Sample images can be independently regenerated from face recognition templates

Andy Adler

School of Information Technology and Engineering

University of Ottawa

Introduction – Biometrics

Biometric authentication

- □ automatic identification, or identity verification, of individuals using behavioural and/or physiological characteristics

Increasing use of biometrics

- □ In National I.D. systems since Sept. 11, 2001
- □ For example: USA requires a biometric on passport from VISA waiver countries – almost all have chosen face recognition

Biometric processing

- acquisition of a biometric sample image
- conversion of the sample image to a biometric template
- comparison of the new (or "live") template to previously stored templates
  - calculate a match score.
  - High match scores indicate a likelihood that the images are from the same individual.

Biometric Template

- compact digital representation of the essential features of the sample image.
  - typically vendor specific
- Template is stored
  - In national I.D. database
  - On passports or I.D. documents
- Templates are considered non-identifiable
  - Often treated like a password “hash” – and distributed to non-trusted parties

Biometric Template

- Biometric vendors have uniformly claimed that it is impossible or infeasible to recreate the image from the template.
  - 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 sample image
  - proprietary nature of the storage format makes templates infeasible to "hack".

This work: algorithm to regenerate image from templates

- Begun with suspicion that templates could be “hacked” to retrieve information
- Then, noticed that the *match score* gives out significant information
- Idea: begin with a guess, make small modifications, and play *colder, warmer, hot* to optimize guess

Algorithm: preprocessing

Given

Person ID in FR database

Preprocessing

Normalize local image database: Img[i]

Calculate eigenface representation: EF[k]

Determine starting image, Im[0]

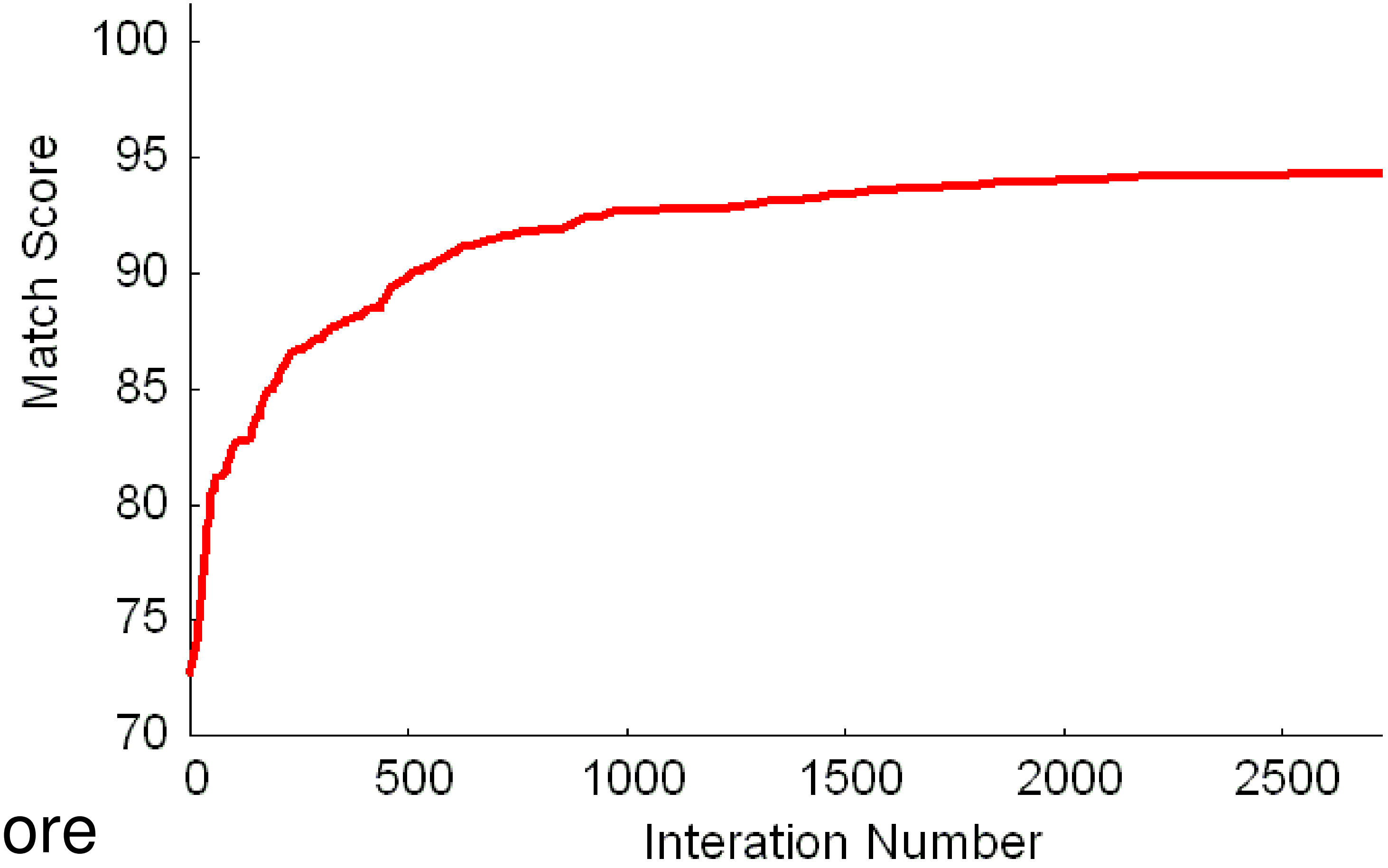
Find image in local database with maximum match score against target

