Biometrics & Authentication Technologies: security issues

Andy Adler

Systems and Computer Engineering

Carleton University, Ottawa

What are Biometrics

Automatic identification of an individual based on behavioural or physiological characteristics

What are Biometrics

Automatic identification of an individual based on behavioural or physiological characteristics

← Computer based ie. fast

Forensics is the science of humans identifying humans

What are Biometrics

Automatic identification of an individual based on behavioural or physiological characteristics

Two types:

1. Verification

2. Identification

What are Biometrics

Automatic identification of an individual based on behavioural or physiological characteristics

Biometrics is **only** about identity of individual. Other technologies manage security

What are Biometrics

Automatic identification of an individual based on behavioural or physiological characteristics

Behavioural biometrics:

- Gait
- Voice
- Typing dynamics
- Signature

What are Biometrics

Automatic identification of an individual based on behavioural or physiological characteristics

Physiological Biometrics

- Fingerprint
- Face
- Iris
- Retina
- Hand Geometry
- Dental shape
- DNA

…

What is Biometrics security

- Somewhat difficult to define
    - Biometric systems implicitly have an “attacker”
- My definition: biometrics security is against
    - Stronger attacks than zero-effort impostors
    - Does not include underlying computer security

Taxonomy

Presentation attacks (spoofing)

- appearance of the biometric sample is physically changed or replaced.

Biometric processing attacks:

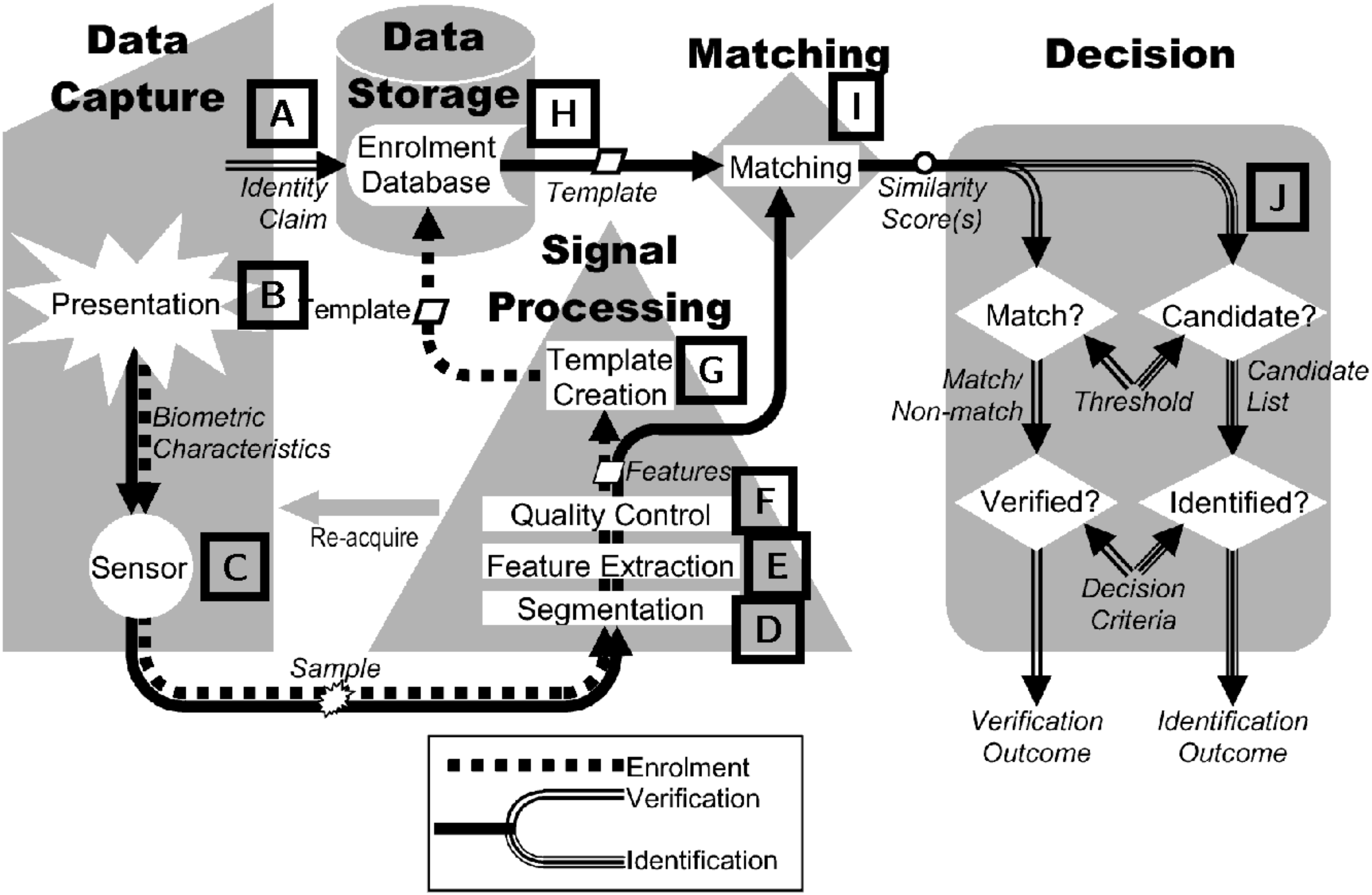
- an understanding of the biometric algorithm is used to cause incorrect processing and decisions,

Software and networking vulnerabilities:

- based on attacks against the computer and networks on which the biometric systems run, and

Social attacks:

- in which the authorities using the systems are fooled.



ISO Biometrics Concept Diagram

Biometrics Vulnerabilities

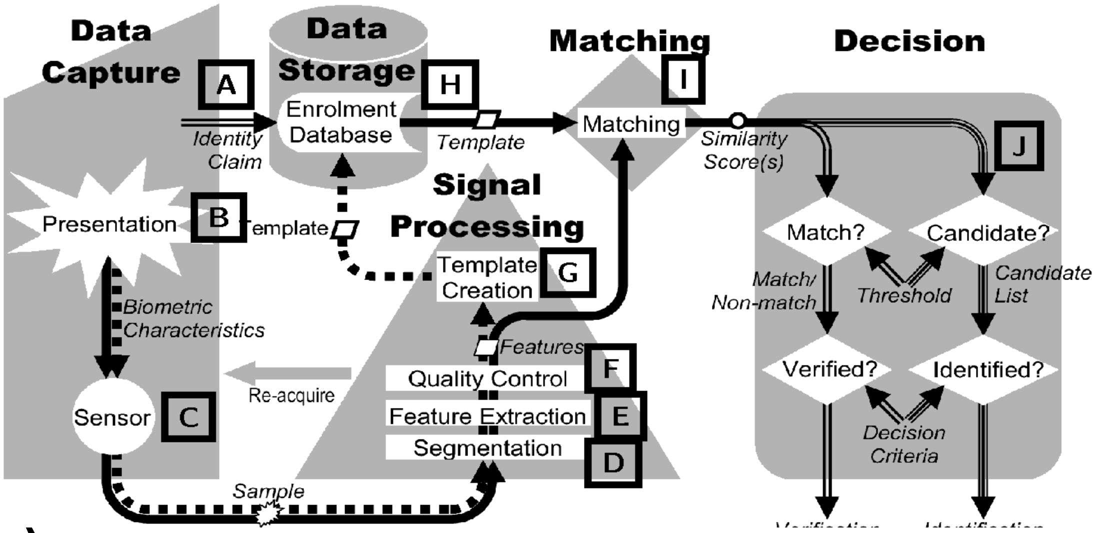
Taxonomy (from Maltoni et al, 2003):

- □ Circumvension
- □ Covert acquisition
- □ Collusion / Coercion
- □ Denial of Service

Biometrics Security Issues

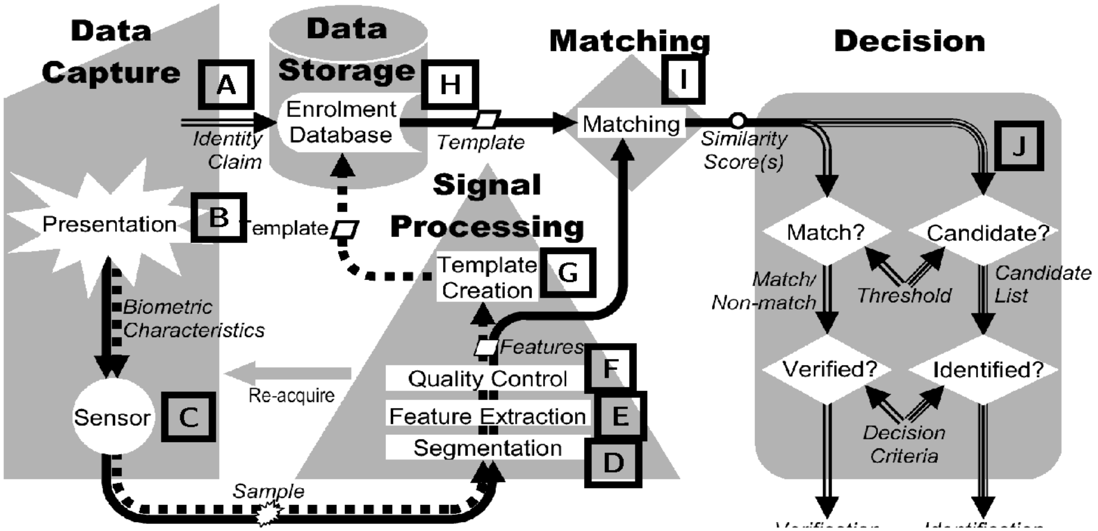
- Biometrics are not secrets
- Biometrics cannot be revoked
- Biometrics have secondary uses

IdentityClaim [A]



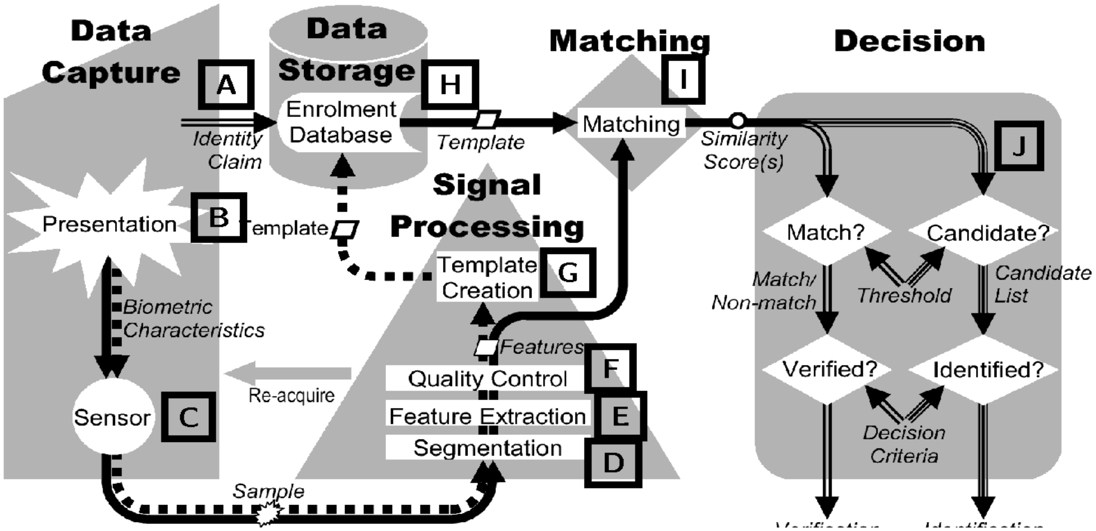
- ID Claim (via token)
  needed for most biometric functions
- Vulnerable to all ID document fraud

Presentation [B]



- Makeup / tilt head / cut fingerprints
- Avoid detection (False Neg) easier than Masquerade (False Pos)

