Biometrics (CSE 40537/60537)

Lecture 1: What is biometrics?

Adam Czajka

Biometrics and Machine Learning Group
Warsaw University of Technology, Poland

Fall 2014
University of Notre Dame, IN, USA

Lecture 1: What is biometrics?

- Basics
- Biometric characteristics (modalities)
- Expected properties of biometrics
- (Brief) history of biometrics
- Biometric system in a nutshell
- Biometric errors and decision making

Two meanings of biometrics

1. Biometrics in a wider sense
   - etymology: measurement of the properties of living beings (in Greek: *bios* = "life", *metron* = "measurement")
   - aim of the measurement: not defined (e.g. medical diagnostics)

Two meanings of biometrics

2. Biometrics as a part of Computer Science
   - measurement of physical of behavioral properties of human beings
   - aim of the measurement defined: automatic identity recognition

Two meanings of biometrics

2. Biometrics as a part of Computer Science
   - measurement of physical of behavioral properties of human beings
   - aim of the measurement defined: automatic identity recognition

Biometrics = use of physical or behavioral properties of human beings for automatic identity recognition

Biometrics as a part of Computer Science

Use of physical or behavioral properties of human beings for automatic identity recognition, that is:

- something that is characteristic only to me (e.g. my face, my handwriting, my voice)
- not something that I know (e.g. a password or PIN)
- not something that I have (e.g. a key or credit card)

Biometrics as a part of Computer Science

Use of physical or behavioral properties of human beings for automatic identity recognition, that is:

- the data that we process must result from a measurement of a living subject
- hence, the biometric sensors must deliver authentic biometric samples
- we need liveness detection (a.k.a. presentation attack detection) to call the system a biometric system

Biometrics as a part of Computer Science

Use of physical or behavioral properties of human beings for automatic identity recognition, that is:

- it is not a human expert who decides, but applied methodologies may be inspired by the expert's experience, knowledge or skills,
- typical requirements: uncontrolled environment, high processing speed, repeatability, predictability

Biometric recognition types

Wayman, Jain, Maltoni, Maio, 2005

1. Positive recognition
   $\Rightarrow$ verification of the hypothesis: a sample represents the subject known to the system (i.e. already registered)
2. Negative recognition
   $\Rightarrow$ verification of the hypothesis: a sample represents the subject unknown to the system (not yet registered)

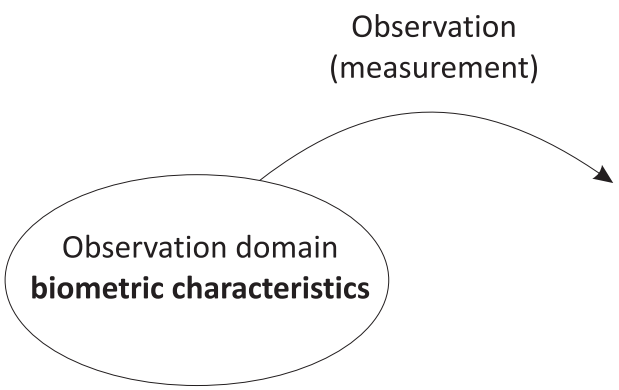
Authentication types

1. Classic
   - verification
     (confirmation of the identity claim, a.k.a. 1:1 comparison)
   - identification
     (searching for match, a.k.a. 1:N comparison)

Authentication types

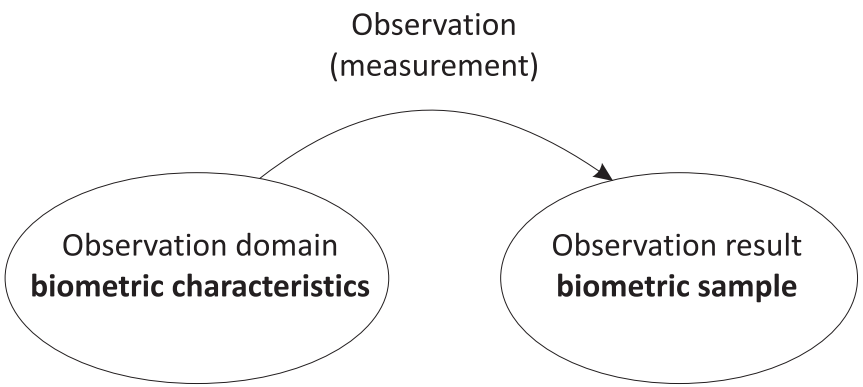
1. Classic
   - verification
     (confirmation of the identity claim, a.k.a. 1:1 comparison)
   - identification
     (searching for match, a.k.a. 1:N comparison)
2. New (due to biometrics)
   - negative authentication
     - negative identification: I'm not a member of the group X
     - negative verification: I'm not the subject X
   - elimination of "multiple identities"

