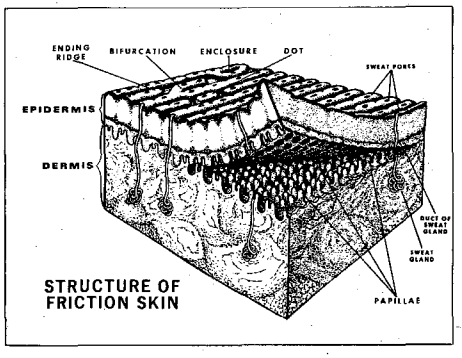
Lab Notes

Every fingerprint has 100 or more minutiae. As a rule of thumb, the positive matching of a dozen or so minutiae is sufficient to establish identity in American courts.



A cross section of fingertip skin. The epidermis (outer skin) is anchored to the dermis (inner skin) by a double row of peglike papillae that determine how the fingerprint's ridges run. Thus superficial injuries to the epidermis and the normal sloughing of skin cells do not change a person's fingerprints.

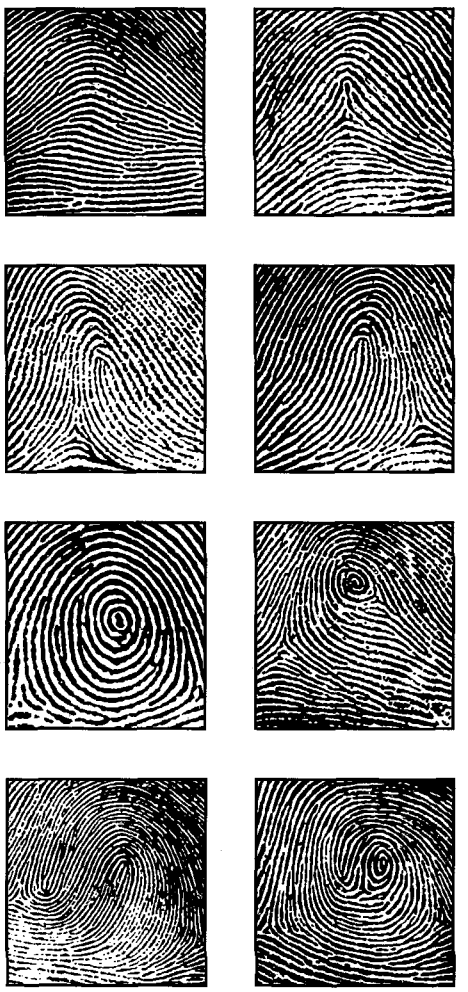
Digital Computing

Watch practically any science-fiction movie, and there will be a scene where some high-security area—the master computer room, or the vault where the viruses are kept—is protected by a palm lock. Someone walks up to the door, raises a hand, and presses the palm against a plate set in the wall. The device recognizes the palm print and opens the door—or refuses to, once the bad guys have seized control.

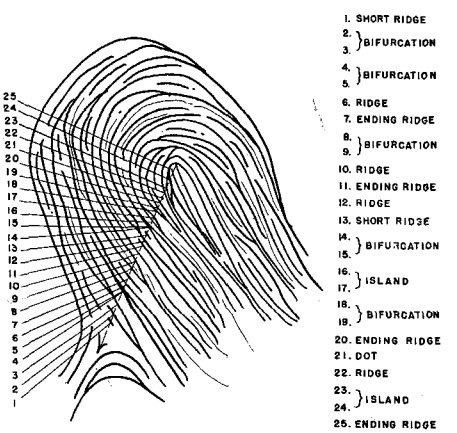
This stock feature of 21st-century technology is taking form today, and here it is only 1993. Pierre Baldi (PhD '86), a member of the technical staff at JPL and visiting associate in biology at Caltech, and Yves Chauvin, a visiting research associate at Stanford University's psychology department, have developed a system that scans a fingerprint, compares it against a set of stored fingerprints, and decides whether it matches one of the stored ones. In addition to keeping bad guys (or good guys) out of supersecret labs, the technology could be applied to door locks of all sorts, and expanded to such other workaday applications as credit-card, check, or passport verification.

