Information Security
Identification and authentication

Advanced User Authentication II
2019-02-08

Amund Hunstad

Guest Lecturer, amund@foi.se

LiU expanding reality

Agenda for lecture I within this part of the course

Background

Statistics in user authentication

Biometric systems

Tokens

Authentication✓

eID✓

ePassports✓

Biometrics in general✓

Fumy, W. and Paeschke, M. Handbook of eID Security

A. Jain, A. Ross and K. Nandakumar, Chapters 1 in "Introduction to Biometrics"

Agenda for lecture II within this part of the course

Background

Statistics in user authentication

Biometric systems

Tokens

- Statistics
- Generic biometric system
- Design cycle

A. Jain, A. Ross and K. Nandakumar, Chapters 1, 6 & 7 in "Introduction to Biometrics"

Agenda for lecture II within this part of the course

Background

Statistics in user authentication

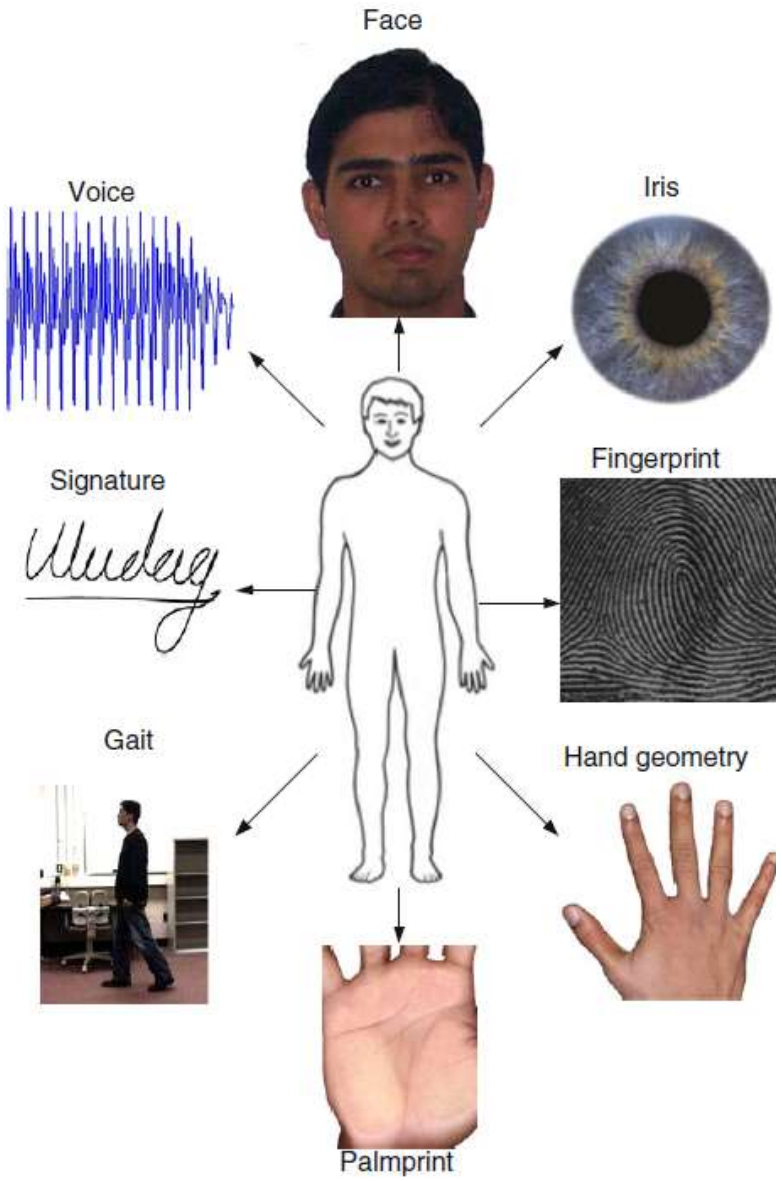
Biometric systems

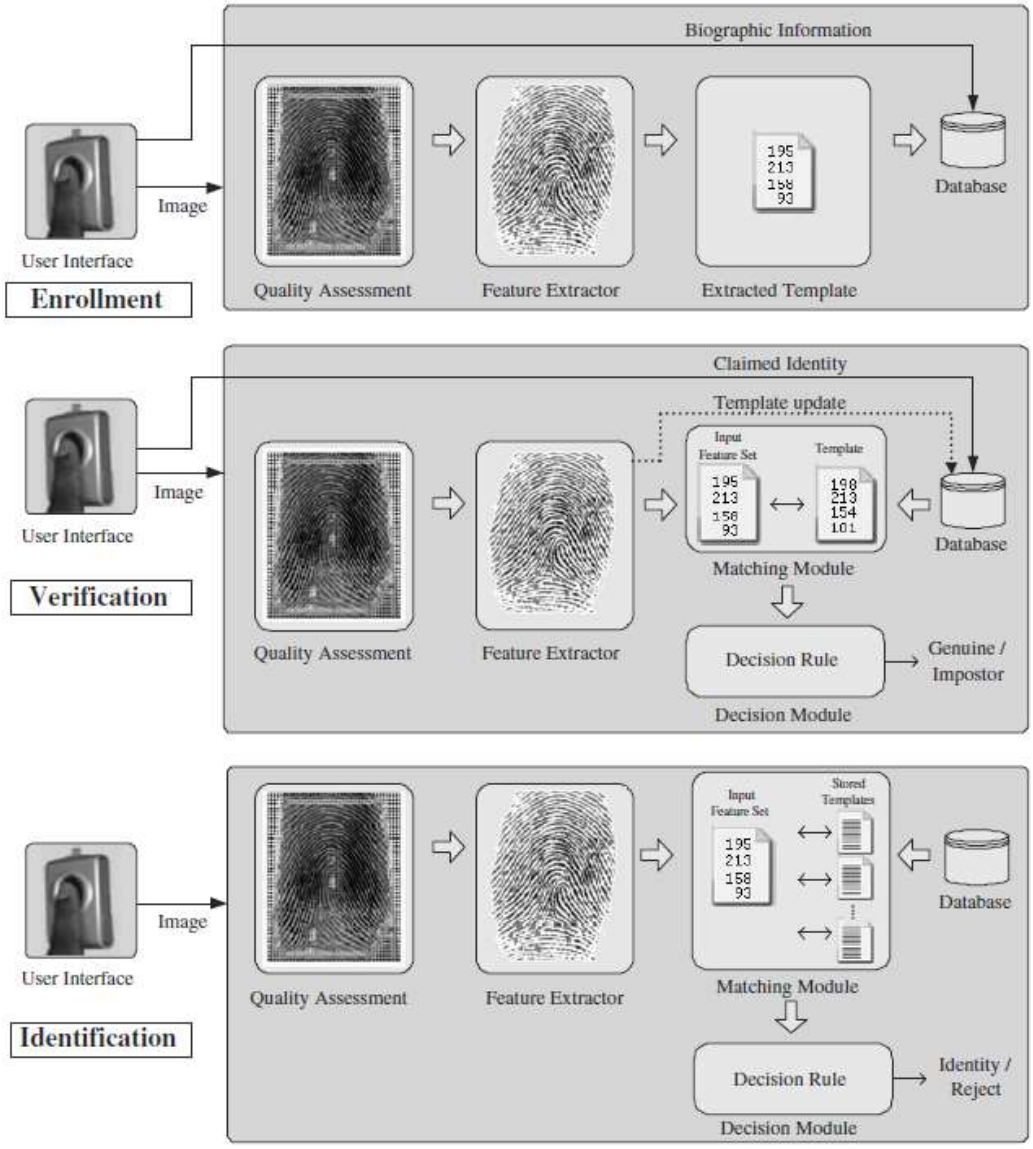
Tokens

Security threats
Attacks
Multibiometrics
Biometric traits, examples
Attacks on tokens

A. Jain, A. Ross and K. Nandakumar, Chapters 6 & 7, 2-5 in "Introduction to Biometrics"

Ross Anderson, Security Engineering, Chapter 16


Face
Voice
Iris
Signature
Fingerprint
Gait
Hand geometry
Palmprint

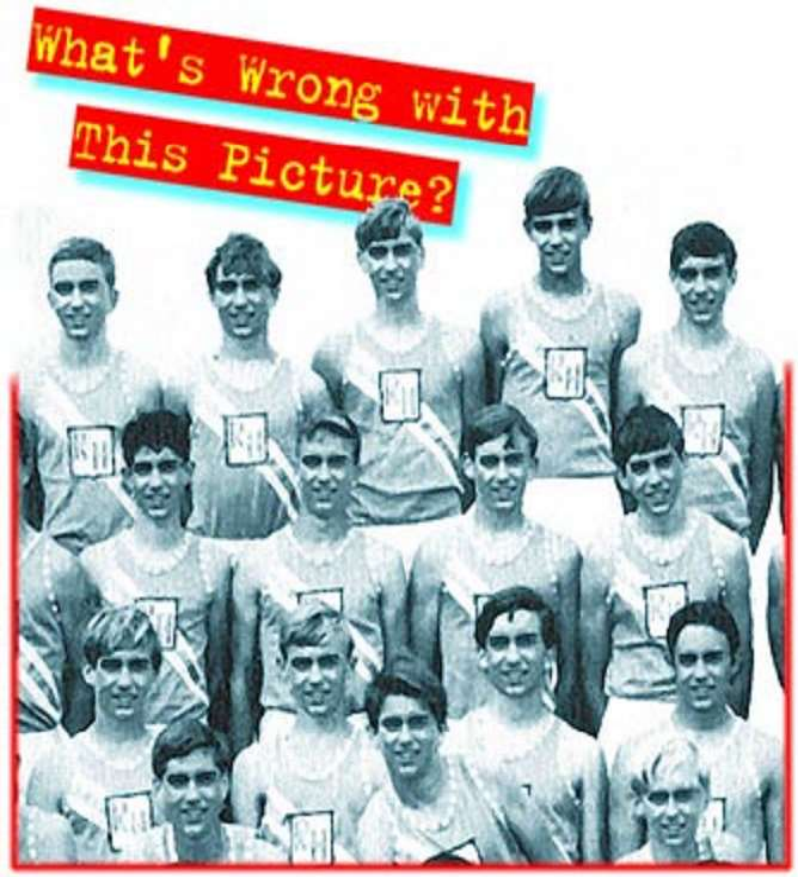

Biographic Information
Image
User Interface
Enrollment
Quality Assessment
Feature Extractor
Extracted Template
Database
Claimed Identity
Template update
Input Feature Set
Template
Verification
Matching Module
Decision Rule
Genuine / Impostor
Decision Module
Stored Templates
Identification
Identity / Reject

Identification

"Who am I?"

Comparisons are made with every template in the database

The result is an identity (name or user ID) or "NO MATCH"



Identity verification = Authentication

"Am I the person who I claim I am?"

The user claims to have a certain identity (e.g. by specifying a user name)

Comparisons are made only with one template.

The result is TRUE/FALSE

Matching, decision regions, hypothesis testing

A typical system has a threshold parameter which determines the allowed variance

Statistical theory for hypothesis testing enables analysis

It is necessary to balance user population statistics against intended use

More about this …

Statistics in user authentication

Problems and unexpected effects

Matching, decision regions, hypothesis testing

A typical system has a threshold parameter which determines the allowed variance

Statistical theory for hypothesis testing enables analysis

It is necessary to balance user population statistics against intended use

More about this …

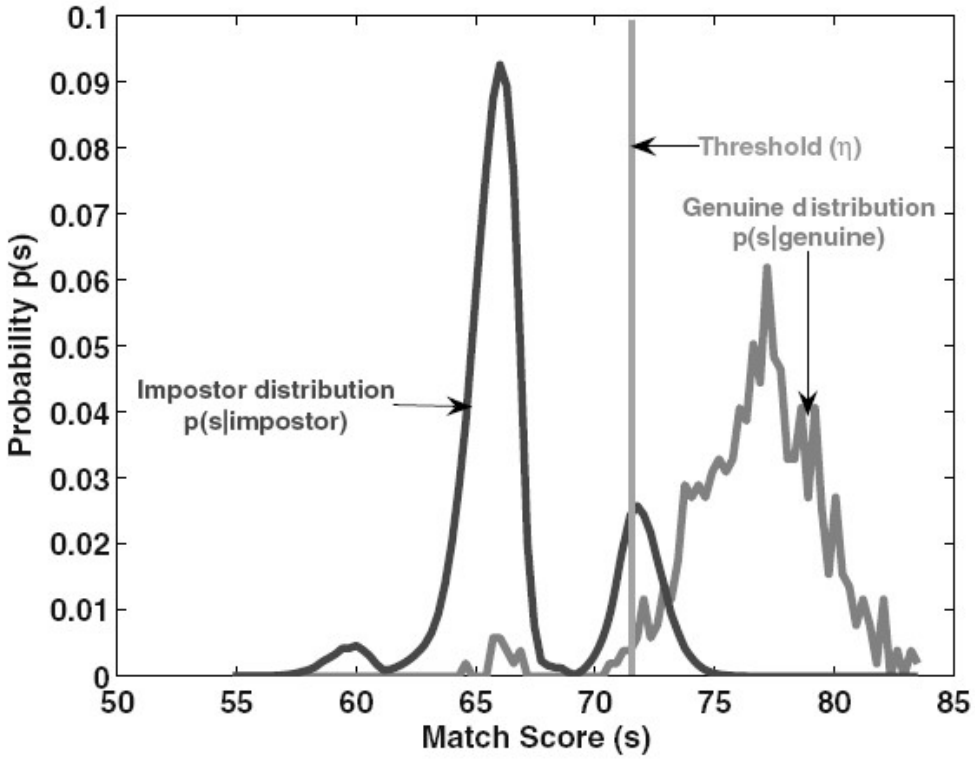
Statistics in user authentication

For identification, you must consider the probabilities that two persons ever have matching authentication data

For verification, you must estimate the probability that an impostor can guess a victim's parameter value and imitate it

Statistics in biometrics

- A typical system has a threshold parameter which determines the allowed variance
- Use statistical theory for hypothesis testing
- Balance user population statistics against intended use plus importance of each of the CIA criteria, and set thresholds accordingly



Failure rates

Admitting a person under the wrong identity

FAR – False Acceptance Rate, also called

FMR – False Match Rate

Rejecting a person claiming correct identity

FRR – False Rejection Rate, also called

FNMR – False Non-Match Rate

Failure rate effects

Remember:

Admitting a person under the wrong identity means damaged Confidentiality and/or Integrity

Rejecting a person claiming correct identity means damaged Availability

Identification effects

Hypothesis testing answers “True” or “False”

Hypothesis can be “this is person X”

Highly unbalanced in the sense that most subjects are not person X

Creates effects that surprise some

Identity testing problems

Suppose there are 10,000 persons on a “no fly” list

An airport uses identification devices with FAR=0,1% and FRR=5%. Reasonable values?

- A terrorist has a 5% chance of passing the check of the “no fly”-list. Send 20 and one will succeed
- A typical airport like Arlanda (≈ 50 000 passengers per day) will detain 50 innocent people each day

Traps in using FRR

False Rejection Rate is a mean value over a trial population

It does not (necessarily) give the general probability that a given user is rejected

Usually there is a subset of users who get most of the rejections

It is not valid for users deliberately trying not to be recognised

Conditional vs mean values

If the correct user is often rejected due to anomalies, attempts at false acceptance as that user may fail often and vice versa. This distorts “true” values

If the attacker knows the statistics of single users, the most likely victim can be chosen

Example 1

A user population has two sets of users, X with excellent characteristics for the biometric system and Y with bad characteristics. 1% belong to Y

A user from X has FAR 0.5%

A user from Y has FAR 50%

Total FAR ≈ 1%

An attack deliberately at a Y person still has 50% probability of succeeding

Example 2

A user population has two sets of users, X with good characteristics for the biometric system and Y with bad characteristics. 1% belong to Y

A user from X has FRR 0.5%

A user from Y has FRR 50%

Total FRR ≈ 1% (looks good, you must re-authenticate only once for every 100 attempts on the average)

Users from Y must re-authenticate every other time when using the system. And they must make three attempts one out of four times etc.

General statistics

How large is the set of possible values?

Are some more likely than others?

How large is the user population?

How many guessing attempts can be made per time unit?

Are there restrictions on the possible number of attempts against the same user?

Are there general restrictions on the number of attempts?

