

Chuck Jorgensen, Chief Scientist, Ames Research Center, Moffett Field, CA

Chuck Jorgensen, Chief Scientist for the Neuro Engineering Lab at NASA Ames Research Center, in Moffett Field, CA, currently works with bioelectrical interfacing and the detection of human emotion. His research in subvocal speech was a 2006 finalist for the Saatchi & Saatchi international prize for world changing ideas.

NASA Tech Briefs: What are some of the applications for bioelectrical interfacing?

Chuck Jorgensen: If you put someone in a constrained suit, like a space suit or a firefighter or hazmat suit, the pressurization that's occurring from the breathing apparatus, as well as the limitations on finger movement in a pressurized suit, make doing tasks like typing or small joystick control very difficult to do, or actually dealing with, say, an external robotics device that you might want to control with this system.

We began to ask several questions: Could we intercept the neurological signals prior to actually initiating a full movement and utilize those signals in ways that would send command to devices? The first work that we did, and this was probably about 6 years ago, was to take a look at the electromyographical signals, the surface measurement of muscle innervation that's occurring down an arm like when you clench your fist, the electrical signals cause the muscles to contract; there's basically an electrical signal activity that can be picked up with external sensors, and the magnitude of those and the timing of those and how they behave can be intercepted, recorded, and turned into patterns that can be sent to a machine.

What we looked at first was: Could we intercept those neural commands without requiring something like a pilot's joystick for an airplane? The general idea would be you reach into the air, you grab an imaginary joystick, and you would fly a plane simply by clenching your fist and moving your wrist in different directions, as though you were magically touching real hardware. The answer to that was yes, and we demonstrated it to Administrator Bolden a number of years ago by landing a full Class-4 simulator at San Francisco's airport by actually reaching in the air and flying the plane.

The next question that arose then was: If we could handle those fairly coarse muscle commands for something like grabbing a joystick, could we move it further? So the question then became: Can we intercept these electromyographic signals and type without a keyboard? And the answer to that was also yes. We demonstrated that we could use a numeric keyboard by picking up the commands of individual fingers -- picking the information up actually off the outside of the arm, the electromyographic signals on the arm, before they got to the hand. That was important because in certain kinds of tasks you might want to have gloves on, or the hand might impact a surface built for an astronaut or something, say, in a combat situation where the hand would take impact. So we wanted to pick up the signals before they got to the hand.

That finally led to subvocal speech. If we can get signals that tiny on the arm, what about the tiny signals that are sent to the tongue and the larynx and the voice box? Can we pick those up? The implication being that we might be able to understand what somebody is going to say even if they didn't say it out loud. And the answer again to that was yes.

We started developing a technology that let us have a person do the equivalent of whisper, or actually only move the mouth or simulate the articulation of words without making any audible sound at all, and pick that speech up. We demonstrated it for a number of situations. We demonstrated it in a pressurized firesuit. We demonstrated it in pressurized underwater diving equipment. These are both environments that would be analogous to a space suit where you have a lot of aspirator hits, or you have a background noise.

One of the big interests in subvocal speech was not only the ability to communicate silently. But also to communicate in the presence of extreme noise. An example of that would be someone at an

airport near a large jet engine, where normally you wouldn't be able to talk on a cell phone or communicator. You'd pick it up either auditorily as with a traditional microphone, but if that was overwhelmed with a sudden noise, you'd be able to pick it up through the actuation of the subvocal signals.

NTB: How limited is the vocabulary? What are the limitations of what you can communicate?

Jorgensen: The original work that we were doing had a fairly limited vocabulary because the types of information you could extract from surface signals without any kind of invasive behavior had us initially starting with very small numbers of words: Ten or fifteen words, for example, for things like police-ten codes. Later, we began to take a look at not just recognizing whole words, which is the way we originally started out: left, right, forward, backward, words that would control a robot flat-form, for example. We began to wonder: Can we pick up the vowels and the consonants, the building blocks of many words?