Presentation [B]

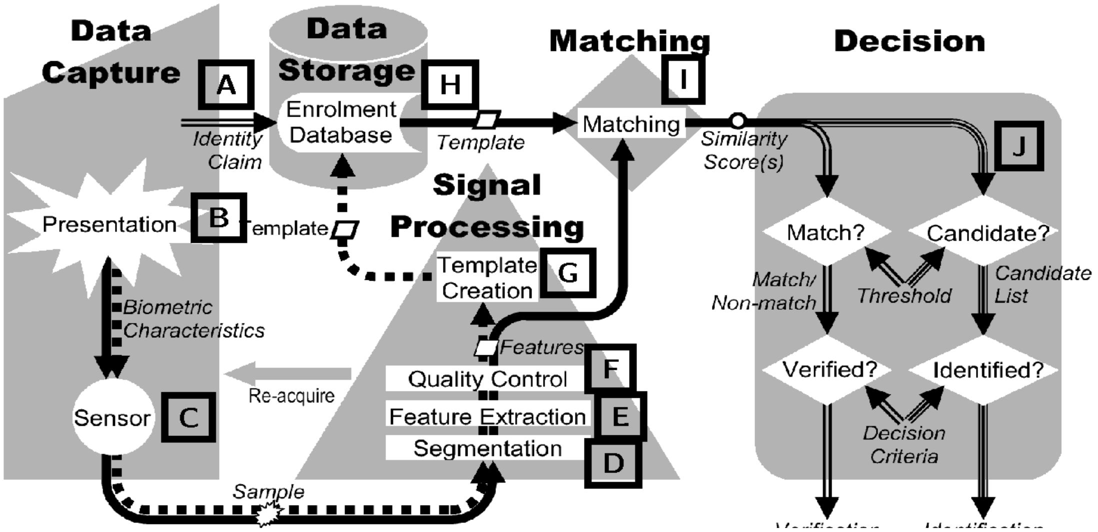


Spoofing: *Attempt to fool biometric system with artificial biometric*

- Fingerprint: gummy, etching, mould
- Face, Iris, Voice

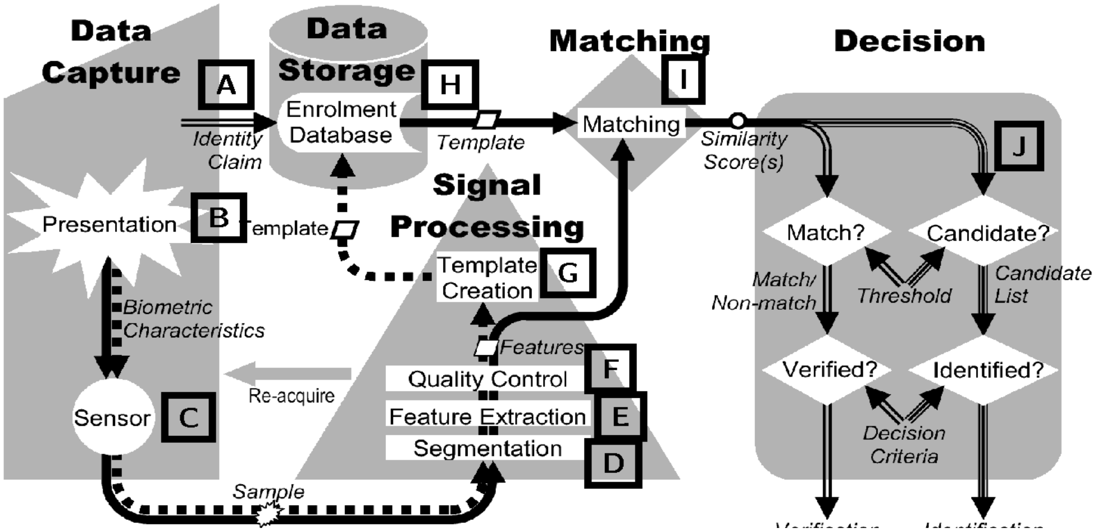
Liveness: Approach to detect spoofing attempts

Sensor [C]



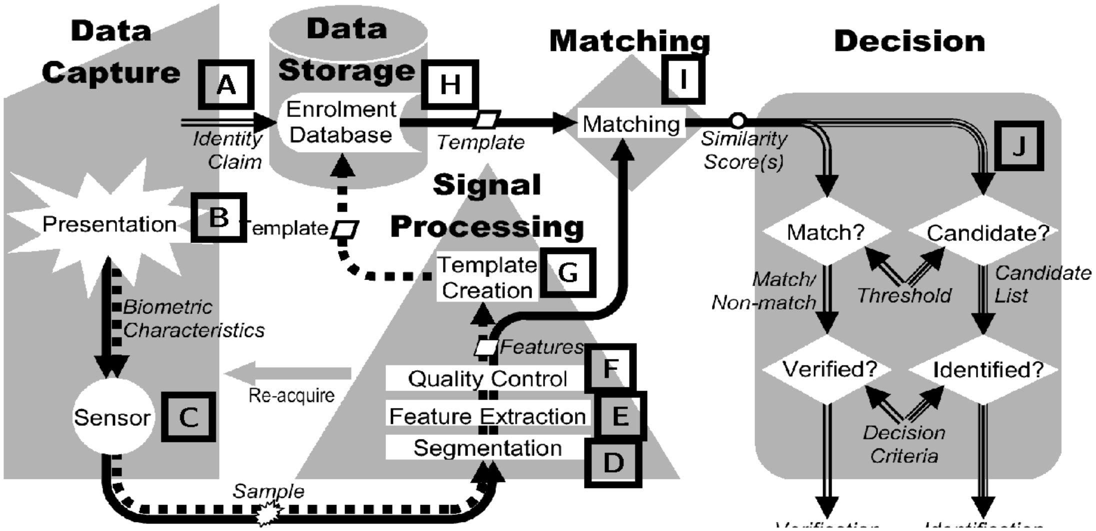
- Subvert or replace sensor hardware
- Eavesdropping / replay
- Bypass biometric completely

Segmentation [D]



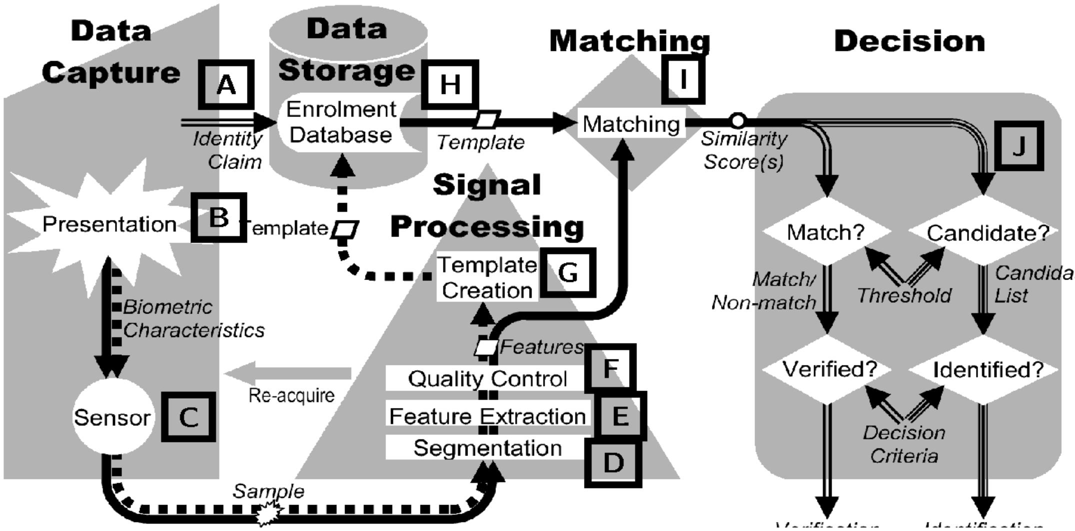
- Segmentation isolates biometric image from background
- Damage fingerprint core / cover one eye

Feature Extraction [E]



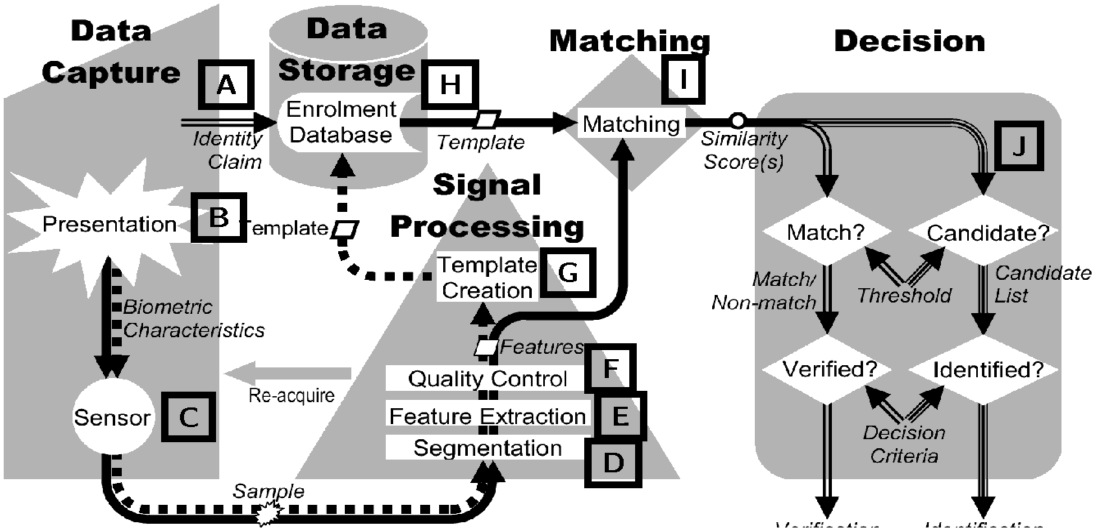
- Use knowledge of algorithm to construct "features" to confuse algorithm
- Biometric "Zoo"
  - □ Sheep – system performs well
  - □ Goats – difficult to recognize
  - □ Lambs – easy to imitate
  - □ Wolves – likely to identify as another

Quality Control [F]



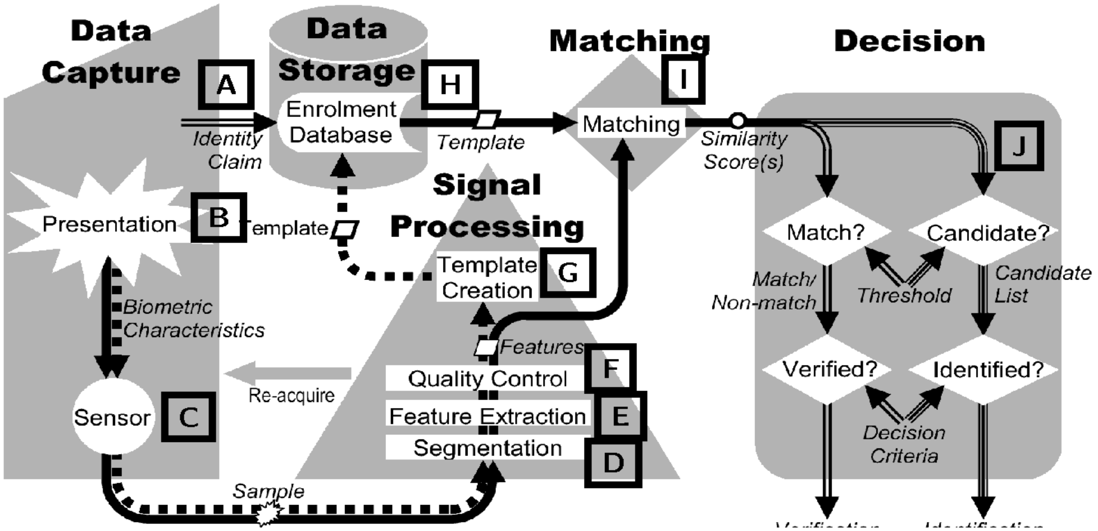
- Quality used to prevent enrolment of poor images
- Misclassify as good – force decrease of internal thresholds
- Misclassify as poor - DoS