Basic vocabulary



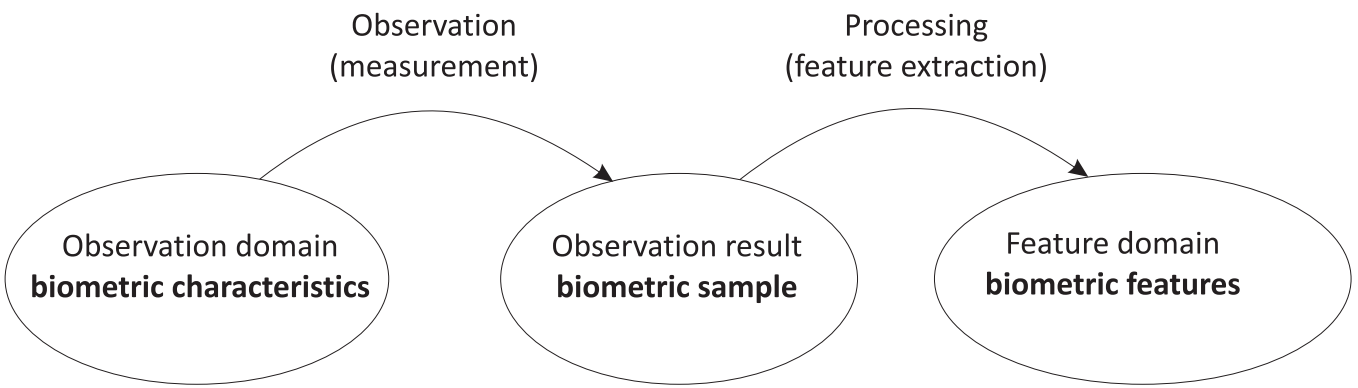
1. Subject (human being)
2. Observation domain
   - **biometric characteristics** or **modality** – a single physical or behavioral property that we use for biometric recognition
   - examples: face appearance, hand shape, finger veins

Basic vocabulary



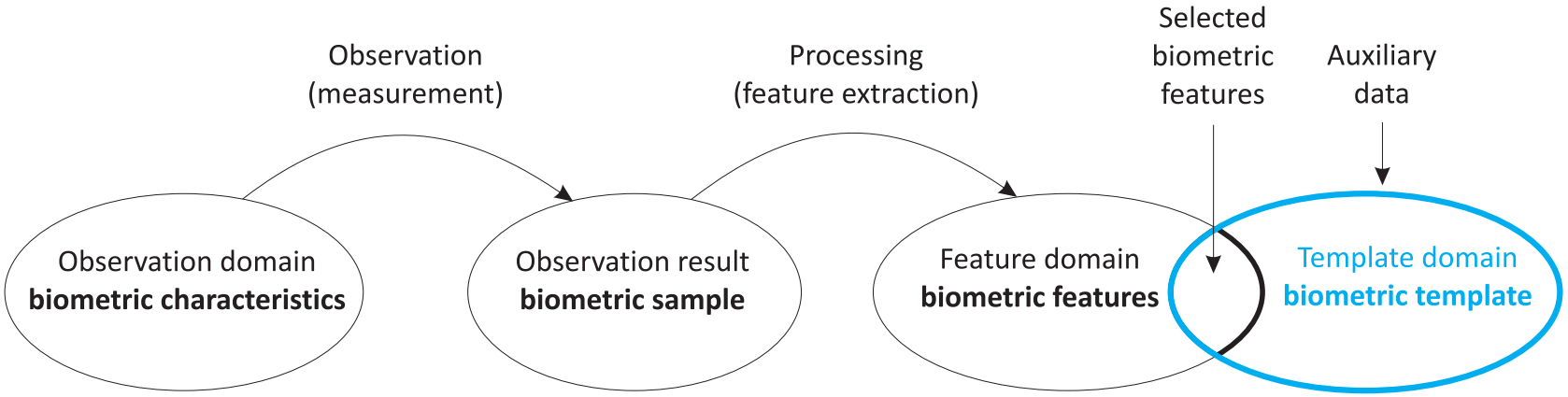
3. Observation result
   - **biometric sample** – measurement results, raw or pre-processed
   - examples: face 3D image, hand 2D image, image of finger veins

Basic vocabulary



4. Feature domain
   - **biometric features** – representation of a sample, typically much shorter than a sample
   - examples: distance between characteristic points in a face image, length or/and width of fingers, a single location of finger vein intersections

Basic vocabulary



5. Template domain
   - **biometric template** – all or selected biometric features calculated for a given sample and stored as a reference
   - biometric template may contain some auxiliary data necessary to perform the biometric comparison
   - examples: set of distances among characteristic points in a face image, lengths and widths of fingers, map of finger vein intersections

Typical division

Typical division

1. Characteristics related to physical properties
   (our appearance or anatomical properties)
   - momentary observation and immediate biometric samples
   - static measurement, time relations not used in recognition

Typical division

1. Characteristics related to physical properties
   (our appearance or anatomical properties)
   - momentary observation and immediate biometric samples
   - static measurement, time relations not used in recognition
2. Characteristics related to behavioral properties
   (our behavior due to habits, training, genetics, or mixture of those)
   - observation of a conscious action performed by the subject
   - dynamic measurement (in time), time relations used in recognition

Physical characteristics

Physical characteristics

1. Commercially used
   - finger: fingerprints, finger veins
   - hand: hand geometry, palm veins, palm print (the latter popular in forensic science)
   - face: use of 2D and 3D face images, face tracking
   - eye: iris, veins (located on the sclera, or underneath the retina and called "retina recognition")
2. Emerging, possible to be used in practice
   - finger: finger 3D geometry, thermal data, structure of the tissue underneath the nails
   - hand: thermal data
   - face: thermal data
   - ear: geometry, thermal data, ear prints (the latter popular in forensic science)
   - other: DNA, scent

