1 Motivating Problem
- Criminal breaks into a window using a screwdriver leaving behind a tool mark.
- A suspect is identified and found with a screwdriver in his possession; was that the screwdriver used in the crime?
- Grinding process of tools creates unique striae that are left behind when the tool is scraped against a hard surface, analogous to fingerprints.

2 Measurement Technique
Stylus profilometer reads across a tool mark recording the depths of the striae. The depths become a numerical dataset when read pixel-to-pixel.

Comparing two digitized tool marks results in a single numerical index value of similarity.

3 Data
- Available tool marks
  - 1 field mark (from crime scene): $x_0$
  - $n$ lab marks made by the suspect tool; $x_1, \ldots, x_n$
  - $y_j$ is the comparison value from comparing $x_j$ with $x_0$
- Resulting data after all pairwise comparisons are made:
  - $y_j$ for $j = 1, \ldots, n$: all comparisons of the field mark to a lab mark
  - $y_{ij}$ for $i,j = 1,\ldots,n$ and $i < j$: all pairwise comparisons of lab marks
- Regard $y_{ij}$ as approximately normal

4 Interpreting the Data
Did the suspect tool create the field mark?
- i.e.: Does the field mark match the lab marks?
  - If $y_{ij}$ are comparable to $y_j$ then “no evidence” that the tool marks are different; i.e.: match
  - If $y_{ij}$ are small relative to $y_j$ then “evidence” that the tool marks are different; i.e.: no match

5 Basic Model
- $E(y_{ij}) = \mu_0$ for $j = 1, \ldots, n$; mean for comparison of field mark to lab marks
- $E(y_{ij}) = \mu_i$ for $i,j = 1, \ldots, n$ and $i < j$; mean for comparison of lab marks
- $	ext{Var}(y_{ij}) = \sigma^2$; variance for all comparisons
- $\text{Corr}(y_{ij}, y_{kj}) = 0$ if $i \neq k$, $i \neq j$, $i \neq k$, and $j \neq i$
- $\text{Corr}(y_{ij}, y_{kj}) = 1$ if $i = k$ or $i = j$ or $j = k$, or $j = l$
- That is, comparisons involving a common tool mark are correlated
- Parameters estimated using MLE based on weighted least squares
- Likelihood ratio test used to test $H_0$: $\mu_0 = \mu_1$

6 Angle Influence
- Angle at which the tool is held when making the mark changes the appearance of the tool mark
- Boxplots show comparison values for tool marks made by the same tool at varying angles: 30°, 45°, 60°, 75°, & 85°
- Tool angles differing by more than 10° result in small comparison values indicative of non-matches
- Angle of a tool mark should be incorporated into the model

7 Model with Angle
- $a_i$ = the tool angle (in degrees) at which tool mark $x_i$ is made
- $d(a_i, a_j) = \exp(-\theta(a_i - a_j)^2)$, a positive similarity measure; shown in figure below
- We chose $\theta = 0.01$ so that $d(a_i, a_j) = 1$ when tool angles match, is small when $|a_i - a_j| > 10$ and approaches 0 when $|a_i - a_j| > 10$
- With more data, we would estimate $\theta$

8 Analysis
Still want to know if the same tool was used to make the field mark and the lab marks. Use a LRT with the Angle Models defined as
- Null Model
  - $y_j = N(\mu_0, \sigma^2)$ where
  - $\mu_0 = \mu_1 + \alpha_i d(a_i, a_j)$ for $i,j = 1, \ldots, n$ and $i < j$
- Alternative Model
  - $y_j = N(\mu_i, \sigma^2)$ where
  - $\mu_0 = \mu_1 + \alpha_0 d(a_0, a_j)$ for $j = 1, \ldots, n$ and $i < j$
- Correlation structure is the same as described in Block 5. MLEs can be obtained for $\mu_0, \mu_i, \sigma^2, \alpha_0$ and $\alpha_i$ provided values of $a_0$ and $\rho$ which are estimated using a grid search.

9 Matching Results
- Matching datasets consist of only tool marks made by the same tool: 20 tool marks – 4 at each of the 5 angles
- Each mark chosen to be the field mark one-at-a-time
- Matches were found to be sensitive to flaws in the data and the matching process so “bad” tool marks and “bad” matches were removed
- LRT from Block 8 was performed on remaining tool marks – total of 108 LRTs. Histogram shows the resulting p-values
- Since tool marks match, we expect most tests should Reject $H_0$ and have small p-values which is shown in the graph

10 Estimation Results
The following histograms display the behavior of the estimate of $a_i$ from the matching data. Each graph shows the estimated value of $a_i$ for each of the five angles, the red line shows the true field angle.

11 Non-matching Results
- Non-matching datasets consist of 20 lab tool marks made by the same tool – 4 at each of the 5 angles – plus one field mark made by a different tool; one chosen from each angle from every remaining tool
- LRT from Block 8 was performed on datasets
- Repeated for all 6 tools – total of 150 LRTs; p-values are shown in the histogram
- Since the tool marks do not match, we expect most of the tests should Reject $H_0$ and have small p-values which is shown in the graph

12 Conclusions
- The angle at which the tool is used can have a major effect on the similarity of tool marks
- For known matches, the field angle is predicted within 5° of the true angle in over 80% of the likelihood ratio tests
- Angle prediction is accurate as long as there are lab angles available within 10° of the field angle
- More data is necessary to include $\theta$ in the estimation process
- Correct match determination is sensitive to flaws in the tool marks as well as in the initial matching process
- The LRT based on the unknown-angle model effectively discriminates between matching and non-matching tool mark pairs

Funded by the National Institute of Justice Grant # 2009-DN-R-119