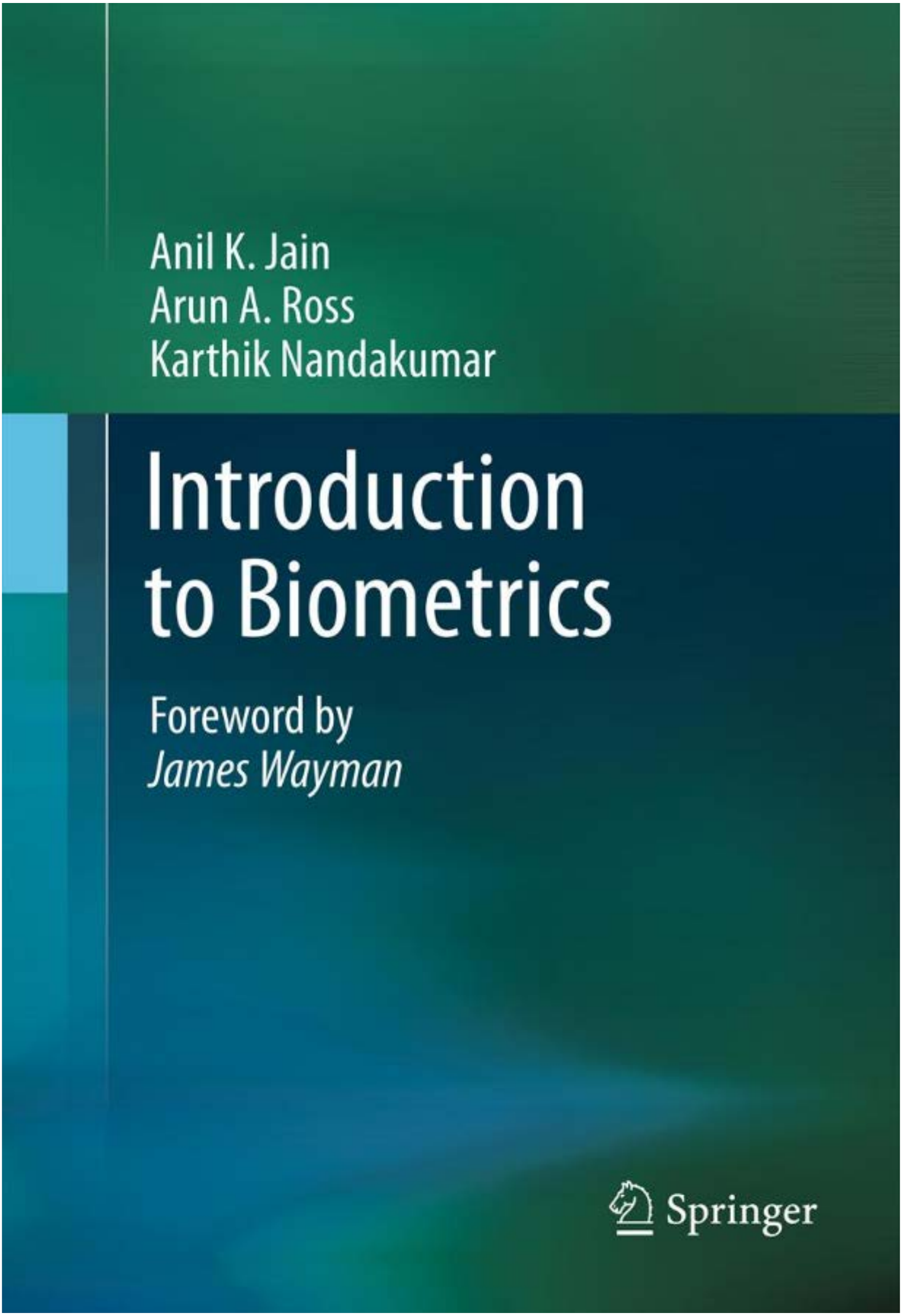


Introduction to Biometrics

Anil K. Jain • Arun A. Ross • Karthik Nandakumar

Introduction to Biometrics

Foreword by James Wayman

Springer

Prof. Anil K. Jain
Department of Computer Science
and Engineering
Michigan State University
East Lansing, Michigan
USA
jain@cse.msu.edu

Dr. Arun A. Ross
Lane Department of Computer Science
and Electrical Engineering
West Virginia University
Morgantown, West Virginia
USA
arun.ross@mail.wvu.edu

Dr. Karthik Nandakumar
Institute for Infocomm Research
A*STAR
Fusionopolis
Singapore
knandakumar@i2r.a-star.edu.sg

ISBN 978-0-387-77325-4 e-ISBN 978-0-387-77326-1
DOI 10.1007/978-0-387-77326-1
Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2011942231

Contents

Chapter 1
INTRODUCTION

Sur un procédé d'identification permettant de retrouver le nom d'un récidiviste au moyen de son seul signalment, et pouvant servir de cadre pour une classification de photographies à la préfecture de police, à la sûreté générale, au ministé de la justice, etc.

Alphonse Bertillon, 1881[1].

About an identification process that enables finding the name of a repeat offender based on his description only, and that can be used in the context of a classification of photographies in the police headquarters, in the national security office, at the ministry of justice, etc.

English translation

The ability to identify individuals uniquely and to associate personal attributes (e.g., name, nationality, etc.) with an individual has been crucial to the fabric of human society. Humans typically use body characteristics such as face, voice, and gait along with other contextual information (e.g., location and clothing) to recognize one another. The set of attributes associated with a person constitutes their personal *identity*. In the early days of civilization, people lived in small communities where individuals could easily recognize each other. However, an explosion in population growth accompanied by increased mobility in modern society has necessitated the development of sophisticated identity management systems that can efficiently record, maintain, and obliterate personal identities of individuals.

Identity management plays a critical role in a number of applications. Examples of such applications include regulating international border crossings, restricting physical access to important facilities like nuclear plants or airports, controlling logical access to shared resources and information, performing remote financial transactions, or distributing social welfare benefits. The proliferation of web-based services (e.g., online banking) and the deployment of decentralized customer service

[1] A. Bertillon, "Une application pratique de l'anthropométrie, Annals de démographie internationale, 1881"

A.K. Jain et al., *Introduction to Biometrics*, DOI 10.1007/978-0-387-77326-1_1,

centers (e.g., credit cards) have led to the risk of identity theft[2]. Rising magnitude of identity theft and heightened concerns about national security have reinforced the need for reliable identity management systems.

1.1 Person Recognition

The fundamental task in identity management is to establish the association between an individual and his personal identity. One must be able to determine a person's identity or verify the identity claim of an individual whenever required. This process is known as person recognition. A person can be recognized based on the following three basic methods (see Figure 1.1): (a) what he knows, (b) what he possesses extrinsically, and (c) who he is intrinsically. While the first method relies on the fact that the individual has exclusive knowledge of some secret information (e.g., password, personal identification number, or cryptographic key), the second method assumes that the person has exclusive possession of an extrinsic token (e.g., identification card, driver's license, passport, physical key, or personal device such as a mobile phone). The third method establishes the person's identity based on his inherent physical or behavioral traits and is known as biometric recognition. Formally, biometric recognition can be defined as the science of establishing the identity of an individual based on the physical and/or behavioral characteristics of the person either in a fully automated or a semi-automated manner.

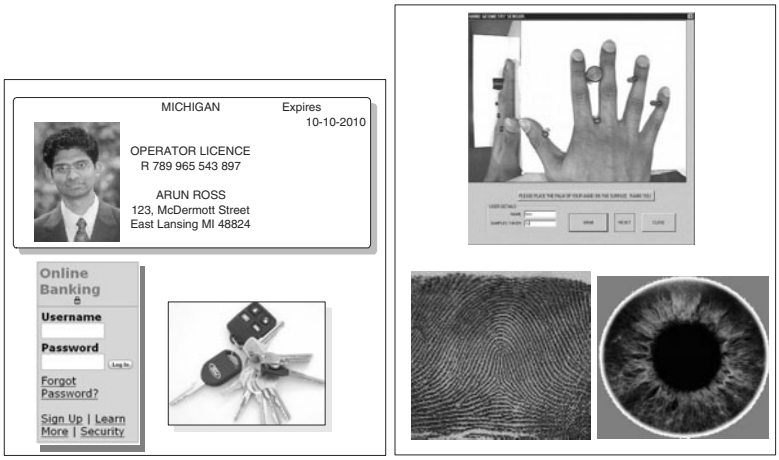
Knowledge-based and token-based person recognition rely on surrogate representations of identity such as passwords or ID cards, which can be easily forgotten/lost, guessed/stolen, or shared. Moreover, they cannot provide vital identity management functions like non-repudiation and detecting multiple enrollments by the same person under different identities. For example, individuals can easily deny (repudiate) using a service by claiming that their password had been stolen or guessed. Individuals can also conceal their true identity by presenting forged or duplicate identification documents. In addition, traditional mechanisms like passwords and tokens do not provide strong evidence for post-event person recognition, such as suspect identification at a crime scene. Therefore, it is becoming increasingly apparent that knowledge-based and token-based mechanisms alone are not sufficient for reliable identity management.

Biometric recognition, or simply biometrics[3], offers a natural and more reliable solution to the problem of person recognition. Since the biometric identifiers are inherent to an individual, it is more difficult to manipulate, share, or forget these traits. Hence, biometric traits constitute a strong and reasonably permanent link between a person and his identity.

---

[2] Identity theft or identity fraud occurs when a person usurps the identity of another individual or claims a false identity in order to access resources or services to which he is not entitled.

[3] The term *biometric recognition* is perhaps more appropriate than *biometrics*, because the latter has been historically used in the field of statistics to refer to the analysis of biological (particularly medical) data. For the sake of brevity, we use these two terms interchangeably in this book.



(a) (b)

Fig. 1.1 Three basic approaches to person recognition. (a) Traditional schemes use passwords ("what you remember") and ID cards or keys ("what you possess extrinsically") to validate individuals and ensure that system resources are accessed only by a legitimately enrolled individual, (b) with the advent of biometrics, it is now possible to establish an identity based on "who you are intrinsically".

Any person who presents his biometric identifier to a biometric system for the purpose of being recognized can be called a *user* of the system. Since biometric systems require the user to be present at the time of authentication, they can also deter users from making false repudiation claims. Moreover, only biometrics can establish whether a certain individual is already known to the identity management system, although the individual might deny it. This is especially critical in applications such as welfare disbursement, where an impostor may attempt to claim multiple benefits (i.e., double dipping). Due to the above reasons, biometric recognition is being increasingly adopted in a number of government and civilian identity management applications either to replace or to complement existing knowledge-based and token-based mechanisms.

1.2 Biometric Systems

A biometric system measures one or more physical or behavioral characteristics (see Figure 1.2), including fingerprint, palmprint, face, iris, retina, ear, voice, signature, gait, hand vein, odor, or the DNA[4] information of an individual to determine or ver-

[4] DNA refers to deoxyribonucleic acid, which contains the genetic information necessary for the development and functioning of living organisms.

ify his identity. These characteristics are referred to by different terms such as *traits*, *indicators*, *identifiers*, or *modalities*. In this chapter, the various building blocks of a generic biometric system and the issues involved in the design, implementation, and evaluation of such a system will be discussed. The details on implementing a biometric system based on specific biometric traits will be dealt with in the subsequent chapters.

1.2.1 Enrollment and recognition phases

How does a biometric system identify a user based on his physical and/or behavioral traits? This process consists of two main phases, namely, enrollment and recognition (see Figure 1.3). During the *enrollment* phase, the biometric data is acquired from the individual and stored in a database along with the person's identity. Typically, the acquired biometric data is processed to extract salient and distinctive features. In many cases, only the extracted feature set gets stored, while the raw biometric data is discarded. During the *recognition* phase, the biometric data is re-acquired from the individual and compared against the stored data to determine the user identity. Thus, a biometric system is essentially a pattern recognition (or a pattern matching) system consisting of four basic building blocks, namely, (a) sensor, (b) feature extractor, (c) database, and (d) matcher as shown in Figure 1.3. These four modules will now be discussed in turn.

1.2.2 Sensor module

A suitable user interface incorporating the biometric sensor or reader is needed to measure or record the raw biometric data of the user. For example, an optical fingerprint sensor may be used to image the friction ridge pattern at the tip of the finger. The design of a good user (or human-machine) interface is critical for the successful implementation of a biometric system. An intuitive, ergonomic, and easy to use interface may facilitate rapid user habituation and enable the acquisition of good quality biometric samples from the user.

The quality of the raw biometric samples also depends on the characteristics of the sensor used. For most biometric modalities, the raw biometric data is in the form of two-dimensional images (e.g., fingerprint, face, iris, etc.). Exceptions include voice (1-dimensional amplitude signals), online signature (pen pressure, position, and velocity), odor and DNA (chemical-based). For image-based data, factors like resolution, frame rate, and sensitivity of the camera play an important role in determining the image quality. Figure 1.4 shows fingerprint images at two different resolutions obtained using different fingerprint sensors. One may also need to consider the demographic characteristics of the target population like age and gender, and other cultural issues (e.g., some users may be averse to touching a sensor sur-

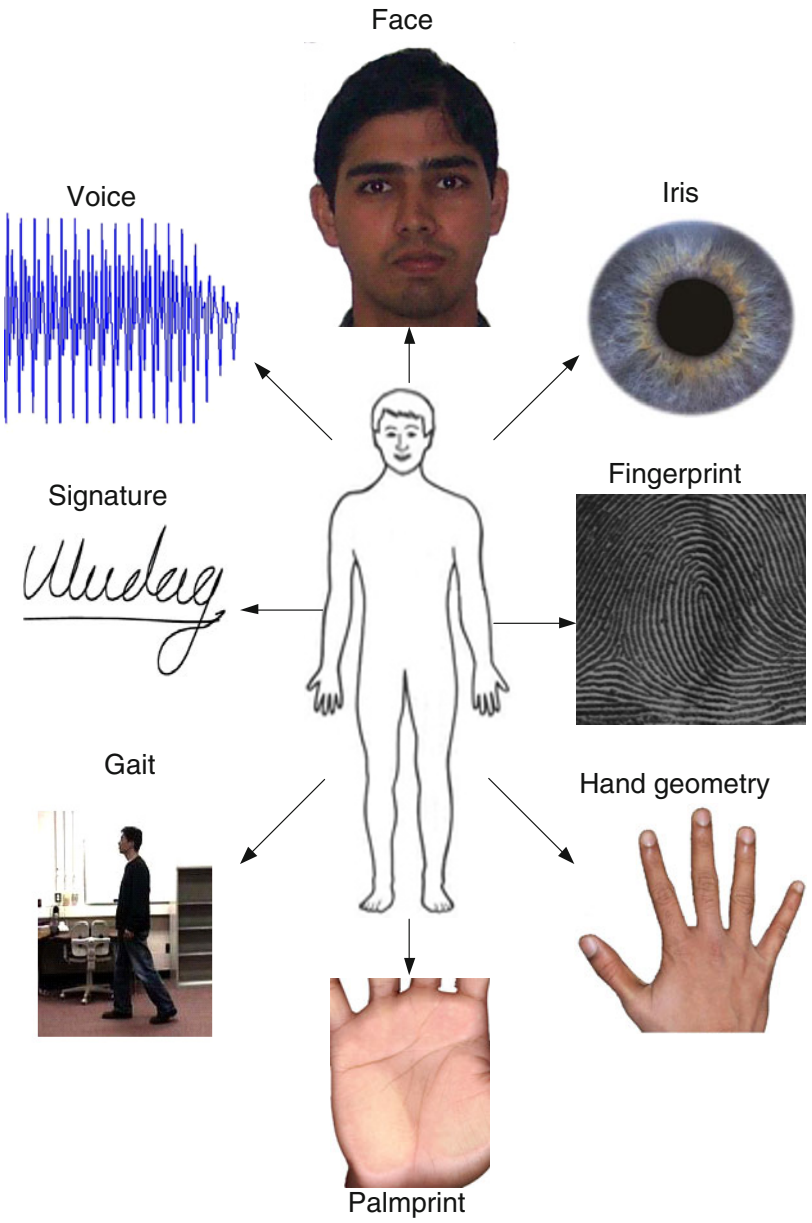


Fig. 1.2 Examples of body traits that have been used for biometric recognition. Physical traits include face, fingerprint, iris, palmprint, hand geometry, voice, and ear shape, while gait, signature, and keystroke dynamics are some of the behavioral characteristics. The distinction between a physical trait and a behavioral characteristic is actually not very important. This is because the biometric data captured from an individual is typically a manifestation of both the physical and behavioral aspects of the person. For example, while fingerprint is a physical trait, the fingerprint image acquired from a person also depends on how he interacts with the sensor, i.e., the user's behavior. Similarly, while gait may be a behavioral trait, it is to some extent defined by the physical characteristics of the human body.