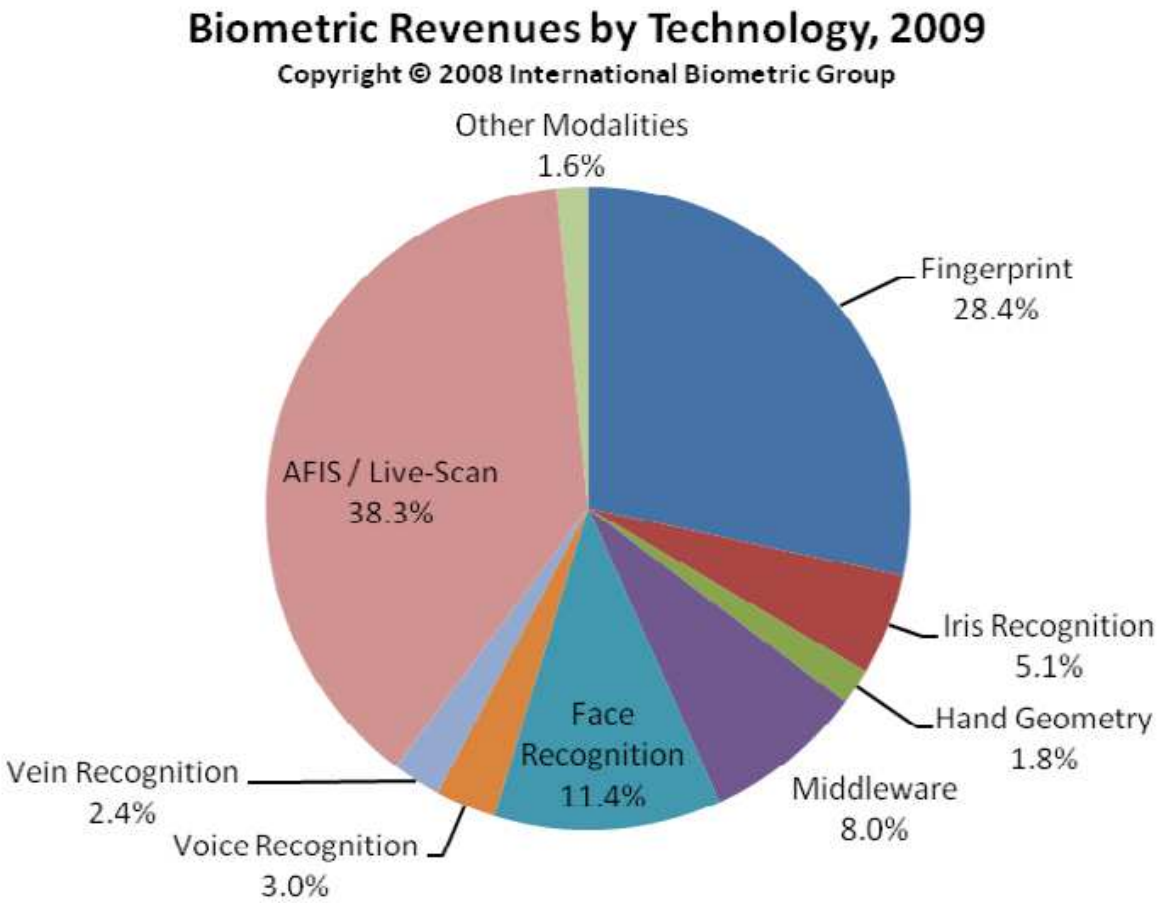
Behavioral characteristics

Behavioral characteristics

1. Commercially used
   - handwriting: signature
     (in particular on-line, i.e. registered on the graphical tablet)
   - voice: speaker recognition
     (who has said something, not what has been said)
   - hand: keystrokes dynamics (recently appeared again)
2. Emerging, possible to be used in practice
   - handwriting: use of any legible text (not limited to a signature)
   - gait: dynamics measured by mobile devices, silhouette
     movement measured at the distance
   - eye: pupil dynamics, eyeball dynamics
   - evoked potentials: EEG (brain) and ECG (heart) signals
   - other: lip movement (often merged with speaker recognition)

