Adapting the Chumbley Score to Bullet Striations

Ganesh Krishnan, Heike Hofmann

April 21, 2018

Objective and Motivation

- Same Source Matching of Bullet lands
- Evaluate performance of Chumbley Score method when used for Bullet Striations
- Bullet striations have curvature, not present in toolmarks
- Identify Error rates and effect of different parameters on them (In short finding the best error rates possible)

Structure

- Data being used
- What is the Chumbley Score Method?
- Identifying Best parameter Settings for Bullets
- Modifications to the Algorithm
- Results using Hadler and Morris [2017] method and results using Modified method

Variations of Chumbley score method and Error Rates for toolmarks

Research paper	Method	Data Source	False Positives	False Negatives
Faden et al. [2007]	Maximized Correlation	Screwdrivers	-	-
Chumbley et al. [2010] (Same-Surface Same-Angle)	Randomized Chumbley Score	Screwdrivers	2.3%	8.9%
Grieve et al. [2014]	Randomized Chumbley Score	Slip-joint	-	-
Hadler and Morris [2017] (Same-Surface Same-Angle)	Deterministic Chumbley Score	Screwdrivers	0%	6%

Table 1: Error Rates for Toolmarks using variations of the chumbley score method

Digitized Striation Marks

- Data
  - Ruger P85s Bullet Lands, or Hamby scans (Hamby et al. [2009]) provided by NIST (85,491 comparisons)
  - Bullet striation marks $\approx$ 2mm
  - Screwdriver marks $\approx$ 7mm (all chumbley score papers)
- Let $x(t_1)$, $t_1 = 1, 2, ... T_1$ and $y(t_2)$, $t_2 = 1, 2... T_2$ be two digitized marks (where T_1 and T_2 are not necessarily equal).
- T_1 and T_2 are the final pixel indexes of each marking. Therefore give the respective lengths of the markings.
- Signatures/ Profiles (NIST- Hamby) $\approx$ 1200 pixels (2 mm) Screwdriver toolmarks (Chumbley Papers) $\approx$ 9000 pixels (7 mm)

Chumbley Score

Step 0 : Defining a coarseness parameter

- Used to remove drift and (sub)class characteristics from individual markings
- Lowess or Loess fit residuals = Signatures
- Removes topographic structure (curvature)
- Improve the signal to noise Ratio

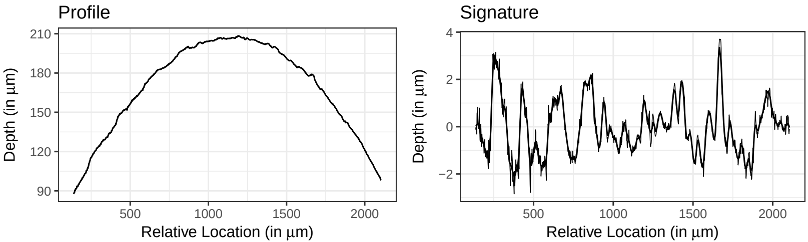


Figure 1: Bullet land profile (left) and the corresponding signature (right).

Algorithm

- Two steps: Optimization (1^{st}) and Validation (2^{nd}).
- Windows $\implies$ short segments of the markings
  - Have *predefined sizes*. $(T_1$ or $T_2 >>> w_o$ & w_v$)

    1. w_o used in the Optimization step
    2. w_v used in the Validation step

Optimization step

- **Goal :**Align markings horizontally as best as possible
- Correlation Matrix of all possible windows of size w_o between $x(t_1)$ and $y(t_2)$ computed
- *Identify lag for horizontal alignment*
  Window Pair with maximized correlation $\implies$
  Optimal vertical (in-phase) shift of $t_1^o - t_2^o$
  - For aligning the two markings.

$$(t_1^o, t_2^o) = \underset{1 \leq t_1 \leq T_1, 1 \leq t_2 \leq T_2}{\arg\max} \operatorname{cor}\left(x^{w_o}(t_1), y^{w_o}(t_2)\right)$$

where t_1^o, t_2^0 are the respective starting points of w_o in $x(t_1)$ and $y(t_2)$

- Let t_1^* and t_2^* be relative optimal locations, where $t_i^* = t_i^o/(T_i - w_o)$ for $i = 1, 2$, such that $t_1^*, t_2^* \in [0, 1]$.
- Once (sub-)class characteristics are removed, these locations have uniform distribution in $[0, 1]$

Validation Step

- Two sets of windows of size w_v chosen from both markings (see Figure 2)
- First set or **Same Shift**
  - pairs of windows are extracted from the two markings using the optimal vertical shift. $t_1^o - t_2^o$
- Second set or **Different Shift**
  - the windows are extracted using a different (out-of-phase) shift.

