How Quality Influences Human-Computer Face Recognition

Dr. P. Jonathon Phillips
National Institute of Standards and Technology
Acknowledgements

• *In collaboration with*
  – Alice O’Toole
  – Fang Jiang
  – Nils Pénard
  – Janet Ayyad
  – Hervé Abdi

• *supported by NIJ (JP) & TSWG*
Overview

• Rationale
• Background on the FRGC
• Testing humans
• Results
• Conclusions and implications
Problem

• Are face recognition algorithms ready for applications?
  – enormous improvements over last decade
  – accuracy of algorithms tested intensively

• How accurate do they have to be to be useful?
  – meet or exceed human performance
Why?

- **Humans are the competition!**
  - Human-machine comparisons *virtually* never done

- **Putting algorithms in the field**
  - Impact on security?

- **Relative level of performance**
  - “Easy” images
  - “Hard” images
Face Recognition Grand Challenge

Phillips, Flynn, Scruggs, Bowyer, Worek 2006
The primary objective of the FRGC is to:

Develop still and 3D algorithms to improve performance an order of magnitude over FRVT 2002.
Select Point to Measure

• Verification rate at:
  - False accept rate = 0.1%

• Current:
  - 20% error rate (80% verification rate)

• Goal:
  - 2% error rate (98% verification rate)
FRGC Modes Examined

- Single Still
- Multiple Stills
- Outdoor/Uncontrolled
- 3D Full Face
- 3D Single view
FRGC Experiments

Exp 1: Controlled indoor still versus indoor still

Exp 2: Multiple still versus multiple still

Exp 3: 3d versus 3D
   3t - Texture only
   3s - Shape only

Exp 4: Uncontrolled still versus indoor still
## Size of Experiments

<table>
<thead>
<tr>
<th>Exp.</th>
<th>Target set size</th>
<th>Query set size</th>
<th>No. Sim Scores (million)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>16,028</td>
<td>16,028</td>
<td>257</td>
</tr>
<tr>
<td>2</td>
<td>4,007</td>
<td>4,007</td>
<td>16</td>
</tr>
<tr>
<td>3</td>
<td>4,007</td>
<td>4,007</td>
<td>16</td>
</tr>
<tr>
<td>4</td>
<td>16,028</td>
<td>8,014</td>
<td>128</td>
</tr>
</tbody>
</table>
FRGC Progress

Experiment
Verification rate @ FAR = 0.001
Baseline
Median
Max

University Jan 2005

University Aug 2005

Exp 1 Exp 2 Exp 3 Exp 3t Exp 3s Exp 4

Experiment

17 11 10 4 5 12
Human-Computer Comparison

O’Toole, Phillips, Jiang, Penard, Ayyad, Abdi 2005
Human-Machine Comparisons

• Same image pairs from Exp. 4

• Seven state-of-the-art algorithms
  – 4 from industry
  – 3 from academic institutions

• Comparisons
  – 120 difficult face pairs
  – 120 easy face pairs
Sampling

- homogeneous
  - caucasian males/females 20-30 yrs
  - comparisons made on identity not
    - age, race, sex
Comparing Humans and Algorithms

• problem
  – 128 million face pairs?

• sample face pairs
  – most difficult
  – easiest
Easy and Difficult

* PCA Baseline Algorithm
  - scaled and aligned images (SAIC)
  - available and widely used since the 90’s
  - but not state-of-the-art
Selecting Easy/Difficult Pairs

• **“easy” match pairs**
  – 2 “similar” images of same person
    • similarity scores > 2 sd *above* mean similarity of match pairs

• **“difficult” match pairs**
  – 2 “dissimilar” images of same person
    • similarity scores < 2 sd *below* mean similarity of match pairs

• **“easy” no-match pairs**
  – 2 “dissimilar” images of different people
    • similarity scores < 2 sd *below* mean similarity of no-match pairs

• **“difficult” no-match pairs**
  – 2 “similar” images of different person
    • similarity scores < 2 sd *above* mean similarity of no-match pairs
Methods

• Stimuli
  – 240 pairs of faces
    • 120 male pairs
      – 60 easy
      – 60 difficult
    • 120 female pairs
      – 60 easy
      – 60 difficult
Human subject raters respond…

- 1. sure they are the same person
- 2. think they are the same person
- 3. not sure
- 4. think they are not the same person
- 5. sure they are not the same person
Identity Matching for Difficult Face Pairs

Verification Rate vs. False Accept Rate

- NJIT
- CMU
- Viisage
- Human Performance
- Algorithm A
- Algorithm B
- Algorithm C
- Algorithm D
- Chance Performance
Results Summary

• 3 algorithms surpass humans!
  – NJIT (Liu, IEEE: PAMI, in press)
  – CMU (Xie et al., 2005) (In three talks)
  – Viisage (Husken et al., 2005)

• 4 less accurate than humans
Conclusions

- Algorithms compete favorably with humans on the difficult task of matching faces across changes in illumination
  - some algorithms are better than humans on “difficult” face pairs
  - nearly all are better than humans on “easy” face pairs
We Have Quality