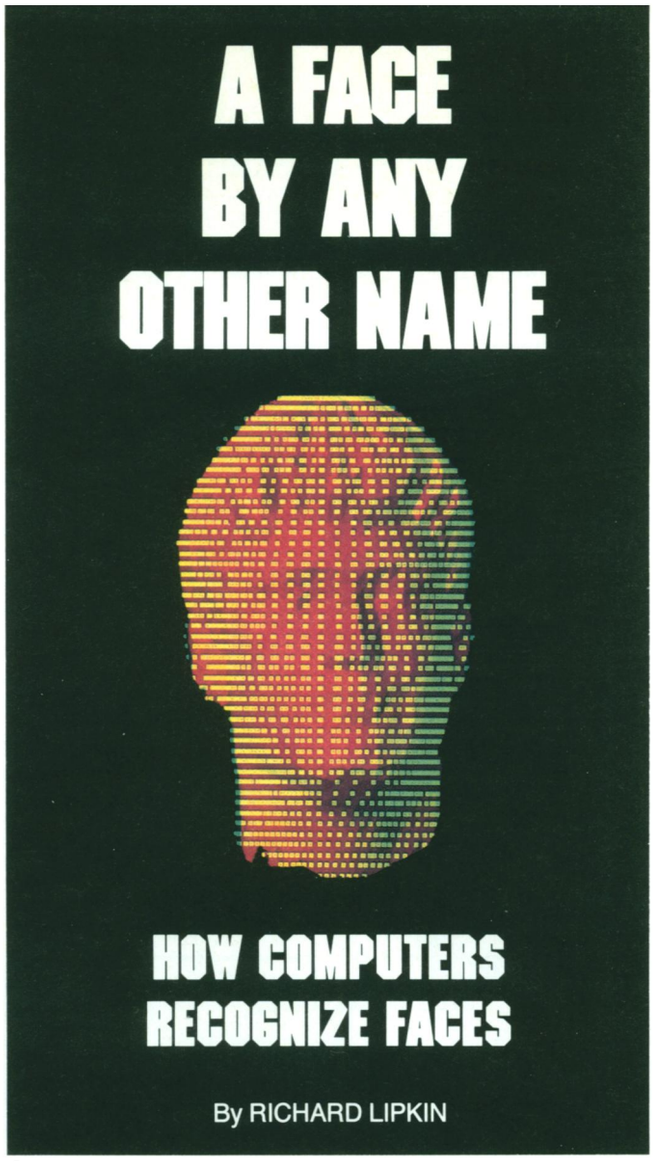
A FACE BY ANY OTHER NAME



HOW COMPUTERS RECOGNIZE FACES

By RICHARD LIPKIN

A computer recognizes and models a face in steps. It begins with a rough overview, then looks at facial details.

Images: Pentland *et al.*

Between each and every pair of ears lies a unique world: a face.

How well we know our own faces — even more the faces of those we love. But how do we distinguish one face from another, quickly finding a familiar one in a crowd? How does a mother find her daughter instantly in a nursery or a man know his brother after 20 years' passage? The human brain has mastered this feat through evolution. And yet, we can be fooled. Color someone's hair, add a beard, and even a spouse may pass as a stranger.

The mystery of face recognition poses a by no means trivial problem. Using the latest techniques, psychologists and neuroscientists are only now getting a hint of how the brain recognizes images. Incrementally, they are finding that the secret lies not in one neurological process but in many. A battery of neurons must fire before one person can recognize another. Some combination of fuzzy, holistic neuronal matchings captures the overall picture, and thousands of detail-monitoring nerve cells note a subtle skin tone or a mouth's distinct angle.

Presumably, if the human brain can recognize a face in a split second, a computer can too. The question, though, is how. What must a computer do to identify and verify a particular face? Answering this complex question will yield strong returns in better security systems and perhaps even marvelous new animation techniques.