So there was so preliminary work done on that, and the answer was: Yes, we can pick up some of those vowels and consonants, but not all of them, because not everything that you're doing with the muscles reflects what goes on with speech. An example of that would be what they call applosives, which are the popping type of sounds that you make by closing your lips and pressurizing your mouth (Peter, Paul, Pickled Peppers, etc.). Those types of applosive noises are not represented. We did some work also at Carnegie Mellon connecting it to a classical speech recognition engine, except the front end of it was now a subvocal pickup. I believe that work got up into the 100s to possibly 1000-2000 word capability. That was probably the most advanced work using that specific approach to subvocal speech.

NTB: Where are we at now? Is it in use currently?

Jorgensen: The NASA budget for that work was terminated, partly to do with a termination of a broader program for the Extension of the Human Senses. The idea has been picked up worldwide, and there's a very large group in Germany working on it now, and there were a number of worldwide activities. I'm still getting calls from different people around the world that are pursuing it in their laboratory. Our ultimate goal on this, and I still think that there's work that can be done, was to develop a silent cell phone, so that we would be capable of communicating either auditorily or silently on a cell phone using the same type of technology.

What does it look like, and is it a user-friendly technology?

Jorgensen: It's mixed. It's easier to implement with the coarser muscle movements like with, for example, the control stick area of that technology. It's very straightforward. That can be a sleeve that is slid over your arm. Something like subvocal speech requires picking up signals on different areas around the mouth and the larynx. And although I imagine that a qualified Italian designer could make it look cool, the reality of it is that you have to still put in place sensors in different areas of the face to pick it up.

We were originally doing our work with the classical type of wet electrode sensors that you would see if you want to have an electro-cardiogram in a doctor's office. They're bulky. They're patchy. We later did work on dry electrodes, which didn't require that moisture, and the most advanced work currently out there that we had also initiated was capacitive sensors, which picked up the tiny electromagnetic fields without requiring direct contact with the skin. These sensors were brought down to the level of about the size of a dime, and they've continued to shrink since then. That was an

important part of the puzzle. We needed to both have the sensor technology to collect the signals in a non-obtrusive way and the processing algorithms to do something with it. We focused more on the processing algorithms. The Department of Defense has advanced the sensor side of it quite heavily. They, in fact, have entire helmets that have been populated with microsensors. The components are there, but so far it wouldn't be a "drop it on." There would have to be individual training and customization.

NTB: What were your biggest technical challenges when you were designing this type of sensor technology?

Jorgensen: The sensor technology itself was not designed at NASA. We subcontracted it. It was based on an earlier technology that initially was developed by IBM called SQUID (Superconducting Quantum Interference Device). That patent was picked up by a company in southern California, Quasar Corp., that solved a number of processes that IBM was not able to solve. They've advanced that technology substantially, as well as several other people that have begun to do the same thing with nanosensors in gaming systems. So you'll see a lot of the children's gaming systems are beginning to get pretty sophisticated in terms of what they can pick up, in terms of the same kinds of signals.

NTB: What is your day-to-day work? What are you working on currently?

Jorgensen: I'm a chief scientist at NASA Ames, and I started what is now referred to as the Neuroengineering Laboratory. My current projects are actually focused in a slightly different area. They're taking a look at the detection of human emotions. We're looking at a number of ways to extract the human emotional responses from various characteristics of the speech signal, particularly the characteristics called prosody.

We've been looking at the capability for example of using prosody as a way of detecting fatigue in pilot communications or air traffic controller communications, also the detection of emotional states (fear, anger, happiness) by analyzing typical microphone acoustic signals and determining what the emotional state of the individual is. We've also been looking at the automation of various systems that are looking at the overall human behavior: things like pupil dilation, for example, eye tracking, other areas that all reflect emotional states.

NTB: Pupils and the eyes?