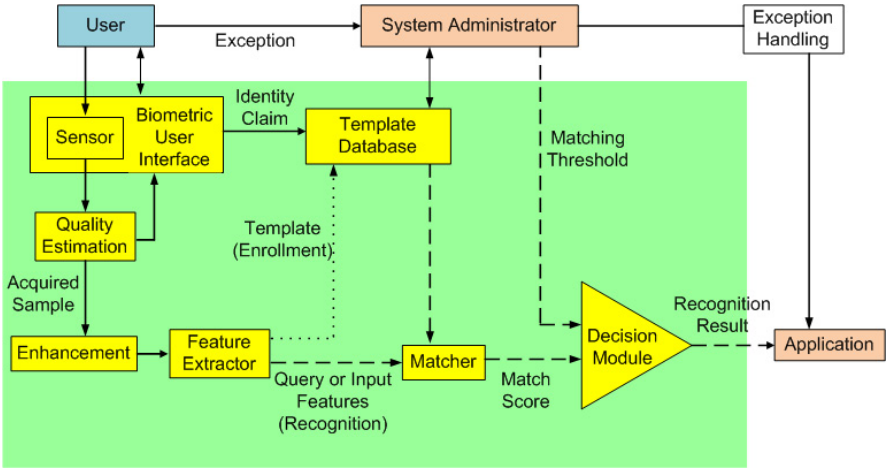


Fig. 1.3 Basic building blocks of a generic biometric system.

face) while designing the sensor module. Furthermore, factors like cost, size, and durability also impact the sensor design.

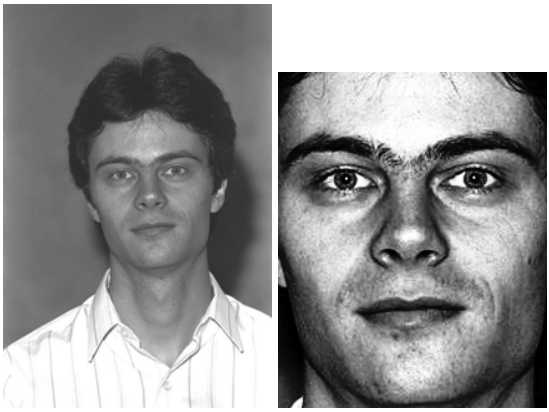
1.2.3 Feature extraction module

Usually, the raw biometric data from the sensor is subjected to pre-processing operations before features are extracted from it. The three commonly used pre-processing steps are (a) quality assessment, (b) segmentation, and (c) enhancement. First, the quality of the acquired biometric samples needs to be accessed to determine its suitability for further processing. If the raw data is not of sufficient quality, there are two options. One can either attempt to re-acquire the data from the user or trigger an exception (failure alarm) alerting the system administrator to activate suitable alternate procedures (typically involving some form of manual intervention by the system operator). The next pre-processing step is known as *segmentation*, where the goal is to separate the required biometric data from the background noise. Detecting a face in a cluttered image is a good example of segmentation. Finally, the segmented biometric data is subjected to a signal quality *enhancement* algorithm in order to improve its quality and further reduce the noise. In the case of image data, enhancement algorithms like smoothing or histogram equalization may be applied to minimize the noise introduced by the camera or illumination variations. Figure 1.5 shows a face image obtained after segmentation and quality enhancement based on histogram equalization. In some cases, the above pre-processing steps may be inseparable from the actual feature extraction step. For example, quality assessment in itself may entail the extraction of some features from the acquired biometric data.

(a) (b)

Fig. 1.4 Fingerprints scanned at (a) 1000 points per inch (ppi) and (b) 500 points per inch using different fingerprint sensors. The intricate details of the finger such as location of sweat pores can be more easily observed in the higher resolution fingerprint image shown in (a) as compared to the lower resolution image in (b).

Feature extraction refers to the process of generating a compact but expressive digital representation of the underlying biometric trait, called a *template*. The template is expected to contain only the salient discriminatory information that is essential for recognizing the person. For example, the position and orientation of minutia points (locations where the friction ridges in a fingerprint pattern exhibit some anomalies) are believed to be unique for each finger. Therefore, detecting the minutia points in a fingerprint image is a key feature extraction step in most fingerprint-



(a) (b)

Fig. 1.5 Face segmentation and enhancement. (a) A face image of a person as captured by the camera and (b) the processed face image obtained after segmentation (removal of the background and other non-face regions such as hair and regions below the chin) and contrast enhancement based on histogram equalization.

based biometric systems. Figure 1.6 shows the commonly extracted features used to represent fingerprint, iris, and face images.

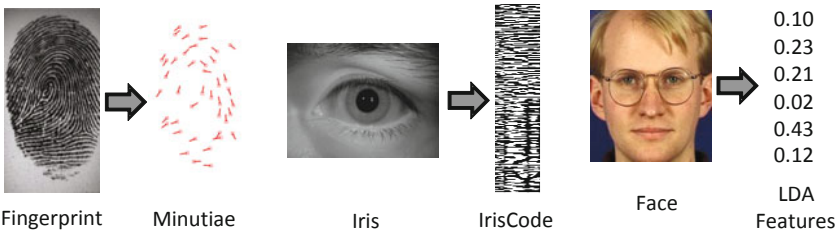


Fig. 1.6 Common features extracted from fingerprints, iris, and face. Fingerprint is commonly represented as a set of points depicting the minutiae; iris is represented as a binary vector depicting the binarized response of an input image to Gabor filters; face is commonly represented as a vector of real numbers depicting, say, the coefficients of Linear Discriminant Analysis (LDA).

During enrollment, the template gets stored either in the central database of the biometric system or is recorded on a token (e.g., smart card) issued to the individual based on the nature of the application.

At the time of recognition, the template is retrieved from the database, and matched against the feature set extracted from the new biometric sample acquired from the user. This new feature set obtained in the recognition phase is usually referred to as the *query* or *input*. In many image-based biometric systems (e.g., face or

fingerprint), the raw biometric images may also be stored in the database along with the templates during enrollment. Such images are often known as *gallery images*, *reference images*, *stored images*, or *enrollment images*. The images acquired during recognition are known as *probe images*, *query images*, or *input images*.

The template of a user can be extracted from a single biometric sample, or generated by processing multiple samples acquired during enrollment. Thus, the minutiae template of a finger may be extracted after mosaicing (combining) multiple impressions of the same finger. Some systems store multiple templates in order to account for the large variations that may be observed in the biometric data of a user. Face recognition systems, for instance, may store multiple templates of an individual, with each template corresponding to a different facial pose with respect to the camera.

1.2.4 Database module

The biometric system database acts as the repository of biometric information. During the enrollment process, the feature set extracted from the raw biometric sample (i.e., the template) is stored in the database along with some personal identity information (such as name, Personal Identification Number (PIN), address, etc.) characterizing the user. One of the key decisions in the design of a biometric system is whether to use a centralized database or a decentralized one. Storing all the templates in a central database may be beneficial from a system security perspective, because the data can be secured through physical isolation and by having strict access control mechanisms. On the other hand, compromise of a central database would have far greater implications than the compromise of one of the sites in decentralized database. This is because malicious individuals (corrupt administrators or hackers) can abuse the biometric information stored in the database to compromise the privacy of innocent users.

1.2.5 Matching module

The purpose of a biometric matcher is to compare the query features against the stored templates to generate match scores. The match score is a measure of the similarity between the template and the query. Hence, a larger match score indicates greater similarity between the template and the query. If a matcher measures the dissimilarity (instead of the similarity) between the two feature sets, the score is referred to as a distance score. A smaller distance score indicates greater similarity. In a fingerprint-based biometric system, the number of matching minutiae between the input and the template feature sets can be considered as the degree of similarity (match score). The match score may also be moderated based on the quality of the presented biometric data. The matcher module also encapsulates a decision making

module, in which the match scores are used to either validate a claimed identity or provide a ranking of the enrolled identities in order to identify an individual.

1.3 Biometric Functionalities

A biometric system can provide two types of identity management functionalities, namely, *verification* and *identification*. Throughout this book, the generic term recognition will be used when we do not wish to make a distinction between the verification and identification functionalities. Moreover, the term authentication will be used as a synonym for verification. Figure 1.7 shows the enrollment and recognition phases of a biometric system operating in the verification and identification modes.

1.3.1 Verification

In verification, the user claims an identity and the system verifies whether the claim is genuine, i.e., the system answers the question "Are you who you say you are?". In this scenario, the query is compared only to the template corresponding to the claimed identity (a one-to-one match). The identity claim is usually made through the use of a Personal Identification Number (PIN), a user name, or a token (e.g., smart card). If the user's input and the template of the claimed identity have a high degree of similarity, then the claim is accepted as "genuine". Otherwise, the claim is rejected and the user is considered an "impostor". In the biometrics literature, the terms "client" or "authentic" are sometimes used in place of the term "genuine". Verification is typically used in applications where the goal is to prevent unauthorized persons from using the services.

Formally, verification can be posed as the following two-category classification problem: given a claimed identity I and a query feature set $\mathbf{x}^A$, we need to decide if $(I, \mathbf{x}^A)$ belongs to "genuine" or "impostor" class. Let $\mathbf{x}_I^E$ be the stored template corresponding to identity I. Typically, $\mathbf{x}^A$ is compared with $\mathbf{x}_I^E$ and a match score s, which measures the similarity between $\mathbf{x}^A$ and $\mathbf{x}_I^E$, is computed. The decision rule is given by

$$(I, \mathbf{x}^A) \in \begin{cases} \text{genuine}, & \text{if } s \geq \eta, \\ \text{impostor}, & \text{if } s < \eta, \end{cases} \tag{1.1}$$

where η is a pre-defined threshold. If a distance score is used in place of the similarity or match score, the inequalities in the decision rule shown in equation (1.1) should be reversed. When the identity claim is deemed to be "genuine", the user is allowed to access the services provided by the system.

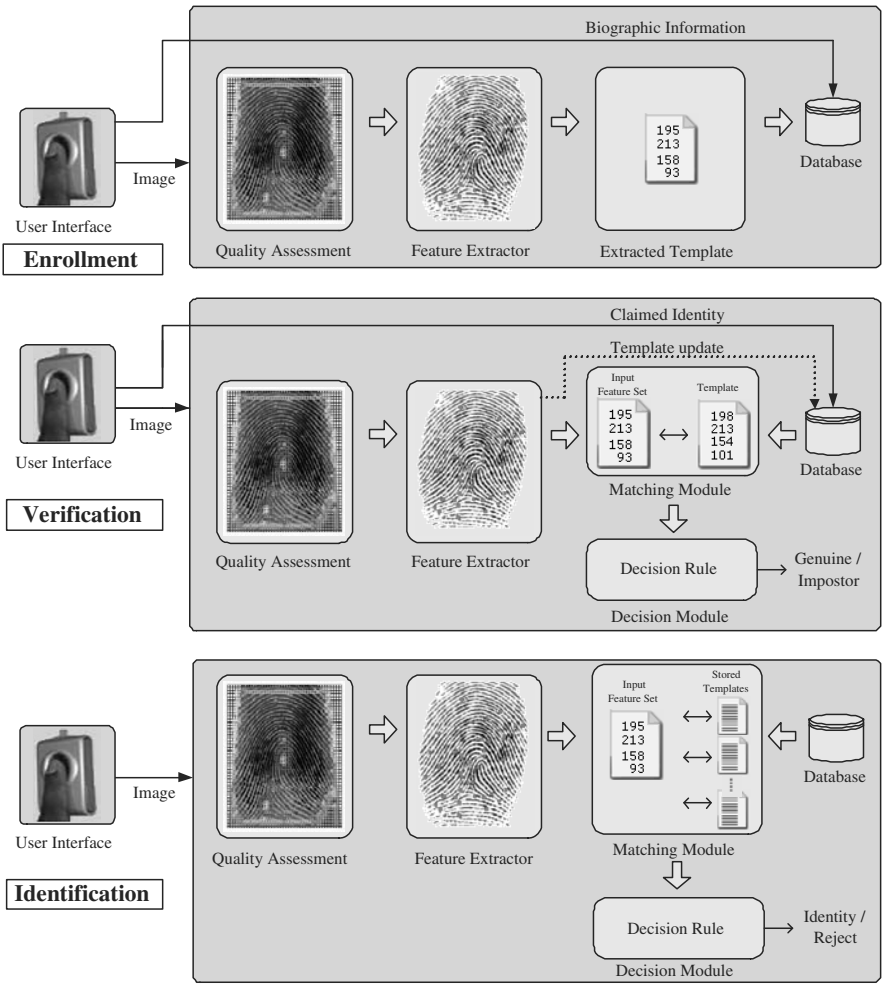


Fig. 1.7 Enrollment and recognition stages of a biometric system operating in the verification and identification modes. The dotted line in the verification module is an optional operation to update a specific user's template.

1.3.2 Identification

Identification functionality can be further classified into positive and negative identification. In positive identification, the user attempts to positively identify himself to the system without explicitly claiming an identity. A positive identification system answers the question "Are you someone who is known to the system?" by determining the identity of the user from a known set of identities. In contrast, the user in a negative identification application is considered to be concealing his true iden-

tity (either explicitly or implicitly) from the system. Negative identification is also known as screening and the objective of such systems is to find out "Are you who you say you are not?".

The purpose of negative identification is to prevent a single person from using multiple identities. Hence, screening can be used to prevent the issue of multiple credential records (e.g., driver's licence, passport) assigned to the same person or to prevent a person from claiming multiple benefits under different names (a problem commonly encountered in welfare disbursement applications). Screening is also often used at airports to verify whether a passenger's identity matches with any person on a "watch-list".

In both positive and negative identification, the user's biometric input is compared with the templates of all the persons enrolled in the database and the system outputs either the identity of the person whose template has the highest degree of similarity with the user's input or a decision indicating that the user presenting the input is not an enrolled user. Formally, the problem of identification can be stated as follows: given a query feature set $\mathbf{x}^A$, we need to decide the identity I of the user, where $I \in \{I_1, I_2, \cdots, I_N, I_{N+1}\}$. Here, $I_1, I_2, \cdots, I_N$ correspond to the identities of the N users enrolled in the system and I_{N+1} indicates the case where no suitable identity can be determined for the given query. If $\mathbf{x}^E_{I_n}$ is the stored template corresponding to identity I_n and s_n is the match score between $\mathbf{x}^A$ and $\mathbf{x}^E_{I_n}$, for $n = 1, 2, \cdots, N$, the decision rule for identification is,

$$\mathbf{x}^A \in \begin{cases} I_{n_0}, & \text{if } n_0 = \arg\max_n s_n \text{ and } s_{n_0} \geq \eta, \\ I_{N+1}, & \text{otherwise,} \end{cases} \tag{1.2}$$

where η is a pre-defined threshold. The above decision rule is commonly known as *open set identification*, because it is possible to return a result indicating that the user presenting his biometric trait is not among the N enrolled users. Almost all practical biometric identification systems (including screening systems) use open set identification. It is also possible to force the system to return one among the N enrolled identities, irrespective of the value of s_{n_0}. Such a scenario is called *closed set identification*.

In some practical biometric identification systems (e.g., latent fingerprint matching), identification is semi-automated. A semi-automated biometric system outputs the identities of the top t matches ($1 < t \ll N$) and a human expert manually determines the identity (among the t selected identities) that best matches the given query. The value of t could be determined based on the availability and throughput of the human expert(s). Against a large database such as the FBI's Integrated Automated Fingerprint Identification System (IAFIS), which has approximately 60 million users enrolled, the typical value of t could range from 20 to 50. Another approach is to return all identities whose corresponding match scores exceed the threshold (η) in equation (1.2). Since the number of enrolled users in the database can be quite large (e.g., FBI-IAFIS), the identification task is significantly more challenging than verification.

1.4 Biometric System Errors

The science of biometric recognition is based on two fundamental premises, namely, *uniqueness* and *permanence* of the underlying biometric trait. A biometric identifier is said to be unique only if any two persons in the world can be differentiated based on the given identifier. A biometric trait is permanent if it does not change over the lifetime of an individual. However, these two premises are seldom true in practical biometric systems. This can be primarily attributed to two reasons.

Firstly, the physical trait itself may not be unique. For instance, when fingerprint recognition systems gained popularity at the beginning of the 20^{th} century, press reports claimed that fingerprints were truly unique.

