Quantitative Firearms and Toolmark Analysis: New Developments and Software

Petraco Group
City University of New York, John Jay College
Outline

• 3D toolmark data, pre-processing and feature extraction: \texttt{x3pr, feature2}

• The statistics:
  • Identification Error Rates
  • “Match” confidence estimate from Conformal Prediction Theory: \texttt{cptID}
  • “Match” probability estimates from Empirical Bayes: \texttt{fdrID}
  • “Match” probability estimates from CMS data and Bayesian Networks
Data Acquisition For Toolmarks

Confocal Microscope

Focus Variation Microscope
Screwdriver Striation Patterns in Lead

2D profiles
3D surfaces (interactive)
Bullets

Bullet base, 9mm Ruger Barrel
Toolmark Surface Data

- A growing database $^{Zheng}$:

- Put in your two cents: OpenFMC$^{Lillien}$
Toolmark Surface Data

- Standardizing file format: .x3p

x3p C++/ Petraco library, WindowsBrubaker

x3pr for, Any OS
Pre-processing Surface Data

- 3D tool mark data usually needs (a lot of…) preprocessing
  - We use a combination of R and C++/OpenCV (via Rcpp) = feature2

Possibly fill holes
Preprocessing Surface Data

- 3D tool mark data usually needs (a lot of…) preprocessing
  - In feature2:

Possibly remove “long range” behavior (leveling, form removal)

Crop out areas of interest

Bandpass filters via:

waveslim Whitcher
Good Features are the Key!

- We need a tool mark feature set that is:
  - Large in number
  - (possibly) transnationally invariant
  - (possibly) rotationally invariant
  - Mostly statistically independent
  - DISCRIMINATORY!
Aperture primer shear on a 9mm cartridge case fired from a Glock 19

Toolmark Features

Mean total profile:

Mean “waviness” profile:

Mean “roughness” profile:
FAST-Consecutive Matching Striae (CMS)-Space

Biasotti-Murdock dictionary: “Closest Match Ref Set”

Repp and parallelR-core packages are great for quick and easy speedups

Database/queries

• Find best matching “word” in query to each “dictionary word”
• Similarity metric is arbitrary
• We used ccf and dtw
• Cut-offs should be introduced
• Process produces a registration free/translation/rotation-invariant multivariate feature vector

Rcpp and parallelR-core packages are great for quick and easy speedups
3D PCA of 1740 real and simulated mean profiles of striation patterns from 58 screwdrivers:

- How many PCs should we use to represent the data??
  - No unique answer

FIRST we need an algorithm to I.D. a toolmark to a tool

- ~45% variance retained
Support Vector Machines

- Support Vector Machines (SVM) determine efficient association rules
  - *In the absence of specific knowledge of probability densities*
Error Rate Estimation: Machine Learning

• **Cross-Validation**: hold-out chunks of data set for testing
  • Known since 1940s
  • Most common: **Hold-one-out**

• **Bootstrap**: Randomly selection of observed data (with replacement)
  • Known since the 1970s
  • Can yield *confidence intervals around error rate estimate*

• **The Best**: Small training set, BIG test set
Error Rate Estimation: Pair-Wise Comparisons

- **Univariate** approaches compute estimates of similarity score distributions for **Known Matches (KM)** and **Known Non-Matches (KNM)**

\[
\text{Call: Matches} = \frac{\text{#False Matches}}{\text{#KNM-Comparisons}}
\]

\[
\text{Call: Non-Matches} = \frac{\text{#False Non-matches}}{\text{#KM-Comparisons}}
\]

\[
\text{Error Rate} = \frac{(\text{#False Non-matches} + \text{#False Matches})}{\text{#Comparisons}}
\]
How good of a “match” is it?

