CSE 40537 / 60537: Biometrics

Biometrics Basics 5
Anticipating Homework #1....

If you do not have access to a computer that you can easily install software libraries or a VM hypervisor on, let me know.
Measurement Trade-Offs

Problem 2: Biometric (Mis)Identification – Why am I delayed as a “suspect”?

Large watch lists exacerbate the problem.

• Example: Faces checked against terrorist watch list

• Assume a system that checks each person’s face against a watch list of 1,000 suspects. Assume Newark Airport: 5,000 people per hour/14hr day

• The system has an FMR of 0.1%

   Over 70,000 false matches will occur per day from 1K watch list

• Note: 2011 US TSC TSDB list was > 450K
Measurement Trade-Offs

Problem 3: Biometric Identification – “Who can I be today?”

Biometric databases are a security problem.

Example 3: Faces checked against government database

• Criminals gain access to a large face database, and start looking for someone their gang can use to steal an identity.

• With an FMR of 0.1%, a single face will match \((.001 \times 6,000,000) = 6,000 \) people in the DC area. With a “gang” of 10 or 100 what can they do?

• Note: Colorado DMV records have fingerprint, photo and all driving information
Measurement Trade-Offs

Doppelganger Attack

System Parameter:
FMR = 1/X Attempts

Enrollment Data

Verification Algorithm

Match?

Yes

Doppelganger Detected

No

Try another print from the “Dictionary”
# Measurement Trade-Offs

False Matches During Duplicate Checks Require Additional Processing.

<table>
<thead>
<tr>
<th>Number of False Hits <strong>Per Search</strong> (i.e., each person being checked)</th>
<th>Total <em>false matches</em> that must be resolved in determining duplicates</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FAR</strong></td>
<td><strong>50 Million</strong></td>
</tr>
<tr>
<td>1%</td>
<td>500,000</td>
</tr>
<tr>
<td>0.1%</td>
<td>50,000</td>
</tr>
<tr>
<td>0.01%</td>
<td>5,000</td>
</tr>
<tr>
<td>0.001%</td>
<td>500</td>
</tr>
<tr>
<td>0.0001%</td>
<td>50</td>
</tr>
</tbody>
</table>

Two ways to address these false hits are:

- **Manually?** (Note: 2.5 Billion minutes is about 4,750 years!)
- **Automatically use another biometric and hope to reduce FAR**
Correct Decisions

GMR = 1 - FNMR

Maximize

GNMR = 1 - FMR

Minimize
Error estimators as a function of \( \tau \)
Failure to Acquire

Falsely rejected biometric samples

Problem at acquisition time:

Wei et al., “Robust and Fast Assessment of Iris Image Quality,” ICB 2006
Failure to Enroll

Falsely rejected biometric samples at enrollment time

“Immigration Delay Disease”

Image Credit: Eli Sprecher, American Journal of Human Genetics
Receiver Operating Characteristic Curve (ROC)

Genuine Match Rate (%)

False Match Rate (%)

Image Credit: Jain et al., Handbook of Biometrics, 2008
Detection Error Tradeoff Curve (DET)

Image Credit: Jain et al., *Handbook of Biometrics*, 2008
Cumulative Match Curve (CMC)

Image Credit: Jain et al., *Handbook of Biometrics*, 2008
Rise of Machine Learning

Why the sudden surge in ML vision applications?

Advances in parallel computing
Large Memories
Web-scale Data

Additional error statistics...
Accuracy

Accuracy =

\[
\frac{|\text{genuine matches}| + |\text{genuine non-matches}|}{|\text{genuine matches}| + |\text{genuine non-matches}| + |\text{false matches}| + |\text{false non-matches}|}
\]

Error = \( E = 1 - \text{Accuracy} \)
Cross Validation

\( N \) training samples
\[ X_y = \{(x_1, y_1), \ldots, (x_N, y_N)\} \]

Empirical error on the training set
\[ M_y = f(X_y, \varphi_M), \text{ s.t. } E_{\text{train}} \text{ is minimized} \]

How do we estimate?