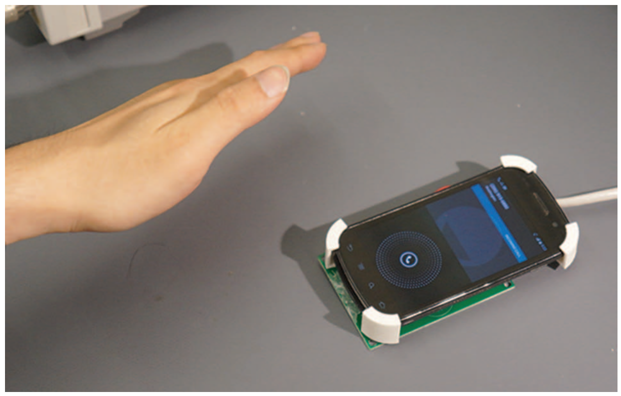
John Edwards

Human–Machine Interfaces: Methods of Control

If you are worried that artificial intelligence enabled systems are well on their way toward assuming total command of the planet, you can take some heart in the fact that there is still a great deal of important research being done in human–machine interfaces (HMIs), much of it involving signal processing. Making certain that various types of systems do precisely what their human masters demand lies at the heart of most HMI research.

The current HMI field is very competitive, and academic, government, and commercial researchers are working hard to create advanced technologies that are both useful and marketable. The major trends driving the sector include an ever-increasing demand for enhanced user efficiency; rapid growth in information technology and telecom sectors; and a continuing expansion of electronic, mobile, computer, and electromechanical applications.



[FIG1] University of Washington researchers have created a new type of low-power wireless sensing technology that promises to allow users to "train" their smartphones to recognize and respond to specific hand gestures made near the phone. (Photo courtesy of the University of Washington.)

COMMAND BY GESTURE

Smartphones have become increasingly affordable and more widely used over the past several years. Yet smartphones and their applications are difficult to control in situations where the user lacks direct access to the touchscreen, such as while driving a car, cooking a meal, or exercising. While voice recognition technology promises a partial solution to the problem, such systems are far from foolproof and particularly unreliable in noisy environments.

In an effort aimed at creating an alternate "hands off" control technology, University of Washington researchers have created a new type of low-power wireless sensing technology that promises to allow users to "train" their smartphones to recognize and respond to specific hand gestures (Figure 1). The new SideSwipe technology developed in the labs of Matt Reynolds and Shwetak Patel, both associate professors of electrical engineering and of computer science, uses the phone's own wireless transmissions to sense and recognize nearby hand gestures.

"Current smartphones use a variety of different built-in sensors, such as accelerometers and gyroscopes, that can track the motion of the phone itself," Reynolds says. "We have created an entirely new type of sensor that uses the reflection of the phone's own wireless signal to detect nearby gestures, allowing users to interact with their phones even when they are not holding the phone, looking at the display, or touching the screen."

Whenever someone uses a smartphone for voice or data communication, the device transmits radio signals on a 2G, 3G, or 4G cellular network to communicate with a nearby cellular base station. SideSwipe takes advantage of the fact that when a user's hand moves through space near the phone, the user's body reflects some of the transmitted signal back toward the device. SideSwipe uses multiple small antennas to capture changes in the reflected signal and classify the changes to detect the specific type of gesture performed. The result is that hovering, tapping, and sliding gestures can be associated with various phone commands, such as silencing a ring, changing a song, or muting the speakerphone.

"The GSM signal that we are working with was originally designed for communication, but we are analyzing the signal in a different way, from the perspective of pulling out gestures," Reynolds remarks. He adds that he and coresearcher Patel were inspired by radar technology. "In the case of radar, you have a controlled emitter of an electromagnetic wave that bounces off an aircraft or a ship or something like that and comes back," he says. "We realized that the same thing is happening all the time to the cell phone transmission."

Reynolds says that signal processing is essential to the technology. There are multiple phases of signal processing, ranging from very simple filtering to more sophisticated machine learning, he says. "We used signal smoothing, bandpass filtering to extract the frequency band that has useful gesture information in it, which tends to be a very low-frequency signal," he explains. "Then we use

Digital Object Identifier 10.1109/MSP.2015.2412128
Date of publication: 15 June 2015

a machine-learning technique called the *support vector machine* to do the classification of features into specific gestures."

When developing their prototype, the researchers added a receiver with four directionally sensitive antenna elements to the smartphone's case. "Then we looked at the signal waveforms that were coming from that receiver and figured out a way of ignoring the fact that the signal was originally designed for communication and instead looked at its envelope to see changes in reflection," Reynolds says.

[FIG2] Jan Scheuermann (right) reaches out with the thought-controlled robot arm to touch Jennifer Collinger's hand. (Photo courtesy of the University of Pittsburgh Medical Center.)

SideSwipe leverages the unmodified GSM bursts that inherently exist when someone is using a smartphone. When the user performs a particular type of hand gesture, the antennas pick up the fluctuation in their respective propagation paths. By combining the signals from four antennas, the researchers were able to identify unique patterns for different gestures.