Conformal Prediction $^{\text{Vovk}}$

- Can give a judge or jury an easy to understand measure of reliability of classification result
- **Confidence** on a scale of 0%-100%
- **Testable claim**: Long run I.D. error-rate should be the chosen significance level

- This is an orthodox “frequentist” approach
  - Roots in Algorithmic Information Theory
- Data should be IID but that’s it
How Conformal Prediction works for us

• Given a “bag” of obs with known identities and one obs of unknown identity

  • Estimate how “wrong” labelings are for each observation with a non-conformity score (“wrong-iness”)

  • For us, one-vs-one SVMs: $t_i = \frac{1}{k-1} \sum_{j=1}^{k(k-1)/2} \lambda_{i,j}$

$\lambda = 0$

• Correctly classified and behind margins
• Shouldn’t contribute to “wrong-iness”

$0 < \lambda < C$

• Correctly classified but SVs or marginal
• Should contribute something to “wrong-iness”

$\lambda = C$

• Wrong
• Should contribute most “wrong-iness”
How Conformal Prediction works for us

• Given a “bag” of obs with known identities and one obs of unknown identity\textsuperscript{Vovk}

  • Looking at the “wrong-iness” for all the known observations in the bag:
    • Ask: Does labeling-\textit{i} for the unknown have an unusual amount of “wrong-iness”??:

Given “wrong-iness” for labeling-\textit{i} of unknown, number of obs with at least as much “wrongi-ness”

\[
p_{\text{possible-ID}_i} = \frac{\# \left\{ j \in \{1, 2, \ldots, n\} : t_j^{\text{possible-ID}_i} \geq t_{\text{test-pattern}}^{\text{possible-ID}_i} \right\}}{n} \quad i \in \{1, 2, \ldots, k \text{ I.D.s}\}
\]

• If not:
  • \[ p_{\text{possible-ID}_i} \geq \text{chosen level of significance } \alpha \]
  • Put ID \textit{i} in the (1 - \alpha)*100% confidence interval: \( \Gamma^{1-\alpha} \)

\[
\text{ID}_i \in \Gamma^{1-\alpha} \quad \text{if } p_{\text{ID}_i} \geq \alpha
\]
Conformal Prediction

- For 95%-CPT (PCA-SVM) confidence intervals will not contain the correct I.D. 5% of the time in the long run
- Straight-forward validation/explanation picture for court

Empirical Error Rate: 5.3%

Theoretical (Long Run) Error Rate: 5%

14D PCA-SVM Decision Model for screwdriver striation patterns
How good of a “match” is it?

Efron Empirical Bayes

• An I.D. is output for each questioned tool mark
  • This is a computer “match”

• What’s the probability the tool is truly the source of the tool mark?

• Similar problem in genomics for detecting disease from microarray data
  • They use data and Bayes’ theorem to get an estimate
Empirical Bayes

• From Bayes’ Theorem we can get\textsuperscript{Efron}:

\[
\hat{\Pr}(S^- \mid z) = \frac{\hat{p}(z \mid S^-)}{\hat{f}(z)} \hat{\Pr}(S^-)
\]

Estimated probability of not a true “match” given the algorithms' output z-score associated with its “match”

Names: **Posterior error probability (PEP)**\textsuperscript{Kall}
**Local false discovery rate (lfdr)**\textsuperscript{Efron}

• Suggested interpretation for casework:

\[
1 - \hat{\Pr}(S^- \mid z) = \text{Estimated “believability” that the specific tool produced the tool mark}
\]
Fit local-fdr models

Validation set p–values

All validation set z–values

$\Phi^{-1}(p\text{-values})$

JAGS generated sample of 20 f(z)

Use locfdr

Fit classic Poisson regression for $f(z)$

Use modified locfdr/JAGS or Stan

Fit Bayesian hierarchical Poisson regressions
Bayesian Hierarchical Poisson Regression Details

- To run the Bayesian Estimation we use JAGS^{Plummer} or Stan^{Gelman}:

DAG for the Poisson Regression

\[ \sigma_\beta \]

\[ \beta_0 \]

\[ \sigma_\epsilon \]

\[ \lambda_k \]

\[ \epsilon_k \]

\[ y_k \]

\[ \beta_i \quad h_i(x_k) \text{ or } (x_k)^i \quad k = 1:K \quad i = 1:d \]

\[ \sigma_\beta \sim \text{Uniform}(0, 100) \]

\[ \sigma_\epsilon \sim \text{Uniform}(0, 100) \]

\[ \beta_0 \sim \text{Normal}(0, \sigma_\beta) \]

\[ \epsilon_i \sim \text{Normal}(0, \sigma_\epsilon) \]

\[ \beta_j \sim \text{Normal}(0, \sigma_\beta) \]

Suggested by Gelman
A Bayesian Hierarchical Model: Believability Curve

Posterior Association Probability (Believability...)

\[ P(S+|z) = \text{tdr}(z) \]

Implemented in \texttt{fdrID}^{\text{Petraco}}

for \[ R \]

JAGS MCMC Bayesian over-dispersed Poisson with intercept, on test set
Empirical Bayes’

- Model’s use with crime scene “unknowns”:

Estimated Posterior Error Probabilities (Local FDRs)

| Pr(S=z|est. | fdrID for: |
|---|---|
| -4.8 | 0.0000 |
| -4.6 | 0.0005 |
| -4.4 | 0.0010 |
| -4.2 | 0.0015 |

This is the est. post. prob. of no association
= 0.00027 = 0.027%

Computer outputs “match” for:
unknown crime scene toolmarks-with knowns from “Bob the burglar” tools
Likelihood Ratios from Empirical Bayes

- Using the fit posteriors and priors we can obtain the likelihood ratios

\[
\widehat{\text{LR}}(z) = \frac{\hat{\Pr}(H_p|E)}{\hat{\Pr}(H_d|E)} = \frac{\hat{\Pr}(H_p)}{\hat{\Pr}(H_d)} = \frac{\hat{\text{tdr}}(z) \hat{\pi}_0}{\hat{\text{fdr}}(z) (1 - \hat{\pi}_0)}
\]
Bayesian Match Probabilities from CMS

2007 Neel and Wells study Neel, Wevers, Buckleton:

- Count the number of each type of CMS run for KM and KNM comparisons
- A CMS type is its run length:
  - 4X means 4 matching adjacent lines in a comparison of two striation patterns

<table>
<thead>
<tr>
<th>Number observed</th>
<th>CMS run lengths:</th>
<th>Number observed</th>
<th>CMS run lengths:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2X</td>
<td>3X</td>
<td>4X</td>
</tr>
<tr>
<td>0</td>
<td>508</td>
<td>612</td>
<td>694</td>
</tr>
<tr>
<td>1</td>
<td>186</td>
<td>172</td>
<td>135</td>
</tr>
<tr>
<td>2</td>
<td>109</td>
<td>59</td>
<td>43</td>
</tr>
<tr>
<td>3</td>
<td>39</td>
<td>29</td>
<td>19</td>
</tr>
<tr>
<td>4</td>
<td>21</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>14</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>&gt;8</td>
<td>13</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Model each column of counts as arising from a multinomial distribution with Dirichlet prior
Bayesian Match Probabilities from CMS

- Updated CMS run length probabilities:

<table>
<thead>
<tr>
<th>Number observed</th>
<th>CMS run lengths:</th>
<th>KM comparisons</th>
<th>KNM comparisons</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2X</td>
<td>3X</td>
<td>4X</td>
</tr>
<tr>
<td>0</td>
<td>0.550</td>
<td>0.663</td>
<td>0.752</td>
</tr>
<tr>
<td>1</td>
<td>0.202</td>
<td>0.187</td>
<td>0.147</td>
</tr>
<tr>
<td>2</td>
<td>0.119</td>
<td>0.065</td>
<td>0.047</td>
</tr>
<tr>
<td>3</td>
<td>0.043</td>
<td>0.032</td>
<td>0.022</td>
</tr>
<tr>
<td>4</td>
<td>0.024</td>
<td>0.018</td>
<td>0.019</td>
</tr>
<tr>
<td>5</td>
<td>0.012</td>
<td>0.011</td>
<td>0.003</td>
</tr>
<tr>
<td>6</td>
<td>0.005</td>
<td>0.011</td>
<td>0.002</td>
</tr>
<tr>
<td>7</td>
<td>0.012</td>
<td>0.008</td>
<td>0.004</td>
</tr>
<tr>
<td>8</td>
<td>0.016</td>
<td>0.003</td>
<td>0.001</td>
</tr>
<tr>
<td>&gt;8</td>
<td>0.015</td>
<td>0.002</td>
<td>0.002</td>
</tr>
</tbody>
</table>