"Only once during the existence of our solar system will two human beings be born with similar finger markings" - *Harper*'s headline, 1910.

"Two like fingerprints would be found only once every 10^{48} years" - *Scientific American*, 1911.

Such claims were accepted over time, not because of rigorous scientific evidence in their favor, but rather due to a lack of contradiction and relentless repetition. In the last two decades, the claims about the uniqueness of fingerprints have been challenged by both the scientific and legal communities. Similarly, the uniqueness or individuality of other biometric modalities has not been clearly established.

The genetic similarity between related individuals (e.g., twins, father and son) may also contribute to the lack of uniqueness of some biometric traits. For example, the facial appearance of identical twins is almost the same. Modalities such as DNA, where the genetic constitution of the individual largely determines their biometric characteristics are referred to as genotypic factors/features. In contrast, the modalities whose characteristics are determined by other sources of randomness in nature (e.g., fingerprints) are referred to as phenotypic factors/features. Figure 1.8 shows the fingerprint, face, and iris images obtained from identical twins.

Furthermore, the notion that the biometric traits are permanent is also not an established scientific fact. The effects of body growth (especially during childhood and adolescence) on common biometric identifiers like face, fingerprint, or iris, have not been studied in detail. Even casting aside possible natural changes in the physical traits, practical biometric systems face a much more challenging problem. Biometric systems rely only on the digital measurements of the body characteristics, and not the real physical traits. This process of measurement (sensing) introduces variations in the samples of the same biometric trait of a user obtained over a period of time. Consequently, the feature sets obtained from different samples of the same biometric trait of a user are seldom identical.

The variability observed in the biometric feature set of an individual is known as *intra-user variations* or *intra-class variations*. This variability may be due to reasons like imperfect sensing conditions (e.g., noisy fingerprint due to sensor malfunction), alterations in the user's biometric characteristic (e.g., respiratory ailments impacting

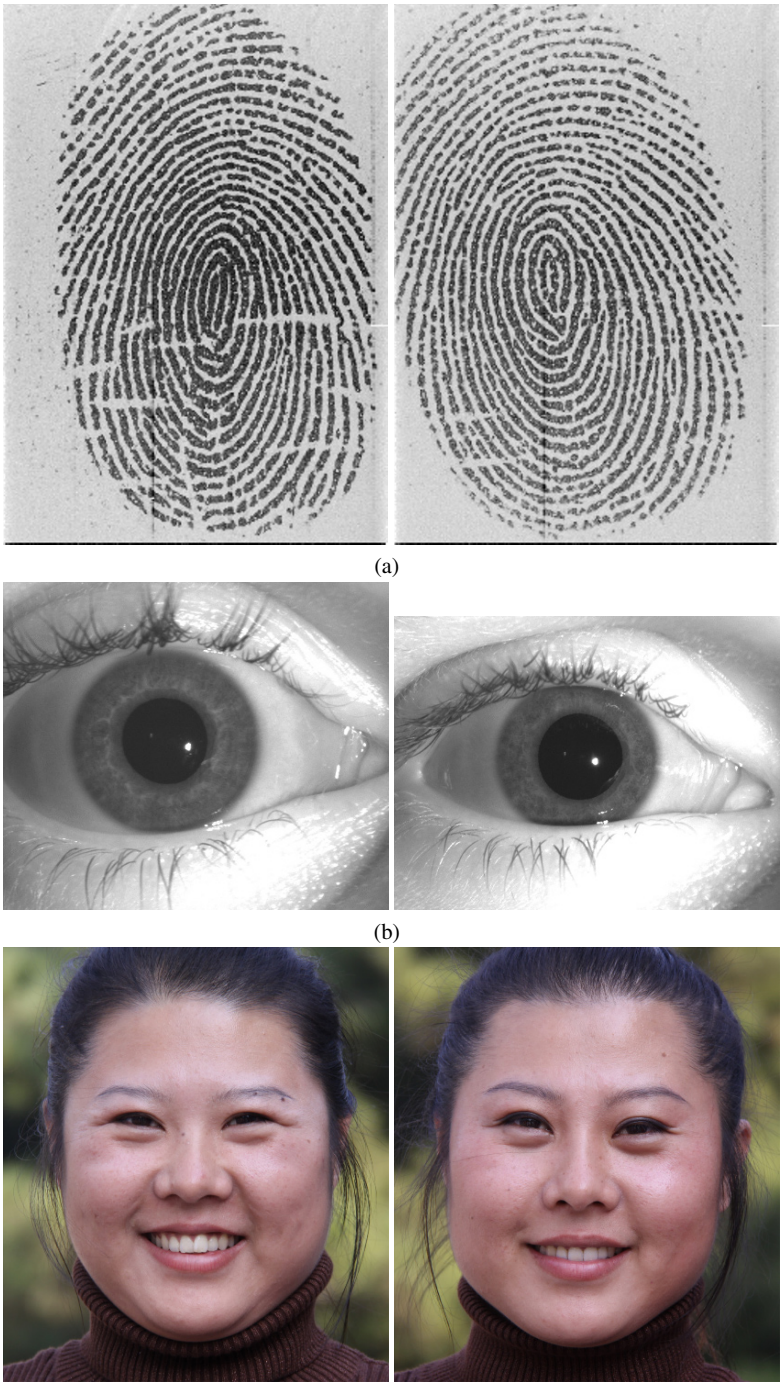


Fig. 1.8 Biometrics of twins. (a) The right index fingerprints of a pair of twins; (b) right eyes of a pair of twins; (c) face images of a pair of twins.

speaker recognition), changes in ambient conditions (e.g., inconsistent illumination levels in face recognition applications), and variations in the user's interaction with the sensor (e.g., occluded iris or partial fingerprints). As an illustration, consider two impressions of the same finger obtained on different days shown in Figure 1.9. These impressions differ in terms of the geometric distortions and amount of overlap caused by factors such as placement of finger on the sensor, applied finger pressure, skin condition, and feature extraction errors. Similarly, Figure 1.10 shows the intra-user variations in the face images of a person due to changes in pose and other attributes such as facial hair.

Fig. 1.9 Illustration of intra-user variability in fingerprints. Two different impressions of the same finger obtained on different days are shown with minutia points marked on them. Due to differences in finger placement and distortion introduced by finger pressure variations, the number and location of minutiae in the two images are different (33 and 26 in the left and right images, respectively). The number of corresponding/matching minutiae in the two images is only 16. Some of these correspondences have been indicated in the figure. The concept of minutia points will be described in the next chapter.

Intra-user variations are even more prominent in behavioral traits since the varying psychological makeup of an individual might result in vastly different behavioral characteristics at different time instances. For example, depending on the stress level of an individual, the voice sample presented by the person at the time of authentication may be significantly different from the enrolled template. Similarly, an inebriated person's gait and signature may be substantially altered.

Given the variability in the acquired biometric traits, it is factitious to expect a perfect match between any two biometric feature sets, even if they come from the same individual. In fact, if two feature sets are indeed identical, it may be a strong indication that the biometric data actually comes from an adversary who is replaying the data recorded at an earlier time. Therefore, there is a fundamental difference between password-based authentication systems and biometric systems.

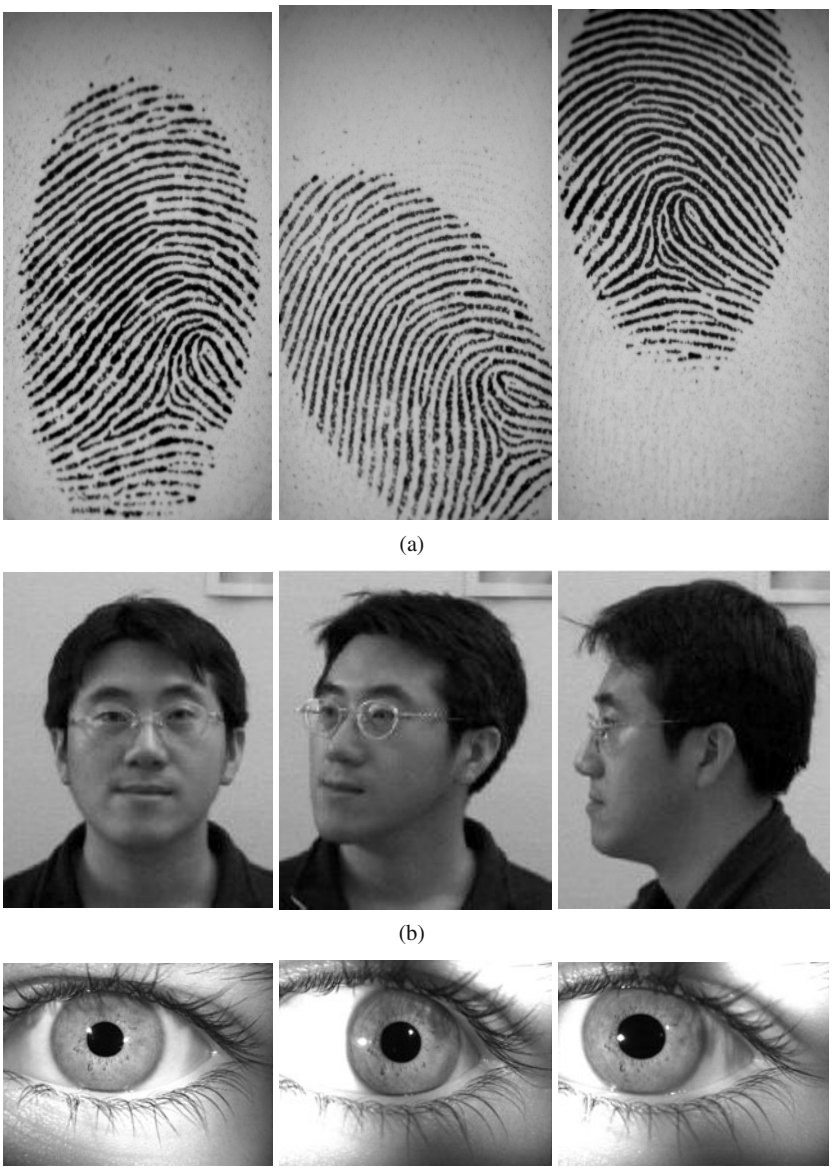


Fig. 1.10 Intra-user variations. (a) Variations in the fingerprints of the same person due to difference in position and orientation of the finger printed, (b) variations in the face images of the same person due to changes in pose, and (c) variations in the iris images due to difference in dilation and gaze direction.

In a password-based system, a *perfect* match between two alphanumeric strings is necessary to validate a user's identity. On the other hand, biometric systems mostly decide on a person's identity based on a *close* match between the template and the query, where the strength of the match (or the degree of similarity) is represented by the match score.

An ideal biometric feature set must exhibit small inter-user similarity and small intra-user variations. In practice, both these conditions may not be fully met either due to inherent *information limitation* (lack of uniqueness) in the underlying biometric trait or due to *representation limitation* (problems in feature extraction). Practical feature extraction systems, typically based on simplistic models of biometric data, fail to capture the richness of information in a realistic biometric input resulting in the inclusion of redundant or spurious features, and the exclusion of salient features.

Due to large inter-user similarity and large intra-user variations, a biometric system can make two types of errors, namely, *false non-match* and *false match*. When the intra-user variation is large, two samples of the same biometric trait of an individual (mate samples) may not be recognized as a match, and this leads to a false non-match error. A false match occurs when two samples from different individuals (non-mate samples) are incorrectly recognized as a match due to large inter-user similarity.

1.4.1 Performance measures

The basic measures of the accuracy of a biometric system are *False Non-Match Rate* (FNMR) and *False Match Rate* (FMR). FNMR refers to the expected probability that two mate samples (samples of the same biometric trait obtained from the same user) will be falsely declared as a non-match. FMR is the expected probability that two non-mate samples will be incorrectly recognized as a match.

A False Non-Match Rate of 5% indicates that on average, 5 in 100 authentication attempts by genuine users will not succeed. A majority of the false non-match errors are usually due to incorrect interaction of the user with the biometric sensor and can be easily rectified by allowing the user to present his biometric trait again. This scenario is similar to the case where the user in a password-based authentication system makes a mistake while entering a password and is allowed to re-enter the password.

A False Match Rate of 0.02% indicates that on average, 1 in $5,000$ authentication attempts by random impostors are likely to succeed. It is quite natural to consider how the FMR of a biometric system compares with the security provided by a password-based system. Consider a simple knowledge-based authentication system that uses a four-digit numeric PIN. Since a 4-digit PIN can take up $10,000$ different values, $5,000$ impostor attempts will be required, on average, to correctly guess the PIN. Does this mean that the security of a biometric system operating at 0.02% FMR is equivalent to the security provided by a 4-digit PIN? The answer is *no* because of two reasons. Firstly, it should be noted that the effective security provided by a

4-digit PIN is typically much less than 1 success in 5,000 impostor attempts, because most users tend to use numbers that are easy to remember (e.g., 1234, year of birth, etc.) and such PINs can be easily guessed by the adversary in a few attempts. Secondly, while a single adversary can theoretically attempt any number of guesses for a PIN, he has only a limited number of biometric samples (say ten fingers or two irides) that can be tried physically. To overcome this limitation, the adversary can make use of an off-line database of biometric samples or templates. However, in order to input these samples/templates, he must circumvent a physical component in the biometric system (sensor, feature extractor, or communication channels). This circumvention can be made very difficult by securing the physical infrastructure of the biometric system.

1.4.1.1 Verification system error rates

In the context of biometric verification, FNMR and FMR are generally referred to as False Reject Rate (FRR) and False Accept Rate (FAR), respectively. Strictly speaking, FMR and FNMR are not always synonymous with FAR and FRR, respectively. This is because while FNMR and FMR are measured as a proportion of the number of biometric matching attempts, FAR and FRR are application level metrics that measure the proportion of successful or failed transactions (a transaction may involve one or more matching attempts). However, in this book we treat them as being equivalent.

A match score is termed as a *genuine* or *authentic* score if it indicates the similarity between two mate samples. An *impostor* score measures the similarity between two non-mate samples. As discussed in section 1.3, a verification system makes a decision by comparing the match score s to a threshold η. Therefore, given a set of genuine and impostor match scores, FRR can be defined as the proportion of genuine scores that are less than the threshold η and FAR can be defined as the fraction of impostor scores that are greater than or equal to η.

Consider a scenario where the biometric data (e.g., right index fingerprint) corresponding to N users is acquired. Further, assume that each user is asked to provide t samples of their biometric data. To generate a genuine match score, a pair of samples from the same user have to be compared using the matcher; to generate an impostor match score, a pair of samples from two different users have to be compared. Thus, using this biometric data, a total of $Nt(t-1)/2$ genuine scores and $(N(N-1)t^2)/2$ impostor scores can be generated by the matcher. Here, it is assumed that the matcher is symmetric in the sense that comparison of sample A against B gives the same score as the comparison of B against A.

In the subsequent mathematical notations, we will use the labels ω_0 and ω_1 to denote the impostor and genuine classes, respectively. Let $p(s|\omega_1)$ and $p(s|\omega_0)$ be the probability density functions of the genuine and impostor scores, respectively. Figure 1.11 illustrates the genuine and impostor match score distributions corresponding to a face biometric system. The FAR and FRR of the biometric system are given by

system because the familiarity of users with the system can affect recognition accuracy since a habituated user is likely to provide good quality biometric data.

4. **Attended** versus **unattended operation**: Attended versus unattended classification refers to whether the process of biometric data acquisition in an application is observed, guided, or supervised by a human (e.g., a security officer). Furthermore, an application may have an attended enrollment operation but unattended recognition operation. For example, a banking application may have a supervised enrollment when an ATM card is issued to a user, but the subsequent uses of the biometric system for the ATM transaction are not attended.
5. **Controlled** versus **uncontrolled operation**: In a controlled environment, ambient environmental conditions such as temperature, pressure, moisture, lighting conditions, etc. can be moderated during the operation of a biometric system. Typically, indoor applications such as computer network login operate in a controlled environment, whereas outdoor applications such as keyless car entry or parking lot surveillance operate in an uncontrolled environment. This classification is also important for the system designer as a more rugged biometric sensor is needed for an uncontrolled environment.
6. **Open** versus **closed system**: If a person's biometric template can be used across multiple applications, the biometric system can be considered as open. For example, a user may use a fingerprint-based recognition system for entering secure facilities, computer network login, electronic banking, and bank ATMs. When all these applications use separate templates (databases) for each application, the system is considered closed. A closed system may be based on a proprietary template whereas an open system will need standard data formats and data compression methods to exchange and compare information between different systems (most likely developed by different commercial vendors).

