

Will Gesture-Recognition Technology Point the Way?

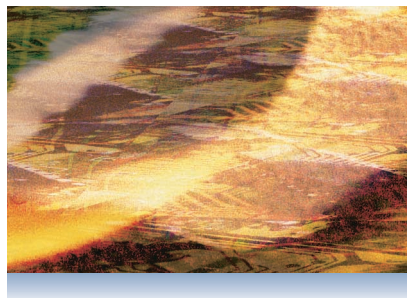
David Geer

When playing most video games, speed is of the essence. Manipulating a joy stick, mouse, or other input device slows a player's reaction time. Players would prefer to control game activities by movements or gestures.

Physically disabled users, who frequently have trouble providing the strength or precision necessary to use traditional computer input devices, would also benefit from being able to control devices and enter information via eye blinks, head motions, or other gestures.

For these and other reasons, considerable research has gone into computer-related gesture-recognition technology. Now, this research is bearing fruit as the technology increasingly appears in commercial products such as Canesta's Virtual Keyboard for PDAs; iMatte's iSkia projector-based presentation technology; and Cybernet System's GestureStorm for weather reporting, NaviGaze head- and eye-movement-based cursor and mouse interface technology, and UseYour-Head game controller.

Gesture-recognition systems identify human gestures and use them to convey information such as input data or to control devices and applications such as



computers, games, PDAs, browsers, cell phones, and MP3 audio players. For example, eye movements could initiate mouse clicks or hand gestures could manipulate computer graphics.

Researchers continue to improve gesture-recognition technology—for example, by making algorithms faster, more robust, and more accurate.

Proponents say gesture recognition has many potential new uses, such as helping surgeons perform operations and improving security, surveillance, and military applications.

However, the technology still faces major challenges. For example, gesture-recognition devices such as motion-tracking gloves are too intrusive for mainstream use. In addition, the video processing that records user movements in some gesture-recognition products is resource intensive.

“Commercially, gesture recognition must prove it can yield results that

existing peripherals can't already achieve, or users won't see the point in spending the time and money on the technology,” said Jackie Fenn, a Fellow in emerging trends and technologies for Gartner, a market research firm.

GESTURE-RECOGNITION TECHNOLOGY

In the early 1960s, users could move a light-emitting pen to control the Sketchpad computer-aided design system. Several subsequent commercial systems also worked with light-emitting pens.

Research into camera-based computer vision for gesture recognition began in earnest in the early 1990s at places such as the Massachusetts Institute of Technology Media Lab, Japan's Advanced Telecommunications Research Institute International, and the University of Zürich.

Since then, a few companies have sold gesture-recognition software. Until now, though, the technology hasn't had a significant commercial impact.

Gathering gesture data

Users create gestures by a static hand or body pose or by a physical motion—including eye blinks or head movements—in two or three dimensions. Software translates the gestures into letters or words, or simple or complex commands. The computer then acts based on the input or command.

Several image- or device-based hardware techniques gather information about gestures. Image-based techniques detect a gesture by capturing pictures of a user's motions during the course of a gesture, such as via a camera, as Figure 1 shows. The system sends these images to computer-vision software, which tracks them and identifies the gesture.

Device-based techniques use a glove, stylus, or other position tracker, whose movements send signals that the system uses to identify the gesture.

For example, instrumented gloves house sensors that relay information about the wearer's hand and finger

positions. Styli interface with display technologies to record and interpret gestures like the writing of text. Finger-based sensors detect finger positions, and some tablet PCs work with electromagnetic-resonance pens.

Position trackers also use ultrasound emissions and infrared light to identify the movements that make up a gesture. For example, changes in ultrasound waves could measure the changes in a finger's position relative to a fixed point.

Recognizing gestures

A key issue for gesture-recognition systems is interpreting which gesture a series of motions actually represents. The systems generally do this by applying statistical modeling to a set of movements.

Some systems track gesture movements through a set of critical positions. When a gesture moves through the same critical positions as does a stored gesture, the system recognizes it. Other systems track the body part being moved, compute the nature of the motion, and then determine the gesture.

These systems generally recognize and identify gestures using hidden Markov models, a statistical technique designed to cope with unknown parameters. With HMMs, the challenge is to determine the most probable hidden parameters from the observable parameters. A system can use the extracted parameters for further analysis, such as the pattern recognition required for gesture identification.

NEW DEVELOPMENTS

A significant factor making gesture recognition more practical for widespread use is that hardware and processing costs have decreased considerably over time, noted Richard Marks, Sony Consumer Entertainment's special projects manager for research and development.

