From the Iriscode to the Iris: A New Vulnerability Of Iris Recognition Systems

Javier Galbally

Biometrics Recognition Group - ATVS
Escuela Politécnica Superior
Universidad Autónoma de Madrid, SPAIN

http://atvs.ii.uam.es
Outline

1. Introduction: Biometrics and Security
2. Biometrics
3. Iris Recognition
4. The Reconstruction Method
5. Experimental Protocol
6. Results: Performance
7. Results: Appearance
8. Conclusions
1. Introduction: Biometrics and Security
FAQ when dealing with IT solutions for security applications:

- How secure is this technology?
- Why should I trust it?
- Who assures the level of security offered by this system?
- ...

Independent Security Evaluation

How is this being implemented in BIOMETRI CS?
Security Evaluation in Biometrics

• Projects:

• Competitions:

• Standards:

Constant need to search for new vulnerabilities
2. Biometrics
Biometric systems

ENROLLMENT

Verification

Identity claim

Biometric Sample

Pre-Processing & Feature Extraction

DATABASE

User Template $T_i$

Matcher (1:1 Matching)

Decision Threshold

Accept Reject
Biometric modalities

**BEHAVIOURAL**
(signature, voice, gait...)

**PHYSIOLOGICAL**
(fingerprints, iris, face, hand geometry...)

- Characteristics:
  - **Universality**: everybody should possess it
  - **Distinctiveness**: should have enough intervariability
  - **Permanence**: should not vary through time
  - **Collectability**: should be easy to acquire
  - **Performance**: should have good error rates
  - **Acceptability**: user should not be reluctant to use it
  - **Circumvention**: difficult to bypass
Attacks to Biometric Systems

• Possible points of attack to a biometric system.

DIRECT ATTACKS
(Spoofing, mimicry)

INDIRECT ATTACKS
(Trojan Horse, Hill Climbing, Brute Force, channel interception, replay attacks, masquerade attacks...)

1. Sensor
2. Feature extractor
3. Matcher
4. Database
5. Threshold
6. Accept/Reject
7. 8. 9. 10.
Objective: Inverse Biometrics

• Inverse Biometrics:
  Can we reconstruct the sample from the template?

• Traditional answer → NO!
• However…

Is this possible for the iris?
3. Iris Recognition
Iris Recognition

- Very low error rates
- Long-term permanence
- Many commercial solutions
- ...
- Vulnerabilities?
Iris Recognition: How does it work?

Acquisition + Detection
Iris Recognition: How does it work?

1. **GENUINE SAMPLE**
2. **Segmentation**
3. **Normalization**

**TEMPATE ➔ IRISCODE**

```
010101010100010101010101010001010101010101010100100101010101010101010100
010101010101000101010101010101010101010101010101010101010101010101010101
010001010101010001010101010101010101010101010101010101010101010101010101
010101010101010001010101010101010101010101010101010101010101010101010101
010101010101010101010101010101010101010101010101010101010101010101010101
010101010101010101010101010101010101010101010101010101010101010101010101
```
4. The Reconstruction Method
The Problem (I)

How do we know that an iris image is the reconstruction of a given template?

Because it is positively matched to the genuine template by iris recognition systems

- Find an iris image: $I_R$
  - Any iris image? $\rightarrow$ NO!
- Such that:
  - It’s associated template $B_R$
  - When compared to the template $B$ (the one being reconstructed)
  - Using a matching function $J$
  - Gives a score higher than a certain threshold $\delta$
The Problem (II)

How do we find such an iris image?

Use a GENETIC ALGORITHM to look for it (i.e., optimize the score = optimize the fitness function)

- GENETIC ALGORITHMS:
  - Heuristic search tool
  - ITERATIVELY applies certain rules inspired in natural evolution
  - To a population of individuals (possible solutions)
  - According to a given fitness function which has to be optimized
**Assumption:** we have access to $s$ for several $IR$
The Solution: The Algorithm (I)

• **STEP 1**: Generate initial population $P_0$ with $N$ individuals ($I_R$)
• **STEP 2**: Compute the $N$ scores $s_i$
• **STEP 3**: Generate the next generation $P_n$ according to four rules (which use $s_i$):
  - **Elite**: two individuals
  - **Selection**: stochastic universal sampling
  - **Crossover**: scattered crossover
  - **Mutation**: random changes
• **STEP 4**: Redefine $P_0 = P_n$ and go back to step 2.

