Images can be regenerated from quantized biometric match score data

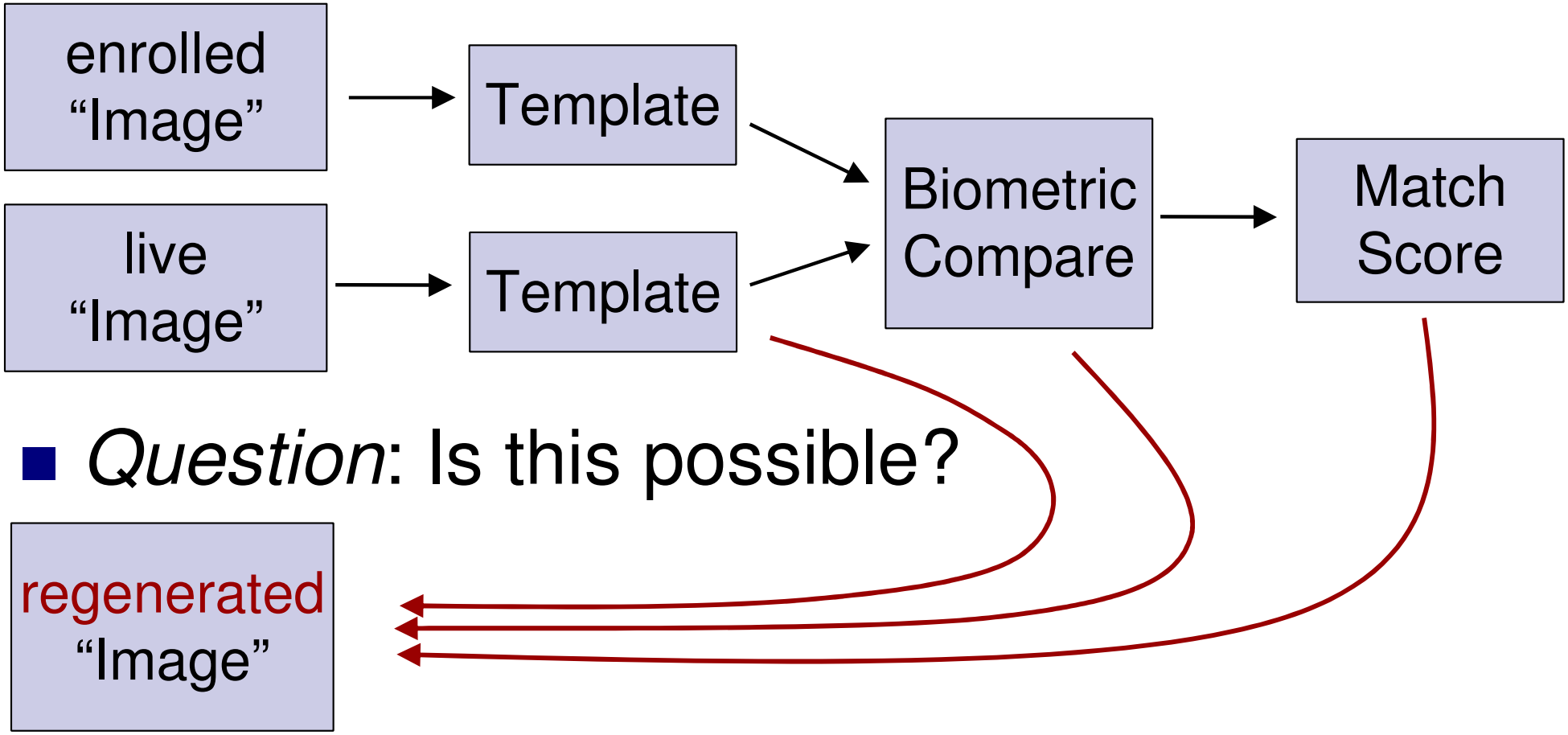
Andy Adler
School of Information Technology and Engineering
University of Ottawa

Problem: Biometrics security

- *Biometric authentication*:
  identification of individuals using behavioural and/or physiological characteristics:
  - Fingerprint, iris image, face recognition, gait, …
- Applications:
  - Identity cards and systems (ie. border control)
  - Authentication for login / security
  - Time and attendance
- Biometric systems vulnerabitilies?
  - Obviously, they can be exposed to all traditional cryptographic attacks
  - Are they vulnerable to *image based* attacks

Images can be regenerated …?

