Face Recognition Vendor Test 2006
Experiment 4 Covariate Study

Dr. J. Ross Beveridge
Dr. Geof H. Givens
Dr. Bruce Draper
Mr. Yui Man Lui
Colorado State University

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

The work was funded in part by the Technical Support Working Group (TSWG) under Task T-1840C.
Motivation

*Factors that influence face recognition*
Motivation - Attributes of People

What makes recognition harder/easier?

Young

...

Young
Motivation - Attributes of People

Gender?

Age
Motivation - Attributes of People

Race?

Age
Gender
Motivation - Smile?

Expression?

Age
Gender
Race
Motivation - Environment

Control: Mugshot vs. posed indoor or outdoor

Age  Gender  Race  Expression
Motivation - Glasses

Glasses in uncontrolled imagery.
Motivation - Recap

Age

Gender

Race

Expression

Uncontrolled

Glasses
But Wait, There’s More, Quality

- You cannot do much about
  - Gender, Age, Race, …
- Some control over
  - Setting, Glasses, Expression, …
- What about measurable image properties?
  - Resolution, Focus, …

ISO SC 37 "Biometrics" - Factors Affecting Face Image Quality Imaging
ACQUISITION PROCESS AND CAPTURE DEVICE PROPERTIES
  2. physical properties (e.g. resolution and contrast)
Covariate Analysis
For Analysis We Need …

- Lots of Performance Data
  - FRVT 2006

Specific Problem
- Uncontrolled frontal still against mugshot gallery

Methodology
- Generalized Linear Mixed Effect Model

\[
\log \left( \frac{p_{padj}}{1-p_{padj}} \right) = \mu + \gamma_a + \gamma_b B + \gamma_j + \gamma_{aj} + \pi_p
\]
Introduction - More on FRVT

22 participants, 10 countries

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>China</td>
<td>Denmark</td>
<td>Germany</td>
<td>Israel</td>
<td>Japan</td>
</tr>
<tr>
<td>Romania</td>
<td>Spain</td>
<td>South Korea</td>
<td>United Kingdom</td>
<td>United States</td>
</tr>
</tbody>
</table>

10 US, 7 foreign, 5 with offices in and outside the US

Executing Agency

Sponsoring Agencies

- National Institute of Standards and Technology
- Science & Technology Directorate
- Transportation Security Administration
- Federal Bureau of Investigation
- National Institute of Justice
- Intelligence Technology Innovation Center
- Homeland Security
Introductory - Progress

2006 - Falsely turn away 1/100 people, when only admitting 1/1000 imposters.

For controlled frontal still images
Our Focus - Uncontrolled Stills
Uncontrolled to Controlled Still

2006 - Falsely turn away 10/100 to 40/100 people, when only admitting 1/1000 impostors.

FRVT 2002 Performance (Controlled vs Controlled)

Turn Away 20/100 (at 1/1,000 FAR)
FRVT Covariate Analysis

- **Algorithm -** score fusion of 3 top performers.

- **Imagery -** Uncontrolled match to Controlled.

- **Subset of FRVT 2006 Experiment 4**

- **345 subjects and 110,514 match scores.**
Performance Variable

Verification Outcome: Success / Failure

| Verified | FAR   | Gender | Race   | ...
|----------|-------|--------|--------|-------
| Yes      | 1/100 | Female | Asian  | ...

FAR is a factor

| Verified | FAR   | Gender | Race   | ...
|----------|-------|--------|--------|-------
| No       | 1/1,000 | Male | Asian  | ...

Levels 1/100 1/1,000 & 1/10,000

| Verified | FAR   | Gender | Race   | ...
|----------|-------|--------|--------|-------
| Yes      | 1/10,000 | Male | White | ...

Distributed across pairs.

There are 110,514 pairs like this!

* Outcomes for illustration purposes only.
Face Region In Focus Measure

FRIFM: Sum of Sobel edge magnitude inside an ellipse bounding the face.

“Active Computer Vision by Cooperative Focus and Stereo” by Eric Krotkov.
Face Region In Focus Measure

Low FRIFM examples

High FRIFM examples
Fitting a Statistical Model

Focus
Glasses
Race
Head Tilt

Age
Gender
Resolution

FAR

Verification Outcomes

Generalized Linear Mixed Model

4/21/2008
FRVT 2006 Experiment 4 Covariate Analysis
Using the Statistical Model

Fitted Generalized Linear Mixed Model

\[ P(\text{success} \mid \text{covariates}) \]
Let $A$ and $B$ be 2 covariates that might influence algorithm performance. For example, $A=$ gender (categorical) and $B=$ Query-Eye-Distance (continuous).

- Let $a$ index levels of $A$.
- Let $j$ index the FAR setting, $\alpha_j$

$Y_{pabj}$ is

- 1 if Person $p$ is verified correctly, 0 otherwise.

$Y_{pabj}$ depends on:

- person $p$, covariates $A$ and $B$, and
- false alarm rate $\alpha_j$.
GLMM Model Continued ...