Biometric characteristics: market perspective

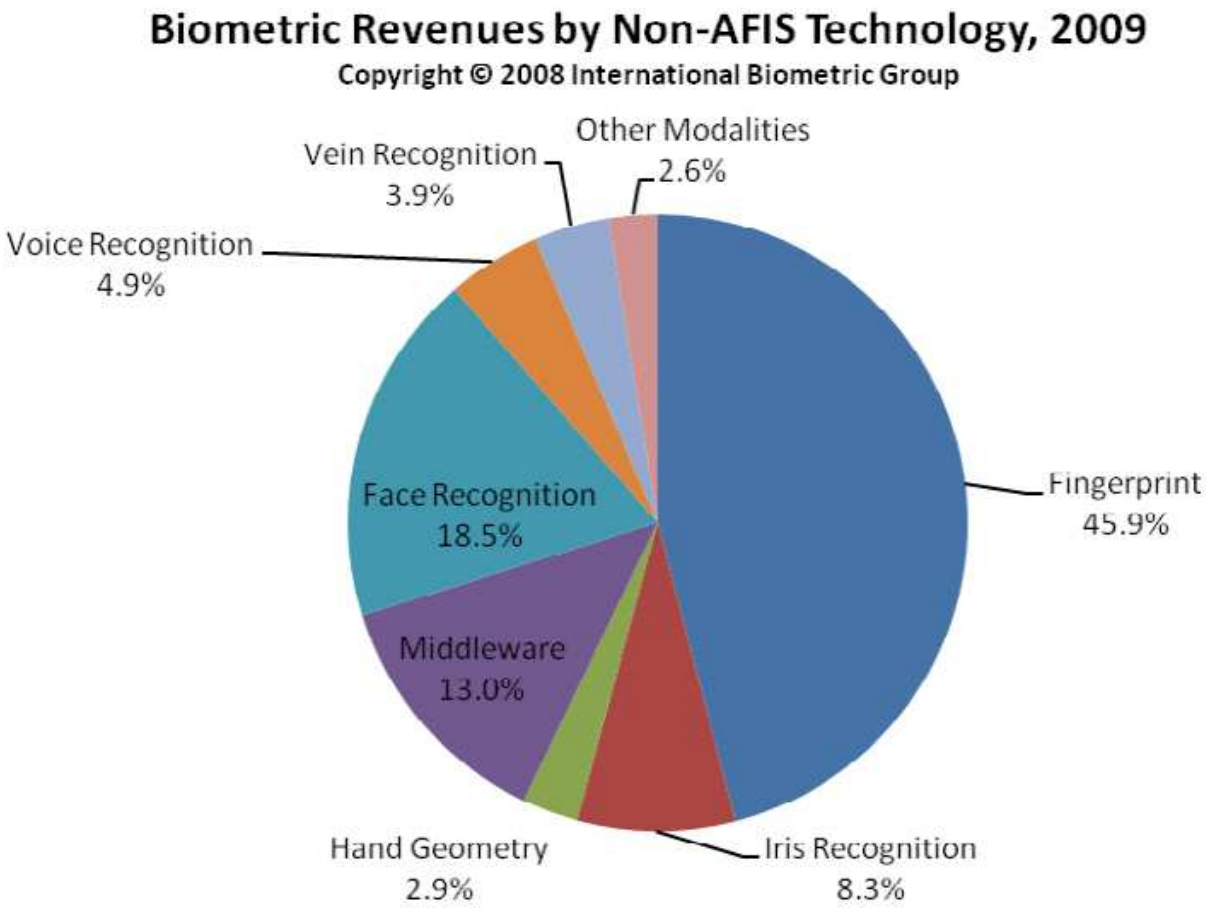
(including AFIS: Automated Fingerprint Identification Systems)



Source: IBG Biometrics Market and Industry Report 2009-2014

Biometric characteristics: market perspective

(without AFIS)



Source: IBG Biometrics Market and Industry Report 2009-2014

Information capacity

Requirement: diversification of subjects

1. Genotype and genetic penetrance
2. Twins: 1 in 80 births; identical twins: 1 in 240 births → 0.8% of recognition errors, if we use only genotype-based properties
3. We have no formal proof for uniqueness of the biometric properties; our beliefs are based only on experiments

Estimation how selected biometric characteristics depend on the genotype

Stability over time

Requirement: correct identity recognition after some long time (e.g. a few years)

1. Resistance to health conditions
2. Template aging
   - ongoing research for different modalities, apparent decrease in matching performance, yet no definitive answer related to the impact on operational scenarios
   - data collection extremely difficult, only a few datasets exist: MORPH (face; time interval up to 29 years), KFRIA Ageing DB (fingerprint; time interval: 1 year), BioBase II NASK/PW (multimodal: face, iris, fingerprint, handwritten signatures and hand shape; time interval: 7 years), University of Notre Dame (iris; time interval: 4 years), NIST OPS-XING and OPS-FIELD (iris; time interval: 4 years)

User acceptance

Requirement: generate no fears related to health, social aspects or religion

1. Technology used to increase user's comfort, with minimum (or no) cooperation required (but lest we forget about abuses)
2. Personal or sensitive data protection; depends on local law, but we always would like to know:
   - who processes our biometric data, how long and for what purposes
   - what kind of information is revealed along with our biometric data (health issues, physiological predisposition, age, gender, race, etc.)
   - if our biometric data can be generated, copied or captured without our will (e.g. latent fingerprints on glass, face photos from facebook, handwritten signature on the credit card bill, synthetic data)

User acceptance

Requirement: generate no fears related to health, social aspects or religion

3. "Big brother" syndrome
   - can our biometric data be used without our will or knowledge, e.g. for search in forensic databases?
   - biometric science requires research databases (for development and testing of algorithms)
   - technology starts to surround us: "ubiquitous computing", "pervasive computing", "cloud computing", ...
4. Machine's decision $\rightarrow$ mistrust
5. Unobtrusive and comfortable for users

