

Security and privacy issues in biometric systems

Andy Adler

School of Information Technology and Engineering

University of Ottawa

Newsflash! Biometrics

- eight fingerprints and face required to get new US Visa (US VISIT)
- New ICAO passport standard requires biometric data in document
- UK will issue biometric based ID card
- Sea-farer's ID card will incorporate two fingerprints

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:

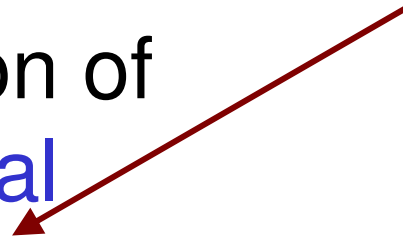
1. Verification

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

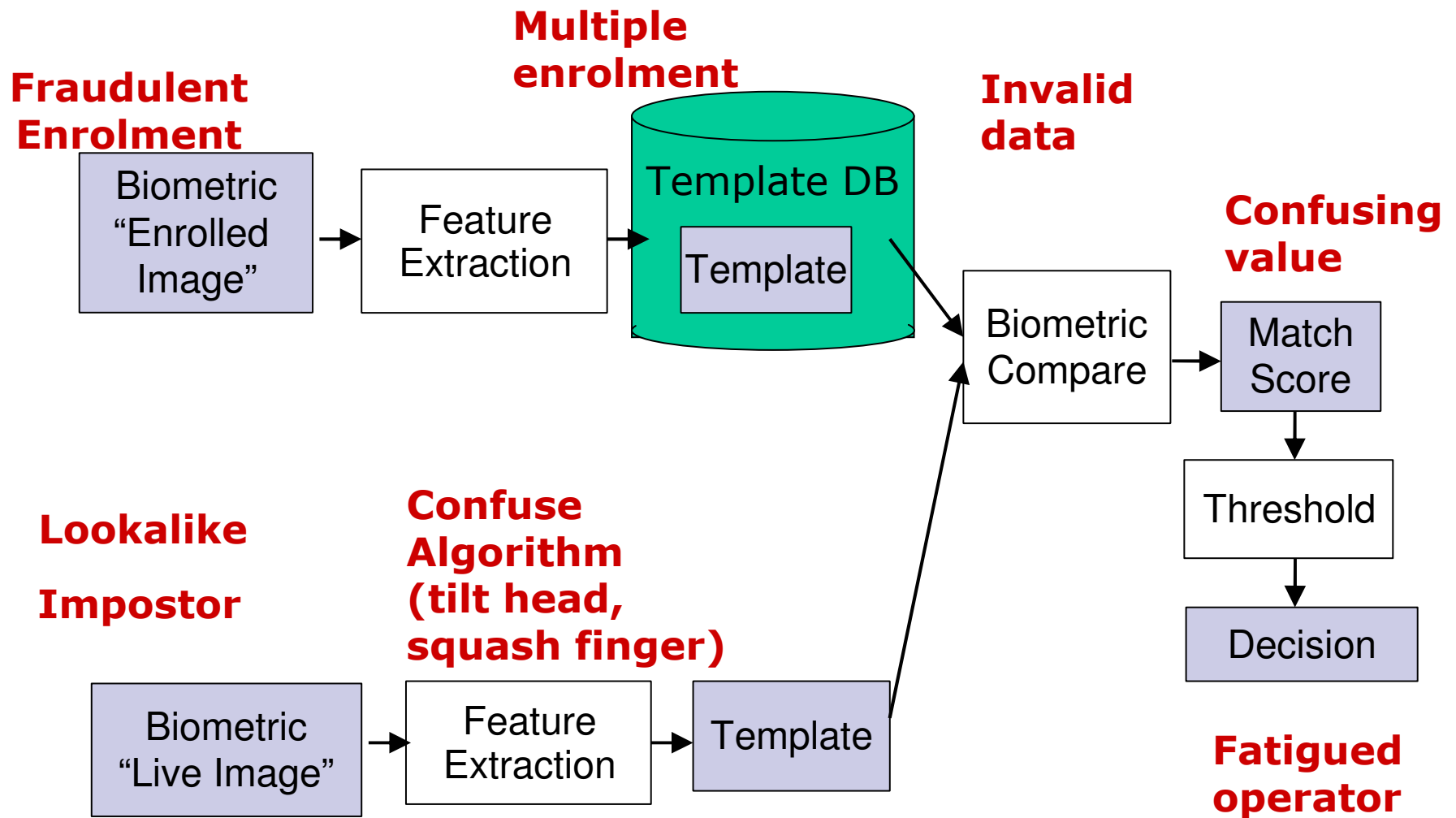
Automatic identification of an individual based on behavioural or **physiological** characteristics

Physiological Biometrics

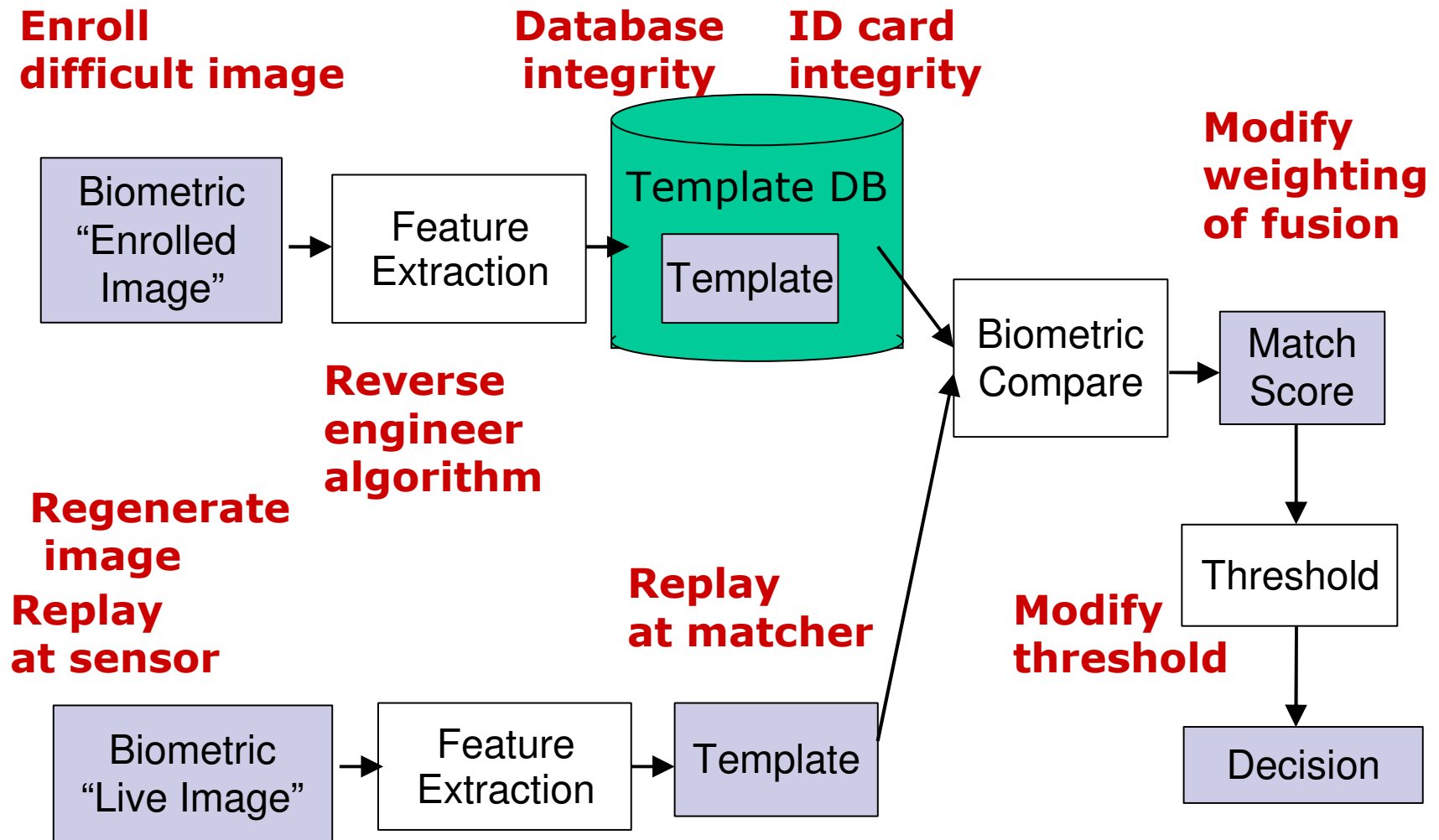
- Fingerprint
- Face
- Iris
- Retina
- Hand Geometry
- Dental shape
- DNA

...