Algorithm: Regenerate image

Optimize image estimate, Im[k]

```
for k = 0 to optimization_tries
    □ Select an eigenface (EF)
    □ find c to minimize:
        match_score( Img[k] + c×EF, target )
    □ Im[k+1] = Img[k] + c×EF
    □ crop Im[k+1] if values outside image bounds
```

Results: Match score

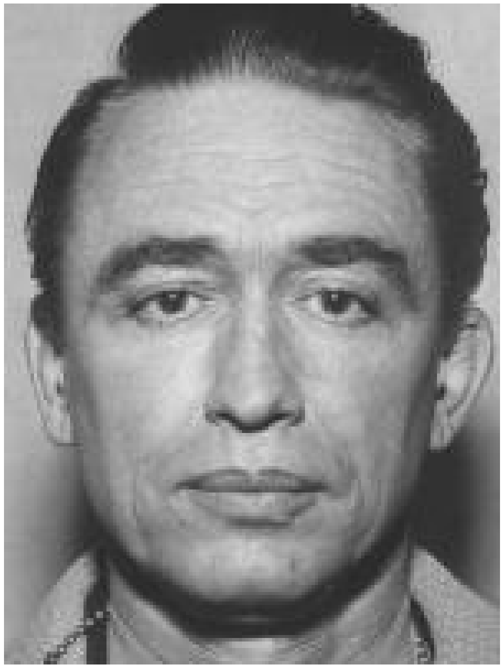
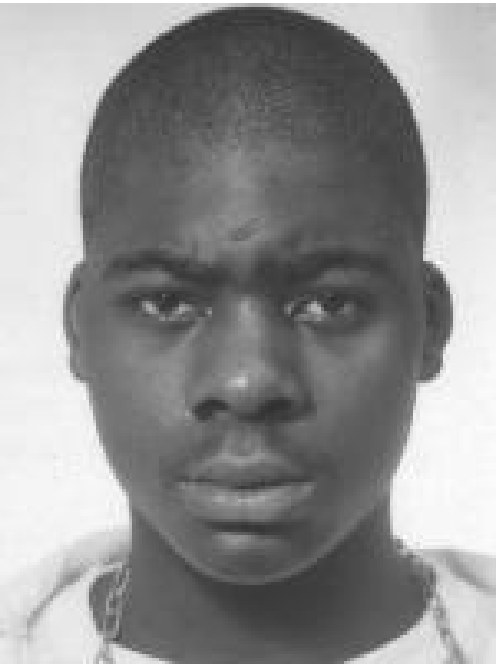
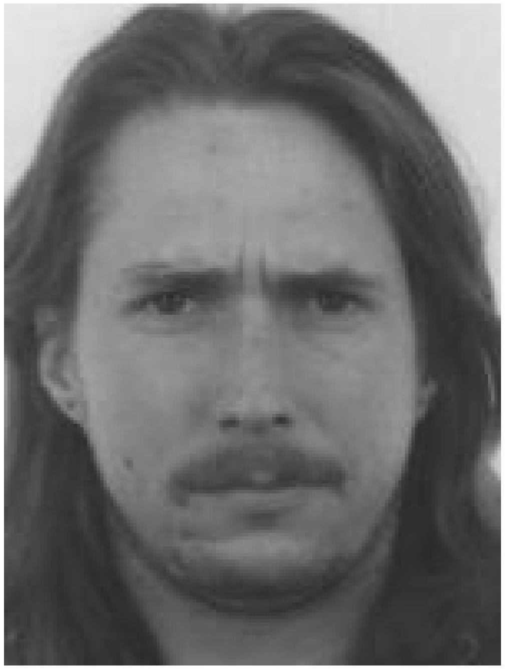


