Feature-based calculation of likelihood ratios for forensic comparison of fired cartridge cases

*Nabanita Basu*¹

*Rachel S Bolton-King*¹,²

*Geoffrey Stewart Morrison*¹,³

¹ Forensic Data Science Laboratory, Aston University
² Department of Society, Crime & Environment, Staffordshire University
³ Forensic Evaluation Ltd
Acknowledgements

• This research was supported by Research England’s Expanding Excellence in England Fund as part of funding for the Aston Institute for Forensic Linguistics 2019–2023.

• Thanks to the organizations and individuals who donated the fired cartridge cases.

• Thanks to ScannBI Technology Europe GmbH for the loan of the Evofinder® imaging system.

Disclaimer

• All opinions expressed are those of the presenter and, unless explicitly stated otherwise, should not be construed as representing the policies or positions of any organizations with which the presenter is associated.
Contents

- Scenarios and hypotheses
- Database
- Preprocessing
- Feature extraction
- Statistical models
- Validation results
- Conclusion and future plans
Scenarios and Hypotheses
Scenario 1

- One or more firearms are fired at a crime scene and the cartridge cases are ejected.

- Crime-scene investigators later recover two fired cartridge cases.

- A forensic practitioner compares the two questioned-source cartridge cases with one another.

- The forensic practitioner draws an inference with respect to whether they were fired by the same firearm or not.
Hypotheses 1

• $H_s$: The two cartridge cases were fired by the same firearm.

• $H_d$: The two cartridge cases were fired by different firearms from the same population.
Scenario 2

- A firearm is fired at a crime scene and the cartridge case is ejected.

- Crime-scene investigators later recover the fired cartridge case.

- Police investigators seize a firearm from a suspect.

- A forensic practitioner fires multiple cartridges from the seized firearm and collects the ejected cartridge cases.
  
    - 3 cartridges
    - 9 cartridges
Scenario 2

- The forensic practitioner compares:
  - the fired cartridge case recovered from the crime scene
    (the questioned-source cartridge case)
  with
    - the cartridge cases fired from the suspect’s firearm
      (the known-source cartridge cases)
- The forensic practitioner draws an inference with respect to whether the questioned-source and known-source cartridge cases were fired by the same firearm or not.
Hypotheses 2

• $H_s$: The questioned-source cartridge case and the multiple known-source cartridge cases were fired by the same firearm.

• $H_d$: The questioned-source cartridge case and the multiple known-source cartridge cases were fired by different firearms from the same population.
Hypotheses

- Relevant population:

  - Semi-automatic pistols that:
    - fire 9 mm diameter centre-fire Luger-type ammunition
    - have hemispherical firing pins
    - have parallel breech-face marks
Database
Database

- Existing databases tend to have:
  - a very large number of fires from a small number of firearms
  - a small number of fires from a moderate number of firearms

- For training a likelihood-ratio model, we need:
  - a relatively large number of fires from a relatively large number of firearms of the same class
  - representing within-source and between-source variability
Database

- Database of 3D digital images of the bases of fired cartridge cases:
  - ~300 firearms from relevant population
    - 10 cartridges fired from each firearm
  - Aim was 1,000 firearms, but COVID

- Firearms were in the possession of a number of operational forensic laboratories, law-enforcement agencies, military units, and private individuals in Barbados, Canada, France, Germany, UK, and USA
Database

- Database of 3D digital images of the bases of fired cartridge cases:
  - Digitally imaged using Evofinder®
    - 3D surface topography
    - exported as a matrix of values $z(x,y)$ in x3p format
Database

- Example 3D digital image
Preprocessing
Preprocessing

• Segmentation:
  • Separation of the firing-pin impression and the breech-face region from the remainder of the image and from each other.

• Illumination correction:
  • Correction for non-uniformities in illumination. This can include planar-bias correction.

• Noise removal:
  • Removal of imaging artifacts.

• Registration:
  • Rotation and alignment.
Preprocessing

- Segmentation

(a)  (b)  (c)  (d)
Feature Extraction
Feature extraction

• Previous attempt to calculate likelihood ratios for fired cartridge cases have used similarity scores.
  
  • Similarity scores do not take account of typicality with respect to the relevant population.
  
  • They are not suitable for calculating likelihood ratios addressing hypotheses of interest in a case.
Feature extraction

- Extract fixed-length feature vectors:
  - One feature vector from each image
  - Suitable for calculating common-source likelihood ratios
Feature extraction

- Feature sets:
  - central moments
  - circle-moment invariants
  - Legendre moments
  - coefficients of Fourier series fitted to concentric circles
  - Zernike moments
Feature extraction

- Zernike moments

<table>
<thead>
<tr>
<th>n \ m</th>
<th>-4</th>
<th>-3</th>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th>+1</th>
<th>+2</th>
<th>+3</th>
<th>+4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></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>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></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>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Feature extraction

- Extracted from different segmentations:
  - firing-pin impression [66]
  - breech-face region [231]
  - whole region of interest (excluding flowback) [297]
  - firing-pin impression + breech-face region (feature concatenation) [297]
  - firing-pin impression + breech-face region (score-level fusion) [66]+[231]
  - whole region of interest (including flowback) [297]

[number of Zerinke moment magnitude and phase features extracted]
Statistical Models
Statistical models

- Based on backend pipeline used in forensic voice comparison.