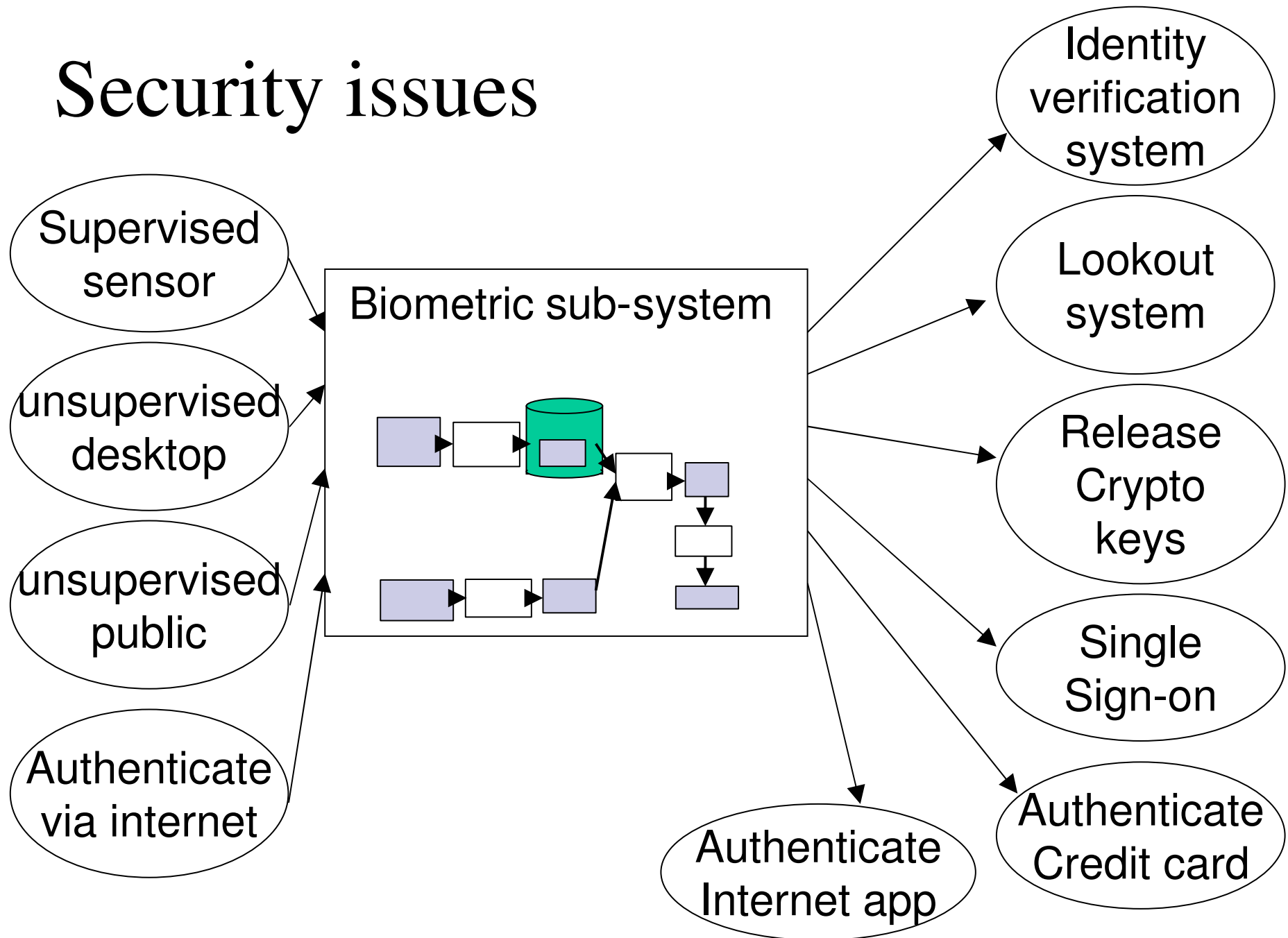
Security issues



Security issues



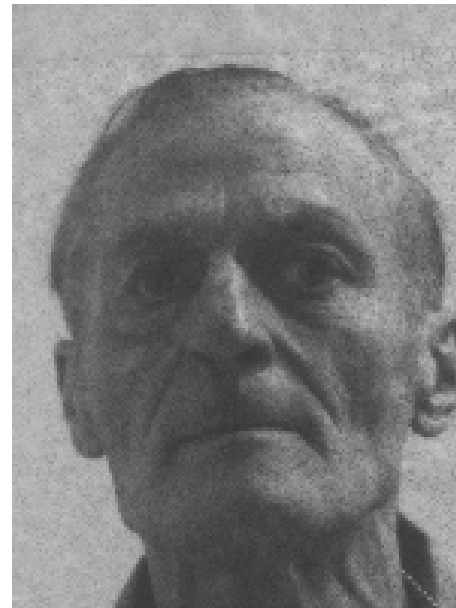
Security issues



Security issues

- Biometrics only provides identity
 - Need to be coupled to a system
- These systems are also vulnerable to all of the traditional security threats
 - as well as all sorts of new ones
 - and interactions between old and new ones

Face Recognition: Human vs. Automatic Performance



same person?

Same person? **Yes**

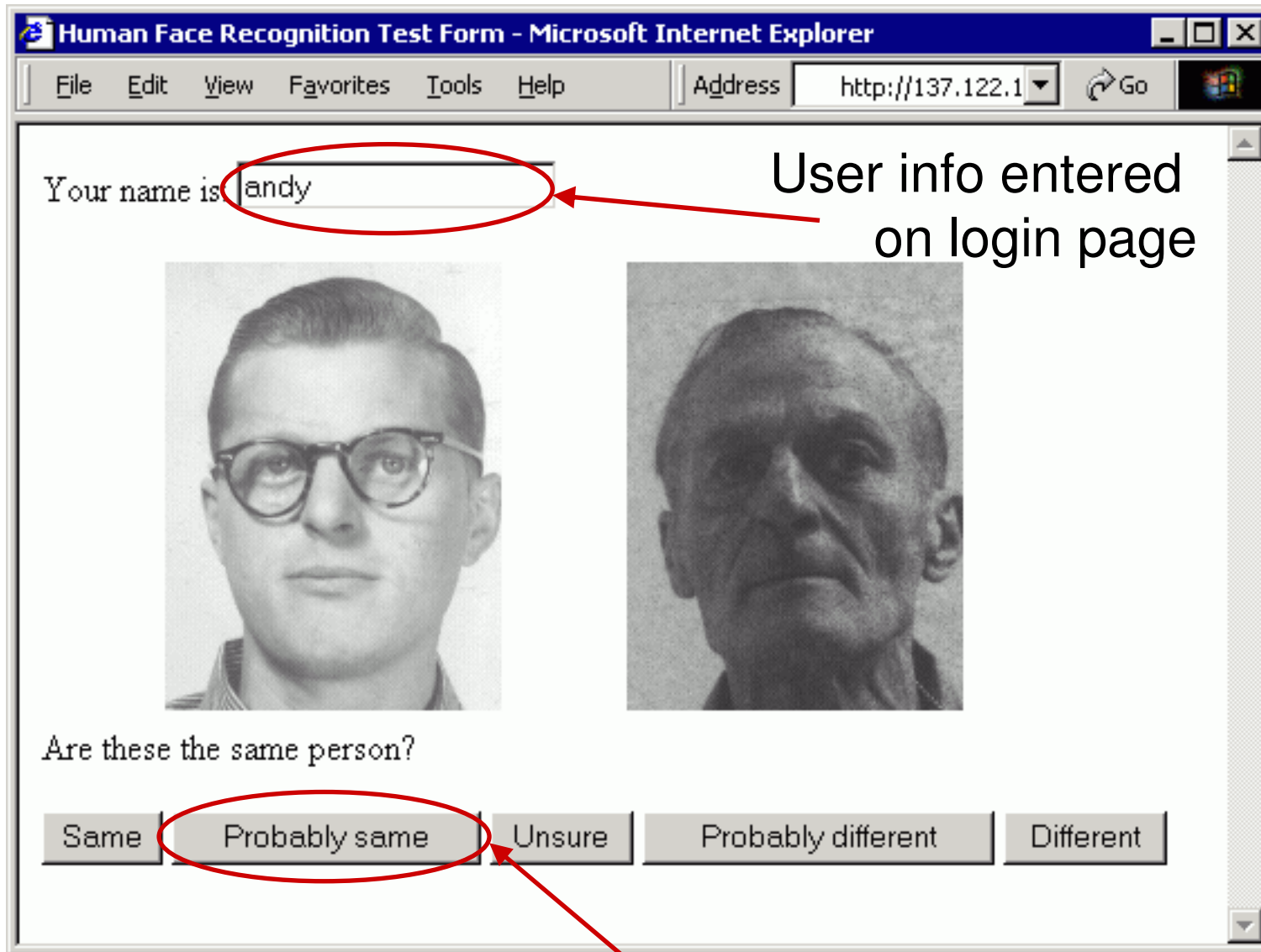
- I have just demonstrated a massively parallel face recognition computer
- Of all biometric modalities, automatic face recognition is most often compared to human performance
- Surprisingly little work has been done to quantify these levels of performance