\( Y_{pabj} \) is Bernoulli Random Variable with success probability \( p_{pabj} \).

\[
\log \left( \frac{p_{pabj}}{1-p_{pabj}} \right) = \mu + \gamma_a + \gamma_b B + \gamma_j + \gamma_{aj} + \pi_p
\]

- \( \mu \) = grand mean
- \( \gamma_a \) = effect of setting \( a \) of factor \( A \)
- \( \gamma_b \) = effect of covariate \( B \)
- \( \gamma_j \) = effect of \( \alpha_j \), i.e. a FAR setting
- \( \gamma_{aj} \) = interaction effect between \( A \) and FAR
- \( \pi_p \) = subject id. random effect (next page)
The outcomes, i.e. verification success/failure, are uncorrelated when testing different people but correlated when testing the same person under different configurations.

The Mixed in Generalized Linear Mixed effect Model.

\[ [\pi_1, \ldots, \pi_n]^T \text{ Multivariate Normal where} \]
\[
E(\pi_p) = 0, \text{ Variance } \pi_p = \sigma^2_\pi,
\]

\[
\text{Cor } (y_{pabj}, y_{p'a'b'j'}) = \begin{cases} 
\phi & \text{if } p = p' \\
0 & \text{if } p \neq p'
\end{cases}
\]

This means:

The outcomes, i.e. verification success/failure, are uncorrelated when testing different people but correlated when testing the same person under different configurations.
Vendor Test Covariate Analysis Findings

From the highly expected …

… to the unexpected.
Finding 1: False Accept Rate

Solid = Indoors
Dashed = Outdoors

Probability of Verification

1/FAR

10000
1000
100
Finding 2: Gender

Solid = Indoors
Dashed = Outdoors
Finding 3: Race

- Solid = Indoors
- Dashed = Outdoors

Race: Asian (30k), Hispanic (3k), Unknown (5k), White (74k)

Probability of Verification: 1.0, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4
Finding 4: Glasses

Solid = Indoors
Dashed = Outdoors

Probability of Verification

0.4 0.5 0.6 0.7 0.8 0.9 1.0

N  Y

Query Glasses
(Target Never Wore Glasses)
Finding 5: Distance Between Eyes, Query Image

Small

Medium

Large

Query Location

Indoor
Target FRIFM
25 35 45
10 30 50 70
Query FRIFM

Outdoor
Target FRIFM
25 35 45
10 30 50 70
Query FRIFM

Probability of Verification

0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
Finding 5:

Distance Between Eyes, Query Image

Size of query image (distance between eyes)
Finding 5: Distance Between Eyes, Query Image

Query Location
- Indoor
- Outdoor

Query Environment

Target FRIFM

Probability of Verification

0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0

FRVT 2006 Experiment 4 Covariate Analysis
Finding 5: Distance Between Eyes, Query Image

Boundary of observed data

Probability of Verification

0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
Finding 5: Distance Between Eyes, Query Image

Small

Medium

Large

Large $P_V$ range

~0.90 – ~0.10

Probability of Verification

0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
Finding 5: Distance Between Eyes, Query Image

Low FRIFM good; even for one image
FRIFM Conclusion

- Large performance variation.
  - Indoors [>0.95, ~0.70]
  - Outdoors [~0.90, ~0.10].
- Interaction between covariates
  - Environments (indoors, outdoors)
  - Query image size
  - Target and query FRIFM
- Low FRIFM good
  - Effect if control for only one image
- Outdoors: query size very important
FRIFM Conclusion

- According to this analysis
- Out of focus is higher quality
- Remember, edge density surrogate for focus
  - Is this really quality, …
  - Or other environmental factors, …
  - Or algorithm aberration?
### Factors Affecting Face Image Quality

<table>
<thead>
<tr>
<th>Character</th>
<th>Behavior</th>
<th>Imaging</th>
<th>Environment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FACE</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. anatomical characteristic (e.g. head dimensions, eye position)</td>
<td>1. closed eyes</td>
<td>1. image enhancement and data reduction process</td>
<td>1. dynamic characteristics of the background like moving objects</td>
</tr>
<tr>
<td>2. injuries and scars</td>
<td>2. (exaggerated) expression</td>
<td>2. physical properties (e.g. resolution and contrast)</td>
<td>2. variation in lighting and relate potential defects as</td>
</tr>
<tr>
<td>3. ethnic group</td>
<td>3. hair across the eye</td>
<td>3. optical distortions</td>
<td>• deviation from the symmetric lighting</td>
</tr>
<tr>
<td>4. impairment</td>
<td>4. head pose</td>
<td>4. static properties of the background (e.g. wallpaper)</td>
<td>• uneven lighting on the face area</td>
</tr>
<tr>
<td>5. Heavy facial wears, such as thick or dark glasses</td>
<td>5. makeup</td>
<td>5. camera characteristics • sensor resolution</td>
<td>• extreme strong or weak illumination</td>
</tr>
<tr>
<td>6. subject posing (frontal / non-frontal to camera)</td>
<td>6. scene characteristics • geometric distortion</td>
<td>6. scene characteristics • geometric distortion</td>
<td>3. subject posing, e.g.:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• too far (face too small), or too near (face too big)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• out of focus (low sharpness)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• partial occlusion of the face</td>
</tr>
</tbody>
</table>

From Covariates to Quality Measures ….
Conclusion

- Quality is NOT in the eyes of the beholder
- It is in the performance numbers

- Model quantifies performance change.
  - Turn the knob.
  - Read off the change in performance.
  - Interaction between covariates.
- Tells us where to put our efforts
  - Indoors it is FRIFM.
  - Outdoors it is Query Image Size.
- These models are used in other fields.
  - e.g., Biomedical.
- Studies of Biometrics should use them.
Thank You