Template Creation [G]



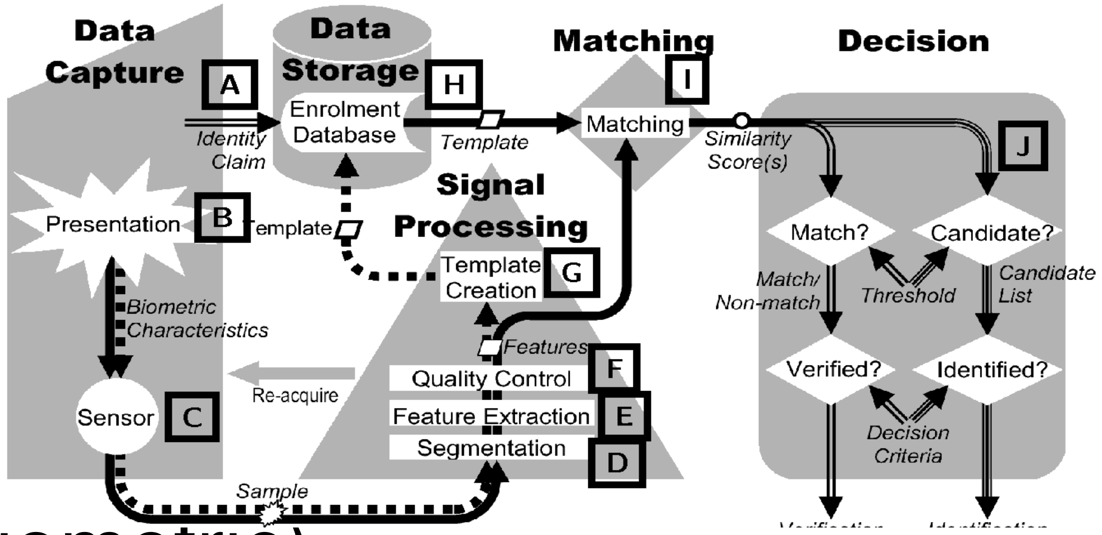
- Regeneration of images from template storage

Data Storage [H]



- Storage in:
  - □ Government database
  - □ ID card
  - □ Electronic Devices
- Vulnerable to all flaws in computer system

Matching [I]

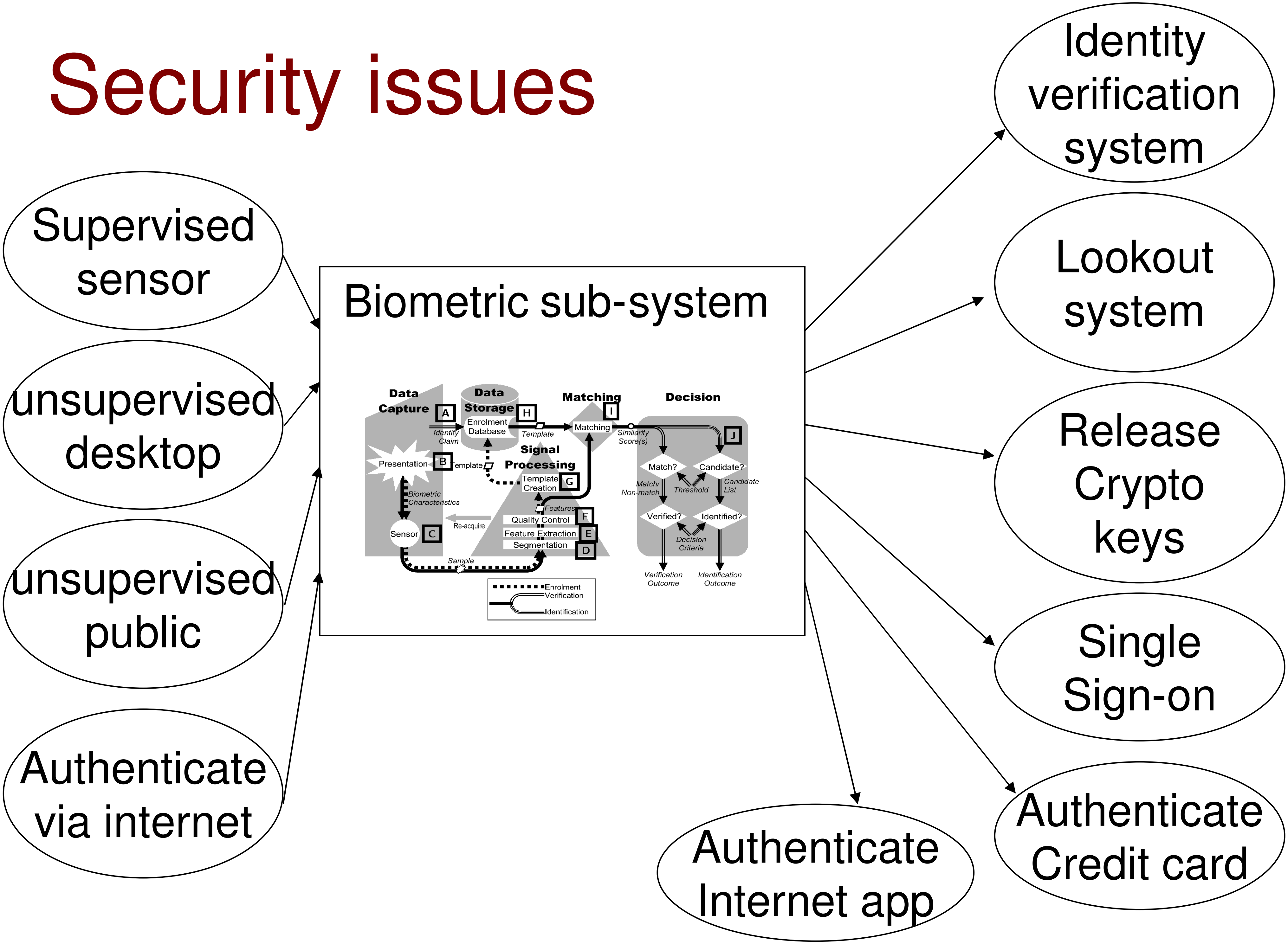


- Need
  - □ threshold (single biometric)
  - □ fusion parameters (multiple biometrics)
- Modify threshold choices by specific template enrolments

Decision [J]

- Fatigue of human operators

Security issues



Supervised sensor

unsupervised desktop

unsupervised public

Authenticate via internet

Biometric sub-system

Identity verification system

Lookout system

Release Crypto keys

Single Sign-on

Authenticate Credit card

Authenticate Internet app

Biometric template security [E]

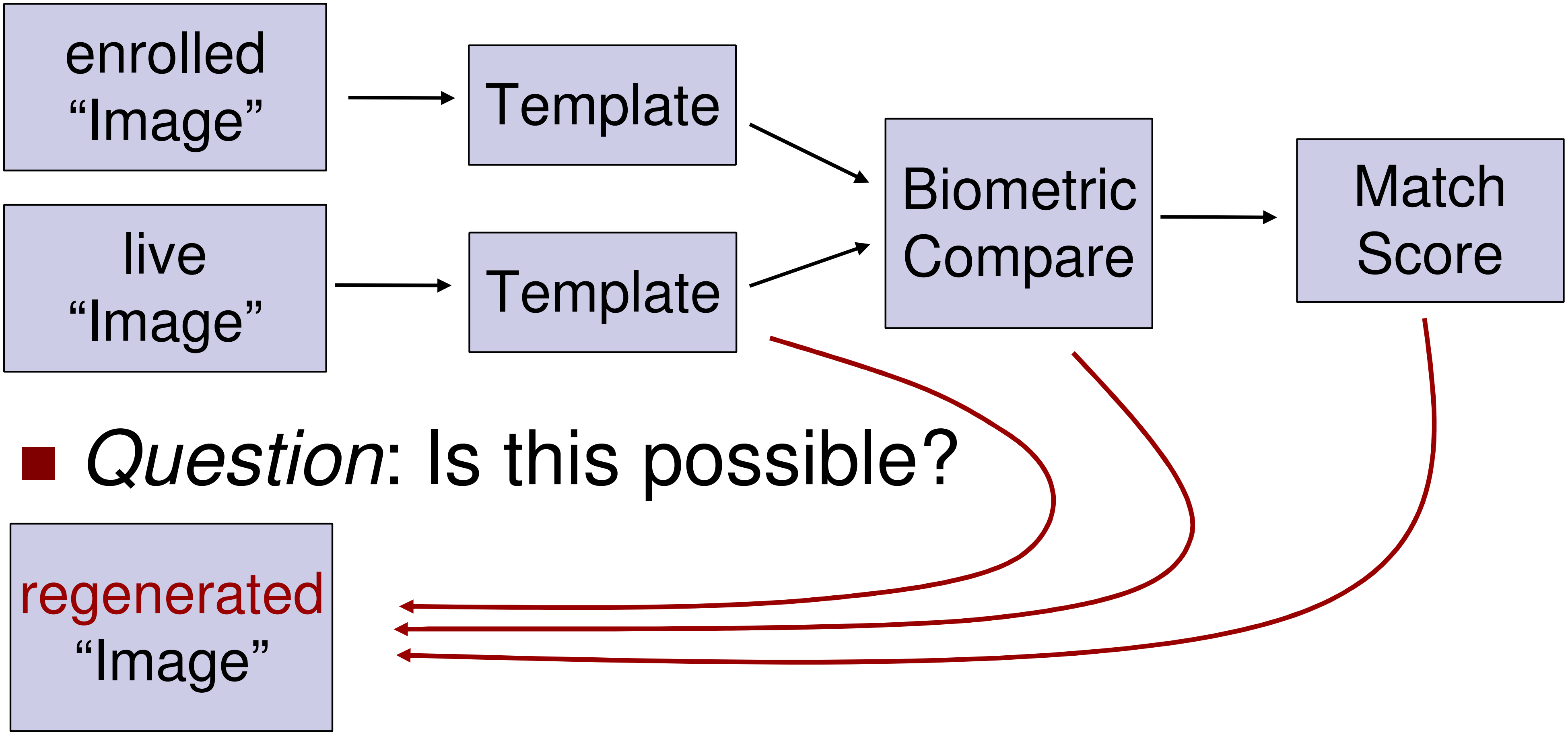
It is claimed to be impossible or infeasible to recreate the enrolled image from a template.

Reasons:

- □ templates record features (such as fingerprint minutiae) and not image primitives
- □ templates are typically calculated using only a small portion of the image
- □ templates are much smaller than the image
- □ proprietary nature of the storage format makes templates infeasible to "hack".

Images can be regenerated …?