Other studies

- Kemp et al. (1997) analyzed supermarket cashiers identifying shoppers credit card photos
 - Results show poor performance.
- Chang Hong et al. (2003) analyzed people matching poor-quality video to high-quality photographs
 - Results show high performance.
- Burton et al. (1998,2001) compared PCA based and graph-matching algorithms against human ratings
 - Primarily to elucidate aspects of human memory not to evaluate algorithms

Test Design

Participants	Employees of 3M Security Systems Division (then AiT) in Ottawa, Canada
Participation	Voluntary – announcement at company weekly meeting
Participant demographics	16 Male, 5 Female, ages 20-40, predominantly Caucasian
Test format	Web based: subject participated from their office
Instructions	Focus on accurate results



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

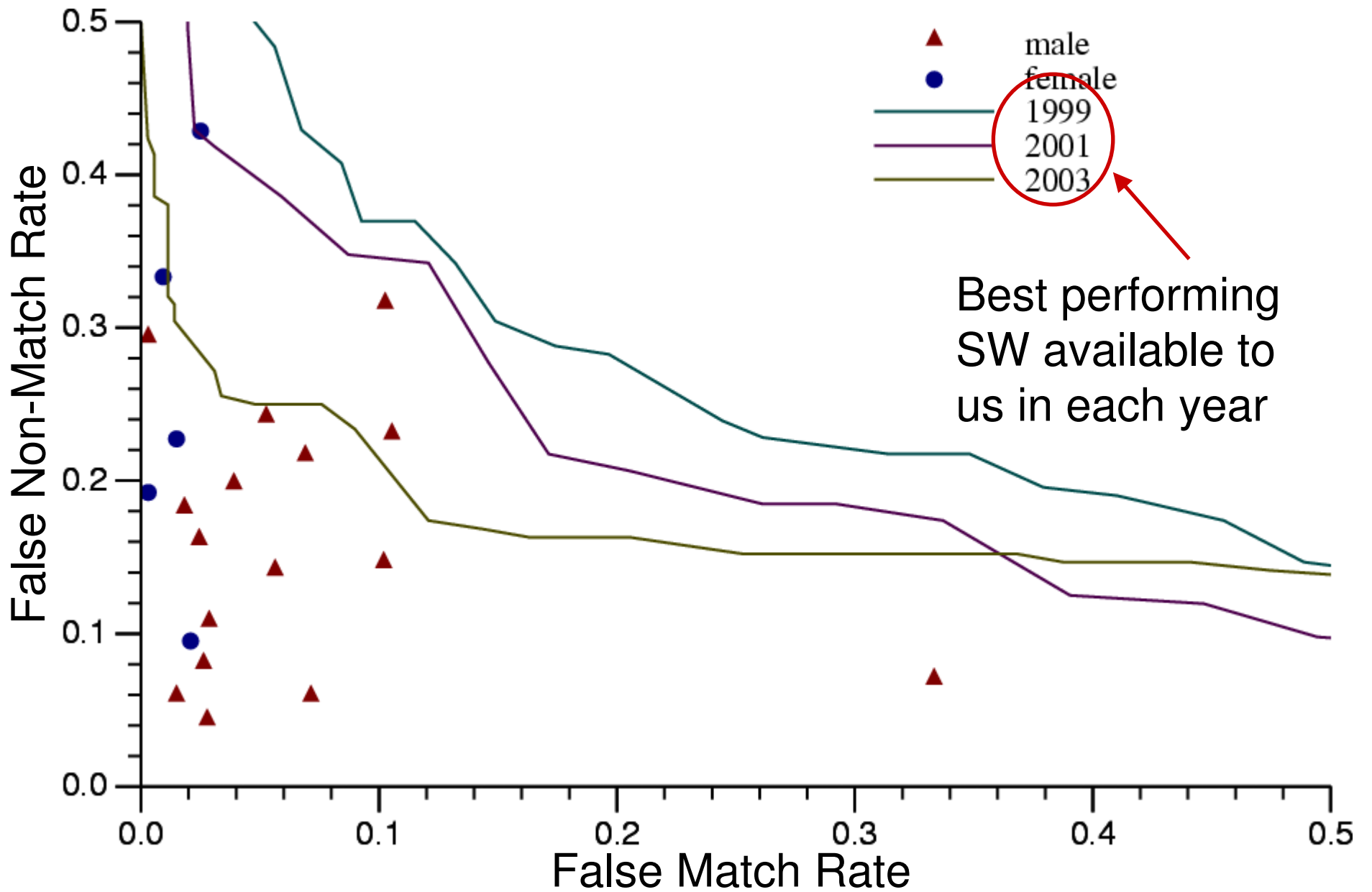
Analysis

■ Human results

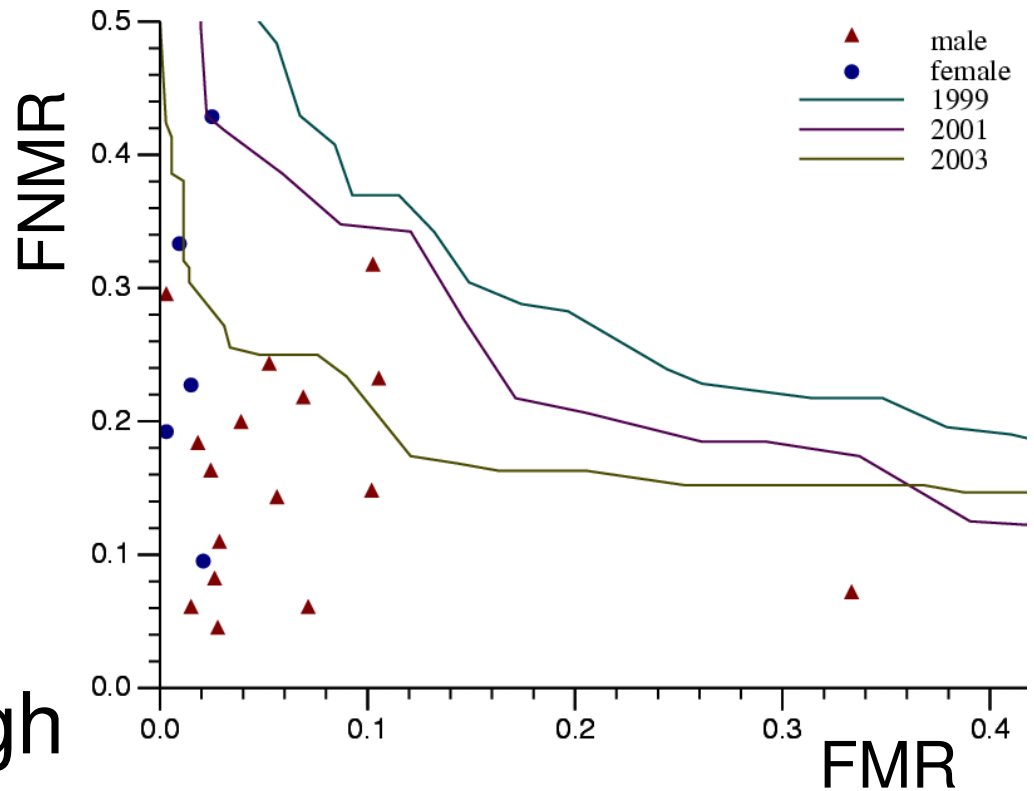
- Post-processed to choose optimal “threshold” for them
- An operating point FMR/FNMR calculated

