Introduction to Biometrics

Arun Ross

Professor

Michigan State University

rossarun@cse.msu.edu

http://www.cse.msu.edu/~rossarun

Pre-requisites | Lectures

Prerequisites:

CSE 331 and STT 351: A basic knowledge of statistics, linear algebra, and programming is expected. A background in image processing will be helpful, but not required. Please contact the instructor if you have any questions.

Lecture Details:

Time: Monday and Wednesday, 12:40pm – 2:00pm

1. This is primarily an in-person course (in EB 1145)
3. D2L Link for the course: https://d2l.msu.edu/d2l/home/1579304

Instructor and Office Hours:

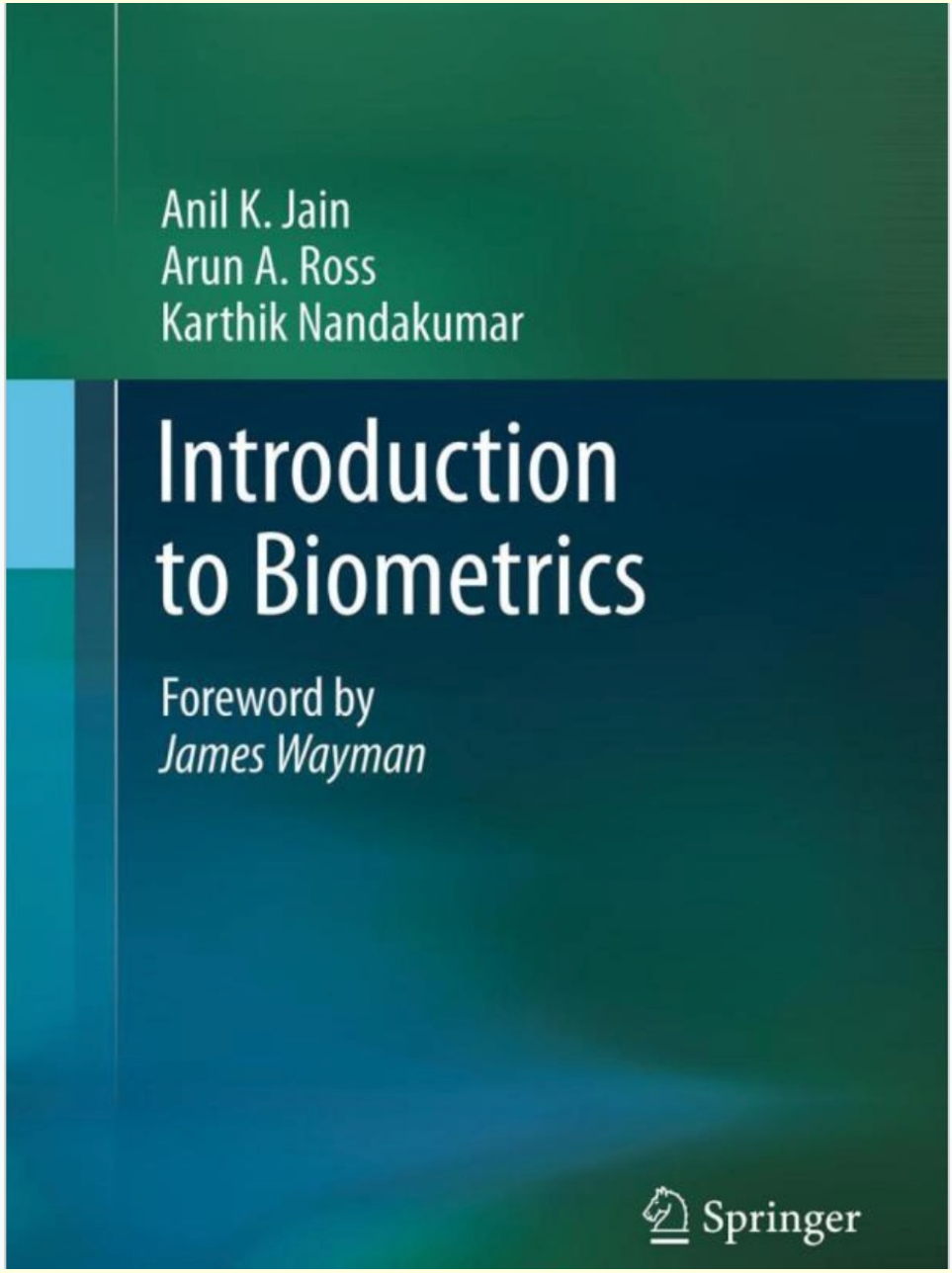
Dr. Arun Ross
(rossarun@cse.msu.edu)

Office Hours:

- Tuesdays 10:00am – 11:00am or by appointment only
- EB 3142

Textbook:

Introduction to Biometrics, Jain et al, Springer 2011.
ISBN: 978-0-387-77325-4.



Description

- Biometrics refers to the science of recognizing humans by utilizing the physical (e.g., fingerprint, face, iris) or behavioral (e.g., gait, signature) traits of an individual.
- This course will discuss several of these traits and the automated techniques used for feature extraction and matching. The focus will be on automated face, fingerprint and iris recognition.
- The error metrics used to evaluate the performance of a biometric system will be discussed.
- Topics related to multimodal biometrics, protecting biometric templates, and biometric data privacy will also be presented.
- The programming projects will be geared toward implementing basic feature extraction and matching algorithms.

Objectives

Objectives:

To equip students with a good knowledge of:

(a) the design and working of a generic biometric system;

(b) the features used to represent and match individual biometric traits;

(c) the performance metrics used to evaluate a biometric system;

(d) the socio-legal implications of biometrics.

The concepts will be explained from a pattern recognition and image processing perspective.

Outline of Topics

- **Introduction:**
  - What is Biometrics?
  - History of Biometrics
  - Applications
  - Biometrics as a Pattern Recognition System
  - Characteristics of a Biometric System

- **Performance Evaluation:**
  - Error Rates
  - ROC, DET, CMC Curves

- **Feature Extraction and Matching:**
  - Fingerprint Recognition
  - Face Recognition
  - Iris Recognition
  - Speaker Recognition

- **Other topics:**
  - Multibiometrics, Security, Privacy

Grading

The *tentative* weight associated with each grading component is as follows:

- Homework – 30%
- Project – 45%
- Lab Exercise – 5%
- Exam/Quiz – 20%

Assignments and tests will be based on (a) topics covered during lectures, (b) PowerPoint files used in lectures, (c) contents of the textbook, and (d) reading material assigned by the instructor. Students are expected to take down notes during the lecture.

Grading Policy

- Assignments must be turned in via D2L before the deadline.
- Late assignments will not be accepted.
- Make-up for exams will only be granted under exceptional circumstances.
- Instructor reserves the right to deny requests for make-up.

Spartan Pledge

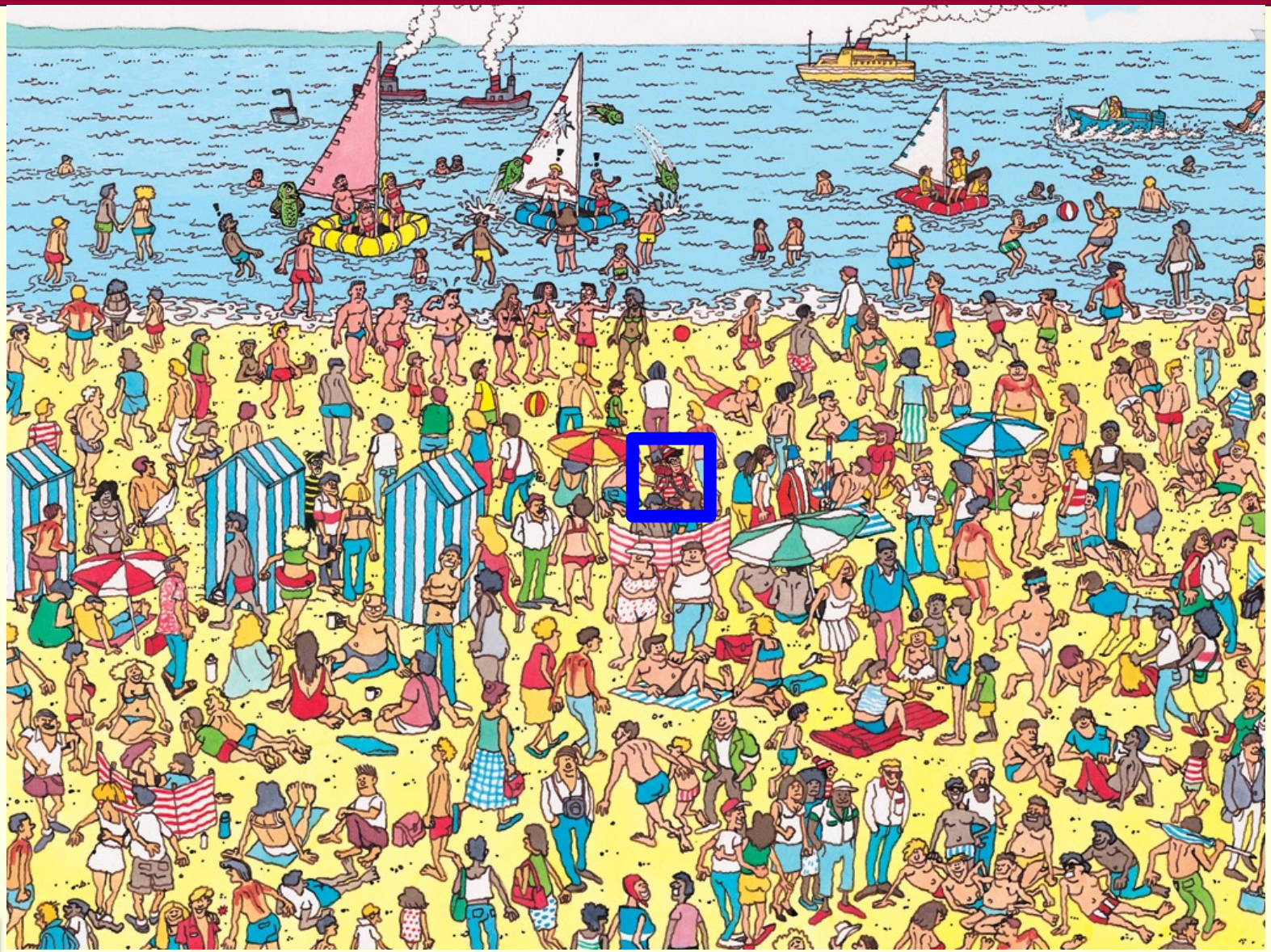
The Spartan Code of Honor Academic Pledge

As a Spartan, I will strive to uphold values of the highest ethical standard. I will practice honesty in my work, foster honesty in my peers, and take pride in knowing that honor in ownership is worth more than grades. I will carry these values beyond my time as a student at Michigan State University, continuing the endeavor to build personal integrity in all that I do.

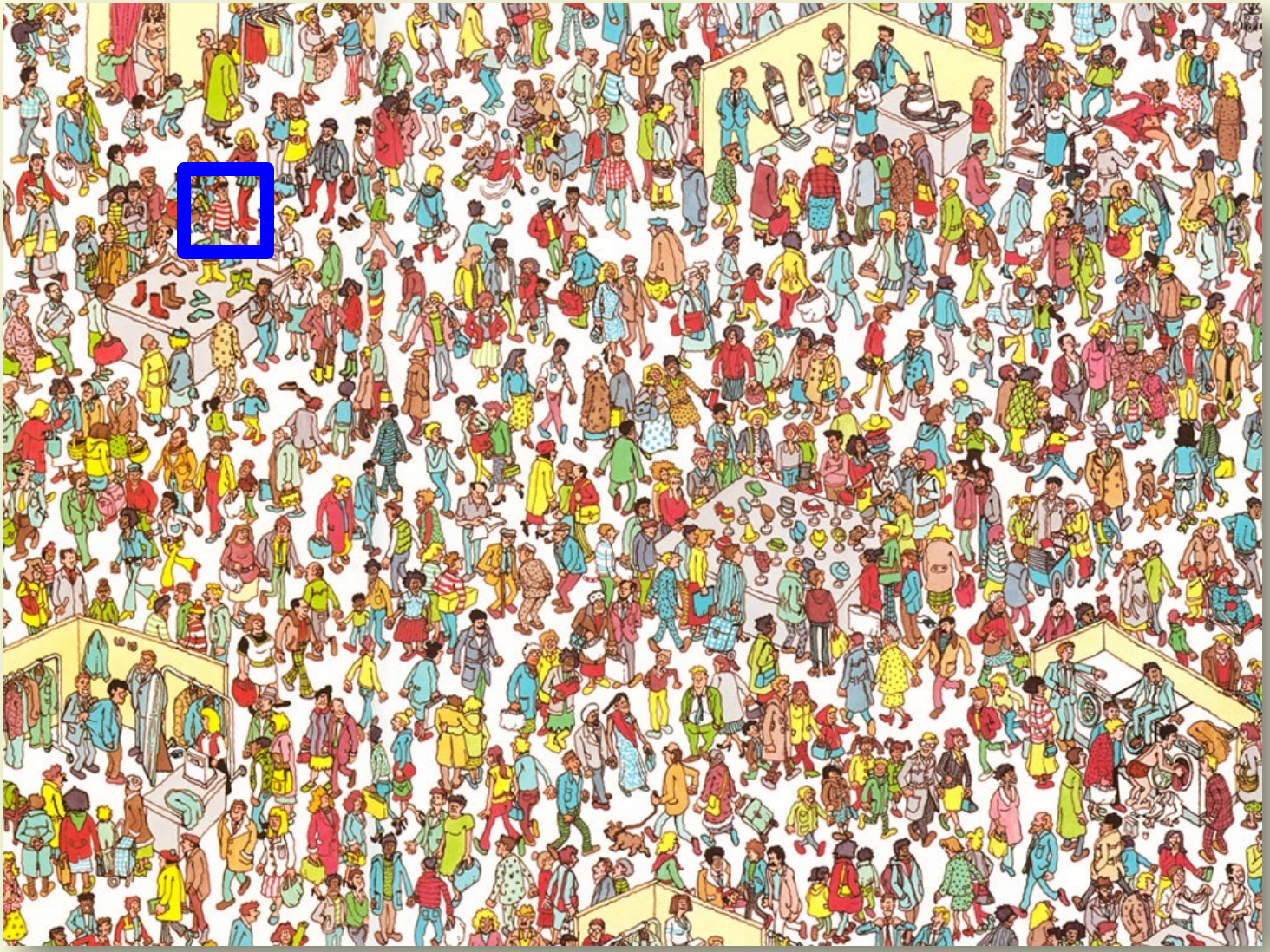
Where's Waldo?