A group of ten study participants tested SideSwipe with 14 different hand gestures, such as tapping, hovering, and sliding, at various positions and distances from the smartphone. The smartphone was calibrated to its user's hand movements prior to each test. The smartphone recognized gestures with about 87% accuracy, Reynolds says.

"We are interested in interaction in cases where you are not holding the phone," Reynolds states. "If you think about the use of the phone during the day, most of the time people have the phone in their pocket or in a handbag or, let's say, on a table." Because the SideSwipe sensor is based on low-power receivers and relatively simple signal processing when compared with something like camera video, Reynolds expects that SideSwipe will have only a minimal impact on battery life.

Reynold feels that the technology is still at a preliminary stage. "We are facing a much longer series of research leading to even more efficient ways of extracting gesture information, whether that is different signal processing strategies or new machine-learning algorithms," he says. "Currently, we are using extra receivers that are built into something like a snap-on phone case, but it is likely that the hardware could eventually be built into the phone itself—that would require the cooperation of the phone manufacturer."

DOING BY THINKING

Futurists and science fiction writers have long predicted that people one day will be able to control various types of devices by thought alone. Such technology would allow individuals to operate remote vehicles, machinery located inside mines and other dangerous places, and a variety of other simple and complex devices conveniently and across any distance. Even more importantly, a thought-driven HMI would allow people who have lost the use of their limbs to control robotic systems that provide mobility or the ability to grasp and manipulate various types of objects, ranging from eating utensils to doorknobs to light switches.

University of Pittsburgh researchers are investigating an HMI technology that promises to allow people to operate a robotic arm, capable of mimicking natural arm and hand movements, simply by thinking about whatever task that needs to be performed. Working with Jan Scheuermann, a 55-year-old Pittsburgh woman who has been paralyzed from the neck down since 2003 due to a neurodegenerative condition, the researchers have been able to increase the robotic arm's maneuverability from seven dimensions to ten dimensions over the past three years (Figure 2).

The additional dimensions result from four hand movements—finger abduction, a scoop, thumb extension, and a pinch—allowing Scheuermann to pick up, grasp, and move a range of objects much more precisely. "She can now pinch the fingers, flex them all together, spread the fingers apart, and then move the thumb independently," says Jennifer

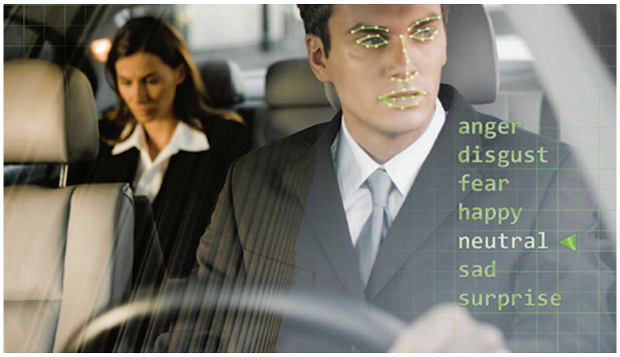
Collinger, an assistant professor in the University of Pittsburgh's Department of Physical Medicine and Rehabilitation, a project lead investigator. "This greatly increased the amount of function in the hand."

In 2012, Scheuermann underwent surgery in the regions of her brain responsible for right arm and hand movements. The operation fitted her with a pair of quarter-inch electrode grids, each containing 96 tiny contact points. After the electrode grids on Scheuermann's brain were linked to a computer, creating a brain–machine interface (BMI), their contact points could detect electrical pulses firing between the brain's neurons.

"In terms of signal processing, I would say our system is maybe a little bit simple," Collinger says. "It is more on the data analysis and computational side where it gets complex." Signal filtering is used to remove as much noise as possible. "Then we use linear regression techniques to find a relationship between the firing rates and the movement parameters we are trying to decode," Collinger explains.

Algorithms decode the firing signals and also identify the patterns associated with a particular arm movement, such as lifting the arm or twisting the wrist. "Our results show that individual motor cortical neurons encode many parameters of movement, that object interaction is an important factor when extracting these signals, and that high-dimensional operation of prosthetic devices can be achieved with simple decoding algorithms," Collinger says.

When developing the algorithms, Collinger and her coresearchers followed a unique approach based on neurobiological principles. They first optimized a population vector algorithm that encodes prosthetic movements on the basis of direction-dependent tuning of the motor cortex neuronal ensemble. The decoded movement was then sent to a shared controller that integrated the user's intent, position feedback, and various constraining task-dependent features to optimally guide the robotic arm movements.



[FIG3] Researchers in the LTS5 of the EPFL are investigating a technology that reads and identifies facial expressions indicating various moods, such as anger, disgust, fear, happiness, sadness, and surprise. (Photo courtesy of EPFL.)

In the lab, the robotic arm is set up next to Scheuermann on a stand. "We positioned it close enough to her to allow her to feed herself, to take a drink, those kinds of things," Collinger says. "I certainly think that it could be mounted to her wheelchair."