■ Software results

- Same images presented to FR software (worked with 13 packages and versions)
- ROC curve calculated



Results



- Error rates are high
- Significant improvement in SW 1999-2003
- Most motivated, attentive humans can outperform face recognition software
- No significant difference male/female

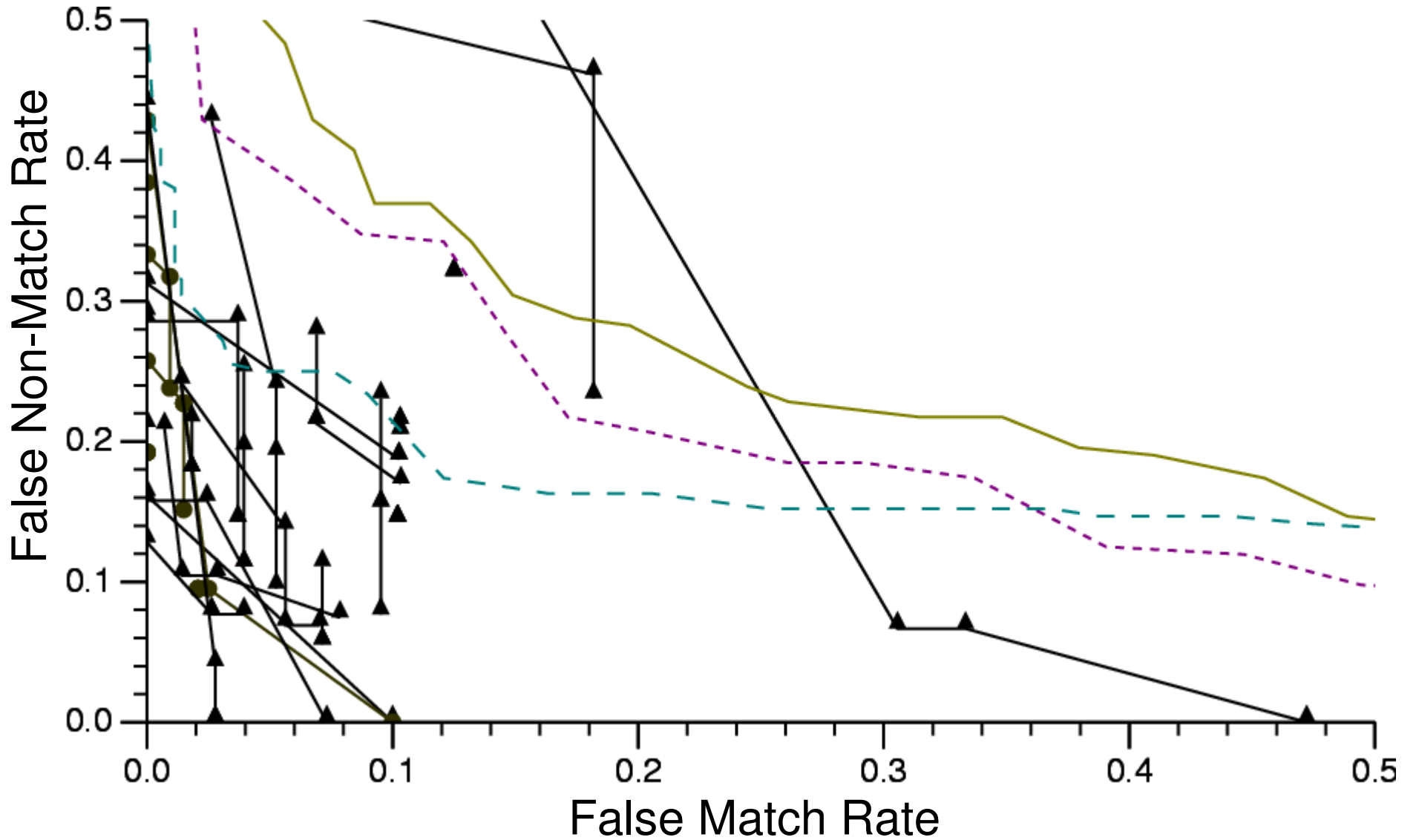
Conclusion

- Currently, most people are able to significantly outperform FR software on difficult data sets
 - Unlimited time (took 10 s avg.)
 - Motivated staff
- Thus, we have perhaps measured some kind of *upper limit* for HFR

