Forensic Biometrics: quantifying forensic evidence from biometric traces

Didier Meuwly
Principal scientist, NFI
Forensic biometrics chair, UT
Halmstad - Sweden, 16th June 2016
OUTLINE

1. Netherlands Forensic Institute

2. Definitions

3. Biometric modalities

4. Forensic biometric applications
   Forensic investigation (FP) – Forensic evaluation (Body height - Face - Speech)

5. Validation (FP)

6. Conclusion
1. NETHERLANDS FORENSIC INSTITUTE

@NFI.nl

Focuses on criminalistics

Makes use of science for criminal investigation and evidence evaluation

An integrated model

- Products and services
- R&D Innovation
- Education & Training
ORGANISATION

- Digitale Technologie en Biometrie: 1500
- Front Office: 200
- Forensisch Chemisch Onderzoek: 6100
- Humane Biologische Sporen: 42500
- Microsporen: 1400
- Medisch Forensisch Onderzoek: 4700
# MAIN TASKS

## Products & services
- R&D
- Education and training

## Netherlands
<table>
<thead>
<tr>
<th>Products &amp; services</th>
<th>Netherlands</th>
<th>International</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Prosecution</td>
<td>United nations</td>
</tr>
<tr>
<td></td>
<td>Law enforcement</td>
<td>International criminal court</td>
</tr>
<tr>
<td></td>
<td>Defense ministry</td>
<td>International forensic institutes</td>
</tr>
<tr>
<td></td>
<td>Secret services</td>
<td>International Criminal Tribunal Former Yugolavia</td>
</tr>
<tr>
<td></td>
<td>Antitrust</td>
<td>Special Tribunal Lebanon</td>
</tr>
<tr>
<td></td>
<td>Customs</td>
<td>Court special Sierra Leone</td>
</tr>
<tr>
<td></td>
<td>Tax services</td>
<td>International agency for Nuclear Energy</td>
</tr>
<tr>
<td></td>
<td>Immigrations services</td>
<td>International organization for migration</td>
</tr>
</tbody>
</table>
MAIN TASKS

Products & services

R&D

Education and training

Real time, on site chemical identification
Forensic recognition and individualisation
Big Data and Intelligent Data Analysis
From source to activity
Advancing Forensic Medicine
CSI Innovations
CBRN Forensics
Emerging technologies
Process optimization in Forensic Investigation
MAIN TASKS

Products & services
NFI and other practitioners

Dutch and international law enforcement practitioners (police and legal)

University students (MSc, PhD) Amsterdam, Leiden, Twente

R&D

Education and training
European Network of Forensic Science Institutes (ENFSI)
2. DEFINITIONS

Forensic science

Forensic science has an object: the study of crime and its traces.

Traces are the most elementary information that result from crime.

They are silent witnesses that need to be detected, seen, and understood to make reasonable inferences about criminal phenomena, investigation or demonstration for intelligence, investigation and court purposes.
DEFINITIONS

**Biometric recognition**

Human-based and automated recognition of individuals based on their physical/biological/chemical and behavioural characteristics. Biometric recognition allows to differentiate human beings and to recognise them to a certain degree depending on the:

- modality
- quality of the data
- application

DEFINITIONS

**Forensic biometrics**

Application of human-based and computerised biometric recognition methods and technologies to:

**Analyse** biometric traces and reference specimens and to answer questions:

- about the origin of a biometric trace (source level inference)
- on the activity that leads to a biometric trace (activity level inference)
- and if this activity is constitutive of an offence (offence level inference)
3. BIOMETRIC MODALITIES
BIOMETRIC MODALITIES

- genetic profile
- face
- keystroke
- size and shape of hard tissues
- iris

- papillary ridges
- voice
- touch screen dynamics
- thermo-grams
- retinal

- manuscript
- gait
- biometric authentication
- vein

- signature
- body measurements
- teeth
- results of surgery
- vein patterns

- earmark
- geometry of the body
- tattoo – scar
- results of dentistry

- bitemark
- extremities
- body odour

- lipmark
- 
PHYSICAL - BEHAVIOURAL

- genetic profile
- papillary ridges
- manuscript
- signature
- earmark
- bitemark
- lipmark
- face
- voice
- gait
- body measurements
- geometry of the body extremities
- keystroke
- touch screen dynamics
- biometric authentication
- size and shape of hard tissues
- teeth
- results of surgery
- results of dentistry
- tattoo – scar
- body odour
- iris
- thermo-grams
- retinal vein pattern
- other vein patterns
Didier Meuwly
UNIVERSITY OF TWENTE. EEMCS

