

# Can sample images be regenerated from biometric templates?

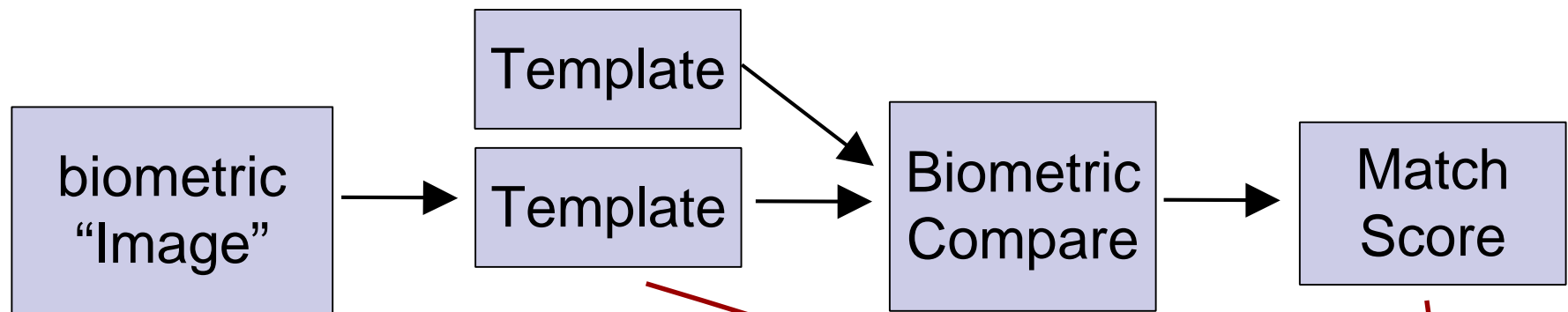
Andy Adler

School of Information Technology and Engineering

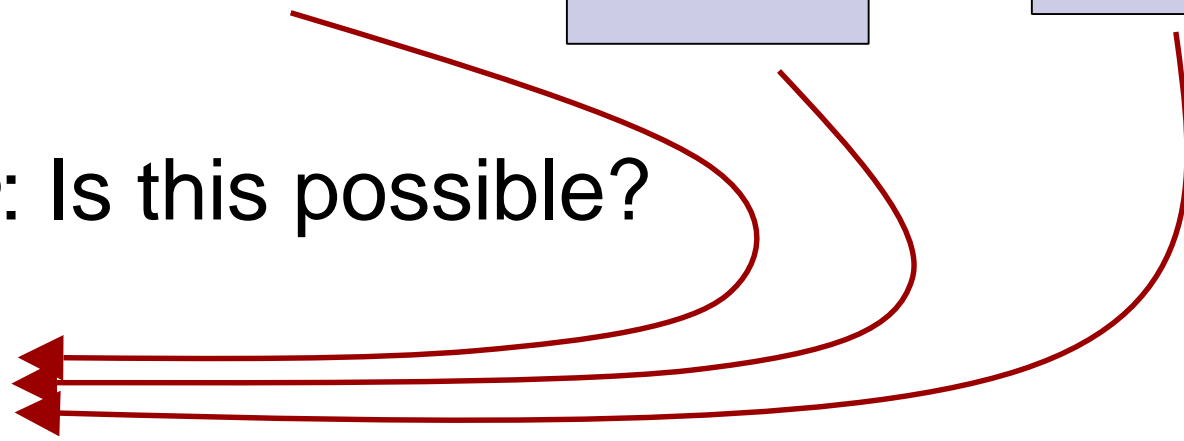
University of Ottawa

# What do we mean *regenerated* ?

## ■ Typical Biometric processing



## ■ *Question*: Is this possible?



# *Regenerated* images

International Biometric Group defined the following possibilities:

- Feature
  - an image which fools biometric algorithm
- Generic image
  - a rough resemblance to the original
- Image
  - virtually identical to the original

Source: [www.ibgweb.com/reports/public/reports/templates\\_images.html](http://www.ibgweb.com/reports/public/reports/templates_images.html)