In-phase and Out-of-phase

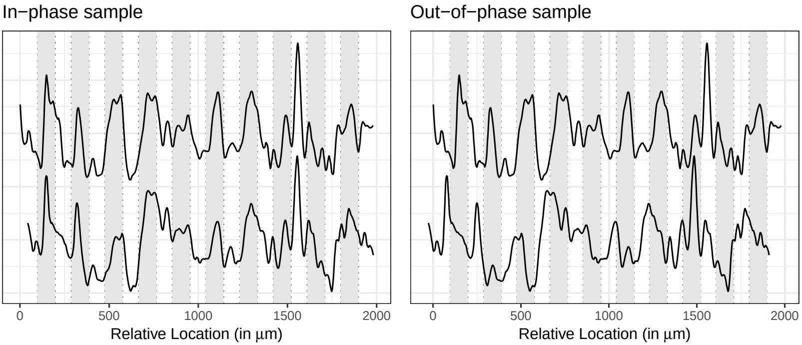


Figure 2: Two markings made by the same source. For convenience, the markings are moved into phase on the left and out-of phase on the right. In-phase (left) and out-of-phase (right) samples are shown by the light grey background. The Chumbley-score is based on a Mann-Whitney U test of the correlations derived from these two sets of samples.

- Both same- and different-shift pairs correlations between the markings are calculated.
- For Same-Source markings, correlations
  - for the in-phase shift should be high
  - for out-of-phase shift should be low.
    - Provide a measure for the base-level correlation to which in-phase shift correlations can be compared.
- The Chumbley score is the Mann Whitney U statistic computed by comparing between in-phase sample and out-of-phase sample.

Block Diagram

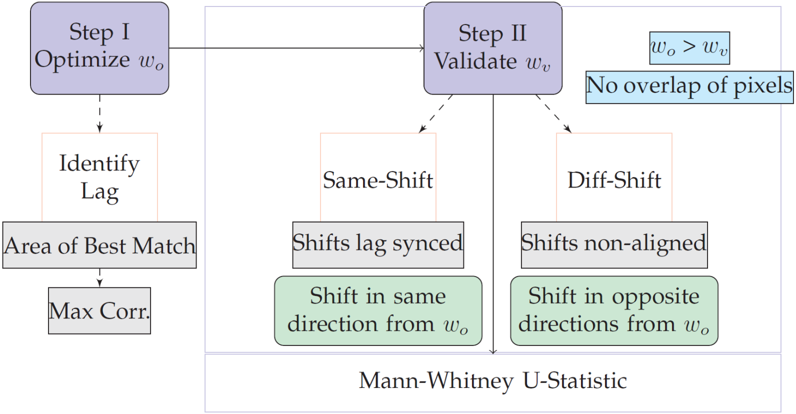


Figure 3: An overview of the adjusted chumbley score method as given by Hadler and Morris [2017]

Starting Points

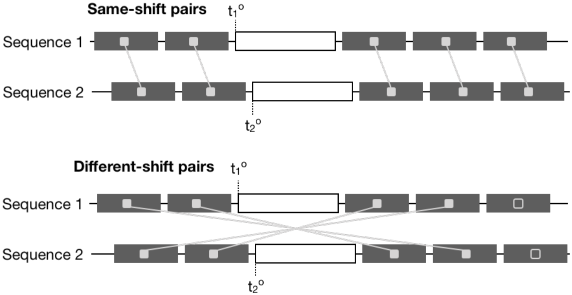
More precisely, let us define starting points of the windows of validation $s_i^{(k)}$ for each marking $k = 1, 2$ as

$$s_i^{(k)} = \begin{cases} t_k^* + iw_v^* & \text{for } i < 0 \\ t_k^* + w_o^* + iw_v^* & \text{for } i \geq 0, \end{cases} \tag{1}$$

for integer values of i with $0 < s_i^{(k)} \leq 1 - w_v^*$

The Hadler and Morris [2017] method (CS1)