User acceptance

Requirement: generate no fears related to health, social aspects or religion

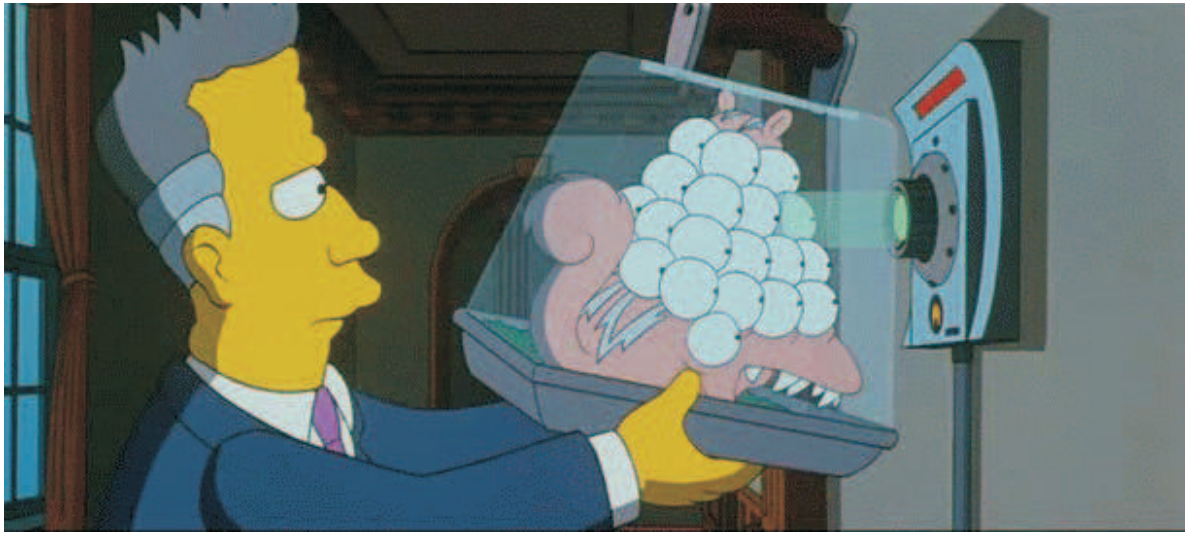
3. "Big brother" syndrome
   - can our biometric data be used without our will or knowledge, e.g. for search in forensic databases?
   - biometric science requires research databases (for development and testing of algorithms)
   - technology starts to surround us: "ubiquitous computing", "pervasive computing", "cloud computing", ...
4. Machine's decision $\rightarrow$ mistrust
5. Unobtrusive and comfortable for users

Example of unobtrusiveness and comfort of use: iris recognition in *Minority report* movie

Resistance to presentation attacks

Requirement: delivering of authentic biometric samples

1. Hazardous or impossible physical alterations (e.g. mutilations)
2. Difficult imitation or disguise
3. Possibility to construct effective liveness detection tests



Amenable to implement

Requirement: reasonable balance between costs and performance

1. Universality of the selected characteristics (e.g. tattoo signs)
2. Easiness of measurement
3. Repeatability of observation and measurement
4. Reasonable costs of manufacturing

First references to some biometric properties




Source: Wikipedia

Source: thelinebreak.wordpress.com


Source: A. Jain, Introduction to biometrics, in: Biometrics: Personal Identification in Netw. Society, Kluwer, 1998

1. approx. 31 000 years B.C.
   - Chauvet cave in France, the oldest hand prints, 'signatures' of the painters (hypothesis)
2. approx. 2 000 years B.C.
   - caves on Gavrinis island in France, sculptures showing fingerprints
3. approx. 700 years B.C.
   - Ancient China, Assyria and Babylon
   - relation between people and transactions, use of clay tablets with fingerprints

First automatic systems

1. XX century
   - '60s: speaker and fingerprints recognition
   - '70s: hand geometry
   - '80s: retina and handwritten signatures
   - '90s: iris recognition
5. Nowadays
   - mass applications of biometrics (e.g. e-passports, AADHAAR)
   - biometric mobile devices (e.g. tablets, smart phones, watches)
   - capture without cooperation
   - security of biometrics (liveness detection and template protection)

Tasks of the biometric system

1. User enrollment
   - generation of the biometric reference template based on biometric sample(s)
   - typically the connection between the template and other personal data is performed at the enrollment
   - storage of the biometric reference template for future use (i.e. authentication)
2. User authentication
   - matching of the biometric reference template with a temporary template (calculated for authentication purposes)

Tasks of the biometric system

Enrollment

1. Supervised by an operator (yet exceptions happen)
2. Multiple measurement $\Rightarrow$ several biometric samples
   - restrictive quality control
   - selection of the best sample(s), or merging the information at the sample level
3. Feature extraction and template generation
   - cohesion of features vs. variance of features (if we have feature sets originating from multiple samples)
   - optional: estimation of feature variances and storing them within the template

Tasks of the biometric system

Enrollment (*continued*)

4. Storage of the biometric template
   - central database or personal carriers (e.g. a smart card) may be used
   - building the relation between biometric template and personal data (required to authenticate the person)
5. Priority in the enrollment: delivering good quality reference template, hence it may take some time (typically a few minutes)

Tasks of the biometric system

Authentication

1. Not supervised by an operator (again: exceptions happen)
2. Reading of the biometric reference template from a database or other carrier (e.g. smart card, if matching off card)
3. Single measurement ⇒ single biometric sample
   - optional quality control
   - liveness test (due to no human inspection)
4. Feature extraction and template generation

Tasks of the biometric system

Authentication (*continued*)

5. Matching and decision
   - calculation of the matching score among templates
   - comparison of the matching score with the acceptance threshold
   - optional use of feature variance and/or individual acceptance threshold (if present within the template)
6. If something goes wrong multiple attempts (measurements) are allowed (typically three), and the process is repeated
7. Priority in the authentication: delivering reliable and fast decision, hence it must be fast (typically a fraction of a second for verification, identification may be slower depending on the application)

