Evaluation of Biometric Spoofing in a Multimodal System

Ricardo N. Rodrigues, Niranjan Kamat and Venu Govindaraju

rnr4@buffalo.edu
Security of MMS

- Multimodal systems (MMS)
- combine two or more biometrics
  - Improve performance
  - Improve universality
  - Increase security
    - Security aspects in MMS relatively unexplored
      - WORK IN PROGRESS!
Are multimodal systems more secure?

• Quotes from literature (but with no evidence.)
  – “An intruder would have to break more than one mode to successfully break the whole system.”
  – "Further, multibiometric systems provide anti-spoofing measures by making it difficult for an intruder to spoof multiple biometric traits simultaneously.“
  – "In general, by fusing several classifiers, the overall error rate (the false accept rate and the false reject rate) is known to go down and also reduces spoof attacks on the biometric system."

• Our claim: MMS may NOT be more secure
More vulnerability points to attack

More choice on modalities to attack

Can go for weakest link
Motivation

- Explore security of Multimodal Biometric Systems
- Skilled Forgery = Spoofing
  - Underestimate FAR when only considering random imposters
  - Signature verification tests have always dealt with skilled forgeries
- Intuition: a student needs to cheat and score high in one subject to pass if the passing threshold is based on average scores in all subjects.
Examples of spoofing
Skilled Forgery

- Fingerprint

- Face

Recognition stage

Face spoof by 2D image convincing

Fingerprint spoof better than random imposter
Our Multimodal System (MMS)

- Combine face and fingerprint modalities
- Fusion at matching score level
- Fingerprint: extension of NIST system (Jea 2004)
- Face: undisclosed commercial system
Fusion Schemes

• **Weighted Sum (WSum)**
  \[ S_{mm} = \alpha S_{fi} + S_{fa} \]
  Alpha weight fixed iteratively on training data in steps of 0.1
  weakest security mode could get high weight for accuracy

• **Likelihood Ration (LLR)**

• **Bayesian Likelihood Ratio (Bayes)**
  – Same as LLR but with different
  – \( T \): presence of spoof attack
  – \( S \): spoof attack successful

\[
p(s|I) = \sum_{T \in \{0,1\}} \sum_{S \in \{0,1\}} p(s, T, S|I)
\]
Bayesian LLR

- **Joint distribution:**

\[
p(s, T, S | I) = p(s, S | T, I)p(T | I)
= p(s | T, S, I)p(S | T, I)p(T | I)
= p(s | S, I)p(S | T, I)p(T | I)
\]

\[
p(T = 1 | I) = 0.5 \quad p(S = 1 | T = 1) = 1
\]

\[
s_{mm} = \frac{p(s_{fa} | G)p(s_{fa} | G)}{p(s_{fa} | I)p(s_{fa} | I)}
\]

(Fingerprint spoofed, Face spoofed;  Fingerprint real; Face spoofed)

\[
p(s_{fi} | I)p(s_{fa} | I) = 0.25p(s_{fi} | S = 1, I)p(s_{fa} | S = 1, I) + 0.25p(s_{fi} | S = 1, I)p(s_{fa} | S = 0, I) + 0.25p(s_{fi} | S = 0, I)p(s_{fa} | S = 1, I) + 0.25p(s_{fi} | S = 0, I)p(s_{fa} | S = 0, I)
\]

(Fingerprint spoofed, Face real;  Fingerprint real, Face real)
Datasets

• **Fingerprint Liveness Detection Competition 2009**
  – Crossmatch subset – 217 users, 4000 fingerprints
  – Contains real and spoofed samples
  – Ignored data from other sensors

• **Face: NIST Biometric Score Set**
  – 517 users, 534578 matching scores
  – Assumption: spoofed sample = genuine sample.

• **‘Virtual’ Multimodal Dataset**
  – Random associations of fingerprint and face data
Experiments

Randomly create 217 'virtual' multimodal users (x 5)

Training (40% virtual users)

- **Genuine** scores: real sample vs real sample of same user
- **Impostor** scores: real sample vs real of different users
- **Spoofed** scores: real sample vs spoof of same user

- Estimate fingerprint - genuine, impostor, spoof distributions.
- Same for face - except genuine = spoof distribution
Fingerprint Score Distributions

Spoof scores: match between spoofed sample of a user with template of same user

Estimation using Gaussian kernel and parzen window
Face Score Distributions

Spoof scores: match between spoofed (genuine) sample of a user with template of same user