- Same-shift pairs of length w_v are all pairs that start in:
$$(s_i^{(1)}, s_i^{(2)}) \quad \forall\, i \in \mathbb{Z}$$
for which both $s_i^{(1)}$ and $s_i^{(2)}$ are defined.
- Different-shift pairs are defined as
$$(s_i^{(1)}, s_{-i-1}^{(2)}) \quad \forall\, i \in \mathbb{Z}$$
where both $s_i^{(1)}$ and $s_{-i-1}^{(2)}$ are defined (see fig. 4).



Failed Tests

- By definition (equation 1), some number of tests fail to produce a result
- Either because the number of eligible same-shift pairs is 0, or the number of different-shift pairs is 0.
- t_1^o, t_2^o not necessarily independent
  - **same-source:** Assume high dependence, $\mathrm{corr}(t_1^o, t_2^o) \approx 1$
    - Example: $w_o = 120$, coarseness (c) $= 0.3$, $\mathrm{corr}(t_1^o, t_2^o) = 0.85$
  - **diff-source:** Assume independence of t_1^o, t_2^o
    - Example: $w_o = 120$, coarseness (c) $= 0.3$, $\mathrm{corr}(t_1^o, t_2^o) = 0.12$

Failure Rate

$$P\left(t_1^o < w_v \bigcap t_2^o > T_2 - w_o - w_v\right) +$$
$$P\left(t_1^o < T_1 - w_o - w_v \bigcap t_2^o < w_v\right).$$

Same-shift failure

- Same-source ≈ 0
- Different-source $\approx 2\ P(t_i < w_o)^2 = \frac{2w_v^2}{(T_1 - w_o)(T_2 - w_o)}$

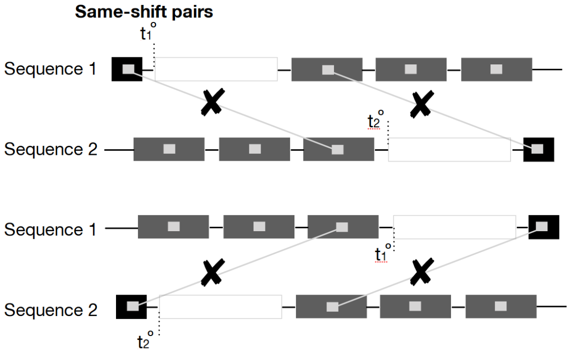


Figure 5: Sketch of same-shift pairings (top) when the lag is too large to accomodate a a vaildation window in either of the two signatures

Different-Shift Failure

- Same-source (Assuming $t_1^o \approx t_2^o$) $\approx 2w_v/(T_i - w_o)$

$$P(t_1^o < w_v \cap t_2^o < w_v) + P(t_1^o < w_v \cap t_2^o < w_v) = 2P(t_o^1 < w_v)$$

- Different-source $\approx 2P(t_i < w_o)^2 = \frac{2w_v^2}{(T_1 - w_o)(T_2 - w_o)}$

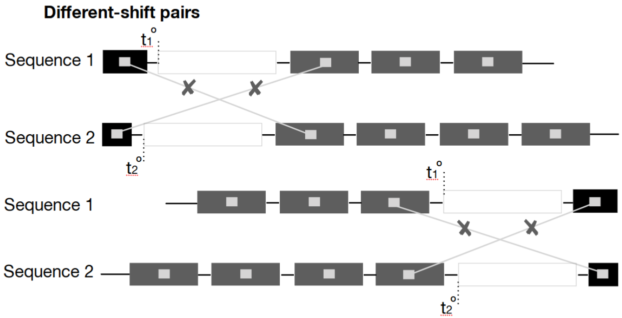


Figure 6: Sketch of diff-shift pairings (top) when the number of diff-shift computations is likely to 0

Proposed Modification

- Failures due to missing Same-shift pairs unavoidable
- Failures due to missing different-shift pairs preventable

Define same-shift pairs identical to Hadler and Morris [2017] as pairs

$$(s_i^{(1)}, s_i^{(2)}) \quad \forall\, i \in \mathbb{Z}$$

where the boundary conditions of both sequences are met simultaneously.