Forensic Biometrics: quantifying forensic evidence from biometric traces
16.06.2016

PHYSICALLY COLLECTED - DIGITISED - DIGITAL

- genetic profile
- papillary ridges
- manuscript
- signature
- earmark
- bitemark
- lipmark
- face
- voice
- gait
- body measurements
- geometry of the body extremities
- keystroke
- touch screen dynamics
- biometric authentication
- size and shape of hard tissues
- teeth
- results of surgery
- results of dentistry
- iris
- thermo-grams
- retinal vein pattern
- other vein patterns

body odour

tatoo – scar
<table>
<thead>
<tr>
<th>Genetic Profile</th>
<th>Face</th>
<th>Keystroke</th>
<th>Size and Shape of Hard Tissues</th>
<th>Iris</th>
</tr>
</thead>
<tbody>
<tr>
<td>Papillary Ridges</td>
<td>Voice</td>
<td>Touch Screen Dynamics</td>
<td>Thermo-grams</td>
<td>Retinal Vein Pattern</td>
</tr>
<tr>
<td>Manuscript</td>
<td>Gait</td>
<td>Biometric Authentication</td>
<td>Teeth</td>
<td>Other Vein Patterns</td>
</tr>
<tr>
<td>Signature</td>
<td>Body Measurements</td>
<td>Tatoo – Scar</td>
<td>Results of Surgery</td>
<td></td>
</tr>
<tr>
<td>Earmark</td>
<td>Geometry of the Body Extemities</td>
<td>Body Odour</td>
<td>Results of Dentistry</td>
<td></td>
</tr>
</tbody>
</table>
Core biometric technology

**Feature extraction**

<table>
<thead>
<tr>
<th>Reference specimen</th>
<th>Trace specimen</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Object (R)</strong></td>
<td><strong>Object (T)</strong></td>
</tr>
<tr>
<td>Recording</td>
<td>Recording</td>
</tr>
<tr>
<td><strong>Feature (r)</strong></td>
<td><strong>Feature (t)</strong></td>
</tr>
<tr>
<td>Template</td>
<td>Test</td>
</tr>
</tbody>
</table>
## FORENSIC PROCESS

<table>
<thead>
<tr>
<th>Crime-scene</th>
<th>Analysis</th>
<th>Interpretation</th>
<th>Reporting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial issues</td>
<td><strong>Case assessment</strong></td>
<td>ID verification</td>
<td><strong>Decision</strong></td>
</tr>
<tr>
<td>What, where, when, how, who</td>
<td>- Reporting</td>
<td>Identification</td>
<td>(1:1, 1:N, 1:N+1)</td>
</tr>
<tr>
<td><strong>Scene investigation</strong></td>
<td>- Interpretation</td>
<td>Intelligence</td>
<td><strong>Cases links</strong></td>
</tr>
<tr>
<td>- Refine scenarios and hypotheses</td>
<td>- Analysis</td>
<td><em>Intelligence</em></td>
<td>(M:N links)</td>
</tr>
<tr>
<td>- Trace recovery</td>
<td><strong>Analytical methods</strong></td>
<td><strong>Investigation</strong></td>
<td><strong>Rank list</strong></td>
</tr>
<tr>
<td><strong>Crime scene assessment</strong></td>
<td>- Indicative methods</td>
<td><strong>Evaluation</strong></td>
<td>(M:N sources)</td>
</tr>
<tr>
<td>- Description</td>
<td>- Human-based</td>
<td>- Human-based</td>
<td><strong>Strength of evidence</strong></td>
</tr>
<tr>
<td>- Requests</td>
<td>- <strong>Computer-assisted</strong></td>
<td>- Computer-based</td>
<td>(Likelihood ratio)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Analytical methods**

- Indicative methods
- Human-based
- **Computer-assisted**

**Comparative methods**

- Human-based
- **Computer-based**

**Scene investigation**

- Refine scenarios and hypotheses
- Trace recovery

**Crime scene assessment**

- Description
- Requests
## 4. FORENSIC BIOMETRIC APPLICATIONS

<table>
<thead>
<tr>
<th>Forensic scenario</th>
<th>Comparison ... with ...</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID verification</td>
<td>Reference from suspect</td>
<td>Reference database</td>
</tr>
<tr>
<td>Intelligence</td>
<td>Trace</td>
<td>Trace</td>
</tr>
<tr>
<td>Investigation</td>
<td>Trace</td>
<td>Reference database</td>
</tr>
<tr>
<td>Evaluation</td>
<td>Trace</td>
<td>Reference from suspect</td>
</tr>
</tbody>
</table>
FORENSIC BIOMETRIC APPLICATIONS

