How Do You Feel? Your Computer Knows

Interfaces can sense your mood, if you let them.

NEARLY A DECADE after its retirement, the advice-spewing "Clippy" remains one of technology's most hated characters. As part of Microsoft's Office Assistant help system, the paperclip-faced avatar proposed help based on Bayesian probability algorithms: start a word-processing document with "Dear," and it offered to help you write a letter. Express exasperation, and Clippy would gleefully continue pestering you: it could not sense your mood.

Perhaps Clippy would still be with us if it had employed affective computing, a growing field that attempts to determine a user's mood or emotion through visual, auditory, physiological, and behavioral cues—cues that comprise one's "affect." An affect-enabled Clippy might see your look of disgust and make itself scarce; conversely, it might pop up sooner when you furrow your brow in confusion.

Affective computing could go well beyond desktop help systems. Researchers cite possibilities for educational, medical, and marketing applications, and have begun to commercialize their efforts. They are building systems that attempt to measure both emotions—short-term, responsive phenomena—and longer-term moods. Even as they surmount technical barriers, they will have to face ethical questions as they help machines access a formerly hidden area of your mind.

Deconstructing Moods and Emotions

Scientific interest in how mood and emotion affect interaction appears to be universal: Sun Tzu referred to "the art of studying moods" for military strategy in his sixth-century-BC classic *The Art of War*. However, it was not until the 19th century that empiricists rigorously connected emotion to its physical underpinnings. Notable books included Sir Charles Bell's *The Anatomy and Philosophy of Expression* (1824) and Charles Darwin's *The Ex-pression of the Emotions in Man and Animals* (1872) —both of which held that emotions are physiological and universal in nature. William James' seminal 1884 article "What is an Emotion?" posited that some emotions have "distinct bodily expressions"—affects—that can be categorized, measured, and analyzed.

Mid-20th-century researchers largely discarded such theories of universal, measurable emotions, in favor of those that held emotions to be learned and culturally determined. They enjoyed a revival with the work of Paul Ekman, whose 1978 publication of the "Facial Action Coding System" (FACS) provided a basis for deconstructing affect in facial expressions.

It was Rosalind Picard's 1995 paper "Affective Computing" that first fully asserted the value of affect in computing systems. An unusually philosophical paper, Picard said it was "roundly rejected" when submitted to a journal for publication. (One reviewer wrote, "this is the kind of stuff that fits in an
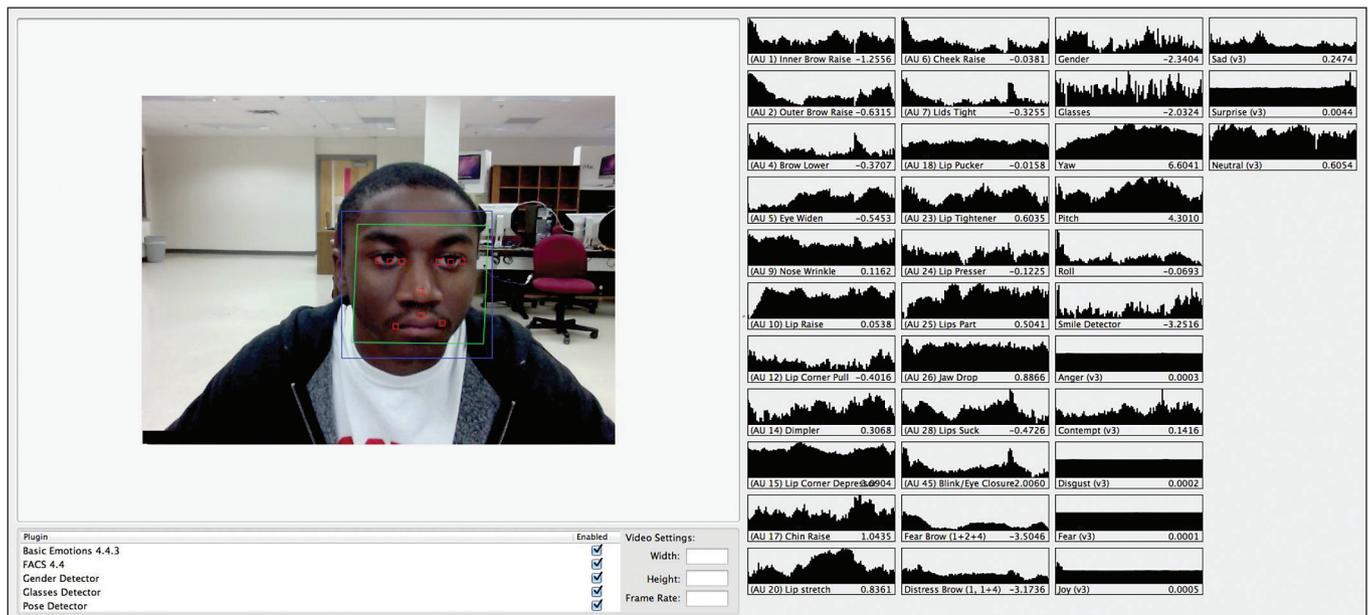
A screenshot showing video processing by the Computer Expression Recognition Toolbox (CERT).

