

January 1, 2011

Computers That See You and Keep Watch Over You

By STEVE LOHR

I. Hundreds of correctional officers from prisons across America descended last spring on a shuttered penitentiary in West Virginia for annual training exercises. Some officers played the role of prisoners, acting like gang members and stirring up trouble, including a mock riot. The latest in prison gear got a workout — body armor, shields, riot helmets, smoke bombs, gas masks. And, at this year's drill, computers that could see the action.

II. Perched above the prison yard, five cameras tracked the play-acting prisoners, and artificial-intelligence software analyzed the images to recognize faces, gestures and patterns of group behavior. When two groups of inmates moved toward each other, the experimental computer system sent an alert — a text message — to a corrections officer that warned of a potential incident and gave the location.

III. The computers cannot do anything more than officers who constantly watch surveillance monitors under ideal conditions. But in practice, officers are often distracted. When shifts change, an observation that is worth passing along may be forgotten. But machines do not blink or forget. They are tireless assistants.

IV. The enthusiasm for such systems extends well beyond the nation's prisons. High-resolution, low-cost cameras are proliferating, found in products like smartphones and laptop computers. The cost of storing images is dropping, and new software algorithms for mining, matching and scrutinizing the flood of visual data are progressing swiftly.

V. A computer-vision system can watch a hospital room and remind doctors and nurses to wash their hands, or warn of restless patients who are in danger of falling out of bed. It can, through a computer-equipped mirror, read a man's face to detect his heart rate and other vital signs. It can analyze a woman's expressions as she watches a movie trailer or shops online, and help marketers tailor their offerings accordingly. Computer vision can also be used at shopping malls, schoolyards, subway platforms, office complexes and stadiums.

VI. All of which could be helpful — or alarming. "Machines will definitely be able to observe us and understand us better," said Hartmut Neven, a computer scientist and vision expert at Google. "Where that leads is uncertain." Google has been both at the forefront of the technology's development and a source of the anxiety surrounding it. Its Street View service,

which lets Internet users zoom in from above on a particular location, faced privacy complaints. Google will blur out people's homes at their request.

VII. Google has also introduced an application called Goggles, which allows people to take a picture with a smartphone and search the Internet for matching images. The company's executives decided to exclude a facial-recognition feature, which they feared might be used to find personal information on people who did not know that they were being photographed.

VIII. Despite such qualms, computer vision is moving into the mainstream. With this technological evolution, scientists predict, people will increasingly be surrounded by machines that can not only see but also reason about what they are seeing, in their own limited way. The uses, noted Frances Scott, an expert in surveillance technologies at the National Institute of Justice, the Justice Department's research agency, could allow the authorities to spot a terrorist, identify a lost child or locate an Alzheimer's patient who has wandered off.

IX. The future of law enforcement, national security and military operations will most likely rely on observant machines. A few months ago, the Defense Advanced Research Projects Agency, the Pentagon's research arm, awarded the first round of grants in a five-year research program called the Mind's Eye. Its goal is to develop machines that can recognize, analyze and communicate what they see. Mounted on small robots or drones, these smart machines could replace human scouts. "These things, in a sense, could be team members," said James Donlon, the program's manager.

X. Millions of people now use products that show the progress that has been made in computer vision. In the last two years, the major online photo-sharing services — Picasa by Google, Windows Live Photo Gallery by Microsoft, Flickr by Yahoo and iPhoto by Apple — have all started using face recognition. A user puts a name to a face, and the service finds matches in other photographs. It is a popular tool for finding and organizing pictures.

XI. Kinect, an add-on to Microsoft's Xbox 360 gaming console, is a striking advance for computer vision in the marketplace. It uses a digital camera and sensors to recognize people and gestures; it also understands voice commands. Players control the computer with waves of the hand, and then move to make their on-screen animated stand-ins — known as avatars — run, jump, swing and dance. Since Kinect was introduced in November, game reviewers have applauded, and sales are surging. To Microsoft, Kinect is not just a game, but a step toward the future of computing. "It's a world where technology more fundamentally understands you, so you don't have to understand it," said Alex Kipman, an engineer on the team that designed Kinect.

'Please Wash Your Hands'

XII. A nurse walks into a hospital room while scanning a clipboard. She greets the patient and washes her hands. She checks and records his heart rate and blood pressure, adjusts the intravenous drip, turns him over to look for bed sores, then heads for the door but does not

wash her hands again, as protocol requires. "Pardon the interruption," declares a recorded women's voice, with a slight British accent. "Please wash your hands."