Illustration example, card PIN

A card PIN has 10,000 possible values

The probability to guess a PIN in the usually allowed three consecutive attempts is thus only one in more than 3000

If 3500 cards are stolen each year, at least one misuse through correctly guessed PIN should be expected per year

With 5000 stolen cards, it is more likely that one of them gets its PIN guessed in the first attempt, than that none gets that effect

Remember

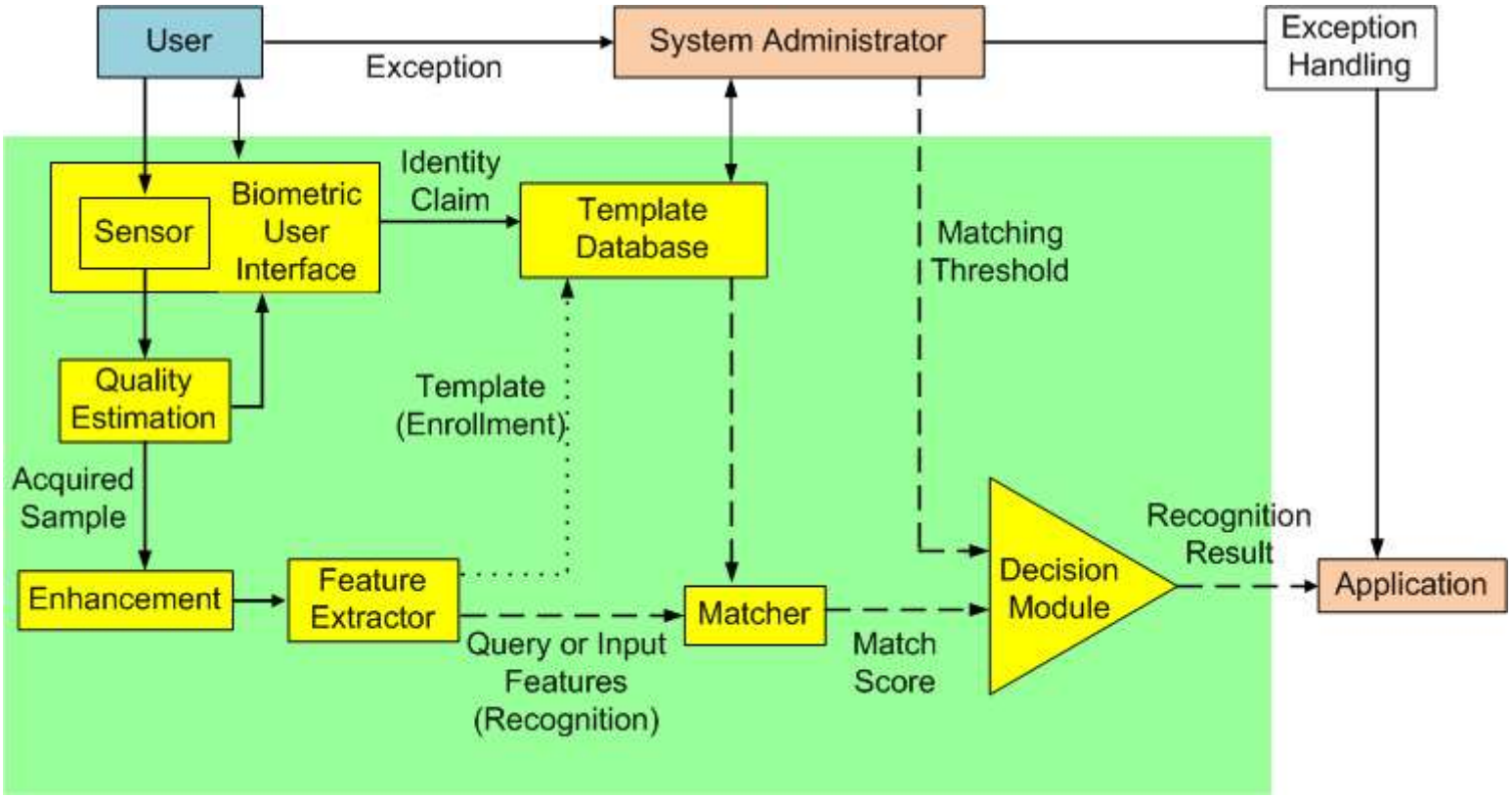
Balance risks against population characteristics, like size but not only size

Average risks can be much higher for subsets of users than for the total population

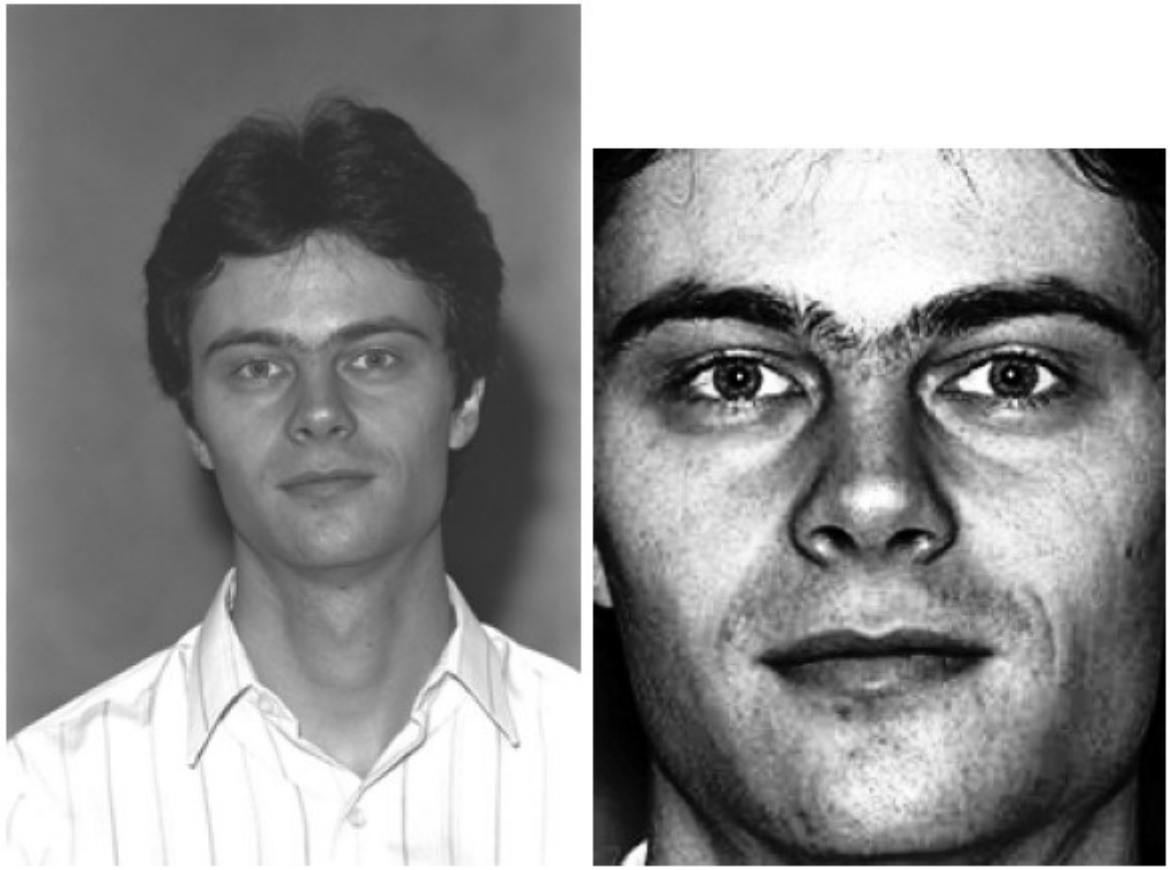
If one single customer is hit, it does not matter to that customer that the average risk per customer was very low

If some customers are at high risk, the organisation is bound to get hit eventually

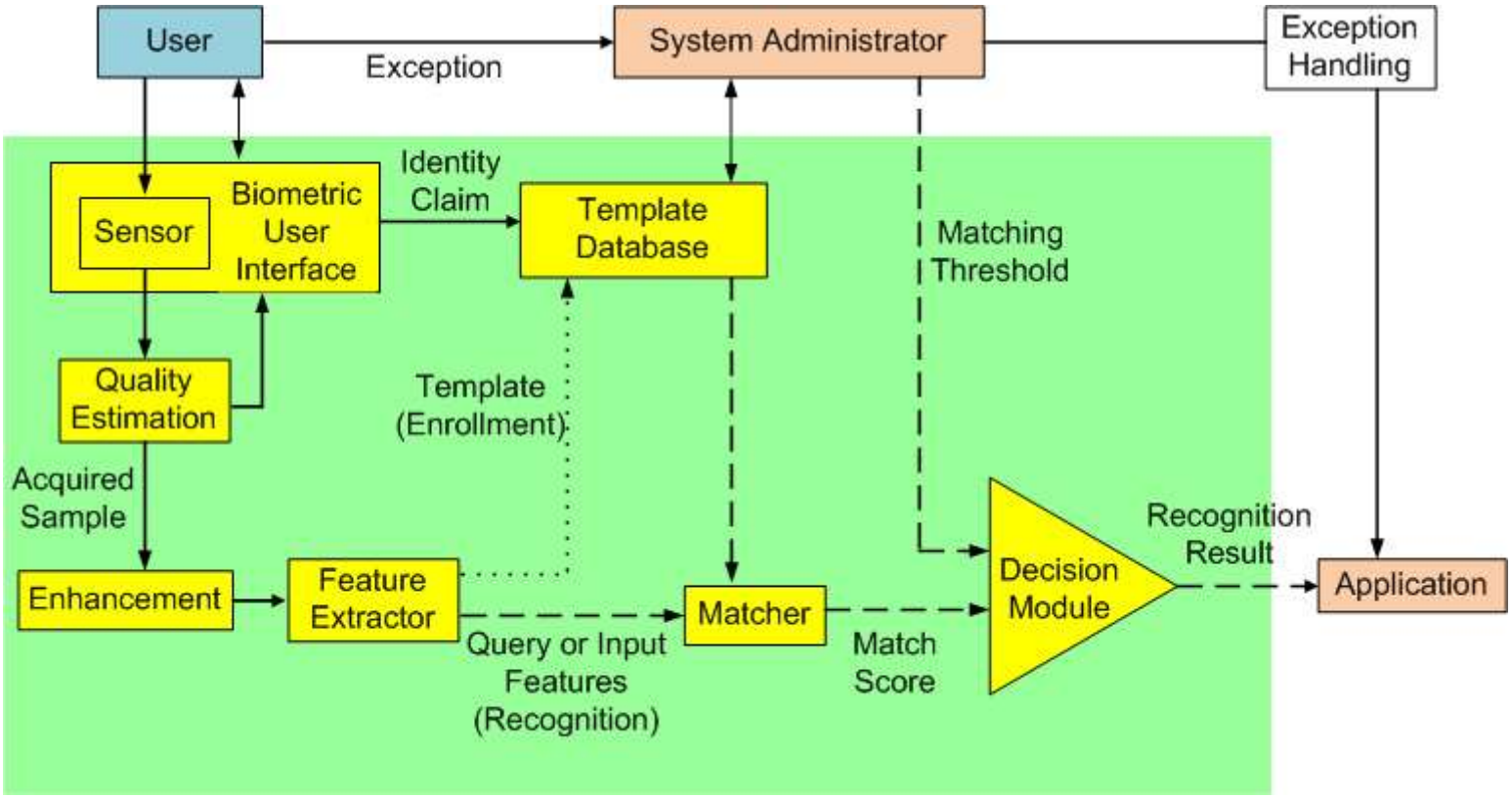
Generic biometric system: Building blocks



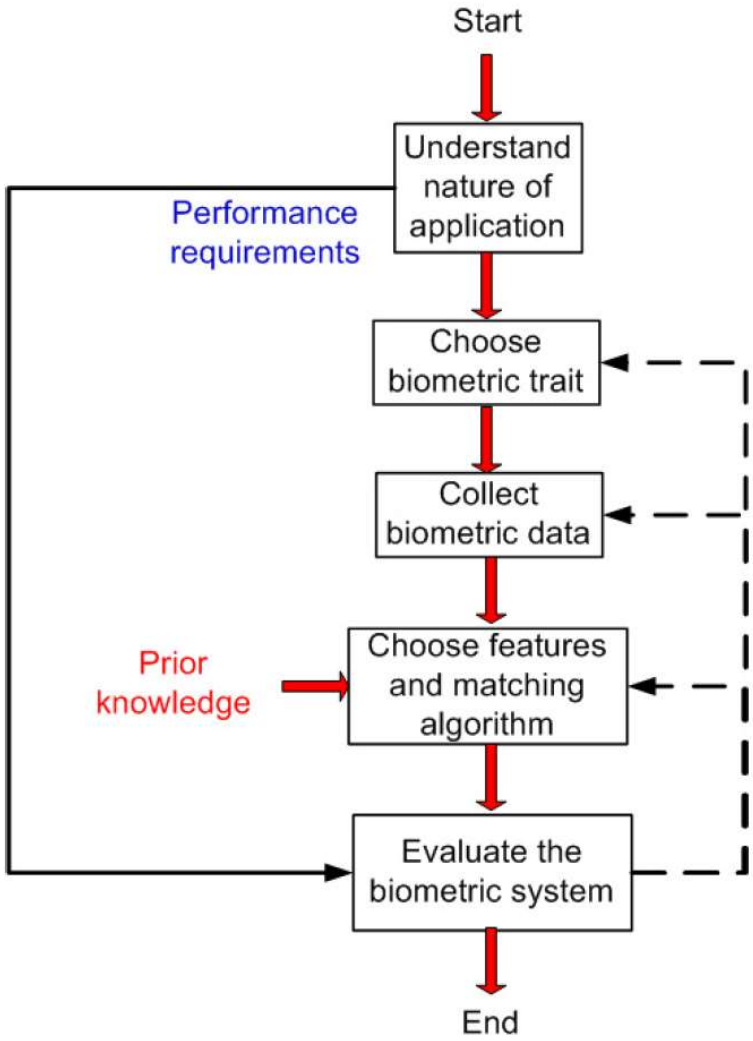
Feature extraction: Segmentation and enhancement



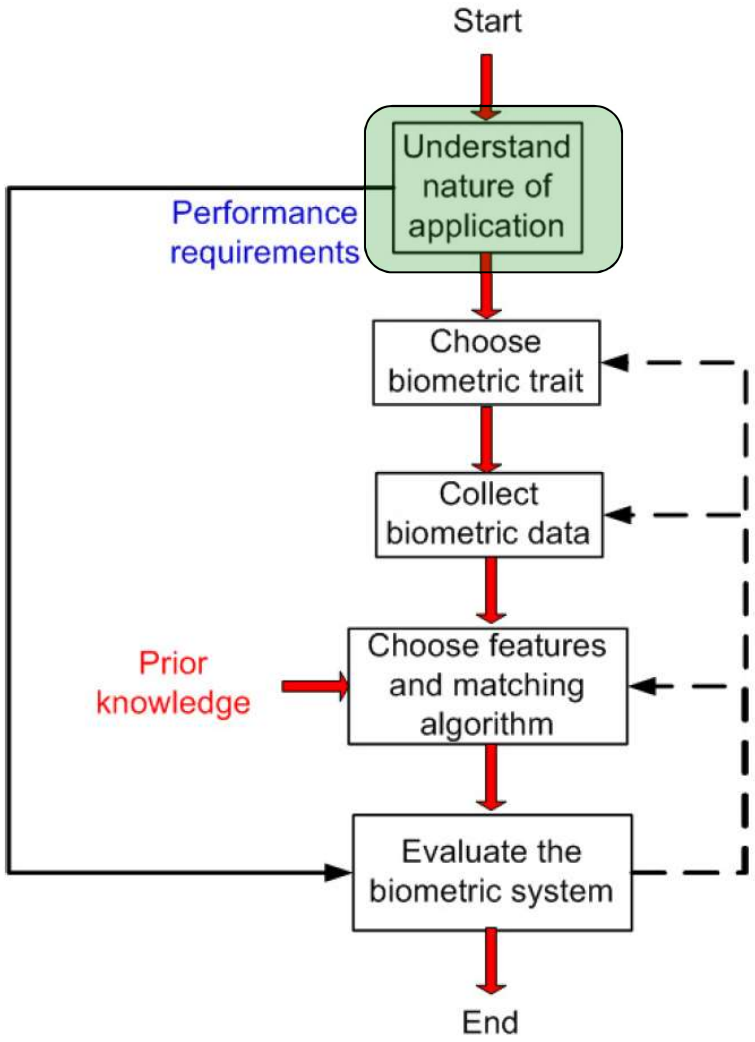
Generic biometric system: Building blocks