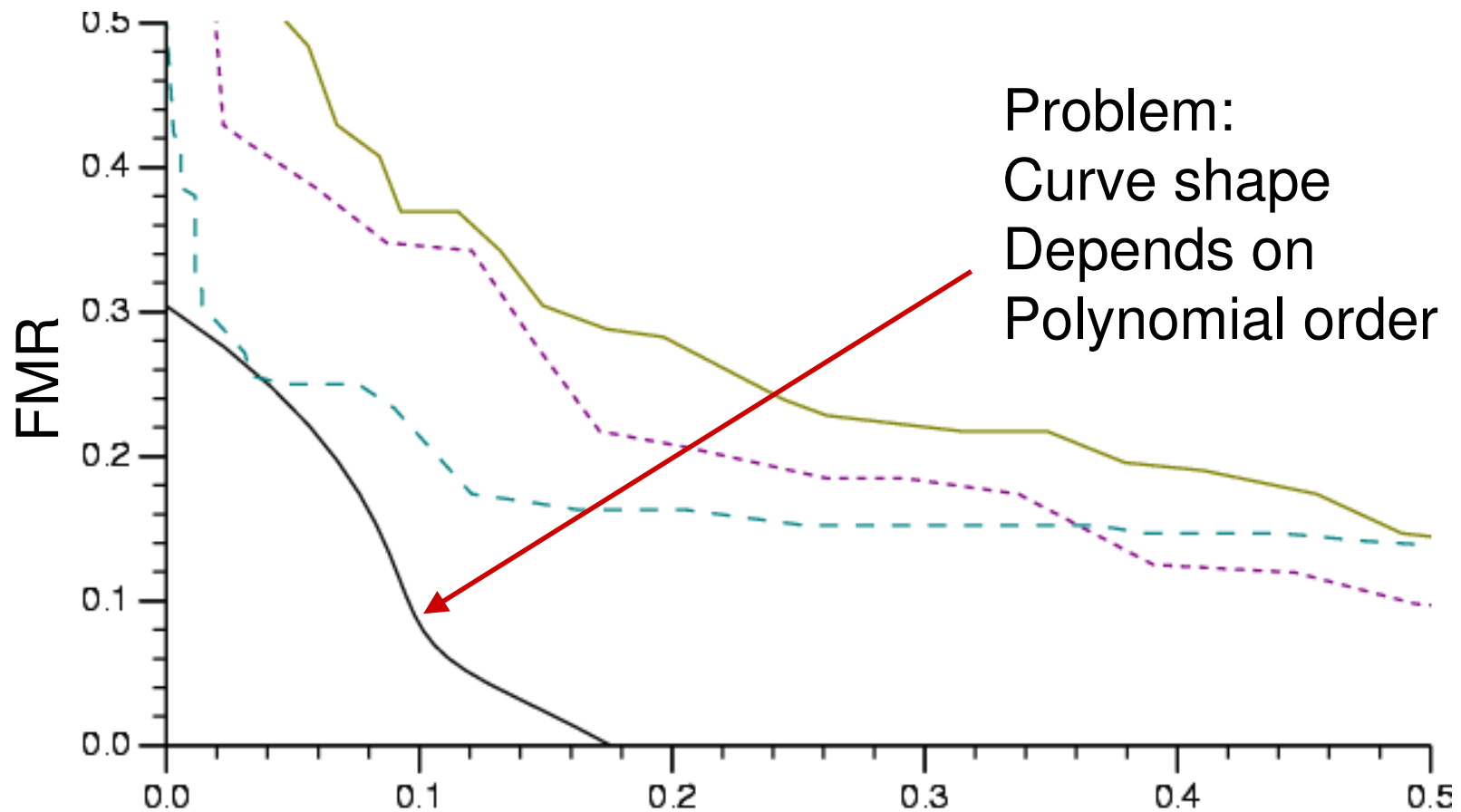
Unanswered questions

1. How do humans perform as familiarity increases?
2. What is the effect of motivation, routine and boredom?
3. Do experts outperform untrained recognizers?
4. What distinguishes good recognizers from poor ones?
5. What if a live subject is available?

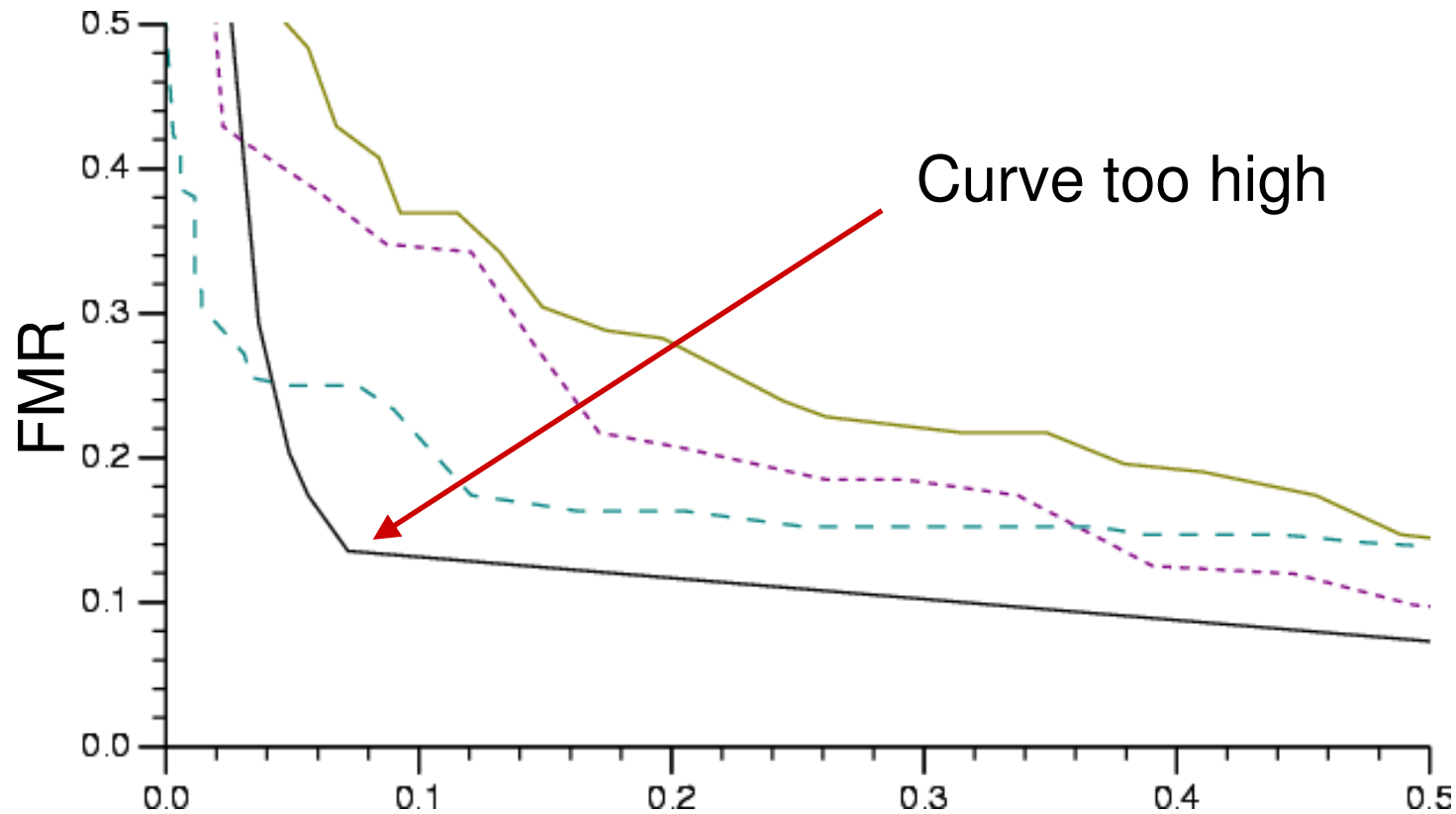
Questions: Average HFR?



Idea #1: Convert to polar coords: fit to polynomial in (r, θ)



Idea #2: Collect all match score data for humans and calculate average



Problem: Match score values mean different things for different people; can't legitimately take ensemble

Security of biometric templates

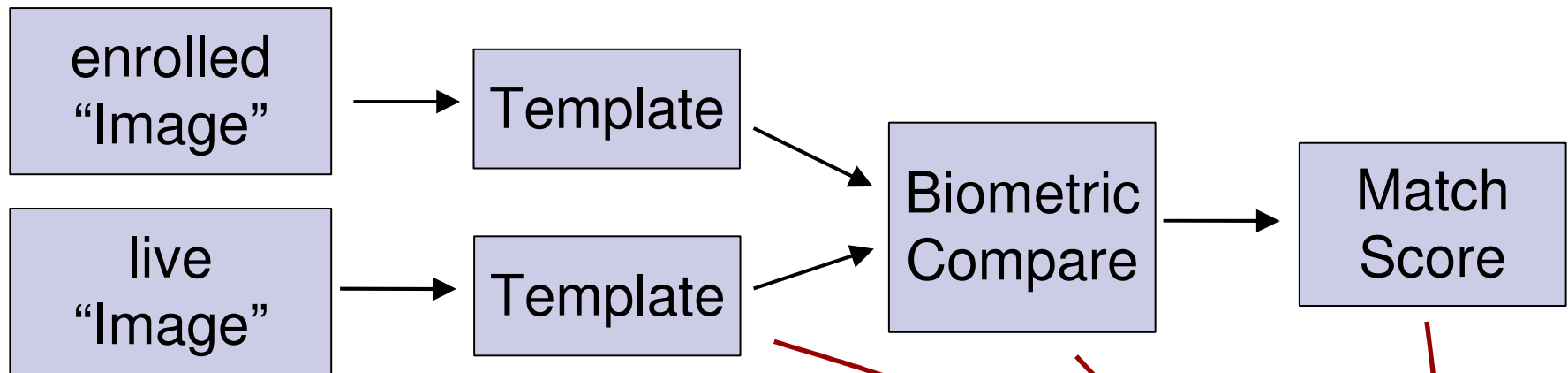
Most biometric vendors have claimed its impossible or infeasible to recreate the enrolled image from a template.

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".