Jorgensen: There's a large interest now in the commercial community as well as developing interest in NASA in determining what people's emotional reactions are from a commercial standpoint. For example, advertisers are very interested if you're on the phone. They want to determine if somebody's getting unhappy with their service or whether they're reacting positively to the pitch that they might be getting over the Internet, or to some kind of vocal communication. They want that kind of feedback that sometimes we're missing in emails, where somebody has to put the little smiling icon in the email. They'd rather know automatically what somebody is really feeling when they're saying these things. Those human communication aspects, which are sort of the focus that I'm really most involved in now, are broadcast on many channels. Those channels are things like your facial expression, the timbre of your voice, the dilation of your pupils, the rate of movement of the eyes, and the rate at which the body position changes in time.

NTB: How do you respond to someone who might be skeptical saying that a machine couldn't possibly detect emotion as well as a human?

Jorgensen: It's certainly not at that state, but the interesting thing is that what we've observed, for example, with actors attempting to show different emotions and have the machine detect it, is that the human raters of what emotion is being expressed don't agree at much higher percentage than what some of our machine evaluations are. So the humans themselves can't always agree on what emotion is being expressed. The person can say "I'm trying to express a happy emotion," but the observer can be confused, whether they're grimacing sometimes, or whether they're laughing. It's surprising. It's hard to establish what truth is when someone says how well a machine is doing and how well a person is doing.

NTB: What do you see as the easiest application for this type of technology?

Jorgensen: Within NASA, what I'm currently most interested in trying to do would be something that would help in the identification of pilot fatigue, where pilots may be reaching fatigue states and not be consciously aware of it themselves. but it begins to show up in various characteristics of their performance in their voice or in their emotional or neurological response.

NTB: What are your biggest challenges there? Does your work involve constantly calibrating that technology?

Jorgensen: Parts of this are fairly cutting-edge. There was a lot of stuff done a few years ago that's often confused with what we're looking at called voice stress analysis, which you see on television. They're saying that someone is lying because their voice stresses. The scientific studies actually have pretty well proven that that does not work. It's a bogus approach and it's based on certain assumptions of muscle tensions that really aren't there. What we're looking at is a little bit different. We're actually looking at the full spectral response of the vocal system, not a specific frequency, and across many variables.

For example, in our current work, we're looking at over 988 variables extracted from just the human voice alone, and the challenges there are formidable when determining which variables are actually the drivers for the different emotions, and how they have to be combined mathematically into different models since we have the pattern recognition questions.

We're actually looking at some other aspects of it as well, which is how to turn those patterns into visual images or having all those variables draw a picture. The picture can be recognized as the emotion anger or the emotion happiness, or something else, to actually have the data themselves tell you the state of the system. This has applications beyond just emotions, and it can be used for system health monitoring, for example.

NTB: What would you say is your favorite part of the job?

Jorgensen: Definitely trying to do something that's cutting-edge. My background is sort of a weird combination called mathematical psychology. And what's interesting to me is to try and take the soft sciences of psychology and social science and overlay a hard engineering mathematics basis for it. I find that a very fascinating combination because one side of it is rather intuitive, and the other side has to be very hardnosed and analytical. Where the two meet makes for some interesting research challenges.

NTB: Do you see this field growing, and do you see more support for these types of technologies?

Jorgensen: Actually, I do, for a variety of reasons. I'll give you a non-NASA example. The Department of Defense, for example, is quite interested now in trying to see whether or not American technology can deal with problems that are associated with social understanding of groups: How do groups form? How do you identify who the leader of a group is based on a series of sociological cues, and somehow quantify that and understand the underlying structure of the network? Similarly, there's a range of social sciences of small group theory, network theory, and how cognitive learning can be improved to rapidly absorb new languages. These are all aspects of things that are associated with teaching and learning, but to have the next step or the next advancement of these requires moving into something that really hasn't occurred before, which is this translation of the soft sciences into the hard sciences.

To listen to a downloadable podcast of the full interview, visit www.techbriefs.com/podcast. For more information, contact amber.n.philman@nasa.gov.