Biometric Functionalities and System Errors

Guodong Guo

8/29/2017
Two types of functionalities

• A biometric system can provide two types of identity management functionalities:
  – Verification
  – Identification

• The generic term “recognition” will be used when we do not wish to make a distinction between the verification and identification functionalities

• authentication = verification
**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?”
- **One-to-one match**: the query is compared only to the template corresponding to the claimed identity.
- **How to claim an identity?**
Claim an identity

• 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)

• **Genuine** and **Impostor**: 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”

• Other terms: “client” or “authentic” == “genuine”
Application scenarios

• Verification is typically used in applications where the goal is to prevent unauthorized persons from using the services
  – Such as ...?
two-category classification

- Verification can be considered as a two-category classification problem
- Given a claimed identity \( I \) and a query feature set \( x^A \), we need to decide if \((I, x^A)\) belongs to “genuine” or “impostor” class
- Let \( y \) be the stored template corresponding to identity \( I \). Typically, \( x^A \) is compared with \( y \) and a match score \( s \), which measures the similarity between \( x^A \) and \( y \), is computed. The decision rule is given by
(Cons.)

• \((I, x^A) \in \text{genuine, if } s \geq \eta,\)
  
impostor, if \(s < \eta,\)

where \(\eta\) 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 should be reversed.

• When the identity claim is deemed to be “genuine”, the user is allowed to access the services provided by the system.
Identification

• Identification functionality can be further classified into positive and negative identification

• **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.

• **Negative identification**: The user is concealing his true identity from the system either explicitly or implicitly.
  
  – 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?”
negative identification

• The purpose of negative identification is to prevent a single person from using multiple identities.

• Examples:
  – Screening can be used to prevent the issue of multiple credential records (e.g., driver’s license, passport) assigned to the same person;
  – Screening is used to prevent a person from claiming multiple benefits under different names (a problem commonly encountered in welfare disbursement applications);
Formal description of the identification problem

• Given a query feature set $x^A$, we need to decide the identity $I$ of the user, where $I \in \{I_1, I_2, \ldots, I_N, I_{N+1}\}$. Here, $I_1, I_2, \ldots, I_N$ 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 $x_n$ is the stored template corresponding to identity $I_n$ and $s_n$ is the match score between $x^A$ and $x_n$, for $n = 1, 2, \ldots, N$, the decision rule for identification is,

$$x^A \in I_j, \quad \text{if } j = \text{argmax}_n s_n \text{ and } s_j \geq \eta,$$

$$I_{N+1}, \quad \text{otherwise},$$

where $\eta$ is a pre-defined threshold.
Open set Identification

• 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*, why?
Closed set Identification

• Force the system to return one among the $N$ enrolled identities, irrespective of the value of $s_j$. Such a scenario is called *closed set identification*.
Semi-automated 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)

• Why do we need this semi-automated identification?
Example

• 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 ($\eta$), since the number of enrolled users in the database can be quite large.
Biometric System Errors

• The science of biometric recognition is based on two fundamental premises, namely, *uniqueness* and *permanence* of the underlying biometric trait.

• **Uniqueness**: A biometric identifier is said to be *unique* only if any two persons in the world can be differentiated based on the given identifier.

• **Permanence**: A biometric trait is *permanent* if it does not change over the lifetime of an individual.
uniqueness and permanence

• These two premises are seldom true in practical biometric systems
  – The physical trait itself may not be unique
    • Example: Fingerprint, face, or iris.
  – The uniqueness or *individuality* of 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
  – Example: The facial appearance of identical twins is almost the same
Two types of features

• **Genotypic factors/features**: Modalities such as DNA, where the genetic constitution of the individual largely determines their biometric characteristics are referred to as **genotypic factors/features**

• **Phenotypic factors/features**: The modalities whose characteristics are determined by other sources of randomness in nature (e.g., fingerprints) are referred to as **phenotypic factors/features**
Permanence

• The notion that the biometric traits are permanent is not an established scientific fact, either

• 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

• Biometric systems rely only on the digital measurements of the body characteristics, and not the real physical traits
Biometric samples are seldom identical

• **Intra-user Variations**: 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 speaker recognition)
  – changes in ambient conditions (e.g., inconsistent illumination levels in face recognition applications)
  – variations in the user’s interaction with the sensor (e.g., occluded iris or partial fingerprints)
No perfect match

• 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

• **When can we have a perfect match?**
  
  – e.g., in password-based authentication systems, but *not* in a biometric system
Close match in biometric systems

- 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

- It 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
  - *representation limitation* (problems in feature extraction)