Features Editor:
Rebecca L. Deuel
rdeuel@computer.org

ALSO FEATURED THIS ISSUE

Trivial Pursuit? Game's Research Value Challenged

Silencing Voice Recognition Technology

Laurianne McLaughlin

Imagine voice recognition software that knows what you're going to say even before you say it. It sounds futuristic, but NASA scientists are having early success developing prototype technology that recognizes and processes silent speech for use with a PC. Instead of grabbing words from spoken sounds, the technology grabs words from nerve signals the human brain sends to the throat and tongue.

The NASA technology includes in-house-developed software and small stick-on sensors that the person doing the "speaking" wears on the chin and throat (see Figure 1). But this system doesn't require your mouth or lips to move. For instance, as you read to yourself silently, without facial expression, the software works in concert with the sensors to read the signals and "hear" the word.

This prototype software and its companion sensor technology could benefit people other than astronauts: NASA envisions possible applications for those who are physically challenged or recovering from illnesses or accidents.

"The potential of helping folks, not only in NASA but also people who are injured or need help, is a big plus," says Chuck Jorgensen, the scientist leading the effort at the NASA Ames Research Center. Jorgensen, whose background is in mathematical psychology, says the project fascinates him because the team is learning much about the human nervous system's functionality and how the brain operates and computes. But the technology's potential to help others inspires him and his team to keep trying to solve the project's considerable technical challenges.

Quiet idea, big goals

Jorgensen says NASA had good reasons for starting its silent speech program, which began around 2000 as an extension of a 1998 project. For example, astronauts in very low pressures or microgravity have a hard time communicating. Also, outside a space station or in other outdoor settings, abundant background noise makes traditional voice recognition technology impractical.

The project also aims to save weight and design costs on PC systems and wearable computers for astronauts, Jorgensen says. "The future of technology may be tiny, so you don't want a huge keyboard. This means extra weight in space, and design constraints."

Finally, if an astronaut is injured or has muscle weakness, he or she might not have the physical strength to speak or turn knobs. Silent speech commands could be the answer to all of these concerns.

Although the project is still young, the NASA team has proven that recognition software can indeed capture and process silent, or "subvocal," speech.

"We're measuring the electronic signal [EMG] being sent by the brain to the muscles of the tongue and throat," Jorgensen says. "When the brain has made a decision, we're intercepting the signal."

How it works

The NASA team uses silver chloride, stick-on sensors to intercept those brain signals. The signals must be amplified, using a standard signal amplifier, to 20 to 500 hertz. (Brain signals are at 20 hertz or lower.) Currently, the system works using two sensors: one under a person's chin and one on the throat.

The two signals are differentially measured, subtracting one from the other, and rectified. The system then applies a dual-tree wavelet transform to the rectified wave form. This generates a matrix you can plot, a matrix of wavelet coefficients. If you colored it, you'd get a picture.

"We operate on that picture to create a series of pattern examples," Jorgensen says. "The pattern is used as a training signal to a neural network. The neural network learns a functional mapping between the pattern features and a category like a word or an action."

In other words, the team is taking samples of subvocally spoken words and mapping them to specific results.

"That's one of the interesting things about this technology," he says. "Suppose you wanted to make raising your eyebrow mean 'yes' and sticking out your tongue mean 'no.' We could map that."

Not surprisingly, Jorgensen has already received a great deal of interest in the possibility of using these methodologies for injured or physically challenged people.

1094-7167/04/$20.00 © 2004 IEEE
Published by the IEEE Computer Society

But NASA must refine the technology considerably before it would be practical for those applications.

"What we've achieved so far is roughly comparable to the early stage of voice recognition," Jorgensen says. "We're at the stage of individual speakers and individual words."

Signal processing riddles

This unusual project fits in the AI domain, given the modern view of AI that includes work with neural networks and that the NASA scientists are working to understand a cognitive process.

"We're asking, 'What's the language of the neurological system?' This is clearly a human-machine interaction technology but at a very synergistic level," Jorgensen says.

The signal-processing issues present some of the NASA team's thorniest challenges. For starters, when the brain sends signals to the tongue and throat to make words, thousands of neurons are firing. "It's a very tough signal-processing problem since we are measuring the signals on the surface of the skin," Jorgensen says.

It's a different set of problems than those faced by developers of commercially available voice recognition software. Current voice recognition software uses hidden models involving acoustics and timing to address audio problems—obstacles the NASA system doesn't have to tackle. But underlying both types of software are riddles in signal processing and pattern recognition.

With speech recognition software, the program converts words to vowels and consonants and applies techniques such as using context and frequency of sounds in English words to guess the word being spoken. Sounds feed the software's front end.