Within the gadget-filled offices of the Media Laboratory at the Massachusetts Institute of Technology, Alex Pentland tinkers with a computer system that can single out one face among thousands with surprising accuracy. Given a database of 7,562 images (variations of the faces of 3,000 people), Pentland's system can ferret out an individual purely by decoding the person's "mug shot" — a flat, head-on snapshot.

Even when people shift position or expression, don new hairstyles or sunglasses, the program succeeds. In one test of 200 random faces, the computer topped 95 percent accuracy when asked to find the most similar face in the image base.

Pentland, a mathematically inclined computer scientist, has designed this system, called Photobook, to treat mug shots not as images per se, but as visual information. Thus the computer never really "sees" someone's face. Instead, it interprets each picture as a grid of information, as defined by a branch of mathematics called information theory. An image of a face — as of a house or a tree — imparts a unique set of information to a viewer. This computer program analyzes the content of that information and compares it with the image database.

Photobook uses a two-tiered method to recognize faces — a holistic view and feature analysis. On the holistic side, the computer gives a facial image a quick overview, ascertaining how the face fits together as a whole. Then, by treating the image as a matrix of information, it searches for eigenvectors, or mathematical patterns, characteristic of that particular face.

These eigenvectors (the German prefix "eigen" means "own" or "individual") describe precisely how that face differs from other stored facial images. "A face's key features, in terms of eigeninformation, may or may not relate to what we call facial features, like eyes, nose, lips, and hair," Pentland says. "But they are markers that denote unique characteristics of that face."

Pentland calls this approach "eigenface," based on mathematical eigenvalues in "face space," the computer's three-dimensional storage space. By working with a fixed set of facial images and treating them as one huge matrix of information, the computer finds the main features of the faces in its database and combines them to form one face.

In essence, the computer takes all the stored faces and averages them, generating a single, ghostly looking eigenface — a sort of fuzzy everyface. Photobook then ranks an individual face as a unique variation of the eigenface. Thus, each face becomes a unique version of a known type of object.

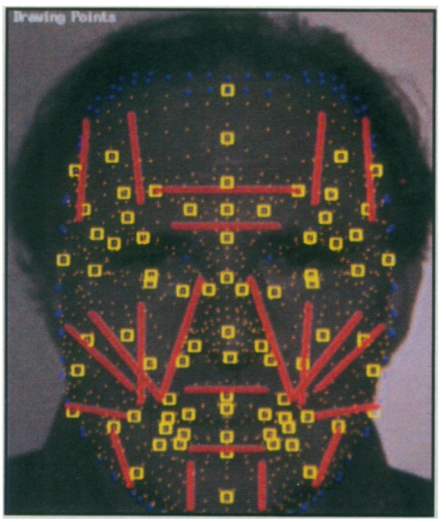
Though this analysis carries the cool edge of digital processing, it may not operate too far afield from the human brain. When a woman gazes at her lover's face, that image occurs first as mere scattered light on her retinas. Of course, random retinal pulses mean nothing until

Science Service, Inc. is collaborating with JSTOR to digitize, preserve, and extend access to *Science News*. www.jstor.org

they become linked, through some subconscious route, to the implicit notion that faces exist. Once her brain has registered that it sees a face, and not something else, it can begin to appreciate the uniqueness of that face.

Underlying this fleeting cognitive process is the tacit knowledge that human beings wear faces on the front of their heads; that faces serve well for identifying people, and that faces have features to look for — eyes, ears, nose, and mouth. Such knowledge refines the plethora of possibilities that any image presents, narrowing the field for a human brain to interpret.

A computer face-recognition system does this too.



Analyzing facial muscles.

One of the biggest problems in digital recognition is finding the face in an image, Pentland says. "Once the computer finds the face, you're halfway home." Photobook has become fairly nimble at finding faces in pictures. But then, it looks at ordinary mug shots.