Fingerprint matching for identity verification dates back to at least 1901, when Sir Edward Henry introduced his system of fingerprint classification at Scotland Yard. (Today, variants of Henry's system are used by many agencies, including the FBI.) Now, as then, fingerprints were taken by pressing an inked finger against a white index card. After sorting the fingerprints into such classes as loops, arches, and whorls, final matches are made by comparing minutiae—quirks of individual lines in the print, such as bifurcations, trifurcations, line endings, and islands. Every fingerprint has 100 or more minutiae. As a rule of thumb, the positive matching of a dozen or so minutiae is sufficient to establish identity in American courts. But for your average computer to make those dozen matches entails locating every minutia in both prints—a complex task in its own right—and comparing all ten-thousand-plus possible pairings of those minutiae. To further confound J. Edgar Computer, ink impressions are usually blurry, dirty, and replete with gratuitous ink spots. And minutiae are unreliable—a quick

The computational problems are horrendous, and minutiae-based matching programs haven't fared too well.



Fingerprint classes and subclasses, according to the FBI's system.
Top row: The arch (left), and tented arch. About five percent of all fingerprints fall in this class.
Second row: Loops come in two varieties. Radial loops open toward the thumb, while ulnar loops open toward the pinkie. Loops account for roughly 65 percent of all prints, with the ulnar loop being the single most common print.
Third row: The whorl (left), and the central pocket loop, a subclass of whorls. Whorls make up some 30 percent of all prints.
Bottom row: Double loops (left) are also a whorl subclass. Accidentals (right) include whatever doesn't fit elsewhere in the system.



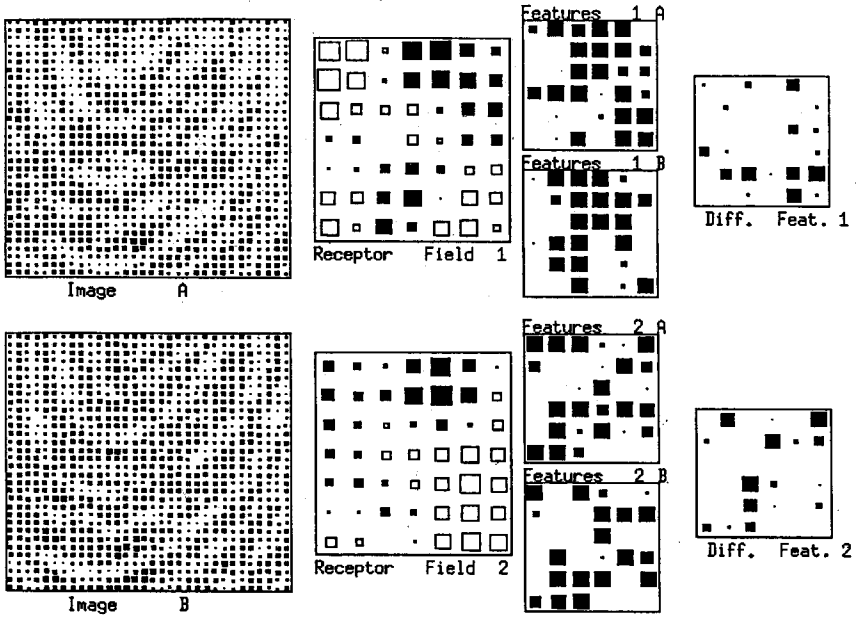
Final classifications are made by counting the number of ridges that intersect a very precisely defined line. From here, the only thing needed to establish identity is the minutiae match. This sketch shows a 25-ridge print. Its numbered ridges contain several minutiae—two short ridges, five bifurcations, four line endings, two islands, and a dot.

twitch can create a bifurcation on the index card where none existed on the finger. The computational problems are horrendous, and minutiae-based matching programs haven't fared too well.

