Biometric Identification Techniques

Hunny Mehrotra
Indian Institute of Technology Kanpur, India

Pratyush Mishra
Indian Institute of Technology Kanpur, India

Phalguni Gupta
Indian Institute of Technology Kanpur, India

INTRODUCTION

In today's high-speed world, millions of transactions occur every minute. For these transactions, data need to be readily available for the genuine people who want to have access, and it must be kept securely from imposters. Some methods of establishing a person's identity are broadly classified into:

1. *Something You Know:* These systems are known as knowledge-based systems. Here the person is granted access to the system using a piece of information like a password, PIN, or your mother's maiden name.
2. *Something You Have:* These systems are known as token-based systems. Here a person needs a token like a card key, smartcard, or token (like a Secure ID card).
3. *Something You Are:* These systems are known as inherited systems like biometrics. This refers to the use of behavioral and physiological characteristics to measure the identity of an individual.

The third method of authentication is preferred over token-based and knowledge-based methods, as it cannot be misplaced, forgotten, stolen, or hacked, unlike other approaches. Biometrics is considered as one of the most reliable techniques for data security and access control. Among the traits used are fingerprints, hand geometry, handwriting, and face, iris, retinal, vein, and voice recognition.

Biometrics features are the information extracted from biometric samples which can be used for comparison. In cases of face recognition, the feature set comprises detected landmark points like eye-to-nose distance, and distance between two eye points. Various feature extraction methods have been proposed, for example, methods using neural networks, Gabor filtering, and genetic algorithms. Among these different methods, a class of methods based on statistical approaches has recently received wide attention. In cases of fingerprint identification, the feature set comprises location and orientation of ridge endings and bifurcations, known as a minutiae matching approach (Hong, Wan, & Jain, 1998). Most iris recognition systems extract iris features using a bank of filters of many scales and orientation in the whole iris region. Palmprint recognition, just like fingerprint identification, is based on aggregate information presented in finger ridge impression. Like fingerprint identification, three main categories of palm matching techniques are minutiae-based matching, correlation-based matching, and ridge-based matching. The feature set for various traits may differ depending upon the extraction mechanism used.

The system that uses a single trait for authenticity verification is called unimodal biometric system. A unimodal biometric system (Ross & Jain, 2003) consists of three major modules: sensor module, feature extraction module, and matching module. However, even the best biometric traits face numerous problems like non-universality, susceptibility to biometric spoofing, and noisy input. Multimodal biometrics provides a solution to the above mentioned problems.

A multimodal biometric system uses multiple sensors for data acquisition. This allows capturing multiple samples of a single biometric trait (called multi-sample biometrics) and/or samples of multiple biometric traits (called multi-source or multimodal biometrics). This approach also enables a user who does not possess a particular biometric identifier to still enroll and authenticate using other traits, thus eliminating the enrollment problems. Such systems, known as multimodal biometric systems (Tolba & Rezq, 2000), are expected to be more reliable due to the presence of multiple pieces of evidence. A good fusion technique is required to fuse information for such biometric systems.

Depending on the application context, a biometric system may operate either in verification or identification mode (Jain, Bolle, & Pankanti, 1999a). In verification mode, the system validates a person's identity by comparing the captured biometric data with his or her own biometric template stored in the system database. In such a system, an individual who desires to be recognized claims an identity (usually via PIN), a user name, or a smartcard, and the system conducts one-to-one comparison to determine whether the claim is true or not. In the identification mode the system recognizes the individual by searching the templates of all the users in the

database for a match (Ross, Nandakumar, & Jain, 2006). The time required by the biometric system to claim identification is critical for many applications. Apart from good accuracy, response time, and retrieval, efficiency plays an important role in the identification mode.

BACKGROUND

A biometric system is essentially a pattern recognition system that recognizes a person based on a feature vector derived from a specific physiological or behavioral characteristic that the person possesses (Prabhakar, Pankanti, & Jain, 2003). A brief overview of the field of biometrics and a summary of some of its advantages, disadvantages, strengths, limitations, and related privacy concerns are presented in Jain, Ross, and Prabhakar (2004).

Multimodal biometric systems are those that utilize more than one physiological or behavioral characteristic for enrollment, verification, or identification. Ross and Jain (2003) have presented an overview of multimodal biometrics and have proposed various levels of fusion, various possible scenarios, the different modes of operation, integration strategies, and design issues. Different fusion strategies as well as its performance are given by Allano et al. (2006). The performance of multimodal biometric authentication systems using state-of-the-art commercial off-the-shelf (COTS) fingerprint and face biometric systems on a population approaching 1,000 individuals is examined by Snelick, Uludag, Mink, Indovina, and Jain (2005).

The identification system is designed in such a way that the search space and data retrieval time reduce to the minimum. Classification techniques have been introduced to reduce the search time of identification systems. There exist several classification techniques like classification of face images based on age (Kwon & Lobo, 1999), where input images can be classified into one of three age-groups: babies, adults, and senior adults. Gender classification from frontal facial images using genetic feature subset selection is considered in Sun, Bebis, Yuan, and Louis (2002). Further, nonlinear support vector machines (SVMs) are investigated for appearance-based gender classification with low-resolution thumbnail faces from a FERET (FacE REcognition Technology) face database (Moghaddam & Yang, 2002). Classification of fingerprints into five categories—whorl, right loop, left loop, arch, and tented arch—is done using a novel representation (FingerCode) based on a two-stage classifier (Jain, Prabhakar, & Hong, 1999c). Ern and Sulong (2001) give a good account of fingerprint classification techniques, and Jain, Murty, and Flynn (1999b) overview pattern clustering methods from a statistical pattern recognition perspective, with a goal of providing useful advice and references to fundamental concepts accessible to the broad community of clustering practitioners. Another approach for search space reduction is to index the database using some data structures like B+ trees, pyramid technique (Mhatre, Palla, Chikkerur, & Govindaraju, 2005), or hash functions. However the main concern behind the use of suitable indexing technique is to catalog the database in such a way that the identification template falls in the bin of the enrolled template.

IDENTIFICATION TECHNIQUES

Biometrics identification is expensive in terms of time as it involves a large number of comparisons in the database. As database size increases, data retrieval and search times increase significantly. Thus to overcome the problem arising from large biometric records, there must be some method to reduce the search space for a matcher to operate. The identification system is designed in such a way that the search space and data retrieval times reduce to the minimum. The biometric database can be segmented using one or all of the following three approaches in the hierarchy discussed below.

Classification

Classification refers to assigning an object physically into one of a set of predefined categories. The main idea behind the use of a classification algorithm is to divide the database into groups where each group has homogenous characteristics. The important step is to design a classifier based on texture, pattern, or some soft biometric attribute. Some commonly used classification techniques for well-known traits include face, fingerprint, and iris classification.

Face Classification

Face images are used to classify a person based on age, gender, ethnicity, and so forth. Age classification is used to classify the input images into one of three age groups: babies, young adults, and senior adults (Kwon & Lobo, 1999). The computations are based on skin wrinkle analysis.

Automated gender classification is possible with the use of statistical classification algorithms that are trained with a set of known images. With a large number of known images, it may be possible to directly train the classification algorithms, with each grayscale image representing a point in the image space. However, it is often much better to reduce the number of dimensions, in an efficient way, allowing accurate training with less observations. In applications using face images, a well-established dimension reduction technique is to use the Karhunen-Loeve Transform on an ensemble of images, to generate an eigenspace composed of eigenfaces. The idea is to project images into this eigenspace and use