What happens when a live video camera monitors a scene, looking for someone randomly entering a room? "This is a much bigger problem," says Baback Moghaddam, an MIT computer scientist. "The computer doesn't even know where to look. So we must build into it mechanisms for detecting heads and facial features, so it knows where to look. For instance, you don't generally look for a head on the floor."

Finding a face in a crowd would pose a problem for a hidden airport security system automatically scanning passers-by for known terrorists or for an office clearance system that admits only key employees. Working on an experimental system called Face-Rec, Moghaddam is tackling the problem that arises when someone randomly walks up to the video eye of a computer identification system — how to find that person's face among the visual clutter.

Once the computer finds and sizes up a face, it must determine who's there — that is, identify the face. Photobook has distinguished itself from other face-recognition programs by accurately identifying people from among a large number of images. In a test using 2,500 mug shots, Pentland and his colleagues varied the lighting, size, and head orientation of 16 male graduate students. It correctly identified 96 percent of them despite changes in lighting, 85 percent despite a turned head, and 64 percent despite adjustments in size. Overall, the test bore out the system's strength and accuracy.

Once a person has been identified, there's a final problem: verification. In this process, the computer must ask and accurately answer the question, "Are you really who you say you are?"

"Most security systems these days rely on verification, which is an inherently easier problem than identification," says Pentland. "You're dealing with a much smaller set of possibilities. The person says who [he or she is], and then the system decides if that's true."

Bank cash machines do this, asking for a personal identification number before doling out dollars. A more complex setting, such as a courtroom, may require fingerprints as an identifier. Yet fingerprints generally prove more useful for verifying than for identifying a person.

Faces also work surprisingly well as verifiers. Photobook can verify individuals in less than 10 seconds with an accuracy of nearly 97 percent, falsely rejecting someone less than 2 percent of the time and falsely verifying someone less than once in 10,000 times.

In contrast, computerized fingerprint scans showed no false verifications but falsely rejected people's identity 9 percent of the time. Verification systems using vocal patterns, handprints, or eye retinal patterns turned in slower and poorer results than the eigenface system.

To shore up the computer's accuracy, Moghaddam is adding eigenfeature templates to it — things like eigeneyes, eigennoses, and eigenmouths. These help keep the system from getting fooled when someone sports a new hairdo, grows a beard, puts on glasses, or just alters facial expression. With eigenfeatures added to eigenfaces, recognition accuracy hovers around 98 percent.

Both Photobook and Face-Rec can learn new faces on their own. When presented with a new face, the computer checks it out repeatedly in face space, then decides whether the person is unrecognizable or bears a new face. If the latter, the system enters the new face and averages it into the eigenface.

Pentland believes that with this degree of accuracy, real-world applications become feasible, as in police stations, which must maintain huge files of mug shots for quick suspect identification. Or a customs center, which must screen for outlaws passing the border. Or voter registration. The Mexican government, for example, wants to assemble a cache of 50 million facial images to stem the problem of double balloting.

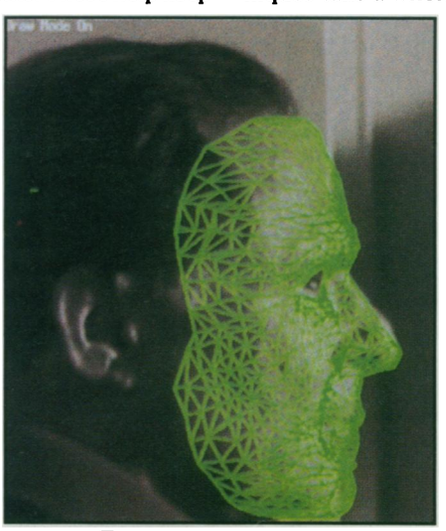
Yet to achieve such power, a computer must be able to handle many views of someone's head, such as a profile or three-quarter view. This requires facial modeling and, at some level, an understanding of facial expressions.

"When you look at a photo, you can tell if someone's happy, sad, contemptuous, or angry," says Irfan A. Essa, a research assistant at MIT. "We want to make computers that can detect known facial patterns, like a smile or frown. Or the difference between a real or fake smile."