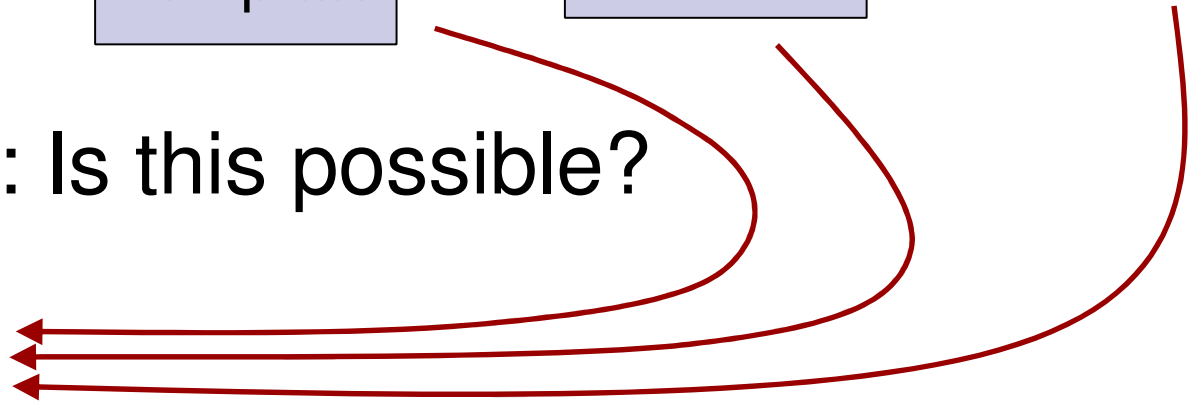
Images can be **regenerated** ...?

■ Typical Biometric processing



■ *Question:* Is this possible?

regenerated
"Image"



Automatic image *regeneration*

Question: is it possible to have generic software to regenerate images from biometric templates?

Answer: Yes

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

Requirement: access to a match scores

“Hill-climbing” Algorithm

Preprocessing:

- *Obtain Local Database (LD) of face images:*
Images are rotated, scaled, cropped
- *Eigenface decomposition of LD:*
 i th eigenimage is represented by EFi .
- *Initial image selection (IM_0):*

“Hill-climbing” Algorithm

Iterative estimate improvement: (for $i \dots$)

- Randomly select eigenimage: EF_k
- Iterate for a range of values c_j :

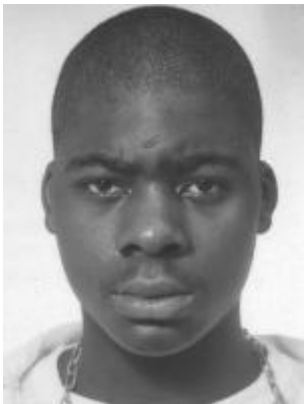









$MS_j = \text{biometric_compare}(IM_k + c_j \times EF_k, IM_{targ})$

- $j_{max} = j$ for which MS_j is maximum
- $IM_{i+1} = IM_i + c_{j,max} \times EF_k$
- Truncate IM_{i+1} to image limits (ie. 0 to 255)

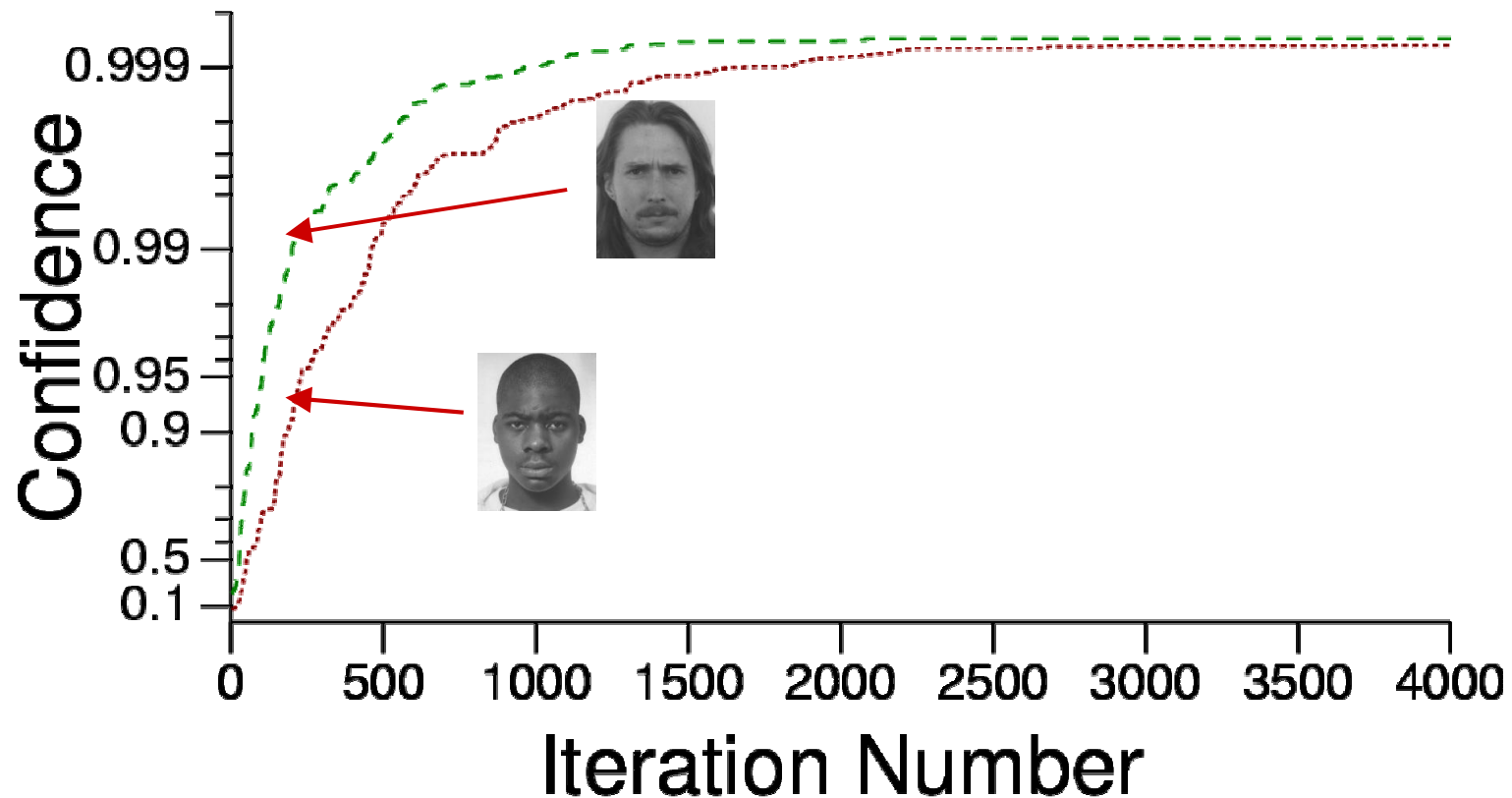
Results

- Tests were performed against three different face recognition algorithms
 - All are recent products by well known commercial vendors of biometric systems.
 - Two of the vendors participated in the 2002 face recognition vendor test
- For all images and all biometric algorithms, the regenerated image compared at over 99.9% confidence

