Proc. Natl. Acad. Sci. USA
Vol. 96, pp. 11065–11066, September 1999
From the Academy

This paper is a summary of a session presented at the first Chinese-American Frontiers of Science symposium, held August 28–30, 1998, at the Arnold and Mabel Beckman Center of the National Academies of Sciences and Engineering in Irvine, CA.

Automated biometrics-based personal identification

WEICHENG SHEN* AND TIENIU TAN†

*Identification Technology Division, EER Systems Inc. McLean, VA 22102; and †National Laboratory of Pattern Recognition (NLPR), Institute of Automation, Chinese Academy of Sciences, Beijing 100080, China

Why Use Biometrics-Based Personal Identification?

Biometrics-based personal identification attempts to answer the questions "Who are you?" and "Are you who you claim to be?" Personal identification, regardless of method, is ubiquitous in our daily lives. For example, we often have to prove our identity to gain access to a bank account, to enter a protected site, to draw cash from an ATM, to log in to a computer, to claim welfare benefits, to cross national borders, and so on.

Conventionally, we identify ourselves and gain access by physically carrying passports, keys, badges, tokens, and access cards or by remembering passwords, secret codes, and personal identification numbers (PINs). Unfortunately, passports, keys, badges, tokens, and access cards can be lost, duplicated, stolen, or forgotten; and passwords, secret codes, and PINs can easily be forgotten, compromised, shared, or observed. Such loopholes or deficiencies of conventional personal identification techniques have caused major problems to all concerned. For example, hackers often disrupt computer networks; credit card fraud is estimated at $2 billion per year worldwide; and in the USA, welfare fraud (by double dippers) is believed to be in excess of $4 billion a year. Robust, reliable, and foolproof personal identification solutions must be sought to address the deficiencies of conventional techniques.

Right at the frontier of such solutions is biometrics-based personal identification. What are biometrics? Biometrics are personal physical or biological measurements about an individual. Some frequently used measurements are height, weight, hair color, eye color, and skin color of an individual. As one may easily observe that these particular measurements can give a "correct" description about an individual, but more than one individual can fit such a description. To uniquely identify an individual based on biometrics data, the following characteristics of biometrics data are desirable: highly unique to each individual, easily obtainable, time-invariant (no significant changes over a period of time), easily transmittable, able to be acquired as nonintrusively as possible, and distinguishable by humans without much special training. The last characteristic above is helpful to manual intervention, when deemed necessary, after an automated biometrics-based identification/verification system made a decision.

Automated biometrics-based personal identification systems can be classified into two main categories: identification and verification. In a process of verification (1-to-1 comparison), the biometrics information of an individual, who claims certain identity, is compared with the biometrics on the record that represent the identity that this individual claims. The comparison result determines whether the identity claims shall be accepted or rejected. On the other hand, it is often desirable to be able to discover the origin of certain biometrics information to prove or disprove the association of that information with a certain individual. This process is commonly known as identification (1-to-many comparison).

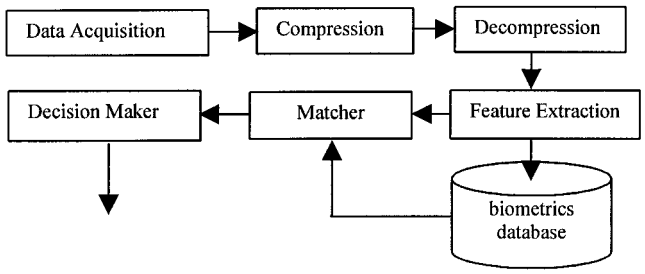

FIG. 1. A generic biometrics-based system.

A Generic System

How can a computer compare two sets of biometrics data and make a decision about whether they belong to the same individual or not? In a strict sense, a computer is not able to do that based on the biometrics alone. A computer can provide only a similarity measurement, a numerical score, which informs the operator about the similarity of the pair of underlying biometrics data. In addition, it can generate a list of pair-wise biometrics data comparisons that are in an ascending order, commonly known as the candidate list. Note that a candidate list is produced only in an identification process.

A typical automated biometrics-based identification/verification system consists of the six major components depicted in Fig. 1. The first component of an automated biometric identification/verification system is a data acquisition component that acquires the biometric data in digital format by using a sensor. For fingerprints, the sensor is typically a scanner; for voice data, the sensor is a microphone; for face pictures and iris images, the sensor is typically a camera. The quality of the sensor has a significant impact on

FIG. 2. A fingerprint.

PNAS is available online at www.pnas.org.