Baldi and Chauvin use a piece of software, called a neural network, that recognizes the gestalt of a fingerprint rather than picking it apart. Neural nets, which are loosely modeled on protoplasmic brains, have proven quite adept at rapid pattern recognition—a function basic to survival. Instead of having a few very complex processing units performing sequential calculations in isolation, a neural net has many simple, interconnected processors. Like relatives at a family reunion, the neurons all talk at once, and feedback between them pushes the network into a steady state—the "answer"—almost as fast as the news of Cousin Debbie's engagement travels round the dinner table. The network's programs and memories are encoded in the connections between its neurons. Learning a new task, or creating new memories, involves adjusting the connections' strengths.

The network learns to match fingerprints by trial and error—it is given pairs of images and asked whether they are the same or not. Its connections are tweaked a bit after each guess until it picks the right answer every time. To do this, of course, you need lots of fingerprints. "We phoned the FBI, and that didn't work, for obvious reasons," Baldi recalled dryly. "They wanted to investigate us. 'Who are you? What do you want it for? What is your telephone number?' Plus, some of their prints are of very poor quality, because they're taken from people who don't want to be fingerprinted. We ended up constructing our own data base." The FBI's fingerprint files wouldn't have helped much, anyhow—in order to teach print matching, you need multiple prints from the same finger, taken at different times under different conditions, so that the network learns how much a print can vary and still constitute a match. So Baldi and Chauvin rounded up twenty of the usual suspects—colleagues—taking one index-finger print from each person on each of five different days. From this pool of 4,950 possible pairs of

Right: The frame grabber creates a 512 × 464 pixel image of the fingerprint to be matched. The computer draws a rectangle around what it thinks are the print's edges. It then draws a 105 × 105 pixel square (black), the midpoint of whose top line is the rectangle's center. This square is slid over a 65 × 65 pixel square (white) of the reference image until they line up. Below, from left: The computer compresses each 65 × 65 image to 32 × 32—Images A and B. Two filters—Receptor Field 1 and 2—scan each image. The squares symbolize the filters' internal connections. Black is positive, white is negative, and a square's size reflects the connection's strength. The more black in the Features column, the more that filter recognizes what it sees. Upon subtraction (Diff. Feat.), the more white, the better the match. The last column comes up black for a match, white for a mismatch.

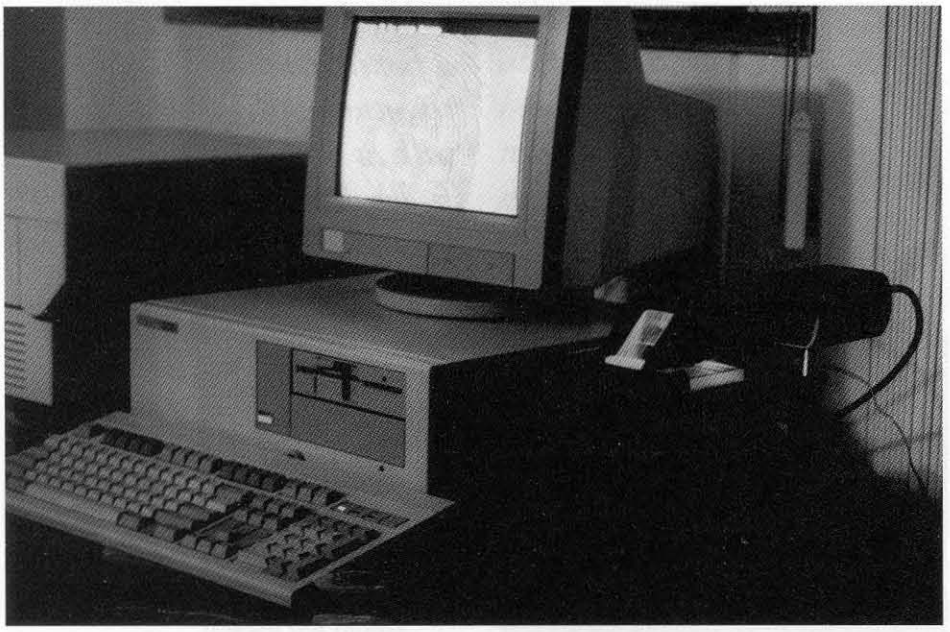


Like relatives at a family reunion, the neurons all talk at once, and feedback between them pushes the network into a steady state—the "answer"—almost as fast as the news of Cousin Debbie's engagement travels round the dinner table.

prints, they used 300 pairs for training, saving the remaining 4,650 to verify that the network had mastered the task.