- Series of children's books created by Martin Handford
- Illustration of a large number of people engaged in various activities
- Readers are challenged to find a character named Waldo

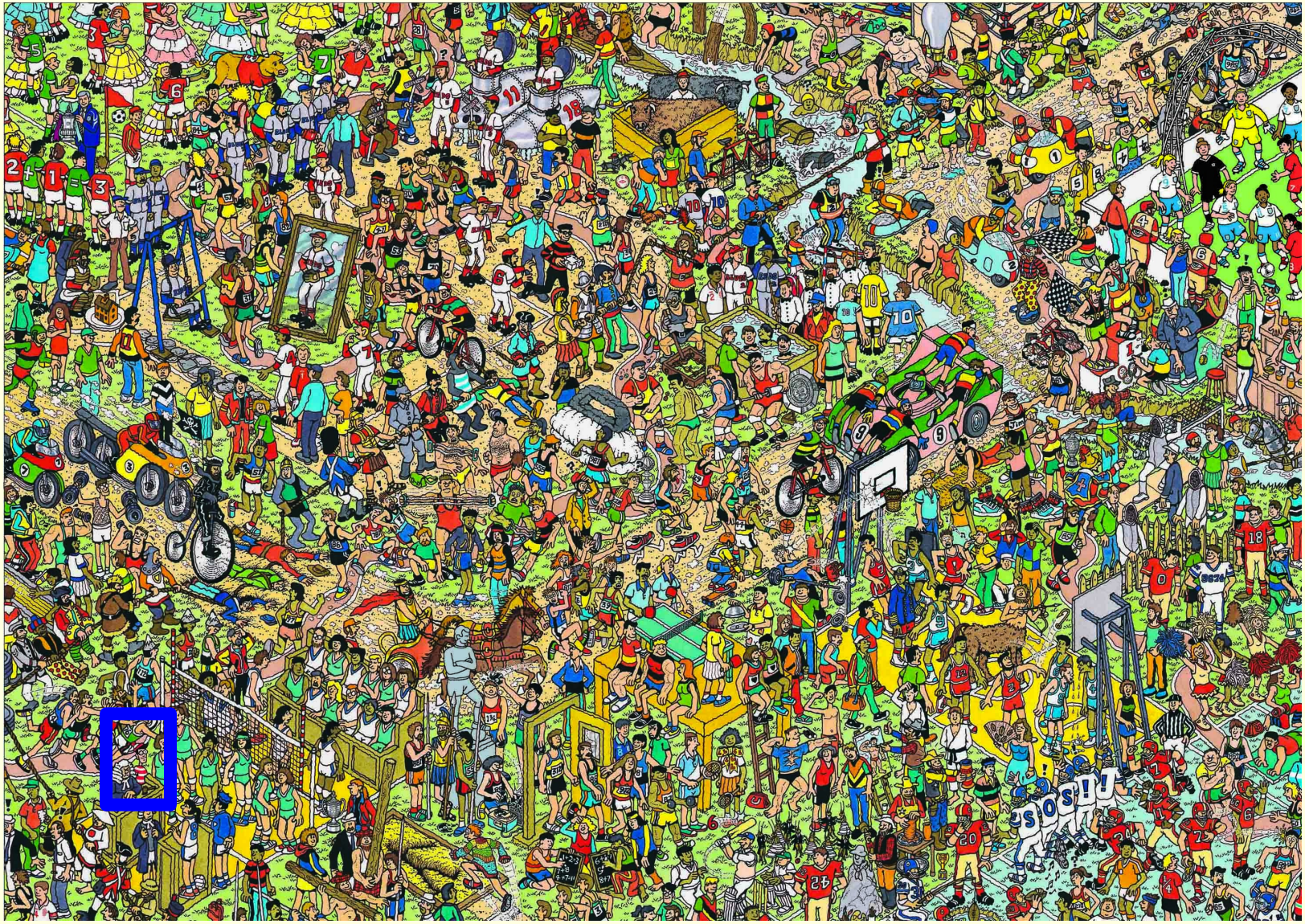
Where's Waldo?



Where's Waldo?



Where’s Waldo?



Other Characters Added

- **Odlaw**, **Wizard Whitebeard**, **Wilma**, **Wenda**, **Woof**

So Who Is Waldo?

- Waldo defined by facial appearance, clothing, accessories
- Waldo can be occluded
- Challenging environment:
  - different types of environment
  - large number of people
  - some people look similar
  - people engaged in diverse activities
  - variations in illumination and pose

Biometric Recognition

- Automated recognition of individuals based on their biological and behavioral characteristics
- Biological and behavioral characteristic of an individual from which distinguishing, repeatable biometric features can be extracted

H.T. F. Rhodes, Alphonse Bertillon: Father of Scientific Detection, Harrap, 1956

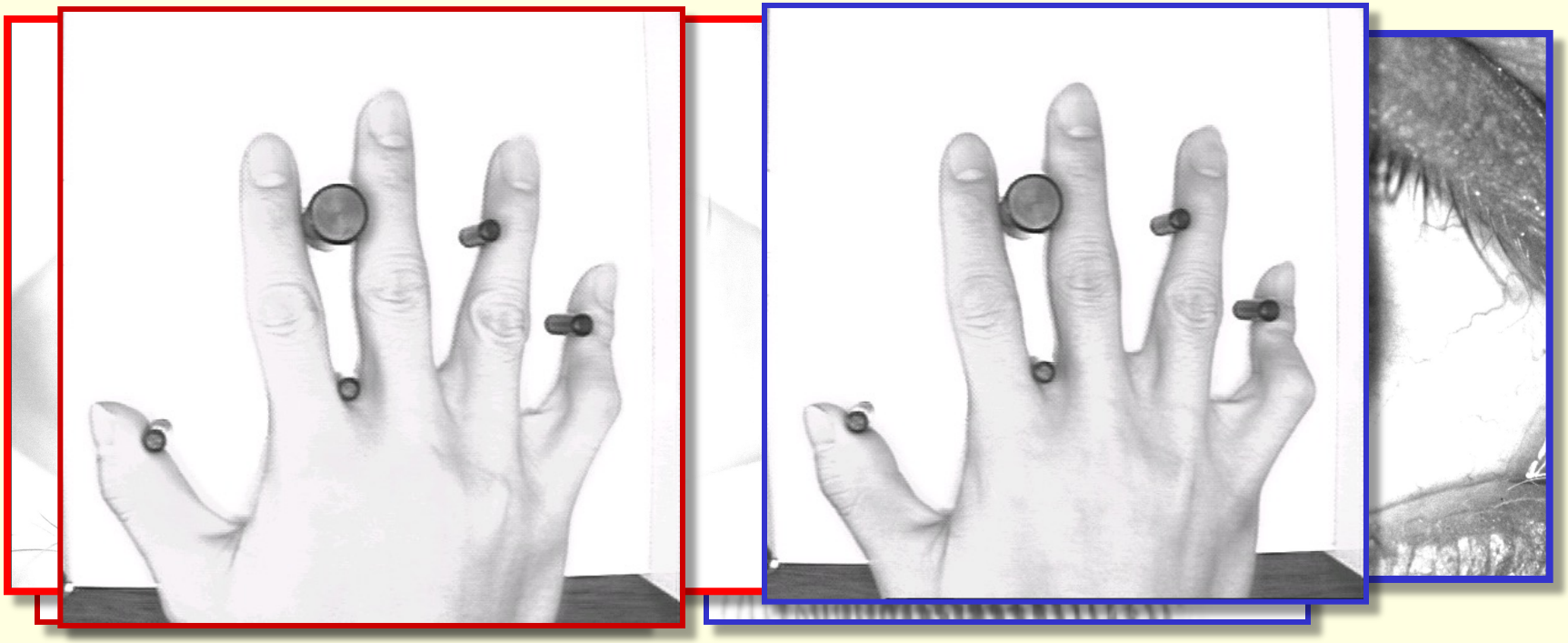
An Early Biometric System

Identify repeat offenders

H.T. F. Rhodes, Alphonse Bertillon: Father of Scientific Detection, Harrap, 1956

Biometric Matching

- When **observing** two biometric samples, estimate the ratio of the likelihood that they are of the same person to the likelihood that they are of different people



Biometric Recognition

- Verification

  Are you who you say you are? (1:1 match)

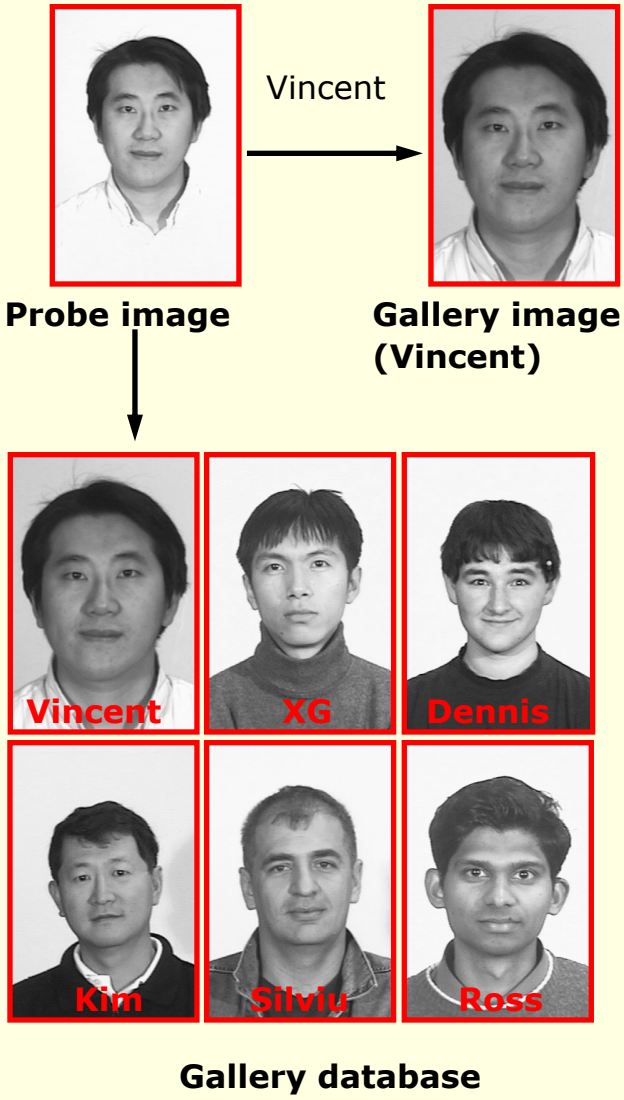
- Identification

  Who is this person? (1:N match)

- Watch List

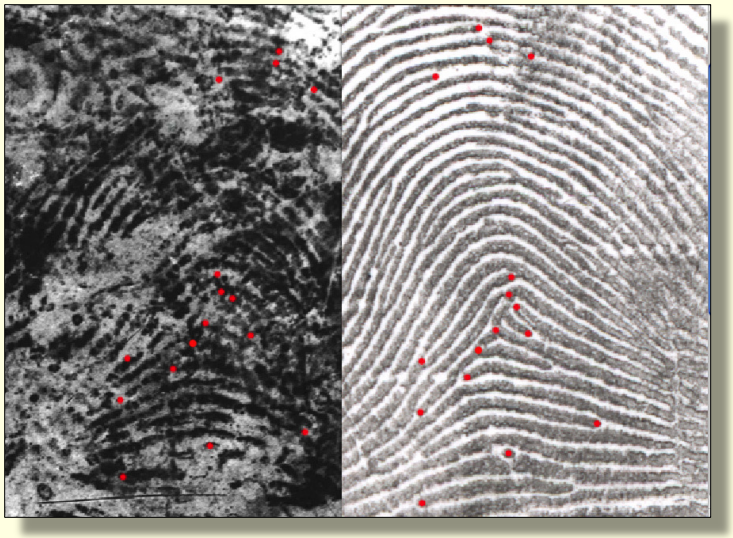
  Is this a wanted person?

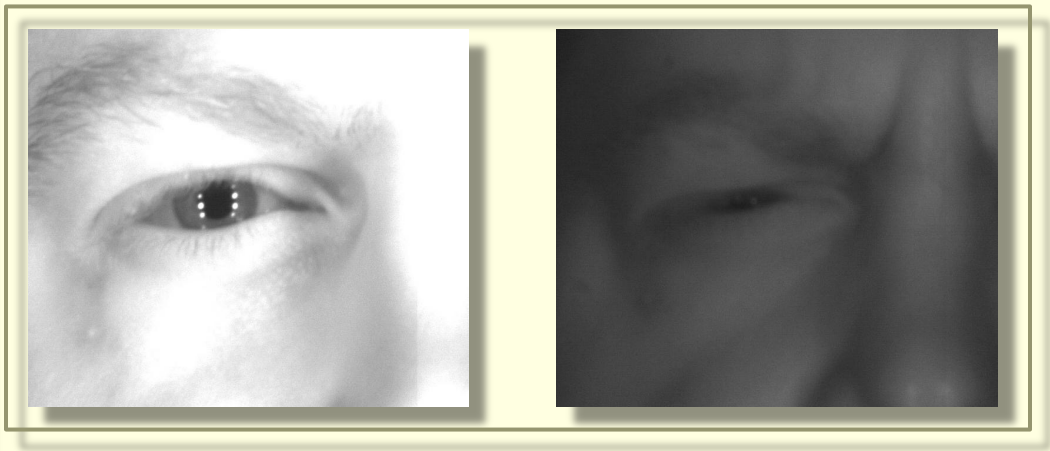
Biometrics can provide negative identification (i.e., I am not he) capability. It can search for multiple enrollments by the same individual



Real-world Comparison

- Compute the similarity or dissimilarity between two instances of biometric data corrupted by noise



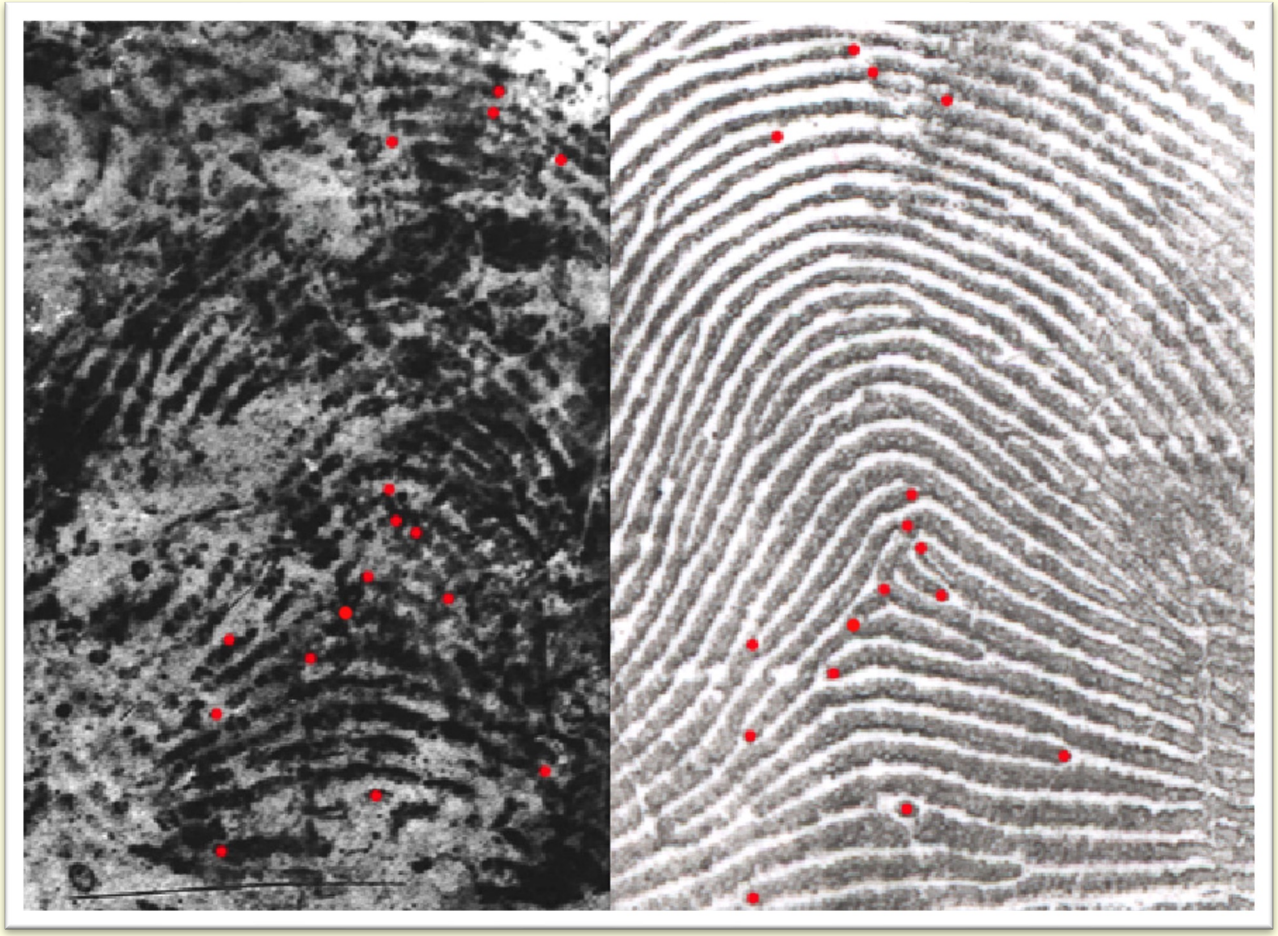


- Compute the ratio between the likelihood of the two images originating from the same individual to the likelihood that they originate from different individuals

Criminal Investigation

Are These The Same?