Estimation using Gaussian kernel and parzen window
Experiments

Testing Scenarios (60% virtual users)

1. No spoofing
   - real face sample and real fingerprint samples

2. Face spoofed
   - spoofed face sample and real fingerprint samples

3. Fingerprint spoofed
   - real face sample and spoofed fingerprint samples

4. Both spoofed
   - spoofed face sample and spoofed fingerprint samples
Scenario 1 – No spoofing

ROC

All fusion schemes show gains
LLR slightly better than Bayes
Results
(Store thresholds for different FARs from Scenario 1)

• Thresholds at $FAR_1=1\%$

<table>
<thead>
<tr>
<th>System</th>
<th>Thrs.</th>
<th>$FAR_1$</th>
<th>$FRR_1$</th>
<th>$FAR_2$</th>
<th>$FAR_3$</th>
<th>$FAR_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fing</td>
<td>0.13</td>
<td>1.00</td>
<td>19.91</td>
<td>1.01</td>
<td>57.56</td>
<td>57.58</td>
</tr>
<tr>
<td>Face</td>
<td>0.61</td>
<td>1.00</td>
<td>8.28</td>
<td>91.63</td>
<td>1.00</td>
<td>91.63</td>
</tr>
<tr>
<td>Bayes</td>
<td>-0.85</td>
<td>1.00</td>
<td>1.76</td>
<td>88.60</td>
<td>28.71</td>
<td>96.34</td>
</tr>
<tr>
<td>LLR</td>
<td>-0.45</td>
<td>1.00</td>
<td>1.75</td>
<td>88.15</td>
<td>42.32</td>
<td>96.21</td>
</tr>
<tr>
<td>WSum</td>
<td>0.62</td>
<td>1.00</td>
<td>1.94</td>
<td>91.13</td>
<td>24.07</td>
<td>95.62</td>
</tr>
</tbody>
</table>

• Thresholds at $FAR_1=0.1\%$

<table>
<thead>
<tr>
<th>System</th>
<th>Thrs.</th>
<th>$FAR_1$</th>
<th>$FRR_1$</th>
<th>$FAR_2$</th>
<th>$FAR_3$</th>
<th>$FAR_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fing</td>
<td>0.71</td>
<td>0.10</td>
<td>69.47</td>
<td>0.10</td>
<td>4.60</td>
<td>4.58</td>
</tr>
<tr>
<td>Face</td>
<td>0.64</td>
<td>0.10</td>
<td>15.92</td>
<td>84.17</td>
<td>0.10</td>
<td>84.17</td>
</tr>
<tr>
<td>Bayes</td>
<td>1.42</td>
<td>0.10</td>
<td>12.23</td>
<td>8.23</td>
<td>5.42</td>
<td>72.19</td>
</tr>
<tr>
<td>LLR</td>
<td>5.29</td>
<td>0.10</td>
<td>10.84</td>
<td>60.33</td>
<td>8.07</td>
<td>81.95</td>
</tr>
<tr>
<td>WSum</td>
<td>0.68</td>
<td>0.10</td>
<td>9.41</td>
<td>73.99</td>
<td>7.18</td>
<td>85.30</td>
</tr>
</tbody>
</table>
Scenario 2
Spoofed face + Real fingerprint

NOT ROC
(thresholds from Sc1
All FARs increase
Except fingerprint
Bayes mod more robust than LLR

Inversely proportional to robustness
Scenario 3
Spoofed fingerprint + Real face

- NOT ROC
- (thresholds from Sc1
- All FARs increase
- Except face
- Bayes mod more robust than LLR
Scenario 3
Spoofed face + Real fingerprint

Fusion I FAR (security) is lower than fingerprint alone
Scenario 4

Spoofed face + Spoofed fingerprint

- Face is the weak link
  - as its forgeries (spoof) are good
- Fingerprint robustness hurt by fusion
Summary

• Multimodal Systems (MMS) may not be more secure than the component modalities

• Need to pay attention to skilled forgery

• Intruder needs to only spoof the ‘weakest link’.
  – Vulnerable when an easy to spoof modality such as “voice print” or “face” is one of the fused modalities
Future Work

• Experiment with other biometric modalities

• Quantify the intrinsic security of each biometric modality participating in the fusion
  – Can we empirically measure how hard is to circumvent a biometric system?

• When are MMS less secure than its constituent modes?
  – Probabilistic analysis
  – Trade-off: Security x Performance
    • When is it safe to fuse?