All the above factors profoundly influence the design of a biometric system. Most of the commercial applications of biometrics, such as access to secure facilities, have the following attributes: verification, cooperative, overt, habituated, attended enrollment and non-attended authentication, and closed.

1.5.2 Choice of biometric trait

A number of biometric traits are being used in various applications. Each biometric trait has its pros and cons and, therefore, the choice of a biometric trait for a particular application depends on a variety of issues besides its recognition performance. In general, seven factors must be considered to determine the suitability of a physical or a behavioral trait to be used in a biometric application.

1. **Universality**: Every individual accessing the application should possess the trait. This factor determines the failure to enroll (FTE) rate of the biometric system.

2. **Uniqueness**: The given trait should be sufficiently different across individuals comprising the user population. Otherwise, the false match rate (FAR or FPIR) of the biometric system would be unacceptably high.
3. **Permanence**: The biometric trait of an individual should be sufficiently invariant over a period of time with respect to the matching algorithm. A trait that changes significantly over time is not a useful biometric because it will lead to a high false non-match rate (FRR or FNIR).
4. **Measurability**: It should be possible to acquire and digitize the biometric trait using suitable devices that do not cause undue inconvenience to the individual. Furthermore, the acquired raw data should be amenable to processing in order to extract discriminative feature sets. This factor significantly impacts the frequency of FTE and FTA failures and the recognition accuracy.
5. **Performance**: Apart from the recognition accuracy (FMR, FNMR, FTE, and FTA), the computational resources required to achieve that accuracy and the throughput (number of transactions that can be processed per unit time) of the biometric system should also meet the constraints imposed by the application.
6. **Acceptability**: Individuals in the target population that will utilize the application should be willing to present their biometric trait to the system.
7. **Circumvention**: This refers to the ease with which the trait of an individual can be imitated using artifacts (e.g., fake fingers), in the case of physical traits, and mimicry, in the case of behavioral traits. It also refers to the process of obfuscation, where a user deliberately alters his biometric trait to evade recognition.

No single biometric is expected to effectively meet all the requirements (e.g., accuracy, practicality, cost) imposed by all applications (e.g., forensics, access control, government benefits programs, etc.). In other words, no biometric is *ideal* but a number of them are *admissible*. The relevance of a specific biometric to an application is established depending upon the nature and requirements of the application, and the properties of the biometric characteristic. A brief introduction to some of the commonly used biometric characteristics (also shown in Figure 1.16) is given below:

1. **Fingerprint**: Humans have used fingerprints for personal identification for many decades. A fingerprint is the pattern of ridges and valleys on the surface of a fingertip whose formation is determined during the first seven months of fetal development. While fingerprints have been in use in forensic applications for over 100 years, the advent of low-cost and compact fingerprint scanners has spawned a large number of commercial applications in the past ten years. In applications requiring large-scale identification involving millions of identities, multiple fingerprints of a person (e.g., ten-prints used in Automated Fingerprint Identification Systems (AFIS)) can be used to improve the matching performance, though at the cost of more computational resources. Finally, fingerprints of a small fraction of the population may be unsuitable for automatic identification because of genetic factors, aging, environmental, or occupational reasons (e.g., manual laborers may have a large number of cuts and bruises on their fingerprints). Chapter 2 of this

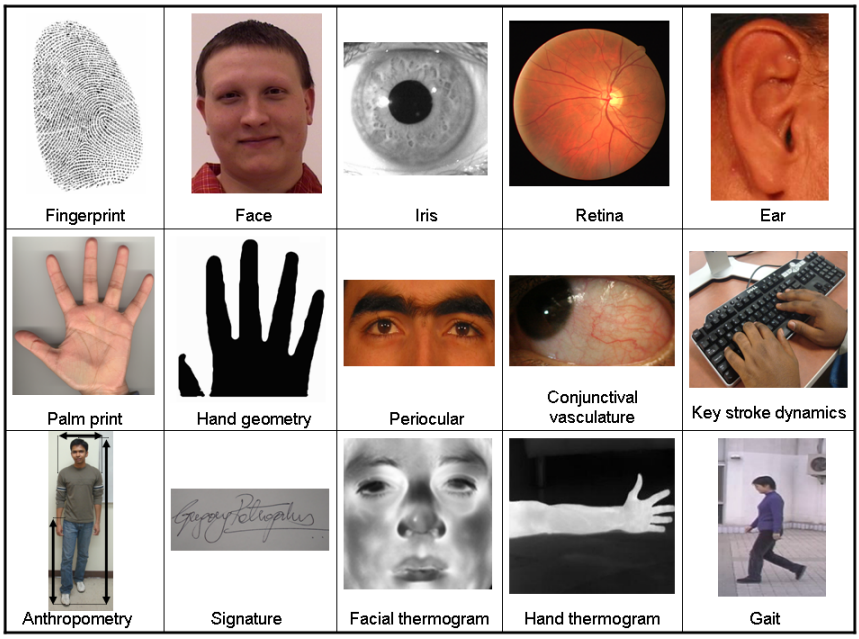


Fig. 1.16 A set of commonly used biometric traits.

book addresses the various issues concerning the design and implementation of automated fingerprint recognition systems.

2. **Palmprint**: The palms of the human hands contain pattern of ridges and valleys much like the fingerprints. The area of the palm is much larger than the area of a finger and, as a result, palmprints are expected to be even more distinctive than the fingerprints. Since palmprint scanners need to capture a large area, they are bulkier and more expensive than fingerprint sensors. Human palms also contain additional distinctive features such as principal lines and wrinkles that can be captured even with a lower resolution scanner, which would be less expensive. Finally, when using a high-resolution palmprint scanner, all the features of the hand such as geometry, ridge and valley features (e.g., minutiae and singular points such as deltas), principal lines, and wrinkles may be combined to build a highly accurate biometric system. There is a growing interest in palmprint matching, particularly latent palmprint matching, among the law enforcement agencies. More details about palmprints are presented in Chapter 2 along with the fingerprints.
3. **Iris**: The iris is the annular region of the eye bounded by the pupil and the sclera (white of the eye) on either side. The visual texture of the iris is formed during fetal development and stabilizes during the first two years of life (the pigmentation, however, continues to change over an extended period of time). The complex iris texture carries very distinctive information useful for personal recognition.

The accuracy and speed of currently deployed iris-based recognition systems is promising and can support large-scale identification. Although early generation iris-based recognition systems required considerable user participation and were expensive, the newer systems have become more user-friendly and cost-effective. Iris recognition is the primary focus of Chapter 4 of this book.

4. **Face**: Face recognition is a non-intrusive method, and facial attributes are probably the most common biometric features used by humans to recognize one another. The applications of facial recognition range from a static, controlled "mugshot" authentication to a dynamic, uncontrolled face identification in a cluttered background. While the authentication performance of the face recognition systems that are commercially available is acceptable for use in some applications, they impose a number of restrictions on how the facial images are obtained, often requiring a fixed and simple background with controlled illumination. These systems also have difficulty in matching face images captured from two different views, under different illumination conditions, and at different times. The design of face recognition systems and the associated challenges will be discussed in more detail in Chapter 3.
5. **Hand geometry (shape)**: Hand geometry recognition systems are based on a number of measurements taken from the human hand, including its shape, size of palm, and the lengths and widths of the fingers. Environmental factors such as dry weather or individual anomalies such as dry skin do not appear to adversely affect the authentication accuracy of hand geometry-based systems. However, the geometry of the hand is not known to be very distinctive and hand geometry-based recognition systems cannot be scaled up for systems requiring identification of an individual from a large population. Furthermore, hand geometry information may not be invariant during the growth period of children. In addition, individuals wearing jewelry (e.g., rings) or with limitations in dexterity (e.g., from arthritis), may pose challenges in extracting the correct hand geometry information. The physical size of a hand geometry-based system is large, and it cannot be embedded in certain devices like laptops. There are a few commercially available systems that are based on measurements of only a few fingers (typically, index and middle) instead of the entire hand. These devices are smaller than those used for hand geometry, but still much larger than those used for acquiring other traits like fingerprint, face, and voice. Hand shape biometric is discussed in Chapter 5.
6. **Gait**: Gait refers to the manner in which a person walks, and is one of the few biometric traits that can be used to recognize people at a distance. Therefore, this trait is very appropriate in surveillance scenarios where the identity of an individual can be covertly established. Most gait recognition algorithms attempt to extract the human silhouette in order to derive the spatio-temporal attributes of an individual who is walking. Hence, the selection of a good model to represent the human body is pivotal to the efficient performance of a gait recognition system. Some algorithms use the optic flow associated with a set of dynamically extracted moving points on the human body to describe the gait of an individual. Gait-based systems also offer the possibility of tracking an individual over an extended period of time. However, the gait of an individual is affected by sev-

eral factors, including the choice of footwear, nature of clothing, affliction of the legs, and walking surface. The segmentation problem is particularly severe for gait-based recognition. More details are provided in Chapter 5.

7. **Ear**: It has been suggested that the shape of the ear and the structure of the cartilegenous tissue of the pinna are distinctive. The ear recognition approaches are either based on matching the distance of salient points on the pinna from a landmark location on the ear or based on the appearance of the ear. Ear recognition could be useful for identifying a person based on a profile photograph. See Chapter 5 for more details on ear biometric systems.
8. **Voice**: Voice is a combination of physical and behavioral biometric characteristics. The physical features of an individual's voice are based on the shape and size of the appendages (e.g., vocal tract, mouth, nasal cavities, and lips) that are used in the synthesis of the sound. These physical characteristics of human speech are invariant for an individual, but the behavioral aspects of speech change over time due to age, medical condition (such as common cold), emotional state, etc. Voice is also not very distinctive and may not be appropriate for large-scale identification. A text-dependent voice recognition system is based on the utterance of a predetermined phrase. A text-independent voice recognition system recognizes the speaker independent of what she speaks. A text-prompted system prompts the user to repeat a phrase generated dynamically, which offers more protection against fraud. A disadvantage of voice-based recognition is that speech features are very sensitive to factors like background noise and microphone characteristics. Speaker recognition is most appropriate in telephone-based applications but the voice signal is typically degraded in quality by the communication channel.
9. **Keystroke**: It is hypothesized that each person types on a keyboard in a characteristic way. This biometric is not expected to be unique to each individual but it may be expected to offer sufficient discriminatory information to permit identity verification. Keystroke dynamics is a behavioral biometric; one may expect to observe large intra-class variations in a person's typing patterns due to changes in emotional state, position of the user with respect to the keyboard, type of keyboard used, etc. The keystrokes of a person could be monitored unobtrusively as that person is keying in information. This makes it possible to "continuously verify" an individual's identity over a session, after the person logs in using a stronger biometric such as fingerprint or iris.
10. **Signature**: The way a person signs her name is known to be a characteristic of that individual. Although signatures require contact with the writing instrument and an effort on the part of the user, they have been accepted in government, legal, and commercial transactions as a method of authentication. With the proliferation of PDAs, Tablet PCs, and smartphones, on-line signature may emerge as the biometric of choice in these devices. Signature is a behavioral biometric that changes over a period of time and is influenced by the physical and emotional conditions of the signatories. Signatures of some people vary substantially: even successive impressions of their signature are significantly different. Further, professional forgers may be able to reproduce signatures that can fool the signature verification system.

11. **DNA**: DNA refers to deoxyribonucleic acid that contains the genetic information necessary for the development and functioning of living organisms. DNA is the one-dimensional unique code for one's individuality - except for the fact that identical twins have the same DNA pattern. It is, however, currently used mostly in the context of forensic applications for suspect and victim identification. Three issues limit the utility of DNA for several other applications: (a) contamination and sensitivity: it is easy to steal a piece of DNA from an unsuspecting subject that can be subsequently abused for an ulterior purpose; (b) automatic real-time recognition issues: the state-of-the-art technology for DNA matching requires cumbersome chemical methods (wet processes) involving an expert's skills and is not yet geared for on-line non-invasive recognition; (c) privacy issues: information about susceptibilities of a person to certain diseases could be gained from the DNA pattern and there is a concern that the unintended abuse of genetic code information may result in social discrimination, e.g., in hiring practices.
12. **Facial, hand, and hand vein infrared thermograms**: The pattern of heat radiated by the human body is a characteristic of an individual and can be captured by an infrared camera in an unobtrusive way much like a regular (visible spectrum) photograph. The technology could also be used for covert recognition. A thermogram-based system does not require contact and is non-invasive, but image acquisition is challenging in uncontrolled environments, where heat emanating surfaces (e.g., room heaters and vehicle exhaust pipes) are present in the vicinity of the human body. A related technology using near infrared (NIR) imaging is used to scan the back of a clenched fist to determine hand vein structure. Infrared sensors are currently expensive, which is a factor inhibiting widespread use of the thermograms.
13. **Odor**: It is known that each object exudes an odor that is characteristic of its chemical composition and this could be potentially used for distinguishing various objects. A whiff of air surrounding an object is blown over an array of chemical sensors, each sensitive to a certain group of (aromatic) compounds. A component of the odor emitted by a human (or any animal) body is distinctive to a particular individual. It is not clear if the invariance in the body odor could be detected despite deodorant smells, and varying chemical composition of the surrounding environment.
14. **Retinal scan**: The retinal vasculature is rich in structure and is supposed to be distinctive for each individual and each eye. It is claimed to be the most secure biometric since it is not easy to change or replicate the retinal vasculature. The image acquisition requires a person to peer into an eye-piece and focus on a specific spot in the visual field so that a predetermined part of the retinal vasculature could be imaged. The image acquisition involves cooperation of the subject, entails contact with the eyepiece, and requires a conscious effort on the part of the user. All these factors adversely affect the public acceptability of the retinal biometric. Retinal vasculature can reveal some medical conditions, e.g., hypertension, which is another factor deterring the public acceptance of retinal scan-based biometrics.

Chapter 3
Face Recognition

"To ask how a human goes about face identification is probably an intractable question at present. But to ask how well and with what cues identification can occur is indeed a tractable question. So, too, is the matter of achieving effective machine identification and retrieval."

Goldstein, Harmon and Lesk, *Proceedings of the IEEE*, May 1971.

Human face images are useful not only for person recognition, but for also revealing other attributes like gender, age, ethnicity, and emotional state of a person. Therefore, face is an important biometric identifier in the law enforcement and human-computer interaction (HCI) communities. Detecting faces in a given image and recognizing persons based on their face images are classical object recognition problems that have received extensive attention in the computer vision literature. While humans are perceived to be good at recognizing familiar faces, the exact cognitive processes involved in this activity are not well-understood. Therefore, training a machine to recognize faces as humans do is an arduous task. However, general methods used in object recognition such as appearance-based, model-based, and texture-based approaches are also applicable to the specific problem of face detection and recognition. This chapter provides an overview of methods that have been developed for automated face recognition and discusses some of the challenges encountered by these systems.

3.1 Introduction

The face is the frontal portion of the human head, extending from the forehead to the chin and includes the mouth, nose, cheeks, and eyes. Being the foremost part in one's interactions with the outer world, the face houses most of the fundamental sensory organs necessary for perceiving the world around, namely, eyes for seeing, nose for smelling, mouth for tasting, and ears for hearing. The face is considered to be the most commonly used biometric trait by humans; we recognize each other and, in many cases, establish our identities based on faces. Hence, it has become a standard practice to incorporate face photographs in various tokens of authentication such as ID cards, passports, and driver's licenses.

A.K. Jain et al., *Introduction to Biometrics*, DOI 10.1007/978-0-387-77326-1_3,

Face recognition can be defined as the process of establishing a person's identity based on their facial characteristics. In its simplest form, the problem of face recognition involves comparing two face images and determining if they are of the same person. While humans seem to be adept in determining the similarity between two face images acquired under diverse conditions, the process of automated face recognition is beset with several challenges. Face images of a person may have variations in age, pose, illumination, and facial expressions (see Figure 3.1) as well as exhibit changes in appearance due to make-up, facial hair, or accessories (e.g., sunglasses). Training a machine to recognize face images exhibiting such unconstrained intra-user variations is a difficult task, especially since the exact cognitive and neural processes involved in humans for the task of face recognition (and recollection) is still not completely known. Moreover, there may be similarities between the face images of different persons (see Figure 3.2), especially if they are genetically related (e.g., identical twins, father and son, etc.). Such inter-class similarities further compound the difficulty of recognizing people based on their faces. Despite these challenges, significant progress has been made in the field of automated face recognition over the past two decades. Techniques for automated face recognition have been developed for the purpose of person recognition from still 2-dimensional (2D) images, video (a sequence of 2D images), and 3D range (depth) images.