in-flight magazine.") Picard, now director of the Affective Computing Research Group at the Massachusetts Institute of Technology (MIT), turned the paper into a book of the same name, found a publisher, and submitted the book with her tenure papers. "Back in '95, there was a lot of conjecture," she said. "We did not have real-time computer vision or vocal analysis: it was before cameras were ubiquitous. But I saw that, if computer agents could see when you are interested, or bored, or confused, that would lead to more intelligent interaction."

First, however, computers would need to learn how to "read" people, and FACS was only part of the puzzle.

Listening and Watching

Emotions spur responses throughout one's entire nervous system, not just in the face. Affective computing researchers have also studied effects on the voice and skin: each returns its own data complex, suitable for differing applications.

Aside from her FACS-based work, Picard has also trained computers to judge emotion based on non-visual signals. She became intrigued by statements from autistic people who said they felt increasing stress before a meltdown, but were frustrated by their inability to express it. The company Picard co-founded, Affectiva, developed a "Q Sensor" bracelet that measures electrodermal activity, and "Q Live" software to chart the results, thereby giving subjects a new way to express their interior lives.

One company attempting to plumb audio data is Tel Aviv, Israel-based Beyond Verbal, which has received two patents for what it calls "Emotions Analytics." Chief Science Officer Yoram Levanon believes the company's automated voice analysis can extrapolate emotional content in speech to help people make better-informed decisions. "Most decisions are made in between one- and three-tenths of a second," he said, "but cognition only begins after half a second. If I talk to you at 120 words per minute, you cannot process it cognitively—you must find another way to focus that." Beyond Verbal's products try to provide that focus by displaying text that comments on the speaker's emotions.

Facial expression recognition has captured the attention of more affective computing researchers. That is partly due to a wealth of developer toolkits that track the facial "Action Units" (AUs) described by FACS and its successors. These tools typically rely on one of two strategies: training a statistical shape model that aligns to characteristic curves of the face (such as eyes, mouth, and nose), as in an Active Appearance Model (AAM) or Constrained Local Model (CLM); or computing low-level vision features such as Gabor filters or Local Binary Patterns (LBP), to train machine-learned models of facial movements.

The Computer Expression Recognition Toolbox (CERT) was developed by the Machine Perception Laboratory at the University of California, San Diego, in collaboration with Ekman. According to Joseph Grafsgaard, a Ph.D. student in the Department of Computer Science at North Carolina State University, "An innovative aspect of CERT is that it is not only identifying facial landmarks, such as the shape of the mouth or eyes. It uses Gabor filters at multiple orientations, which are then mapped to facial action units via machine learning. This allows it to identify fine-grained facial deformations detailed in FACS, such as wrinkling of the forehead, or creasing around the mouth, eyes, or nose."

Another toolkit is produced by Emotient, which was founded by the original developers of CERT and includes Ekman on its board. According to co-founder and lead researcher Javier Movellan, the company's products are unusual in that they rely heavily on neural networks to "learn" different expressions, rather than on watching specific points on the face. He said, "People ask us, 'are you looking for wrinkles?' and we say, 'we don't know!' What we look for is a lot of data so the system can figure it out. Sometimes it takes us a while to figure out how the system is solving the problem."

So Tell Me: How *Do* You Feel?

Affective systems like these benefit greatly from human confirmation. Beyond Verbal "crowdsources" some of that research through its website by asking visitors to rate the results of their demos; Affectiva's facial action

coding experts compare their judgments to the system's. The benefit of crowdsourcing comes mostly during development, and to tune the system's algorithms. A different type of system starts with user queries, then correlates those self-stated moods with environmental measurements to show personal trends.

Emotion Sense started out as a voice-sampling application; speak into it, and it would attempt to determine your emotions. According to the project's Cecilia Mascolo, reader in Mobile Systems at the Computer Laboratory of the University of Cambridge, "After four years of that, we decided to use the human as a sensor for emotions, with something called 'experience sampling;' we ask the user's experience through questionnaires." Now available as an Android app, Emotion Sense matches mood to time of day, location, call patterns, and several other factors. So, for example, you might discover that your mood is generally better in the morning, or while in a certain location.