• **Stopping Criteria**:
  - The best score is higher than $\delta$ (RECONSTRUCTION OK!)
  - Score increase in the last generations is very small
  - Maximum number of generations is reached
The Solution: The Algorithm (II)

Selection (Stochastic Universal Sampling) → PARENTS → Crossover (Scattered) → CHILDREN → Mutation ($p_m$) → MUTATED CHILDREN

Scores ($s$) → Elite (Best 2 individuals) → NEW POPULATION ($P_n$) → Norm. Images ($I_R$) ($R \times C$ pixels)

Normalized Iris Image

20/43
**SELECTION**: Stochastic Universal Sampling

- Select mothers and fathers
- Higher fitness function → higher probability of being selected

Population ($N=5$)

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Fitness Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>N1</td>
<td>$f_1=10$</td>
</tr>
<tr>
<td>N2</td>
<td>$f_2=5$</td>
</tr>
<tr>
<td>N3</td>
<td>$f_3=3$</td>
</tr>
<tr>
<td>N4</td>
<td>$f_4=1$</td>
</tr>
<tr>
<td>N5</td>
<td>$f_5=1$</td>
</tr>
</tbody>
</table>

Fitness Value: $F = f_1 + \ldots + f_5 = 20$

Random $[1, F]$: $p=9$

Step: $S = F/N = 4$

- $p + S = 13$
- $p + 2S = 17$
- $p + 3S = 1$
- $p + 4S = 5$

$p = 9$
The Solution: The Algorithm (IV)

- **CROSSOVER**: Scattered crossover
  - Combine fathers and mothers

<table>
<thead>
<tr>
<th>Father</th>
<th>Mother</th>
<th>Child 1</th>
<th>Child 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 2 3 4</td>
<td>A B C D</td>
<td>1 2 3 4</td>
<td>A 2 3 D</td>
</tr>
<tr>
<td>0 0 1 1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Binary matrix (1 prob. = 0.5)
The Solution: The Algorithm (V)

- **MUTATION:**
  - Mutate children
  - Probability of mutation $p_m$ small

Child

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Best Individual

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>C</td>
<td>D</td>
</tr>
</tbody>
</table>

Binary matrix (1 prob. = $p_m$)

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Mutated child

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>
5. Experimental Protocol
Development and Validation

- Avoid positively biased results
- Publicly available DBs and systems → reproducibility

![Diagram showing the process of development and validation with biosecure DBs and systems.](Image)
Development: DBs (I)

- **Development DBs:**
  - **Biosecure DB:**
    - 420 iris users (right and left iris of 210 users)
    - 2 sessions
    - 2 samples per session
    - Total of 420 x 2 x 2 = 1,680 iris images
    - Available at: http://biosecure.it-sudparis.eu/
  - **Synthetic DB (SDB):** SYNTHETIC database used to generate P0
    - 1000 users
    - 1 session
    - 7 samples per session
    - Total of 7,000 iris images
    - Available at: http://www.citer.wvu.edu/
Development: DBs (II)

- Typical examples from Biosecure DB and SDB.
- Totally different → results are not biased.
Development System: academic implementation. Used to compute scores $s_i$ in the reconstruction algorithm.

- **Segmentation**: iris and pupil boundaries $\rightarrow$ circles
- **Normalization**: rubber sheet model
- **Feature encoding**: based on 1D Log-Gabor filters
- **Matching**: hamming distance

Available at: http://www.csse.uwa.edu.au/pk/studentprojects/libor/sourcecode.html
Validation: DBs

- **Validation DBs:**
  - **Biosecure DB:** REAL database attacked.
  - **Reconstructed Biosecure DB:** SYNTHETIC database used to perform the attacks

- 420 users
- 5 reconstructions of 1 genuine sample per user
- Total of $420 \times 5 = 2,100$ iris reconstructions
**Validation: System**