The search for such subtlety has taken Essa into new territory, using computer vision to model and animate people's expressions. A prototype computer learns how faces express themselves by watching and imitating people. The computer sees how eyes and lips move, which features move together, and how fast each goes.

"Some muscles actuate faster, some slower," Essa says. "For an expression to look real, timing is critical."

From this interactive system, the options begin to mushroom. As the system practices imitating smiles and frowns, Essa sees the potential for realistic animation — the possibility of generating three-dimensional images with emotional depth. "We taught the system to yawn and sneeze," Essa says. "It took 2 minutes. Conventional animation techniques take a whole day."



Fitting a mesh to a face.

Thus, a real-time facial animation system, which maps live movement patterns onto a facial model that understands muscle control, has arrived under the touch of Essa and Trevor Darrell, an MIT computer scientist. While Essa concentrates on details of facial models and muscle control, Darrell forges ahead with real-time facial animation.

Taking its cues

Continued on p.220

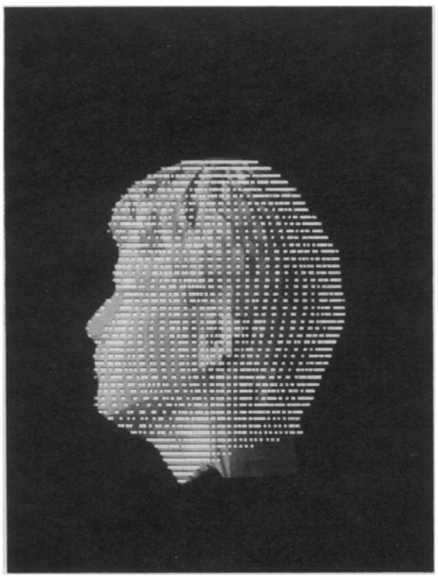
Continued from p.217

from pixel-by-pixel motion detectors, the system marries this input to a simulated face mask rooted in human anatomy. Among its many virtues is its ability to portray an authentic smile by mimicking the raising of eye corners that accompanies the upturning of lips. With a built-in understanding of typical facial gestures, the computer tailors the animated image to an individual's face within a split second. For pure animation, it will generate facial movements.

But why stop at faces? Why not simulate, even automate, whole-body animation? Why not train a computer to watch athletes, dancers, or movie stars and learn their special, subtle moves? A Larry Bird lay-up, a Charlie Chaplin waddle, perhaps a Judy Garland croon. Envision a computer that could take in a great ballet and from the dancers' movements narrate the story.

At the Media Lab, such visions not only raise no eyebrows, they live as bona fide project goals. In a new system called ALIVE, a person wandering before the computer's gaze can watch a replica of himself or herself moving in a virtual world. Within the confines of a virtual 16-foot by 16-foot room, animated autonomous agents roam free in a land of illusion, interacting with other virtual beings.

This project aims, according to Pattie Maes, an MIT computer researcher, to create an artificial environment in which a person can interact, in natural and believable ways, with autonomous, semi-intelligent replicas whose behavior appears equally natural and believable.



Mapping a whole head.

In other words, an automated animation system with no strings attached. Literally. No headgear. No wire-laden data gloves. A system in which a live person's video image unobtrusively feeds a "magic mirror" that interprets that person's silhouette and gestures in real-time, three-dimensional space.

Meanwhile, the user's virtual playmates wander independently in a world they appear to sense, acting on self-generated goals and taking cues from the user's gestures.

In one virtual world, for example, an animated puppet comes over to play, taking the user's virtual hand. When motioned away, the puppet pouts and leaves. When waved back over, the puppet returns giggling. Another virtual setting brings a hamster begging for a meal. Food from a virtual table curbs its appetite, followed by a virtual rub of its virtual tummy. When a predator enters the scene, the hamster scampers away.