# Traditional wisdom

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

# Template are difficult to “hack”

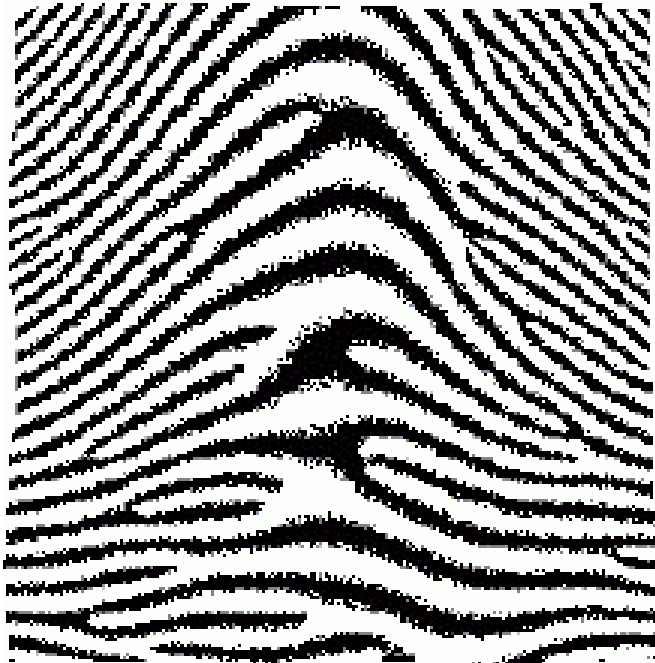
C. Hill reverse engineered the format of a particular (unspecified) fingerprint algorithm.

- Fingerprint images with “controlled differences” were presented and the differences in template observed

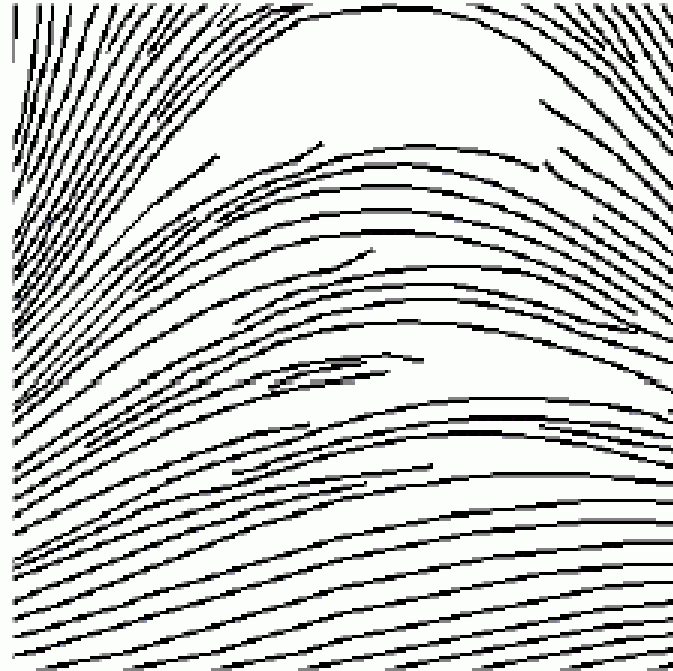
Software was developed to generate an image which would

- compare at high match score with the original,
- show characteristics of the original fingerprint.

# “Hacked” fingerprint images



Original Image



Generated Image

Source: C.J. Hill, 2001 <http://chris.fornax.net/biometrics.html>

# Limits of “hacking” templates

- While this is obviously a powerful technique, it requires
  - Smart people (and lots of their time)
  - Access to the vendor code
- There are some simple countermeasures
  - Template Encryption
    - Note that encrypted templates need a key somewhere. Often it is hidden in the biometric software, and is thus vulnerable

# Automatic image *regeneration*

**Question:** is it possible to have generic software to regenerate image?

- Does not depend on specific vendor file format

**Idea:** begin with a guess, make small modifications; keep modifications which increase the score

# Algorithm: Preprocessing

## **Given**

Person ID in FR database

## **Preprocessing**

Normalize local image database: LDB

Calculate eigenface representation:  $EF[k]$

## **Determine starting image, $Im[0]$**

Find image in local database with maximum match score against target

# Algorithm: Regenerate image

## Optimize image estimate, $Im[k]$ :




- loop (k)
  - Select an eigenface ( $EF$ )
  - find  $c$  to maximize:  
     $match\_score( Im [k] + c \times EF, target )$
  - $Im[k+1] = Im[k] + c \times EF$
  - crop  $Im[k+1]$  if values outside image bounds
- Stop when no further image improvement