A cable connects the electrode grids on Scheuermann's brain to the robotic arm. Yet many BMI researchers foresee the day when user commands are transmitted wirelessly to robotic arms and other electromechanical systems. Steps are already being taken in that direction. Last December, Brown University researchers announced a new high data-rate, low-power wireless brain sensor. The head-mounted, 100-channel transmitter is only 5 cm in its largest dimension and weighs just 46.1 g but can transmit data at up to 200 megabits a second. The technology is currently designed to enable neuroscience research that cannot be accomplished with existing sensors that tether subjects with cabled connections.

During the next research phase, Collinger and her colleagues plan to investigate additional ways of making that arm more controllable. "We plan, for instance, to study whether the incorporation of sensory feedback, such as the touch and feel of an object, can improve neuroprosthetic control," she says.

READING FACIAL EXPRESSIONS

HMIs are advancing to the point where they can function subliminally, potentially saving users from dangerous situations of which they may not even be aware. Showing that HMIs are not only about system control, researchers in the Signal Processing 5 Laboratory (LTS5) of the Ecole Polytechnique Fédérale de Lausanne (EPFL) have created a technology that reads and identifies human facial expressions indicating various moods, such as anger, disgust, fear, happiness, sadness, and surprise (Figure 3). Such a system could prove useful in several fields, including video game development, medicine, and marketing. The EPFL researchers, however, are most interested applying the technology driver safety systems.

"Certain emotional states of the driver, such as stress, rage, or strong euphoria, affect decision making and coordination skills and may cause discomfort and loss of concentration/attention and may give way to accidents," says principal investigator Jean-Philippe Thiran, an EPFL professor. "In this project, we use computer vision and machine-learning methods to detect these emotional states from the drivers' facial expressions." The first studies involved the emotion of stress and the state of fatigue.

Thiran sees significant potential in the research, which is being sponsored by automaker PSA Peugeot Citroën and automotive equipment supplier Valéo. "Some potential applications are an alert system, for instance, in the case of detecting fatigue, or a countermeasure for the certain emotions that may endanger driving quality, such as the activation of a 'driver calming system' using sounds, lights, or odors that are personalized for the driver."

Tests carried out with a prototype proved promising. "We hope that with additional studies we will be able to define the emotional states that are dangerous for driving and also how to reduce them," Thiran says. The researchers' goals are a more comfortable driving experience and fewer accidents attributable to drivers' emotional states. "What is commonly known as 'road rage' or 'aggressive driving,' for example, is a very common cause of car accidents and it is a state that we hope to detect and take measures against," Thiran remarks.

"Signal processing is at the core of research, as in all imaging systems," Thiran says. "[The system] includes robust face detection, feature extraction, and classification," he says, noting that a couple of approaches are used to extract discriminative features and to uncover patterns of different facial expressions. "We investigate approaches based on holistic affine warping and local descriptors," Thiran says.

Holistic affine warping normalizes face images using the coordinates of the left and right eyes. "The locations of the eye centers are derived from the tracked facial landmarks," Thiran says. "After applying an affine transform, the eye centers are fixed in the canonical coordinates in the normalized image." A second approach extracts local descriptors from around the tracked facial landmarks. "It preserves the geometrical information of the facial components and does not introduce additional artifacts," he explains.

"The techniques we are using are common to most facial analysis systems: a face-tracking system that locates the drivers face from the image captured by the in-car system, a head-pose normalization algorithm that compensates for the angle of view, extraction of relevant appearance features from the tracked face, and classifiers to detect whether the expression we are interested in is present," Thiran states.

"Having in-car conditions involves many additional requirements compared to indoor systems or standard human–computer interaction (HCI) systems in which you may restrict the user to face the camera at all times," Thiran says. "As this it not a possibility during driving, we had to come up with a solution to compensate for the up-tilted angle of face-view, that results from the special camera configuration in the car." The system prototype applied a head-pose correction that used a simple three-dimensional model to project the driver's face image on a two-dimensional plane for classification. "We are now working on a facial reconstruction from multicamera system that will enable emotion detection totally regardless of where the driver is looking at," Thiran says.

HMIs ARE ADVANCING TO THE POINT WHERE THEY CAN FUNCTION SUBLIMINALLY, POTENTIALLY SAVING USERS FROM DANGEROUS SITUATIONS OF WHICH THEY MAY NOT EVEN BE AWARE.

Vehicle interior lighting conditions, which change frequently during a journey, also posed a challenge for the researchers. "To tackle that problem, we used near-infrared (NIR) cameras with a special lighting system and adequate filtering so that ambient light effects are at minimum," Thiran says. "We had to rebuild classifiers for expression using NIR images and also applied model adaptation techniques so that we can make use of existing facial expression databases of color images, which are well annotated and validated."

The researchers are now working on detecting other expressions on drivers' faces, such as distraction and lip reading for use in vocal recognition.

AUTHOR

John Edwards (jedwards@johnedwardsmedia.com) is a technology writer based in the Phoenix, Arizona, area.

[SP]

Digital Object Identifier 10.1109/MSP.2015.2434231