- Let us assume that this results in I pairs.
- Let $s_{(j)}^{(k)}$ to be the jth starting location in sequence $k = 1, 2$, i.e. $s_{(1)}^{(k)} < s_{(2)}^{(k)} < ... < s_{(I)}^{(k)}$.

We then define the pairs for different-shifts as

$$\left(s^{(1)}_{(j)}, s^{(1)}_{(I-j+1)}\right) \text{ for } j = \begin{cases} 1, ..., I & \text{for even } I \\ 1, ..., (I-1)/2, (I-1)/2+2, ..., I & \text{for odd } I \end{cases} \tag{2}$$

- For an odd number of same-shift correlations
  - We skip the middle pair for the different-shift correlations (see fig. 7).
- This pairing ensures that the number of different-shift pairings is the same or at most one less than the number of same-shift pairings in all tests.

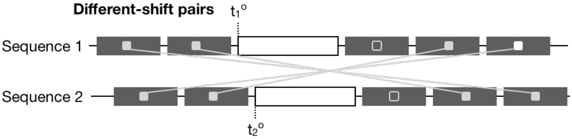


Figure 7: Sketch of adjusted different-shift pairings. At most one of the same-shift pairings can not be matched with a different-shift pair.

Case where CS1 fails but CS2 does not fail

CS1 Hadler and Morris [2017] algorithm
CS2 the suggested modified algorithm

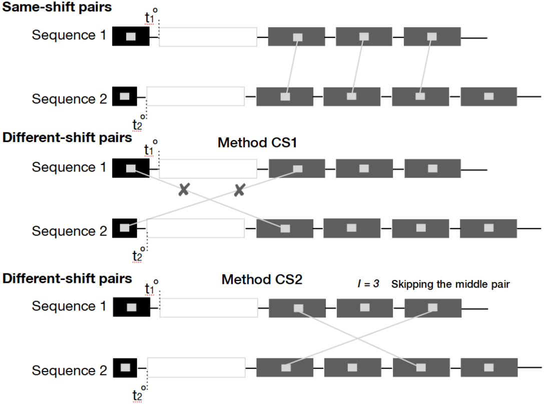


Figure 8: Sketch of a case where CS1 fails but CS2 does not fail

L. Scott Chumbley, Max D. Morris, M. James Kreiser, Charles Fisher, Jeremy Craft, Lawrence J. Genalo, Stephen Davis, David Faden, and Julie Kidd. Validation of tool mark comparisons obtained using a quantitative, comparative, statistical algorithm. *Journal of Forensic Sciences*, 55(4):953–961, 2010. ISSN 1556-4029. doi: 10.1111/j.1556-4029.2010.01424.x. URL http://dx.doi.org/10.1111/j.1556-4029.2010.01424.x.

David Faden, Julie Kidd, Jeremy Craft, L. Scott Chumbley, Max D. Morris, M. James Genalo, Lawrence J.and Kreiser, and Stephen Davis. Statistical confirmation of empirical observations concerning toolmark striae. *AFTE Journal*, 39(2):205–214, 2007.

Taylor Grieve, L. Scott Chumbley, Jim Kreiser, Laura Ekstrand, Max Morris, and Song Zhang. Objective comparison of toolmarks from the cutting surfaces of slip-joint pliers. *AFTE Journal*, 46(2): 176–185, 2014.

Jeremy R. Hadler and Max D. Morris. An improved version of a tool mark comparison algorithm. *Journal of Forensic Sciences*, pages n/a–n/a, 2017. ISSN 1556-4029. doi: 10.1111/1556-4029.13640. URL http://dx.doi.org/10.1111/1556-4029.13640.

James E. Hamby, David J. Brundage, and James W. Thorpe. The Identification of Bullets Fired from 10 Consecutively Rifled 9mm Ruger Pistol Barrels: A Research Project Involving 507 Participants from 20 Countries. *AFTE Journal*, 41(2):99–110, 2009.