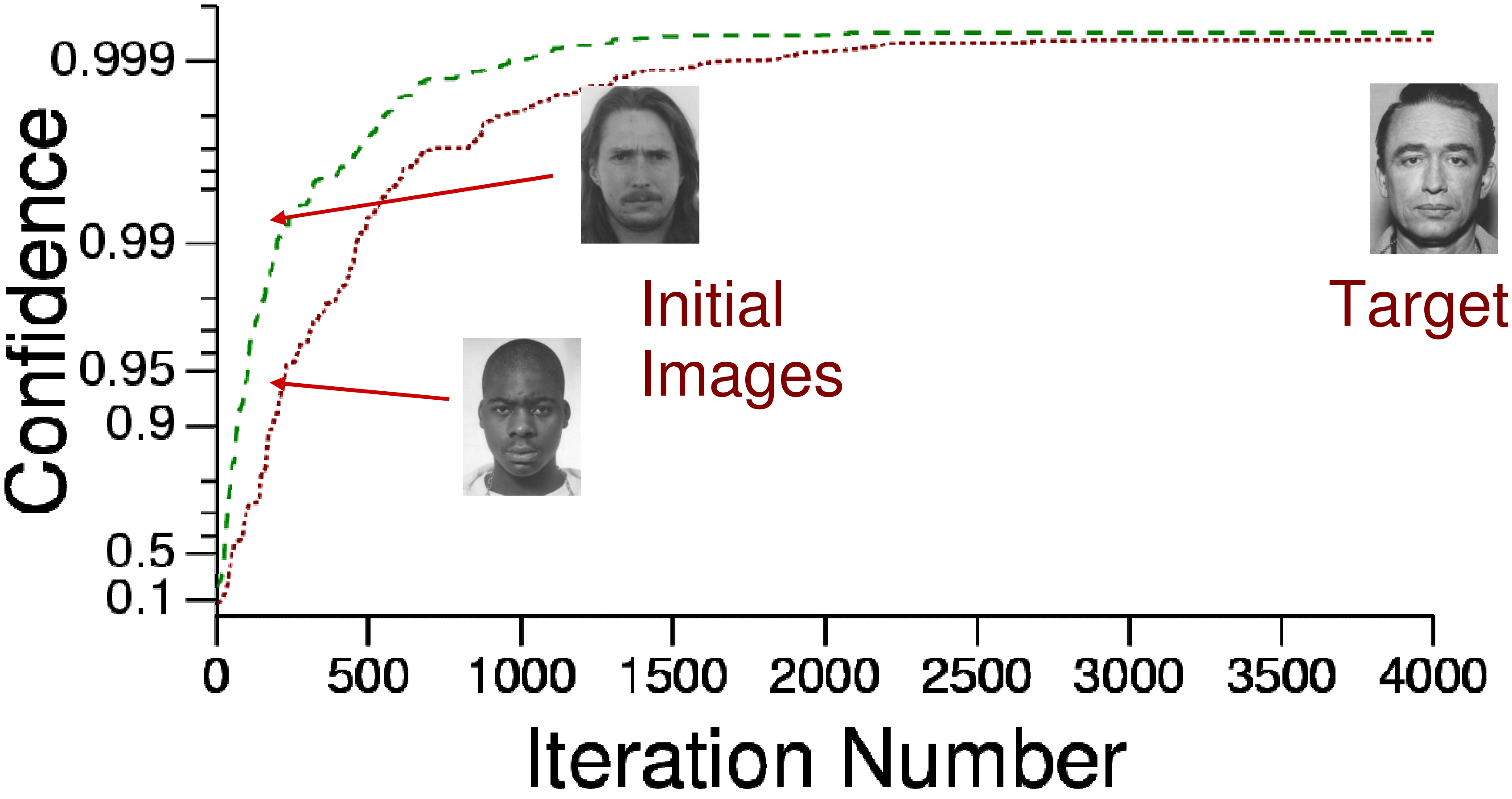
- Typical Biometric processing



- *Question*: Is this possible?

Hill-climbing: begin with a guess, make small modifications; keep modifications which increase the match score

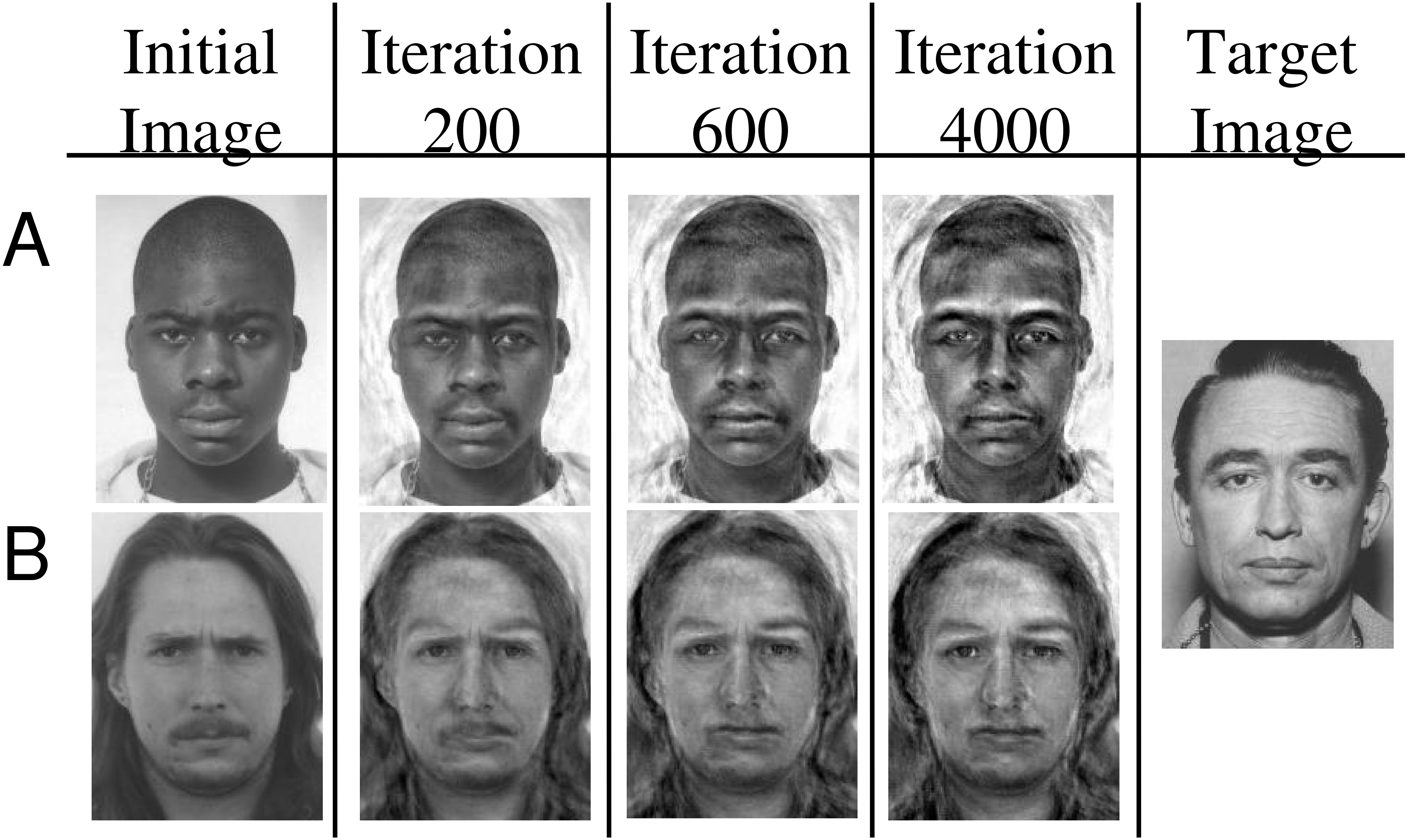
Requirement: access to a match scores



Initial Images

Target

Results

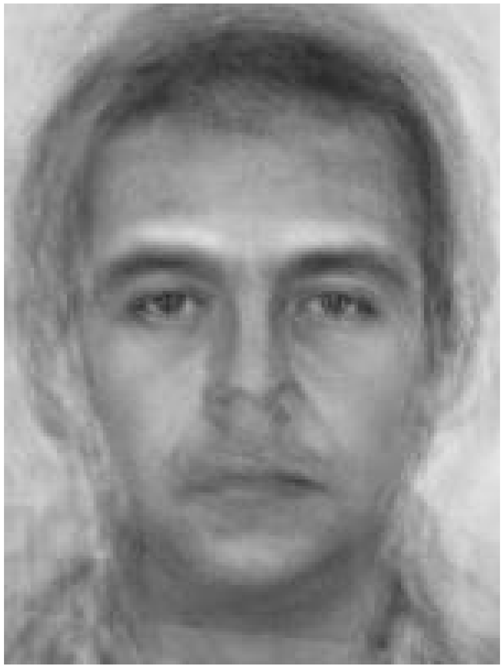


Initial Image | Iteration 200 | Iteration 600 | Iteration 4000 | Target Image

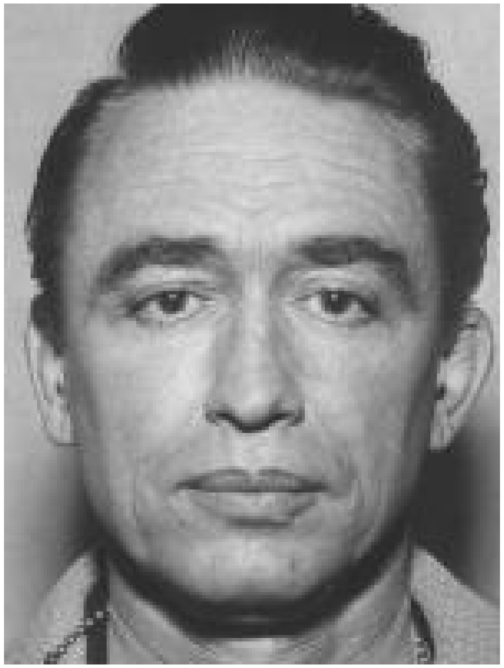
A

B

Improved regenerated image



Average of 10
Best Estimates



Target Image

Extensions to this approach

Recently, this approach has been extended to fingerprint images

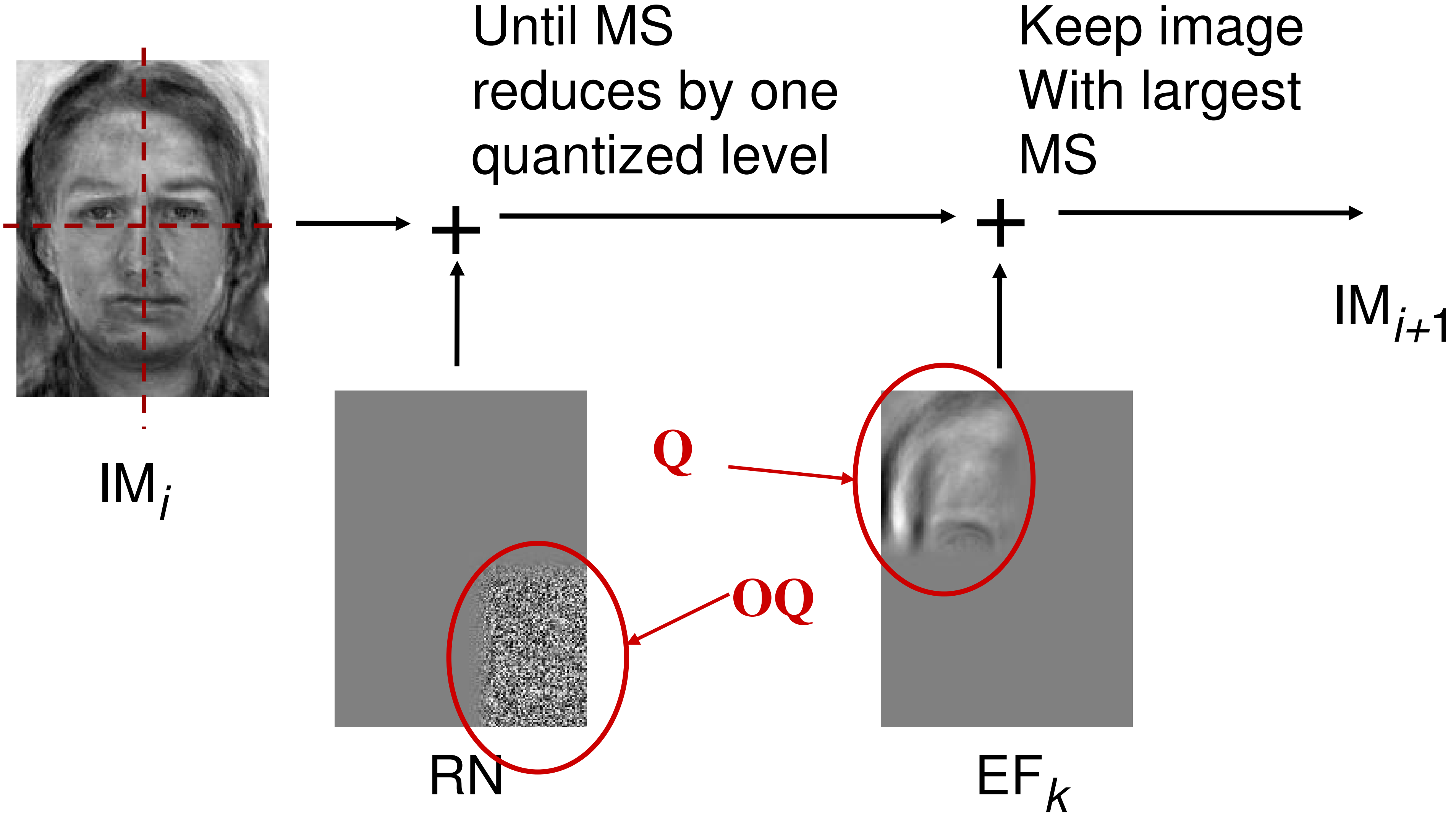
- U.Uludag developed an approach to modify a collection of minutiae
- A.Ross has developed a fingerprint image regenerator