Match Score

- 72.7 indicates 1% confidence image matches
- 94.5 indicates 99.9% confidence

Test images

- intentionally different initial images chosen

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

Iterative “improvement” of estimates

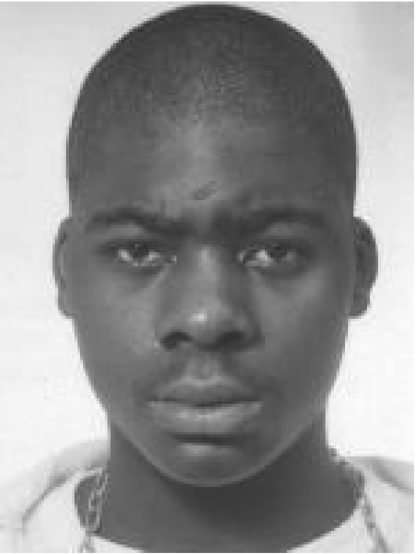
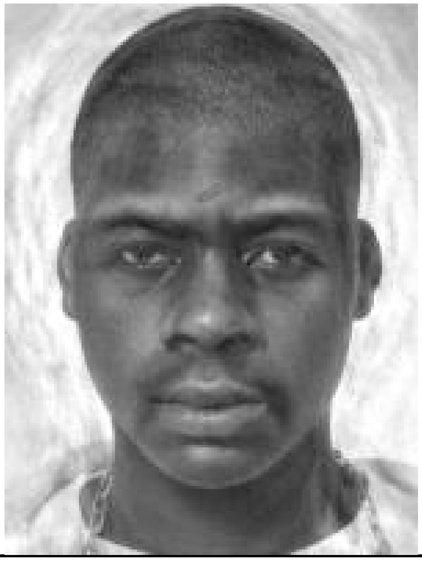
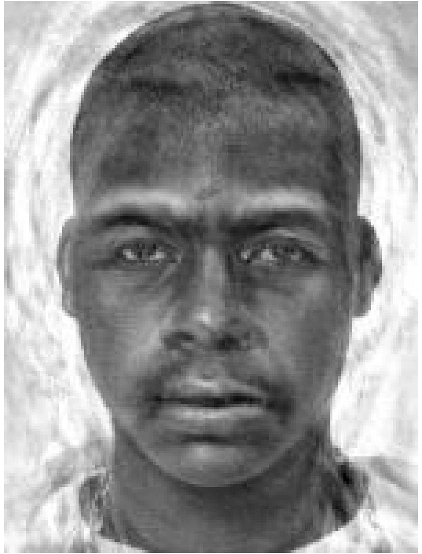
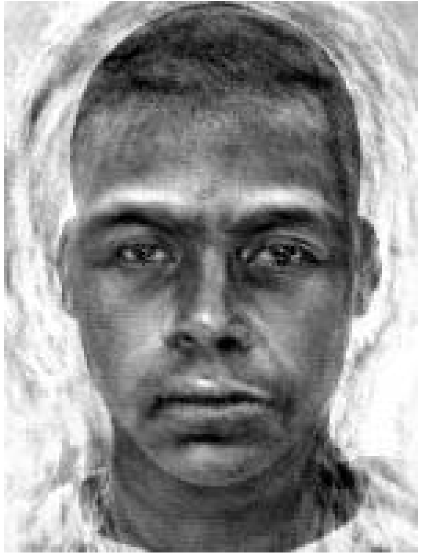
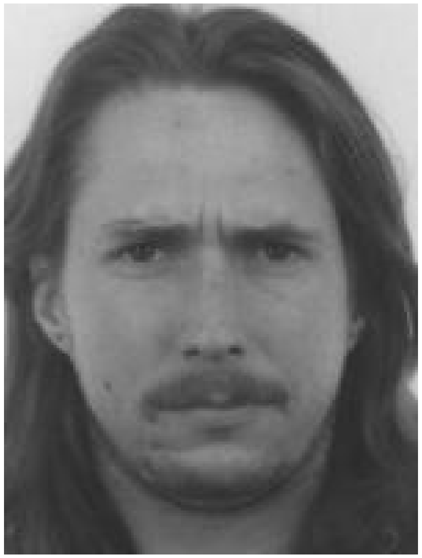
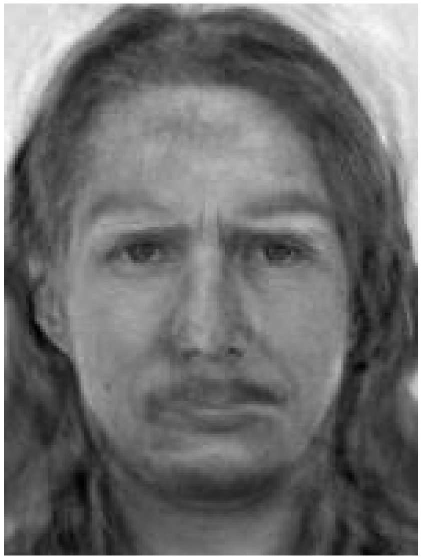
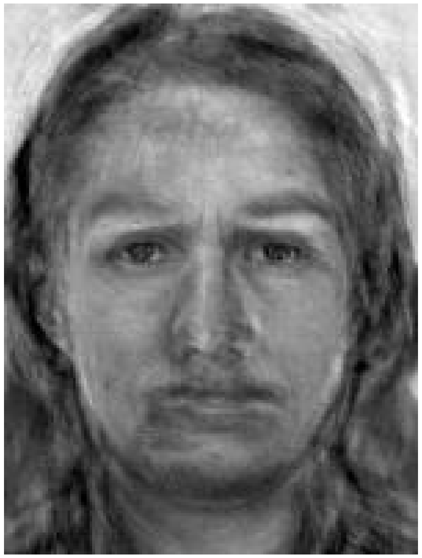
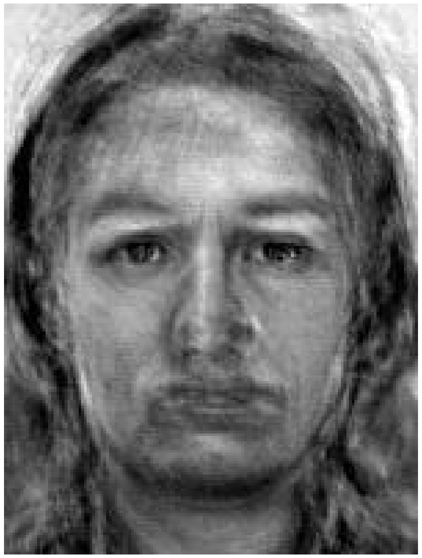
k=0	k=200	k=800	k=4000
			
			