Decision

Identity verification: criminal justice ID management
Identification: Disaster Victim Identification (DVI)
FORENSIC BIOMETRIC APPLICATIONS

Decision

- Identity verification: criminal justice ID management
- Identification: Disaster Victim Identification (DVI)

Selection

- Link cases from biometric traces: forensic intelligence
- Generate short lists of candidates: forensic investigation
FORENSIC BIOMETRIC APPLICATIONS

**Decision**
- Identity verification: criminal justice ID management
- Identification: Disaster Victim Identification (DVI)

**Selection**
- Link cases from biometric traces: forensic intelligence
- Generate short lists of candidates: forensic investigation

**Description**
- Describe the evidential value of the biometric evidence: forensic evaluation
QUESTIONS AND APPLICATIONS

Every biometric process relies primarily on enrolment.
Forensic Biometrics: quantifying forensic evidence from biometric traces
It is impossible for a criminal to act without leaving a trace, considering the intensity of the criminal activity.

28 April 1906

Edmond Locard
QUESTIONS AND APPLICATIONS

Who is a potential suspect? – Forensic investigation (M:N)
FORENSIC INVESTIGATION

Computer-assisted systems

database of N references

references collected

Automated Fingerprint Identification System
Computer-assisted systems

- Trace recovered
- Database of N references
- N pairwise comparisons
- Selection of M candidates from N references
- References collected
- Automated Fingerprint Identification System
Example of scalability of AFIS systems

2004 March 11th Madrid Atocha train station bombing (11-M) – 191 victims killed, 1800 victims wounded
FORENSIC INVESTIGATION

INTERPOL: 183 (196 – 13) countries, access to 10’000’000’000 fingerprints*


*Law enforcement and immigration ten print cards databases. Personal communication, M. Brancheflower, head of Fingerprint Unit, Interpol, April 2nd 2015
Enormous potential population
Madrid – Seattle (8’500km)

Enormous searching database
10’000’000’000 fingerprints

Low quantity of information in the mark
Single / double imposition
Large amount of distortion

Weak link with the suspects
Weak strength of evidence
FORENSIC INVESTIGATION

Look alike?

Sound alike?

Fingerprints: courtesy from Anko Lubach
Face images: diverses sources
Audio recording: courtesy from Hermann Kunzel

Forensic Biometrics: quantifying forensic evidence from biometric traces

16.06.2016
QUESTIONS AND APPLICATIONS

How strong is the evidence? – Forensic evaluation

**Defence hypothesis (Hd)**

The biometric trace originates from another person of the population

**Prosecution hypothesis (Hp)**

The biometric trace originates from the defendant
### FORENSIC EVALUATION

Assign the strength of evidence (Likelihood Ratio)

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hp</strong></td>
<td>Prosecution hypothesis</td>
</tr>
<tr>
<td><strong>Hd</strong></td>
<td>Defense hypothesis</td>
</tr>
<tr>
<td><strong>E</strong></td>
<td>Evidence</td>
</tr>
<tr>
<td><strong>I</strong></td>
<td>Background information</td>
</tr>
<tr>
<td>**Pr (E</td>
<td>Hp, I)**</td>
</tr>
<tr>
<td>**Pr (E</td>
<td>Hd, I)**</td>
</tr>
<tr>
<td><strong>LR</strong></td>
<td>Likelihood Ratio, strength of evidence metric</td>
</tr>
</tbody>
</table>
**FORENSIC EVALUATION**

Assign the strength of evidence (Likelihood Ratio)

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hp</strong></td>
<td>Prosecution hypothesis</td>
</tr>
<tr>
<td><strong>Hd</strong></td>
<td>Defense hypothesis</td>
</tr>
<tr>
<td><strong>Score (s)</strong></td>
<td>Evidence</td>
</tr>
<tr>
<td><strong>I</strong></td>
<td>Background information</td>
</tr>
<tr>
<td>**Pr (s</td>
<td>Hp, I)**</td>
</tr>
<tr>
<td>**Pr (s</td>
<td>Hd, I)**</td>
</tr>
<tr>
<td><strong>LR</strong></td>
<td>Likelihood Ratio, strength of evidence metric</td>
</tr>
</tbody>
</table>
FORENSIC EVALUATION

Assign the strength of evidence (Likelihood Ratio)

Duty of the forensic evaluator

Jules Henri Poincaré (1854-1912)