- Are these two prints from the same finger?



Authenti cation

Is He Allowed Access?

- Is he the owner of this smart phone?

De-duplication

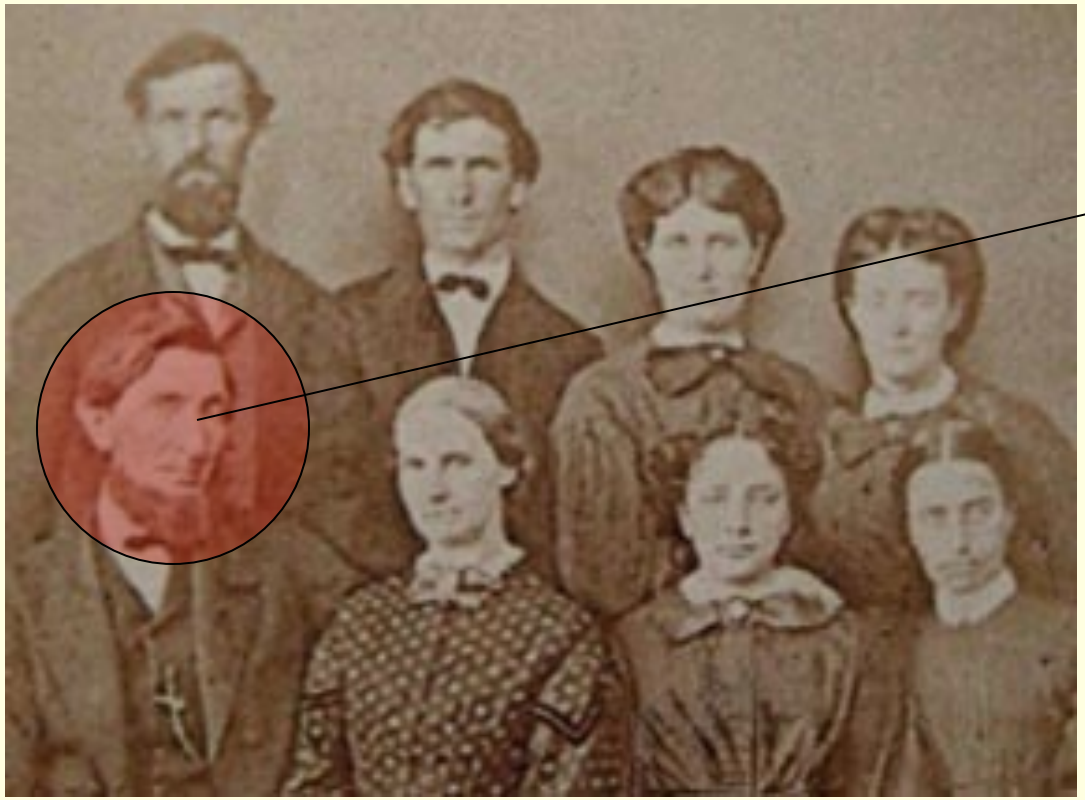
Is She In The Database?

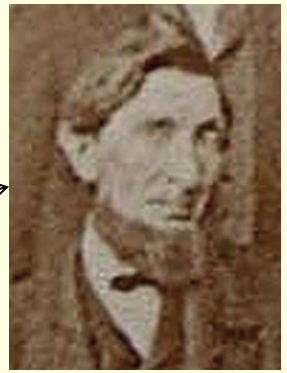
- Does this person already have a driver's license under a different name?

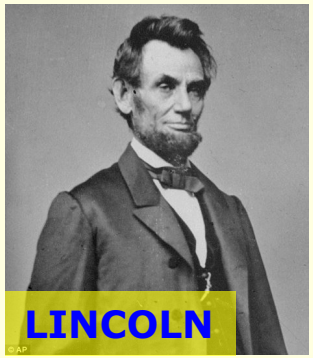
History

Is This Really Him?

- Is this really a photograph of Abraham Lincoln?





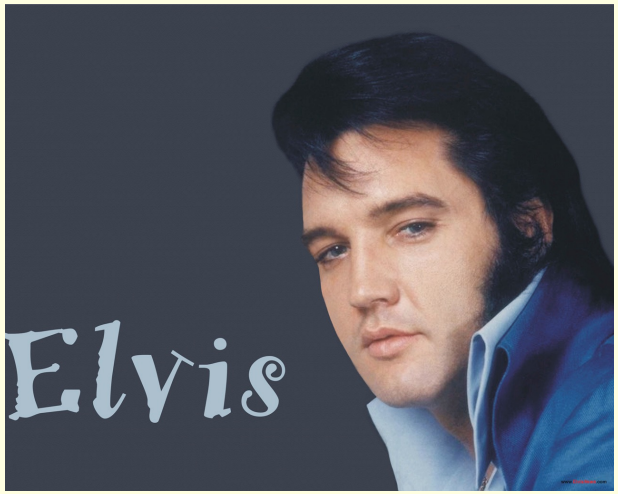


?

Authenti cation

Who is Singing?

- Is this really Elvis Presley's voice? (And if so, is he still alive?!)



https://www.youtube.com/watch?v=HGsssVWiu54

Retrieval

Where is She?

- Find all video frames in which Odette appears

Curiosity

Who is This? (Non-biometric method)

- Who is this person?

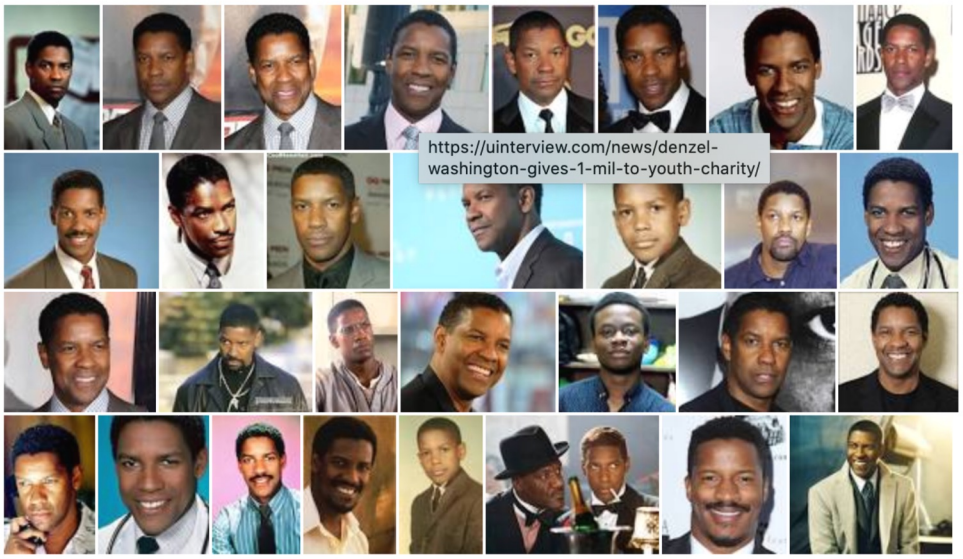
About 339 results (0.89 seconds)

Image size:
197 × 256

Find other sizes of this image:
All sizes - Small - Medium - Large

Possible related search: ***denzel washington young***

Visually similar images

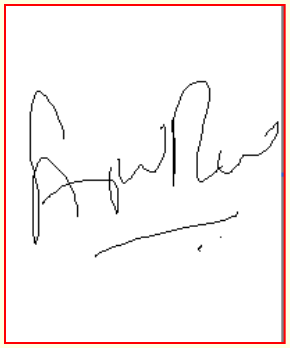


Report images

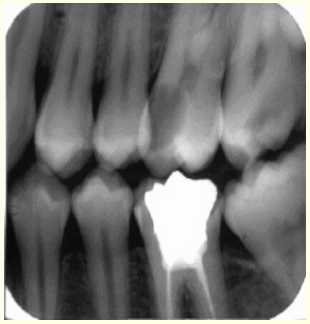
Biometric Traits

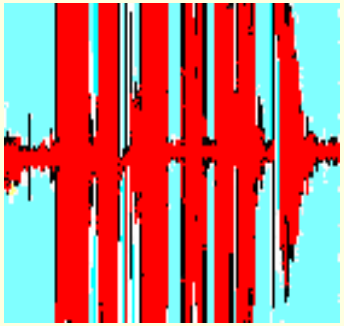
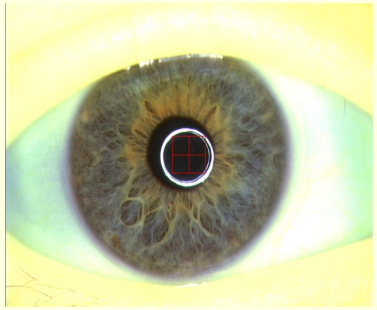






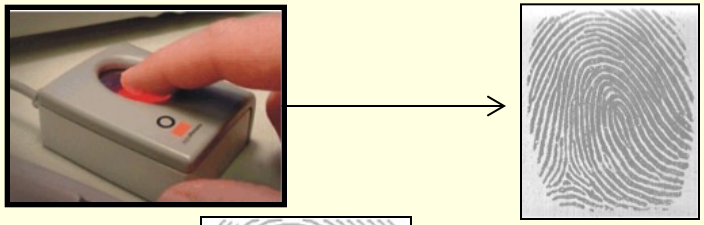


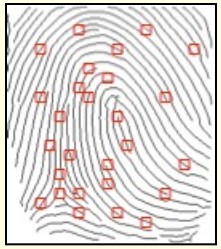



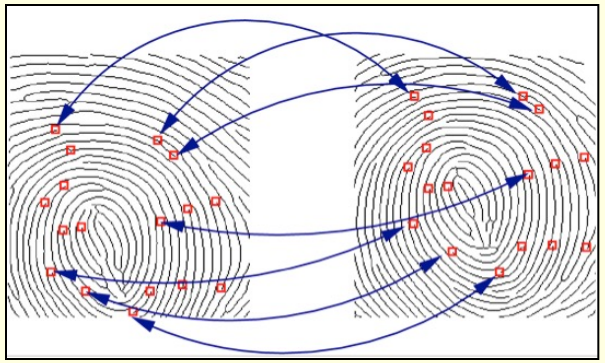



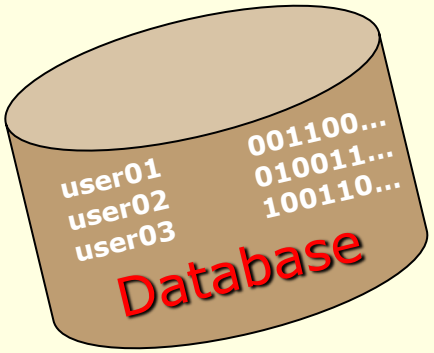

Components of a Biometric System

- **Sensor**: To acquire biometric data from an individual
- **Feature extractor**: To extract a set of discriminative features from the data
- **Comparator (Matcher)**: To compare two extracted feature sets
- **Database**: To store biometric templates of individuals



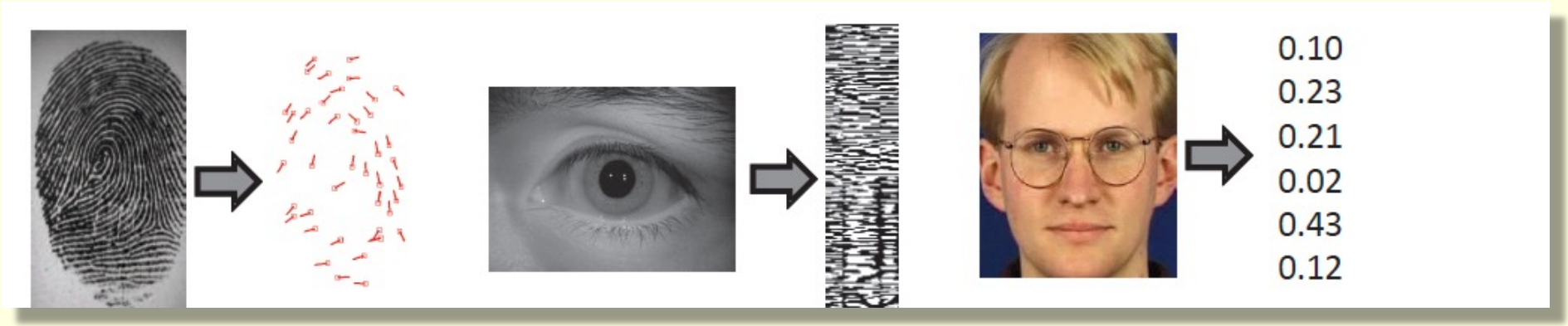






PROBE and GALLERY

What features are extracted?