Mascolo says the ultimate goal is to "find the psychological markers out of all this data we are collecting, so we do not have to ask questions anymore. Maybe we can help psychologists and social scientists better understand what leads to being very happy, what leads to being very sad."

A similar project is Mood Sense (formerly MoodScope), a research collaboration between Microsoft Research Asia and Rice University. Like Emotion Sense, the prototype API queries user feelings, but correlates them only to communication behaviors (text, email, phone calls) and user interactions with applications and the Web browser. That is enough, according to Nicholas Lane, lead researcher at Microsoft Research Asia; "We find that how you use your phone, and what it means about your behavior in general, can be a strong signal of your mood or emotive state. That is the key innovation here." One benefit gained by narrowing the number of sensors monitored: Mood Sense is event-driven, and therefore consumes very little power. Says Lane, "The power savings are enormous over using the camera or audio, so if I wanted to run a psychology experiment, battery life on

"Affective computing was inspired by a desire to show more respect for people's feelings."

the subjects' phones wouldn't change from 40 hours to 10 hours."

To Fear and Love the Ones Who Know You

While smartphone power loss is an inconvenience, those in affective computing point to a greater danger to the field: privacy violations from such technologies being used in a surreptitious or unethical way. Speaking about multi-sensor tracking ("sensor fusion") in the service of affective computing, Freescale Semiconductor Executive Director of Strategy Kaivan Karimi said, "The thing that can promote [sensor fusion] is all the cool stuff that comes along with it. The thing that can kill it is if information starts leaking and getting used the wrong way because there are no boundary conditions set."

Yet researchers seem to agree that some third-party disclosure would be appropriate. Microsoft Research senior research designer Asti Roseway said, "One size does not fit all. You could take certain populations, such as kids with autism or ADHD, or relationships such as teacher to student or parent to child, where having external notification of emotions is beneficial. On the other hand, you could be in a meeting where you get a private poke if you start feeling stressed, and nobody has to know. We have to be really smart about the context, but that takes trial and error."

Affective computing is likely to face many trials. Speaking about his company's face-reading technology, Emotient's Movellan said, "Interest is coming from various places, and some of it is surprising to us. Marketing people want to know if someone *really* likes a product. Educators want to know what parts of a video lecture

people like, and where they are confused or not paying attention. The entertainment industry is also very interested; for example, to build personal robots that respond to your affect. Carmakers could put a camera in front of the driver to tell if you are looking dangerously fatigued and are about to crash. And health care providers can monitor facial expressions and alert caregivers when you have a depressed affect."

Picard acknowledged the potential for privacy abuses. "We have turned down work where people want to read emotions 'from a distance,' without them knowing," she said, "but affective computing was inspired by a desire to show more respect for people's feelings. So it would be incredibly hypocritical to use it in a way that disrespects people's feelings."

Further Reading

Picard, R.W.
Affective computing. M.I.T. Media Laboratory, Perceptual Computing Section Technical Report No. 321 (1995).

Lathia, N., Rachuri, K.K., Mascolo, C., and Roussos, G.
Open source smartphone libraries for computational social science. In 2nd ACM Workshop on Mobile Systems for Computational Social Science (Zurich, Switzerland, 2013)

Wu, T., Butko, N.J., Ruvolo, P., Whitehill, J., Bartlett, M.S., and Movellan, J.R.
Multi-layer architectures for facial action unit recognition. IEEE Transactions on Systems, Man and Cybernetics, Part B: Cybernetics, vol. 42, no. 4, pp. 1027–1038 (2012)

Littlewort, G., Whitehill, J., Wu, T., Fasel, I., Frank, M., Movellan, J., and Bartlett, M.
The computer expression recognition toolbox (CERT). In Proceedings of the 9th IEEE Conference on Automatic Face & Gesture Recognition, pp. 298-305 (Santa Barbara, California, 2011)

LiKamWa, R., Liu, Y., Lane, N. D., and Zhong, L.
MoodScope: Building a mood sensor from smartphone usage patterns. Proceedings of the ACM International Conference on, Mobile Systems, Applications, and Services (MobiSys) (Taipei, Taiwan, 2013)

Karimi, K.
The role of sensor fusion and remote emotive computing (REC) in the Internet of Things. White Paper (document SENFEIOTLFWP), Freescale Semiconductor, Inc. (June, 2013)

Tom Geller is an Oberlin, Ohio-based science, technology, and business writer.