Also, systems are beginning to combine image- and device-based techniques to gather more information about gestures and thereby enable more accurate recognition.

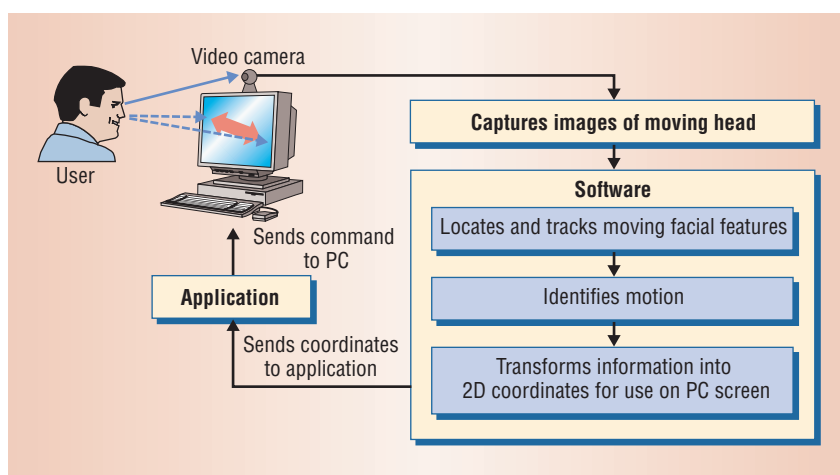


Figure 1. A PC user turns his head across a screen a set distance to issue a command to, for example, move a cursor. A gesture-recognition system uses a video camera to capture images of the head movement. The gesture-recognition software tracks the moving facial features, identifies the motion, and uses statistical modeling to determine the most likely command being issued. The command is then issued as a set of 2D coordinates that show how the cursor should be moved on the screen in response to the command. These instructions are sent to the application being used, which then communicates with the PC.

Algorithms

Matthew Turk, associate professor of computer science at the University of California, Santa Barbara, said, "Probabilistic methods are being developed to make systems more robust and more error tolerant." These methods, which are designed to cope with some degree of uncertainty, more accurately predict the likelihood that a motion is the intended gesture, despite challenges created by such factors as lighting and background.

Francis MacDougall, president of computer-vision vendor Jestertek, said that the company's GroundFX, Jestpoint, and Vivid Group divisions have used heuristics to achieve more robust, accurate, and quicker tracking of gestures. Heuristics is a branch of artificial intelligence that applies experience-derived knowledge to a problem. Systems using the approach learn from the images and motions they analyze and are thus better able to identify subsequent gestures they encounter.

Hardware

Researchers are also upgrading the sensors that relay information about a

user's movements. For example, little sensors make gesture recognition less intrusive by letting vendors put the technology into smaller wearable devices such as rings.

Jestertek is exploring using two or three cameras, rather than just one, to track user motions. Multiple cameras could let systems better analyze gestures in three dimensions and thereby more accurately identify them.

GESTURE RECOGNITION HITS THE MARKETPLACE

Cybernet and other vendors have recently released various types of gesture-recognition products.

For example, iMatte has introduced iSkia, a technology that enables presenters to interact with projectors and screens using gesture recognition. When presenters hold down buttons on a remote control, the iSkia system recognizes the movements of their extended hand and converts them into on-screen drawing or highlighting.

GestureStorm

Cybernet developed GestureStorm, based on a battlefield-command train-

Industry Trends

ing system it designed for the US military, primarily so that TV weather broadcasters can use hand gestures to illustrate their forecasts. Moving a hand one way might make images of raindrops appear, while moving a hand another way might yield an image of a tornado. Broadcasters can also use gestures for purposes such as making images zoom in or out.

The broadcaster makes gestures with a handheld remote control, and GestureStorm tracks the movements. The product uses *image differencing*, explained Chuck Cohen, Cybernet's vice president of research and development. This approach registers two images of the same location at different times and notes the areas where changes have occurred. The system then applies image processing only to the areas with changes, which enables operational efficiency.

Virtual keyboards

This year, Canesta and VKB each plan to debut similar virtual keyboards that let users control PDAs and even automotive equipment, such as navigation systems, with gestures. This is particularly helpful for small devices that have room only for tiny, hard-to-use physical keyboards.

A Canesta-enabled device uses a lens to project an image of a keyboard onto a desk or other flat surface. Users then type on the virtual keyboard.