The face modality has several advantages that make it preferable in many biometric applications. Firstly, unlike fingerprints, face can be captured at a longer stand-off distance using non-contact sensors. Hence, face is a suitable biometric identifier in surveillance applications. Secondly, the face conveys not only the identity, but also the emotions of a person (e.g., happiness or anger) as well as biographic information (e.g., gender, ethnicity, and age). The automated recognition of faces and associated emotions is necessary for designing interactive human-computer interfaces. Thirdly, there are large legacy face databases (e.g., U.S. driver's license repositories covers over 95% of the adult population), which enable large scale analysis of the face modality in terms of individuality or scalability. Finally, compared to other biometric traits like fingerprint and iris, people are generally more willing to share their face images in the public domain as evinced by the increasing interest in social media applications (e.g., Facebook) with functionalities like face tagging. Due to the above reasons, face recognition has a wide range of applications in law enforcement, civilian identification, surveillance systems, and entertainment/amusement systems. Figure 3.3 depicts some of these applications of face recognition.

3.1.1 Psychology of face recognition

Research in the fields of psychology and neuro-cognition indicates that certain parts of the brain are geared toward perceiving the face. Experiments have shown that humans find it difficult to detect or recognize faces that are inverted although they can perceive other inverted objects rather easily. Analysis of patients suffering from *prosopagnosia* (a disorder in which an individual loses his ability to recognize faces

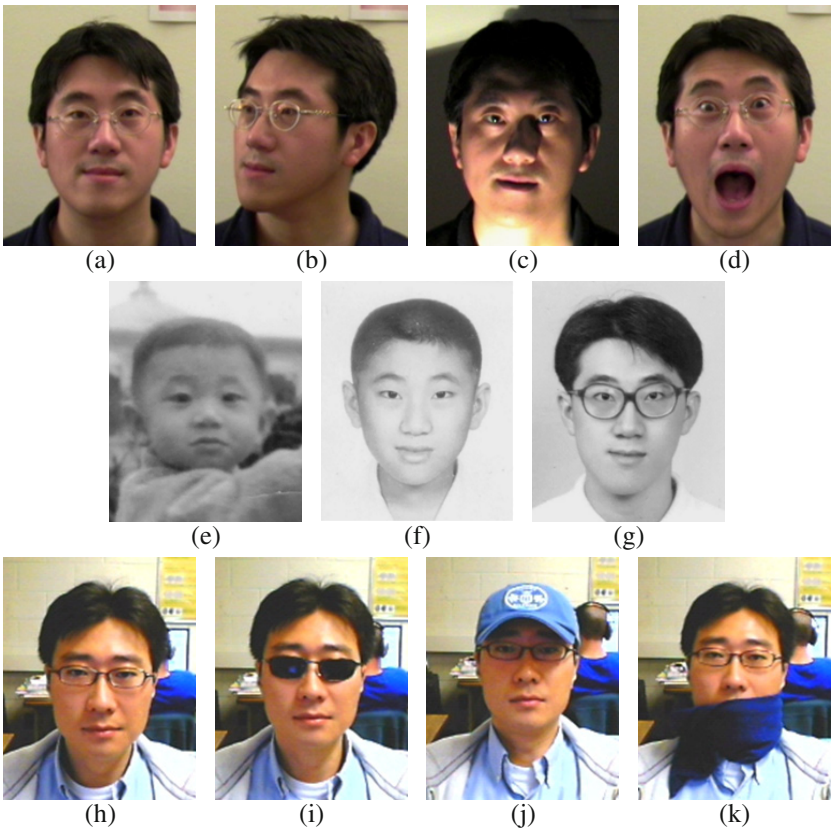


Fig. 3.1 The problem of intra-class (i.e., intra-user) variations is quite pronounced in the context of face recognition. The face image of an individual can exhibit a wide variety of changes that make automated face recognition a challenging task. For example, the face images in (b), (c), and (d) differ from the frontal face image of the person in (a) in terms of pose, illumination, and expression, respectively. The second row shows the variability introduced due to aging. Here, the images in (e), (f), and (g) were acquired when the person in (a) was 32, 21, and 15 years younger, respectively. The third row depicts the problem of occlusion of some facial features due to the person wearing accessories such as (h) prescription glasses, (i) sunglasses, (j) cap, and (k) scarf.

whilst retaining their ability to recognize other non-face objects) has shown that the loss in face recognition capability is caused by lesions in an area of the brain called the temporal cortex. This is also supported by a separate study that recorded an active response in the temporal cortex area of a monkey's brain when presented with images of faces.

The underlying mechanism of face perception in humans has been studied for two purposes: (a) to design machine recognition systems that can mimic the human ability to recognize faces and (b) to understand the neurological or psychological mechanism of brain functions for medical treatment. Because it is difficult to directly observe the brain functions related to face recognition, indirect observations

(a) Twin[1]

(b) Family[2]

[1] www.marykateandashley.com.
[2] news.bbc.co.uk/hi/english/in_depth/americas/2000/us_elections.

Fig. 3.2 The problem of inter-class similarity. The face images of some people (e.g., twins or families) exhibit similarities in appearance that can confound an automated face recognition system.

are commonly made to understand the mechanism supporting human face recognition. For example, based on the observations that human can recognize caricatures and cartoon faces, it is inferred that humans perceive the face based on certain higher-level characteristics. Studies using advanced brain imaging techniques such as functional magnetic resonance imaging (fMRI) are expected to reveal the precise face processing mechanism in the human brain.

3.1.2 Facial features

As indicated earlier, the face is composed of the forehead, eyebrows, eyes, nose, mouth, cheeks, and chin. Anthropometric studies have attempted to characterize the dimensions of the face based on a set of anatomically meaningful landmark or fiducial points. Figure 3.4 shows the representative landmark points used in several anthropometric studies. Anthropometric measurements have been used to study the growth patterns in humans as well as understand characteristics of the face as it pertains to gender and ethnicity. The forensic community has used these landmarks to identify face images. However, these measurements are not extensively used in automated face recognition systems due to their perceived lack of distinctiveness. Moreover, extracting these landmarks in poor quality face images may be challenging.

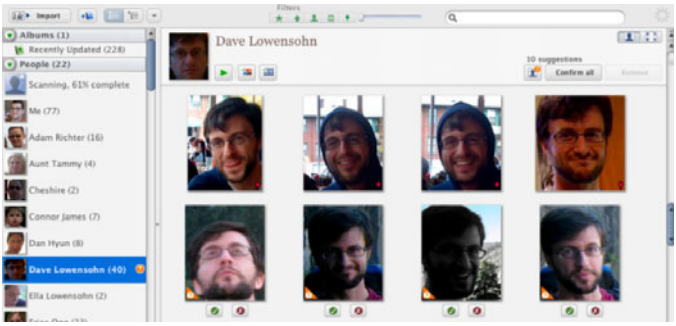
Similar to the case of fingerprints, the facial characteristics can be organized into the following three levels: (see Figure 3.5).

- **Level 1 details** consist of gross facial characteristics that are easily observable. Examples include the general geometry of the face and global skin color. Such features can be used to quickly discriminate between (a) a short round face and

(a)

(b)

(c)



(d)

Fig. 3.3 Applications of automated face recognition: (a) Australia's SmartGate system that facilitates faster immigration clearance for registered travelers; (b) Morpho's driver license solution, where face recognition can be used to prevent a single person obtaining multiple licenses under different names; (c) Microsoft's Kinect device has face recognition capabilities for the purpose of personalizing the XBOX 360 gaming system based on the player's identity; and (d) Google's Picasa and other social network websites offer automated face tagging functionality for easy management of personal photo albums.

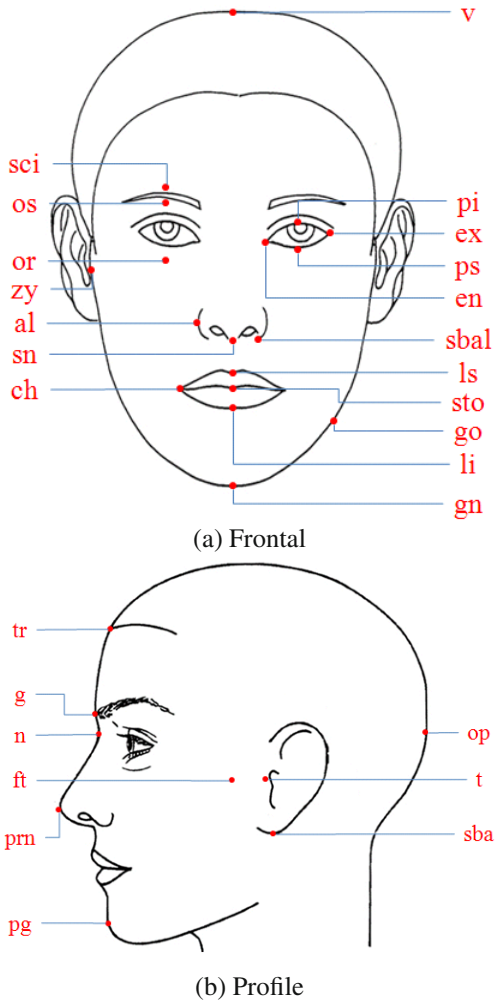


Fig. 3.4 Anthropometric facial landmarks on (a) frontal and (b) profile views of a face (Adapted from the Anthropometry of the Head and Face, 1994).

an elongated thin face; (b) faces exhibiting predominantly male and female characteristics; or (c) faces from different races. These features can be extracted even from low resolution face images ($<$ 30 interpupilary distance (IPD)[1]).

- **Level 2 details** consist of localized face information such as the structure of the face components (e.g., eyes), the relationship between facial components,

[1] IPD stands for Inter-Pupillary Distance, which represents the number of pixels between the centers of the two eyes in the given face image. IPD has been used to measure the image resolution in recent face recognition vendor tests.

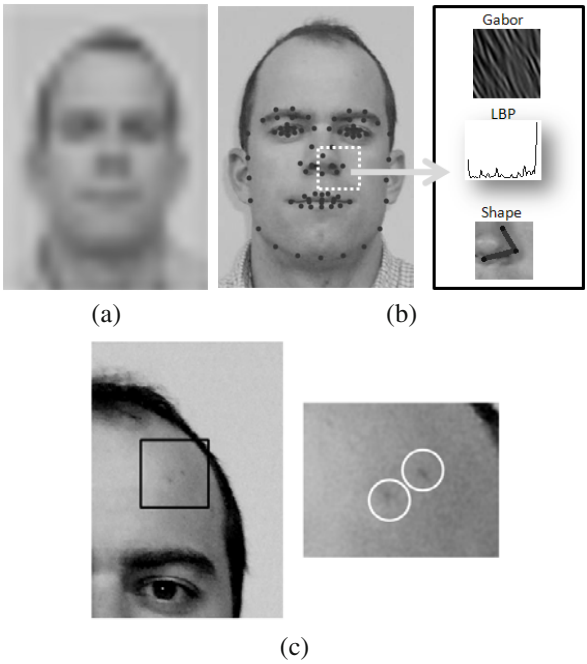


Fig. 3.5 Examples of the three levels of facial features. (a) Level 1 features contain appearance information that can be useful for determining ethnicity, gender, and the general shape of a face. (b) Level 2 features require detailed processing for face recognition. Information regarding the structure and the specific shape and texture of local regions in a face is used to make an accurate determination of the subject's identity. (c) Level 3 features include marks, moles, scars, and other irregular micro features of the face. This information is useful to resolve ambiguities when distinguishing identical twins, or to assist in forensic investigation scenarios.

and the precise shape of the face. These features are essential for accurate face recognition, and they require a higher resolution face image (30 to 75 IPD). The characteristics of local regions of the face can be represented using geometric or texture descriptors.

- **Level 3 details** consist of unstructured, micro level features on the face, which includes scars, freckles, skin discoloration, and moles. One challenging face recognition problem where Level 3 details may be critical is the discrimination of identical twins.

3.1.3 Design of a face recognition system

A typical face recognition system is composed of three modules: (a) image acquisition, (b) face detection, and (c) face matching (see Figure 3.6). The face image

acquired from a sensor can be categorized based on (a) the spectral band (e.g., visible, infrared, and thermal) used to record the image and (b) the nature of the image rendering technique (e.g., 2D, 3D, and video). Since most of the automated face recognition systems make use of 2D images acquired in the visible spectrum, much of this chapter will discuss the processing of this type of images. Face detection (also known as face localization or segmentation) refers to the process by which the face is located in an image and its spatial extent is determined. This task can be significantly challenging when the face object is located in a cluttered background or when multiple face images at different scales are available within the same image. Due to the distinctive characteristic patterns of eyes, most commercial face recognition engines first detect the two eyes prior to localizing the spatial extent of the face. Face detection in 3D images is considered to be an easier problem compared to 2D images because of the availability of depth information. In video streams, face detection can be made robust by tracking the detected faces over a sequence of images. Face matching is usually carried out by comparing the features extracted from the probe and gallery images.

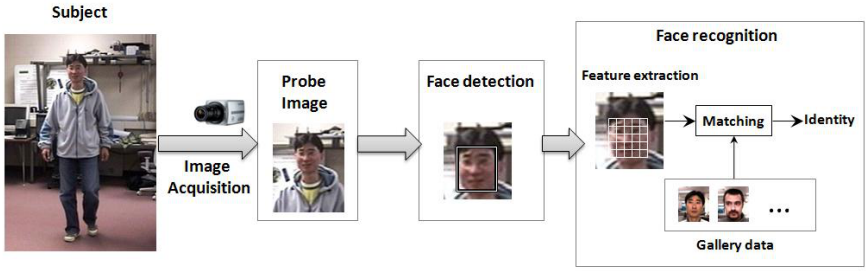


Fig. 3.6 Schematic of the face recognition process.

3.2 Image Acquisition

Automated face recognition requires the face data to be in a machine readable format. Conventional 2D photographs, 3D range or depth images, and videos are the three major types of image formats used in face recognition systems. Sensing technologies are continuously improving in order to increase the image resolution, capture more details by recording the face using multiple spectra (i.e., visible, infrared, and near-infrared), and facilitate real-time operation of 3D sensors.

3.2.1 2D Sensors

Until the development of sophisticated devices for capturing face information in 3D and invisible spectra, two-dimensional photographic images (also known as mugshot or still images) were the only source used by automated face recognition systems. Therefore, a large number of sensors and face recognition techniques have been developed for acquiring and processing 2D face images pertaining to the visible spectrum. Figure 3.7 shows some of the 2D cameras in use today.

(a) [1] (b) [2] (c) [3]

[1] http://www.unisa.edu.au/unisanews/2005/June/biometrics.asp.
[2] http://ws.sel.sony.com/PIPWebServices/RetrievePublicAsset/StepID/SEL-asset-61461/370x251.
[3] http://electrocctv.com/index.php?main_page=product_info&cPath=5_15&products_id=21.

Fig. 3.7 Examples of 2D camera systems. (a) Multiple 2D cameras capturing face images at three different viewpoints, (b) Sony EVI-D70 Pan-Tilt-Zoom camera, and (c) Sony long range infrared camera.

Since the face is a 3-dimensional object, 2D images of a face may occlude some of the facial features. This phenomenon is referred to as self-occlusion. In general, the frontal view of a face contains more details than a profile view and hence, matching frontal views of a face can be expected to provide more accurate person recognition. Multi-camera configurations that capture face images at multiple pose angles have been used to address the pose variation problem. Recognition systems based on 2D face images are also greatly affected by variations in illumination and spatial resolution. To overcome these challenges, new sensors such as high resolution cameras, active pan-tilt-zoom (PTZ) cameras, and infrared cameras are being used. Figure 3.8 shows example images captured in the visible and near-infrared (NIR) spectral bands. A NIR camera can operate even under low illumination conditions because it uses a separate NIR illuminator. Since the NIR illumination is not visible to the human eye, such a camera can be used for covert face acquisition in a dark (e.g., night) environment.

Typical face acquisition systems have a short operating distance limited to approximately 1-2 meters. When the subjects are observed at longer distance, the face

Fig. 3.8 Face images captured in the visible and near-infrared spectra at different wavelengths.

is captured at low resolution (see Figure 3.9), which may cause the face recognition process to fail. One approach to deal with the problem of low spatial resolution is to generate a higher resolution face image from the given low resolution image through a process called super-resolution. The other approach to improve the resolution of face images is to use high resolution cameras or PTZ cameras. A PTZ camera can dynamically zoom in or zoom out to obtain close-up images of objects of interest. However, the field of view of a PTZ camera is severely reduced when it zooms in to an object. Therefore, camera systems with paired static and PTZ cameras have emerged as a promising method to achieve zooming capability in a wide surveillance area. The static camera provides the wide field of view and then directs the PTZ camera to obtain high resolution images of target objects. Figure 3.10 shows an example face acquisition system with a pair of static and PTZ cameras and captured images from the static and PTZ cameras, respectively.

