IFPC 2020

Signal-level fusion for indexing and retrieval of facial biometric data

Pawel Drozdowski

da/sec - Biometrics and Internet Security Research Group, Hochschule Darmstadt

Online, 2020-10-28
Speaker

Dr. Pawel Drozdowski

Senior researcher:
Hochschule Darmstadt

Website:
https://dasec.h-da.de/staff/pawel-drozdowski/

Contact:
pawel.drozdowski@h-da.de
Contents

Motivation

Background

Proposed system

Results

Conclusion
Motivation

Biometric operation modes

Verification

- A biometric claim to an identity is made
- A 1:1 comparison is performed to reach a decision

⇒ Computationally trivial

![Diagram showing the process of biometric verification]

Claimed identity → ID → Comparator → Enrolment DB

Score: 0.95
Decision: Verified

Probe

Reference
Motivation

Biometric operation modes

Identification

- There is no biometric claim
- The decision has to be reached using the biometric data alone

⇒ Computationally expensive (in worst case, exhaustive search)
Motivation

Biometric identification

Computational workload

▶ A system from one of market leaders: 35 million comparisons per second (face) on a high-end computer (i7 processor, 16GB RAM)

<table>
<thead>
<tr>
<th>Location</th>
<th>Population (millions)</th>
<th>1:N time</th>
<th>N:N time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Berlin</td>
<td>3.75</td>
<td>107 ms</td>
<td>3.5 days</td>
</tr>
<tr>
<td>Germany</td>
<td>83</td>
<td>2.37 s</td>
<td>6 years</td>
</tr>
<tr>
<td>EU</td>
<td>512</td>
<td>14.62 s</td>
<td>237 years</td>
</tr>
<tr>
<td>India</td>
<td>1339</td>
<td>38.25 s</td>
<td>1624 years</td>
</tr>
<tr>
<td>World</td>
<td>7800</td>
<td>3.43 m</td>
<td>55121 years</td>
</tr>
</tbody>
</table>

▶ A system for specialised hardware (Multicore Xeon processors, 512 GB RAM): 1.2 billion comparisons per second (face) – ~35 times faster

▶ Infrastructure: more computers/servers to distribute the computations

⇒ High monetary costs (hardware itself, maintenance, and software licenses)
Software- or Hardware-based acceleration speeds up the transactions, but does not yield computational workload reduction

Background

Signal-level fusion

Morphing

- By using image morphing methods, it is possible to create biometric samples which contain biometric information from multiple distinct data subjects. A typical morphing process includes:
  1. Facial landmark detection and triangulation in two or more images
  2. Landmark averaging to a single set of landmarks
  3. Image warping and alpha blending

- Two or more subjects can be morphed together

⇒ An attack vector against automated systems and human experts
Background

Morphing

Why does it work?

![Graph showing comparison scores for mated, non-mated, 2-morph, 4-morph, and 8-morph distributions.](image-url)
Proposed system

The idea

Benefiting from properties of morphed images

- Using morphing to facilitate computational workload reduction
  ⇒ A **vulnerability** is turned into an **asset**

- A two-stage biometric identification system
  1. Signal-level fusion at the first level
     ⇒ Morphs of 2, 4, or 8 subjects
  2. Pre-selection of most promising candidates
     ⇒ Normal comparisons at the second level

- Possible extension to a multi-stage system
  ⇒ In theory, possibility to **reduce computational workload and maintain biometric performance**

⇒ Promising results in a proof-of-concept experimental evaluation

Proposed system

Indexing and retrieval

Overview

Comparison and decision subsystems

1. Comparison with morphs
   \[ \frac{N}{n} \text{ comparison scores: } [0.15 \ 0.35 \ldots \ 0.75] \]

   \( k \) best candidates:

2. Comparison with references
   \[ \frac{k \times n}{n} \text{ comparison scores: } [0.10 \ldots 0.95] \]

   Best candidate:

   Decision: Identified, subject 8

Data storage subsystem

- Enrolment DB
- \( N \) references
- Pair assignment
- Signal-level fusion
- Metadata/statistics

Probe

Unknown identity

?
Proposed system

Indexing
Pair assignment

How to decide **who should get morphed with whom?**
⇒ A combinatorial optimisation problem

Choose a **global or local optimisation?**
⇒ Assignment problem-based definition
⇒ Hungarian algorithm

How to define the **cost function?**
⇒ Random
⇒ Soft-biometrics
⇒ Similarity-score
Results

Evaluation

Experimental setup

- 1024 enrolled subjects (ICAO compliant, FERET + FRGC)
- 4 face recognition systems (OSS and COTS)
- 4 morphing algorithms (OSS and COTS)
- 3 strategies for selection of subjects to be morphed (random, soft-biometric, similarity-score)
- 3 numbers of samples contributing to a morph (2, 4, 8)
- A range of sizes of the pre-selected subset
Results

Evaluation

Why is the computational workload reduced?

- For some parameter configurations $W_{\text{proposed}} < W_{\text{baseline}}$

- $W_{\text{two-stage}} = \frac{N}{n} + k \times n$

  $W_{\text{multi-stage}} = \frac{N}{n_1} + \sum_{l=1}^{\log_2 n_1} 2k_l$

2-stage system

Multi-stage system
Results

Evaluation

Effects of pair assignment strategies

⇒ Mated-morph score distributions significantly shifted towards the mated score distribution as a result of the intelligent pairing
Results

Evaluation

CMCs

Random pairing

Soft-biometric pairing

Similarity-score pairing
Results

Evaluation

Biometric performance and computational workload

<table>
<thead>
<tr>
<th>Recognition system</th>
<th>Computational workload at</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>95% HR</td>
</tr>
<tr>
<td>COTS</td>
<td>18.75%</td>
</tr>
<tr>
<td>OSS</td>
<td>25.78%</td>
</tr>
</tbody>
</table>

⇒ Intelligent pairing methods vastly outperforming random pair assignment. Best results with similarity-score-based pairing method

⇒ Better results with COTS-based recognition, especially at 100% HR

⇒ For some configurations, the computational workload is reduced, while biometric performance (false-negative errors) is maintained

⇒ Pre-selection causes the false positive errors to remain unchanged (worst case) or be reduced (best case)
Conclusion

Proposed system

Summary

- Signal-level fusion-based approach
- Computational workload reduced significantly
- Biometric performance not impaired
- Better results with COTS-based face recognition system
- High impact of the selection of subjects to be morphed on the results of the scheme
Conclusion

General take-away points

Biometric identification systems

- Biometric identification is much more challenging than biometric verification
  - Biometric performance
  - Computational workload

- Increasing computational requirements of the operational systems
  - Growing size, number, and scope of the deployments
  - Direct links to monetary costs and usability

⇒ Computational workload reduction methods
Research opportunities
In Germany and Norway (or remotely)

Interested in working with biometrics or information security?

- Visit websites: https://dasec.h-da.de/ and https://www.ntnu.edu/nbl/
  - Ph.D. positions
  - Internships
  - B.Sc. and M.Sc. theses proposals with the possibility of a stipend
  - Student assistant positions
  - Collaborations with academia, industry, and governmental agencies

- Contact Prof. Dr. Christoph Busch at christoph.busch@h-da.de

Thank you for your attention!

Questions?