But with the NASA software's front end extracting the information directly from neurological signals, some unique issues arise. For instance, when you say a word like *Paul* with a *p* sound, you're building up pressure in the lungs to make the sound. The release of air in the lungs creates the sound. So the *p* sound doesn't show up well when technology is capturing signals from the brain, Jorgensen says.

S noises and nasal sounds (such as you might hear in spoken French) also prove tough to register for the NASA software. This can create missed words and dropouts in patterns.

"The jury's still out if we can get the recognition high enough," Jorgensen says.

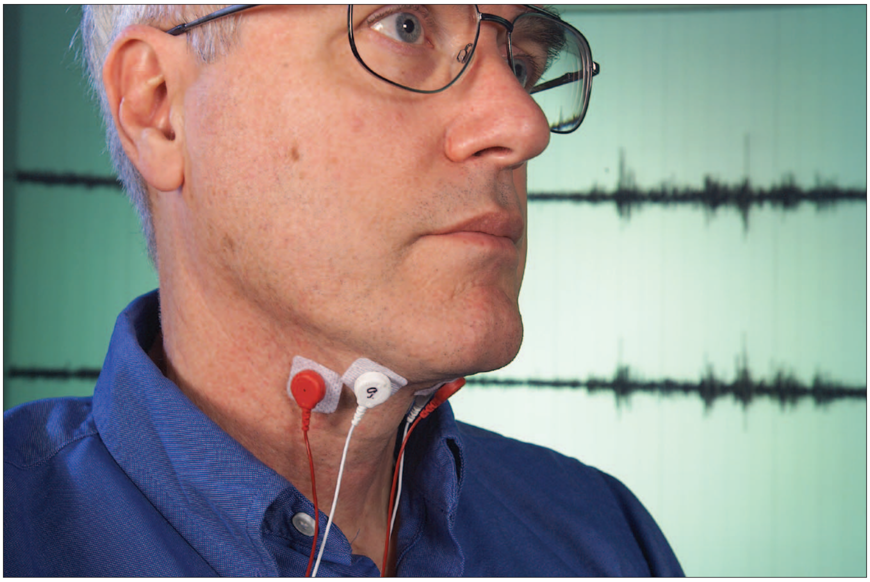

Figure 1. Stick-on sensors read silent signals that can then be processed by recognition software. (photo courtesy of NASA Ames Research Center, Dominic Hart)

A highly accurate word-recognition rate for the NASA software will depend on the signals' richness—as well as what types of sensors are used.

Smart sensors

Along with the software effort, the NASA team devotes much of its time to crafting new sensors to make the system more accurate and practical.

"We don't want everyone running around with sticky sensors and wires running down their throats," Jorgensen says.

The team has recently begun work with noncontact sensors that measure tiny changes in the electric field emanating from the human body. These sensors could be placed in someone's clothing, for example, Jorgensen says.

The sensors must be quite sophisticated because they're tuned to be measuring muscles at specific depths where the brain signals arrive. The fine-tuning's success will control the detail level of the information collected.

But Jorgensen's team (funded by the Computing, Information, and Communications Technology Program, part of NASA's Office of Exploration Systems) has already come a long way with sensor technology in this project. Originally, the group worked on a more industrial-level application, with sensors the size of forklifts, Jorgensen says. Then, within a year and a half, the NASA team reduced the sensors to the size of a beer can, then to a one-centimeter-square box, and then a half-dollar. Today, the sensors are the size of a dime, and shrinking.

Next steps

While the NASA software runs on desktop Unix, Linux, and Windows machines, the team aims for eventual use on wearable devices.

To prove the system's worth, the NASA team hopes to demonstrate real-time Web browsing as its next big step. In the future, the NASA scientists hope to delve deeper into biometrics and explore possibilities such as enabling private communication during a group conference call.

The NASA software can't replace today's voice recognition software for typical dictation to a PC, says William Meisel, a longtime speech technology veteran who runs consulting firm TMA Associates (www.tmaa.com). But perhaps it could find a home in noisy environments such as airplane cockpits, where voice recognition software doesn't work well, or conversely, in cubicle-heavy office environments where quiet is prized, Meisel says. "There are certainly specialized applications for this," Meisel says.