3.2.2 3D Sensors

The inherent pose, expression, and lighting variation problems in 2D face images stem from the 2D rendering of a 3D face object. Efforts in acquiring the face biometric in a 3D format have resulted in the development of 3D face capture systems. There are two types of 3D face capture systems: one is based on laser scanning and the other is based on stereographic reconstruction. It is generally regarded that laser scanners provide more accurate 3D face models, while stereographic cameras provide near real-time capture capability with slight loss in accuracy. Figure 3.11 shows some of the 3D sensors in use today.

The image captured by a 3D sensor typically covers about 120° of the human head and this image is referred to as a 2.5D scan. If a full 3D model of the face is required, it can be constructed by combining approximately three to five 2.5D scans captured from multiple views. 3D face models are usually represented as a polygonal mesh structure (e.g., triangular or rectangular) for computational efficiency (see

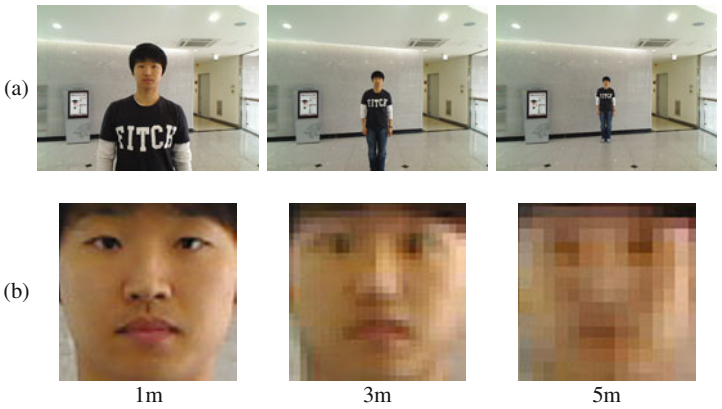
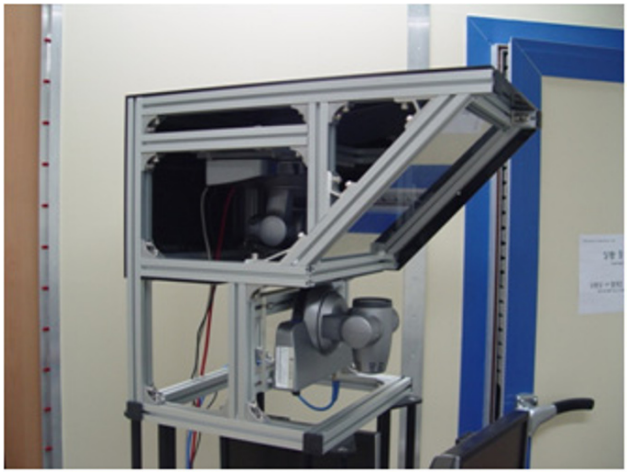


Fig. 3.9 Images recorded by a typical 2D camera with a resolution of 640×480 when the user is at three different distances from the camera (ranging from 1m to 5m). The first row (see (a)) shows the images as acquired by the camera, while the second row (see (b)) shows the face images obtained after face detection and resizing. The inter-pupillary distances (IPDs) are 35, 12, and 7 pixels for the face images obtained at 1m, 3m, and 5m, respectively. This example illustrates the steep decrease in the spatial resolution of the face images when the user is far away from the camera.

Figure 3.12). The 3D mesh structure changes depending on the preprocessing (e.g., smoothing, filling holes, etc.), mesh construction, and imaging process (scanning with laser sensor). Even though the 3D geometry of a face model changes depending on the pose, this change is very small and the model is generally regarded as being pose invariant. Further, the model is also robust to lighting variations. However, 3D face recognition is not invariant to changes in expression, aging, and occlusion. The drawbacks of 3D face acquisition include longer image acquisition time, the large data size of the 3D model (which requires higher computational resources during matching), and the relatively high price of 3D imaging sensors.

3.2.3 Video sequences

A video camera can continuously capture face images, thereby enabling the selection of a good quality face image (e.g., one with frontal pose and neutral expression) for recognition. The drop in price of video cameras has also made them a more viable solution for face recognition systems. Video-based sensors usually provide lower resolution images compared to still 2D sensors in order to handle the large amount of data streaming from the sensor to the storage or processing unit (30 frames per second in the National Television System Committee standard). Video compression techniques are also typically used to handle the large data stream, which can impact the quality of the images acquired in this mode.

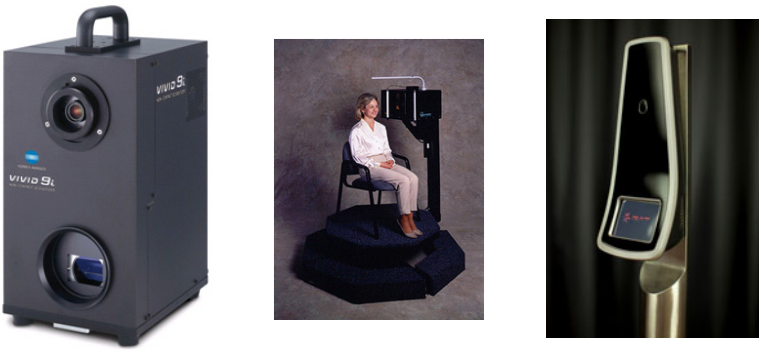


(a)

(b)

(c)

Fig. 3.10 An example of acquiring face images when the subject is far away from the camera. (a) A camera system consisting of a static and a Pan-Tilt-Zoom (PTZ) camera, (b) image of the person captured using the static camera as the person is walking into the room from the right, and (c) a close-up image of the person captured using the PTZ camera.



(a) [1] (b) [2] (c) [3]

[1] http://www.konicaminolta.com/instruments/about/index.html.
[2] http://www.alibaba.com/product-free/262924139/Head_Face_Color_3D_Scanner_Model/showimage.html.
[3] http://www.gat-solutions.sk/en/3d-fastpass.

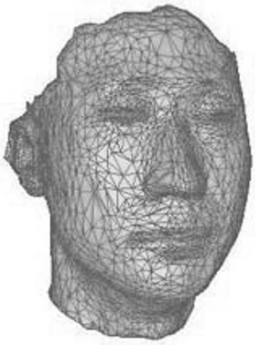
Fig. 3.11 Examples of 3D cameras. (a) Konica Minolta 3D Laser Scanner (VIVID 9i). (b) Cyberware Rapid 3D Scanner. (c) 3D FastPass^{TM} Face Reader from L-1 Identity solutions.

There is a significant interest in developing robust face recognition systems that will accept video streams as an input. Face recognition in video has attracted interest due to the widespread deployment of surveillance cameras. The ability to automatically recognize faces in real-time from video will facilitate, among other things, a covert method for human identification using an existing network of surveillance cameras. Two distinctive pieces of information are provided by a video stream: (a) multiple frames of the same subject and (b) temporal information pertaining to an individual's face. Multiple frames typically depict a variety of poses, allowing for the proper selection of a good quality frame (namely, a high quality face image in near-frontal pose) for superior recognition performance. The temporal information in video corresponds to the dynamic facial motion in the video. However, it is difficult to determine whether there are any identity-related details in facial motion (research in psychology has indicated that facial motion has some discriminatory information that may be useful in establishing identity).

3.3 Face Detection

Face detection is the first step in most face-related applications including face recognition, facial expression analysis, gender/ethnicity/age classification, and face modeling. Variations in pose and expression, diversities in gender and skin tone, and occlusions (e.g., due to glasses) are the typical challenges confounding face detection. While there are a number of approaches for detecting faces in a given image, state-of-the-art face detection methods are typically based on extracting local texture

(a)



(b)

Fig. 3.12 (a) A full 3D model of the human face obtained using a 3D sensor. (b) The 3D face model represented using triangular meshes.

features from the given image and applying a binary (two-class) classifier to distinguish between a face and non-face. This approach follows the seminal work done by Viola and Jones in the field of real-time object detection. The face detection technique proposed by Viola and Jones has been widely used in various studies involving face processing because of its real-time capability, high accuracy, and availability as open-source software under the Open Computer Vision Library (OpenCV). However, the Viola-Jones face detector is not perfect and can produce both false positive and false negative errors as shown in Figure 3.13. A false positive error refers to the detection of a face where none exists, while a false negative error indicates that a face present in the image was not detected.

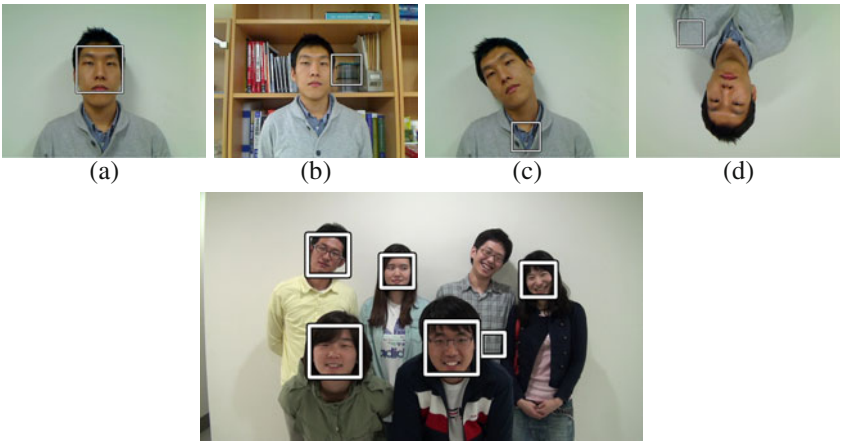


Fig. 3.13 The problem of face detection involves detecting a face in an image. Face detection algorithms have to be robust to variations in illumination, background, rotation, and image resolution. In this figure, the output of the Viola-Jones face detection algorithm, as implemented in the Open Computer Vision Library (OpenCV), is shown for different scenarios: (a) simple background, (b) cluttered background, (c) tilted face, (d) inverted face, and (e) multiple faces. Figures (b) through (e) have both false negatives (faces that are not detected) and false positives (non-face regions are wrongly categorized as faces).

3.3.1 Viola-Jones face detector

The Viola-Jones face detector scans through the input image with detection windows of different sizes and decides whether each window contains a face or not. Figure 3.14 shows the scanning process ranging from small to large windows. In each window, the existence of a face candidate is decided by applying a classifier to simple local features derived using rectangular filters. These rectangular filters can be grouped as two-rectangle, three-rectangle, and four-rectangle filters as shown in

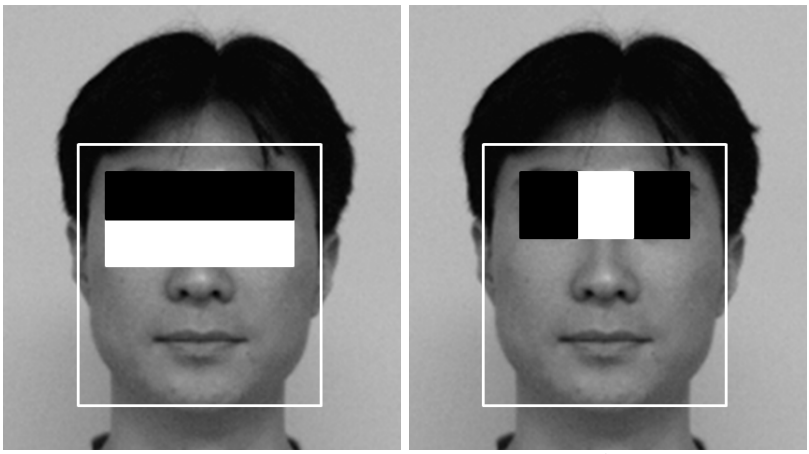


Fig. 3.17 Two most discriminative Haar-like features overlaid on an input image.

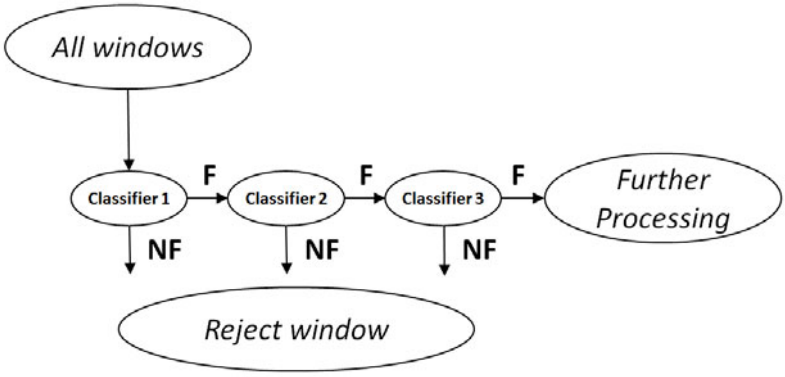


Fig. 3.18 Schematic of a cascaded classifier to speed-up the face detection process. The initial classifier uses only a few features and eliminates a large number of non-faces with minimal processing. Subsequent classifiers consider increasingly more features and further eliminate the remaining non-faces at the cost of additional processing. Here, F indicates that a classifier decides that the tested window contains a face candidate, while NF represents the classifier decision that there is no face candidate within the tested window.

3.4 Feature Extraction and Matching

There are three main approaches to match the detected face images (see Figure 3.20): appearance-based, model-based, and texture-based methods.

- *Appearance-based techniques* generate a compact representation of the entire face region in the acquired image by mapping the high-dimensional face image

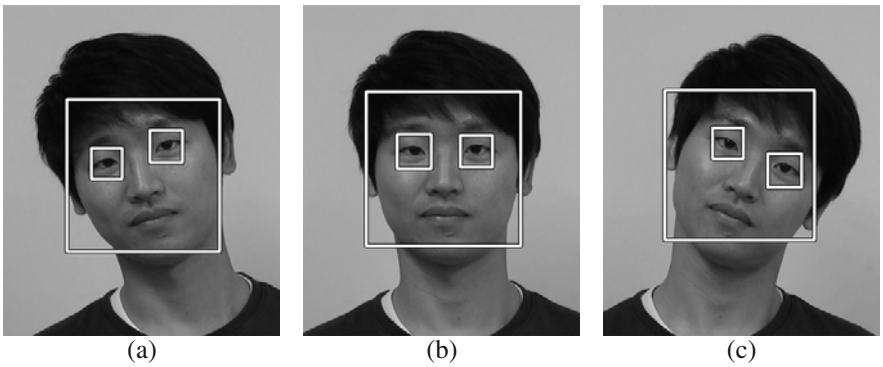


Fig. 3.19 Examples of face and eye detection using the Viola-Jones object detection approach.

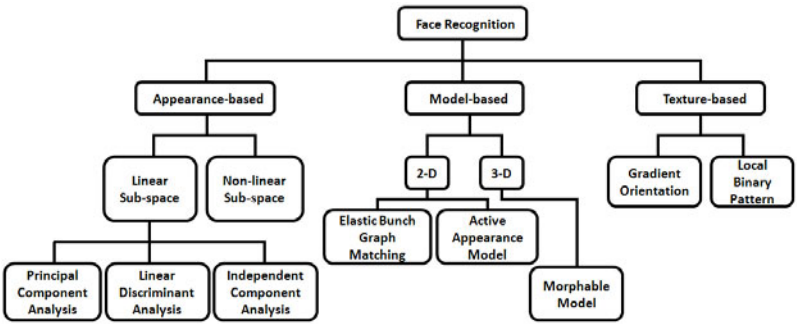


Fig. 3.20 Categorization of face recognition techniques.

into a lower dimensional sub-space. This sub-space is defined by a set of representative basis vectors, which are learned using a training set of images. Though the mapping can be either linear or non-linear, commonly used schemes such as Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), and Independent Component Analysis (ICA) involve linear projections.

- *Model-based techniques* attempt to build 2D or 3D face models that facilitate matching of face images in the presence of pose variations. While the Face Bunch Graphs (FBG) and Active Appearance Model (AAM) are examples of 2D face models, the morphable model is a 3D model.
- *Texture-based approaches* try to find robust local features that are invariant to pose or lighting variations. Examples of such features include gradient orientations and Local Binary Patterns (LBP).

More recently, schemes that make use of 3D models, video input, and micro level details (e.g., freckles, moles, scars) have been developed to improve the accuracy of face recognition systems. While this section describes some of the representa-

tive schemes for matching 2D still images, some of the recent developments are discussed in Section 3.5.