- So what can we use these for??
  - Lot’s of stuff, but we put them into a Bayesian network:
    - BN model for Match/Non-match probabilities given observed numbers of CMS runs
Bayesian Networks

“Prior” network based on Neel and Wells observed counts and Multinomial-Dirichlet model:

Estimate of the “match” probability which can be turned into an LR if so desired

“Instantiated” network with observations from a comparison:
Future Directions

• **Clean up**: cptID, feature2, fdrID

• **GUI modules** for common toolmark comparison tasks/calculations using 3D microscope data

• **2D features** for toolmark impressions

• **Parallel/GPU/FPGA** implementation of computationally intensive routines e.g. ALMA Correlator for astronomy data
  • Especially for retrieving “relevant pop/best match” reference sets

• **Uncertainty for Bayesian Networks**
  • Models, parameters…
References

Lillien:
  • http://www.openfmc.org/
  • http://www.cadreforensics.com/
Brubaker: https://github.com/OpenFMC/x3p
Petraco:
  • https://github.com/npetraco/x3pr
  • https://github.com/npetraco/feature2
  • https://github.com/npetraco/cptID
  • https://github.com/npetraco/fdrID
Edd:
  • https://cran.r-project.org/web/packages/Rcpp/index.html
  • http://dirk.eddelbuettel.com/code/rcpp.html
OpenCV: http://opencv.org/
Whitcher: https://cran.r-project.org/web/packages/waveslim/index.html
R-Core: https://www.r-project.org/
Vovk: http://www.alrw.net/
Efron:
References

locfdr: https://cran.r-project.org/web/packages/locfdr/index.html

JAGS:
  - http://mcmc-jags.sourceforge.net/
  - https://cran.r-project.org/web/packages/rjags/index.html

Stan:
  - http://mc-stan.org/
  - https://cran.r-project.org/web/packages/rstan/index.html

Neel:

Wevers:

Buckleton:

BayesFusion: http://www.bayesfusion.com/
Acknowledgements

• Professor Chris Saunders (SDSU)
• Professor Christophe Champod (Lausanne)
• Alan Zheng (NIST)
• Ryan Lillien (Cadre)
• Scott Chumbley (Iowa State)
• Robert Thompson (NIST)
• Research Team:
  • Mr. Daniel Azevedo
  • Ms. Tatiana Batson
  • Dr. Martin Baiker
  • Ms. Julie Cohen
  • Dr. Peter Diaczuk
  • Mr. Antonio Del Valle
  • Ms. Carol Gambino
  • Dr. James Hamby
  • Mr. Nick Natalie
  • Mr. Mike Neel
• Ms. Alison Hartwell, Esq.
• Mr. Robert McLean
• Dr. Brooke Kammrath
• Mr. Chris Lucky
• Off. Patrick McLaughlin
• Dr. Mecki Prinz
• Dr. Linton Mohammed
• Ms. Diana Paredes
• Mr. Nicholas Petraco
• Ms. Stephanie Pollut

Collaborations,
Reprints/Preprints:
npetraco@gmail.com
http://jjcweb.jjay.cuny.edu/npetraco/
• Dr. Jacqueline Speir
• Dr. Peter Shenkin
• Mr. Peter Tytell
• Dr. Peter Zoon