Design cycle of biometric systems



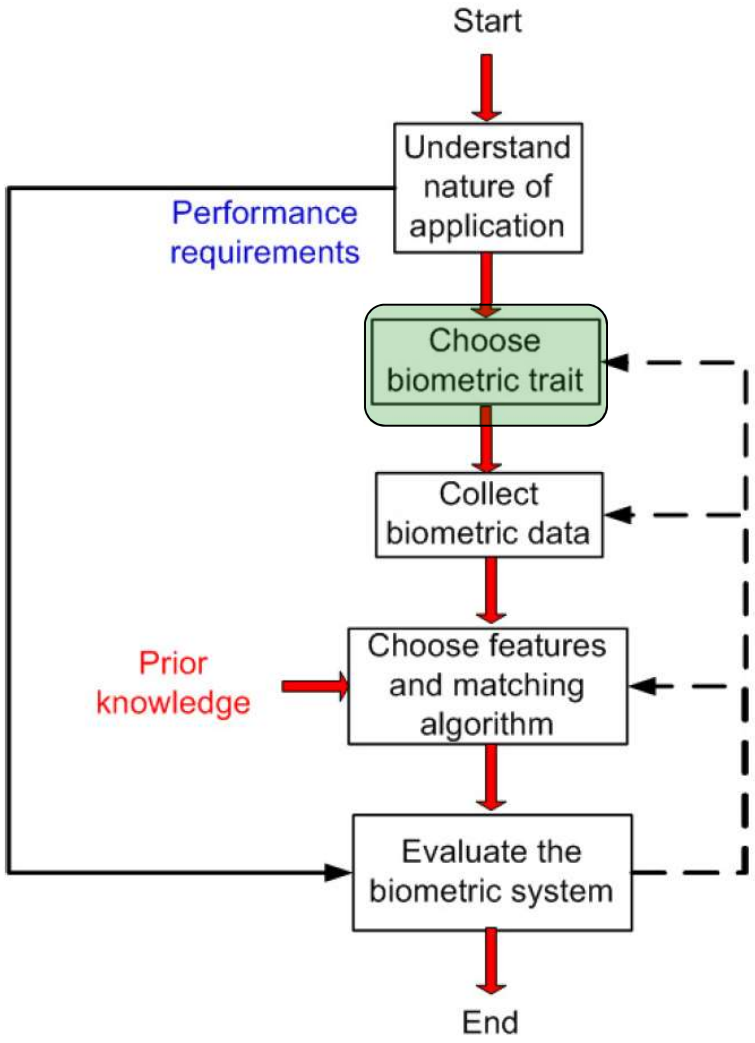
Design cycle of biometric systems



Nature of application

- Cooperative users
- Overt/covert deployment
- Habituated/Non-habituated users
- Attended/Unattended operation
- Controlled/Uncontrolled operation
- Open/Closed system

Design cycle of biometric systems



Choice of biometric trait

- Universality
- Uniqueness
- Permanence
- Measurability (Collectability)
- Performance
- Acceptability
- Circumvention

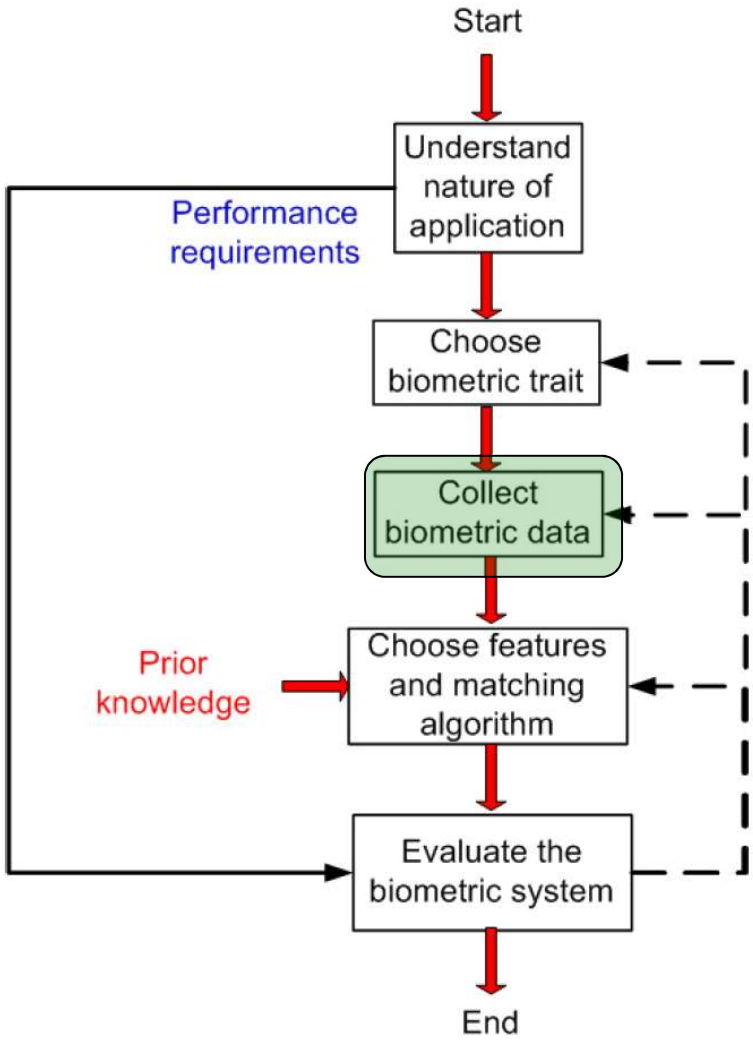
Requirements on biometric traits

Biometrics	Univer-sality	Unique-ness	Perma-nence	Collect-ability	Perfor-mance	Accept-ability	Circum-vention
Face	H	L	M	H	L	H	L
Fingerprint	M	H	H	M	H	M	H
Hand Geometry	M	M	M	H	M	M	M
Keystroke Dynamics	L	L	L	M	L	M	M
Hand vein	M	M	M	M	M	M	H
Iris	H	H	H	M	H	L	H
Retina	H	H	M	L	H	L	H
Signature	L	L	L	H	L	H	L
Voice	M	L	L	M	L	H	L
Facial Thermogram	H	H	L	H	M	H	H
DNA	H	H	H	L	H	L	L

H=High, M=Medium, L=Low

Attempt to classify methods according to how they meet all seven criteria. Valid today? Do you agree in general? Look closely and make your own assessment! There is no “correct” answer…

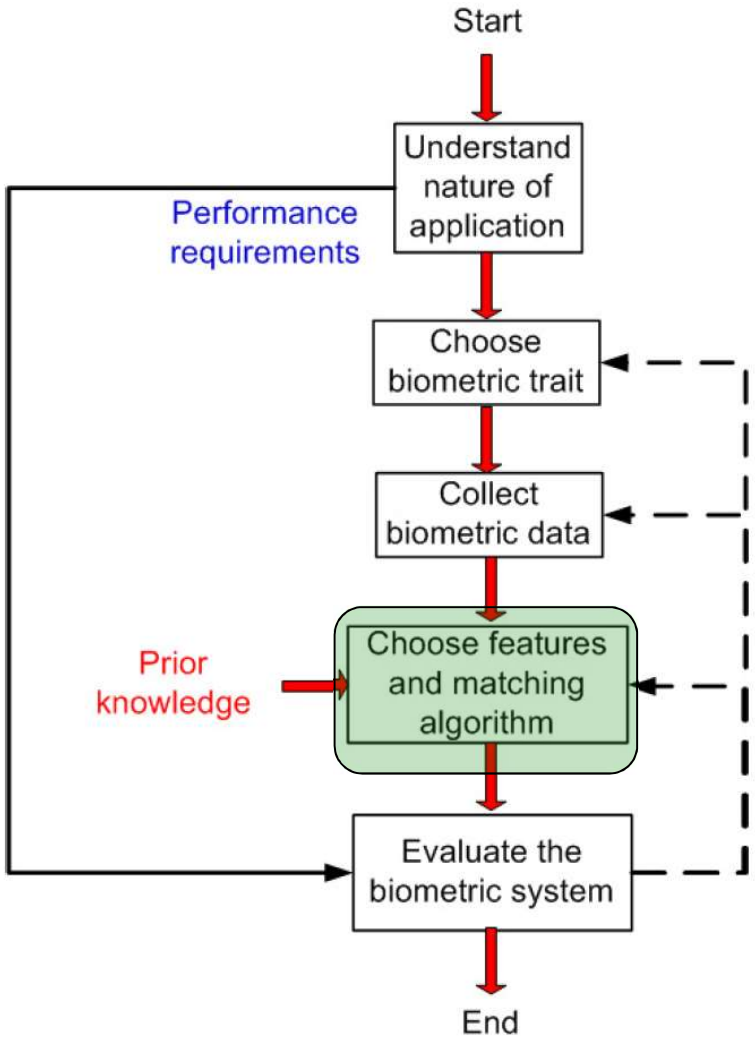
Design cycle of biometric systems



Collecting biometric data

- Appropriate sensors
  - Size, cost, ruggedness, high quality biometric samples
- Collection environment
- Sample population
  - Representative of the population
  - Exhibit realistic intra-class variations
- User habituation
- Legal, privacy & ethical issues

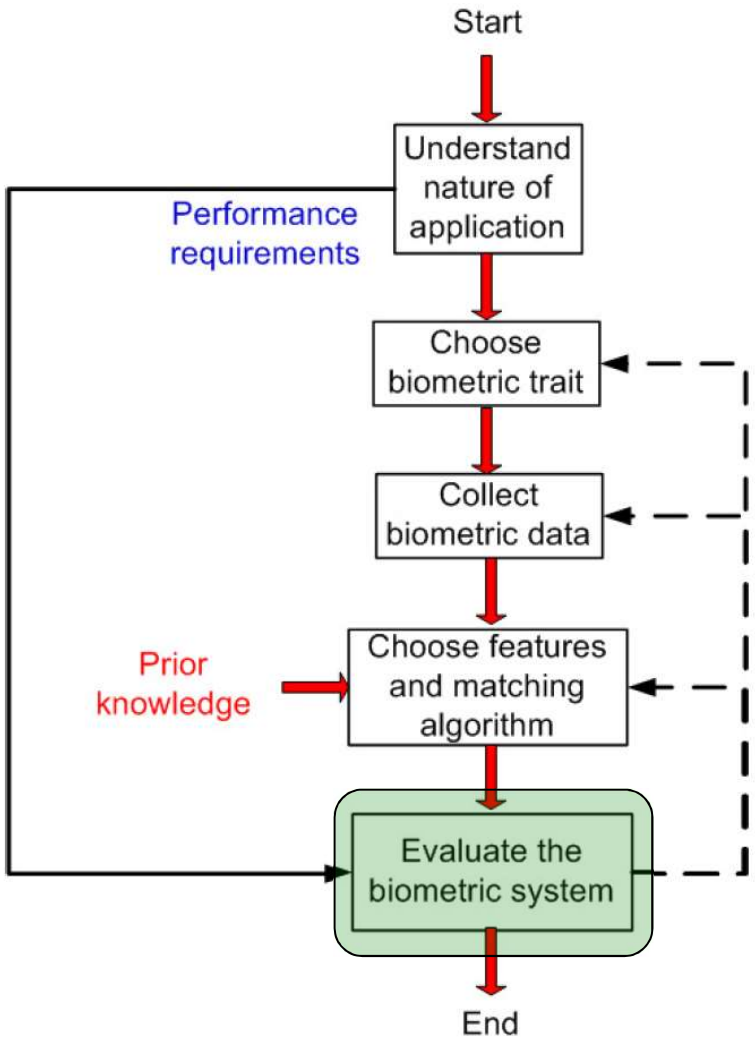
Design cycle of biometric systems



Choice of features/matching algorithm

- Prior knowledge of the biometric trait
  - Uniqueness
- Mimic human ability to discriminate
- Interoperability between biometric systems
- Common data exchange formats …

Design cycle of biometric systems

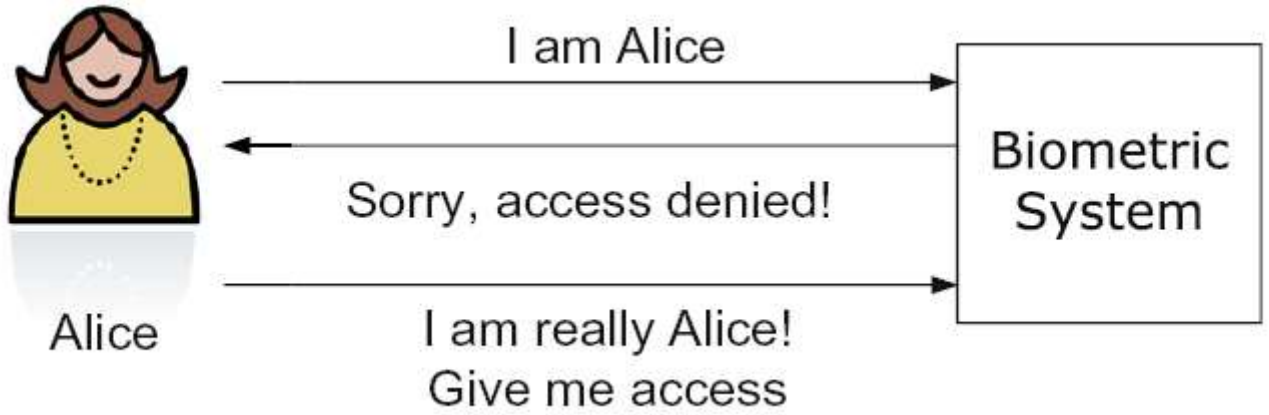


Evaluation of biometric systems

- Technology evaluation
- Scenario evaluation
- Operational evaluation

- Error rates
- System reliability, availability, maintainability
- Vulnerabilities
- User acceptability
- Cost, throughput, benefits
- Return on investment

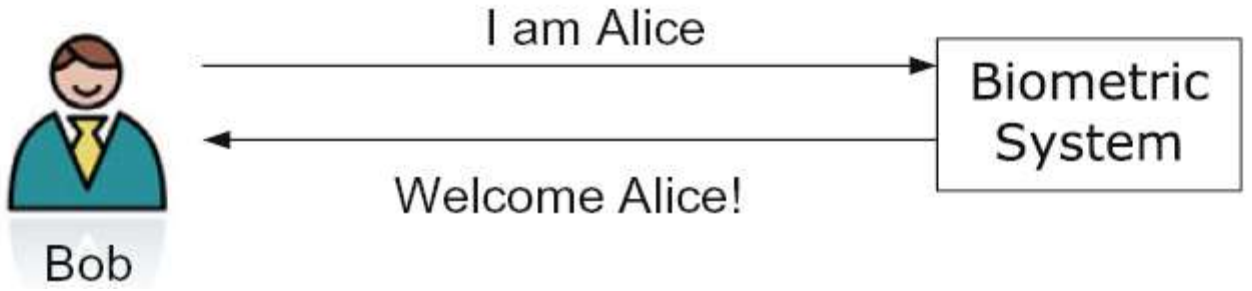
Security threats: Denial-of-service (DoS)



Legitimate users are prevented from obtaining access to the system or resource that they are entitled to

Violates availability

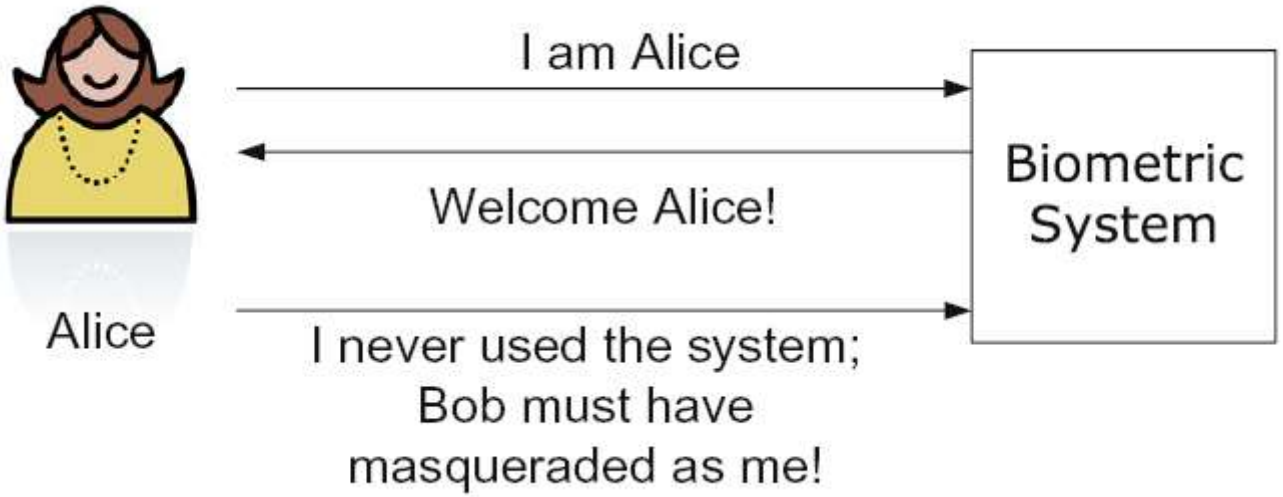
Security threats: Intrusion



An unauthorized user gains illegitimate access to the system

Affects integrity of the biometric system

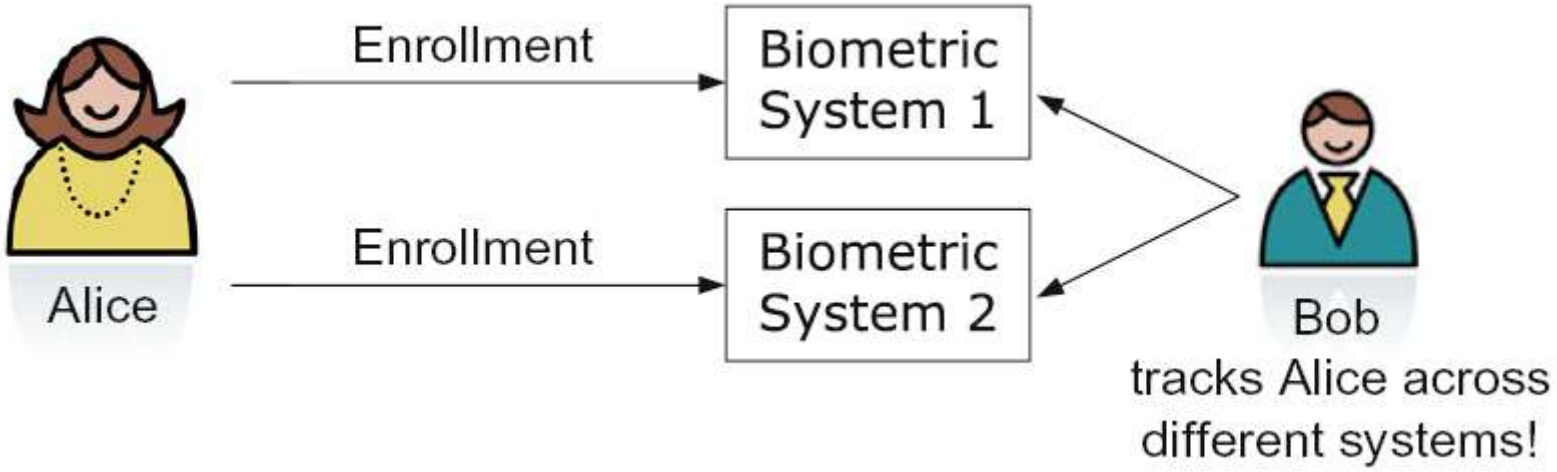
Security threats: Repudiation



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

Security threats: Function creep



An adversary exploits the biometric system designed to provide access control to a certain resource to serve another application, 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

Violates confidentiality and privacy.