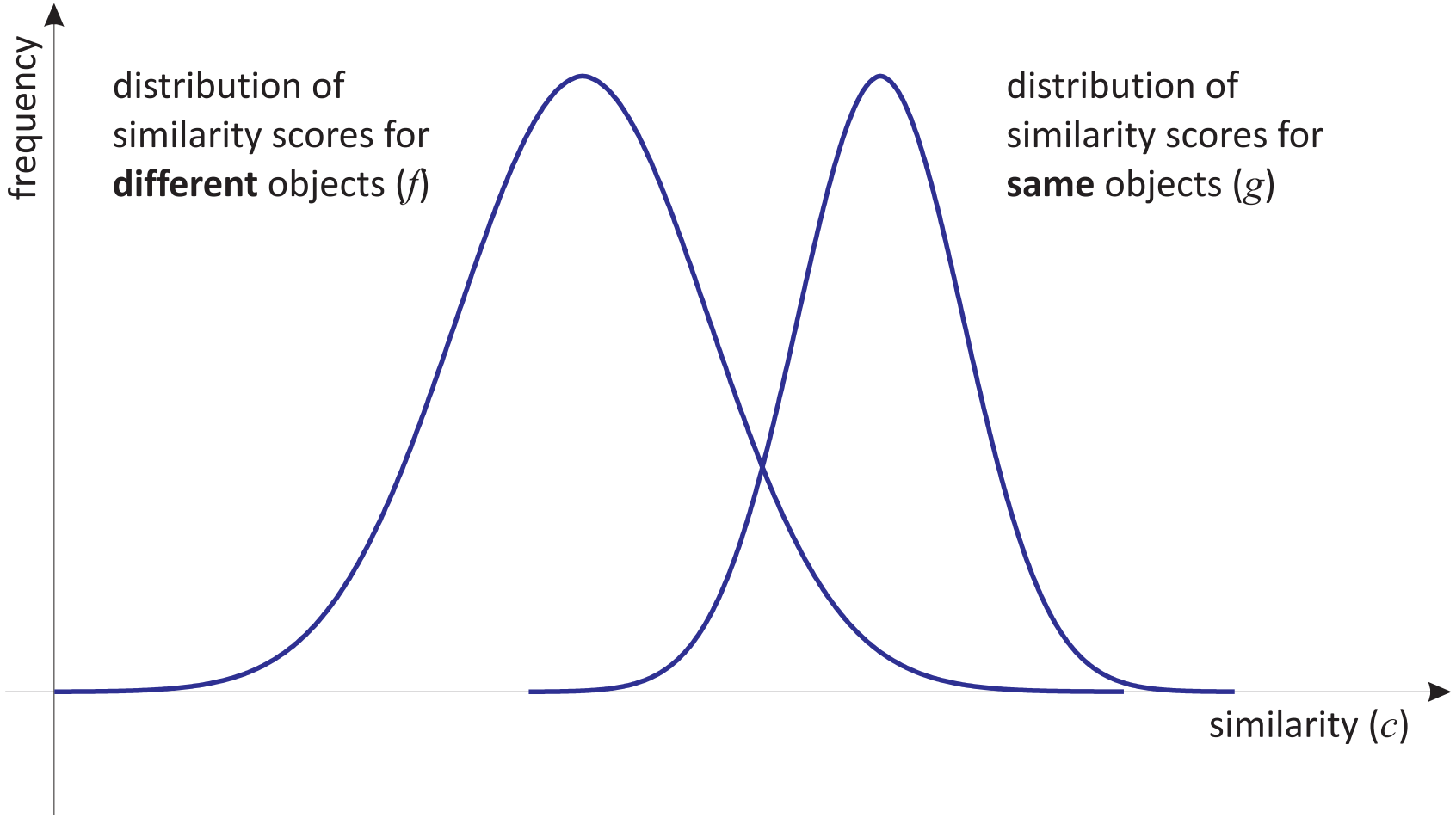
1. Biometric method errors

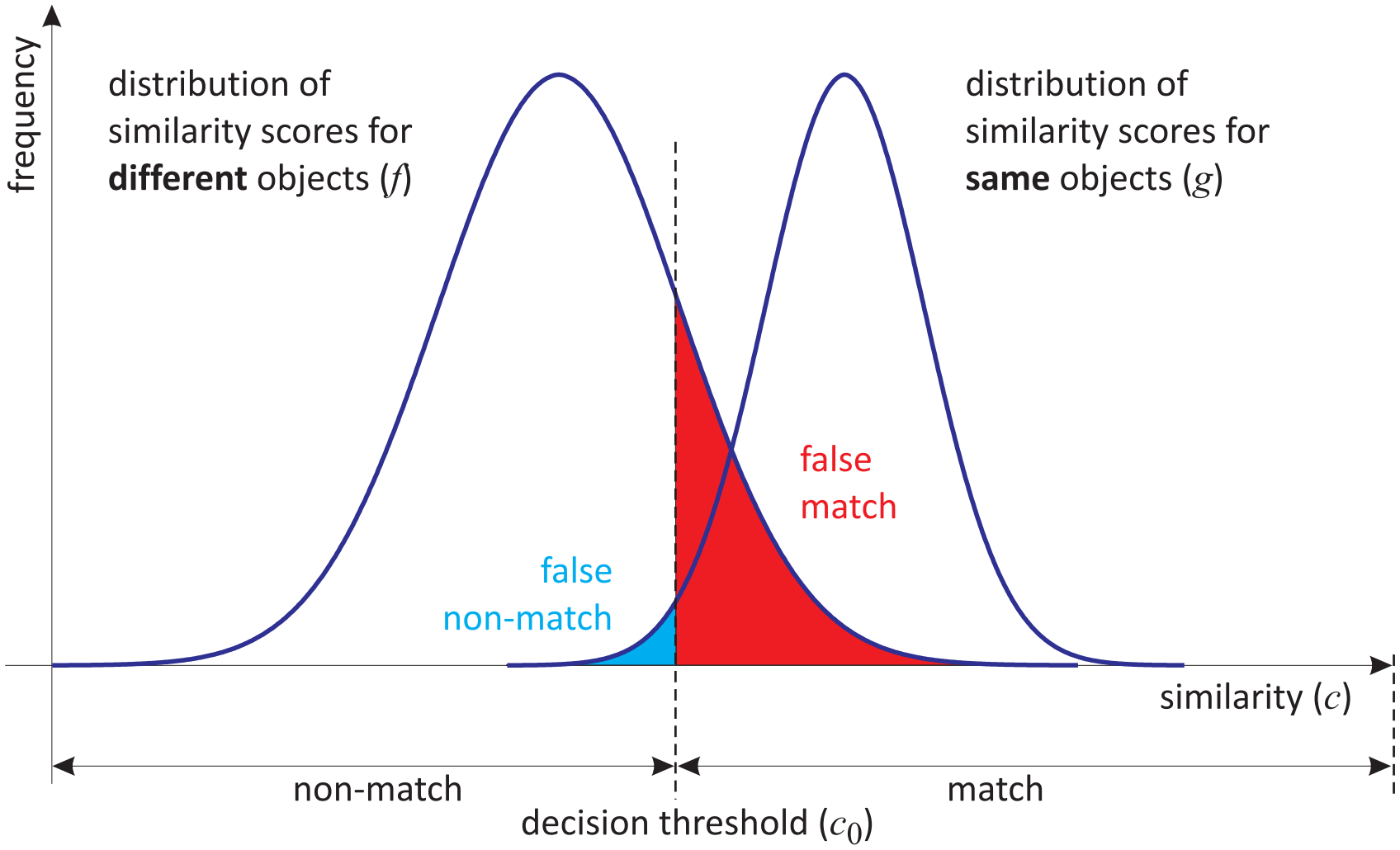
1. Biometric method errors
   - false non-match: a genuine sample did not match a reference template
   - false match: a impostor sample matched a reference template

1. Biometric method errors
   - false non-match: a genuine sample did not match a reference template
   - false match: a impostor sample matched a reference template
2. Matching score
   - similarity or dissimilarity of features calculated during authentication and those included in the reference template

1. Biometric method errors
   - false non-match: a genuine sample did not match a reference template
   - false match: a impostor sample matched a reference template
2. Matching score
   - similarity or dissimilarity of features calculated during authentication and those included in the reference template
3. Possible decisions
   - match or non-match, based on a decision threshold
   - **problem duality:**

     a *match* means that the similarity score exceeded the decision threshold (the similarity was higher than required), or the dissimilarity score was below the decision threshold (the dissimilarity was lower than accepted)


frequency
distribution of similarity scores for **different** objects (f)
distribution of similarity scores for **same** objects (g)
similarity (c)


frequency
distribution of similarity scores for **different** objects (f)
distribution of similarity scores for **same** objects (g)
false match
false non-match
similarity (c)
non-match
match
decision threshold (c_0)

Error probability estimation

False Non-Match Rate – FNMR

If we know g:

$$g_{\mathsf{FNM}}(c_0) = \int_{-\infty}^{c_0} g(c)dc$$

but typically it is not the case, hence:

$$\widehat{g}_{\mathsf{FNM}}(c_0) = \mathsf{FNMR}(c_0) = \frac{\text{number of false non-matches for } c_0}{\text{number of all genuine comparisons}}$$