Appointed to review the Dreyfus case

“... in the impossibility for us [forensic evaluators] of knowing the prior probability, we cannot say: this coincidence proves that the odds of it being a forgery have this or that value. We can only say, by observing the coincidence: the odds become this much larger than before the observation.” Darboux, Appell, Poincaré
Human body height

Observations (E)

A. 181cm ± 10cm
B. 181cm ± 5cm
C. 181cm ± 2.5cm
D. 181cm ± 1cm

Relevant population

A. Worldwide
B. Netherlands

Defendant: 1.82cm ± 0.5cm
Feature-based likelihood ratio methods

- A. 181cm ± 10cm
- B. 181cm ± 5cm
- C. 181cm ± 2.5cm
- D. 181cm ± 1cm

Reference features

182cm ± 0.5cm

Defendant

Trace features

Recovered

Likelihood ratio method

Database of traces

Database of references

Likelihood ratio

Forensic Biometric Evaluation
FORENSIC BIOMETRIC EVALUATION

- Female - World
- Male - World
- Female - NL
- Male - NL

0.1

0.05

0

Female - World
Male - World
Female - NL
Male - NL
# Forensic Biometric Evaluation

**Assign the strength of evidence (Likelihood Ratio)**

<table>
<thead>
<tr>
<th>Observations</th>
<th>A.</th>
<th>B.</th>
<th>C.</th>
<th>D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>[cm]</td>
<td>181±10</td>
<td>181±5</td>
<td>181±2.5</td>
<td>181±1</td>
</tr>
</tbody>
</table>

**Relevant population**

<table>
<thead>
<tr>
<th>Worldwide</th>
<th>Relevant Population</th>
<th>Female</th>
<th>Male</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>160.2 ± 4.9</td>
<td>173.7 ± 5.8</td>
<td></td>
</tr>
<tr>
<td>Netherlands</td>
<td>170.7 ± 5.3</td>
<td>183.8 ± 6.1</td>
<td></td>
</tr>
</tbody>
</table>

Source: [wikipedia.org (2015)](https://wikipedia.org)
Assign the strength of evidence

**Human Body Height (cm)**

<table>
<thead>
<tr>
<th>Strength of Evidence (Log LR)</th>
<th>1 cm</th>
<th>2.5 cm</th>
<th>5 cm</th>
<th>10 cm</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. LR ≅ 32</td>
<td>♂</td>
<td>♂</td>
<td>♂</td>
<td>♂</td>
</tr>
<tr>
<td>B. LR ≅ 12</td>
<td>♂</td>
<td>♂</td>
<td>♂</td>
<td>♂</td>
</tr>
<tr>
<td>C. LR ≅ 6</td>
<td>♂</td>
<td>♂</td>
<td>♂</td>
<td>♂</td>
</tr>
<tr>
<td>D. LR ≅ 3</td>
<td>♂</td>
<td>♂</td>
<td>♂</td>
<td>♂</td>
</tr>
<tr>
<td>A. LR ≅ 17</td>
<td>♂</td>
<td>♂</td>
<td>♂</td>
<td>♂</td>
</tr>
<tr>
<td>B. LR ≅ 7</td>
<td>♂</td>
<td>♂</td>
<td>♂</td>
<td>♂</td>
</tr>
<tr>
<td>C. LR ≅ 3</td>
<td>♂</td>
<td>♂</td>
<td>♂</td>
<td>♂</td>
</tr>
<tr>
<td>D. LR ≅ 1.5</td>
<td>♂</td>
<td>♂</td>
<td>♂</td>
<td>♂</td>
</tr>
</tbody>
</table>
2016 March 22\textsuperscript{nd} Brussels Zaventem Airport bombing – 32 victims killed, 300 victims wounded
FORENSIC BIOMETRIC EVALUATION
Brussels attacks: Freelance journalist believed to be '3rd bomber' faces terror charges

By Ian Johnston
March 26, 2016 15:45 GMT
FORENSIC BIOMETRIC EVALUATION
FORENSIC BIOMETRIC EVALUATION
FORENSIC BIOMETRIC EVALUATION
FORENSIC BIOMETRIC EVALUATION

[ENFSI GUIDELINE FOR EVALUATIVE REPORTING IN FORENSIC SCIENCE]

Strengthening the Evaluation of Forensic Results across Europe (STEOFRAE)

European Network of Forensic Science Institutes

With the financial support of the Prevention of and Fight against Crime Programme of the European Union European Commission - Directorate-General Justice, Freedom and Security

A project funded by the EU ISEC 2010 Agreement Number: HOME/2010/SEC/0/000001759

Outline of the report

Evaluative Statement

General information about an evaluative statement from the laboratory

The findings reported in an evaluative statement are those resulting from the examination and analysis made. These findings are normally evaluated against two propositions: the first proposition (based on the issue, as formulated by the mandating authority) and an alternative proposition usually provided by the defendant.