3.4.1 Appearance-based face recognition

Appearance-based schemes are based on the idea of representing the given face image as a function of different face images available in the training set, or as a function of a few basis faces. For example, the pixel value at location (x,y) in a face image can be expressed as a weighted sum of pixel values in all the training images at (x,y). The set of training images or basis faces forms a subspace and if the given face image is linearly projected onto this subspace, it is referred to as linear subspace analysis. The challenge here is to find a suitable low dimensional subspace that preserves the discriminatory information contained in the face images. In other words, the goal in linear subspace analysis is to find a small set of most representative basis faces. Any new face image can be represented as a weighted sum of the basis faces and two face images can be matched by directly comparing their vector of weights.

3.4.1.1 Principal Component Analysis

Principal Component Analysis (PCA) is one of the earliest automated methods proposed for face recognition. PCA uses the training data to learn a subspace that accounts for as much variability in the training data as possible. This is achieved by performing an Eigen value decomposition of the covariance matrix of the data. Specifically, PCA involves the following five steps.

1. Let $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_N$ be the training set, where each $\mathbf{x}_i$ represents a d-dimensional column vector. Compute the average of the training set as

$$\boldsymbol{\mu} = \frac{1}{N} \sum_{i=1}^{N} \mathbf{x}_i. \tag{3.3}$$

2. Define the data matrix $\mathbf{X}$ as follows: $\mathbf{X} = [(\mathbf{x}_1 - \boldsymbol{\mu})\ (\mathbf{x}_2 - \boldsymbol{\mu})\ \cdots\ (\mathbf{x}_N - \boldsymbol{\mu})]$.
3. Calculate the data covariance matrix as

$$\mathbf{C} = \mathbf{X}\mathbf{X}^T, \tag{3.4}$$

   where $\mathbf{X}^T$ is the transpose of matrix $\mathbf{X}$. Since $\mathbf{X}$ is a $d \times N$ dimensional matrix, the size of the covariance matrix $\mathbf{C}$ is $d \times d$.
4. Compute the Eigen vectors of the covariance matrix $\mathbf{C}$ by solving the following Eigen system.

$$\mathbf{C}\mathbf{E} = \lambda \mathbf{E}. \tag{3.5}$$

Chapter 7
SECURITY OF BIOMETRIC SYSTEMS

"Security is, I would say, our top priority because for all the exciting things you will be able to do with computers.. organizing your lives, staying in touch with people, being creative.. if we don't solve these security problems, then people will hold back. Businesses will be afraid to put their critical information on it because it will be exposed."

Bill Gates (2005)

The primary reasons for using biometric recognition are to apprehend criminals, curtail financial fraud, secure national borders, or control access to physical facilities and logical resources. When the biometric system fails to meet these objectives, the security of the system is said to be breached. This breach of security can be in the form of denial-of-service to legitimate users, intrusion by unauthorized users, repudiation claims by authorized users, or misuse of the biometric data for unintended purposes. Security failures can occur either due to intrinsic limitations of the biometric system or due to explicit attacks by adversaries, who may be insiders (e.g., administrators and legitimate users) or external attackers. The objective of this chapter is to outline the common attacks against biometric systems and discuss techniques that can be employed to counter them. In particular, this chapter will focus on two of the most well-known attacks that are specific to biometric systems, namely, spoofing of biometric traits and leakage of biometric data. Liveness detection and biometric template security algorithms that can mitigate the above two threats will be discussed in detail.

7.1 Introduction

A natural question that arises in biometric recognition is which biometric system is "best" suited for a particular application. Of course, the answer to this question depends not only on technical merits and limitations of the biometric system (e.g., matching accuracy and throughput), but also on other socio-economic factors like user acceptability and system cost. However, given that all other factors are equal, one would obviously prefer a biometric system that has the least probability of failure. But what exactly constitutes a biometric system failure? Recall that in most applications, the primary purpose of using biometrics is to provide non-repudiable

A.K. Jain et al., *Introduction to Biometrics*, DOI 10.1007/978-0-387-77326-1_7,

authentication. Authentication implies that (a) only legitimate or authorized users are able to access the physical or logical resources protected by the biometric system and (b) impostors are prevented from accessing the protected facilities or information. Non-repudiation ensures that an individual who accesses a certain resource cannot later deny using it. Thus, the *integrity* of a biometric system is determined by its ability to guarantee non-repudiable authentication.

From the perspective of the users, there are two additional requirements that a biometric system must meet. Firstly, the legitimate users must have timely and reliable access to the protected resource/service. This is referred to as the *availability* of the biometric system. Secondly, the biometric system and the personal data stored in it must be used only for the intended functionality, which is to control access to a specific resource and not for other unintended purposes. This is known as the *confidentiality* requirement. When one or more of the above three expectations (integrity, availability, and confidentiality) are not met, the biometric system is deemed to have failed.

Failure of a biometric system generally leads to a breach of security in applications or facilities that it is designed to protect. A security threat in a biometric system refers to the possibility of system failure. Depending on the type of failure, these security threats can be classified into four major classes (see Figure 7.1).

- **Denial-of-service (DoS)**: Legitimate users are prevented from obtaining access to the system or resource that they are entitled to, thereby causing inconvenience to genuine users. This violates the availability requirement. Frequent denial-of-service is likely to eventually drive the users towards abandoning the biometric system altogether.

- **Intrusion**: An unauthorized user gains illegitimate access to the system. Since intrusion affects the basic integrity of a biometric system, it is generally considered the most serious security threat.

- **Repudiation**: A legitimate user denies using the system after having accessed it. Corrupt users may deny their actions by claiming that illegitimate users could have intruded the system using their identity.

- **Function creep**: An adversary exploits the biometric system designed to provide access control to a certain resource to serve another application, which the system was never intended to perform. For example, a fingerprint template obtained from a bank's database may be used to search for that person's health records in a medical database. This violates the confidentiality requirement. Although the problem of function creep has been posed primarily as a security threat, it is also widely perceived as a major threat to user privacy.

Public confidence and acceptance of biometric technology will depend on the ability of system designers to guard against all possible security threats. However, no system is likely to be absolutely secure and foolproof. Given the right circumstances and plenty of time and resources, any security system can be broken. Even though

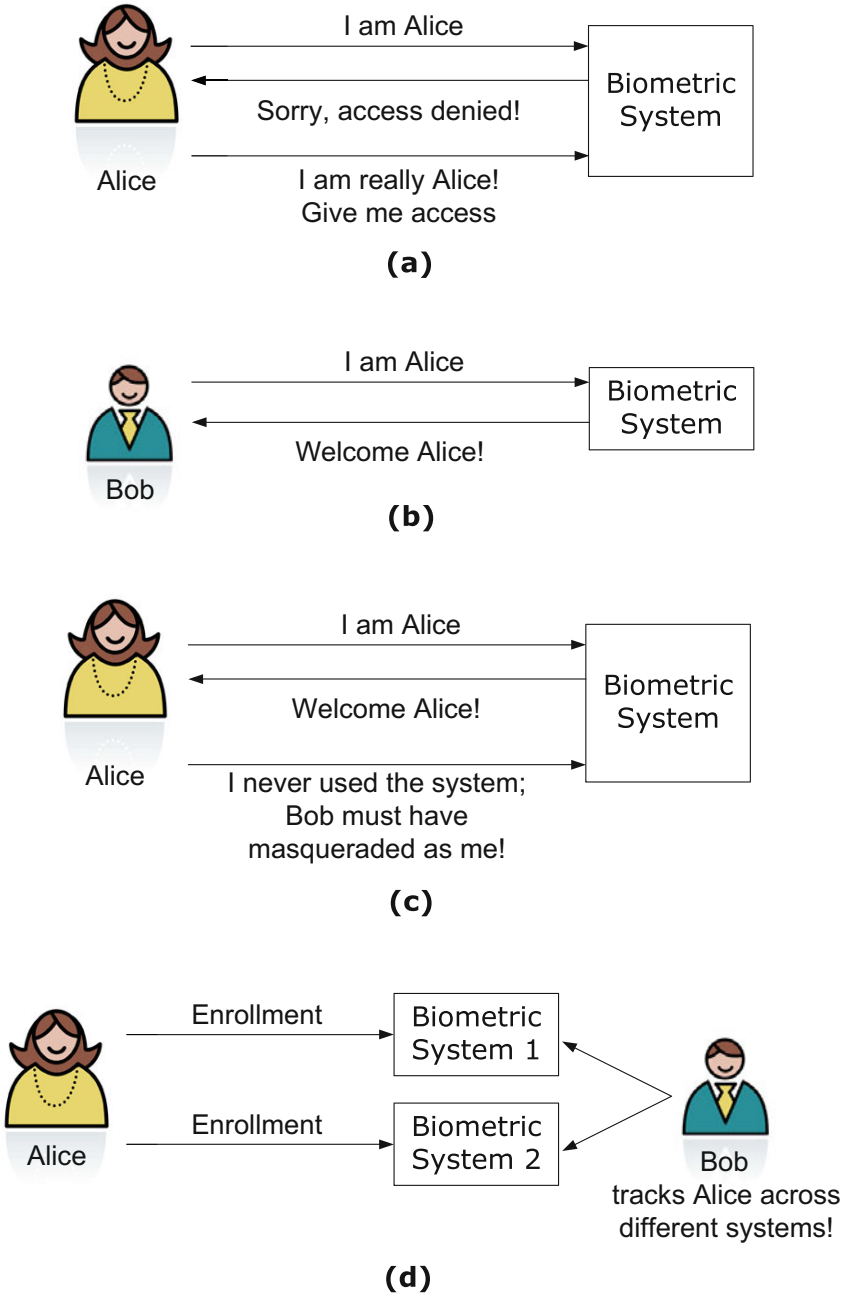


Fig. 7.1 Four major classes of security threats in a biometric system. (a) Denial of service, (b) Intrusion, (c) Repudiation, and (d) Function creep.

7.3 Attacks at the User Interface

In general, any attempt by an attacker to break into the system by presenting a biometric trait can be considered as an attack at the user interface level. At this level, the following attacks and countermeasures are possible.

7.3.1 Impersonation

This refers to the situation where an impostor attempts to intrude the system by posing himself as another authorized user. The impersonation could be either casual or targeted. In *casual impersonation*, the identity to be attacked is chosen randomly and the impostor does not modify his/her own biometric identifiers in any way. The probability of success in such an attack is usually measured by the false match rate (FMR) of the biometric system. This attack can be countered by selecting a very low value of FMR and by restricting the number of failure attempts allowed within a time-frame.

Targeted impersonation occurs when the impostor attacks a specific identity enrolled in the biometric system, which is known to be easier to impersonate (also known as a "lamb" in the Doddington's Zoo). This attack exploits the fact that FMR is not uniform across all users. The impostor may also target an identity whose biometric characteristics are known to be similar to his traits (also known as "Evil Twin" attack). The same countermeasures used against casual impersonation may be employed to limit the success of this type of attack.

Finally, the impostor may also be able to modify his biometric characteristics to match that of the identity under attack. A common name for such an attack is *mimicry*. Examples of this attack include changing one's voice, forging a signature (see Figure 7.4), or mimicking a gait pattern. This threat is more common in systems using behavioral biometric traits and in applications with unattended mode of operation. Countering this attack requires biometric systems that have low false match rate (FMR) under skilled forgery.

(a) (b)

Fig. 7.4 Example of a mimicry attack. (a) Genuine signature samples of a person, (b) skilled forgeries of the signature in (a) created by impostors. (Source: BioSecure Association)

7.3.2 Obfuscation

Any deliberate attempt by an attacker to change his biometric characteristic in order to avoid detection by the biometric system is called obfuscation. Thus, the key difference between mimicry and obfuscation is the motivation behind the attack. Obfuscation is mainly applicable in negative recognition applications, where the attacker wants to hide his true identity. However, it may also be applicable in verification systems that employ a fall-back mechanism to handle false rejects. In this scenario, the adversary may attempt to bypass the biometric system by forcing a false reject decision and then exploit the loopholes in the fall-back mechanism, which may be easier to circumvent.

Obfuscation can be done in a number of different ways. One possibility is to intentionally present a poor quality image or noisy biometric sample (e.g., face with non-neutral expression or a partially open eye) that may not be matched to his/her template in the database. In the case of face recognition, use of makeup, facial hair, and glasses can also lead to a false non-match. Fingerprints can be obliterated through techniques like abrasion, cutting, and burning, or may even be surgically altered or distorted (see Figure 7.5). Similarly, face can be altered using plastic surgery and iris transplants have been depicted in popular science fiction (e.g., in the movie Minority Report). Knowledge of the details of biometric processing algorithms can further facilitate such attacks. For example, if the attacker knows that a particular face recognition system is not robust to pose variations, he can easily circumvent it by presenting only a profile view of the face.

The most effective solution against obfuscation is to improve the robustness of biometric algorithms to intra-user variations in order to achieve a very low false non-match rate (FNMR). It may also be possible to automatically detect some of the alterations such as a non-frontal face or surgically modified fingerprint and subject such users to secondary inspection.

7.3.3 Spoofing

This is the most well-known attack at the user interface level, and it involves the presentation of a spoof biometric trait. A spoof is defined as any counterfeit biometric that is not obtained from a live person (see Figure 7.6). Spoofing includes the presentation of fake or artificial traits (e.g., gummy finger, thin film on top of a finger, photograph or mask of a face, recorded voice, etc.) and things as sinister as dismembered body parts (e.g., a dismembered finger) belonging to a legitimate user to the recognition system. If the sensor is unable to distinguish between spoofed and genuine biometric traits, the adversary can easily intrude the system under a false identity.

This attack requires knowledge of the biometric trait corresponding to the identity to be attacked. This knowledge could be obtained in one of the following four ways: (a) directly colluding with or coercing an authorized user, (b) covert acqui-

sition (e.g., lifting residual fingerprint impressions covertly from the sensor or any surface touched by the authorized user, recording the user's voice, or capturing a photograph of the user's face), (c) estimating a close approximation of the user's biometric template through brute-force or hill-climbing attacks, and (d) stealing the biometric template from a database and reverse engineering the template.

While traditional password-based authentication systems work under the assumption of secrecy (i.e., only the legitimate user knows his password), such an assumption is generally not required for a biometric system to work. In contrast, the strength of biometric authentication is derived from the fact that the biometric characteristic is linked to the user physically. Though an attacker may get hold of a legitimate user's fingerprint pattern, it would not be of much use to the attacker if the sensor can ensure that the scanned fingerprint comes directly from the finger of a live user. Therefore, the solution to counter spoof attacks is to incorporate liveness detection capability in the biometric sensor.

7.3.4 Countermeasure: spoof detection

Spoof detection can be broadly defined as differentiating a real biometric trait presented by a live person from a biometric trait presented through any other source. Spoof detection typically involves checking for signs of human vitality or liveness (e.g., blood pulse), a process known as liveness detection. Despite this subtle difference between spoof detection and liveness detection, the two terms are generally used interchangeably in the biometrics literature, which is also the case in this book. Spoof detection can either be decoupled from or integrated into the biometric recognition process. In a decoupled system, no biometric data is acquired until the spoof detection system is convinced that the biometric trait is presented by a live human user. On the other hand, an integrated system detects the spoof while processing the acquired biometric information (either prior to or during feature extraction).

The susceptibility of a biometric system to a spoof attack depends both on the biometric modality and the specific sensor used to capture the biometric trait. For example, a two-dimensional photograph of a human face may be sufficient to fool a camera used in a face recognition system. However, it is usually very difficult to circumvent an optical or capacitive fingerprint sensor by using a 2-D reproduction of a fingerprint because such sensors inherently depend on capturing the 3-D variations in the ridge-valley structures.

While spoof detection is extremely important to ensure the integrity of a biometric system, it also brings in a few disadvantages. Firstly, almost all spoof detection solutions increase the cost of the biometric system. This is because of the need to have additional hardware to capture new information (e.g., spectral or thermal properties) or a software module to process the biometric data already collected and distinguish between a spoof and a live trait. This additional processing also increases the biometric acquisition time, thereby reducing the throughput of the biometric system. Finally, just like biometric systems that are seldom perfect, spoof detec-

tion systems are also prone to errors. While a spoof detection system may identify and thwart most of the spoofing attempts, it may also incorrectly classify a few real biometric traits as spoofs, leading to an increase in the failure to capture rate.