Error probability estimation

False Match Rate – FMR

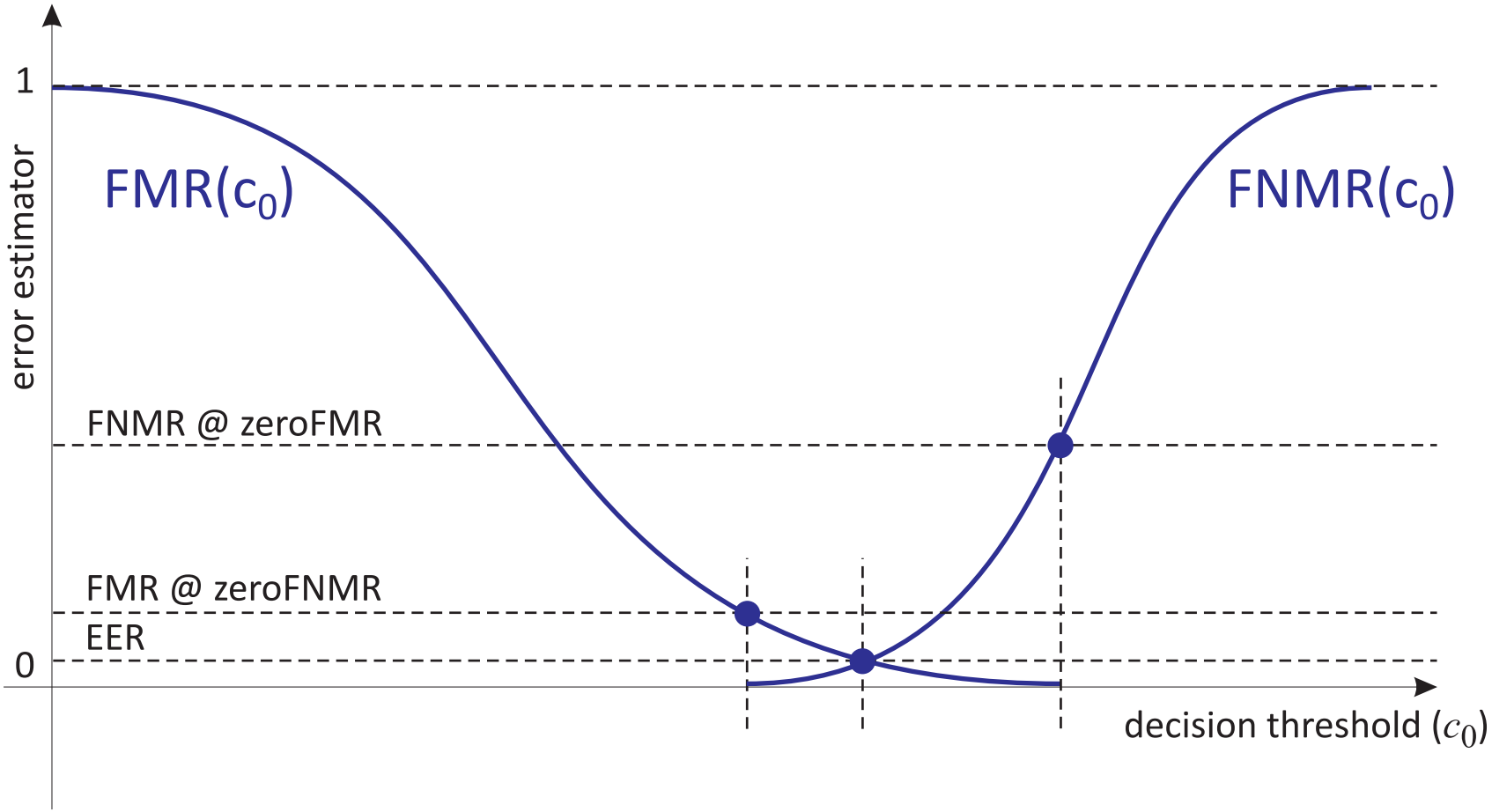
If we know f:

$$f_{\mathsf{FM}}(c_0) = \int_{c_0}^{\infty} f(c)dc$$

but typically it is not the case, hence:

$$\widehat{f}_{\mathsf{FM}}(c_0) = \mathsf{FMR}(c_0) = \frac{\text{number of false matches for } c_0}{\text{number of all impostor comparisons}}$$

Error estimators as a function of the decision threshold



Commonly used error estimators

1. Failure To Acquire – FTA: number of falsely rejected biometric samples (during measurement of the genuine user)

Commonly used error estimators

1. Failure To Acquire – FTA: number of falsely rejected biometric samples (during measurement of the genuine user)
2. Failure To Enroll – FTE: same as FTA, but referring to the reference templates calculated at the enrollment

Commonly used error estimators

1. Failure To Acquire – FTA: number of falsely rejected biometric samples (during measurement of the genuine user)
2. Failure To Enroll – FTE: same as FTA, but referring to the reference templates calculated at the enrollment
3. FMR/FNMR for a given decision threshold

Commonly used error estimators

1. Failure To Acquire – FTA: number of falsely rejected biometric samples (during measurement of the genuine user)
2. Failure To Enroll – FTE: same as FTA, but referring to the reference templates calculated at the enrollment
3. FMR/FNMR for a given decision threshold
4. Equal Error Rate – EER: FMR or FNMR when FMR=FNMR

Commonly used error estimators

1. Failure To Acquire – FTA: number of falsely rejected biometric samples (during measurement of the genuine user)
2. Failure To Enroll – FTE: same as FTA, but referring to the reference templates calculated at the enrollment
3. FMR/FNMR for a given decision threshold
4. Equal Error Rate – EER: FMR or FNMR when FMR=FNMR
5. FMR @ zeroFNMR: we check the minimum FMR that still allows for FNMR=0

Commonly used error estimators

1. Failure To Acquire – FTA: number of falsely rejected biometric samples (during measurement of the genuine user)
2. Failure To Enroll – FTE: same as FTA, but referring to the reference templates calculated at the enrollment
3. FMR/FNMR for a given decision threshold
4. Equal Error Rate – EER: FMR or FNMR when FMR=FNMR
5. FMR @ zeroFNMR: we check the minimum FMR that still allows for FNMR=0
6. FNMR @ zeroFMR: same as for 5., but we demand FMR=0 and search for the minimum FNMR