improve things in a number of ways. One important area is to quantify the chance that the evidence is unique by applying various probability models.

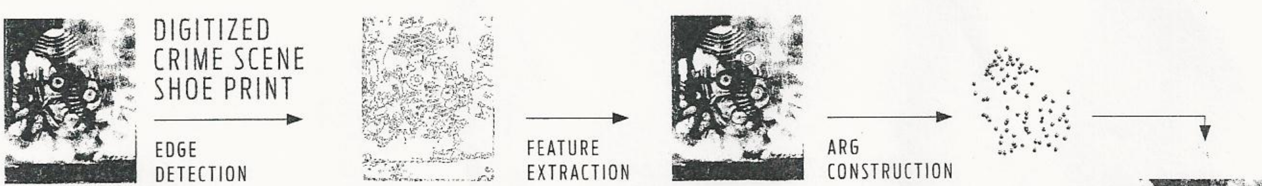
Let's look at the uniqueness problem as it applies to fingerprints. As children, we learn that no two people have the same fingerprints. That may indeed be true, but in a sufficiently large pool of people, the likelihood of two of them having very similar fingerprints is quite high. Beginning in the 1800s, scientists began looking for distinct features within the ridges of fingerprints—such as whorls, arches, single loops, and double loops—that would help distinguish one person's prints from another's. However, such broad features occur with surprising frequency. Left loops, for example, are found in 30 percent of all fingerprints, and right loops are nearly as common. So fingerprint analysts also look at more detailed features, known as minutiae, such as the locations and directions of ridge endings and bifurcations.

Fingerprint-matching algorithms mostly search for and compare minutiae rather than the larger features, because minutiae are less common and yet are still stable and can be reliably extracted and compared by computer. So the key question in quantifying the individuality of fingerprint evidence is, what are the odds that two randomly chosen samples will have a high degree of similarity?

Using statistical modeling based on the frequencies of specific minutiae in the population, the theoretical probability that a random pair of fingerprints will exhibit a match of, say, 12 of 36 minutiae works out to 1 in 100 billion. While that sounds extraordinary, keep in mind that there is almost always uncertainty in the quality of the minutiae data. For instance, a smudged or partial fingerprint can lead a human analyst (or a computer, for that matter) to flag a similarity where none may in fact exist. Also, more research is needed to determine the exact frequencies with which certain fingerprint features occur. At present, nobody really knows how likely it is for two people's prints to match closely or the extent to which trained fingerprint analysts can tell them apart. Indeed, experts sometimes disagree on what constitutes a match.

To address these issues, my research group and others are trying to figure out how to calculate the probability of one person's fingerprints randomly matching those of any other in a given population. That's a key thing to know. If a fingerprint found at the crime scene corresponds reasonably well to a print from the accused, the prosecution hypothesis is that they are from the same person, while the defense hypothesis is that they are from different people. With the right computational tools you can calculate two probability values, one for the prosecution hypothesis and one for the defense hypothesis. And the ratio of the former to the latter shows just how incriminating (or not) this evidence is.

COMPUTATIONAL FORENSICS can also be used to narrow down the range of possible matches against a database of cataloged patterns. To do that, you need a way to quantify the similarity between the



IF THE SHOE PRINT FITS: Shoe prints are among the most common evidence found at crime scenes. They turn up more frequently than fingerprints, perhaps because criminals often think to cover their hands but not their shoes. Shown here is the two-part process by which a computer searches for and compares a digitized crime scene shoe print and a known print taken from a shoe print database. The computer processes the crime scene shoe print by detecting the edges of the print and features such as ellipses and line segments within the tread. From the results, the computer constructs an attribute relational graph (ARG) to show relations between the various components.

The known shoe print is enhanced, and then, as with the crime scene print, its edges and features are detected and an ARG constructed. The two ARGs are then compared, resulting in a score ranging from 0 to 1, with 1 indicating a complete match and 0 indicating no similarity. The comparison can be done with each known print in the database, eventually resulting in a short list of prints that closely match the crime scene print.

CENTER OF EXCELLENCE FOR DOCUMENT ANALYSIS AND RECOGNITION/STATE UNIVERSITY OF NEW YORK AT BUFFALO