FIG. 3. A face image.

the accuracy of the comparison results. The second and third components of the system are optional. They are the data compression and decompression mechanisms, which are designed to meet the data transmission and storage requirements of the system. The fourth component is of great importance, the feature extraction algorithm. Feature extraction algorithm produces a feature vector, in which the components are numerical characterizations of the underlying biometrics. The feature vectors are designed to characterize the underlying biometrics so that biometric data collected from one individual, at different times, are "similar," while those collected from different individuals are "dissimilar." In general, the larger the size of a feature vector (without much redundancy), the higher its discrimination power. The discrimination power is the difference between a pair of feature vectors representing two different individuals. The fifth component of the system is the "matcher," which compares feature vectors obtained from the feature extraction algorithm to produce a similarity score. This score indicates the degree of similarity between a pair of biometrics data under consideration. The sixth component of the system is a decision-maker.

Existing Methods

One of the best known and frequently used biometrics is one's fingerprint (1). Fig. 2 shows a fingerprint captured on a fingerprint card and then scanned into a digital format. Although fingerprints are known to be unique for each individual, using fingerprints to identify or verify an individual's identity requires special fingerprint comparison skill. For a pair of untrained eyes, it may be difficult to distinguish one set of fingerprints from another.

A frequently used biometric, less unique for each individual than fingerprints, is the face (2). Fig. 3 shows a face image. Face pictures often are used as a means for verification, as evidenced by various picture ID cards, because a person's face is an easily accessible and verifiable biometric to human eyes. Nevertheless, faces are known to be ambiguous for identification/verification purposes, as different individuals can have similar facial features. The faces of a pair of identical twins are almost indistinguishable, for example.

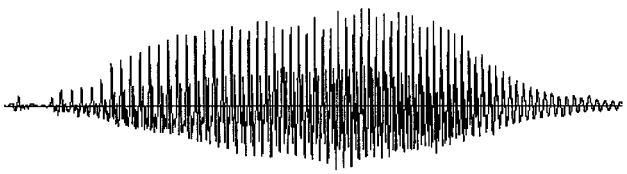

FIG. 4. A segment of speech.

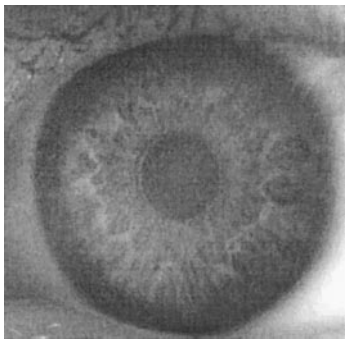

FIG. 5. A human iris.

Another frequently used biometric is one's voice (3). The speech signal waveform for a word is depicted in Fig. 4. Each person's voice has different characteristics, which is the basis for differentiating one from another. Its main advantage is in transmission of the biometrics data. The wide availability of telephone service in the United States provides efficient and reliable data transmission for biometrics data of voice. On the negative side, using voice as biometric data to identify or verify identity also lacks the uniqueness that fingerprint-based identification/verification techniques can provide.

The human iris recently has attracted the attention of biometrics-based identification/verification research and development community (4). Part of the attractiveness of the human iris is that its feature vectors can be compactly represented. A human iris image is depicted in Fig. 5. It has been demonstrated that with a feature vector of relative small size, the human iris exhibited high discrimination power for differentiating different individuals.

Future Research

As discussed above, it appears that none of the biometrics mentioned has all of the desirable characteristics that we described in the beginning. That observation leads to the consideration for hybrid-biometrics system. In other words, we might develop an automated biometrics-based identification/verification system that uses more than one type of biometrics so it can achieve better accuracy. Alternatively, one can develop different biometrics-based systems for different applications because each application has its own specifications.

In conclusion, different applications demand different performances of the biometrics-based systems. Various biometrics have been studied for personal identification and verification purposes. As the demands for high-performance and highly reliable biometrics-based identification/verification systems increase, more hybrid biometrics-based systems are expected to be developed to meet the ever-increasing needs for automated personal identification.

1. Ratha, N., Chen, S. & Jain, A. K. (1996) *IEEE Trans. Pattern Anal. Machine Intell.* **18,** 799–813.
2. Lades, M., Vorbruggen, J., Buhmann, J., Lange, J., Malburg, C. V. D. & Wurtz, R. (1993) *IEEE Trans. Comput.* **42,** 301–311.
3. Campbell, J., Jr. (1997) *Proc. IEEE* **85,** 1437–1462.
4. Wildes, R. (1997) *Proc. IEEE* **85,** 1348–1363.