- Minutiae points from fingerprint
- IrisCode from iris
- Deep learning features from face

These features reside in some algebraic space

Deep Neural Networks

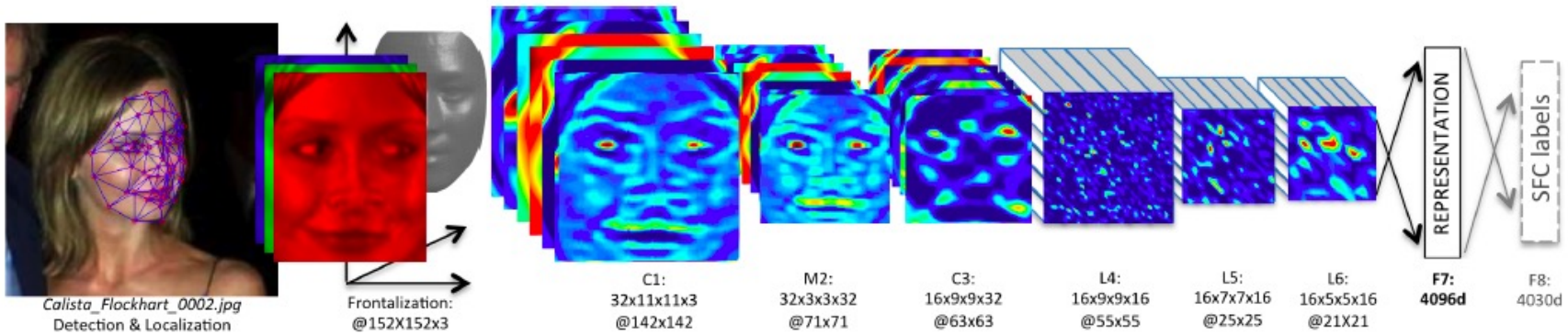


Figure 2. **Outline of the *DeepFace* architecture.** A front-end of a single convolution-pooling-convolution filtering on the rectified input, followed by three locally-connected layers and two fully-connected layers. Colors illustrate outputs for each layer. The net includes more than 120 million parameters, where more than 95% come from the local and fully connected layers.

Match Scores

- The output of a biometric matcher is the match score, (typically a single number), that quantifies the similarity or dissimilarity between two biometric samples
- The higher the score (similarity score), the more certain is the system that the two biometric samples come from the same person

Genuine and Impostor Scores

- Genuine score: Match score obtained when two biometric samples from the same source (e.g., same finger) are compared

- Impostor score: Match score obtained when two biometric samples originating from different sources (e.g., different fingers) are compared

- Therefore, a genuine **similarity** score should be greater than an impostor score in value

- A threshold (or classifier) is used to determine if a score is genuine or impostor

Genuine and Impostor Scores

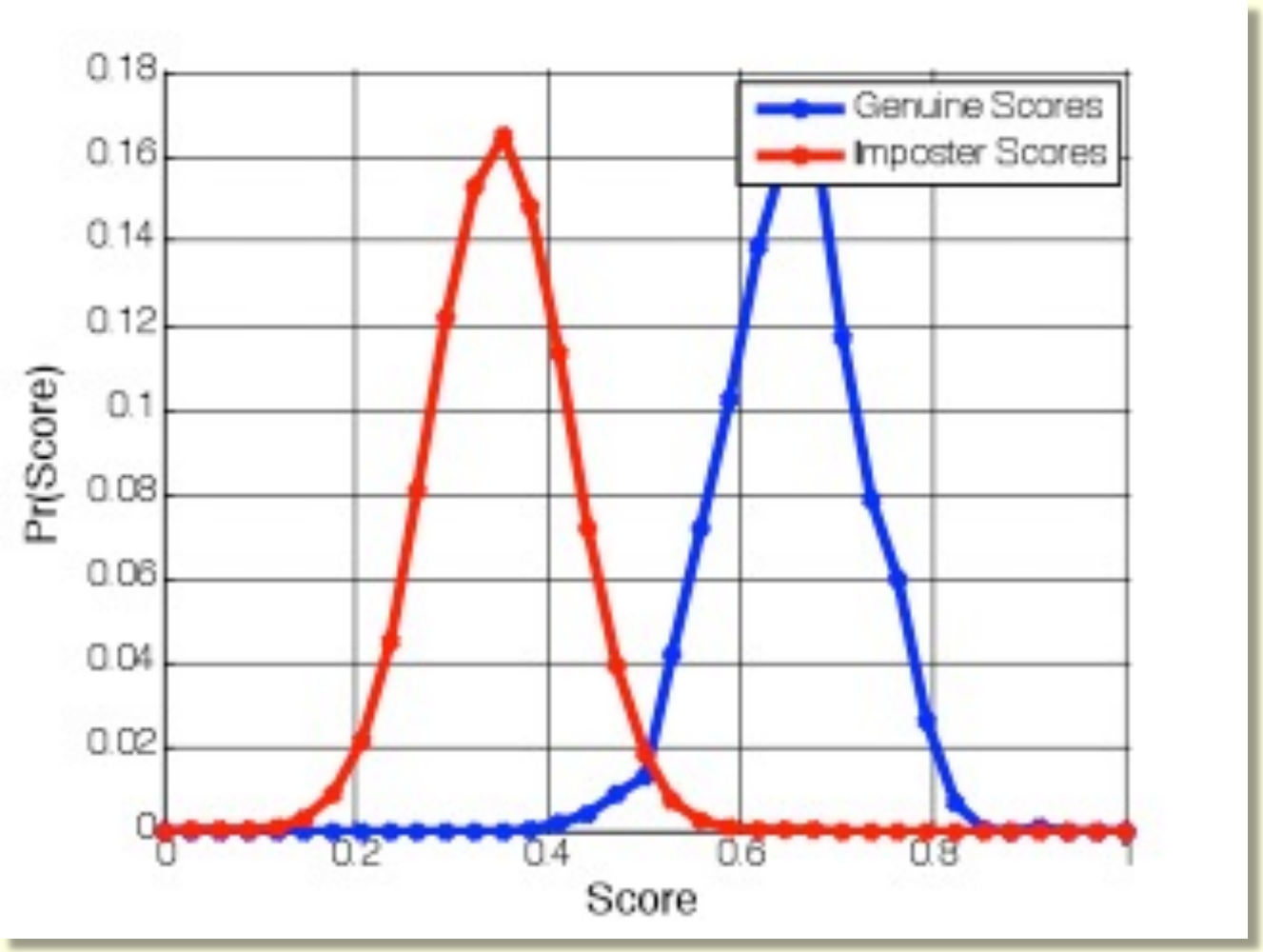
Genuine Score

0.91

0.15

Impostor Score

Score Distribution



Match Score Distributions
Density Estimation Schemes

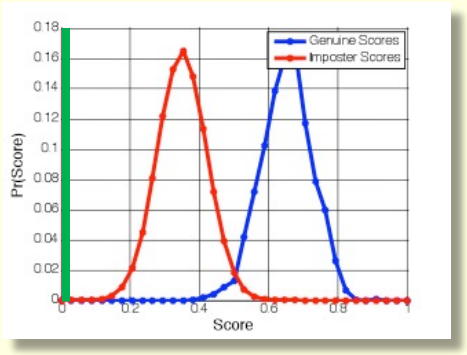
Error Rates (Verification)

- **False Match Rate (FMR):** **False Accept Rate (FAR)**
  - The proportion of impostor scores greater than the **threshold**
  - Probability that an impostor will be incorrectly matched
  - Low FMR required in high secure systems
- **False Non-Match Rate (FNMR):** **False Reject Rate (FRR)**
  - Proportion of genuine scores lower than the **threshold**
  - Probability that a genuine user will be incorrectly non-matched
  - Low FNMR required in systems focusing on user convenience

ROC and DET Curves

- Detection Error Tradeoff (DET) Curve
  - Plot FNMR against FMR at various thresholds
- Receiver Operating Characteristics (ROC) Curve
  - Plot (1 – FNMR) against FMR at various thresholds
  - 1 – FNMR = True Match Rate

Example of DET Curve



False Non-Match Rate (FNMR)

100%

High-security
Access
Applications

Equal Error Rate (FMR = FNMR)

Civilian
Applications

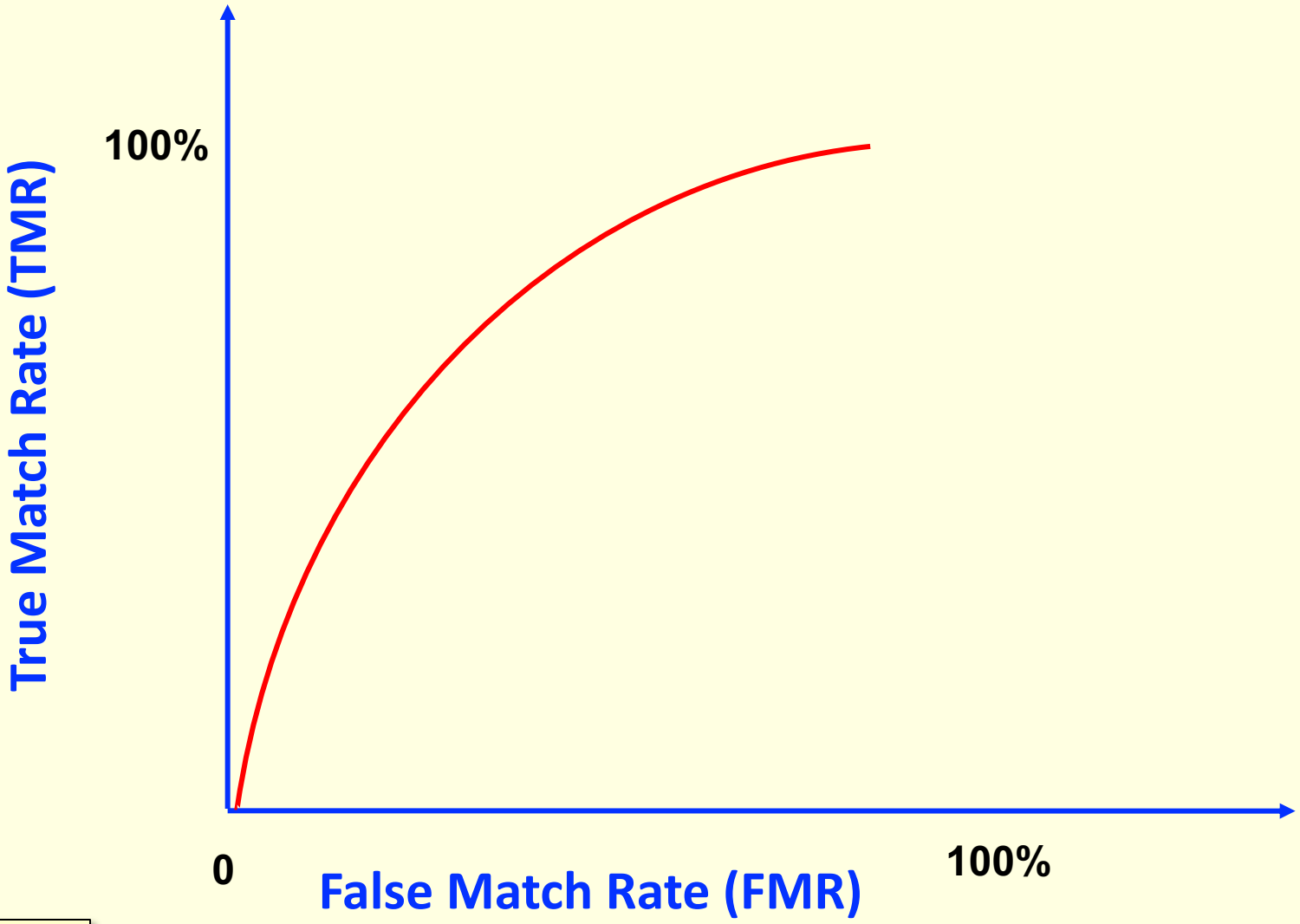
Forensic
Applications

0

100%

False Match Rate (FMR)

Example of ROC Curve



How many genuine and impostor scores?

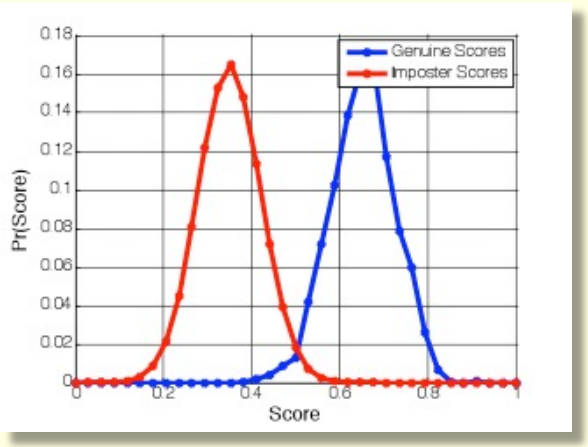
- Consider a biometric experiment
- "N" distinct identities
- "m" face samples per identity
- A biometric matcher is provided to you
  - How many genuine scores can be generated?
  - How many impostor scores can be generated?
  - Assume a symmetric matcher
    - S(a,b) = S(b,a), where "S" is the matcher and "a" and "b" are biometric samples

Computing FMR and FNMR From Actual Scores

- Generate the **genuine** and **impostor** scores
- Let us assume these are **similarity** scores
- At a specified threshold, η:
  - Compute the **number** of impostor scores $\geq \eta$
    - Divide this by the **total number** of impostor scores
    - This will give you the **FMR** at the threshold
  - Compute the **number** of genuine scores $< \eta$
    - Divide this by the **total number** of genuine scores
    - This will give you the **FNMR** at the threshold

Computing FMR and FNMR from Distributions

- Let $p(s|genuine)$ denote the **genuine** distribution
- Let $p(s|impostor)$ denote the **impostor** distribution
- **Assume these are similarity scores**
- At a specified threshold, η:
  - FMR = $\int_{\eta}^{\infty} p(s|impostor)ds$
  - FNMR = $\int_{-\infty}^{\eta} p(s|genuine)ds$



