Machine vision system could help the visually impaired shop for food

May 28 2015, by Cherie Winner

Graduate student Siddharth Advani displays a webcam-equipped haptic glove he helped design to help visually-impaired people shop for groceries. Credit: Patrick Mansell

You're in the mood for pasta, so on the way home from work you stop at

the grocery store and pick up rotini, shaved Parmesan cheese, and the organic tomato sauce you favor. Into the store and back out, 15 minutes, tops. Simple, right?

For those of us who can see, it is. For those of us who are blind or have limited vision, a simple trip to the grocery store can be a major chore.

"You always have to find someone at the store to help you," says Michelle McManus, an IT consultant at Penn State and president of the Happy Valley chapter of the National Federation of the Blind. "Then you have to explain exactly what you want—" and hope the person helping you is diligent about getting it right.

Now researchers at Penn State are leading an effort to help visually-impaired people shop independently. They're creating machines that can interpret a complex visual scene much as the [human brain](#) does. They're making machines that can truly see.

Ambitious vision

This work is part of "Visual Cortex on Silicon," a massive endeavor that spans fields of inquiry ranging from materials design to brain circuitry and includes nearly 50 researchers, from grad students to senior scientists, at Penn State and seven other institutions. Research is under way on many fronts at the same time, with new findings from each field shedding light on the problems in other fields. What neuroscientists learn about the architecture of the mammalian visual cortex helps computer scientists design circuits that reflect the way the brain works.

In 2013 the project won a five-year, $10 million "Expeditions in Computing" award from the National Science Foundation. It is led by Penn State computer scientist Vijay Narayanan, who speaks in rapid-fire bursts and thrives on complex collaborative projects.

"I learn every day from people who work in other fields," says Narayanan. "That's what keeps me running!"

The project's formal name refers to the goal of creating a digital, silicon-based electronic system that performs like the human visual cortex, the part of our brain that processes and interprets visual information.

The project also has an informal name, "Third Eye," inspired by the Hindu god Shiva, whose third eye fills the universe with kindness and spews fire to dispel evil. The name suits both the metaphoric and practical aims of the project: If successful, the project will provide its human operators with additional, often enhanced, visual information that will make their lives easier and safer.

Seeing, shopping, learning

Visual Cortex on Silicon addresses three "domains" or end uses, each of which will augment human vision in particular ways. Third Eye-AR (Augmented Reality) and Third Eye-DA (Driver Assistance) will aid in the recognition of objects and people in a variety of settings, including busy streets and urban battlegrounds. Most of the team's effort in its first year has gone into the third domain, Third Eye-VI, where the aim is to develop a system coupled to a wearable device that will help visually-impaired people do their grocery shopping.

Narayanan, Distinguished Professor of Computer Science and Engineering, says the "million-dollar question" in all three projects is whether the abilities of a cognitive system, be it electronic device or human brain, are due more to its hardware/structure or its software/algorithms. He and his colleagues are exploring multiple solutions to this question, ranging from new software that can run on existing processors to new hardware "fabrics" that have the potential to learn on their own.

He says his team's goal is to develop a system that will recognize that an object it sees is new to it, and store that object in memory. If it encounters the same or similar items enough times, that category will take on more importance. At some point, the system may prompt its human operator to give the item a name and tell the system where it fits in its collection of all known items.

Pay attention

A major challenge in all three domains is to create a system that will know what to pay attention to within a crowded visual field. The human visual cortex has two general modes of attention, says Narayanan. The "bottom-up" mode is akin to browsing, where we take in the scene without looking for a particular item—until something catches our eye because it stands out from its surroundings, like a face we recognize in a crowd or an orange sale sticker on a grocery shelf. In "top-down" mode, we're looking for a specific item and our eyes are drawn to things or qualities (size, color, shape) that we know resemble that item.

Third Eye scientists are trying to devise a machine vision system that can operate in either mode or combine the two, depending on the situation. Their major challenge is how to get the system to deal with a complex scene. For several years now, electronic image systems have been able to pinpoint faces and chunks of text in a scene—unless the scene is too cluttered. We need a system that can direct its attention to significant objects amid a hodgepodge of irrelevant items, as the human visual system does.

But how does the human brain control visual attention? This is where the neuroscientists in the project have provided essential insights.

"If you want to focus on something, you could amplify just the signal, or you can make everything else 'chatter' so the signal is the only voice you

can listen to," says Narayanan. "The brain does it both ways. It amplifies this portion of focus and it also actively suppresses these other things that are not of relevance."

That understanding is a profound advance made possible by the collaboration among scientists from different fields, he says. The challenge now, for him and his colleagues, is to create a machine that can do both.