During the evaluation, probabilities of the findings are assigned assuming in turn that each of the two propositions is true. The ratio of these two probabilities forms the value of evidence and is reported as a graded conclusion in the laboratory's scale of conclusions (attached).

If new information comes about, or if other propositions are requested to be used for the evaluation, this may affect the conclusion.

A graded conclusion shall be considered as a factor that either strengthens (positive scale level), weakens (negative scale level) or leaves unaltered (scale level 0) the prior opinions on the two propositions (prior to the forensic investigation). The laboratory makes no judgment about how probable any of these propositions are.

In a case where a certain exclusion can be made, phrases like "is", "is not" and "can be excluded" are used instead of a graded conclusion.

FORENSIC BIOMETRIC EVALUATION

**Conclusion**
The findings from the examination support the view/proposition that the person indicated on the CCTV still (item1) is the person on the photos and images of items 2-4 (level +2), rather than that the individual is another adult male.

This evaluation is based on my understanding of the relevant circumstances as described above. If any of this information is incomplete or incorrect (in particular if the alternative changes), I will have to re-evaluate my findings.

<table>
<thead>
<tr>
<th>Numerical LR</th>
<th>Verbal LR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>no support</td>
</tr>
<tr>
<td>2 - 10</td>
<td>weak</td>
</tr>
<tr>
<td>10-10(^2)</td>
<td>moderate</td>
</tr>
<tr>
<td>10(^2)-10(^3)</td>
<td>moderately strong</td>
</tr>
<tr>
<td>10(^3)-10(^4)</td>
<td>strong</td>
</tr>
<tr>
<td>10(^4)-10(^6)</td>
<td>very strong</td>
</tr>
<tr>
<td>&gt; 10(^6)</td>
<td>extremely strong</td>
</tr>
</tbody>
</table>
FORENSIC BIOMETRIC EVALUATION

Predicting Performance of a Face Recognition System Based on Image Quality
Abhishek Dutta

FORENSIC BIOMETRIC EVALUATION

Performance of the Cognitec 2D FaceVACS SDK (Commercial Of The Shelf algorithm)

FORENSIC BIOMETRIC EVALUATION

Cognitec 2DFaceVACS SDK performance (Commercial Of The Shelf algorithm)

5. VALIDATION OF COMPUTER-BASED EVALUATION METHODS

<table>
<thead>
<tr>
<th>Crime-scene</th>
<th>Analysis</th>
<th>Interpretation</th>
<th>Reporting</th>
</tr>
</thead>
<tbody>
<tr>
<td>In</td>
<td>Case</td>
<td>Evaluation</td>
<td>Strenght of evidence</td>
</tr>
<tr>
<td>W</td>
<td>- Requests</td>
<td>- Human-based</td>
<td>iso 17025</td>
</tr>
<tr>
<td>h</td>
<td>- Interpretation</td>
<td>- Computer-based</td>
<td>iso 17025</td>
</tr>
<tr>
<td>who</td>
<td>- Analysis</td>
<td>- Refine scenarios and hypotheses</td>
<td>iso 17025</td>
</tr>
<tr>
<td>Scene</td>
<td>- Comparative methods</td>
<td>- Trace recovery</td>
<td>iso 17025</td>
</tr>
<tr>
<td>- Requests</td>
<td>- Human-based</td>
<td>- Refine scenarios and hypotheses</td>
<td>iso 17025</td>
</tr>
<tr>
<td>Crime</td>
<td>- Computer-based</td>
<td>- Trace recovery</td>
<td>iso 17025</td>
</tr>
<tr>
<td>as</td>
<td>- Analytical methods</td>
<td>- Refine scenarios and hypotheses</td>
<td>iso 17025</td>
</tr>
<tr>
<td>- Requests</td>
<td>- Human-based</td>
<td>- Trace recovery</td>
<td>iso 17025</td>
</tr>
<tr>
<td>- Requests</td>
<td>- Analytical methods</td>
<td>- Refine scenarios and hypotheses</td>
<td>iso 17025</td>
</tr>
</tbody>
</table>

Crime-scene assessment
- Description
- Requests

Analysis
- Comparative methods
- Human-based
- Computer-based

Interpretation
- Evaluation
- Human-based
- Computer-based

Reporting
- Strenght of evidence
Does the likelihood ratio method developed for fingermark forensic evaluation computes a correct strength of evidence?
VALIDATION OF COMPUTER-BASED EVALUATION METHODS