Results

	Initial Image	Iteration 200	Iteration 600	Iteration 4000	Target Image
A					
B					

Results: Confidence vs. iteration

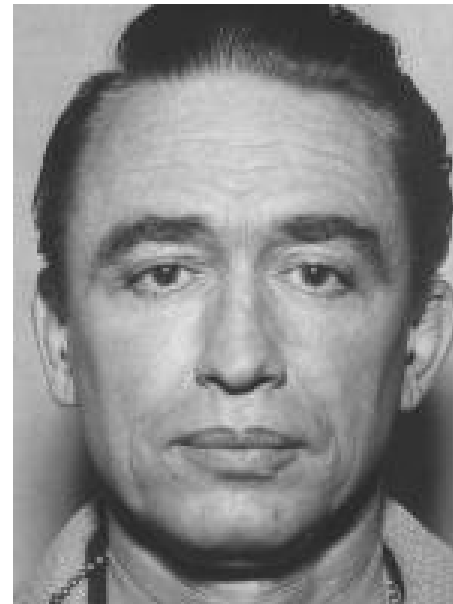


Confidence is the probability of correct verification for a given match score

Improved regenerated image



Average of 10
Best Estimates



Target Image

Extensions to this approach

Recently, this approach has been extended to fingerprint images

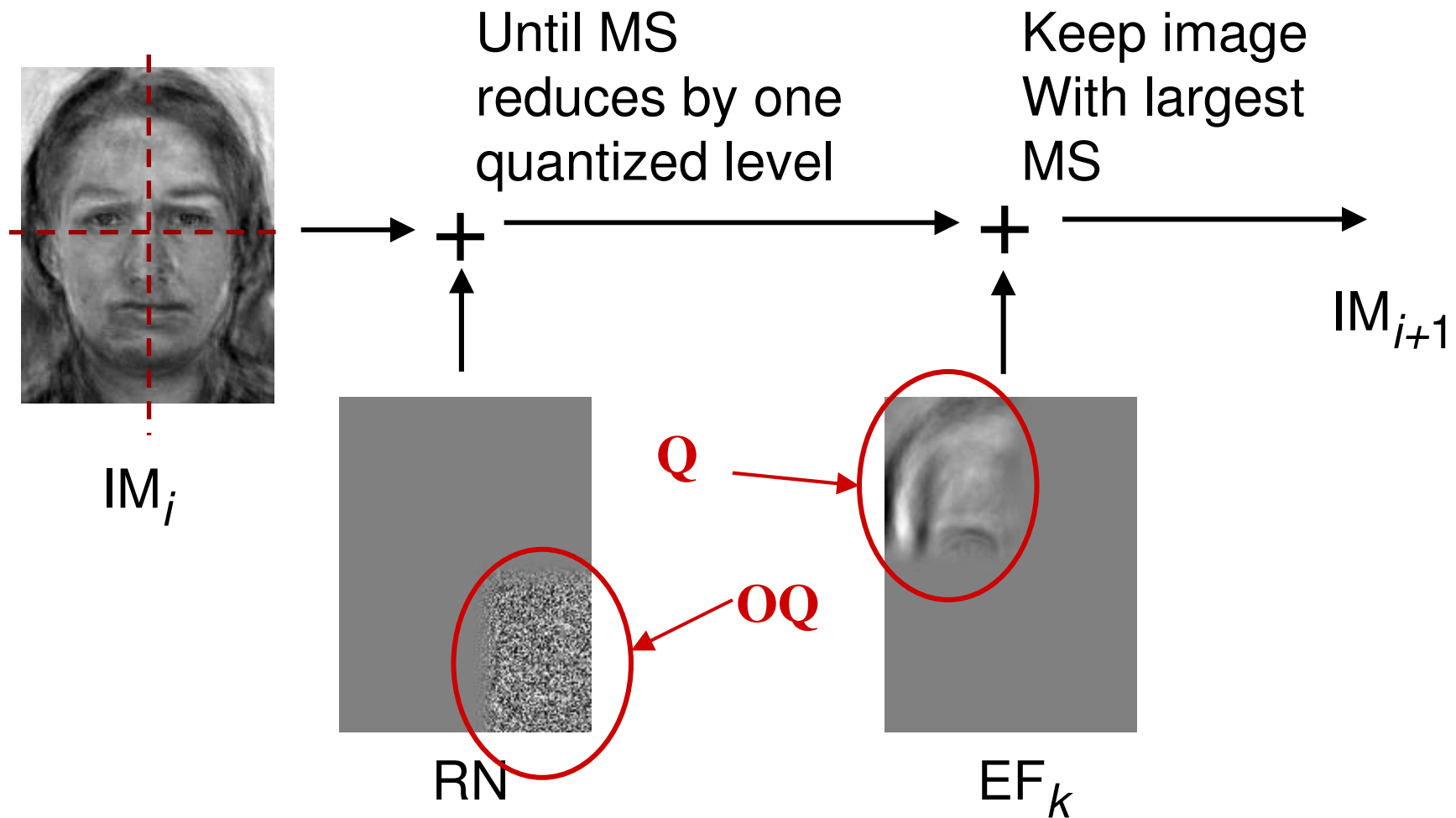
- Uludag developed an approach to modify a collection of minutiae
- Ross has developed a fingerprint image regenerator

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”



Modified “hill-climbing”

Iterative estimate improvement: (for $i \dots$)

- Select eigenimage, EF_k
 - Select quadrant Q . Opposite quadrant is OQ .
 - Generate image RN : noise in OQ and zero elsewhere.
 - Calculate amount of RN to reduce the MS_i by one quantization level.
- New

$MS_i = \text{biometric_compare}(IM_i, IM_{targ})$

$MS_{NI} = \text{biometric_compare}(IM_i + n \times RN, IM_{targ})$

Modified “hill-climbing”

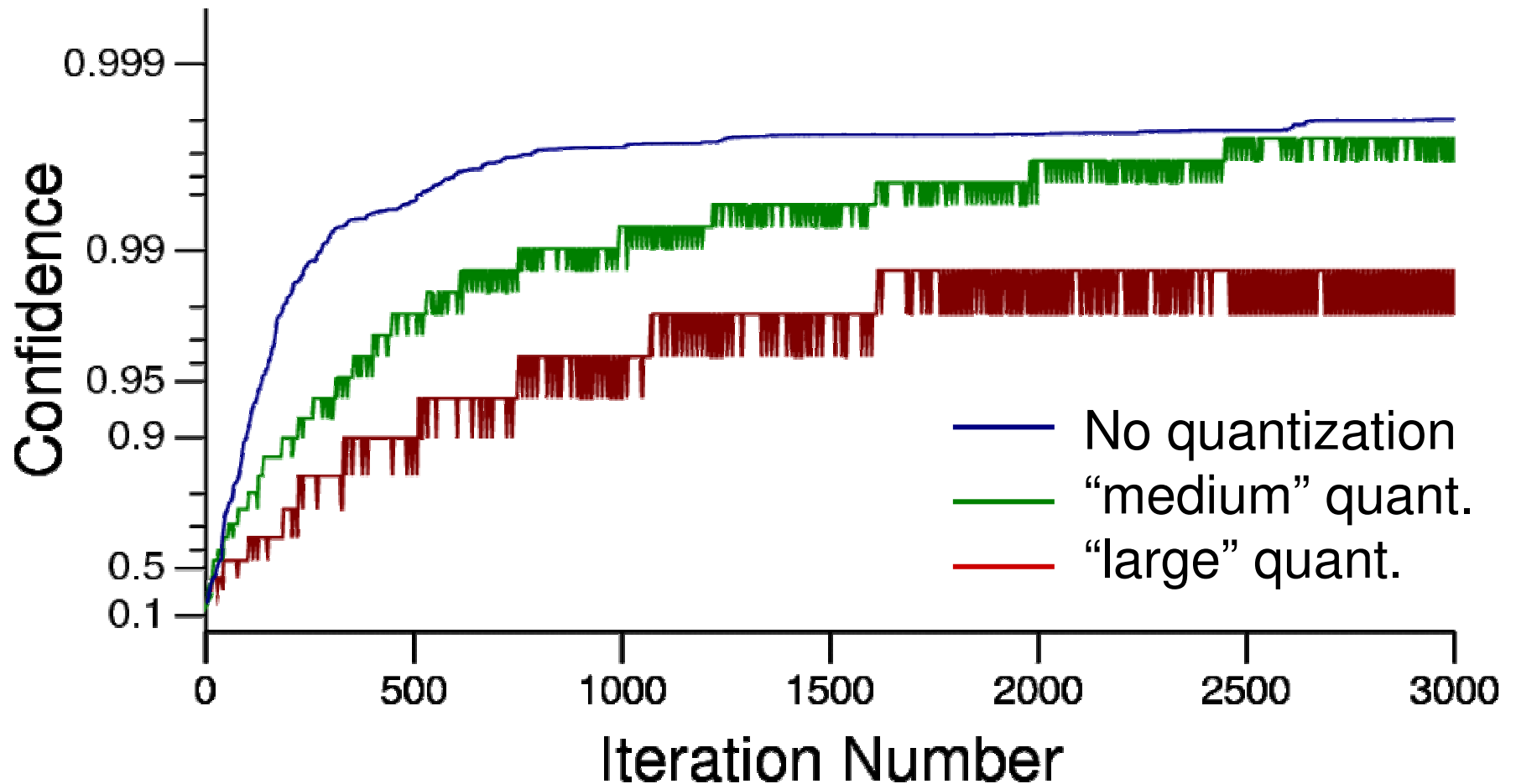
Iterative improvement (continued ...)