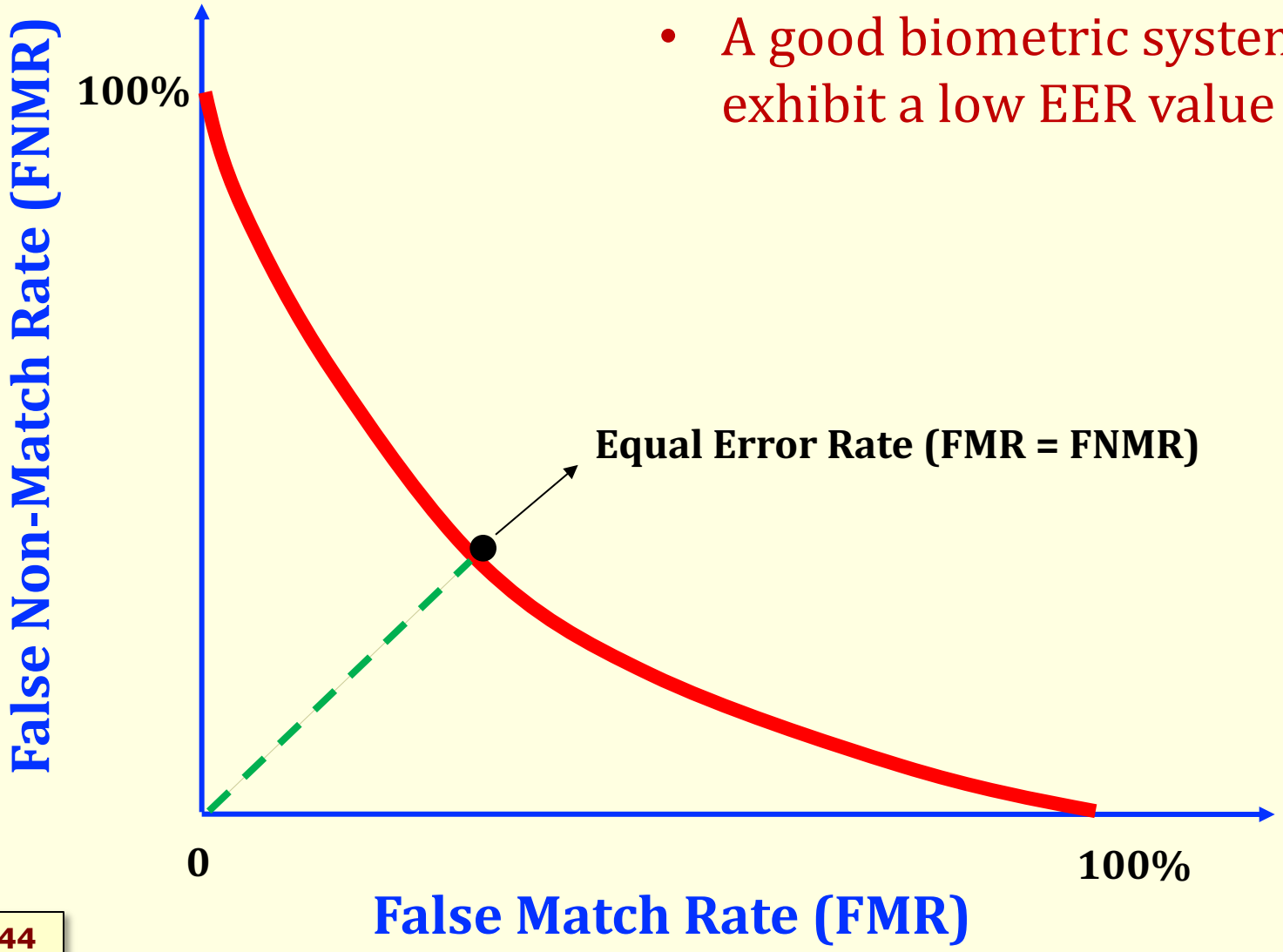
- Here, ∞ refers to maximum score value possible and -∞ refers to minimum score value possible

Computing the DET Curve

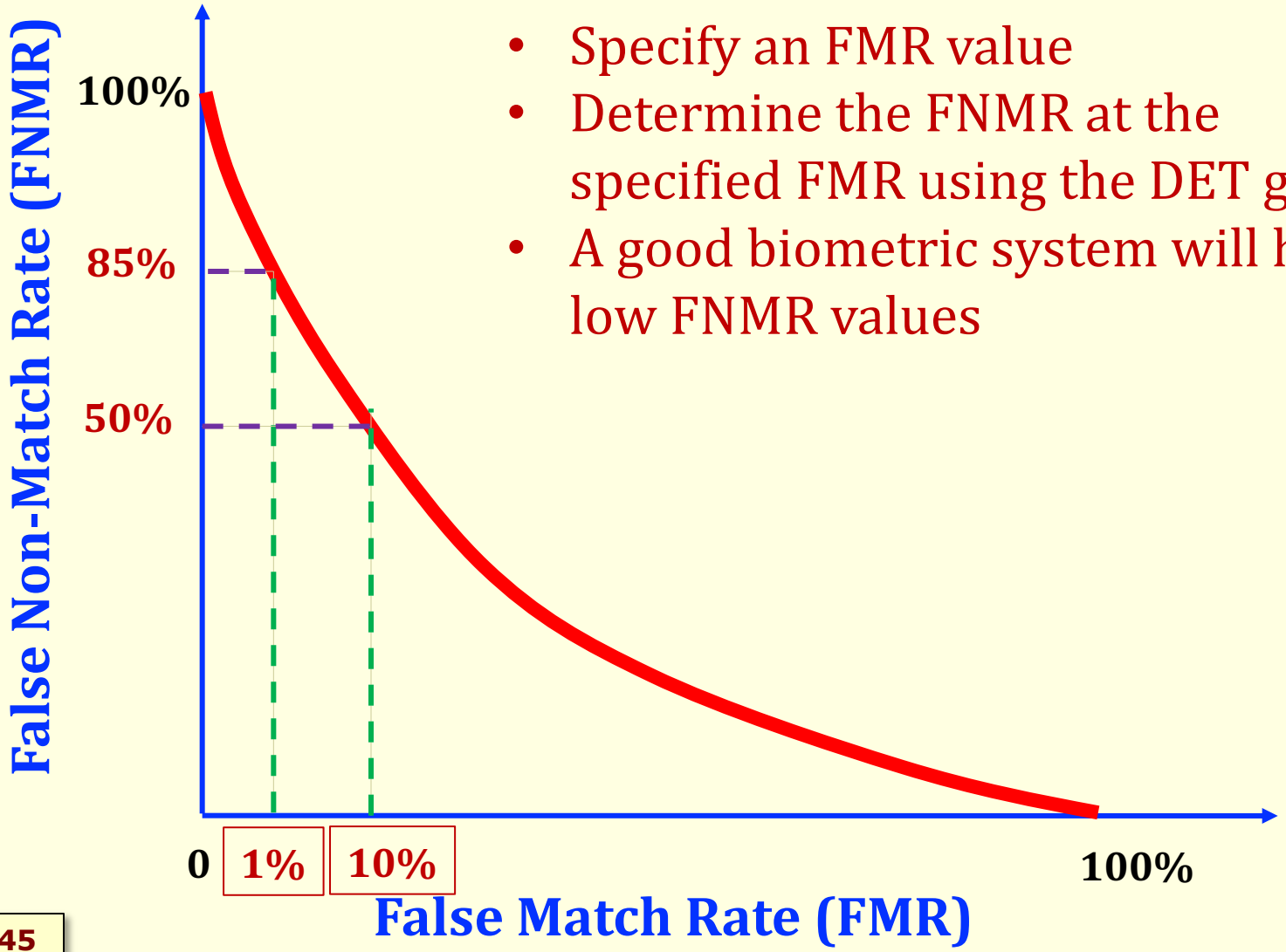
- Vary the threshold η from the minimum of all score values to the maximum of all score values in increments of Δ
- Ensure that you have at least 100 threshold values
- At each value of η compute the FMR and FNMR
- Plot these points, i.e., (FMR, FNMR) in a graph
- This is the DET Curve

Equal Error Rate: EER

- A good biometric system will exhibit a low EER value



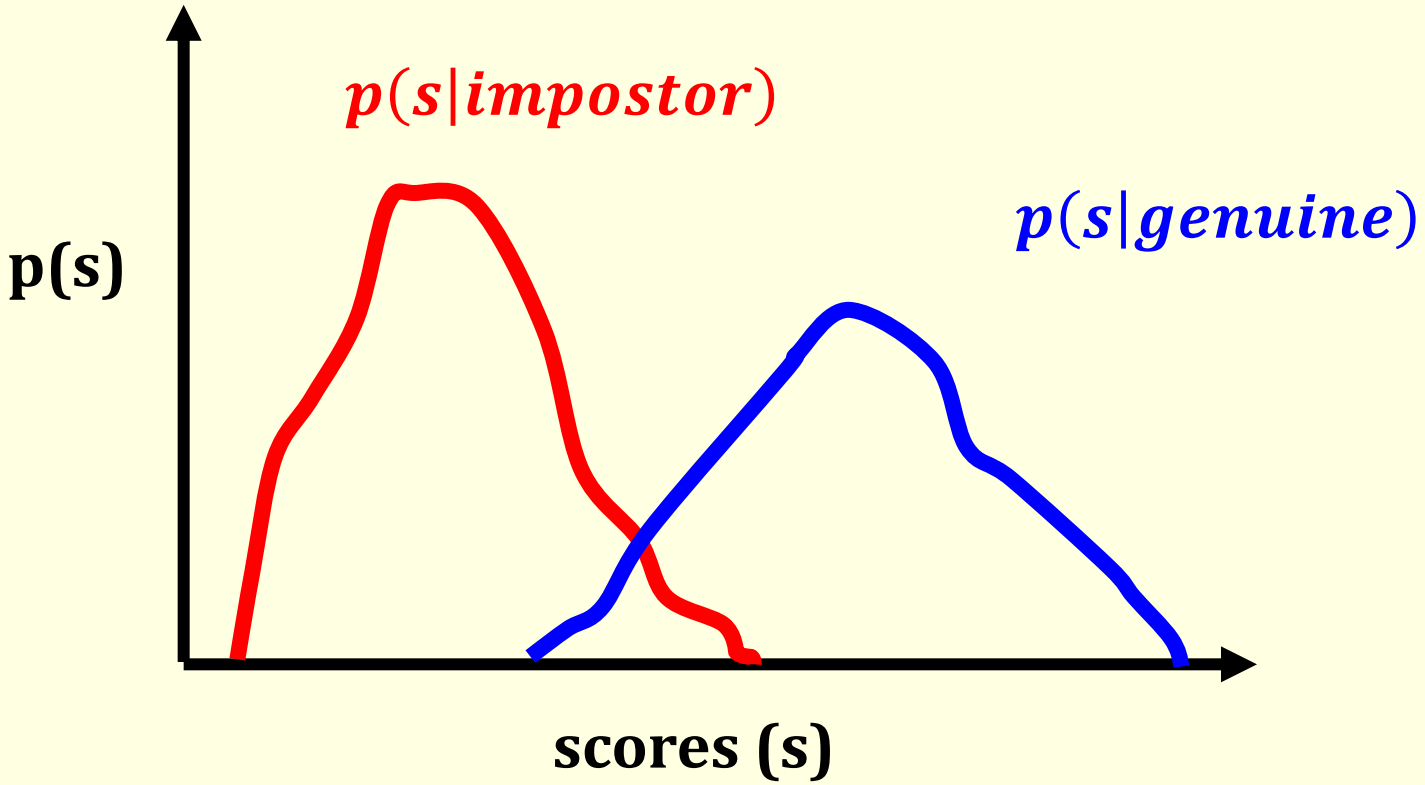
FNMR @ FMR



- Specify an FMR value
- Determine the FNMR at the specified FMR using the DET graph
- A good biometric system will have low FNMR values

Other Performance Metrics

- Let $p(s|genuine)$ denote the **genuine** distribution
- Let $p(s|impostor)$ denote the **impostor** distribution

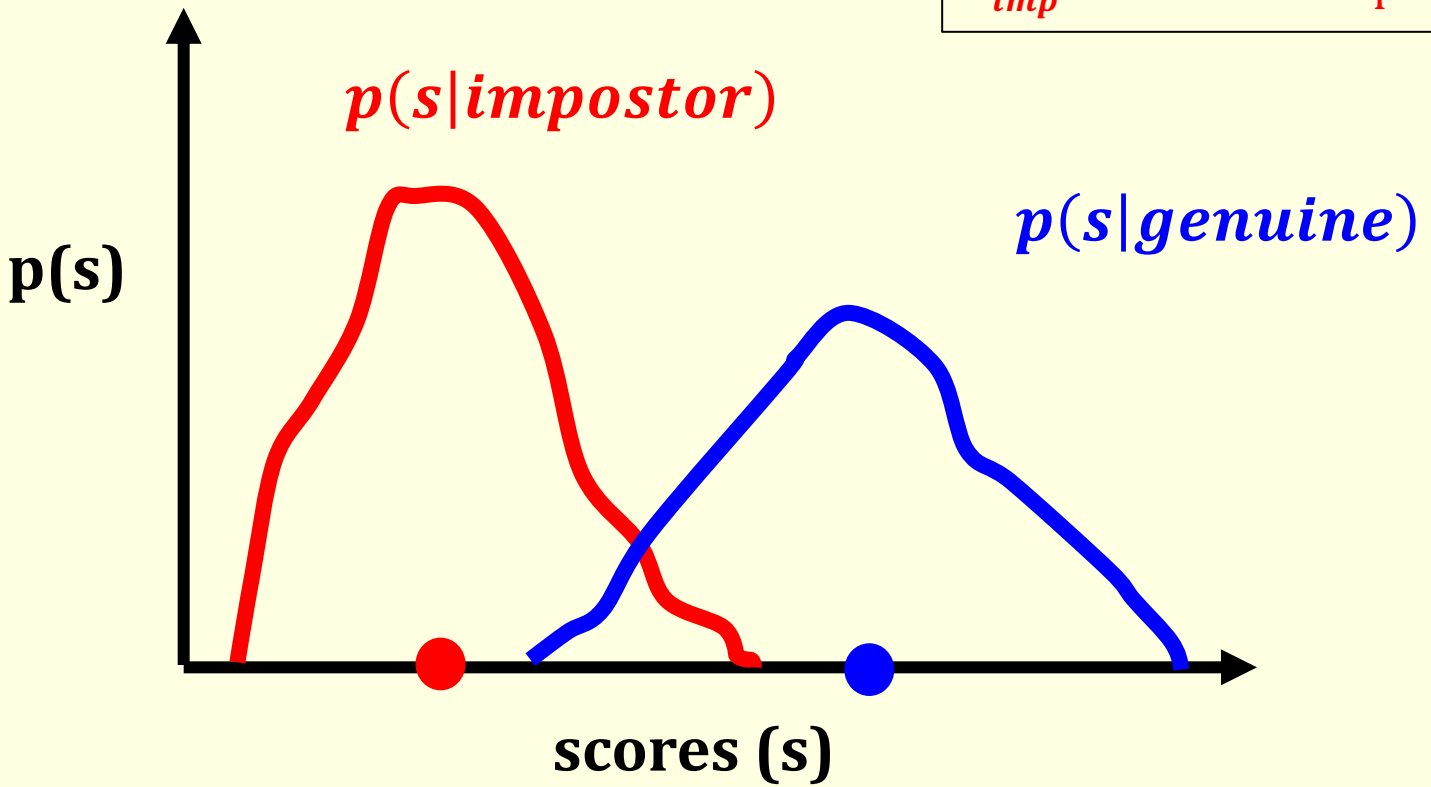


The d-prime Value

- d-prime: $\frac{\sqrt{2}\,|\mu_{gen} - \mu_{imp}|}{\sqrt{\sigma_{gen}^2 + \sigma_{imp}^2}}$