**Primary performance characteristics**

- Directly measure desirable properties of the LR
  - Accuracy – Discriminating power – Calibration

**Secondary performance characteristics**

- Require the primary ones and measure / present how the primary measures vary in different conditions (for instance quality of the data or quantity of information)
  - Robustness – Generalization – Coherence
VALIDATION OF COMPUTER-BASED EVALUATION METHODS

<table>
<thead>
<tr>
<th>Performance characteristics</th>
<th>Performance metrics</th>
<th>Graphical representation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td>$Cllr$, $EER$</td>
<td>$ECE - DET$ plots</td>
</tr>
<tr>
<td>Discrimination power</td>
<td>$Cllr_{\text{min}}$</td>
<td>$ECE_{\text{min}}$ plot</td>
</tr>
<tr>
<td>Calibration</td>
<td>$Cllr_{\text{cal}}$</td>
<td>Tippett plot</td>
</tr>
<tr>
<td><strong>Secondary</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Robustness</td>
<td>$Cllr$, $EER$</td>
<td>$ECE - DET -$ Tippett plots</td>
</tr>
<tr>
<td></td>
<td>LR range</td>
<td></td>
</tr>
<tr>
<td>Coherence</td>
<td>$Cllr$, $EER$</td>
<td>$ECE - DET -$ Tippett plots</td>
</tr>
<tr>
<td>Generalization</td>
<td>$Cllr$, $EER$</td>
<td>$ECE - DET$ plots</td>
</tr>
</tbody>
</table>
VALIDATION OF COMPUTER-BASED EVALUATION METHODS

Discrimination

DET

Detection Error Tradeoff curve

EER

Equal Error Rate
VALIDATION OF COMPUTER-BASED EVALUATION METHODS

**Discrimination**

**Tippett plots**

VALIDATION OF COMPUTER-BASED EVALUATION METHODS

Calibration

Tippett plots

VALIDATION OF COMPUTER-BASED EVALUATION METHODS

Cost log likelihood ratio ($C_{llr}$)

System 1 and 2 are more discriminant than system 3

System 1 produce more calibrated LR than system 2
VALIDATION OF COMPUTER-BASED EVALUATION METHODS

$C_{llr} - \text{Cost Log likelihood ratio} - C_{llr} = C_{llr_{\text{min}}} + C_{llr_{\text{cal}}}$

$E_{CE} - \text{Empirical Cross Entropy (C}_{llr}, \text{for Prior log}_{10} = 0)$

VALIDATION OF COMPUTER-BASED EVALUATION METHODS

VALIDATION OF COMPUTER-BASED EVALUATION METHODS

Coherence

<table>
<thead>
<tr>
<th>Minutiae configuration</th>
<th>EER</th>
<th>Discriminating power Calibrated Cllr_{min}</th>
<th>Accuracy Cllr</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>15.9</td>
<td>0.43</td>
<td>0.5</td>
</tr>
<tr>
<td>6</td>
<td>6.9</td>
<td>0.26</td>
<td>0.28</td>
</tr>
<tr>
<td>7</td>
<td>3.9</td>
<td>0.14</td>
<td>0.16</td>
</tr>
<tr>
<td>8</td>
<td>2.4</td>
<td>0.11</td>
<td>0.13</td>
</tr>
<tr>
<td>9</td>
<td>1.5</td>
<td>0.063</td>
<td>0.075</td>
</tr>
<tr>
<td>10</td>
<td>2.2</td>
<td>0.063</td>
<td>0.074</td>
</tr>
<tr>
<td>11</td>
<td>2.7</td>
<td>0.081</td>
<td>0.1</td>
</tr>
<tr>
<td>12</td>
<td>1.8</td>
<td>0.057</td>
<td>0.084</td>
</tr>
</tbody>
</table>