- Typical Biometric processing



- *Question*: Is this possible?

Traditional wisdom

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

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".

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 biometric server which allows comparison of images to the target

“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 *EF*i.
- *Initial image selection* (*IM*$_0$)*:*

“Hill-climbing” Algorithm

Iterative estimate improvement: (for *i …)*

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

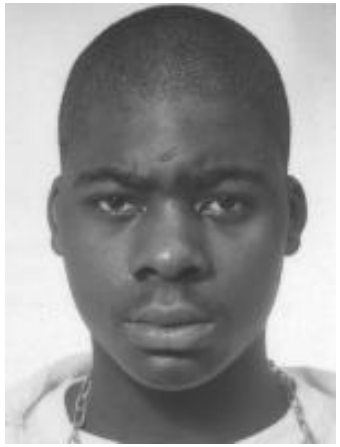
$MS_j = biometric\_compare(IM_k + c_j \times EF_k, IM_{targ})$

- $j_{max} = j$ 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)

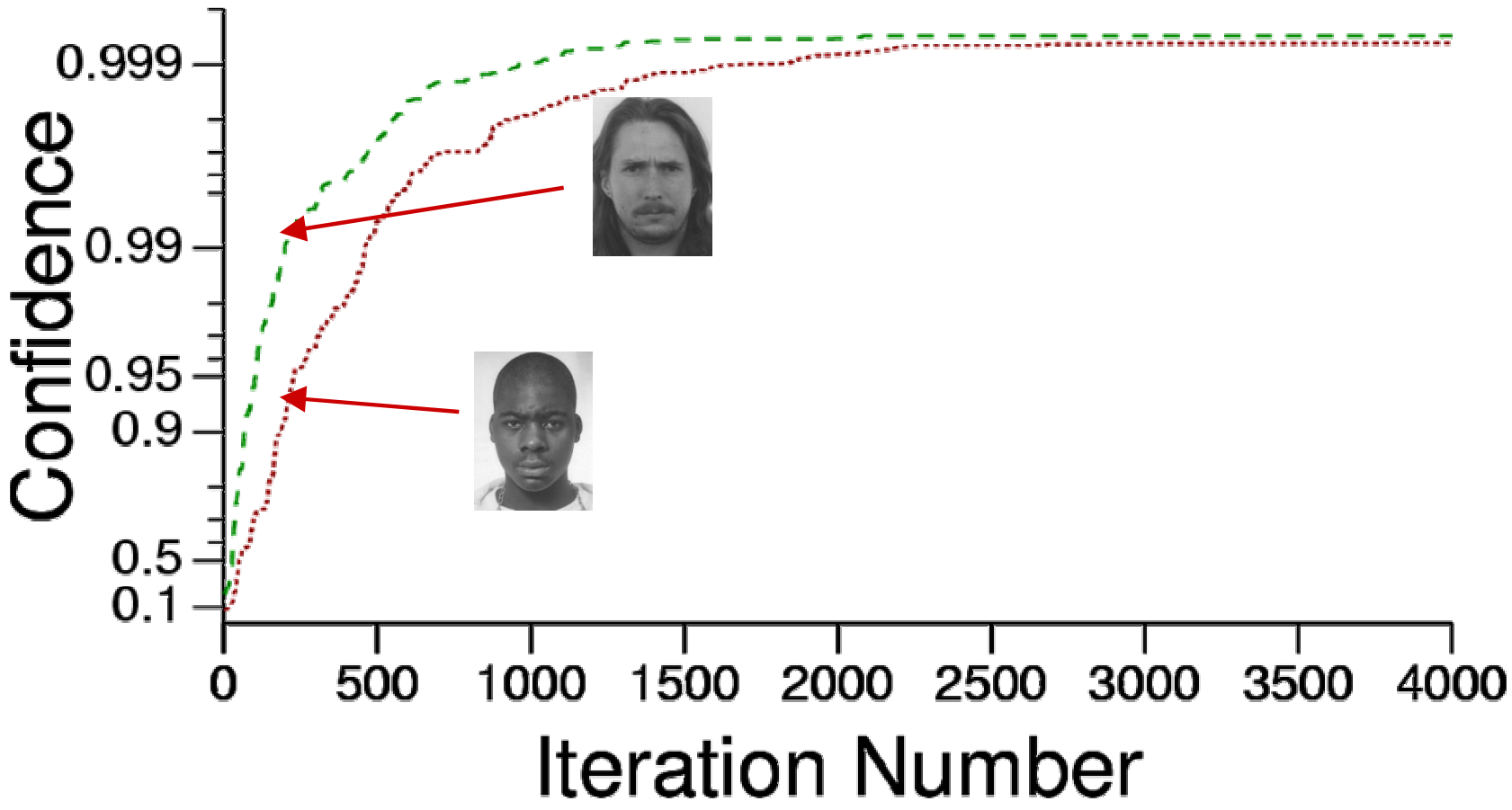
Results

- 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					

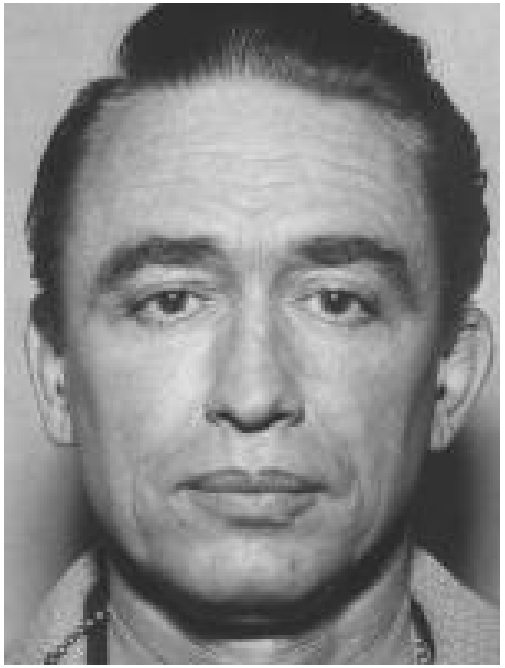
Results: Confidence vs. iteration



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

Improved regenerated image

Average of 10
Best Estimates



Target Image

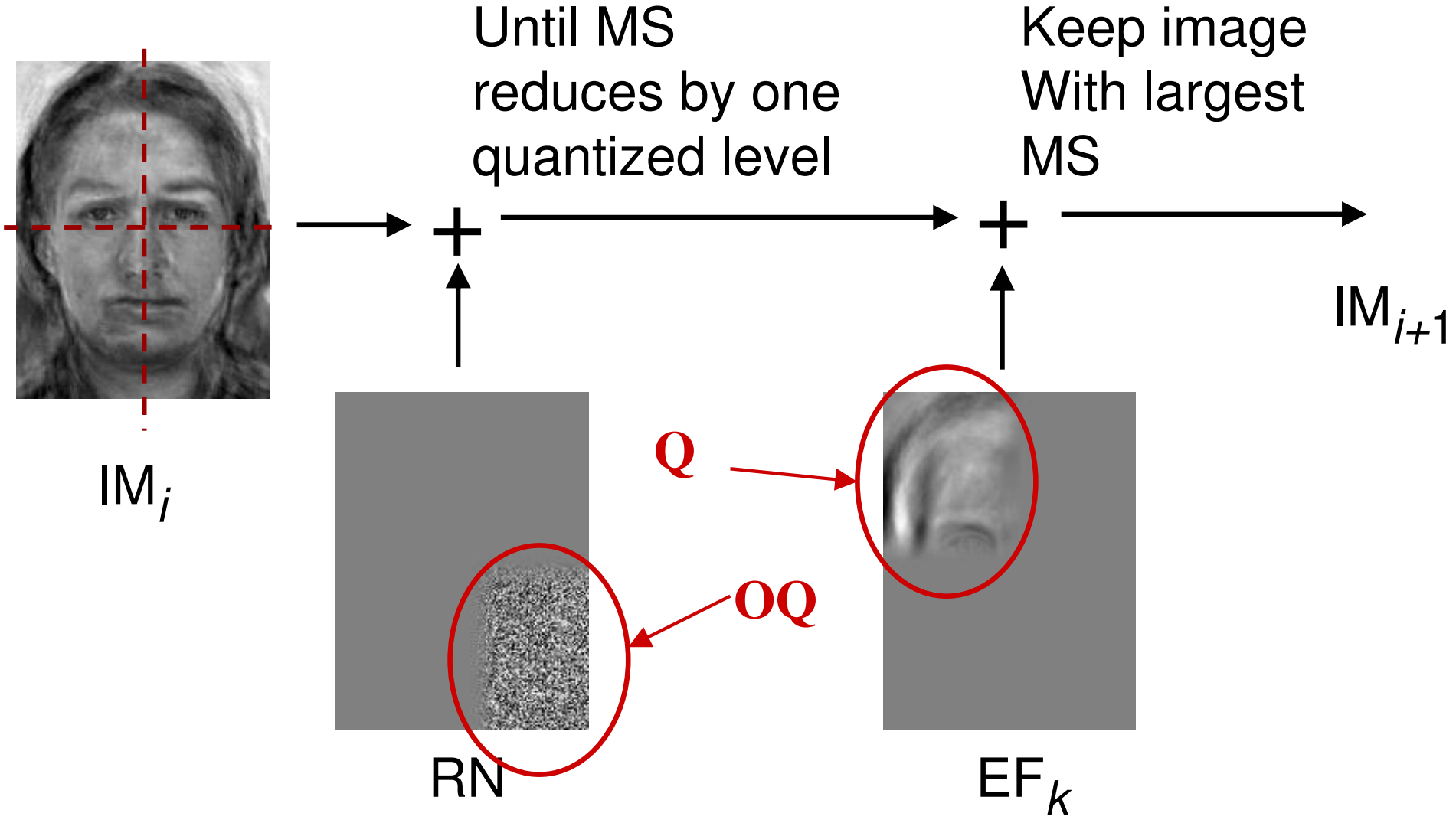
Protection:

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
- This work: We modify the "hill-climbing" algorithm to work with quantized data

Source: BioAPI, version 1.1, p.21, http://www.bioapi.org

Modified "hill-climbing"



Modified "hill-climbing"

Iterative estimate improvement: (for *i …)*

- 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 reduces the MS_i by one quantization level.

New

$MS_i = biometric\_compare(IM_i, IM_{targ})$

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

Modified "hill-climbing"

