The Mug-Shot Search Problem

Ellie Baker and Margo Seltzer
Harvard University, Cambridge, MA., USA
ellie@eecs.harvard.edu

Face Recognition plus Composite Creation

- **Eigenfaces** (Turk & Pentland 1991)
  - Uses PCA to compress images to a low dimensional space of small set of basis vectors called *eigenfaces.*
  - Location in eigenface-space determines the distance between images.
  - Distance from a query image can be used to specify a sort order on a database.
- **Composites**

User Study Goals

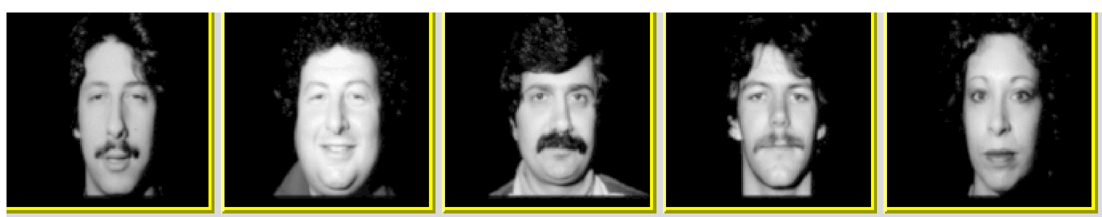
- How well does the eigenface metric correlate with users' assessments of facial similarity?
- Given whatever level of correlation there is between eigenfaces and human users, what search strategies make the best use of it?
- Are the composites helpful?

Prototype System Overview

- **Uses eigenface engine and 4500 image database from Photobook (**Pentland, Picard, Sclaroff - 1994)**.**
- **Queries are either database faces or composites.**
- **Composites are constructed by recombining parts from images in the database.**
- **Interim composites may be used for retrieval and interim retrieval results may likewise be used to update an evolving composite.**

Composite Creation

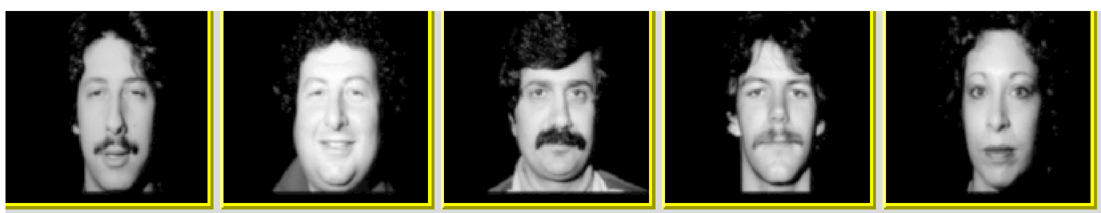
- **Random generation and feature editing**



Register Mental Image

View 100 Random Database Faces

System Generates 10 Random Composites From User's Choices.



User Produces a Composite Via Manual Editing

Evaluation Post-mortem

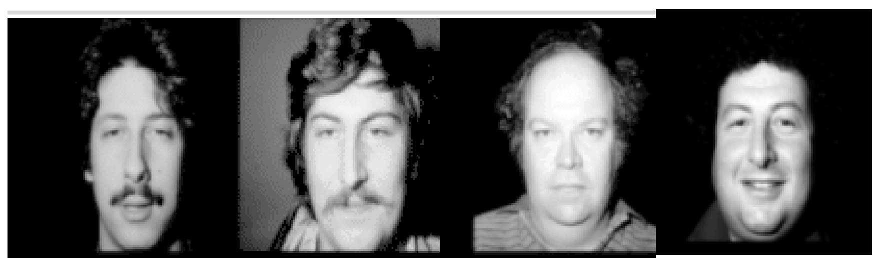
- ***image score*** **= number of image inspections required to find target if that image is used as a query.**
- ***strategy score*** **= number of image inspections required to find target using that strategy.**
- **Determine image scores for each of users':**
  - Top five database choices
  - Random composite choice
  - Final edited composite
- **Which strategies elicit the best average scores over all subjects?**

Best and Worst Case Expected Strategy Scores

- *Worst Case* : sequential search on 4500 images —expected strategy score is 2250.
- Expected image score of closest of N random selections is ~(DatabaseSize/N).
- 4500/100 = 45, so expected score of closest image in random set of 100 is 45.
- *Best Case*: expected strategy score is 100+45 = 145.

Eigenface Best vs. Users' best

40 137 168 206



1230 40 942 137

Results

- **Mean scores for optimal strategies (within a defined class of "reasonable" strategies)**

Target 1: Database images only	323
Target 1: Database + Composites	**260**
Target 2: Database images only	677
Target 2: Database + Composites	**482**

Conclusions

- Eigenface correlation with users' similarity metric exists, but is far from perfect.
- Composites definitely help.
- Hybrid search strategies that use both composites and database images as queries appear to be most successful.