Coherence

VALIDATION OF COMPUTER-BASED EVALUATION METHODS

Limit of use

<table>
<thead>
<tr>
<th>Minutiae configuration</th>
<th>Prior log odds uncalibrated</th>
<th>Prior log odds calibrated</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>&lt;-10, 0.5&gt;</td>
<td>&lt;-10, 2.4&gt;</td>
</tr>
<tr>
<td>6</td>
<td>&lt;-10, 1.4&gt;</td>
<td>&lt;-10, 10&gt;</td>
</tr>
<tr>
<td>7</td>
<td>&lt;-10, 2.4&gt;</td>
<td>&lt;-10, 3.1&gt;</td>
</tr>
<tr>
<td>8</td>
<td>&lt;-10, 2.7&gt;</td>
<td>&lt;-10, 10&gt;</td>
</tr>
<tr>
<td>9</td>
<td>&lt;-10, 10&gt;</td>
<td>&lt;-10, 10&gt;</td>
</tr>
<tr>
<td>10</td>
<td>&lt;-10, 10&gt;</td>
<td>&lt;-10, 10&gt;</td>
</tr>
<tr>
<td>11</td>
<td>&lt;-10, 2.8&gt;</td>
<td>&lt;-10, 3.4&gt;</td>
</tr>
<tr>
<td>12</td>
<td>&lt;-10, 10&gt;</td>
<td>&lt;-10, 10&gt;</td>
</tr>
</tbody>
</table>
## RESULT OF THE VALIDATION OF FINGERMARK EVALUATION METHODS

<table>
<thead>
<tr>
<th>Performance characteristics</th>
<th>Performance metrics</th>
<th>Validation criteria</th>
<th>Experiments</th>
<th>Data</th>
<th>Results</th>
<th>Validation decision</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accuracy</td>
<td>CIlr</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✔</td>
</tr>
<tr>
<td>Discrimination power</td>
<td>CIlr_{min}, EER</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✔</td>
</tr>
<tr>
<td>Calibration</td>
<td>CIlr_{cal}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✔</td>
</tr>
<tr>
<td><strong>Secondary</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Robustness</td>
<td>CIlr, EER, LR range</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✔</td>
</tr>
<tr>
<td>Coherence</td>
<td>CIlr, EER</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✘</td>
</tr>
<tr>
<td>Generalization</td>
<td>CIlr, EER</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✔</td>
</tr>
</tbody>
</table>
FORENSIC BIOMETRIC EVALUATION

Effect of quality

![Graph showing the effect of quality on forensic biometric evaluation.](image)

Courtesy from E. Leitet & N. Appleby (2016) "Image Comparisons at NFC, Sweden" – Workshop on theory and implementation of numerical likelihood ratio methods – NFI 23-24.05
FORENSIC BIOMETRIC EVALUATION

General validation of ASR, durations I

HeatMap EER values duration experiments

Courtesy from D. van der Vloed (2016) "Implementation of ASR in casework" – Workshop on theory and implementation of numerical likelihood ratio methods – NFI 23-24.05
CONTRIBUTE TO AN ISO STANDARD

CONTRIBUTE TO AN ISO STANDARD

NEN, normalisatie en normen
NEN ondersteunt in Nederland het normalisatieproces. Als een partij zich tot NEN richt met de vraag om een afspraak tot stand te brengen, gaan wij aan de slag.

Include the methodology developed in:
Validation of Likelihood Ratio Methods used for Forensic Evidence Evaluation: Application in Forensic Fingerprints in the revision of the ISO/IEC standard 19795

JTC1 - Information technology
SC37 - Biometrics

ISO/I EC 19795-6:2012
Information technology -- Biometric performance testing and reporting -- Part 6: Testing methodologies for operational evaluation
CONTRIBUTE TO AN ISO STANDARD

Methodology and tools for the validation biometric methods for forensic evaluation and identification application
CONCLUSION

A forensic biometric evaluation methods needs to be:

Logical, Balanced, Robust, Validated

Testable and falsifiable

Provide a strength of evidence relative to the hypotheses

Provide reliable results in forensic conditions

Provide the range and limits of use of the method

The strength of evidence computed is:

Relative to the hypotheses tested

Conditional to the method used
Time for questions
ABSTRACT

Forensic biometrics: quantifying forensic evidence from biometric traces

Outline

The tutorial will begin with a short introduction of the Netherlands Forensic Institute (NFI), its tasks, its organisation, its requesters and the role of forensic biometrics within the Institute. Then it will concentrate on the definition of forensic biometrics, the description of the informative value of the different biometric modalities in a forensic context and cover the different forensic applications of biometric technology using operational examples. Then, the validation of forensic evaluation methods used to assess the strength of evidence will be presented in detail. Finally, the tutorial will conclude with a short overview of some current topics of research in forensic biometrics within the NFI.

Definitions

Forensic science has an object of study: the crime and its traces. Traces are the most elementary pieces of information that result from crime. They are silent witnesses that need to be detected, analysed, and understood to make reasonable inferences about criminal phenomena (forensic intelligence), as well as for forensic investigation and court purposes.