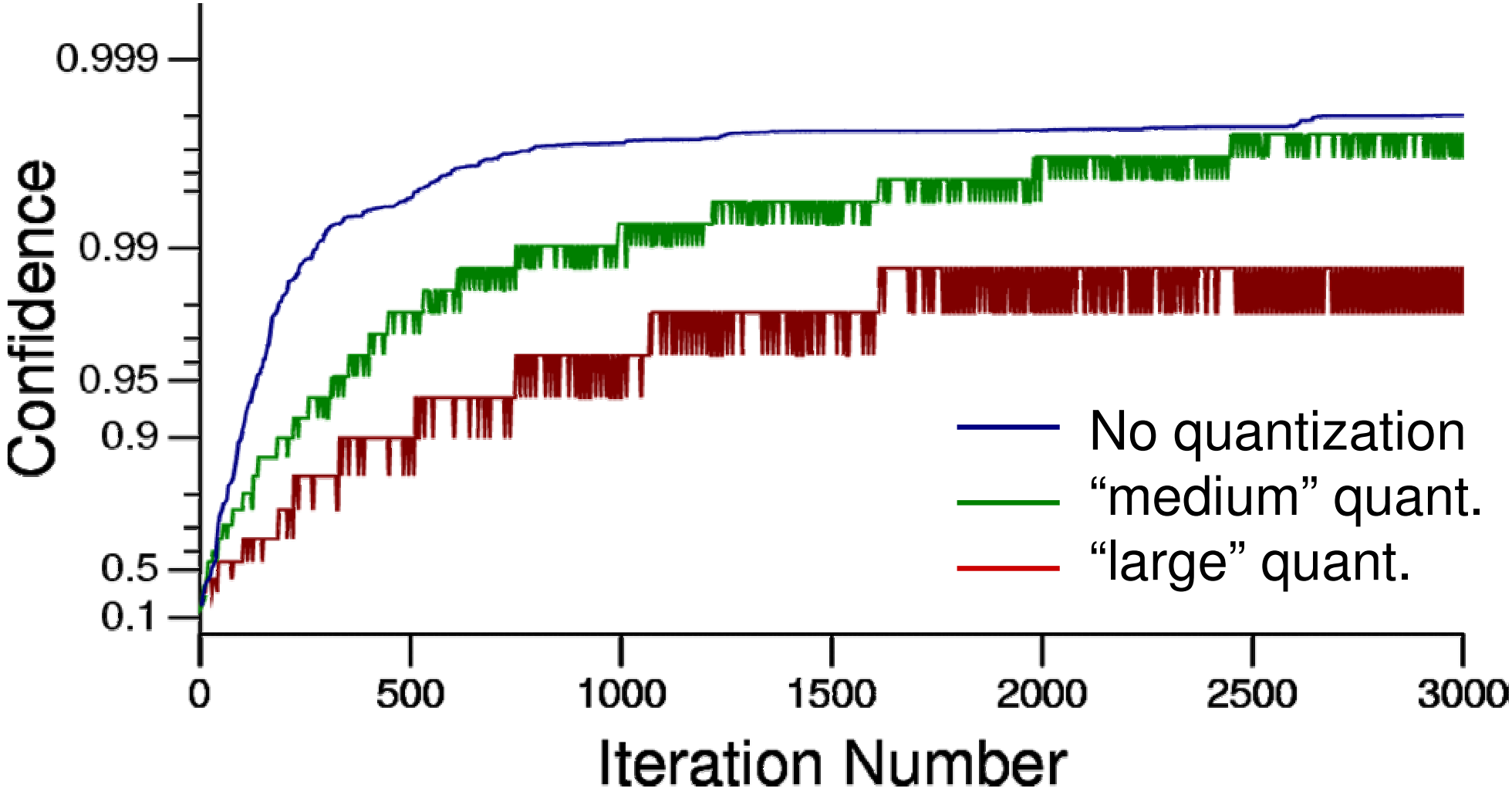
Iterative improvement *(continued …)*

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

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

- $j_{max} = j$ 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

So what?

Approaches shown are:

- Time consuming
  - needs 40,000 biometric comparisons
- Doesn’t produce great images
  - Neither fingerprint / facerec. images look much like the originals

Implications:

- Image regeneration *is* possible
- Smarter people can probably figure out better and faster ways to do it
- Look alike image could be used to
  - masquerade as target
  - Identify target person

Some privacy/security implications:

Biometric Data on ID documents:

- Not an issue for Face Rec. (holders photo is already on the document)
- However, countries may use fingerprint / iris template.

Security agencies may allow searches against watch list:

- Primary agency does not want to distribute images
- However, another agency may access these images through regeneration from match scores

Final thought

- 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