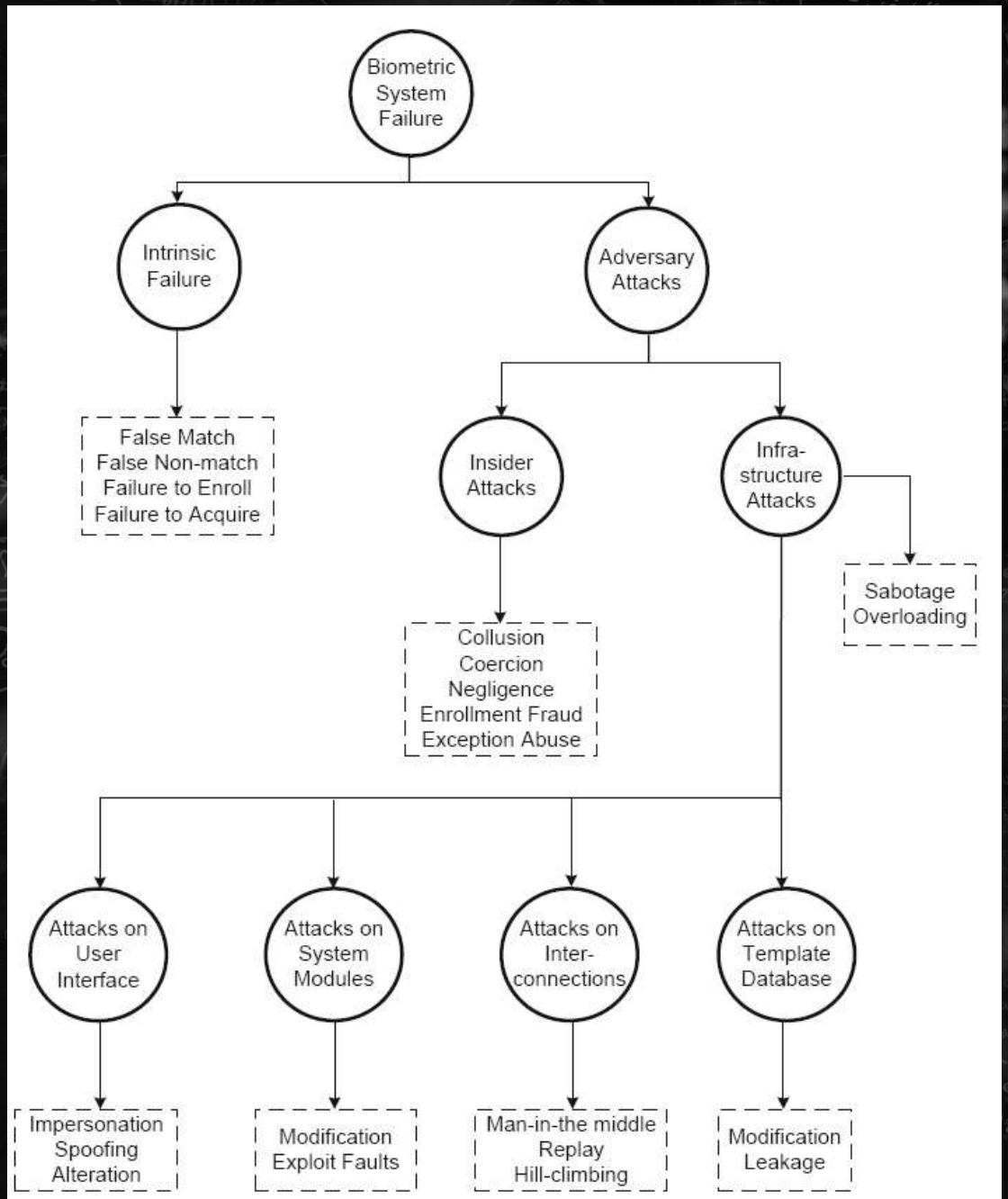
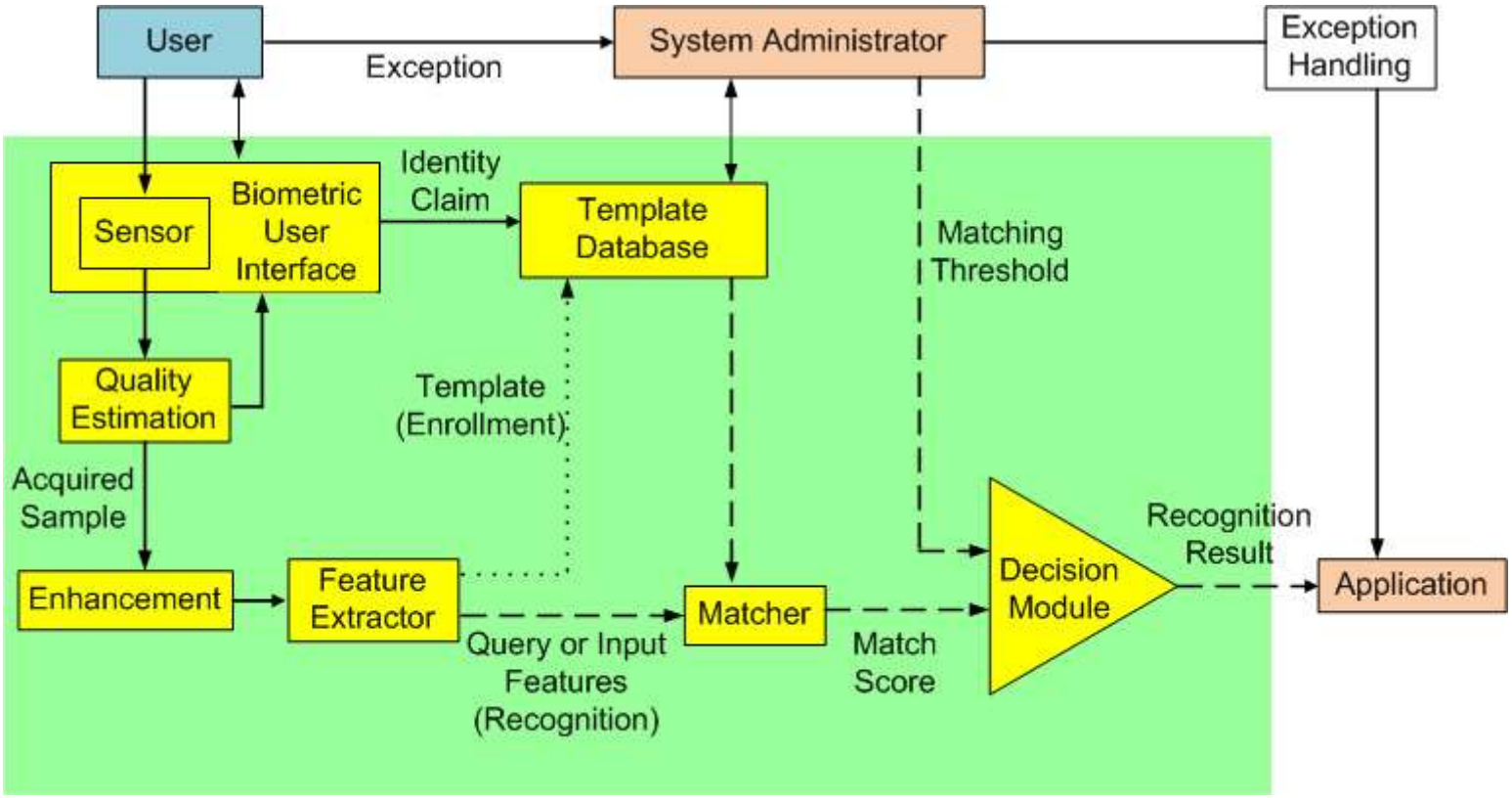


Fig. 7.2 Taxonomy of attacks that can be mounted against a biometric system.

Generic biometric system: Building blocks



Types of adversary attacks

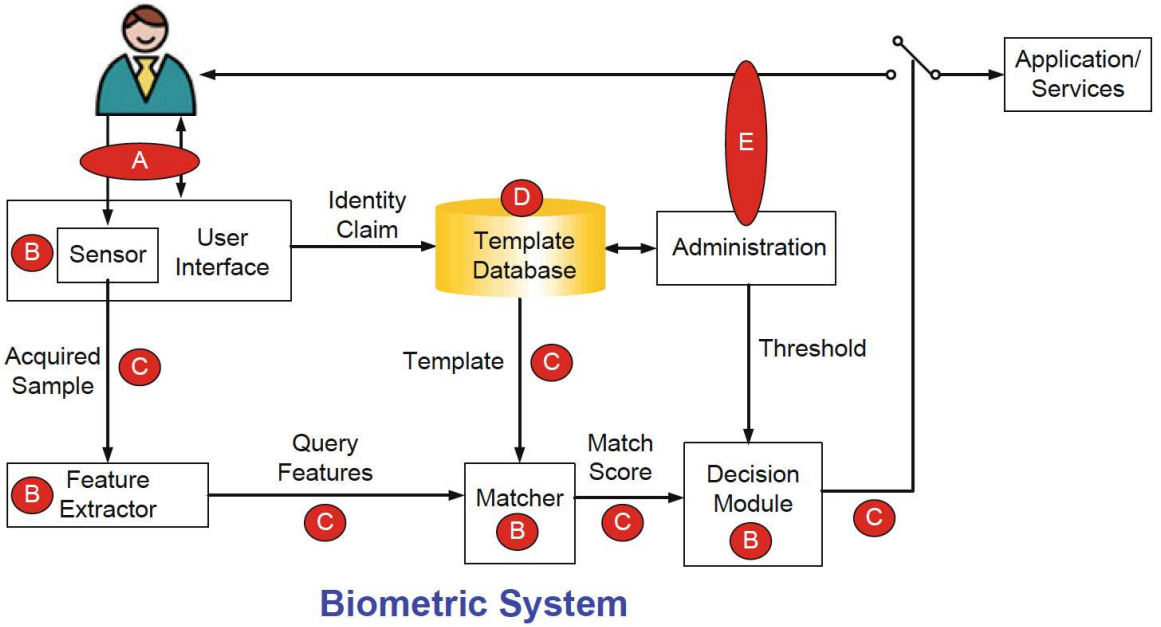
A: User-biometric system interface

B: Biometric system modules

C: Interconnections betweeen biometric modules

D: Templates database

E: Attacks through insiders (admin or enrolled users)



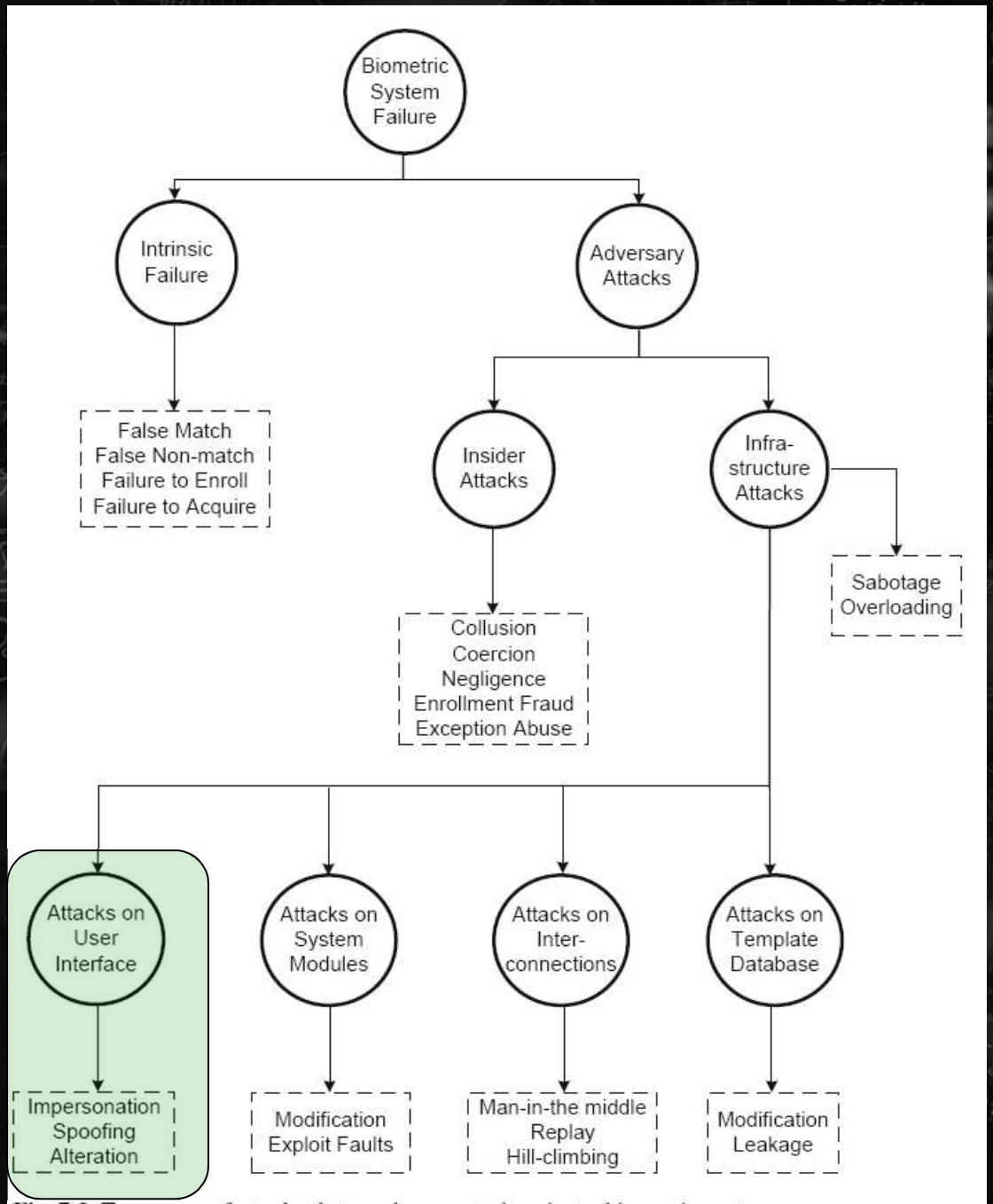
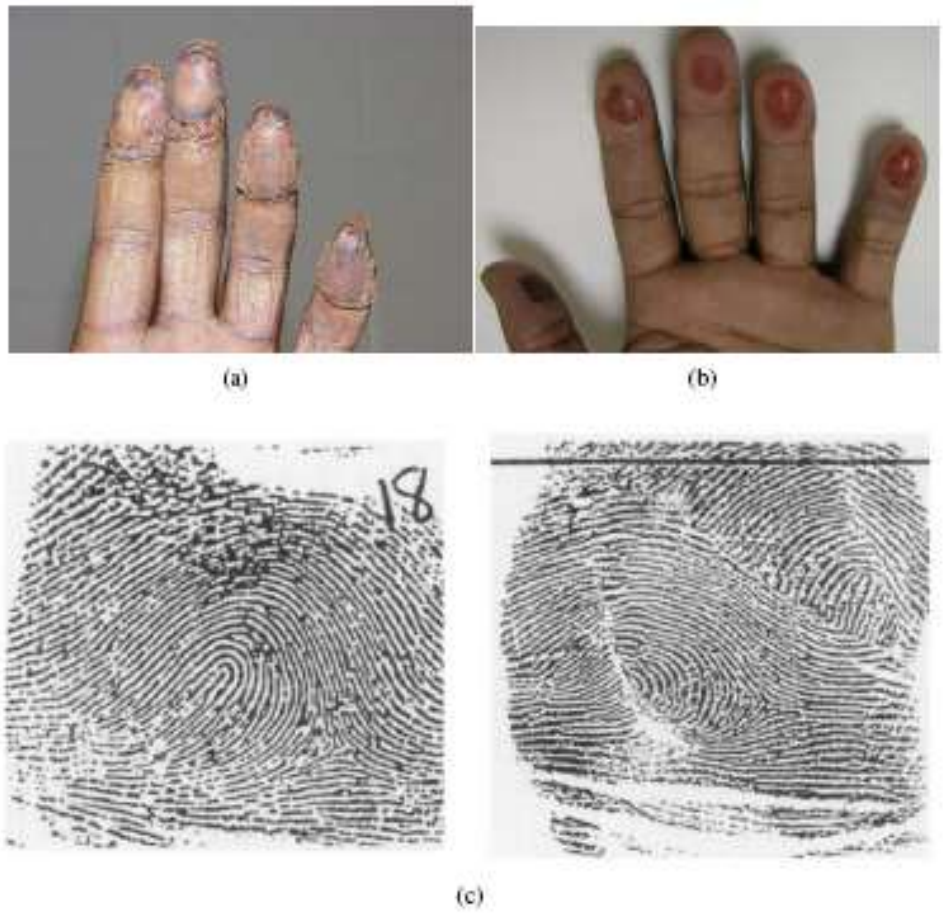