Protection:

According to BioAPI

- “…allowing only discrete increments of score to be returned to the application eliminates this method of attack.”
- Idea: most image modifications will not change the match score

Modified "hill-climbing"



Until MS
reduces by one
quantized level

Keep image
With largest
MS

IM_{i+1}

IM_i

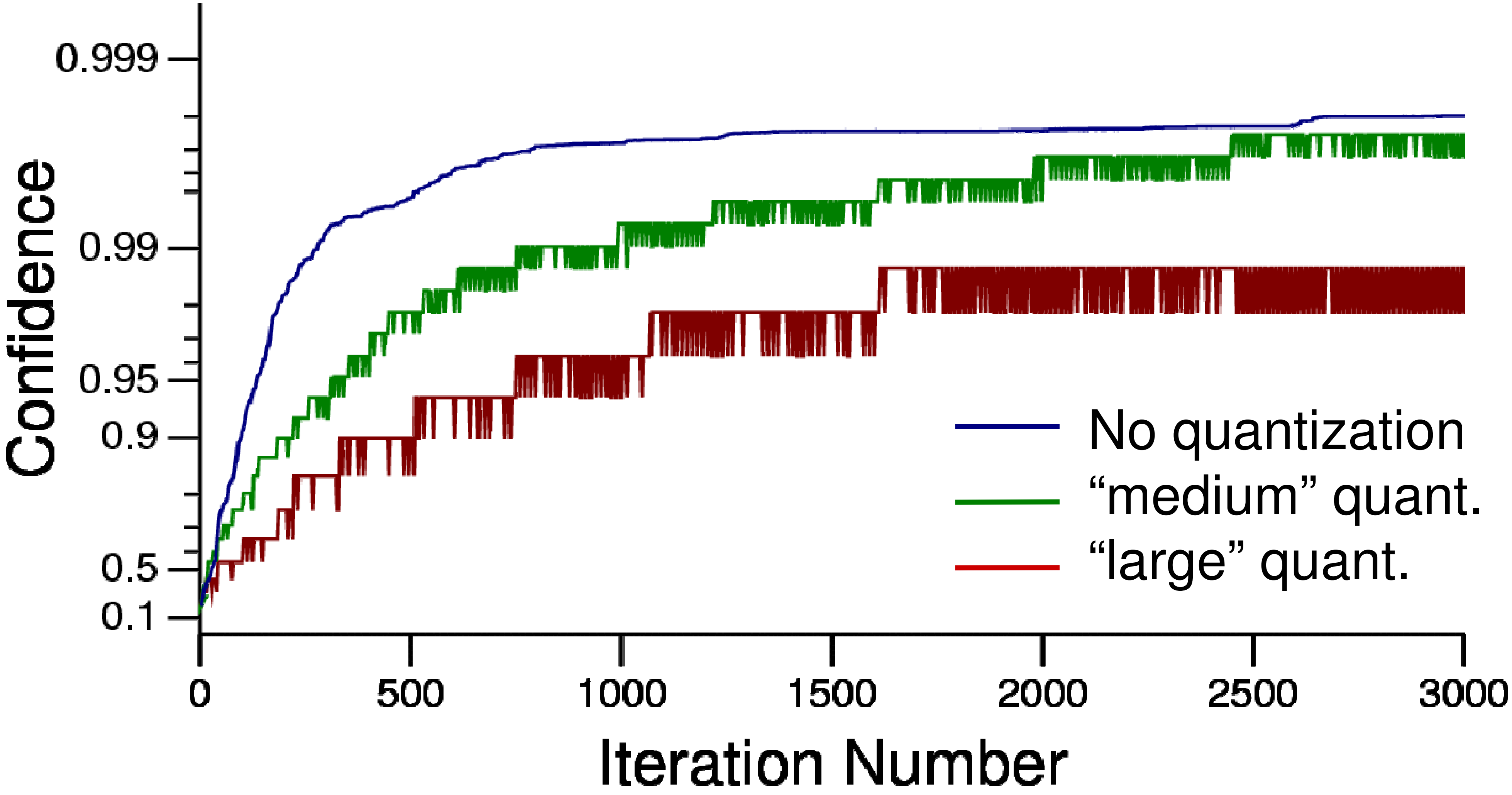
Q

OQ

RN

EF_k

Results: modified “hill-climbing”

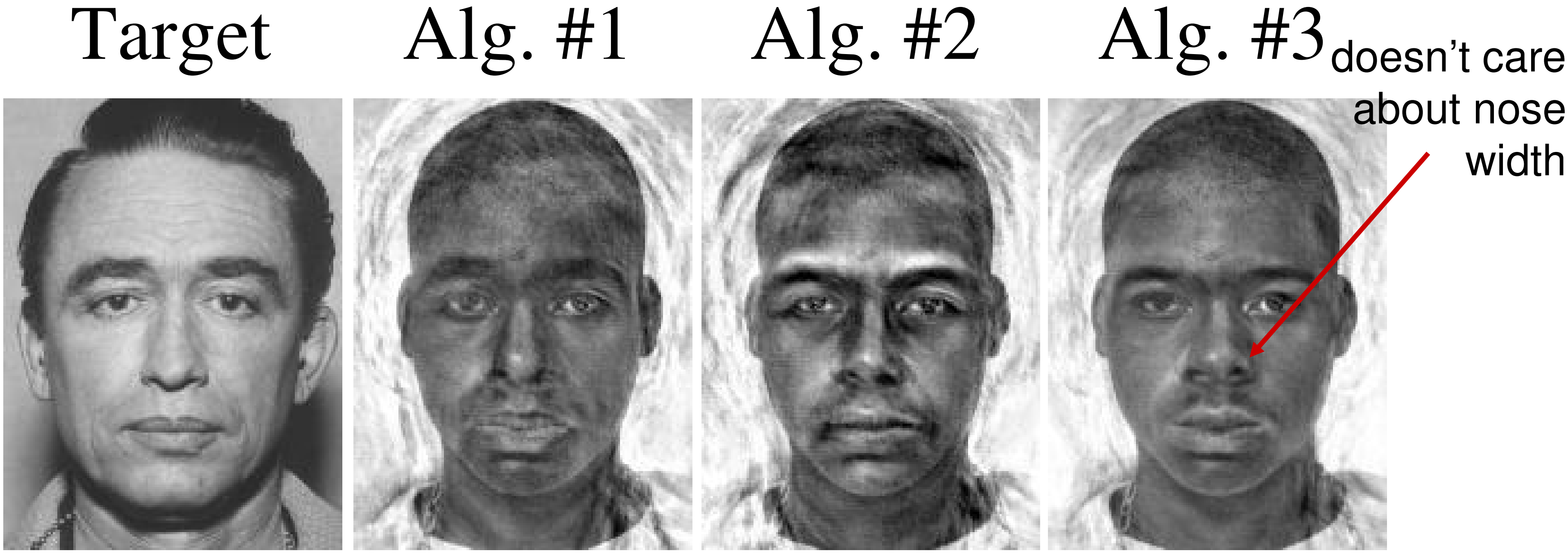


Implications: image regeneration

1. Privacy Implications
   - ICAO passport spec. has templates encoded with public keys in contactless chip
   - ILO seafarer’s ID has fingerprint template in 2D barcode on document

Implications: image regeneration

2. Reverse engineer algorithm
   - Regenerated images tell you what the algorithm ‘really’ considers important



doesn’t care about nose width

Implications: image regeneration

3. Crack biometric encryption

Biometric encryption seeks to embed a key into the template. Only a valid image will decrypt the key

□ Since images vary

Enrolled image + Δ => release key

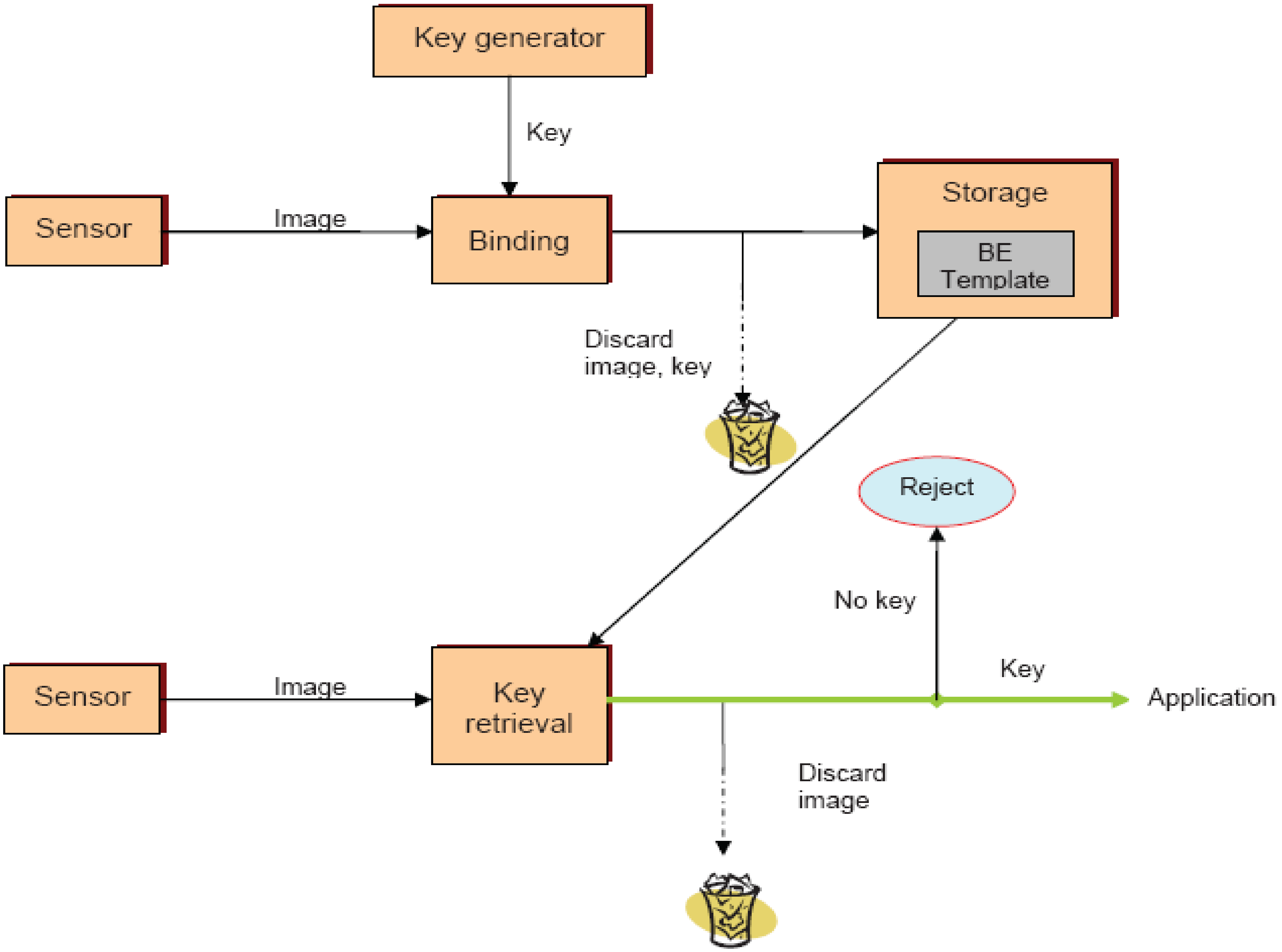
□ However

Enrolled image + Δ + ε => no release

If we can get a measure of how close we are, they we can get a *match score*

Biometric Encryption

- Recent paper by Ontario Information and Privacy Commissioner
  - “Biometric Encryption: A Positive-Sum Technology that Achieves Strong Authentication, Security AND Privacy”
  - A. Cavoukian, A. Stoianov



From: http://www.ipc.on.ca/images/Resources/up-1bio_encryp.pdf

My concern:

- Biometric Encryption (and biometric cryptographic schemes in general) only offer benefits if they are cryptographically secure. If they are not cryptographically secure, then they offer no benefit at all.