XIII. Three months ago, Bassett Medical Center in Cooperstown, N.Y., began an experiment with computer vision in a single hospital room. Three small cameras, mounted inconspicuously on the ceiling, monitor movements in Room 542, in a special care unit (a notch below intensive care) where patients are treated for conditions like severe pneumonia, heart attacks and strokes. The cameras track people going in and out of the room as well as the patient's movements in bed.

XIV. The first applications of the system, designed by scientists at General Electric, are immediate reminders and alerts. Doctors and nurses are supposed to wash their hands before and after touching a patient; lapses contribute significantly to hospital-acquired infections, research shows. The camera over the bed delivers images to software that is programmed to recognize movements that indicate when a patient is in danger of falling out of bed. The system would send an alert to a nearby nurse. If the results at Bassett prove to be encouraging, more features can be added, like software that analyzes facial expressions for signs of severe pain, the onset of delirium or other hints of distress, said Kunter Akbay, a G.E. scientist.

XV. Hospitals have an incentive to adopt tools that improve patient safety. Medicare and Medicaid are adjusting reimbursement rates to penalize hospitals that do not work to prevent falls and pressure ulcers, and whose doctors and nurses do not wash their hands enough. But it is too early to say whether computer vision, like the system being tried out at Bassett, will prove to be cost-effective.

Mirror, Mirror

XVI. Daniel J. McDuff, a graduate student, stood in front of a mirror at the Massachusetts Institute of Technology's Media Lab. After 20 seconds or so, a figure — 65, the number of times his heart was beating per minute — appeared at the mirror's bottom. Behind the two-way mirror was a Web camera, which fed images of Mr. McDuff to a computer whose software could track the blood flow in his face. The software separates the video images into three channels — for the basic colors red, green and blue. Changes to the colors and to movements made by tiny contractions and expansions in blood vessels in the face are, of course, not apparent to the human eye, but the computer can see them. "Your heart-rate signal is in your face," said Ming-zher Poh, an M.I.T. graduate student. Other vital signs, including breathing rate, blood-oxygen level and blood pressure, should leave similar color and movement clues.

XVII. The pulse-measuring project, described in research published in May by Mr. Poh, Mr. McDuff and Rosalind W. Picard, a professor at the lab, is just the beginning, Mr. Poh said. Computer vision and clever software, he said, make it possible to monitor humans' vital signs at a digital glance. Daily measurements can be analyzed to reveal that, for example, a person's risk of heart trouble is rising. "This can happen, and in the future it will be in mirrors," he said.

XVIII. Faces can yield all sorts of information to watchful computers, and the M.I.T. students' adviser, Dr. Picard, is a pioneer in the field, especially in the use of computing to measure and communicate emotions. For years, she and a research scientist at the university, Rana el-Kaliouby, have applied facial-expression analysis software to help young people with autism better recognize the emotional signals from others that they have such a hard time understanding. The two women are the co-founders of Affectiva, a company in Waltham, Mass., that is beginning to market its facial-expression analysis software to manufacturers of consumer products, retailers, marketers and movie studios. Its mission is to mine consumers' emotional responses to improve the designs and marketing campaigns of products.

XIX. John Ross, chief executive of Shopper Sciences, a marketing research company that is part of the Interpublic Group, said Affectiva's technology promises to give marketers an impartial reading of the sequence of emotions that leads to a purchase, in a way that focus groups and customer surveys cannot. "You can see and analyze how people are reacting in real time, not what they are saying later, when they are often trying to be polite," he said. The technology, he added, is more scientific and less costly than having humans look at store surveillance videos, which some retailers do.

XX. The facial-analysis software, Mr. Ross said, could be used in store kiosks or with Webcams. Shopper Sciences, he said, is testing Affectiva's software with a major retailer and an online dating service, neither of which he would name. The dating service, he said, was analyzing users' expressions in search of "trigger words" in personal profiles that people found appealing or off-putting.

Watching the Watchers

XXI. Maria Sonin, 33, an office worker in Waltham, Mass., sat in front of a notebook computer looking at a movie trailer while Affectiva's software, through the PC's Webcam, calibrated her reaction. The trailer was for "Little Fockers," starring Robert De Niro and Ben Stiller, which opened just before Christmas. The software measured her reactions by tracking movements on a couple of dozen points on her face — mostly along the eyes, eyebrows, nose and the perimeter of her lips.