Fig. 7.2 Taxonomy of attacks that can be mounted against a biometric system.

Attacks at the user interface: Obfuscation



Attacks at the user interface: Spoofing

(a)

(b)

(c)

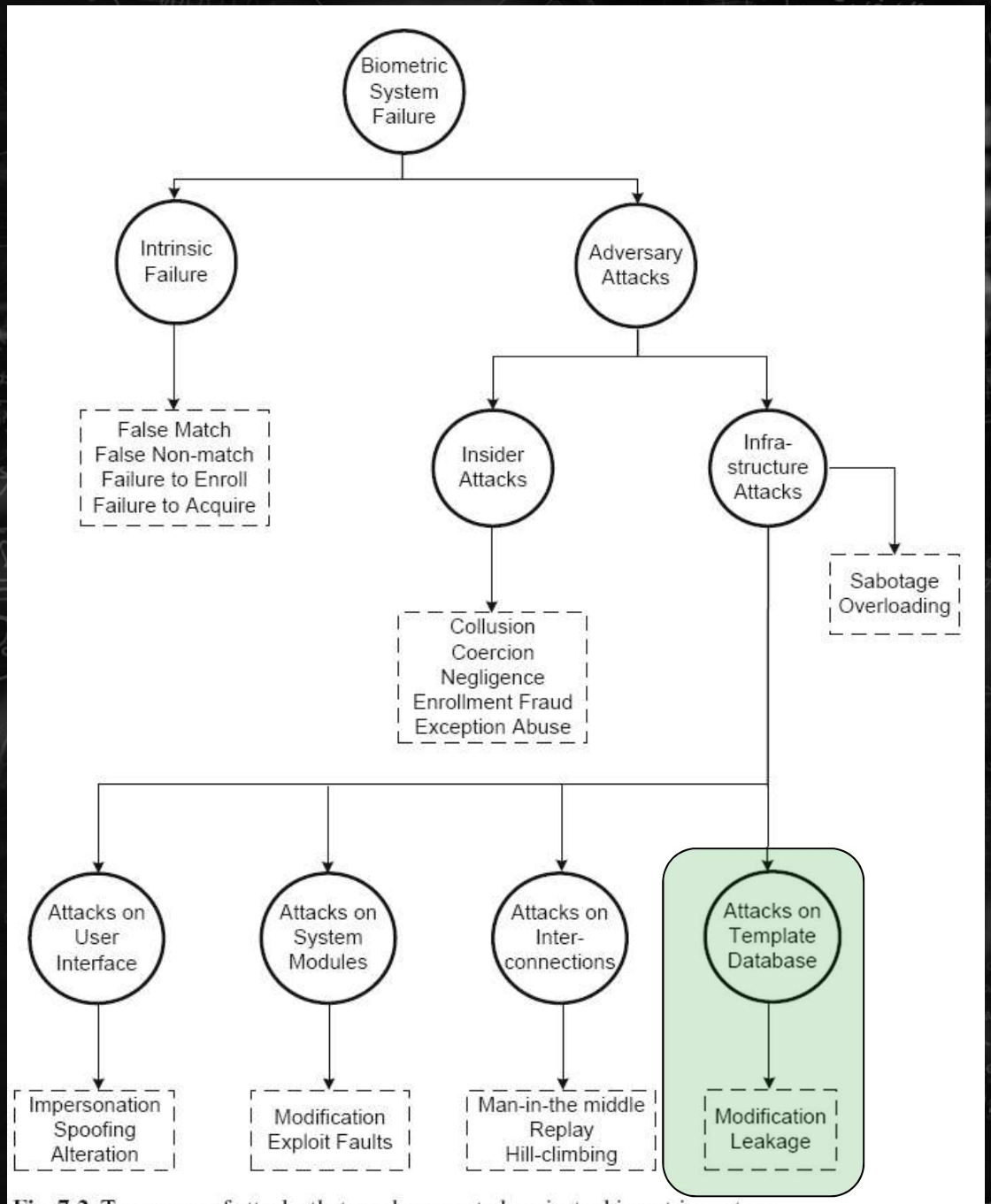


Fig. 7.2 Taxonomy of attacks that can be mounted against a biometric system.

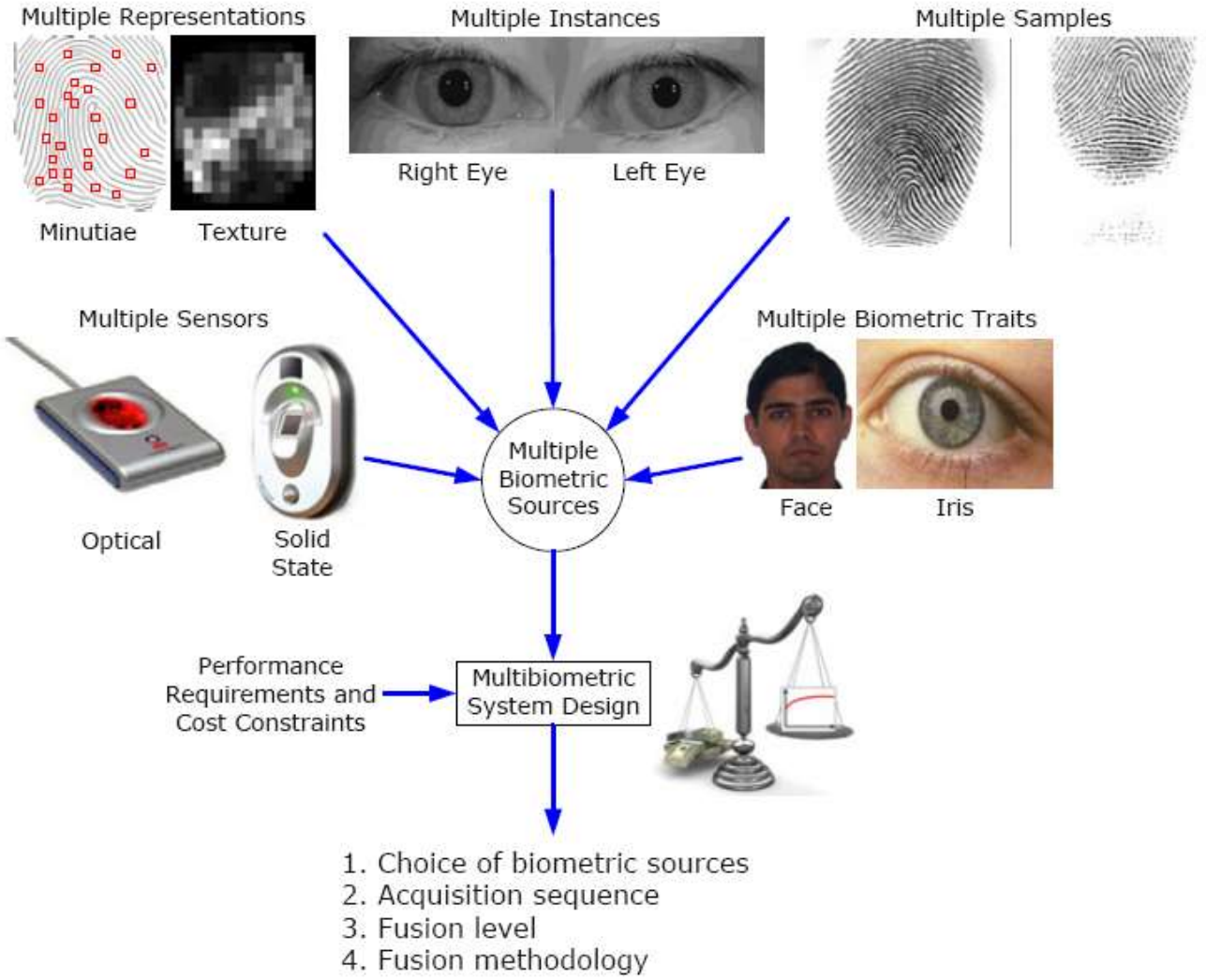
Attacks on the template database

- Gain unauthorized access/Deny access to legitimate users
- Leakage: Stored biometric templates available to adversaries
  - Password-based authentication: Hashed,minor problem
  - Biometrics based: Major problem
    - Biometrics not always secret
    - Physical link user/biometric trait

Attacks on the template database: Leakage

- Obtain biometric & biographic info about large number of users
- Reverse engineer template: Physical spoof
- Replay attack
- Compromised biometric traits: Not possible to replace
- Undermines privacy

Multibiometrics



Multibiometrics: Why?

- More unique (than single)
- Compensate noise, imprecision, inherent drift
- Redundancy
- Fault-tolerance
- Flexibility
- Increase resistance to spoofing
- But: Expensive – Tradeoff cost/benefits

Multi-modal systems

Use two or more different biometric features

AND or OR requirements for each feature

AND increases accuracy and thus protects against false acceptance

OR opens more options and thus protects against too much false rejection

OR is necessary in order to accommodate for physical handicaps

Multiple methods

Use of two or three of the basic categories (what you “know”, “hold” and “are”).

Thus use of something you know or hold in addition to biometrics (or just something you know and something you hold)

Examples:

PIN + card

Fingerprints + card with fingerprint template

Multibiometric sources	Type of information fused				Acquisition architecture		Processing architecture	
	Raw data	Features	Scores	Decisions	Serial	Parallel	Serial	Parallel
Multiple sensors	✓	✓	✓	✓	✓	✓	✓	✓
Multiple representations	×	✓	✓	✓	×	✓	✓	✓
Multiple matchers	×	×	✓	✓	×	✓	✓	✓
Multiple instances	×	✓	✓	✓	✓	✓	✓	✓
Multiple samples	✓	✓	✓	✓	✓	×	✓	✓
Multiple traits	×	✓	✓	✓	✓	✓	✓	✓

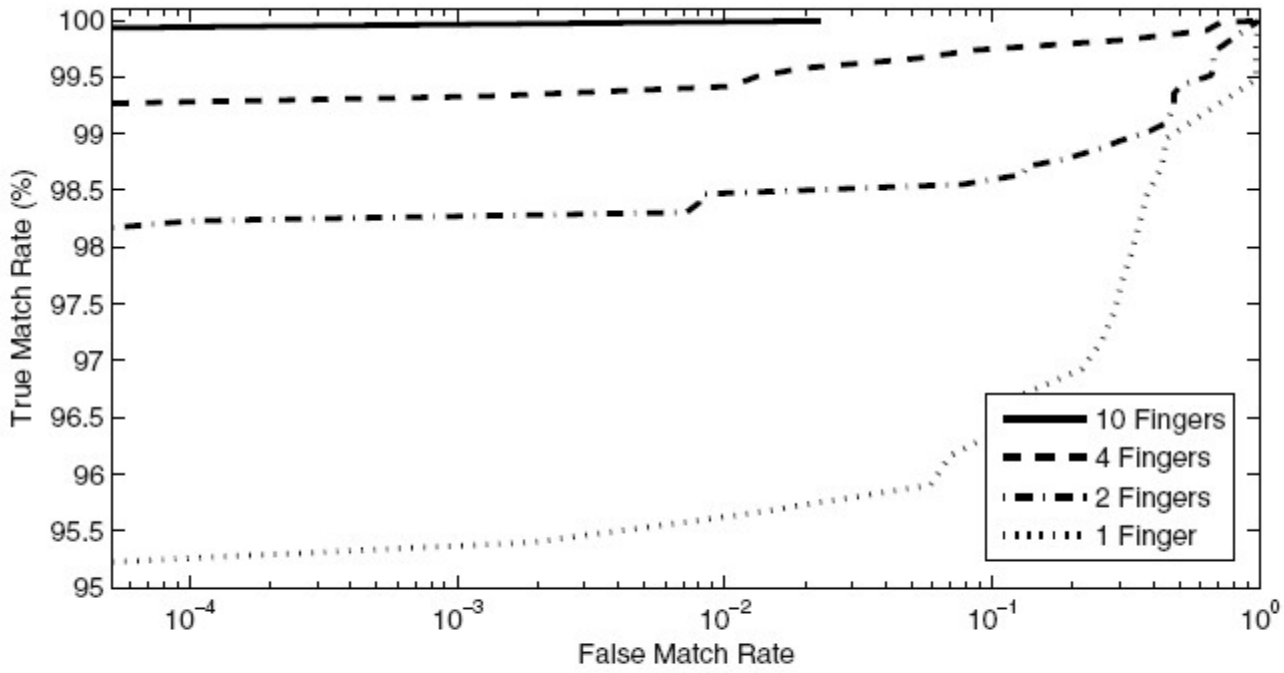
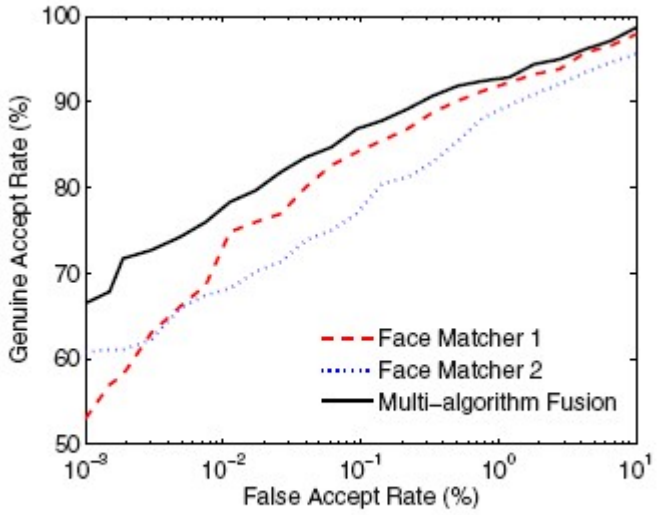
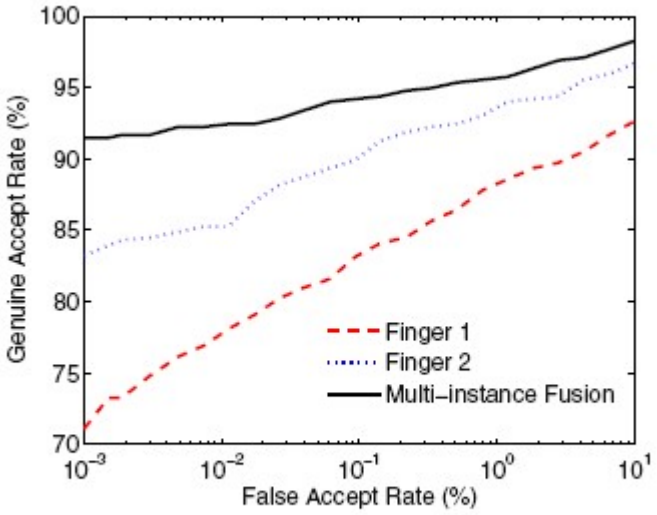


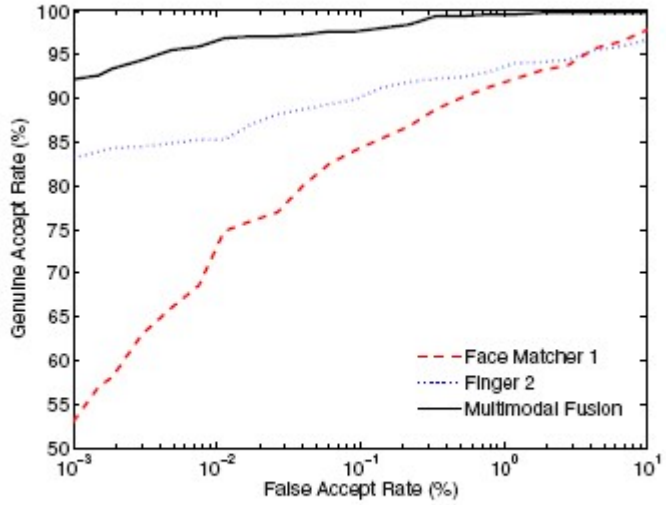
Fig. 6.7 Accuracy improvement in a multi-instance fingerprint verification system. This example extracted from FpVTE 2003 [37] shows that fusion of evidence from the multiple fingers leads to significant improvement in the accuracy.