- Randomly select: EF_k
- Iterate for a range of c_j using quadrant Q

$MS_j = \text{biometric_compare}(IM_k + c_j \times EF_{k,Q}, IM_{targ})$

- $j_{max} = j$ for which MS_j is maximum
- $IM_{i+1} = IM_i + c_{j,max} \times EF_{k,Q}$
- Truncate IM_{i+1} to image limits (ie. 0 to 255)

Results: modified “hill-climbing”



Modified “hill-climbing”

- Discrete match score means less information is available
 - algorithm takes longer
- Image regeneration works because biometric algorithms “sum up” matching characteristics
 - Changes in quadrants are “independent”
 - We degrade image in one quadrant so that match score is in most informative range

Discussion

Images can be regenerated from biometric templates

- will fool biometric algorithm
- visually reflect important features
- The BioAPI recommendation of using quantized match scores does not provide complete protection

Implications: image regeneration

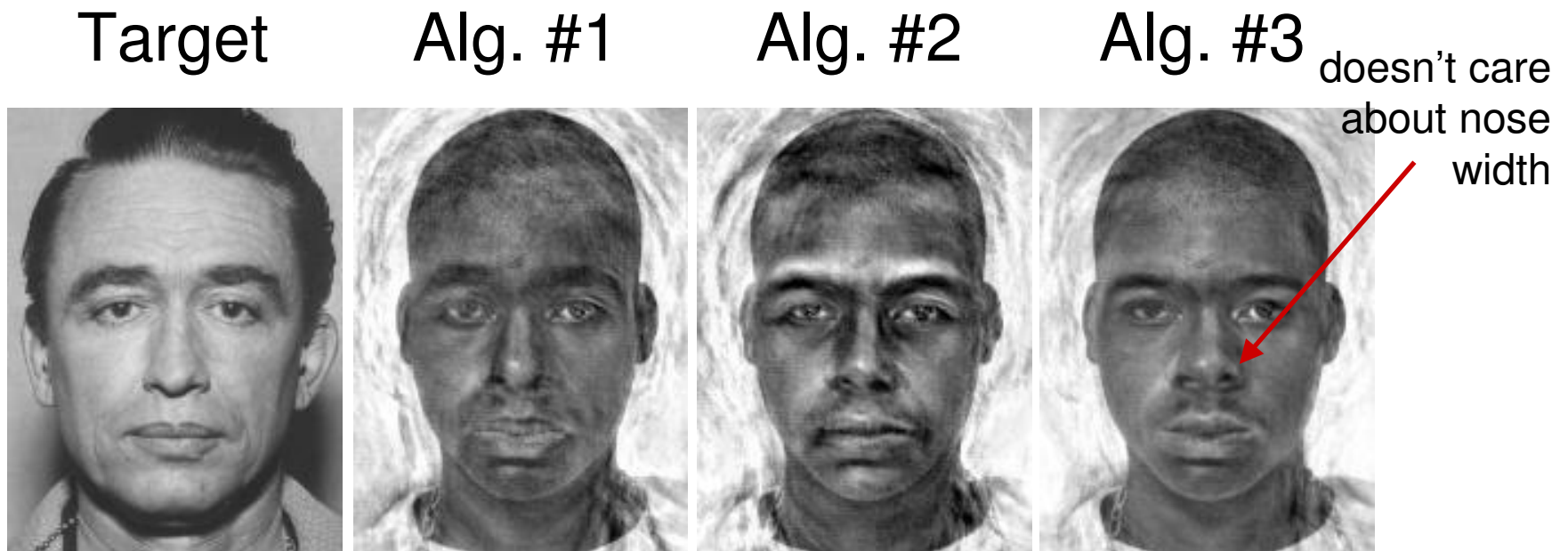
1. Regenerate images for spoofing

- 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



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, then we can get a *match score*

Biometric encryption (Soutar, 1998)

- Average pre-aligned enrolled image (f_0)

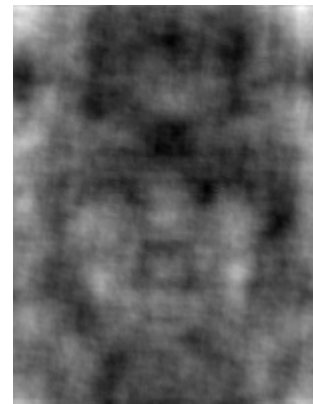


- Calculate template from Wiener filter

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

where R_0 has phase $\pm\pi/2$, ampl = 1

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

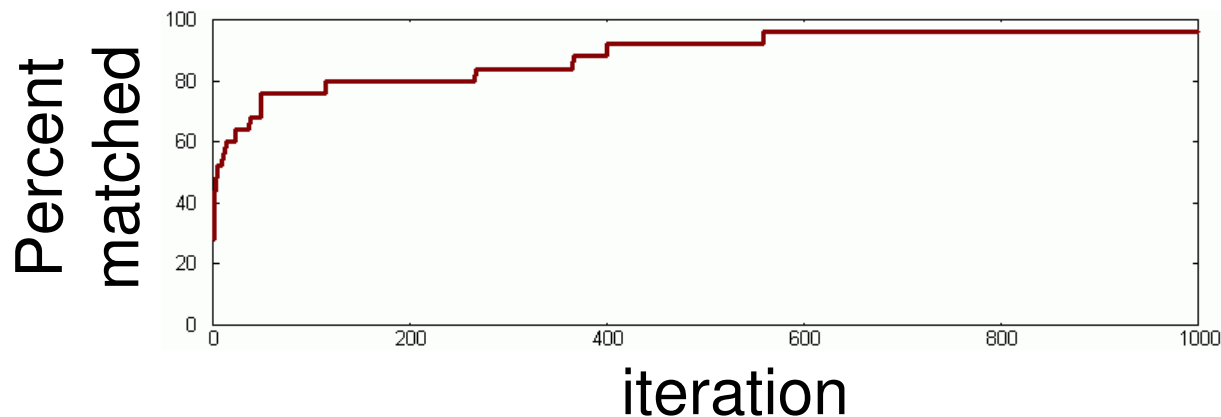


Crack biometric encryption

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



enrolled



Summary

- There is a tendency to use results from cryptography in biometrics security
- However, biometrics images are **not** random data
- Such correlations can probably be exploited to in many biometric systems

Privacy issues

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

Biometrics technology research and privacy?

■ Role of research

- Identify areas where privacy principles are broken
- Develop tests for privacy
- Develop infrastructure to help ensure privacy

Biometrics technology research and privacy?

- Unfortunately, privacy principles are mostly about use of data
- Eg. OECD Privacy Principles:
 - Purpose specification
 - Collection limitation *
 - Use limitation *
 - Security safeguards *
 - Data quality *
 - Accountability
 - Balance security/privacy

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 ... "

Questions?

Security and privacy issues in biometric systems

Abstract:

Biometric authentication technologies form part of sophisticated security systems, which consist not only of biometric sensors and match algorithms, but of databases, communications and cryptographic systems. Little work has been done to study the security and privacy issues of biometric systems in this larger sense, in which well understood characteristics of one part of a system may potentially be exploited elsewhere. One specific interest of mine is to use the fact that biometric images are slightly different each time they are measured, and biometric algorithms must be designed to be tolerant of this variability. This tolerance of variability can possibly be exploited in certain situations to attack an authentication system. For example, it is possible to reconstruct face images from face recognition templates. It also may be possible to use this to attack most biometric encryption methods.

This talk will cover a general overview of some of the security and privacy issues in biometric authentication systems, with a focus on techniques to extract information from biometric templates.