XXII. To the human eye, Ms. Sonin appeared to be amused. The software agreed, said Dr. Kaliouby, though it used a finer-grained analysis, like recording that her smiles were symmetrical (signaling amusement, not embarrassment) and not smirks. The software, Ms. Kaliouby said, allows for continuous, objective measurement of viewers' response to media, and in the future will do so in large numbers on the Web. Ms. Sonin, an unpaid volunteer, said later that she did not think about being recorded by the Webcam. "It wasn't as if it was a big camera in front of you," she said.

XXIII. Christopher Hamilton, a technical director of visual effects, has used specialized software to analyze facial expressions and recreate them on the screen. The films he has worked on include "King Kong," "Charlotte's Web" and "The Matrix Revolutions." Using facial-expression

analysis technology to gauge the reaction of viewers, who agree to be watched, may well become a valuable tool for movie makers, said Mr. Hamilton, who is not involved with Affectiva.

XXIV. Today, sampling audience reaction before a movie is released typically means gathering a couple of hundred people at a preview screening. The audience members then answer questions and fill out surveys. Yet viewers, marketing experts say, are often inarticulate and imprecise about their emotional reactions.

XXV. The software “makes it possible to measure audience response with a scene-by-scene granularity that the current survey-and-questionnaire approach cannot,” Mr. Hamilton said. A director, he added, could find out, for example, that although audience members liked a movie over all, they did not like two or three scenes. Or he could learn that a particular character did not inspire the intended emotional response.

XXVI. Emotion-sensing software, Mr. Hamilton said, might become part of the entertainment experience — especially as more people watch movies and programs on Internet-connected televisions, computers and portable devices. Viewers could share their emotional responses with friends using recommendation systems based on what scene — say, the protagonists’ dancing or a car chase — delivered the biggest emotional jolt.

XXVII. Affectiva, Dr. Picard said, intends to offer its technology as “opt-in only,” meaning consumers have to be notified and have to agree to be watched online or in stores. Affectiva, she added, has turned down companies, which she declined to name, that wanted to use its software without notifying customers.

Darker Possibilities

XXVIII. Dr. Picard enunciates a principled stance, but one that could become problematic in other hands. The challenge arises from the prospect of the rapid spread of less-expensive yet powerful computer-vision technologies. At work or school, the technology opens the door to a computerized supervisor that is always watching. Are you paying attention, goofing off or daydreaming? In stores and shopping malls, smart surveillance could bring behavioral tracking into the physical world.

XXIX. More subtle could be the effect of a person knowing that he is being watched — and how that awareness changes his thinking and actions. It could be beneficial: a person thinks twice and a crime goes uncommitted. But might it also lead to a society that is less spontaneous, less creative, less innovative? “With every technology, there is a dark side,” said Hany Farid, a computer scientist at Dartmouth. “Sometimes you can predict it, but often you can’t.”

XXX. A decade ago, he noted, no one predicted that cellphones and text messaging would lead to traffic accidents caused by distracted drivers. And, he said, it was difficult to foresee that the

rise of Facebook and Twitter and personal blogs would become troves of data to be collected and exploited in tracking people's online behavior.

XXXI. Often, a technology that is benign in one setting can cause harm in a different context. Google confronted that problem this year with its face-recognition software. In its Picasa photo-storing and sharing service, face recognition helps people find and organize pictures of family and friends.

XXXII. But the company took a different approach with Goggles, which lets a person snap a photograph with a smartphone, setting off an Internet search. Take a picture of the Eiffel Tower and links to Web pages with background information and articles about it appear on the phone's screen. Take a picture of a wine bottle and up come links to reviews of that vintage.

XXXIII. Google could have put face recognition into the Goggles application; indeed, many users have asked for it. But Google decided against it because smartphones can be used to take pictures of individuals without their knowledge, and a face match could retrieve all kinds of personal information — name, occupation, address, workplace. "It was just too sensitive, and we didn't want to go there," said Eric E. Schmidt, the chief executive of Google. "You want to avoid enabling stalker behavior."

<http://www.nytimes.com/2011/01/02/science/02see.html?ref=science&pagewanted=print>