Genuine Accept Rate (%)
False Accept Rate (%)
Face Matcher 1
Face Matcher 2
Multi–algorithm Fusion

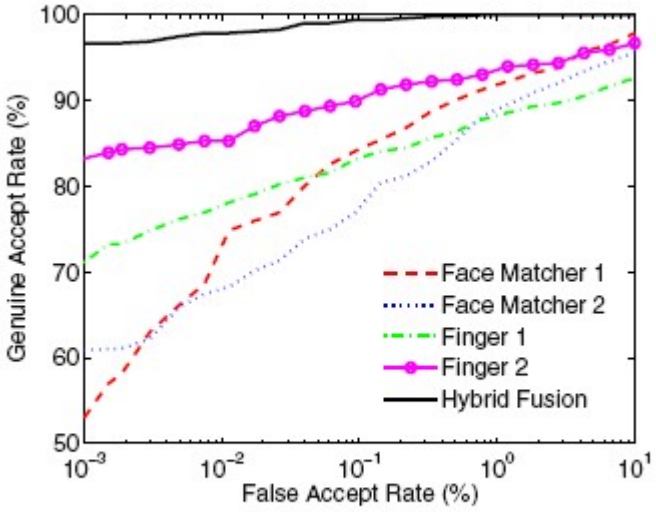
(a)


Genuine Accept Rate (%)
False Accept Rate (%)
Finger 1
Finger 2
Multi–instance Fusion

(b)


Genuine Accept Rate (%)
False Accept Rate (%)
Face Matcher 1
Finger 2
Multimodal Fusion

(c)


Genuine Accept Rate (%)
False Accept Rate (%)
Face Matcher 1
Face Matcher 2
Finger 1
Finger 2
Hybrid Fusion

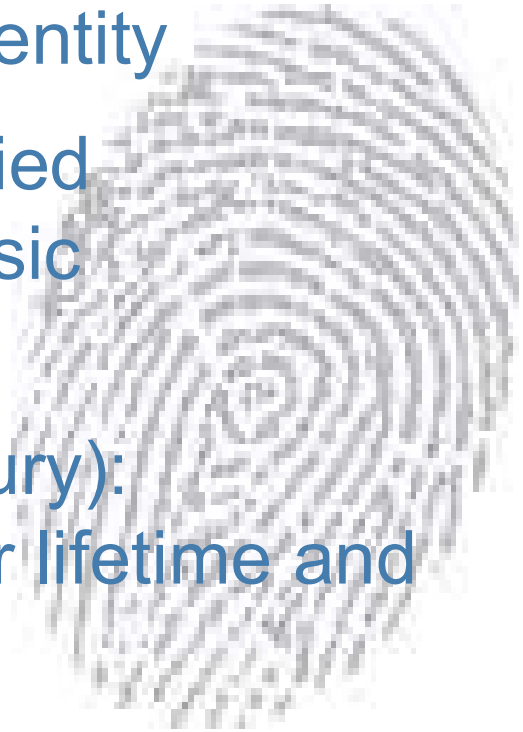
(d)

Fingerprints - history

Already in ancient times fingerprints were used to denote authorship or identity

In 1823 a Czech physician classified fingerprint patterns into nine basic types

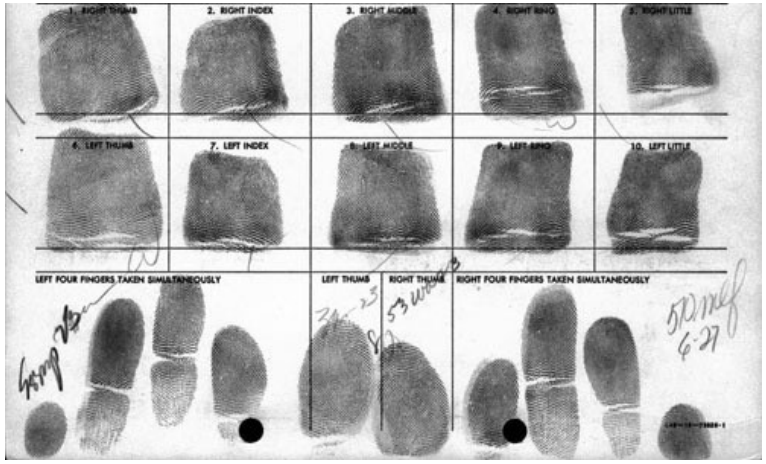
Sir Francis Galton (late 19th century): Fingerprints do not change over lifetime and that no two fingerprints are exactly alike



Fingerprints - history

In 1901 fingerprints were
introduced
for criminal identification
in
England and Wales

The first fingerprint
scanners
were introduced more
than
30 years ago



AFIS installation at Michigan State Police facility. This system was first installed in 1989; the database has 3.2 *million tenprint cards and performs 700,000 searches each year*

Example: Fingerprints

Known and used with formal classification since 19th century.

Cheap readers that are easy to handle

High uniqueness

Fairly easy to make copies

Fingerprints - characteristics

Papillary lines
- ridges
- valleys

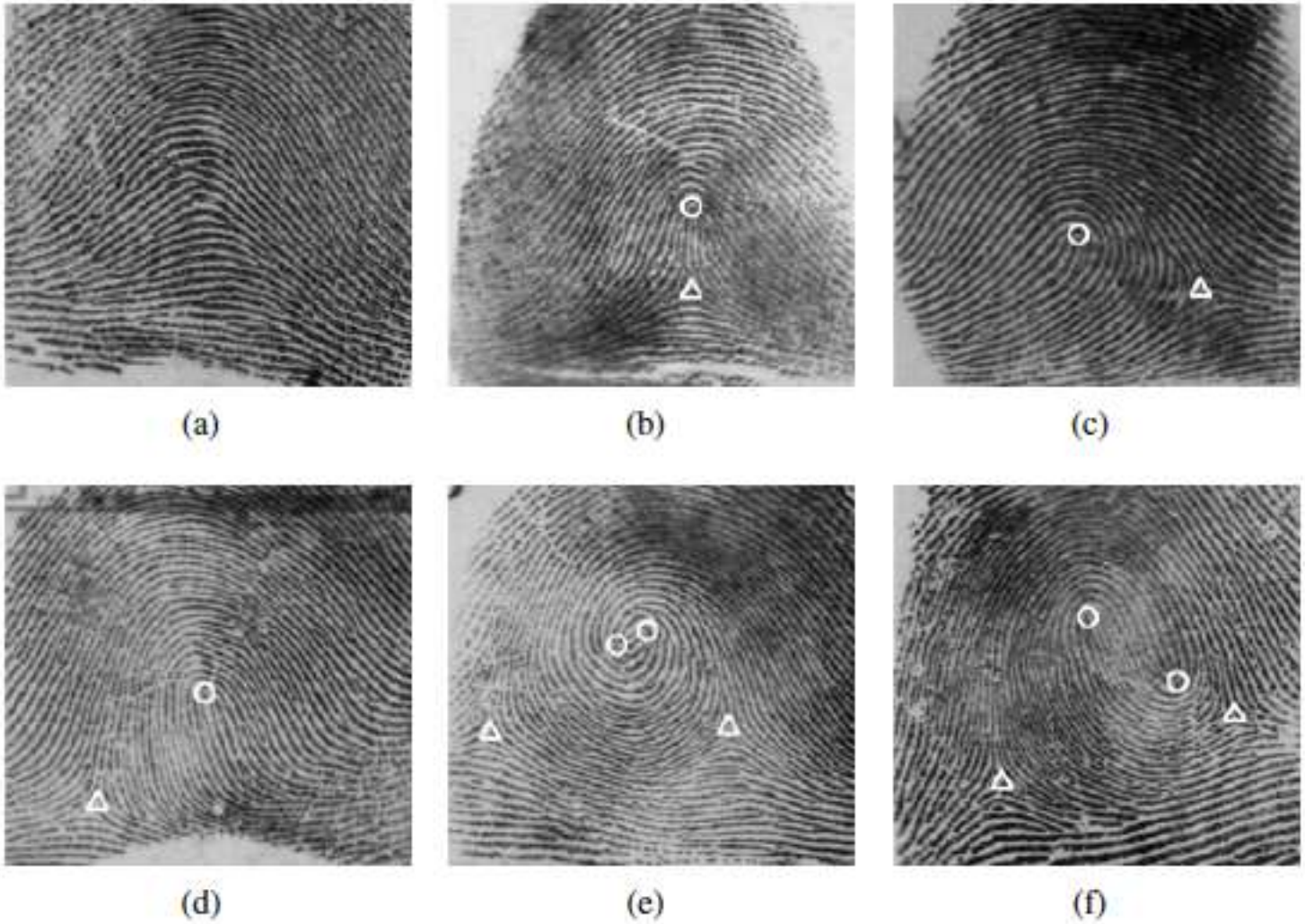


Fig. 2.8 Major fingerprint pattern types. (a) Plain arch, (b) tented arch, (c) left loop, (d) right loop, (e) whorl, and (f) twin loop. A loop is denoted by a circle and a delta is denoted by a triangle. Loop- and whorl-type of fingerprints are found most commonly; about 65% of fingerprints belong to loop-type, and 24% are whorl-type [52]. Twin loop, arch and tented arch account for approximately 4%, 4% and 3% of the fingerprints, respectively.

3 levels of fingerprint features

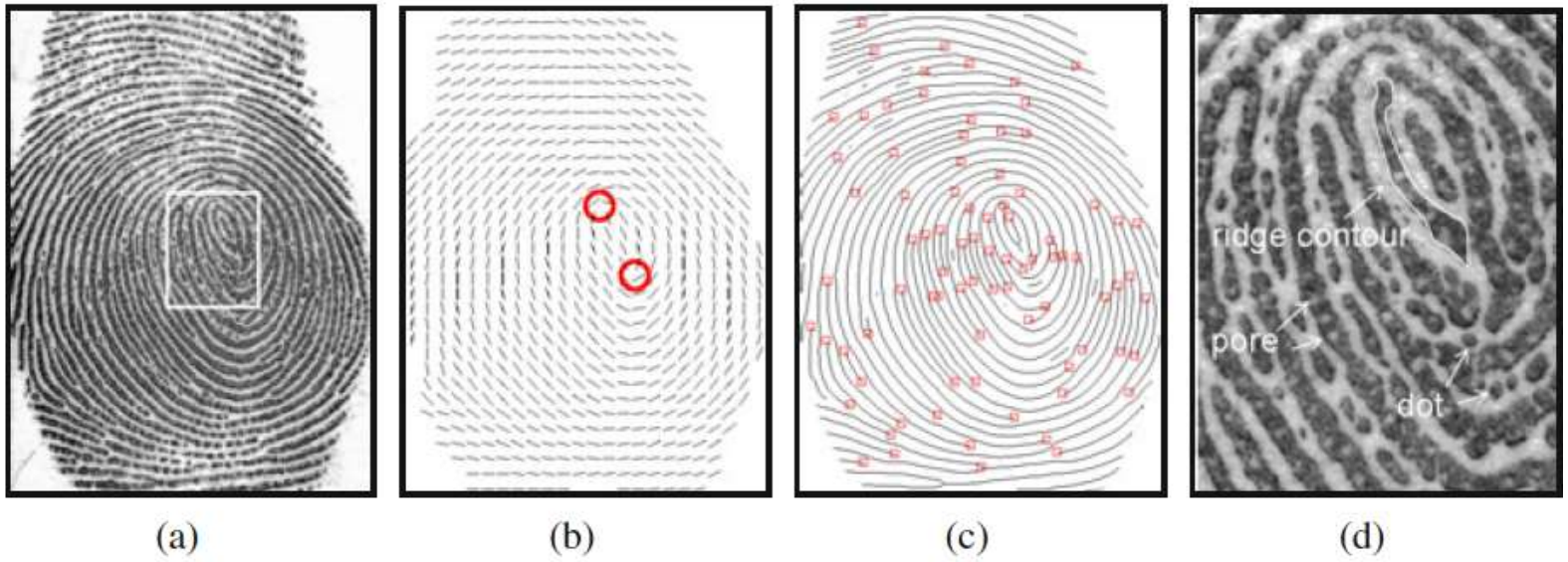


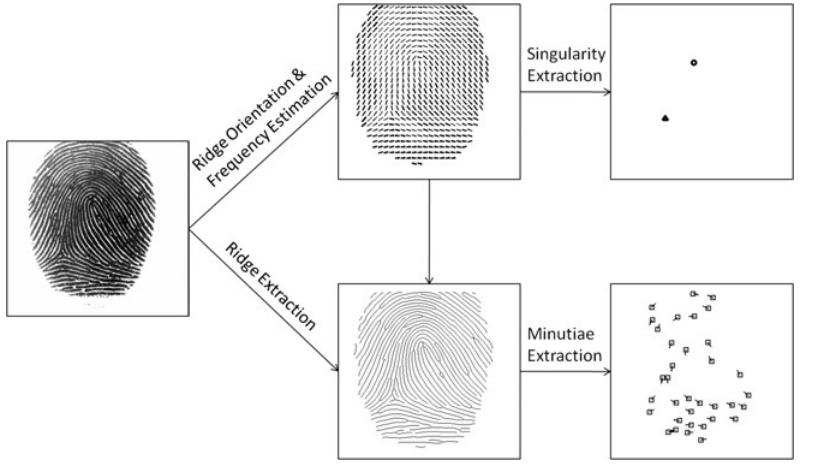
Fig. 2.5 Features at three different levels in a fingerprint. (a) Grayscale image (NIST SD30, A067_11), (b) Level 1 feature (orientation field or ridge flow and singular points), (c) Level 2 feature (ridge skeleton), and (d) Level 3 features (ridge contour, pore, and dot).