To input a print, the subject presses his or her finger against a prism. (Nobody wipes the prism off between takes, adding some real-world noise to the data.) The prism reflects a beam of light into a CCD camera—essentially a video camera. Wherever the fingerprint's ridges touch the glass, however, there is no reflection. The resulting pattern of bright and dark lines seen by the CCD camera is converted to a still picture in digital format by a frame grabber. This digital image contains nearly two million bits of data—rather a lot—so the computer locates the center of the fingerprint and then discards everything but a patch about ten ridge widths square, just below the center of the image.

The computer then slides this square patch over an even smaller square of the reference image in its memory until the two images line up—even if the fingerprints are different, you can still shove them around until they are more or less aligned. The computer discards the part of the test image that sticks out beyond the reference square, and compresses the two overlaid squares into squares one-quarter their size. The compressed squares are scanned by a set of "filters"—



The neural net running on a personal computer. To the right of the computer, the prism glows brightly; the CCD camera lurks behind it in the shadows.

specialized groups of neurons—that look for specific features. (At least two different filters are needed to make reliable matches.) "The network learns what kinds of filters to use as it goes through the training," Baldi noted. "We don't tell it, 'Oh, you should use an averaging filter, or you should use a low-pass filter.' We tell it, 'Use whatever you think is necessary for this task.' That's the essence of neural networks. We don't even know what the filters are looking for. One of them seems to detect lines at a specific orientation, say 45 degrees. It's tempting to say the others are looking for minutiae, but very difficult to prove it." Each filter summarizes its findings as a 6×6 matrix, each element of which ranges from 0 ("I don't see anything remotely resembling what I'm looking for") to 1 ("Ooh, here's a beauty!").

"For the lock on your car, a 25-second wait is annoying, and could be dangerous if you're parked in a bad neighborhood. But for checking passports at the border, it's reasonable."

The network compares the filtered images by subtracting one matrix from the other, with the differences squared to ensure that all the results are positive numbers. A zero difference between two matrix elements means a perfect match in that piece of the image, while a difference of one means not even close. The more zeros (or very small numbers) that come out of this subtraction, the better the two prints match, in that filter's view. The neural net then tots up all the filters' comparisons and decides whether the prints match. The net's final output is a number between one (dead certainty) and zero (no way) that represents the computer's level of confidence that the prints came from the same finger. Not that the system hedges its guesses, mind you—it generally gives its confirmed matches ratings of 0.8 or higher, while the pairs it says are mismatches it rates as less than 0.2. It has good reason to be cocky. In the verification set of 4,650 pairs, it was wrong only 0.5 percent of the time, with the errors about evenly split between false matches and false mismatches. Since letting the wrong person in is usually more catastrophic than keeping the right person out, the system can be adjusted to give no false matches at a penalty of increasing the rate of false mismatches to four percent.

Baldi and Chauvin created the neural net on a midsized computer. The training program took all night to run, but now that the machine is a certified dactyloscopist, it only takes ten seconds to make up its mind. The same net running on a personal computer takes 20 to 25 seconds to decide whether the owner of a finger is who he claims to be. "If that's for the lock on your car, a 25-second wait is annoying, and could be dangerous if you're parked in a bad neighborhood. But for checking passports at the border, it's reasonable. If we ran the same software on a parallel machine, or built a special-purpose chip to run it, we could make matches in milliseconds."

Matching the owner of a fingerprint (or a palm print) against a short list of people authorized to enter a virus vault is a much simpler proposition than identifying an airplane-crash victim by matching a fingerprint in the FBI's files. "If you're working with a small data base, say the five people in your family that drive the car, you can do five matches in quick succession." But the FBI has roughly two billion fingerprints on file, and, even at a millisecond per print, it would take more than three weeks of machine time to try them all. "But most of the interesting commercial applications, like credit cards and passports, are simple validation problems. And yes, there is commercial interest in this process." □ —*DS*