Biometric recognition is the human-based and computerised recognition of individuals, based on their physical/biological/chemical and behavioural characteristics. Computerised approaches consist of feature extraction and feature comparison algorithms, as well as methods for the inference of identity of source. Biometric recognition allows to differentiate between human beings and to recognise them to a certain degree, depending on the modality, application and quality of the data (trace and reference specimens).
ABSTRACT

Forensic biometrics is defined as the application of human-based and computer-assisted biometric recognition methods and technologies to analyse biometric traces and reference specimens, in order to answer questions about the origin of these traces (source level inference). The examination of biometric traces can also answer other forensically relevant questions, about the activity that led to a trace (activity level inference) and whether this activity is constitutive of a criminal offence (offence level inference).

Modalities

In a forensic context, traces like biological traces, fingermarks, earmarks or bitemarks are physically collected on crime scenes. Some others, like face and body images, voice recordings, gait recordings or fingerprints and iris scanned for authentication are digitised with capture devices. Finally, some traces only exist digitally, like keystroke- and touchscreen- dynamics. Some modalities provide traces within physical crimes, some others within cybercrime or post-mortem individualisation. Finally some modalities like retina vein patterns do not provide any forensic trace.

Applications

A short movie describing a robbery is presented as an example to introduce 3 forensic applications in which biometric recognition play a role: forensic intelligence, forensic investigation and forensic evaluation. Forensic intelligence consists in linking criminal cases together; it is introduced using the fingermark and DNA modalities as examples.

Forensic investigation consists of selecting shortlists of candidates that are potentially donors of traces in criminal cases; it is described in detail using the fingermark and face modalities as examples.
ABSTRACT

Forensic evaluation focuses on the description of the strength of evidence that an individual is the donor of a trace in a criminal case. The statistical methodology used to describe this strength of the evidence and assign likelihood ratios is explained using the body height (human-based approach) and the speaker recognition (computer-based approach) modality. Biometric recognition can also be used to reach a decision of identity verification of suspects or a decision of identification (closed-set or open-set) of victims in the context of disaster victim identification (DVI).

Validation

In the last decade, the forensic biometric research has developed computer-assisted methods for the analysis, comparison and evaluation of the evidence to support the forensic practitioners in their quest for more objective methods to report likelihood ratios. According to the EU council framework decision 2009/905/JHA, the forensic service providers carrying out laboratory activities like forensic evaluation of DNA and fingermarks/prints need to be accredited since 2015, for example under the ISO/IEC 17025:2005 standard – General requirements for the competence of testing and calibration laboratories. As a consequence of this EU decision the human-based and computer-assisted methods used for the forensic evaluation of fingermarks/prints need to be validated. Within forensic science, guidelines exist for the validation of the human-based methods used for forensic evaluation. They mainly focus on the education and the competence assessment of the practitioners and therefore are not suitable for the validation of computer-assisted methods.

Methods for the validation of computer-assisted methods have been developed more recently. They are being published and there is an incentive to integrate them in the ISO/IEC 19795:2012 standard – Information technology – Biometric performance testing and reporting – Part 6: Testing methodologies for operational evaluation. The validation strategy is using primary and secondary performance characteristics, identified as relevant to describe the performance and the limits of likelihood ratio methods and related performance metrics are used to measure them. The delicate question of setting validation criteria in a completely new context, in which no baseline exists, will also be discussed.
Didier Meuwly is born in 1968 in Fribourg, Switzerland. After a classical education (Latin/Philosophy), he educated as a criminalist and criminologist (1993) and obtained his PhD (2000) at the School of Forensic Science (IPS) of the University of Lausanne. Currently he shares his time between the Forensic Institute of the Ministry of Security and Justice of the Netherlands (Netherlands Forensic Institute), where he is a principal scientist, and the University of Twente, where he holds the chair of Forensic Biometrics. He specialises into the automation and validation of the probabilistic evaluation of forensic evidence, and more particularly of biometric traces. He was previously the leader of a project about the probabilistic evaluation of fingermark evidence, and responsible of the fingerprint section within the NFI. From 2002 to 2004, he worked as a senior forensic scientist within the R&D department of the Forensic Science Service (UK-FSS), at the time an executive agency of the British Home Office. From 1999 to 2002 he was responsible of the biometric research group of the IPS. He is a founding member of 2 working groups of the European Network of Forensic Science Institutes (ENFSI): The Forensic Speech and Audio Analysis Working Group (FSAAWG) in 1997 and the European Fingerprint Working Group (EFPWG) in 2000. He is still active within the EFWPG. He is also a member of the editorial board and a guest editor of Forensic Science International (FSI).