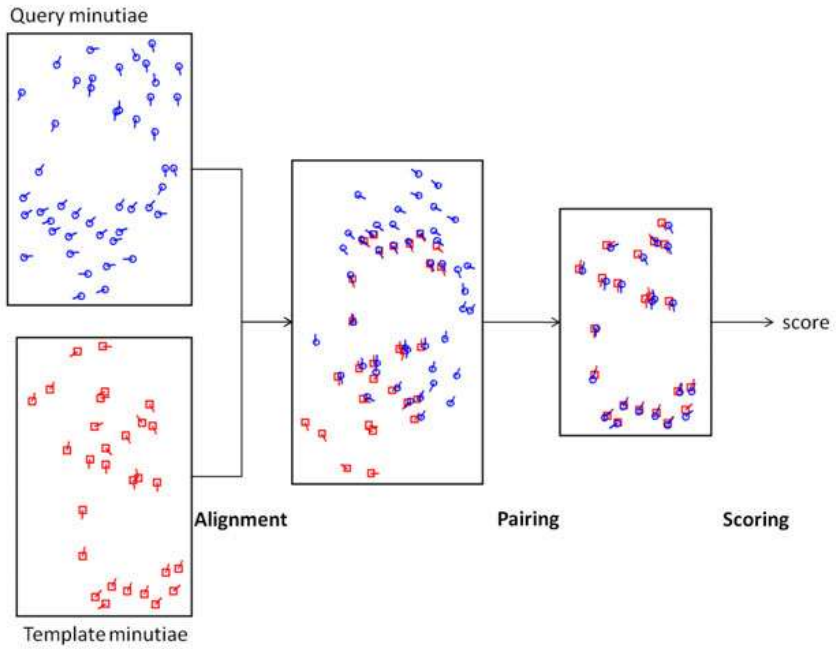
Fingerprints - characteristics

Pattern types
- arches
- loops
- whorls

Core and delta points

Minutiae points





Fingerprints -scanners

Optical scanner

Solid-state scanner
(capacitive sensors)

Ultrasound scanner

Fingerprints – scanners

- Good accuracy
- Used for both identification and verification
- Low cost
- Problem when skin is too dry or too wet
- Problem with dirt

Fingerprints - scanners

Touch (area) sensor

- Quickly becomes dirty
- Problem with latent prints
- Rotation problems
- Area vs cost

Sweep

- Reduced cost
- No dirt or latent prints
- Longer learning time
- Reconstruction of the image is time consuming

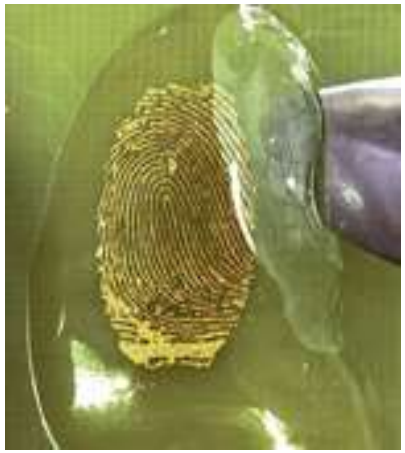
Fingerprints - attacks

Making a user cooperate using force or drugs

Using latent fingerprints

Artificial fingerprint

Gummy fingers



Making an Artificial Finger directly from a Live Finger

How to make a mold

Put the plastic into hot water to soften it.

Press a live finger against it.

The mold

It takes around 10 minutes.

Making an Artificial Finger directly from a Live Finger

Preparation of material

- A liquid in which immersed gelatin at 50 wt.% .



Add boiling water (30cc) to solid gelatin (30g) in a bottle and mix up them.

It takes around 20 minutes.

Making an Artificial Finger directly from a Live Finger

How to make a gummy finger

Pour the liquid
into the mold.

Put it into
a refrigerator to cool.

The gummy finger

It takes around 10 minutes.

- **”Researchers warn of fingerprint theft from ‘peace’ sign”, https://phys.org/news/2017-01-japan-fingerprint-theft-peace.html**
  - **Mobile device w. Camera**
  - **Up to 3 m distance**
  - **Countermeasure: Transparent film with titanium oxide on your fingers!**

- ”Hacker claims you can steal fingerprints with only a camera - Previous attempts to copy fingerprints required specialized tools and the fingerprint itself.”, https://www.cnet.com/news/hacker-claims-you-can-steal-fingerprints-with-only-a-camera/

Gummy fingers results

Real fingerprints	User 1	User 2	User 3
Reader 1	98%	100%	94%
Reader 2	100%	100%	100%
Reader 3	98%	34%	88%

Gummy fingerprint copies	User 1	User 2	User 3
Reader 1	98%	92%	100%
Reader 2	98%	100%	96%
Reader 3	92%	12%	82%

Fingerprint - liveness 1

Skin deformation

Pores

Perspiration

Fingerprint - liveness 2

Temperature

Optical properties

Pulse

Blood pressure

Electric resistance

Detection under epidermis

Example: Iris

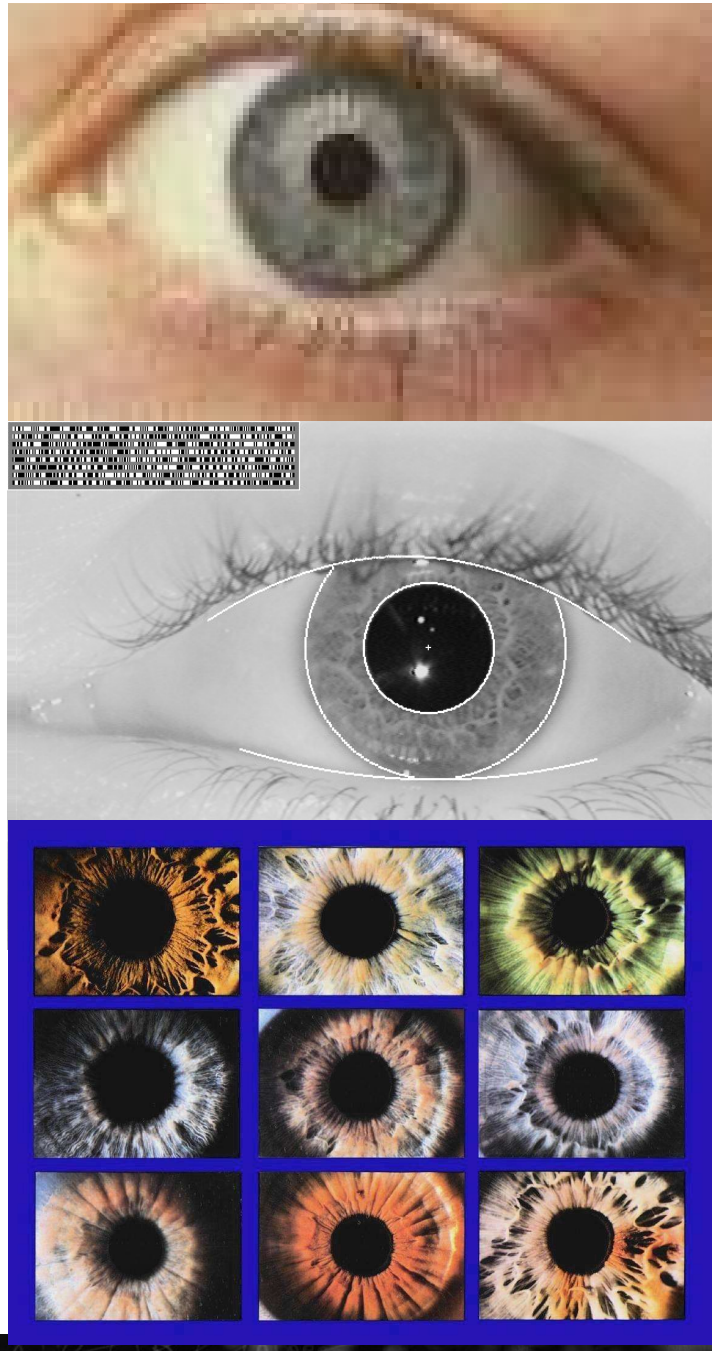
Can be captured from a distance

Monochrome camera with visible and near infra red light

Unique, two eyes and distinguish twins

Liveness detection

Experienced as intrusive



Iris – or actually the rich texture from images of iris

The mesh consists of characteristics such as striations, rings, furrows, etc, giving the iris a unique pattern

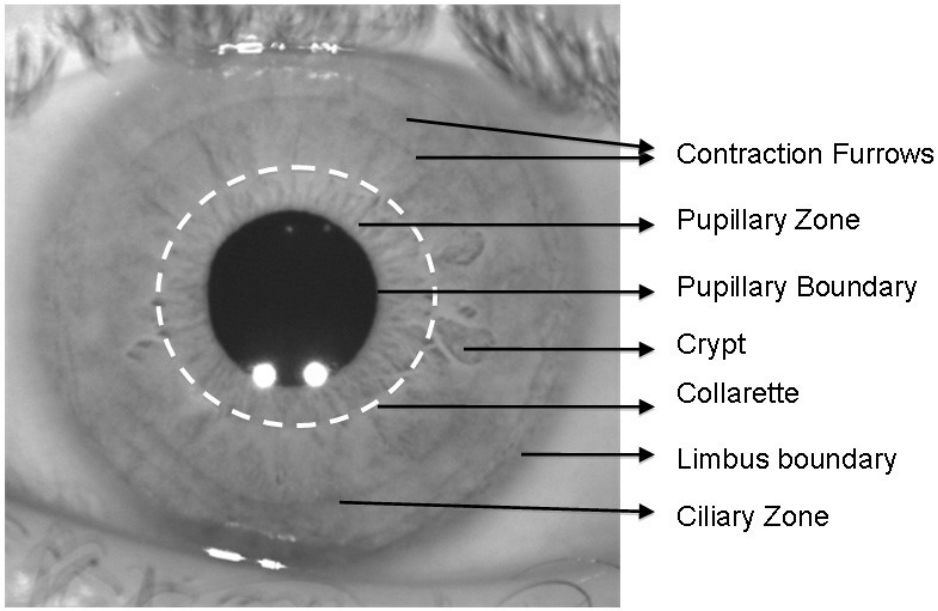
Don’t change with age

Can be captured from up to one meter

Ocular region of the human face

Iris

- Increased use since 1993
- Algorithm patent 1994
  - by Dr. John Daugman
  - used in all iris scanning systems today
- Works even with glasses and contact lenses
- Liveness is checked by using light to change the size of the pupil



NIR image

Iris

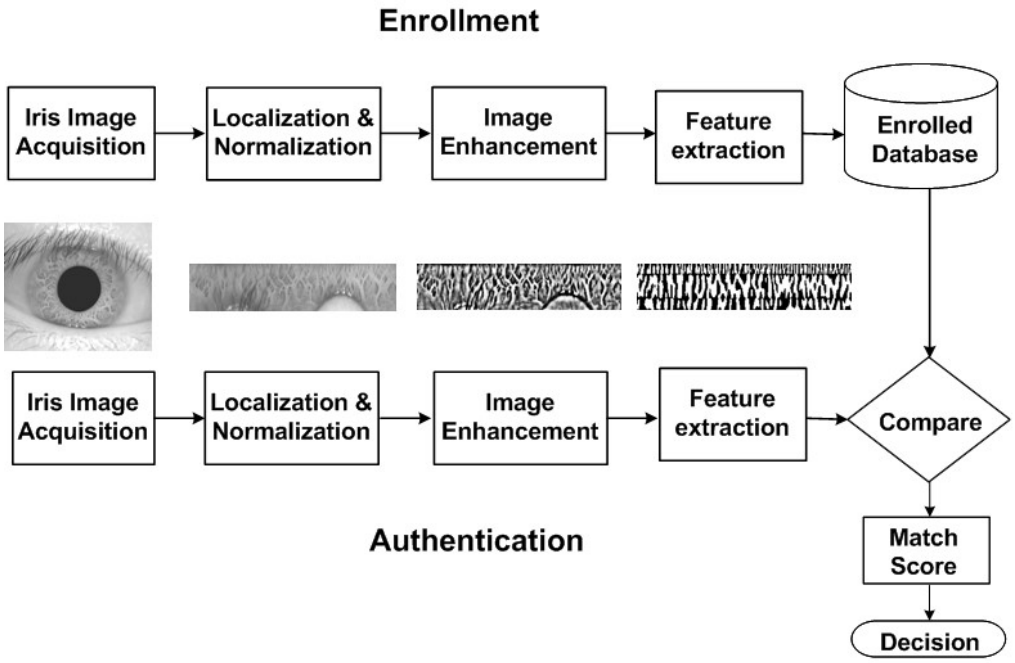
Very accurate, giving low FAR

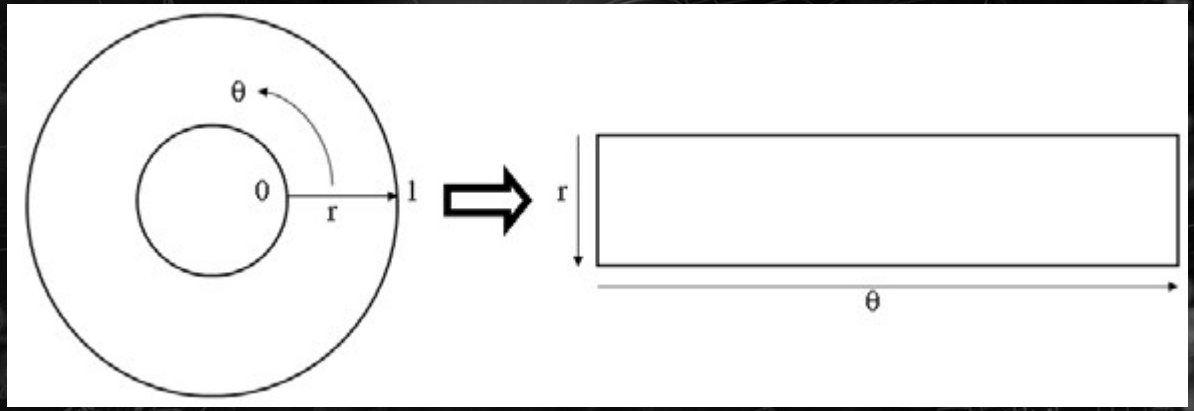
Used for identification and verification

High costs

May suffer from poor lighting and reflections

No human iris experts





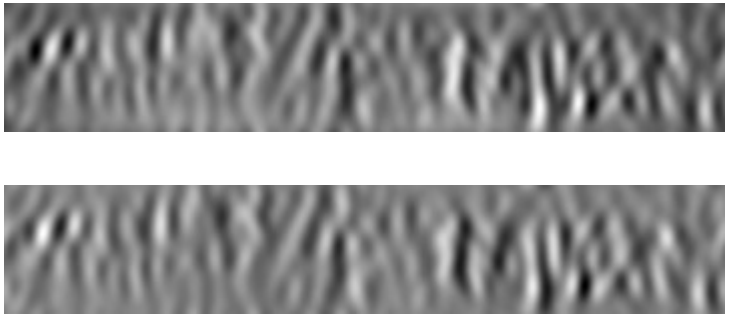
$I(x(r,\theta),y(r,\theta)) \rightarrow I(r,\theta)$

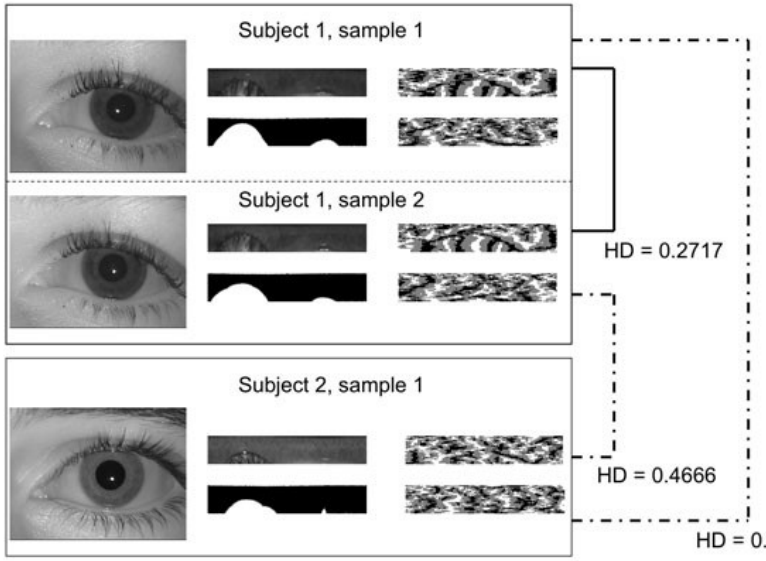
with

$x(r,\theta) = (1-r)xp(\theta)+rxI(\theta)$

and

$y(r,\theta) = (1-r)yp(\theta)+ryI(\theta)$





Iris - attacks

Contact lens with image

Porcelain eye

Photo of an eye

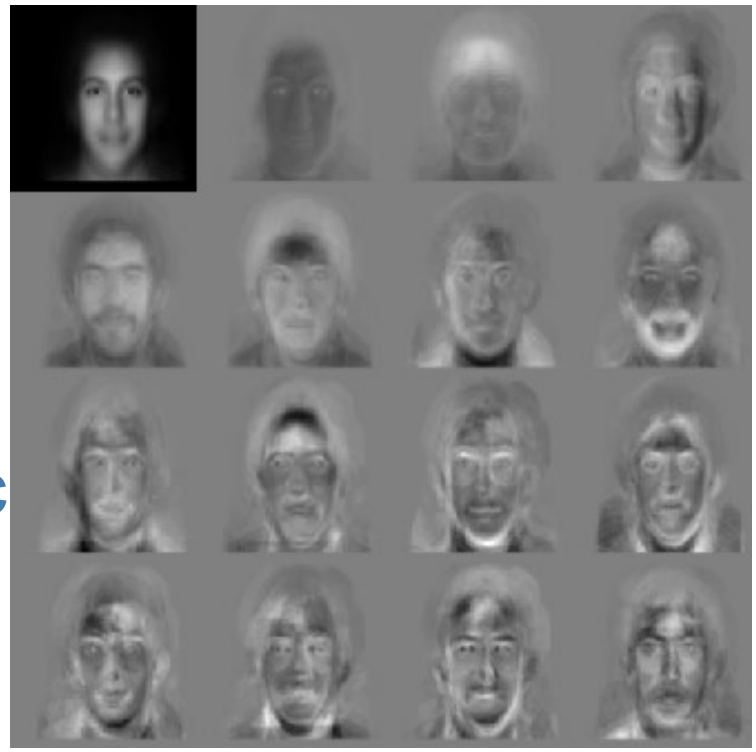
Example: Face

A face image can be acquired using a normal, off-the-shelf camera

Easy to accept by the public

Cost is rather low

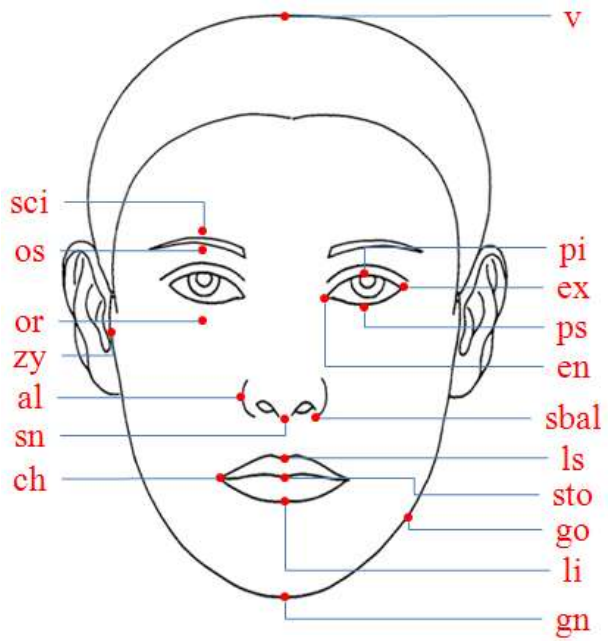
Huge problems with permanence and accuracy