# 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

- intentionally different initial images chosen

Target Person	Initial Image Estimate #1	Initial Image Estimate #2
		

Iteration 0

Iteration 200

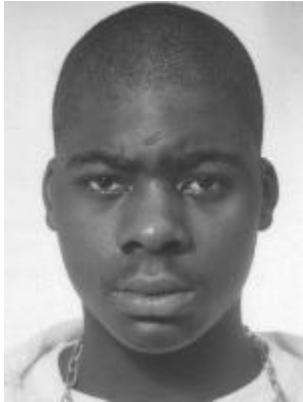
Iteration 600

Iteration 4000

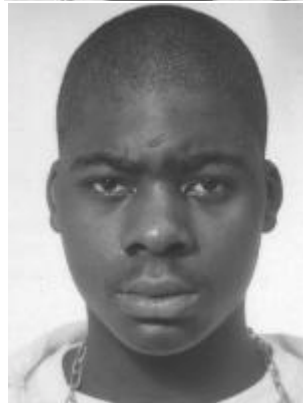
A



B



C



Iteration 0

Iteration 200

Iteration 600

Iteration 4000

A



B



C



# Improved regenerated image: average 10 estimates

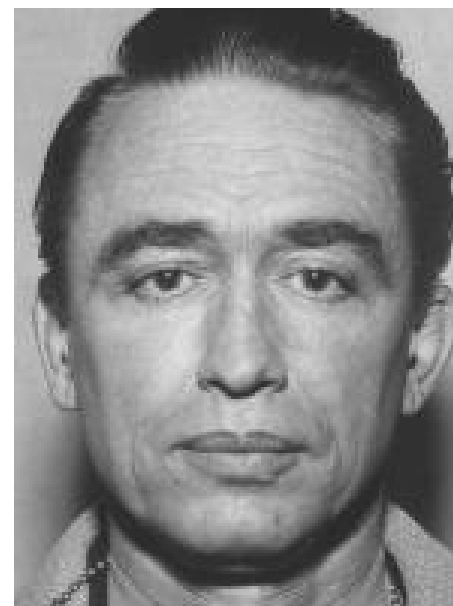
Algorithm A



Algorithm C



Original



# Image estimate changes reflect information in the template

- Feature that are modified
  - Eyebrows
  - Eye shape
  - Nose shape
  - Head shape
  - Mouth expression  $\leq$  *unexpected*
- Features that don't change
  - Hair
  - Moustache / Facial hair region
  - Image background

# Differences between algorithms

- Features modified by some but not other algorithms
  - Hair (Algorithm A)
  - Nose width (Algorithms A,C, not B)
- This probably reflects which features are encoded into the template:
  - This may be a way to reverse engineer details of a proprietary algorithm

# Protection:

## According to BioAPI

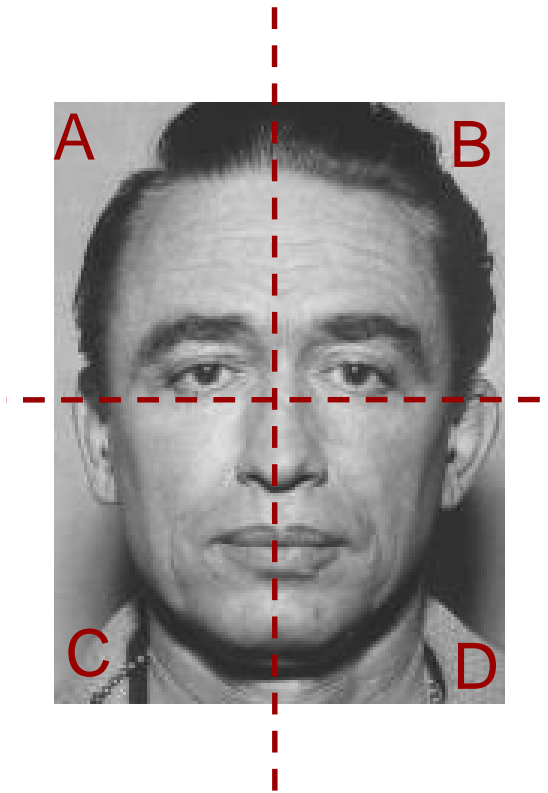
- “...allowing only discrete increments of score to be returned to the application eliminates this method of attack.”
- This makes image regeneration more difficult, but not impossible

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

# Discrete match scores

Preprocessing:

- Step 1: Divide the image into quadrants
- Step 2: From each EF image, calculate the part in each quadrant  $EF_A$ ,  $EF_B$ , etc.



# Discrete scores: regenerate image

loop (k)

- Select an eigenface in quadrant (e.g.  $EF_A$ )
- Select degrade image in opposite quadrant ( $DI_D$ )
- Find  $k$  such that  
     $\text{match\_score}( \text{Im}[k] + k \times DI_D, \text{target} )$   
    is one score lower
- find  $c$  to maximize:  
     $\text{match\_score}( \text{Im}[k] + k \times DI_D + c \times EF_A, \text{target} )$
- $\text{Im}[k+1] = \text{Im}[k] + c \times EF$
- crop  $\text{Im}[k+1]$  if values outside image bounds

# Discrete scores

- Information available is much less, so algorithm takes longer
- Image regeneration works because biometric algorithms “sum up” matching characteristics
  - Changes in quadrants are “independent”
  - 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

Approaches

- “Hack the template”
- “Grind through match scores”

# So what?

Approaches shown are:

- Time consuming
  - needs 40,000 biometric comparisons
- Don'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, some countries may put fingerprint / Iris template, but not be prepared to put image on document.

## 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