NASA believes that, taken far enough, the technology could allow people to "talk" on a cell phone without talking out loud, perhaps for security reasons

Intelligent IEEE Systems

IEEE Computer Society

Publications Office

10662 Los Vaqueros Circle, PO Box 3014
Los Alamitos, CA 90720-1314

STAFF

Lead Editor
Dennis Taylor
dtaylor@computer.org

Group Managing Editor
Crystal R. Shif
cshif@computer.org

Senior Editors
Dale Strok and Shani Murray

Staff Editor
Rita Scanlan

Assistant Editor
Rebecca Deuel

Magazine Assistant
Hilda Hosillos

Production Editor
Monette Velasco

Contributing Editors
Thomas Centrella, Kirk Kroeker, Anne Lear, and Joan Taylor

Design Director
Toni Van Buskirk

Layout/Technical Illustrations
Carmen Flores-Garvey and Alex Torres

Publisher
Angela Burgess

Assistant Publisher
Dick Price

Membership/Circulation Marketing Manager
Georgann Carter

Business Development Manager
Sandra Brown

Senior Production Coordinator
Marian Anderson

Submissions: For detailed instructions and formatting, see the author guidelines at www.computer.org/intelligent/author.htm or log onto *IEEE Intelligent Systems*' author center at Manuscript Central (www.computer.org/mc/intelligent/author.htm). Visit www.computer.org/intelligent for editorial guidelines.

or for politeness on a commuter train. Or, this technology might prove useful for divers communicating under water, Jorgensen says.

Ultimately, other software makers might be able to capitalize on what Jorgensen's team develops and learns, for example, regarding pattern recognition. It's certainly possible that voice recognition software could benefit. Down the road, the NASA team could try to take information gleaned from neurological signals and connect back to a voice recognition program's engine, Jorgensen says.

Trivial Pursuit? Game's Research Value Challenged

Benjamin Alfonsi

A game that hopes to prove humans are more intelligent than computers when it comes to strategy games has yet to take the world by storm. However, it has sparked a debate over whether it has any real research value.

The game's inventor, Omar Syed, a computer scientist and former NASA engineer, contends that advancements in his chess-like computer game, Arimaa (http://arimaa.com), could lead to a major breakthrough in AI.

Not everyone agrees. "I don't see anything that's particularly interesting about the game from a research point of view," says Jonathan Schaeffer, a computer science professor at the University of Alberta. "The motivation here is not research, it's the idea of building a game that is computer resistant."

$10,000 incentive

Still, Syed put his money where his mouth is, offering a US$10,000 reward for the first program capable of defeating the best human player by the year 2020. The game debuted in 2002, and, so far, his challenge stands.

Syed hopes the reward will generate interest in studying the game from a scientific perspective. "It will probably take some time before people start taking the game seriously," Syed says, "but it's starting to happen."

Although Syed based Arimaa on chess principles and components, "it's very different from chess, and it's more difficult," says David Fotland, president of Smart Games Software (www.smart-games.com). Fotland's BOMB program, perhaps the most capable Arimaa program available, fell 8-0 to a human player at the First Annual Arimaa Competition in February 2004.

Syed designed the game to be difficult for automated players. And, for the time being, humans indeed appear invincible.

Programming challenges

What is it about Arimaa that baffles computers? And what, if anything, does the game have to say about AI's capabilities and limitations?

"It boils down to how to deal with the enormous number of possibilities," Syed says. Whereas chess has about 30 possible outcomes in any given turn, Arimaa has between five and 50,000. In addition, players can make up to four moves each turn. So, the number of possibilities rises exponentially, quickly reaching into the hundreds of millions.

In essence, the issue is speed. With Arimaa, today's machines are simply too slow to compete against the human brain. "As computers get faster, [Arimaa] programs will work better," Fotland predicts.

However, Syed says that speed might not be the primary issue. He says the requisite hardware capabilities might already exist, but software applications aren't using them to their optimal effectiveness. "The answer may lie in taking an automated approach, instead of the more commonly used manual approach, to develop an evaluation function that's different from anything we've seen before," he says.

"Arimaa is not a very tactical game. With four steps per move, it is easy for an opponent to escape from a tactical situation," says Jeroen Donkers, a research scientist at the International Computer Games Association (www.icga.org). "Strategic situations ask for a large search depth if they are to be understood by mere game-tree search. Since a large search depth does not appear to be reachable anytime soon, developers are forced to look for these game-specific additions to the standard techniques."

Research potential

Even if a computer program were to defeat a human player in the next human-versus-machine Arimaa showdown, some researchers question if such a victory would qualify as progress in AI research.

"Solving the problem is not a guarantee that you'll have developed good research," Schaeffer says.

Fotland thinks it's worth looking into. He'll likely continue to fine-tune his program and perhaps take another crack at the $10,000 in the next Arimaa competition.

Regarding the 2020 deadline, Fotland says, "If Arimaa becomes a target for AI research, it probably won't even take that long."

Donkers agrees. "In the case of Arimaa, the developing community is closely connected to the population of human players," he says. "In the game room, human players not only challenge each other but also the bots that are available. This means that computers learn from humans, but humans also learn from computers. In this way, a true coevolution takes place in which both parties grow stronger."

"The Arimaa challenge is like the World Computer Chess Championship and the Computer Olympiads," Donkers says. "They drive the scientific development of games research in AI." ■