelligence approach called reinforcement learning. The team, with Ramakrishnan's leadership, developed a method for speeding up the training: building a second agent, called a sidekick, to assist the primary agent.

"Using extra information that's present purely during training helps the [primary] agent learn faster," Ramakrishnan said.

This research was supported, in part, by the U.S. Defense Advanced Research Projects Agency, the U.S. Air Force Office of Scientific Research, IBM Corp. and Sony Corp.

### **Story Source:**

[Materials](#) provided by [University of Texas at Austin](#). *Note: Content may be edited for style and length.*