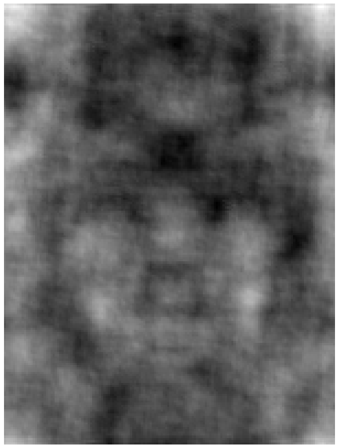
Biometric encryption (Soutar, 1998)

- Average pre-aligned enrolled image (f_0)

- Calculate template from Wiener filter

$$H_0 = F^* R_0^* / (F^* F + N^2)$$

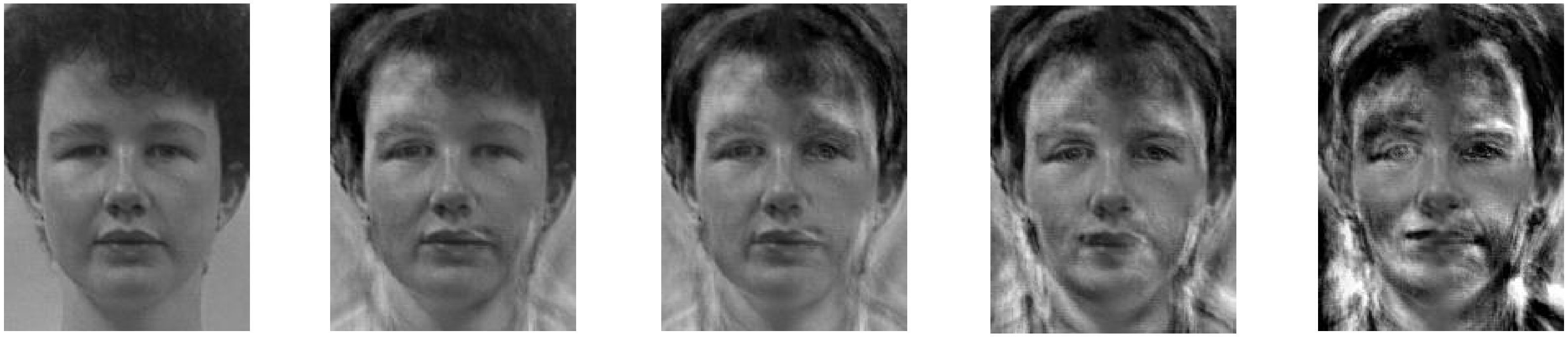
where R_0 has phase ±π/2, ampl = 1

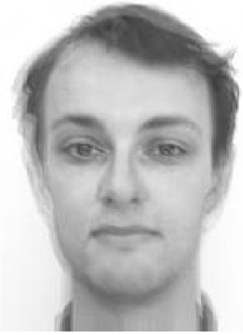


- Each bit of secret is linked to several bits of H_0 with same phase

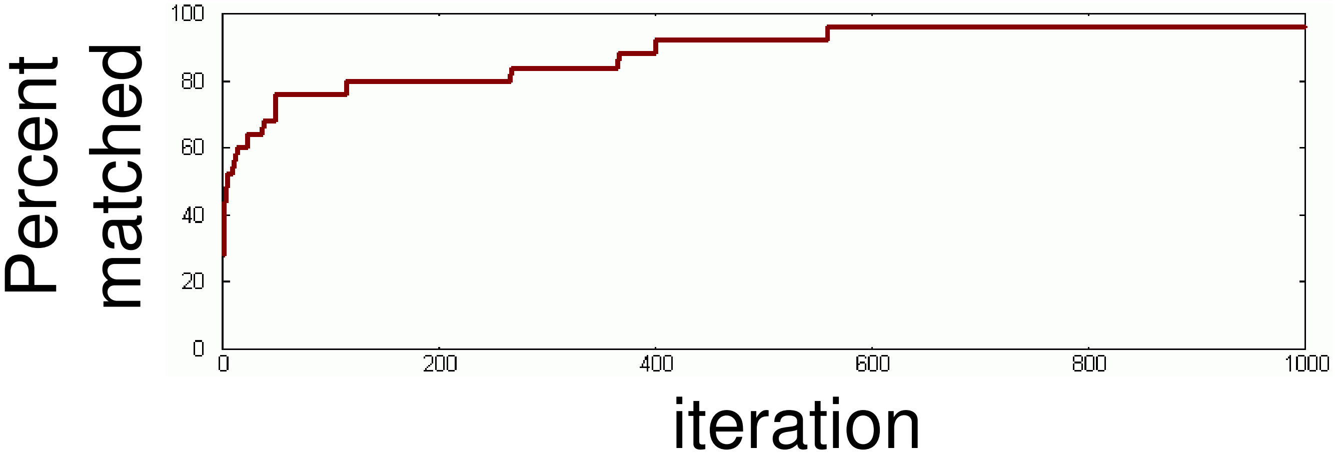
Crack biometric encryption

- Construct *match-score* from number of matching elements in *link table*
- Use quantized template reconstructor

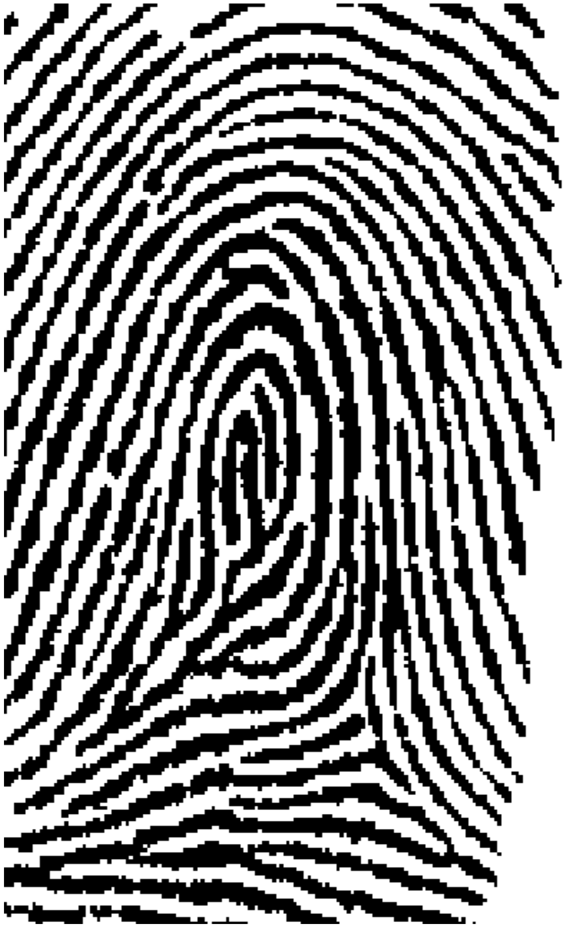




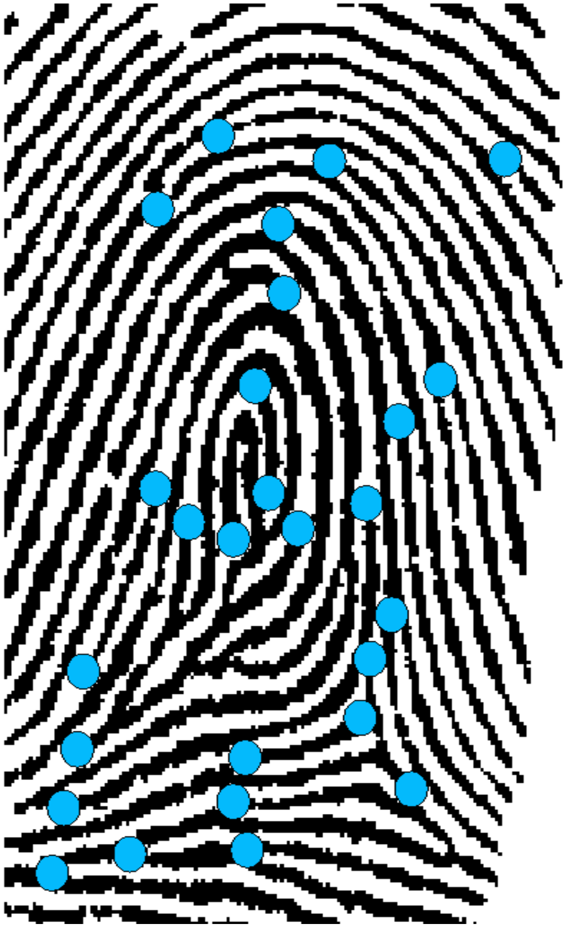
enrolled



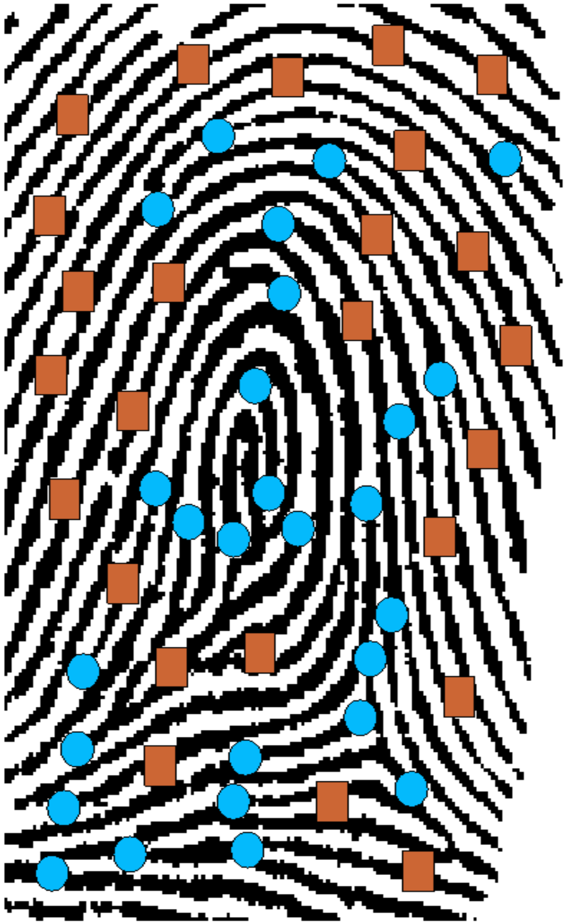
Fuzzy Vaults for fingerprints
(Clancy, 2003)