μ_{gen}: mean of genuine scores
σ_{gen}^2: variance of genuine scores

μ_{imp} :mean of impostor scores
σ_{imp}^2: variance of impostor scores

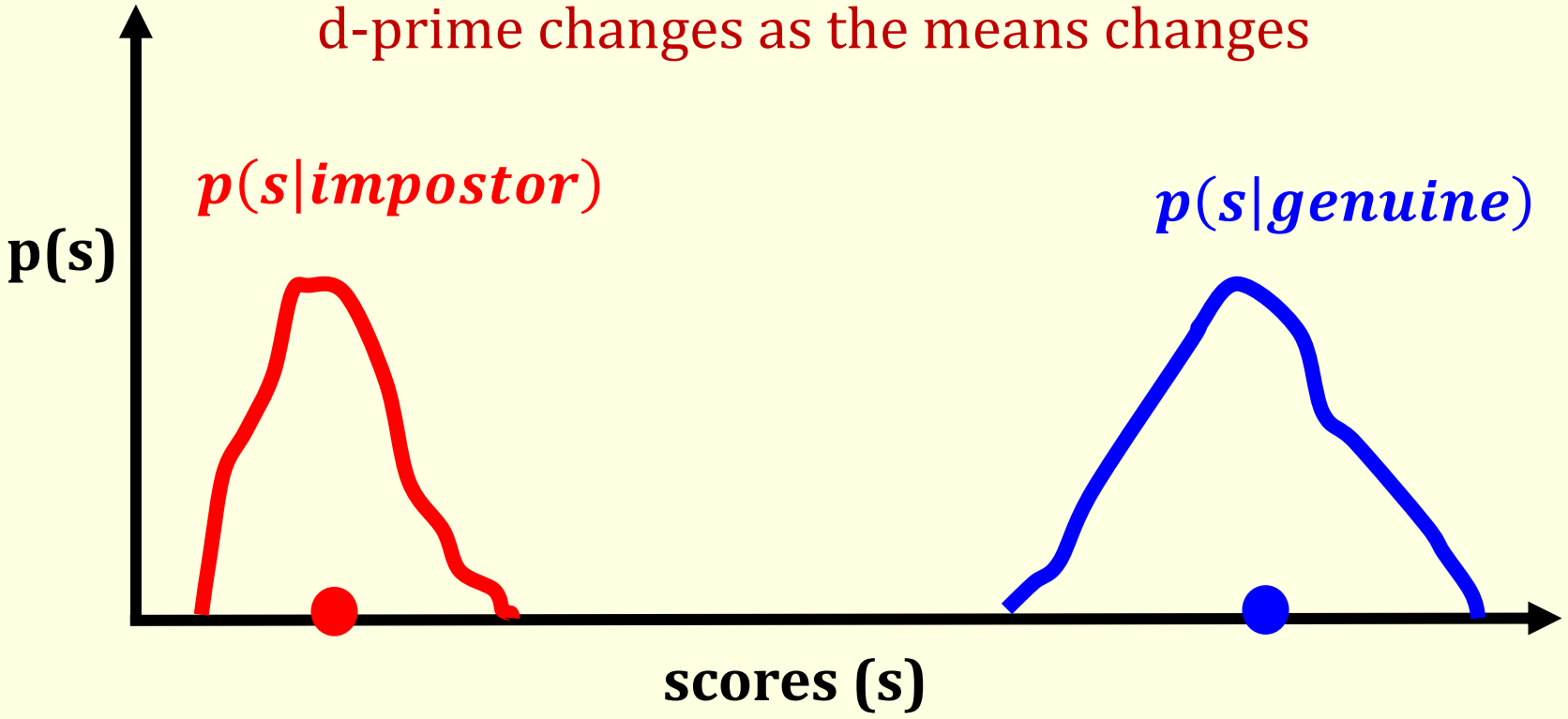


A good biometric system will exhibit a high d-prime value

Other Performance Metrics

- d-prime: $\frac{\sqrt{2}\,|\mu_{gen} - \mu_{imp}|}{\sqrt{\sigma_{gen}^2 + \sigma_{imp}^2}}$

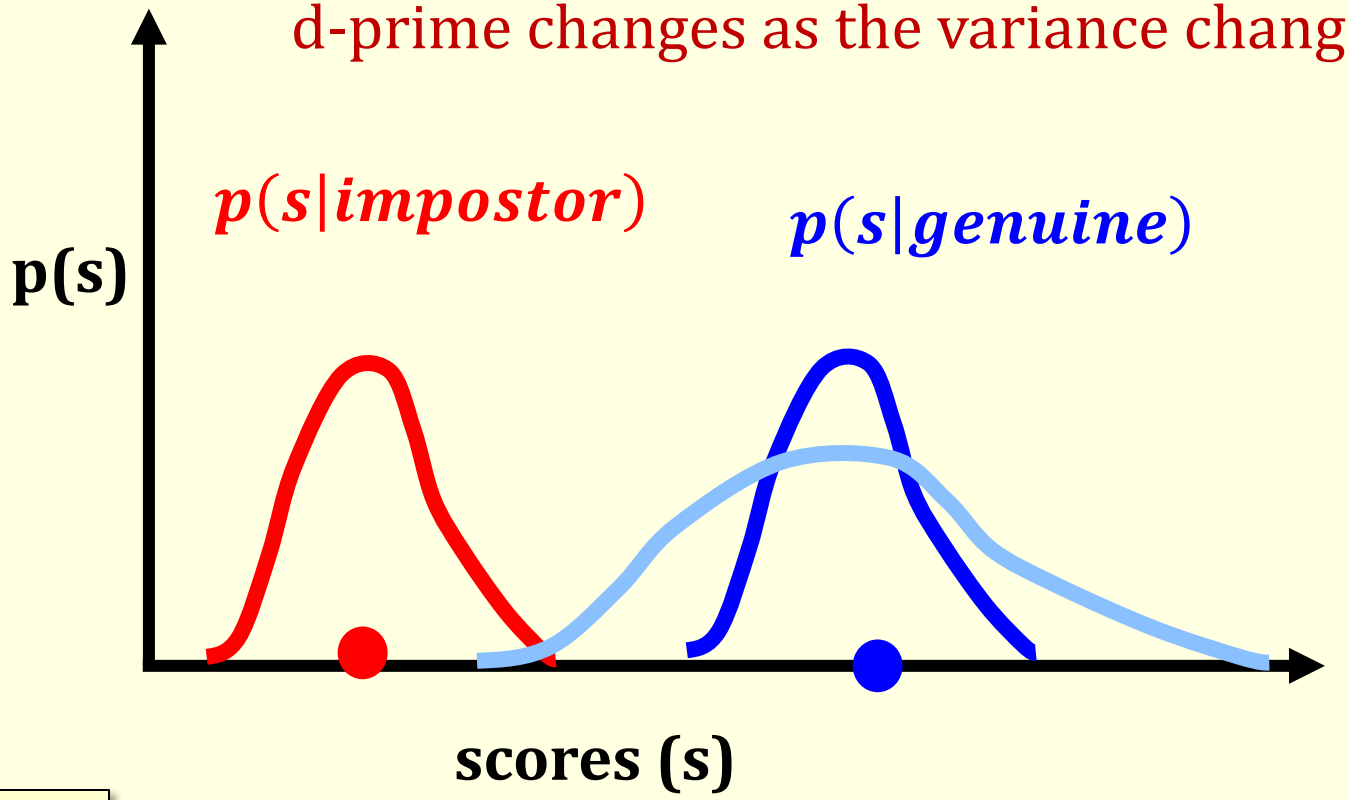
d-prime changes as the means changes



Other Performance Metrics

- d-prime: $\frac{\sqrt{2}\,|\mu_{gen} - \mu_{imp}|}{\sqrt{\sigma_{gen}^2 + \sigma_{imp}^2}}$

d-prime changes as the variance changes

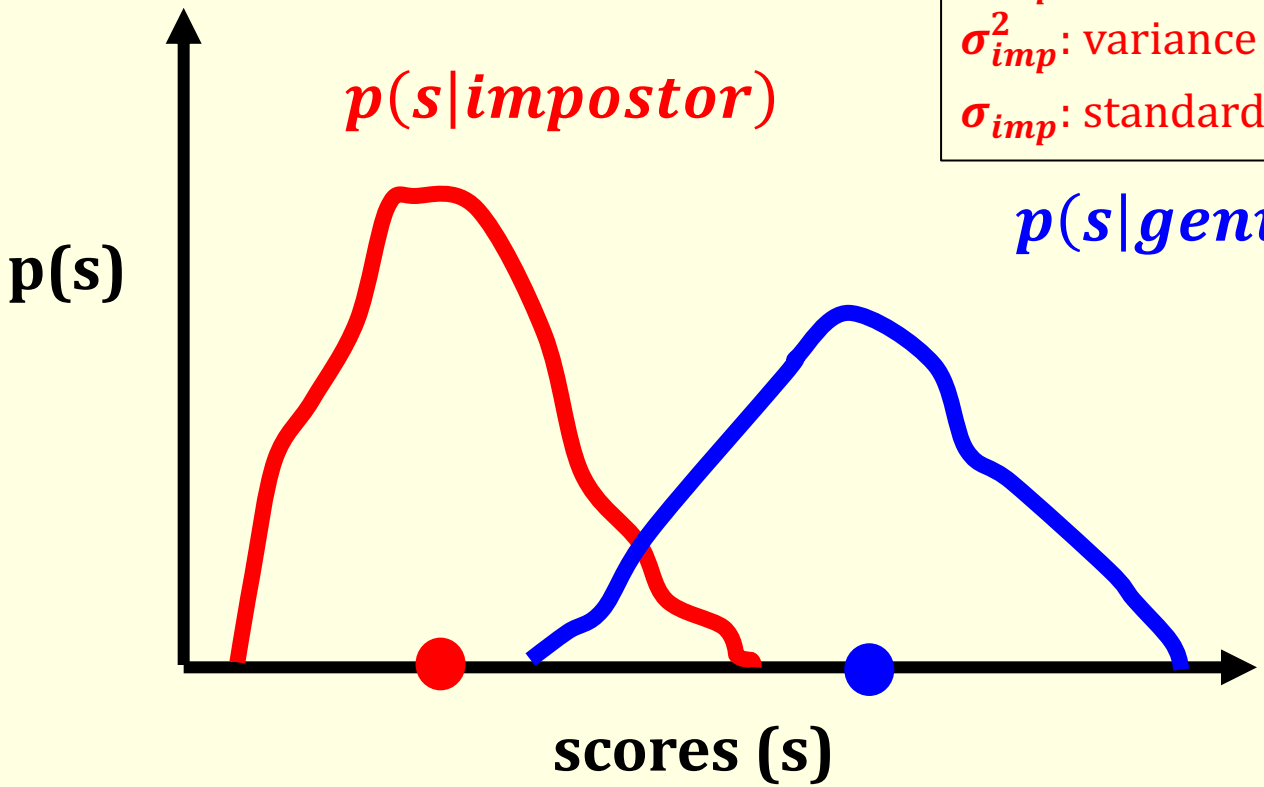


The F-ratio

- F-ratio: $\frac{|\mu_{gen} - \mu_{imp}|}{\sigma_{gen} + \sigma_{imp}}$

μ_{gen}: mean of genuine scores
σ_{gen}^2: variance of genuine scores
σ_{gen}: standard deviation

μ_{imp} :mean of impostor scores
σ_{imp}^2: variance of impostor scores
σ_{imp}: standard deviation



A good biometric system will exhibit a high F-ratio value

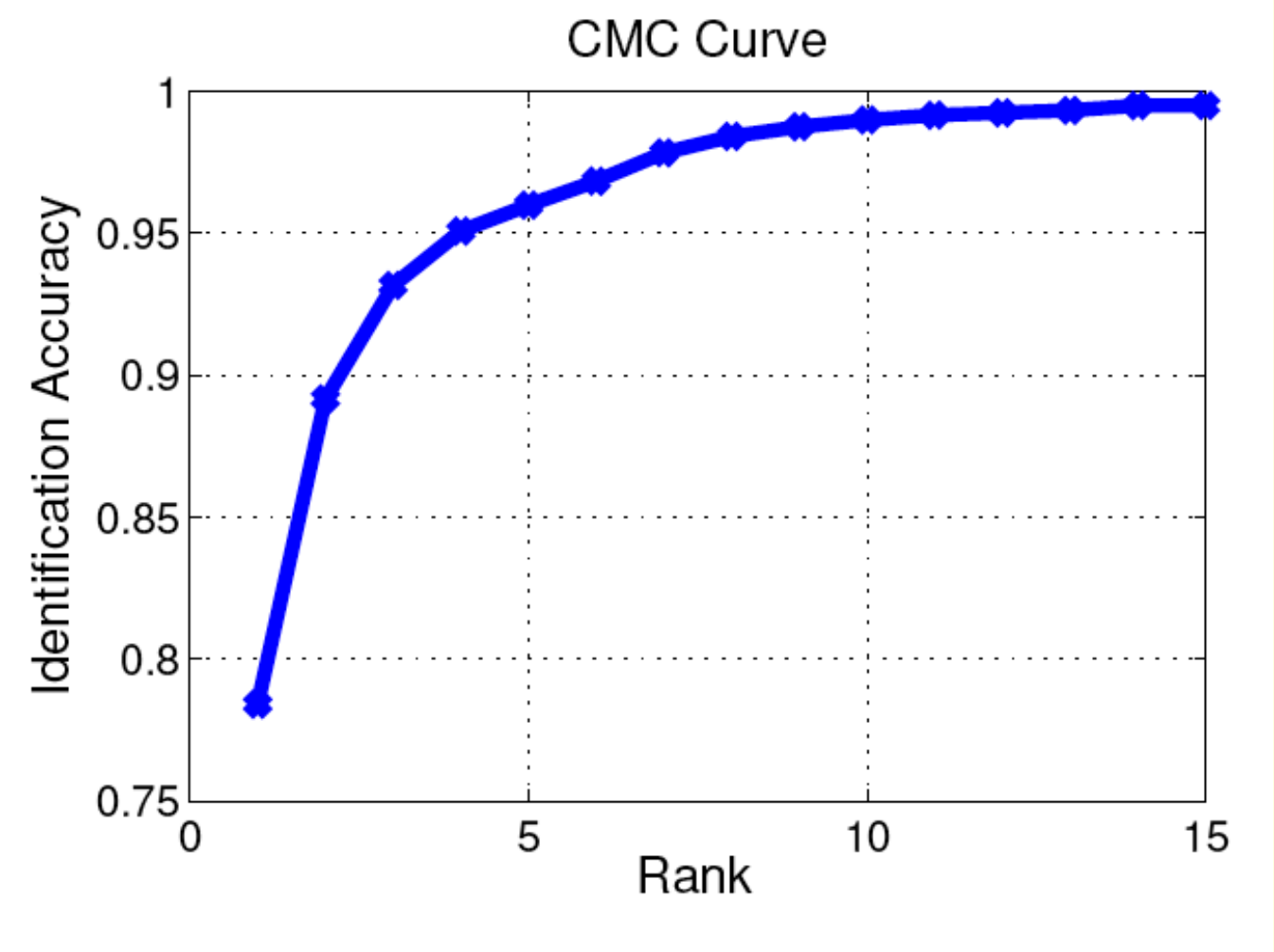
Summary of Performance Metrics

- DET Curve: Closer to the origin, the better
- Equal Error Rate (EER): Lower the better
- FNMR @ pre-defined FMR: Lower the better
- d-prime: Higher the better
- F-ratio: Higher the better