Image improved by averaging estimates

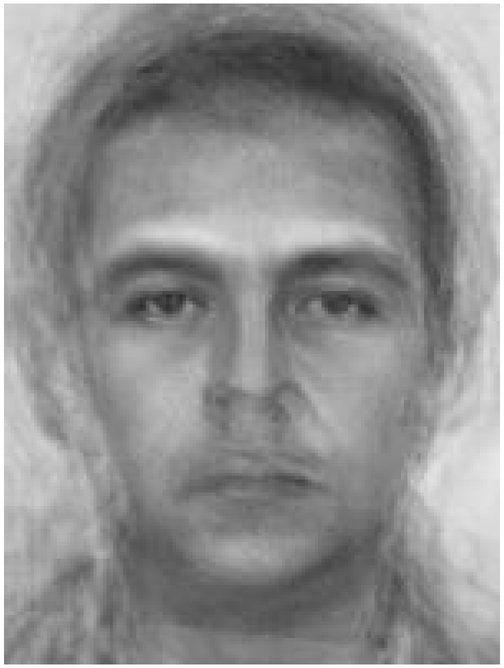
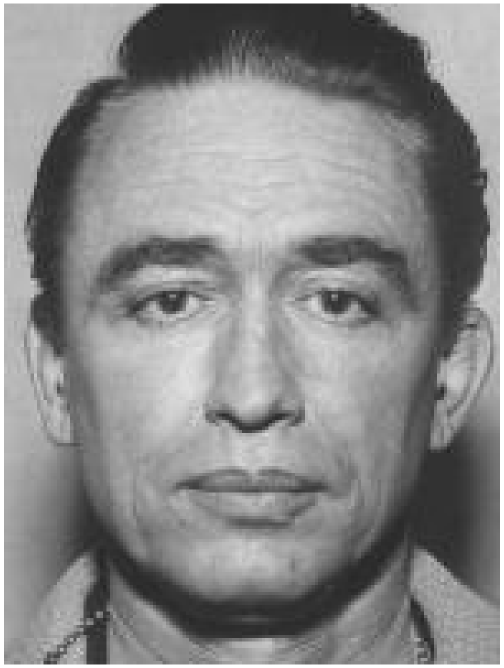
Average of 10 estimates	Original
	

Image estimate changes reflect information in the template

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

Discussion

Developed algorithm to regenerate image from face recognition database

- Image will fool biometric algorithm
- Image visually reflects important features
- Works with algorithms from three different vendors

Discussion: Implications

- Look alike image could be used to
  - □ masquerade as target
  - □ Identify target
- Access to templates or match scores implies access to biometric sample image
  - □ Technique applicable to fingerprint, iris, etc.
- Biometric software systems should provide yes/no only, with no match score values.

Discussion

Approaches to improve image quality

- Use larger number of eigenface vectors
- Use better algorithm
  - Nelder-Mead simplex tried: initial work suggests gives slightly better images for 10x iterations
- Use a better eigenface basis
  - Basis was chosen from Aberdeen database, while sample images from Mugshot, to show that accurate image model not required

Sample Images can be Independently Restored from Face Recognition Templates, Andy Adler (adler@site.uottawa.ca), School of Information Technology and Engineering, University of Ottawa, Ontario, Canada

Abstract

Biometrics promise the ability to automatically identify individuals from reasonably easy to measure and hard to falsify characteristics. They are increasingly being investigated for use in large scale identification applications in the context of increased national security awareness. This paper addresses some of the security and privacy implications of biometric storage. Biometric systems record a sample image, and calculate a template: a compact digital representation of the essential features of the image. To compare the individuals represented by two images, the corresponding templates are compared, and a match score calculated, indicating the confidence level that the images represent the same individual. Biometrics vendors have uniformly claimed that it is impossible or infeasible to recreate an image from a template, and therefore, templates are currently treated as nonidentifiable data. We describe a simple algorithm which allows recreation of a sample image from a face recognition template using only match score values. At each iteration, a candidate image is slightly modified by an eigenface image, and modifications which improve the match score are kept. The regenerated image compares with high score to the original image, and visually shows most of the essential features. This image could thus be used to fool the algorithm as the target person, or to visually identify that individual. Importantly, this algorithm is immune to template encryption: any system which allows access to match scores effectively allows sample images to be regenerated in this way.