Though there are a number of biometric spoof detection algorithms, they can be classified into three main groups based on the mechanism employed for thwarting a spoof attempt. The first approach involves measuring the physiological properties of a live person, which includes blood pulse/pressure, perspiration, spectral/optical properties of the human skin/tissues, electrical/thermal characteristics, and deformation of the muscles/skin. The second approach is based on identifying voluntary or involuntary human behavioral actions like fluctuations in pupil size, blinking, and pupil/eye/head/body movements. The third category is known as the challenge-response mechanism, where the system presents a challenge to the user and measures whether the user responds to the challenge correctly. Examples of challenges include prompting a user to recite a randomly generated phrase/text, asking the user to change his or her facial expression (e.g., smile or frown), and requesting the user to present multiple biometric traits in a randomly generated sequence. Since the last two approaches are fairly straightforward to envision and implement, only the first approach will be discussed in detail.

7.3.4.1 Spoof detection based on physiological properties

While biometric systems are based on physiological characteristics that are unique to each individual (e.g., fingerprint, iris, face), spoof detection algorithms tend to use characteristics that can easily distinguish a human body from innate materials (e.g., silicone gel for fingerprints) used for spoofing. Some of the physiological properties that have been used for spoof detection are discussed below.

- **Pulse rate/ Blood pressure**: This property is generally applicable to biometric traits such as fingerprint and palmprint that require the user to be in physical contact with the sensor. While the pulse rate is a good vitality sign, special hardware may be needed to record this trait. Moreover, the pulse rate and blood pressure vary significantly from one person to another and also within the same person depending on his physical activity and emotional state at the time of acquisition. Furthermore, a single pulse measurement may take up to five seconds. Finally, if a wafer-thin silicone rubber is glued to a real finger, the heartbeat of the underlying finger will result in the detection of a pulse.

- **Perspiration**: Perspiration refers to the sweating process of a live finger. Live fingers exhibit sweating over a period of time whereas fake fingers will not exhibit the sweating process. The perspiration phenomenon starts at the sweat pores on a fingerprint and spreads along the ridge lines, while the valleys do not change. Due to the sweating process in live fingers, the regions around sweat pores can be seen to enlarge over time in a sequence of fingerprint images (see Figure 7.7). One limitation of this procedure to detect a spoof finger is that to observe the sweating process, the finger needs to stay on the fingerprint scanner for a

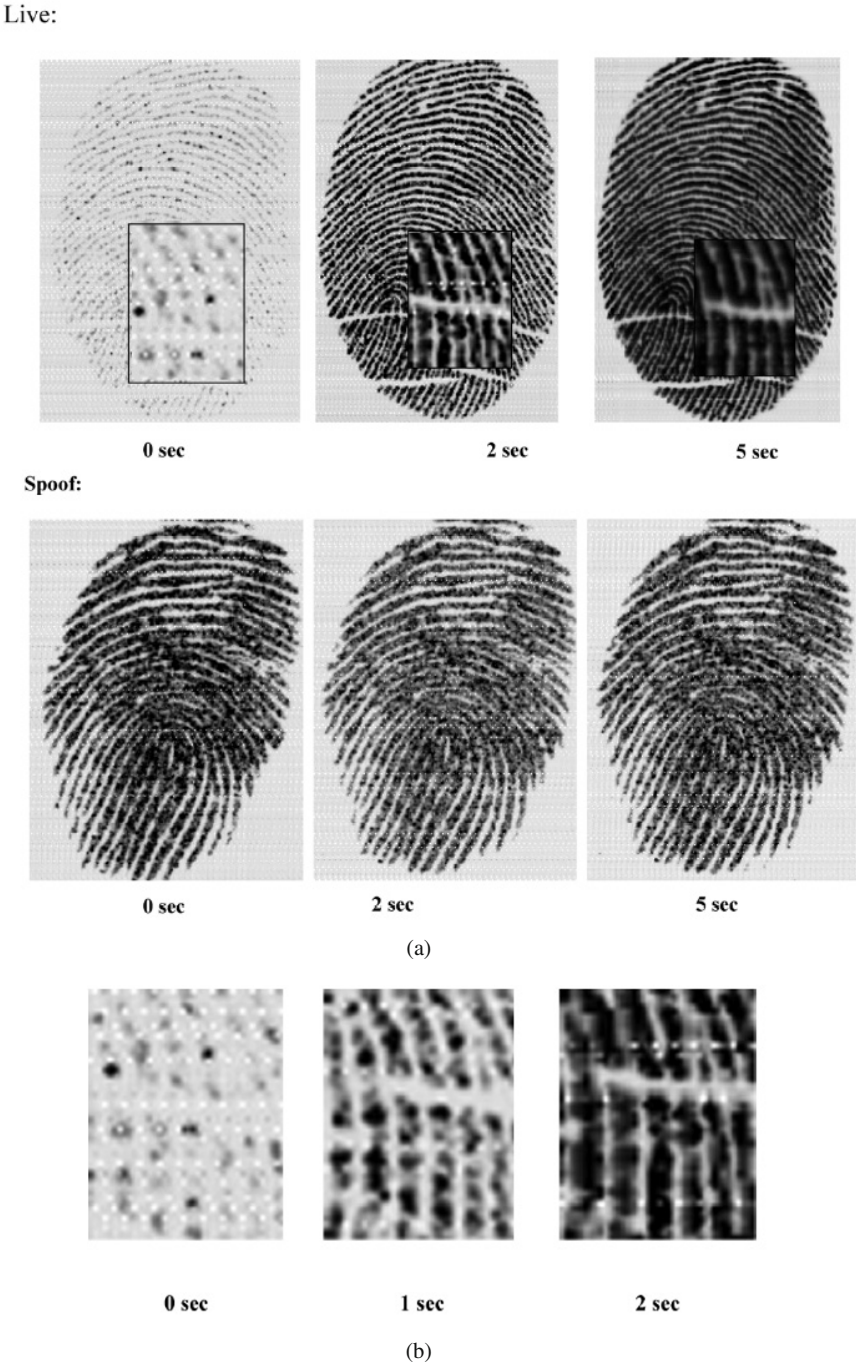


Fig. 7.7 An example of fingerprint spoof detection based on the perspiration pattern of a live finger (adapted from [35]). (a) Example fingerprint images obtained from a live finger (top row) and a fake finger (bottom row) acquired at $0, 2$, and 5 seconds after the finger is placed on the sensor, (b) enlarged fingerprint image sequence that demonstrates progression of a perspiration pattern over time in a live finger.

few seconds. The perspiration-based methods are also expected to have some difficulty in dealing with varying amounts of moisture content occurring in live human fingers.

- **Spectral/optical properties of the human skin:** This is one of the most common characteristics that has been successfully used for spoof detection in many biometric systems, including fingerprint, palmprint, face, and iris. The optical properties that may be measured include the absorption, reflection, scattering, and refraction properties under different illumination conditions (such as wavelength, polarization, coherence). In the case of fingerprints, multi-spectral analysis may be used to measure the surface properties as well as sub-surface properties of a finger since components of blood (oxygenated and deoxygenated hemoglobin) absorb different wavelengths of light. Similarly, the tissue, blood, fat, and melanin pigments in the eyes absorb different wavelengths of light. These properties can be leveraged for liveness detection in fingerprint (see Figure 7.8) and iris recognition systems.

  Eyes have a few additional optical properties that can also be used to detect fake irides. For instance, photographs of an iris can be differentiated from a live iris by detecting phenomena like Purkinje reflections and red eye effect. While Purkinje images are reflections of outside objects against the cornea of the eye, the red eye effect is due to retinal reflection. Moreover, analysis of the two-dimensional Fourier spectrum can also be used to identify contact lenses with a fake iris printed on them (see Figure 7.9).

- **Electrical characteristics**: The electrical conductivity of human tissue differs from conductivity of many other synthetic materials such as silicone rubber and gelatin. The conductivity of the material presented to the fingerprint sensor can be measured to differentiate a live finger from a fake finger. However, the conductivity of live fingers varies a lot depending on environmental conditions such as humidity and temperature. If water or saliva is added to a fake finger, its conductivity may be indistinguishable from that of a live finger.

- **Skin deformation**: The deformation pattern of the human skin can be used for differentiating live fingers from fake fingers. Skin is more flexible than most other materials and the ridges and valleys in a fake finger do not deform like a live fingertip. Real live skin deforms only in a certain way because the skin is anchored to the underlying derma and the deformation is influenced by the position and shape of the finger bone. But measuring these deformation patterns is not easy because it requires capturing a video of the fingerprint at a high frame rate as the finger moves on the sensor surface. This is problematic because most fingerprint sensors are designed for single-touch fingerprint acquisition and the users are trained not to move the finger during capture; excessive deformation will affect the matching accuracy of the system.

One of the common criticisms of the liveness detection algorithms employed in commercial biometric systems is that they are based solely on the principle of *se-*

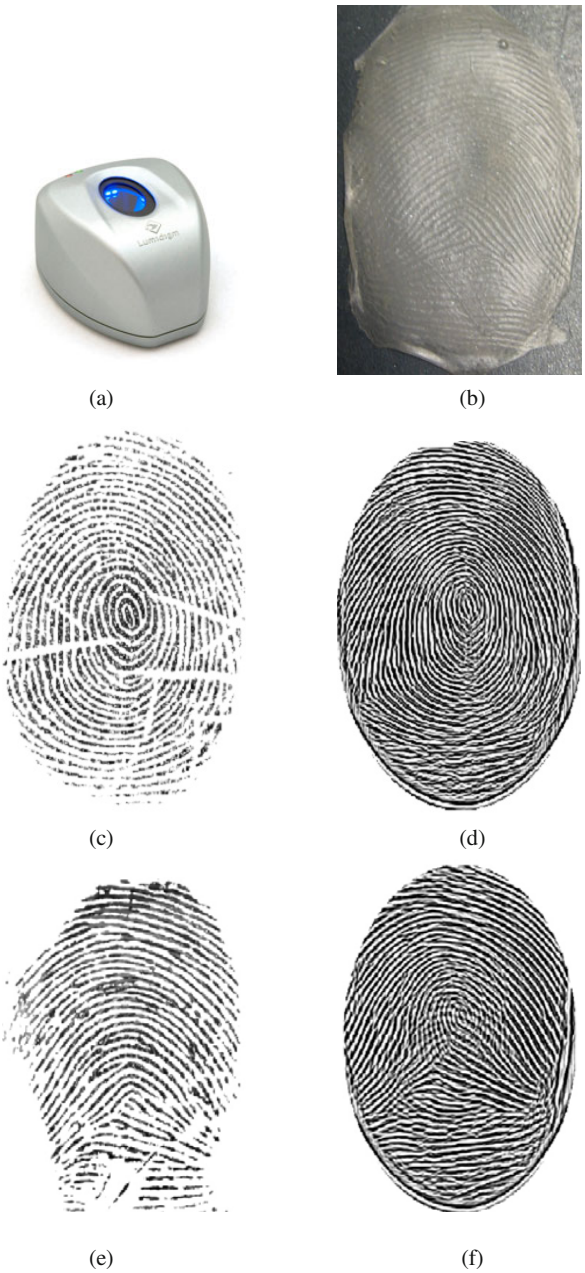


Fig. 7.8 An example of fingerprint spoof detection using the spectral properties of the human tissue. (a) a multi-spectral fingerprint sensor from Lumidigm, Inc. that is capable of capturing the sub-surface properties of a finger, (b) a spoof fingerprint made from glue, (c) an impression of the real finger acquired using a traditional optical fingerprint sensor (based on total internal reflection (TIR) principle), (d) an impression of the real finger acquired using the multi-spectral fingerprint sensor, (e) an impression of the spoof finger (glue spoof overlaid on the real finger) acquired using the optical fingerprint sensor, and (f) an impression of the spoof finger acquired using the multi-spectral fingerprint sensor. It can be observed that the multi-spectral sensor is able to see through the spoof and capture the ridge pattern of the underlying real finger. (Source: Lumidigm, Inc.)

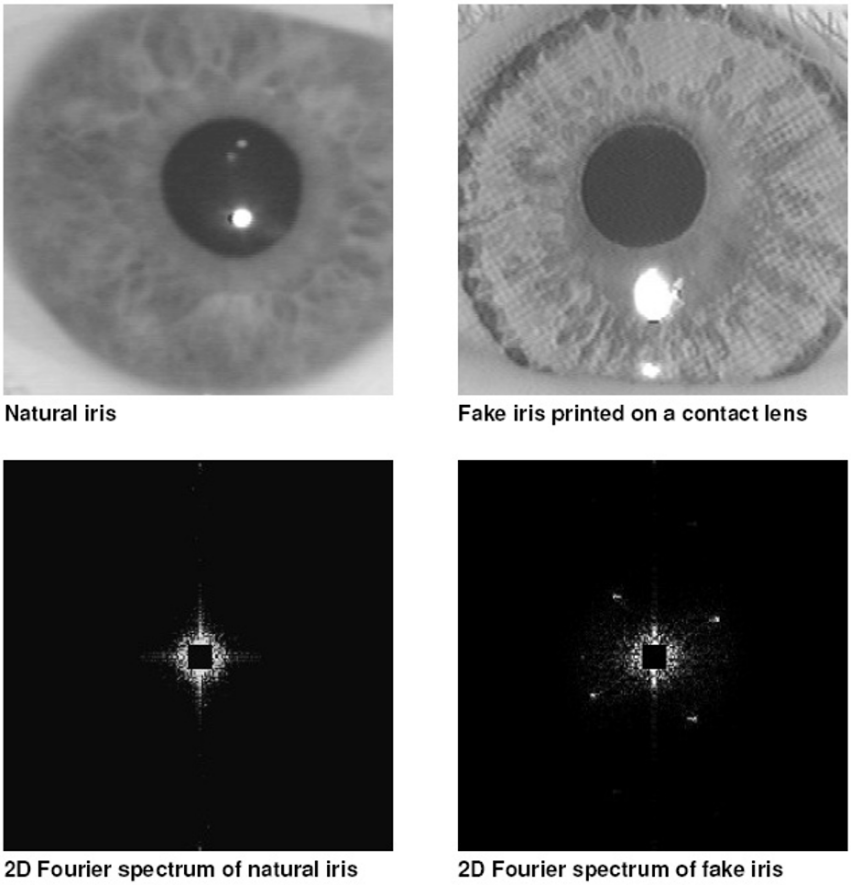


Fig. 7.9 An example of iris spoof detection (adapted from [13]). A printed iris typically exhibits some artifacts that can be detected by analyzing the 2-dimensional Fourier spectrum of the iris image.

curity through obscurity. In other words, biometric vendors do not generally reveal the algorithm or implementation details about their liveness detection methodology because if the specifics of the spoof detection techniques are revealed, the system can be circumvented easily. Experience in cryptographic systems has shown that this approach does not provide satisfactory results over a period of time. Once an attacker identifies a possible vulnerability and successfully carries out a spoof attack, the complete system falls apart. Therefore, one should assume that the attacker has knowledge about the physiological properties used by the system for detecting spoofs. Consequently, it may be possible for the attacker to create a fake finger with the same properties that are verified by the spoof detector. Of course, the addition of more and more physiological characteristics in the spoof detection process

will make it progressively more difficult (though not impossible) for the attacker to fool the system. While biometric sensors should be equipped with as much liveness detection capability as possible, the use of multiple biometric traits (multimodal biometric systems are discussed in Chapter 6) combined with intelligent challenge-response mechanisms may be required to raise the bar to a level that is difficult for an attacker to surmount.

7.4 Attacks on Biometric Processing

The signal processing and pattern matching algorithms that form the crux of automated biometric recognition are implemented in the sensor, feature extractor, matcher, and decision modules. Thus, an attacker can subvert the biometric processing either by directly undermining the core functional modules of the biometric system or by manipulating the communication between these modules. Though the template database is also one of the modules in the biometric system, the motivation and consequences of an attack on the template database are different compared to the other modules. Therefore, the attacks on the template database will be considered separately.

7.4.1 Attacks on the system modules

Attacks on the core functional modules can be mounted either through unauthorized modification or by exploiting the faults in their implementation. The motivation of these attacks could be to cause denial-of-service to legitimate users or facilitate intrusion.

7.4.1.1 Unauthorized modification

The hardware and software components of a biometric system can be modified by attackers. A classic example is the modification of an executable program in a module through a Trojan horse attack. A Trojan horse is malicious software that appears to perform a desirable function for the user, but instead performs some other function that usually facilitates intrusion by unauthorized users. The Trojan horse can disguise itself as one of the modules, bypass that module, and output the values desired by the adversary as input to the subsequent modules. For instance, a Trojan horse program can bypass the feature extractor and send the false features determined by the attacker to the matching module (see Figure 7.10). Similar attacks can also be carried out at the sensing, quality estimation, matching, template database, and decision modules.