Statistical models

- Dimension reduction methods:
  - Principal component analysis
  - Linear discriminant functions

- Resulting number of dimensions:
  - 10 for firing-pin impression
  - 20 for breech-face region
  - 30 for whole region of interest
Statistical models

- Calculation of uncalibrated likelihood ratio:

- Common-source likelihood-ratio model:

\[
\lambda = \frac{f \left( \left[ \nu_q \right] \left| \begin{bmatrix} \hat{\mu}_r \\ \hat{\mu}_r \end{bmatrix}, \begin{bmatrix} \hat{\Sigma}_w + \hat{\Sigma}_b \\ \hat{\Sigma}_b \\ \hat{\Sigma}_b \end{bmatrix} \right) \right)}{f \left( \nu_q \left| \hat{\mu}_r, \hat{\Sigma}_w + \hat{\Sigma}_b \right) f \left( \nu_k \left| \hat{\mu}_r, \hat{\Sigma}_w + \hat{\Sigma}_b \right) \right)}
\]

In the automatic-speaker-recognition literature this model is called the two-covariance version of “probabilistic linear discriminant analysis” (PLDA). In Aitken & Lucy (2004), it is called the “multivariate normal (MVN) procedure.”
Statistical models

- Calibration (and score-level fusion):
  - Regularized logistic regression
    - regularization weight equivalent to one source (i.e., one firearm)

\[
\log(\Lambda) = \beta_0 + \beta_1 \log(\lambda)
\]

\[
\log(\Lambda) = \beta_0 + \beta_1 \log(\lambda_1) + \beta_2 \log(\lambda_2)
\]

- Cross-validation
  - leave-one-source-out / leave-two-sources-out
Statistical models

• Models trained:

  • **Scenario 1:**
    • 1 questioned-source v 1 questioned-source

  • **Scenario 2:**
    • 1 questioned-source v 3 known-source
    • 1 questioned-source v 9 known-source
Statistical models

- Data split:
  - $\frac{2}{3}$ training (dimension reduction and uncalibrated likelihood ratio model)
    - ~200 firearms
  - $\frac{1}{3}$ calibration & validation
    - ~100 firearms
Validation Results
Validation results

• Validation metric and validation graphic commonly used in forensic voice comparison:

  • log-likelihood-ratio cost ($C_{llr}$)

  • Tippett plot

Validation results

- $C_{llr}$
- **Scenario 1:** 1 questioned-source v 1 questioned-source

<table>
<thead>
<tr>
<th>Feature Set</th>
<th>Segmented Region</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>whole region of interest</td>
<td>breech face</td>
<td>firing pin</td>
<td>breech face + firing pin</td>
</tr>
<tr>
<td></td>
<td>including flowback</td>
<td>excluding flowback</td>
<td>feature concatenation</td>
<td>score-level fusion</td>
</tr>
<tr>
<td>Zernike moments magnitude &amp; phase</td>
<td>0.52</td>
<td>0.65</td>
<td>0.69</td>
<td>0.84</td>
</tr>
</tbody>
</table>
Validation results

- $C_{llr}$
- **Scenario 2**: 1 questioned-source v 3 known-source

<table>
<thead>
<tr>
<th>Feature Set</th>
<th>Segmented Region</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>whole region of interest</td>
</tr>
<tr>
<td></td>
<td>breech face</td>
</tr>
<tr>
<td></td>
<td>firing pin</td>
</tr>
<tr>
<td></td>
<td>breech face + firing pin</td>
</tr>
<tr>
<td>including flowback</td>
<td></td>
</tr>
<tr>
<td>excluding flowback</td>
<td></td>
</tr>
</tbody>
</table>

| Zernike moments     | 0.38                                  |
| magnitude & phase   | 0.50                                  |
|                     | 0.55                                  |
|                     | 0.73                                  |
|                     | 0.45                                  |
|                     | 0.48                                  |
Validation results

- $C_{llr}$
- **Scenario 2: 1 questioned-source v 9 known-source**

<table>
<thead>
<tr>
<th>Feature Set</th>
<th>Segmented Region</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>whole region of interest</td>
</tr>
<tr>
<td></td>
<td>including flowback</td>
</tr>
<tr>
<td>Zernike moments magnitude &amp; phase</td>
<td>0.35</td>
</tr>
</tbody>
</table>
Validation results

• Tippett plot

• Scenario 1:
  • 1 questioned-source vs 1 known-source
  • $C_{llr} = 0.52$
Validation results

- Tippett plot

- Scenario 2:
  - 1 questioned-source vs 3 known-source
  - $C_{\text{llr}} = 0.38$
Validation results

- Tippett plot

- Scenario 2:
  - 1 questioned-source v 9 known-source
  - $C_{llr} = 0.35$
Validation results

- Preprocessing **including rotation** versus **not including rotation**:
  
  - Zernike moment magnitude & phase features extracted from the whole region of interest including flowback
  
  - $C_{llr}$ values less than 1% different
Conclusion and Future Plans
Conclusion

• Best feature set:
  • Zernike moment magnitude & phase

• Best segmentation:
  • whole region of interest including flowback

• Rotation not necessary

• In Scenario 2:
  • 3 is probably a sufficient number of cartridges to fire from the seized firearm
Future plans

• Try to improve performance

  • Expand database

  • DNN embeddings
Thank You

Preprint, Data, Software:
http://firearms.forensic-data-science.net/