CMC Curve (closed-set)

- **Closed-set: Gallery contains the correct identity of the probe**
- Each probe biometric sample is compared against all gallery samples
- The resulting scores are sorted and ranked
- Determine the rank at which a true match occurs
- True Positive Identification Rate (TPIR): Probability of observing the correct identity within the top K ranks
- CMC Curve: Plots TPIR against ranks
- CMC Curve: Rank-based metric

CMC Curve


CMC Curve
Identification Accuracy
1
0.95
0.9
0.85
0.8
0.75
0
5
10
15
Rank

Performance (Identification)

- **Open-set: Gallery may or may not contain the correct identity of the probe**
- Each probe biometric sample is compared against all gallery samples
- A set of all gallery samples whose scores exceed a threshold is returned
- True Positive Identification Rate (TPIR): Probability of observing the correct identity in the returned set
- False Positive Identification Rate (FPIR): Probability that the correct identity does not occur in the returned set

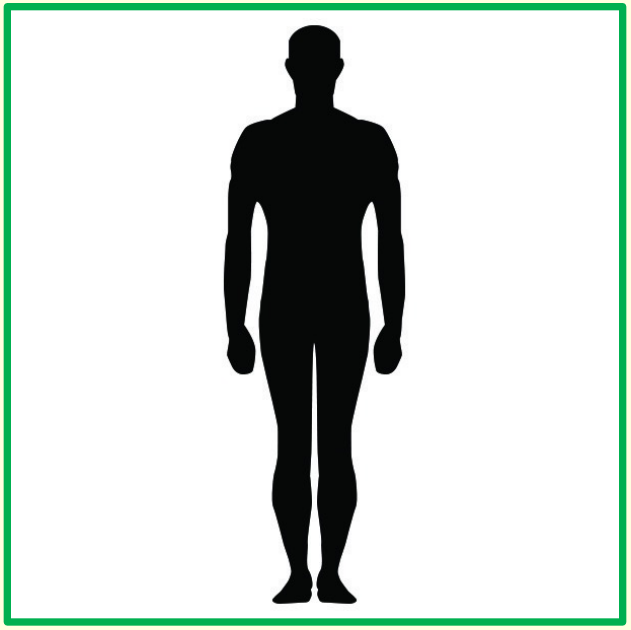
Beyond Pattern Recognition

- Ensuring that the input data is uncorrupted and from a real person
- Protecting the biometric templates in the database
- Ensuring the privacy of an individual

Biometric System

BIOMETRIC TRAIT

HUMAN MACHINE INTERFACE



PERSON

Attributes of a Biometric Trait

- **Uniqueness** (Is it distinctive across individuals?)
- **Permanence** (Does it change over time?)
- **Universality** (Does every subject have it?)
- **Collectability** (Can it be measured quantitatively?)
- **Acceptability** (Is it acceptable to the subjects?)
- **Performance** (Does it meet error rate, throughput, etc.?)
- **Vulnerability** (Can it be easily spoofed or obfuscated?)
- **Integration** (Can it be embedded in the application?)

No biometric trait is "optimal", but many are "admissible"

Attributes of the System

- **Overt vs Covert** (Is the subject aware?)
- **Attended vs Unattended** (Is there operator involvement?)
- **Controlled vs uncontrolled operation** (Are the environment and interface parameters controlled?)
- **Open vs Closed** (Are templates exchanged or shared across applications?)

Attributes of the Person

- **Cooperative vs Non-cooperative** (Is the subject cooperative?)
- **Habituated vs Non-habituated** (Has the subject adapted to the system?)

Challenges in a Biometric System

- **Noise in sensed data**: e.g., defective sensors or unfavorable ambient/physiological conditions
- **Intra-user variations**: e.g., incorrect interaction with sensor, variations in user's biometric trait, sensor characteristics are modified
- **Distinctiveness**: e.g., capacity of biometric template is limited
- **Non-universality**: e.g., all users may not be able to successfully present the trait
- **Presentation attacks**: circumvent the system by using artificial traits or modified traits leading to spoofing or obfuscation

Intra-user variations

© Nostra

FNMR: False Non-Match Rate (False Negative)

Inter-user similarity

TWIN BROTHERS
© Martin Schoeller

MOTHER DAUGHTER
© PleasantonWeekly.Com

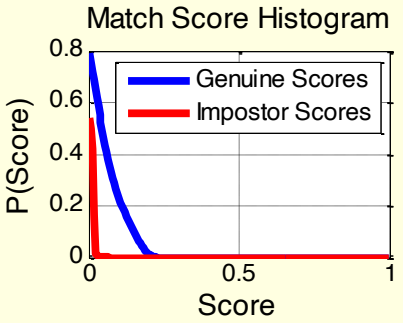
FMR: False Match Rate (False Positive)

Face Similarity – Twins

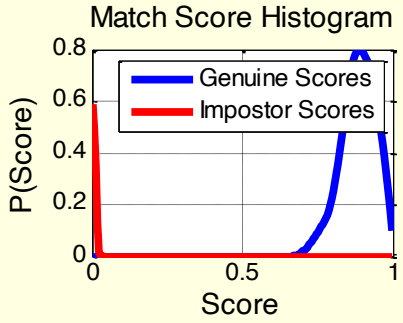
https://www.youtube.com/watch?v=e8-yupM-6Oc

FMR and FNMR Varies Across Subjects

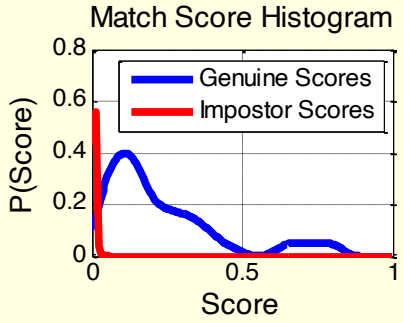
User 1



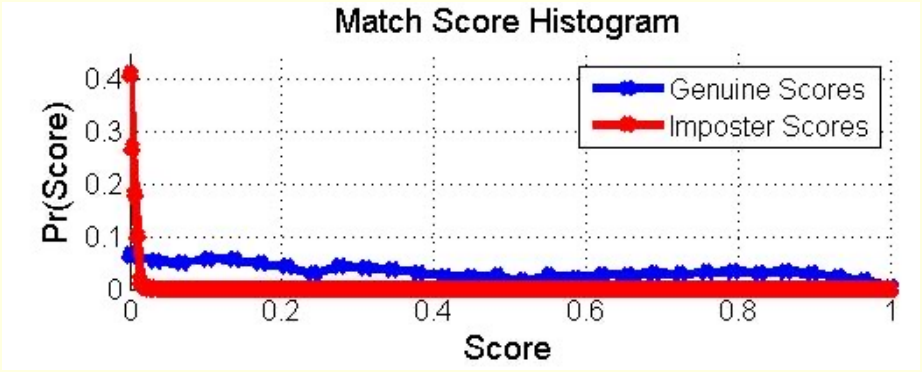
User 2



User 3

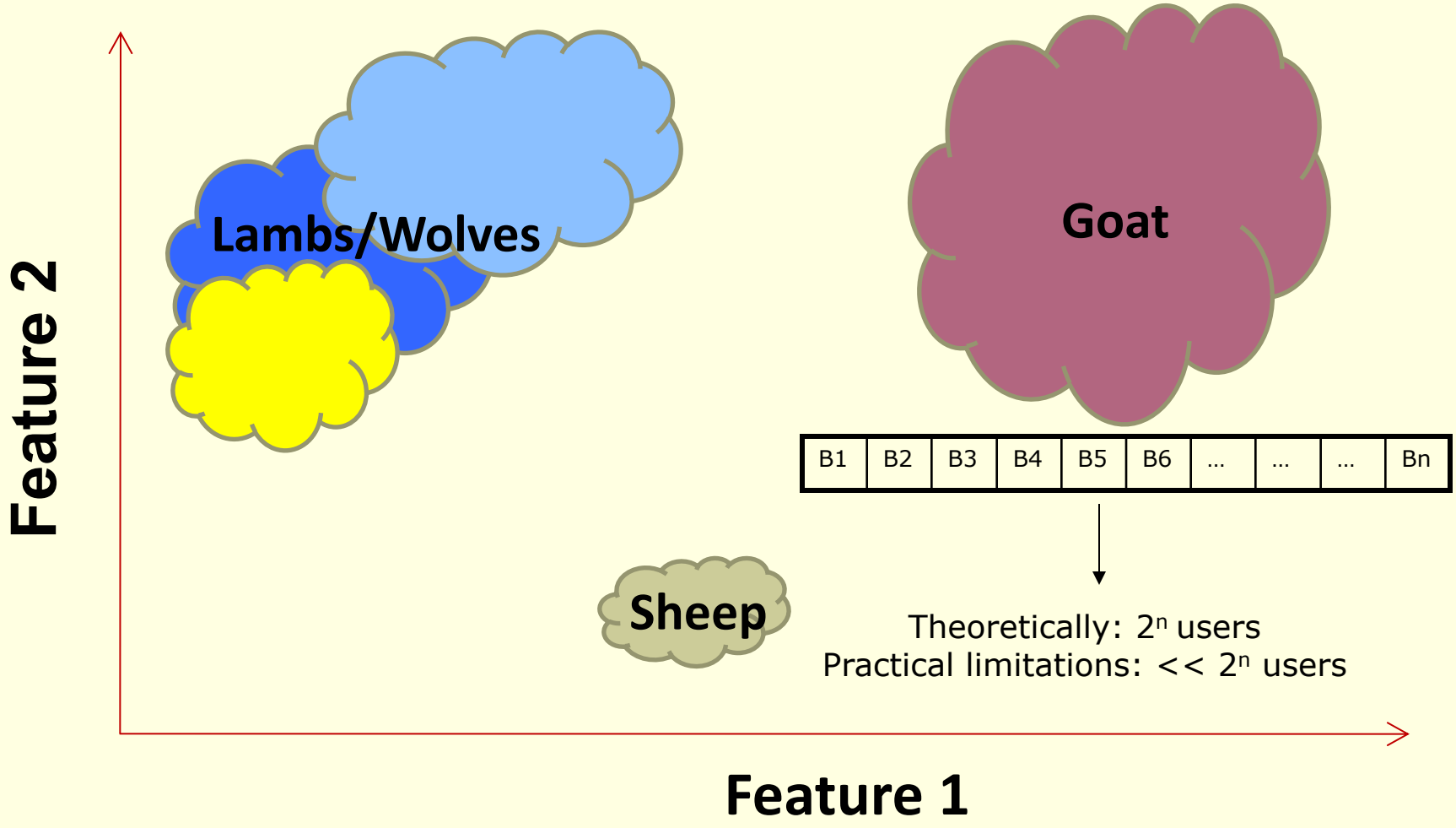


- Subject-dependent FMRs and FNMRs



Capacity of a template

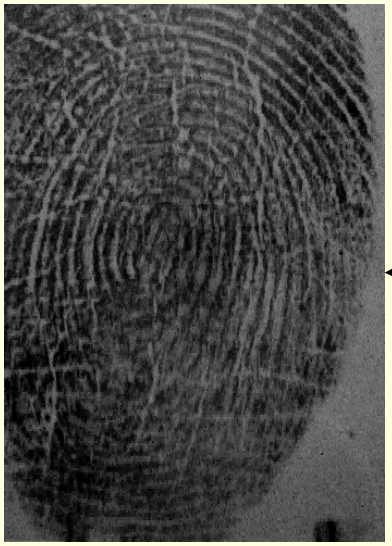
- Existence of a biometric "zoo": Different categories of users impact error rates in a different manner



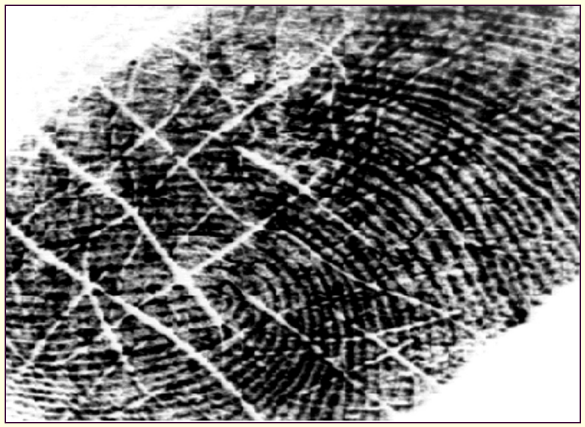
*Ross, Rattani, Tistarelli, "Exploiting the Doddington Zoo Effect in Biometric Fusion," BTAS 2009

Noisy Data

During enrolment



During recognition

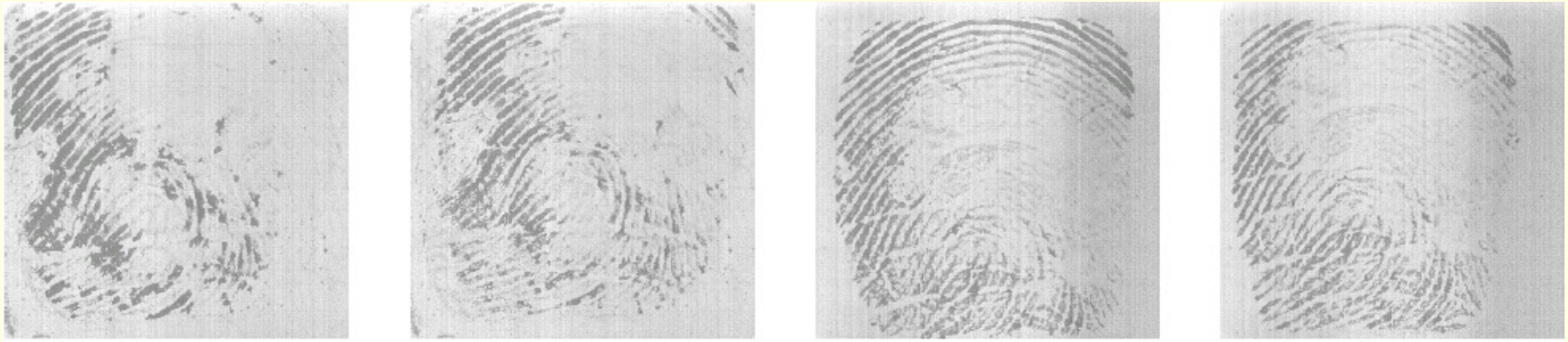


Noise due to smearing, residual deposits, cuts and folds, etc

Can impact both FMR and FNMR

Non-universality

- Some people may consistently offer poor quality fingerprint images which means they have to be identified by some other means

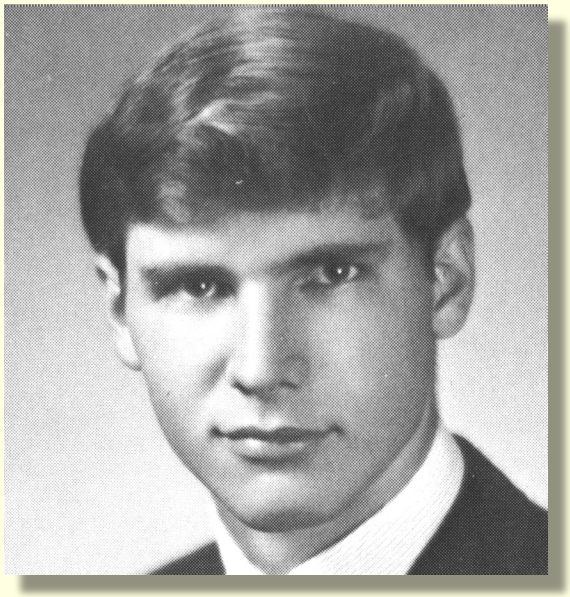


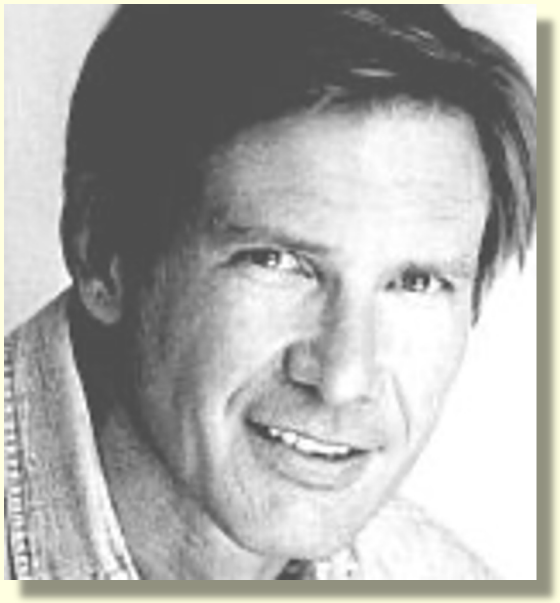
Four impressions of a user's print exhibiting incomplete ridge information

FTE: Failure-to-Enroll Problem

Jain, Prabhakar, Ross, "Fingerprint Matching: Data Acquisition and Performance Evaluation", MSU Technical Report TR99-14, 1999.