Raw Fingerprint



With minutiae



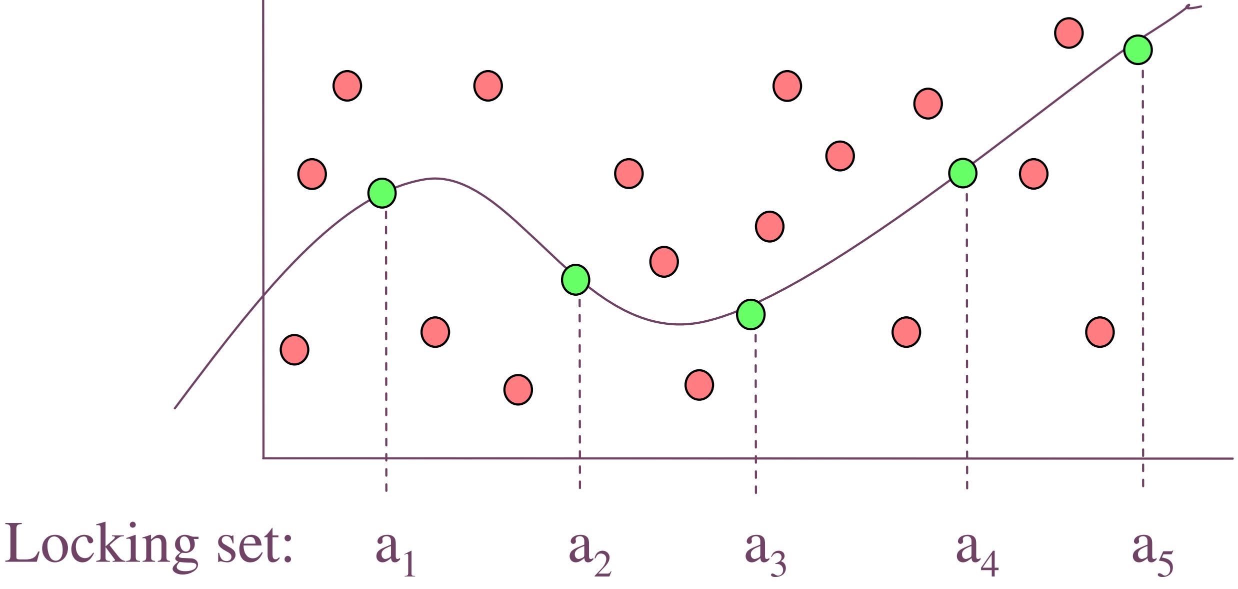
With added “chaff”

Fuzzy Vault encryption

- Encode key $(k_{1,}k_{2,}k_{3,}k_{4)}$ in polynomial coefficients

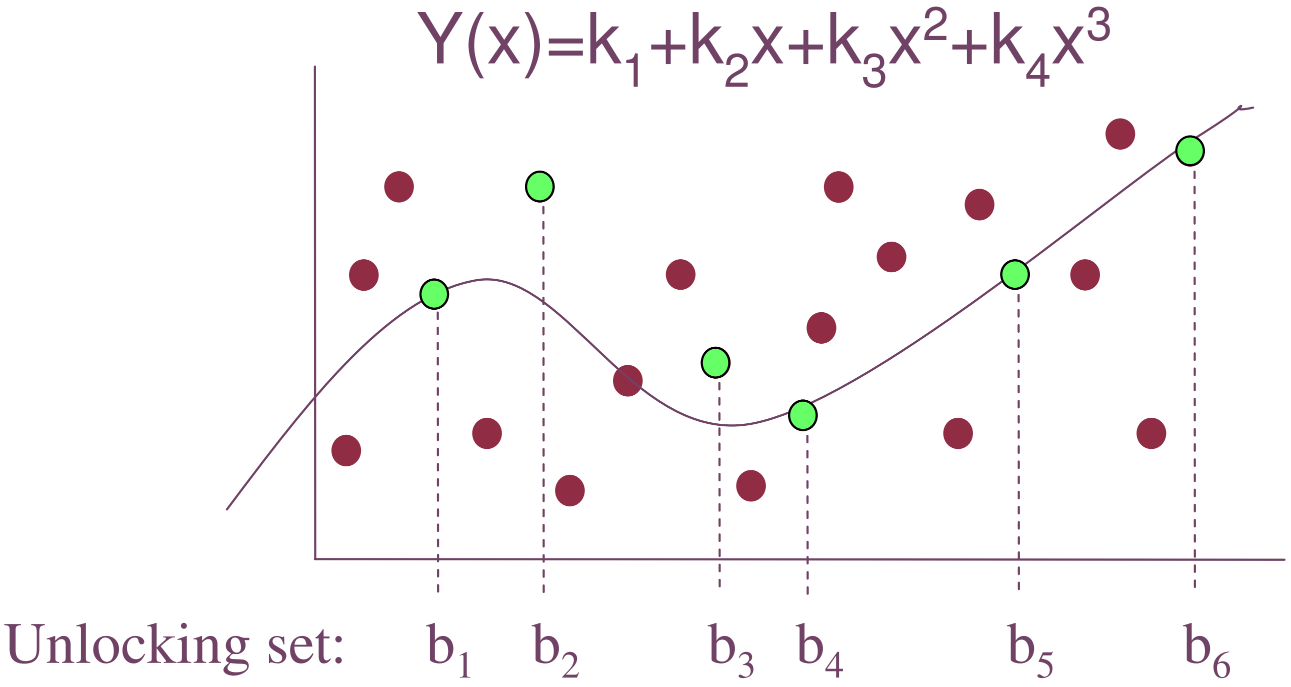
- Template is point co-ordinates

$Y(x)=k_1+k_2x+k_3x^2+k_4x^{3`}$



Fuzzy Vault key-release

- Find polynomial coefficients which best fit to the identified points
- A few wrong points are OK



Collusion Attack

- Users’ fingerprints may be associated with many vaults.
  - Ex: In the smart card implementation, users will likely carry multiple smart cards associated with different companies, each locked with the same fingerprint.

- Is Fuzzy Vault secure when the same fingerprint is used to lock multiple vaults?

Collusion Attack

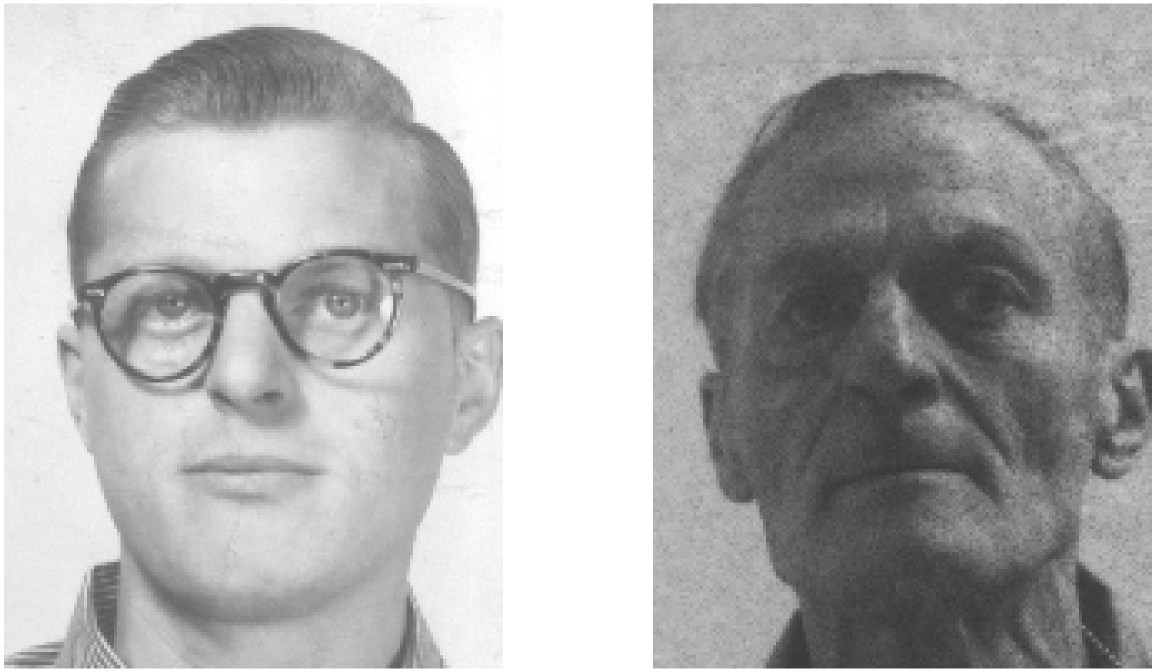
- Multiple vaults with same key, $A_i = A$

V_1			3		**5**		8	9			**12**				**18**			24	**26**		35	**36**	38	49
V_2	1			4	**5**					11	**12**		15	16	**18**			25	**26**			**36**	45	
V_3					**5**	7			10		**12**				**18**	19	21	23	**26**	28		**36**	39	
V_4		2			**5**			9	10	11	**12**	14	15		**18**				**26**	29		**36**		

Summary

- Almost everyone is inventing schemes; very few are breaking them.
- However,
  Anyone can invent a security system that he himself cannot break.
    - B. Schneier.

Face Recognition: Human vs. Automatic Performance



same person?

Same person? **Yes**

- I have just demonstrated a massively parallel face recognition computer

- Of all biometric modalities, automatic face recognition is most often compared to human performance

Choice of images

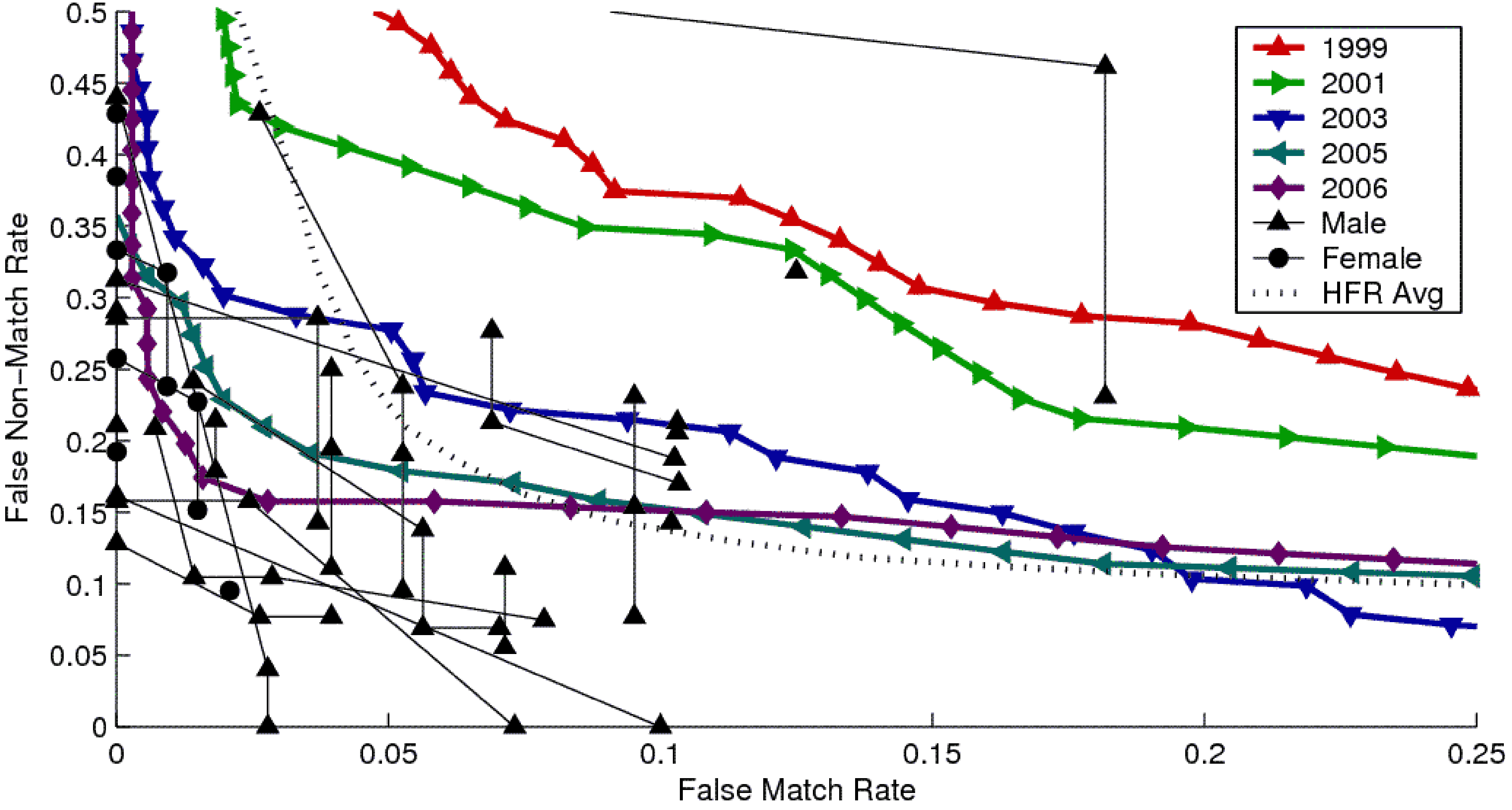
- *Goldilocks* problem:
  Too easy test -> all score 100%
  Too hard test -> all score 0%