In the real world, where most communication occurs without words, such humanized computers represent invaluable learning tools. Since bodies and faces hold such expressive power, one can often glean more about a person's actual moods, intentions, or beliefs from gestures and expressions than from words.

"If a computer has a more human face and is less [emotionally] cool to work with, people can interact with it more naturally," Maes says. "Humanlike agents could train, educate, and motivate people, give personalized feedback, or do tasks for you. But for that to happen, computers must understand facial expressions and gestures as a way of communicating." □

Letters continued from p.211

tion. This hormone is supposed to be safe for humans to eat. However, only if our milk supply is carefully labeled will it be possible to determine the real effects of milk from these cows on humans or other species.

Such labeling is not now required, but it should be.

Joanne Ashley
Manlius, N.Y.

I am a chemical engineer who has been working in hazardous waste management and permitting for almost 23 years, and I have recently spent a great deal of time reviewing Greenpeace's pseudoscientific documents. The documents I have reviewed are always misleading, technically incorrect, biased, and never peer-reviewed. Seeing a Greenpeace report quoted in your magazine, as if it were a credible scientific organization, concerns me greatly because of the implicit approval.

Jim Cudahy
Knoxville, Tenn.

Many toxic substances like DDT build up in fatty tissues. What happens to toxins if a person diets and loses fat? Is the person subjected to a secondary poisoning by the rerelease or megadose of concentrated toxins?

Gregory Hill
Reading, Pa.

The simple answer is yes. As dieting releases fat into the bloodstream, it is recirculated throughout the body. This also explains why nursing – which releases breast fat into milk – can transfer relatively high concentrations of fat-stored toxicants to a baby (see "Organochlorines lace Inuit breast milk," 2/12/94, p.111). – J. A. Raloff

The indictment of chlorine in the environment appears to be alarming, but the inconsequential factoids supporting it make a hash of reasoned (and reasonable) discussion.

What meaning can one attach to the statement that "42 chemicals or classes of chemicals [are] reported to affect the reproductive or hormone system. Twenty-three of the compounds – 55 percent – contain chlorine as an essential ingredient"? How many of these chemicals or classes of chemicals contained such "essential ingredients" as nitrogen, phosphorus, sulfur, sodium, potassium, or half a hundred other elements? Certainly, carbon and hydrogen show up 100 percent of the time.

Abolish chlorine usage? Ask people to stop using bleach and table salt? Ask sewage treatment and water treatment authorities to stop chlorinating? If we abolish the use of chlorine, then plague, pestilence, and famine will sweep across the world. Of course, we'll survive at reduced population levels, but that's OK because our hearts are pure.

Charles J. Sterner
Bethlehem, Pa.

Lifetime vs. annual cancer rate

"Sunscreen can't give blanket protection" (SN: 1/22/94, p.54) states that 1 in 105 Americans will develop melanoma and that 20 percent of them will die from it. At a current U.S. population of approximately 260 million, this would correspond to around 500,000 deaths per year. The article stated that in 1993 there were 6,800 deaths.

Is this an error in the statistics or a very aggressive projection of melanoma cases?

Matthew L. Young
Chaska, Minn.

Not that agressive! Actually, 1 in 105 people in the United States will develop melanoma in the course of a lifetime, not each year. – T. Adler

No hocus about this focus

Talk about hocus-pocus (hocus-focus?). The article on biodiversity ("Biodiversity helps keep ecosystems healthy," SN: 2/5/94, p.84) shows two pictures. One is of the blue sky, the other a closeup chock-full of plants. One shows green and brown grasses, the other a pretty pink flower.

I hope your readers were not as fooled by the psychological photographic trickery as you were.

A. Robert Spitzer
Royal Oak, Mich.

The story is based on data, not photographs. It would only be trickery if the images misrepresented what the data say. – E. Pennisi

Address communications to:
Editor: SCIENCE NEWS
1719 N Street, N.W.
Washington, D.C. 20036
All letters subject to editing.