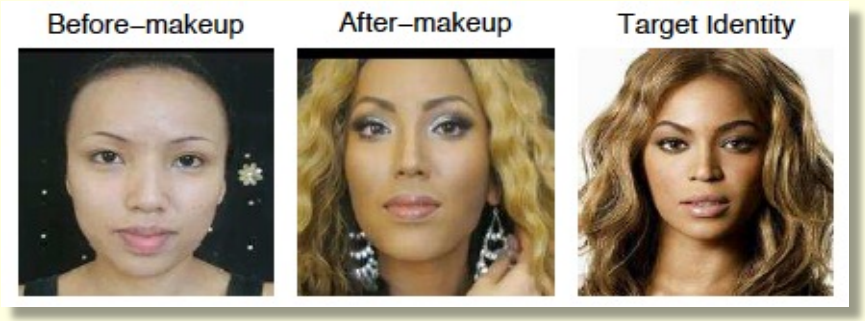
Biometric Ageing

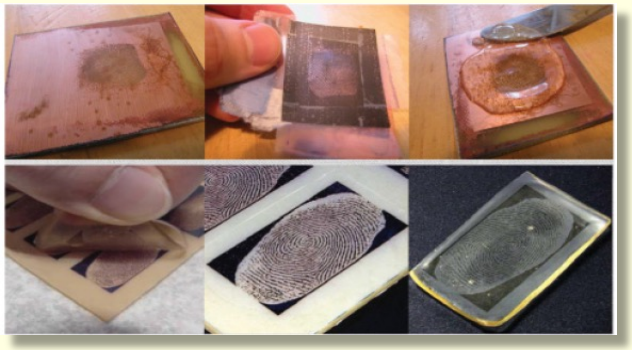


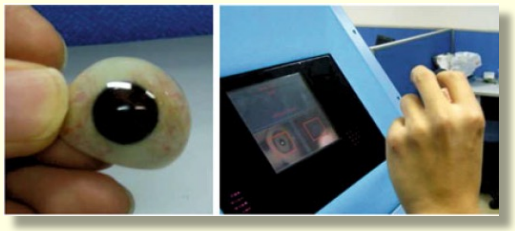


Spoofing: Presentation Attack

- Spoofing: Altering one's trait or creating a physical artifact in order to "spoof" another person's trait

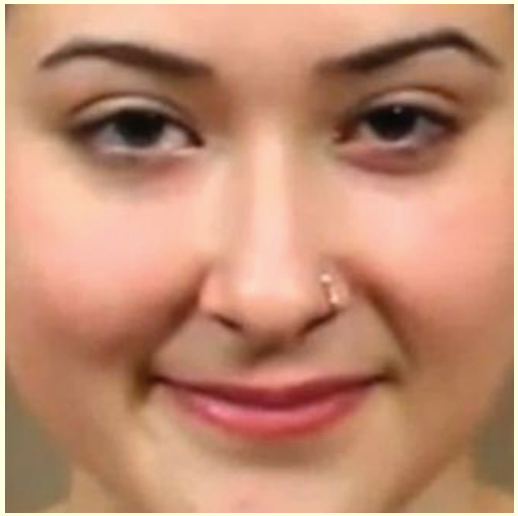






Cosmetics: Spoofing

- Cosmetics: To spoof another person's face image



Before-makeup

After-makeup

Target identity

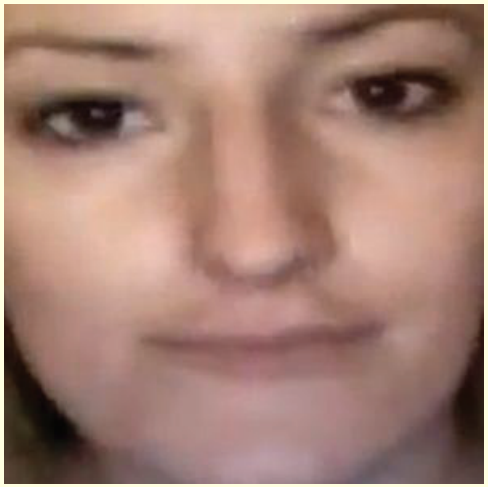
Rank 26 → Rank 6

[13,334 gallery images]

Chen et al, "Spoofing Faces Using Makeup: An Investigative Study", ISBA 2017

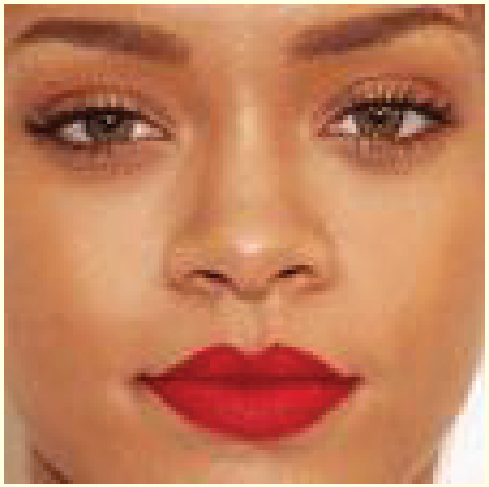
Cosmetics: Spoofing

- Cosmetics: To spoof another person's face image



Before-makeup

After-makeup



Target identity

Rank 734 → Rank 1

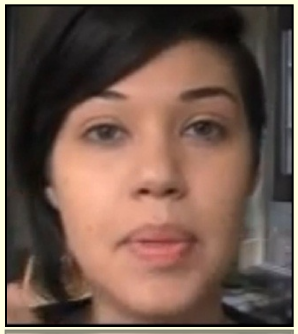
[13,334 gallery images]

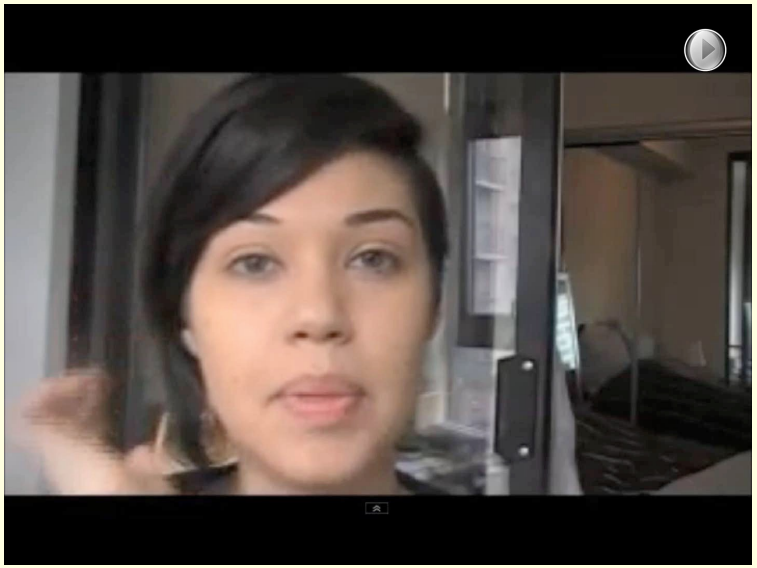
Chen et al, "Spoofing Faces Using Makeup: An Investigative Study", ISBA 2017

Obfuscation: Presentation Attack

- Obfuscation: Masking one's own identity by altering the trait

BEFORE MAKEUP





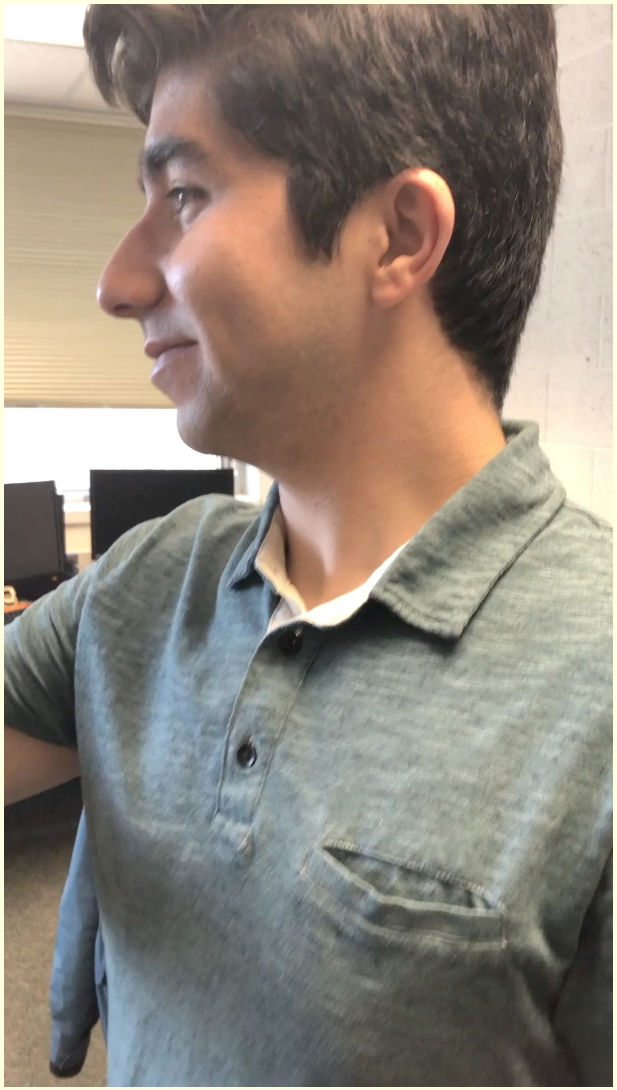
AFTER MAKEUP

Dantcheva et al, "Can Facial Cosmetics Affect the Matching Accuracy of Face Recognition Systems?", BTAS 2012

Fingerprint Anti-spoofing

https://www.youtube.com/watch?v=P8zm1i9gJuE

Face Anti-Spoofing



Atoum et al., "Face Anti-Spoofing Using Patch and Depth-based CNNs", IJCB 2017

Fingerprint Alteration

- 1995: Alexander Guzman was arrested by Florida officials for possessing a false passport
- He was found to have mutilated fingerprints
- After a two-week search based on manually reconstructing the damaged fingerprints and searching the FBI database, the reconstructed fingerprints were linked to the fingerprints of Jose Izquiredo who was an absconding drug criminal

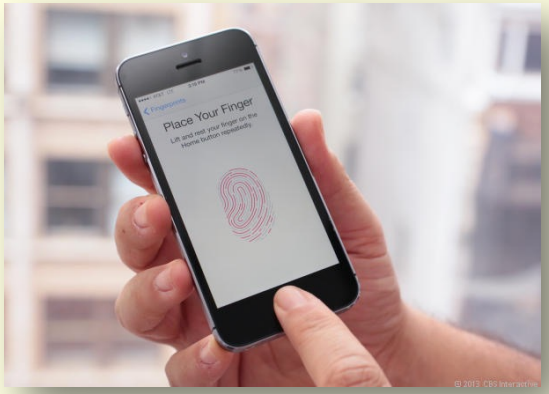
The “Z”-cut

- His fingerprint mutilation process consisted of three steps: making a ‘Z’ shaped cut on the fingertip; lifting and switching two triangles; and stitching them back.

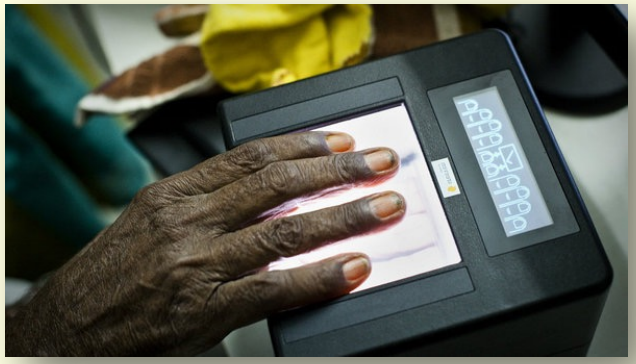
Biometric Applications

Iris: Health Care

Fingerprint: US OBIM



Fingerprint: Apple Touch ID



Fingerprint: Refugee Tracking

Finger Vein: Japan ATMs

Domains

Verification

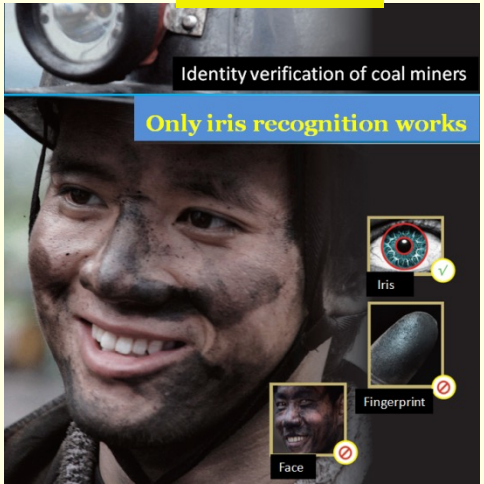
Personalization

Transactions

Healthcare

Convenience

Safety



Travel

Introduction to Biometrics

Arun Ross

Professor

Michigan State University

rossarun@cse.msu.edu

http://www.cse.msu.edu/~rossarun