- Database used: *NIST Mugshot*
  - Large age changes between captures
  - Population that tends to change appearance

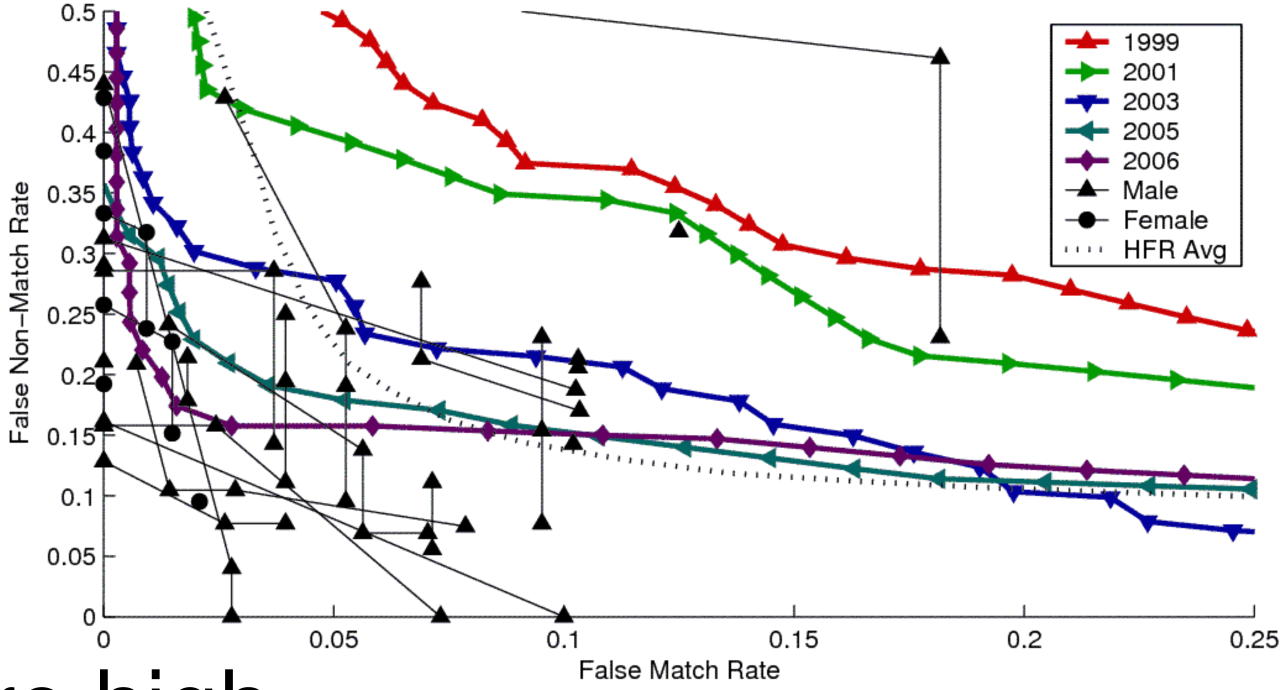
Analysis

- Human results
  - Post-processed to choose optimal “threshold” for them
  - An operating point FMR/FNMR calculated

- Software results
  - Same images presented to FR software (worked with 15 packages – 7 vendors)
  - ROC calculated


0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
False Non−Match Rate
0
0.05
0.1
0.15
0.2
0.25
False Match Rate
1999
2001
2003
2005
2006
Male
Female
HFR Avg

Results



- Error rates are high
- Significant improvement in SW 1999-2006
- Most recent algs outperform about half of people
- No significant difference male/female

information content of a biometric measurement?

Or

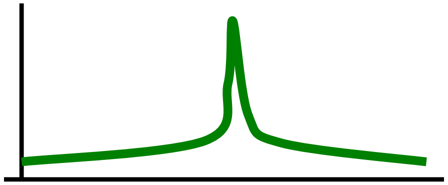
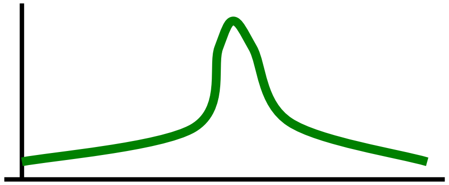
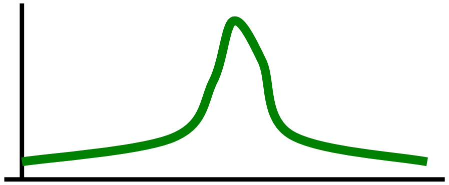
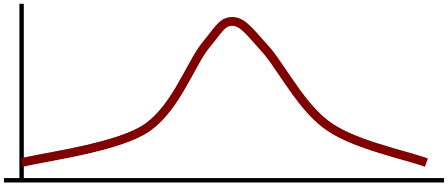
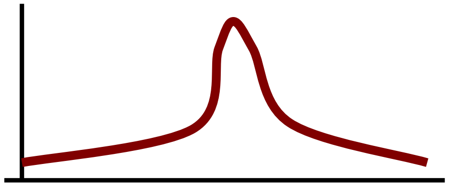
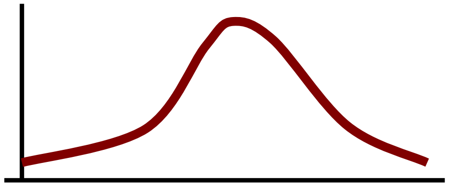
- How much do we learn (about identity) from a biometric image

Or

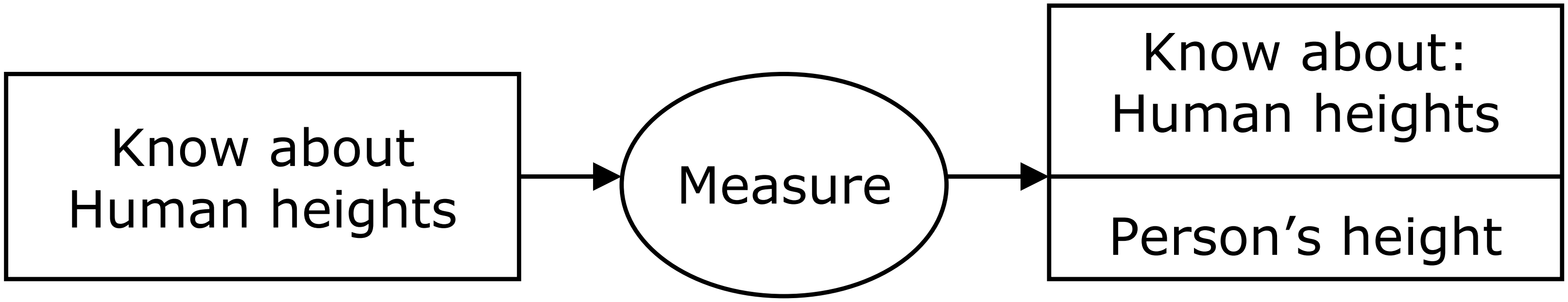
- How much privacy do we loose on releasing a biometric image

Example: measure *Height*

- Measure #1 (at doctor’s office, ie. accurate)
- Measure #2 (via telescope, ie. inaccuate)

Measurement Variability (device errors)	Feature Variability (high heels, carry backpack)	Overall Distribution
		
		

Example: measure *Height*



- How much information learned?

	Average (5½’ tall)	Tall (7½’ tall)
Measure #1	Low	Quite a lot
Measure #2	Almost zero	Low

Proposed measure: *relative entropy $D(p||q)$*

- Given biometric feature vector $\mathbf{x}$
- Distributions
  - □ intra-person distribution, $p(\mathbf{x})$
  - □ inter-person distribution, $q(\mathbf{x})$
- $D(p||q)$ measures inefficiency of assuming q when true distribution is p

Or,

- $D(p||q)$ measures extra information in p than q

Applications: *biometric*

- Meta algorithm
  - Evaluate a new biometric feature
- Biometric Performance limits
  - Template size limits
  - Inherent match performance limits
- Feasibility of Biometric Encryption
  - Limits to Key Length

Applications: *abstract*

- Quantify privacy
  - What is the privacy risk due to the release of certain information?
  - What is the privacy gain in obscuring faces?
- Uniqueness of biometrics
  - Approach to address: “Are faces / fingerprints / irises unique?”

Conclusions

- Approach to measuring information content of a biometric system
- Relative Entropy is appropriate measure
- Help explain *legal, social, performance* issues

Biometrics in Canada (Gov't)

- Passports
- Immigration
- Customs
- Defence
- Natural Resources
- Public Safety

Privacy issues

- There are widespread privacy concerns about biometrics.
- This is not really a biometrics issue. Companies/Governments have proved themselves irresponsible with personal data. Now people are stonewalling.
- Have you ever checked your credit record?
  Mine is about 25% inaccurate.

Epilogue: ***biometrics' future***?

Operator: "Thank you for calling Pizza Hut."

Customer: “Two All-Meat Special..."

Operator: "Thank you, Mr. Smith. Your voice print identifies you with National ID Number: 6102049998"

Customer: (Sighs) "Oh, well, I'd like to order a couple of your All-Meat Special pizzas..."

Operator: "I don't think that's a good idea, sir."

Customer: "Whaddya mean?"

Operator: "Sir, your medical records indicate that you've got very high blood pressure and cholesterol. Your Health Care provider won't allow such an unhealthy choice."

Customer: "Darn. What do you recommend, then?"

Epilogue:

Operator: "You might try our low-fat Soybean Yogurt Pizza. I'm sure you'll like it"

Customer: "What makes you think I'd like something like that?"

Operator: "Well, you checked out 'Gourmet Soybean Recipes' from your local library last week, sir."

Customer: “OK, lemme give you my credit card number."

Operator: "I'm sorry sir, but I'm afraid you'll have to pay in cash. Your credit card balance is over its limit."

Customer: "@#%/$@&?#!"

Operator: "I'd advise watching your language, sir. You've already got a July 2006 conviction for cussing … "

Biometrics & Authentication Technologies: security issues

Andy Adler

Systems and Computer Engineering

Carleton University, Ottawa