Knowing what it sees
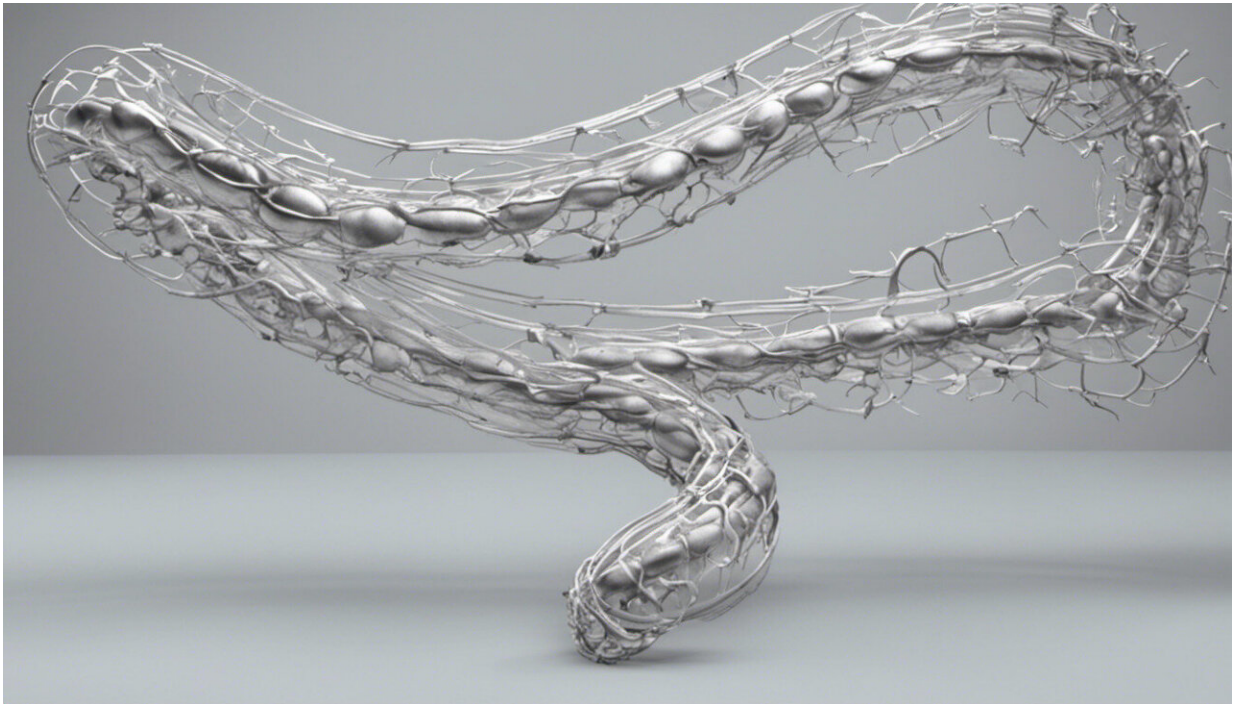
Credit: AI-generated image ([disclaimer](#))

Their system will have to be able to identify, in very specific terms, those objects it recognizes as being important. When the task at hand is

grocery shopping, an obvious way to do that is to use barcodes. The technology for reading them is already well-established, and shopper-assistance devices using it are already being tried.

But that approach is far from perfect. Michelle McManus has little good to say about barcode-based recognition. The scanners work, she says, "but you have to find the barcode!" Every shopper, sighted or not, has probably had the experience of waiting while a cashier struggles to find the barcode on a package and get the scanner to read it. A visually-impaired shopper carrying a scanner would have to take an item from the shelf and keep turning it around until the scanner finds and reads the barcode.

"If the box you show it is not the right thing, you have to try another, and keep trying until you get the right one," says McManus. Multiply the frustration of that process by however many items you're shopping for, and a simple trip to the store becomes a maddening ordeal.

In her view, a better solution is what the Third Eye team is working on—a device that can actually read the labels using recognition skills such as reading and interpreting text and identifying logos and images.

Barcodes can, however, be useful in a supporting role. Jake Weidman, a graduate student in information systems and technology, says the team incorporated barcode recognition into its Third Eye prototype as an optional back-up to give shoppers a way to make sure they had the right item. In their first run-through with the system, he says, visually-impaired shoppers attempted to verify items via barcode about half the time.

Narayanan says that eventually, the Third Eye system will be so good at recognizing products that shoppers will be able to fine-tune the degree of match between an object it sees on the shelf and an object in the

system's memory. With a low degree of match, Third Eye might consider Corn Flakes and Sugar Frosted Flakes similar enough to be the same; with greater stringency, the system would not judge them to match, or might offer them as a potential match the shopper might want to consider.

As of December 2014, the Third Eye: VI system could recognize 87 grocery products, and it recognizes them very precisely. Precision is necessary if the system is to be useful, says Narayanan; most shoppers have strong preferences as to brand and variety.

"If it just says 'cereal' or 'dairy,' it's not going to help anyone," he says. "If you want tomato sauce, we need to know if it's Prego tomato sauce. Is it organic Prego tomato sauce? That's the fine level of detail we need, and that's part of the challenge we face."

First, do no harm

Devising a system that can recognize a useful number of objects within a cluttered visual field is only half the problem. The other half is making sure the system actually helps the people it is meant to help.

For Jack Carroll, Distinguished Professor of Information Systems and Technology, that means asking prospective users about their experience of shopping, and taking their answers seriously.