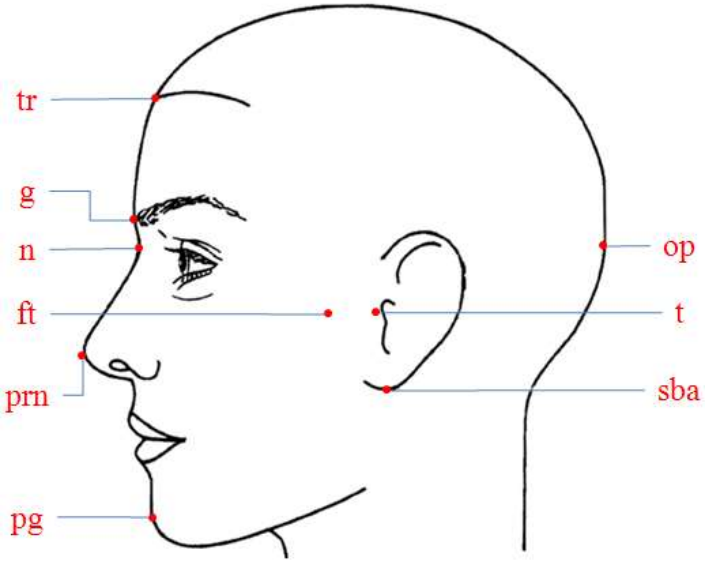


Facial features

Gross facial characteristics, eg general geometry of the face and global skin

Localized face information eg structure of face components or their relations



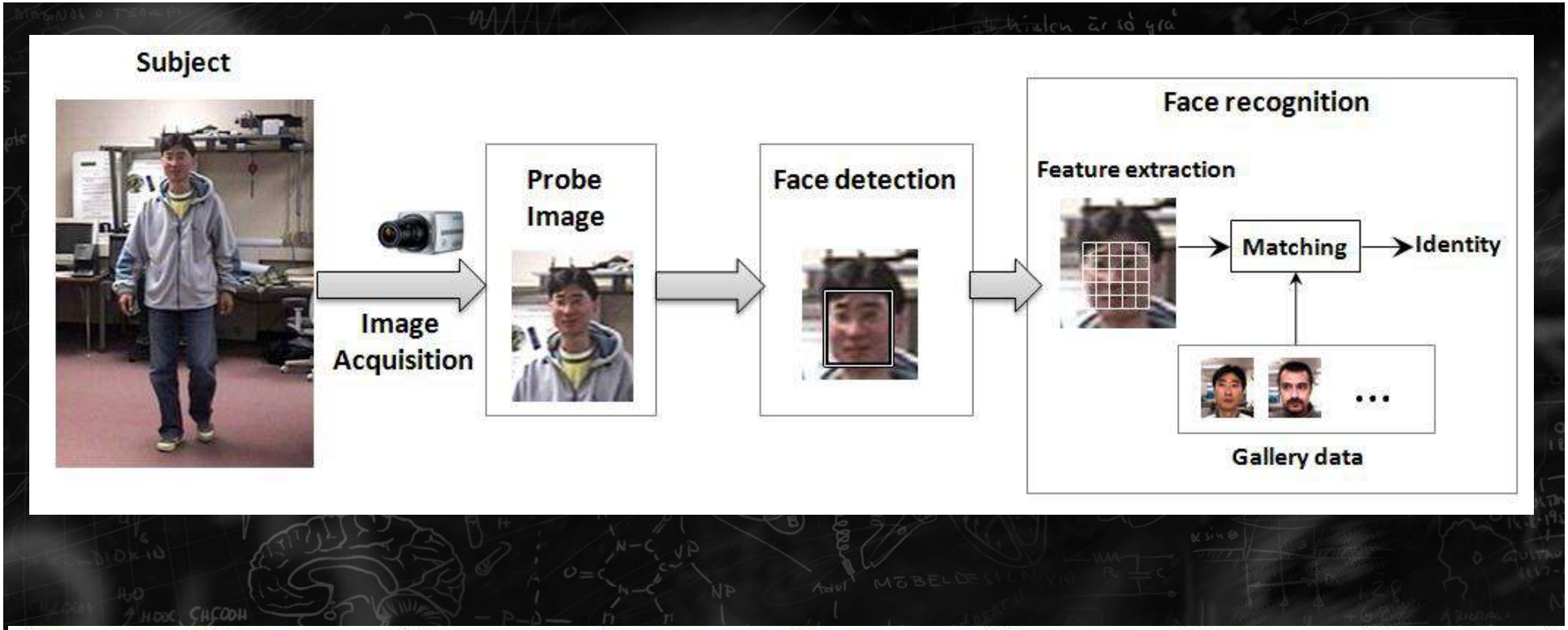


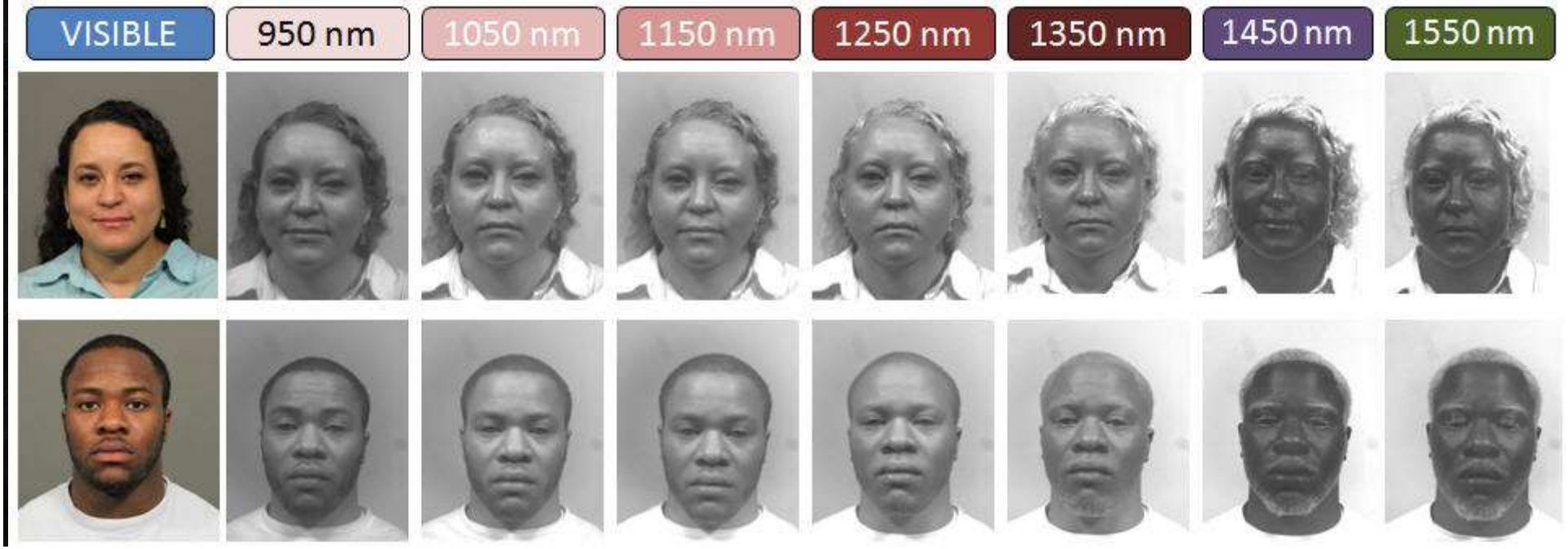
Face recognition algorithms

Global or feature-based approach

Feature-based
- standard points only
- not (too) sensitive to variation in position

Global
- process the entire face
- more accurate
- sensitive to variation in position and scale


Subject
Image Acquisition
Probe Image
Face detection
Face recognition
Feature extraction
Matching
Identity
...
Gallery data


VISIBLE
950 nm
1050 nm
1150 nm
1250 nm
1350 nm
1450 nm
1550 nm

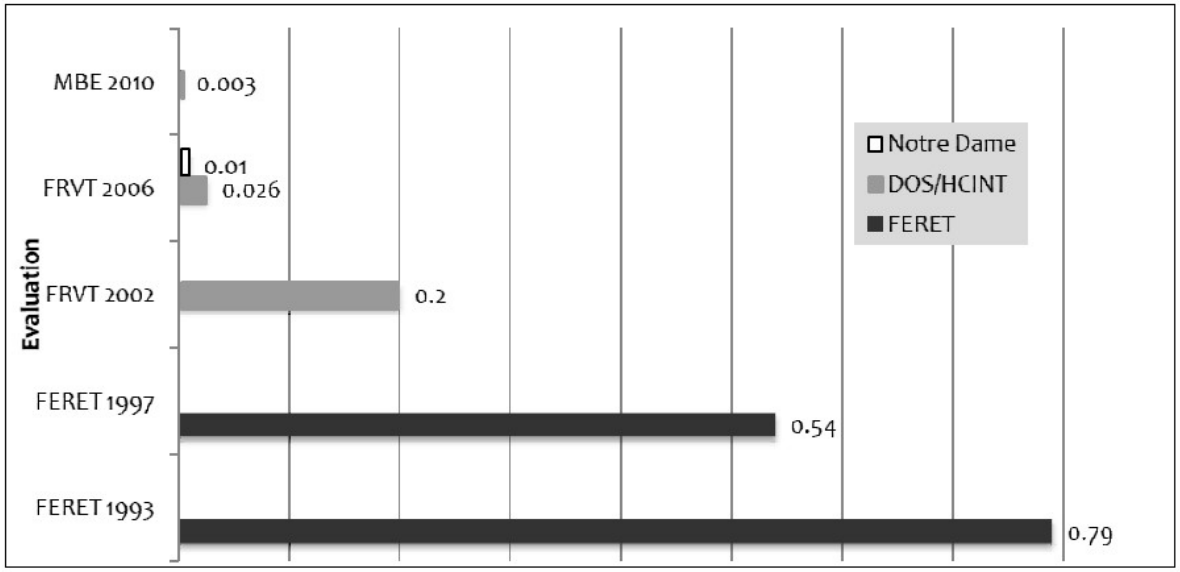
Face - attacks

Photo

Using low uniqueness

Masks or plastic surgery

False Reject Rate at a fixed False Accept Rate in the verification mode



Example: Hand geometry

Usually two views are taken, a top view and a side view.

The system is often bulky.

The hand geometry can change due to age and health conditions.

Example: Voice

Speaker recognition uses a microphone to record the voice.

Text dependent or text independent

Your voice can vary with age, illness and emotions.

Interesting with the increasing use of mobile phones.

Voice

Text dependent or text independent

Dependent
- The text is decided by the system
- Fixed or random
- Cooperation needed

Independent
- Any text can be used
- No cooperation needed
- Much harder

Voice - attacks

Recordings

Computer generated voice

”Tokens”?

”Token” is normally used for any authentication device with processing capacity

Smart cards are a variant

RFID devices (Radio-frequency identification) (ePassports have them!)

Phones with SIM-cards are another example

(Ross Anderson, Security Engineering chapter 16)

Attacking what?

Authentication tokens contain personal keys, which should not be easy to reveal

Loss can be crucial to owner, if the attacker is another person, but usually further use can be blocked

Even more important are **system keys**!!!

System keys may protect data proving payment for services

System keys may enable fabrication of false tokens

Hardware attacks

Studying the equipment

- electro-magnetic signals
- power variations
- time to perform operations

Manipulating the equipment

- probing
- varying power
- inducing errors and stopping operations

Emission, examples

Electromagnetic emissions occur whenever you use an electronic device

Power consumption in the equipment can be measured

Sounds from keyboards can be recorded and analysed

Eavesdropping on tokens

Emissions from processing is usually too weak to intercept without going beyond the cover layer. See probing.

Power for smart cards can easily be eavesdropped at the reader

Power consumption can reveal what processing that goes on, including branches taken after testing internal data

Timing attacks

Speeding up calculations often includes dropping unnecessary steps

Typical example is not doing all the steps when a key bit is zero

Analysis of time to encrypt can directly reveal number of zero bits in key

Combined with power analysis, every key bit can be found

Defence against timing attacks

Do not optimise calculation times

- Multiply with zero and add to total sum
- Branch on values, but always do the same number of steps in both branches

If necessary (no division with zero etc.), insert dummy calculations

Defence against power analysis

Remove timing attacks first

Insert random steps

Defence against eavesdropping

Use sufficient shielding around processors

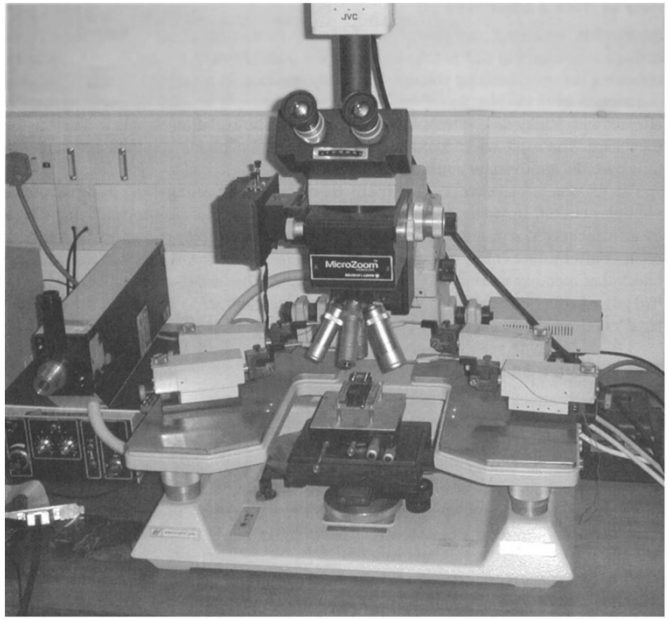
Avoid sending sensitive data, like keys, on internal buses

Probing

Direct contact with the electronics makes direct reading possible

See the literature (Anderson) for details

Also consider remanence! (It can make defences like power removal and erasures futile.)



Defence against probing

Use sufficient shielding around processors

Hardened and shatter-prone epoxy with meshes etc. makes removal of coatings much more difficult and expensive

Avoid sending sensitive data, like keys, on internal buses

Consider internal encryption

Remove power and erase sensitive data, when an attack is detected

Power manipulation

Preventing check data from being written may disable protective checks

Introduction of errors in the processing flow may alter the actual instruction sequence in ways that reveal sensitive data

- Checks can be skipped
- Limits for what can be output may be cancelled

Defence against power manipulation

When writing check data, always check that it is indeed written before proceeding with the calculations

Hide which step the processor executes in the processing flow (see power analysis)

Inducing errors

Carefully designed erroneous inputs can trigger unwanted events

- Similar to using security holes and badly designed protocols in general

Errors can be injected in stored data via particle beams, light on partly revealed surfaces etc.

- manipulate instruction flow
- change control limits
- alter key bits in ways that make analysis possible

Defence against induced errors

Use error detection for stored values, and check before use

Check outputs for consistency, if possible

Check inputs and block everything except meaningful, correctly designed sets