An infrared light beam that the device directs above the projected keyboard detects the user's fingers. The device monitors how long it takes a pulse of infrared light to reflect off the user's moving fingertips and return to a sensor. The gesture-recognition software then calculates both the distance and direction of users' fingers as they move from key to key, determines where they are on the virtual keyboard, and issues the appropriate input to the device.

James Spare, Canesta's vice president of marketing, said the company sells its virtual keyboard to equipment manufacturers for use in their products.

Video game controllers

With video games, gesture recognition could either replace or supplement game controllers, such as joysticks, mice, and keyboards.

This year, for example, Cybernet plans to release UseYourHead 2, which would let game players use head motions to input directional instructions, such as moving a character or piece of equipment or changing a player's field of vision, said Cohen.

The application examines changes in the color and hue saturation of a user's face as the head moves, he explained.

Gesture recognition is starting to appear in more products.

For instance, if the head moves to the left, the technology recognizes this because the colors and hues move to the left in the image plane.

NaviGaze

With Cybernet's free, recently released NaviGaze, users can work with applications by moving cursors with head movements and clicking the mouse with eye blinks. For example, instead of double-clicking, users can double-blink while looking at an icon or file name. Cybernet created the system for disabled people who can use only their head and eyes.

As with UseYourHead 2, the system recognizes head motions by tracking changes in facial color and hue saturation. It also has been programmed to recognize the difference between an open and closed eye and can thus respond to eye blinks.

Once the system recognizes a gesture, it determines which command the motion represents and sends the information to the operating system to initiate the appropriate action.

RECOGNIZING THE HURDLES

One of gesture recognition's key challenges is that the necessary image

processing can be slow, which creates unacceptable latency for fast-moving video games and other applications.

Vendors also want to make gesture-recognition technology less intrusive, such as by eliminating the need for gloves, to encourage more widespread use, noted analyst Joe Laszlo with the Jupiter Media market research firm.

Interface-related issues

A problem the technology faces is that there isn't a common gesture language, specifying the way users should make gestures to make sure they are easily recognized, explained Sony's Marks.

If users are left to make gestures as they see fit, recognition systems will have trouble identifying the motions with the probabilistic methods they currently use. It would be easier to teach people to make a gesture a certain way than to teach a recognition system to recognize many different ways of making the same gesture.

Reliability and performance

Robustness is critical for gesture-recognition technology. Many products don't read motions accurately or otherwise don't function optimally when such factors as the background or lighting changes, said UC Santa Barbara's Turk.

In addition, they don't always properly recognize motions made against busy or otherwise confusing backgrounds, Marks noted.

Gesture-recognition technology, particularly its image processing, demands considerable resources from host systems. This can monopolize resources needed for other system functions or make gesture recognition difficult to run, particularly on PDAs and other resource-constrained devices.

This can even cause problems in larger systems, according to Turk. Therefore, he said, researchers must figure out better ways to enable the technology to work within system resources, perhaps by designing dedicated gesture-recognition chips or cards.

Gesture recognition could be used in many settings in the future. For example, Georgia Institute of Technology researchers have created the Gesture Panel system to replace traditional vehicle dashboard controls. Drivers would change, for example, the temperature or sound-system volume by maneuvering their hand in various ways over a designated area. This could increase safety by eliminating drivers' current need to take their eyes off the road to search for controls.

The Gesture Panel uses infrared LEDs to illuminate a driver's hand. A ceiling-mounted camera then records the hand's changing position, and the

software determines which gesture was made and which command to issue, said Thad Starner, professor in Georgia Tech's College of Computing.

Georgia Tech is trying to patent Gesture Panel but won't release it commercially, at least not in its current form, Starner noted.

During the next few years, according to Gartner's Fenn, gesture recognition will probably be used primarily in niche applications because making mainstream applications work with the technology will take more effort than it's worth.

When implementing gesture recognition, Fenn explained, companies will

have to be imaginative to derive the greatest benefits. Simply overlaying the technology on current applications to perform generic tasks like clicking and menu selection won't maximize its capabilities. ■

David Geer is a freelance technology writer based in Ashtabula, Ohio. Contact him at geercom@alltel.net.

Editor: Lee Garber, Computer,
l.garber@computer.org

Look to the Future

IEEE Internet Computing reports emerging tools, technologies, and applications implemented through the Internet to support a worldwide computing environment.

In 2004–2005, we'll look at

- Homeland Security
- Internet Access to Scientific Data
- Recovery-Oriented Approaches to Dependability
- Information Discovery: Needles and Haystacks
- Internet Media

... and more!

IEEE
Internet Computing

www.computer.org/internet/