"We're studying shopping with visually-impaired people: how they organize the task and how they think about it," he says. "What's difficult about it, what's rewarding about it, what's meaningful about it? Because what you don't want to do in supporting an activity technologically is make it less rewarding, less meaningful, or more challenging."

He and graduate students Jake Weidman and Sooyeon Lee have been

working with the Sight Loss Support Group of Central Pennsylvania, the local chapter of the National Federation of the Blind, and visually-impaired high-school students who came to campus last year for a three-week crash course in independent living. They were pleased to find out that grocery shopping was an excellent choice for the Third Eye's first application.

"It really is a key activity for visually-impaired people," says Carroll. "It's a kind of validation that they are like us, and that they can go into the stores, which are built for us, not for them, and they can cope." More than that, he says, they enjoy it. "Even the visually-impaired kids we talked to said shopping was right at the top of their list of things they like to do and value being able to do."

Browse, or zero in?

One thing the visually-impaired students helped them with was answering the basic question: What's the best way for the Third Eye system guide a visually impaired shopper toward items she might want?

Former Penn State research associate Kevin Irick, a member of the Third Eye:VI design team, uses the haptic glove to scan groceries in the team's lab. Credit: Patrick Mansell

To answer that question, Weidman and Carroll used a "Wizard of Oz" prototype. They had students wear a chest-mounted iPad that would see grocery items on the shelves and transmit the images to Weidman in a nearby control room. Based on what he saw through the iPad's camera, Weidman would give verbal instructions to the student.

"If you remember, in the movie there's a little guy behind a curtain who's creating the appearance of a wizard, but there isn't any wizard, there's just a guy behind the curtain," says Carroll. "In a Wizard of Oz prototype, there is no system. There's the appearance of a system"—in

this case, Weidman giving the shopper verbal feedback as the Third Eye device might do. By following scripts that offered different kinds of information and different wording, the researchers were able to evaluate what kinds of guidance the students preferred.

"We looked at whether it's more desirable to give shoppers more directive feedback with respect to what the items were, where the items were, and where they should be directing their attention, or whether it would be good to give them more open-ended feedback," says Carroll. "There was a clear preference for the browsing dialog."

He says the Third Eye system could eventually do both, giving the shopper general information about what it sees while browsing and then, at the shopper's request, providing guidance to pick up a wanted item.

Guided by touch

Verbal feedback is a good way to go in browsing mode, but for selecting specific products it seems clunky—"Move your hand two inches to the right and six inches forward." So the Third Eye team developed a more subtle, elegant, and private form of direction: a haptic glove that guides the user's hand toward the chosen item by vibrating at different strengths and in different positions on the hand.

So far, people who have tried the glove have learned quickly—"within five minutes," said one—to respond smoothly and accurately to the vibrations.

The glove also gave the team a better place to put the system's camera. Instead of being strapped to the shopper's chest, the small webcam is attached to the glove at the base of the palm. When the hand reaches out, the "eye" sees what the hand is pointing towards and the system gets a continuous view of what's on the shelves near the shopper.

Graduate students Jake Weidman (foreground) and Josh Snyder demostrate the "Wizard of Oz" experiment they used to test what kind of verbal feedback was most helpful to visually-impaired shoppers. A camera on the haptic glove Snyder is wearing relays images of grocery products to Weidman in a nearby control booth. Credit: Patrick Mansell

Carroll, his students, and the glove design team will soon launch a new trial with visually-impaired volunteers to further refine the system. For instance, what's the best way to guide shoppers looking for stacked items such as cans of soup? The shopper needs to pick up the can on top; if he grabs a can in the middle, the stack will come tumbling down.

Looking ahead

In related research, graduate student Sooyeon Lee is working with other volunteers to learn more about how visually-impaired people handle groceries at home: where they store and how they organize goods, how they know when supplies are running low, and how they maintain a list of items to buy on their next trip to the store.

Narayanan is already thinking about how the Third Eye-VI device could be made available to the people who could benefit from it. Businesses might buy one or two of the gloves for their visually-impaired customers to use, just as many stores now have motorized scooter-carts for their customers who have trouble walking. They could keep the devices updated with sale prices and locations of items. When a shopper scans in a list of items to be bought that day, the system might even suggest an alternative if a different brand of a list item is available for less money.

When the Third Eye's vision system recognizes an item the shopper wants, it makes different areas of the glove vibrate to direct the shopper's hand toward that item. Credit: Patrick Mansell

McManus says that from the point of view of the visually-impaired community, the research team is going about the project in exactly the right way.

"Part of the reason we like the Third Eye project is because they get in touch with blind and visually-impaired people before it's developed," she says. "Instead of coming to us after it's developed, and then going, 'Oh, wait a minute, this may not work correctly.' "

Narayanan agrees that listening to the potential users of their device has been a crucial aspect of the program, both to set goals and to keep the project in perspective.

"I do not want to over-promise," he says. "There are certain things that they are extremely good at managing themselves. We do not need to assist them in certain environments. We are just trying to make sure we are sensitive to their needs."

Provided by Pennsylvania State University