- **VeriEye**: commercial application
  - BlackBox: no info about how it works $\rightarrow$ unbiased results
  - It requires as input EYE images (NOT normalized iris images)
  - Available at: http://www.neurotechnology.com/verieye.html
• Performance measure: Success Rate (SR) \( \rightarrow \) \( SR = \frac{A_s}{A_T} \)
  - \( A_s \) = Successful attacks
  - \( A_T \) = Total attacks

• Types of attack

  **Attack 1**
  - Attack 1a: real = reconstructed sample (\( AT = 1 \times 5 \times 420 = 2100 \))
  - Attack 1b: real = NOT reconstructed sample (\( AT = 3 \times 5 \times 420 = 6300 \))

  **Attack 2**
  - Attack 2a: real = reconstructed sample (\( AT = 1 \times 420 = 420 \))
  - Attack 2b: real = NOT reconstructed sample (\( AT = 3 \times 420 = 1260 \))

  **Attack 3**
  - average(4 real) vs 5 reconstructed (\( A_T = 1 \times 420 = 420 \))

**Most likely attacking scenario**
6. Results: Performance
How do we know that an iris image is the reconstruction
of a given template?

Because it is positively matched to the genuine template
by iris recognition systems
(score higher than a certain threshold $\delta$)

FAR=0.1%
Results: Development (II)
• VeriEye (validation system): commercial application
  - It requires as input EYE images (NOT normalized iris images)

• Our EYE images look like…
Results: Validation (I)

- The reconstruction algorithm is validated → very high performance
- Unrealistically high security scenario → 75% of breaking the system
- More likely to break the original sample, than other real sample from the same user.
- Still, very high probability of breaking other real samples.
- For the most likely attacking scenario → 92% SR
- More than one reconstruction → 30% SR increase
- Yet another new vulnerability → black circle + white background = Eye image

<table>
<thead>
<tr>
<th>FAR</th>
<th>$\text{SR}_{1a}$</th>
<th>$\text{SR}_{1b}$</th>
<th>$\text{SR}_{2a}$</th>
<th>$\text{SR}_{2b}$</th>
<th>$\text{SR}_3$</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1%</td>
<td>81.2</td>
<td>66.7</td>
<td>96.2</td>
<td>92.8</td>
<td>96.7</td>
<td>86.7</td>
</tr>
<tr>
<td>0.05%</td>
<td>79.2</td>
<td>63.4</td>
<td>96.2</td>
<td>91.4</td>
<td>95.2</td>
<td>85.1</td>
</tr>
<tr>
<td>0.01%</td>
<td>77.3</td>
<td>60.9</td>
<td>95.2</td>
<td>90.9</td>
<td>93.8</td>
<td>83.6</td>
</tr>
<tr>
<td>0.0001%</td>
<td>69.0</td>
<td>49.1</td>
<td>92.8</td>
<td>82.8</td>
<td>82.9</td>
<td>75.3</td>
</tr>
</tbody>
</table>
6. Results: Appearance
Results: Appearance (1)

What about humans? Are they deceived by the reconstructed irises?
Results: Appearance (II)

- 100 irises (50 real / 50 synthetic)
- 25 non-experts / 15 experts

- Rank: 0 (fully synthetic) - 10 (fully real)
- 15 minutes max.

Over 37% of misclassified irises by non-experts → real-like appearance
FSR/FRR very close → not easier to distinguish one class over the other
Average scoring very close → idem
Not so easy with experts, but still possible
Results: Appearance (III)

- Would you like to try?

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
6. Conclusions
Conclusions

- Can iris images be reconstructed from the iriscode? → YES!
- Can this reconstructed images be used to successfully break iris recognition systems? → YES!
- Is it more dangerous to be able to reconstruct SEVERAL iris images? → YES!
- Should iris recognition systems check that what is being presented is really an eye image? → YES!
- Do the iris reconstructed images look real to the average human? → YES!

- To sum up… do we need to develop specific countermeasures for this new vulnerability? → YES!
  - Cryptography for the templates.
  - Liveness detection for the systems.
Results: Appearance (III)

- Would you like to try?