

Summer 2007

Statistical Confirmation of Empirical Observations Concerning Tool Mark Striae

D. Faden

Iowa State University

J. Kidd

Iowa State University

J. Craft

Iowa State University

Max D. Morris

Iowa State University, mmorris@iastate.edu

Lawrence Genalo

Iowa State University, genalo@iastate.edu

Follow this and additional works at: http://lib.dr.iastate.edu/mse_pubs



Part of the [Applied Statistics Commons](#), [Categorical Data Analysis Commons](#), [Ceramic Materials Commons](#), [Metallurgy Commons](#), and the [Other Materials Science and Engineering Commons](#)

The complete bibliographic information for this item can be found at http://lib.dr.iastate.edu/mse_pubs/212. For information on how to cite this item, please visit <http://lib.dr.iastate.edu/howtocite.html>.

This Article is brought to you for free and open access by the Materials Science and Engineering at Iowa State University Digital Repository. It has been accepted for inclusion in Materials Science and Engineering Publications by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

Statistical Confirmation of Empirical Observations Concerning Tool Mark Striae

Abstract

Too/marks produced by 44 sequentially manufactured screwdriver tips have been characterized for surface roughness using a pro.filometer. Too/marks were produced in lead at angles of 30 o, 60 o, and 85 °. A computer program developed to compare and match the pro.filometer data has been used to show that marks from a single tip produced at similar angles yield much higher correlation values than marks produced from the same tip but at different angles. This analysis provides statistical support for the widely-accepted empirical observation that too/mark striae must be reproduced at similar angles in order to be unambiguously identified as being made by a particular tool.

Keywords

Comparison microscope, screwdriver, statistics, striae, toolmark

Disciplines

Applied Statistics | Categorical Data Analysis | Ceramic Materials | Metallurgy | Other Materials Science and Engineering

Comments

This article is from *AFTE Journal* 39 (2007): 205. Posted with permission.

Authors

D. Faden, J. Kidd, J. Craft, Max D. Morris, Lawrence Genalo, J. Kreiser, and S. Davis

Statistical Confirmation of Empirical Observations Concerning Tool Mark Striae

By: D. Faden, J. Kidd, J. Craft, L.S. Chumbley, M. Morris, L. Genalo, J. Kreiser*, S. Davis Iowa State University, Ames Laboratory Ames, Iowa *Illinois State Police, Retired, Springfield, Illinois

Key Words: comparison microscope, screwdriver, statistics, striae, toolmark

ABSTRACT

Toolmarks produced by 44 sequentially manufactured screwdriver tips have been characterized for surface roughness using a profilometer. Toolmarks were produced in lead at angles of 30°, 60°, and 85°. A computer program developed to compare and match the profilometer data has been used to show that marks from a single tip produced at similar angles yield much higher correlation values than marks produced from the same tip but at different angles. This analysis provides statistical support for the widely-accepted empirical observation that toolmark striae must be reproduced at similar angles in order to be unambiguously identified as being made by a particular tool.

Introduction

In light of the 1993 Daubert vs. State of Florida decision, forensic examiners are under increasing pressure to prove that the methods they routinely employ meet established criteria pertaining to a scientific investigation. It was asserted in this ruling that in order to qualify as "scientific knowledge", four criteria must be met: testability of scientific principle, known or potential error rate, peer review and publication, and general acceptance in a particular scientific community [1]. Thus, proving that basic inferences or assertions held by experts in the field of forensics have a sound scientific basis is a goal of law enforcement agencies and researchers around the country. In support of this goal, Iowa State University has been conducting research into the matching of toolmark striae using quantitative, objective measurements of the surface. Toolmarks have been produced using sequentially manufactured screwdriver tips and the three-dimensional nature of the marked surfaces has been measured. A computer algorithm has been developed that allows rapid comparison of large numbers of data files produced by this characterization method.

In the field of toolmark identification it has long been assumed that optical comparative matching of toolmarks is only valid when the examiner can reproduce in the laboratory setting, at least in some approximate way, the angles that existed when the original mark was made. In the case of a screwdriver pry mark this might involve producing a number of marks with the

suspect tool held at varying angles for each side. The expertise of the examiners, and their ability in pattern recognition, even when the match may be somewhat tenuous, is the driving force to make repeated test marks at various angles in order to allow them to make an unequivocal statement concerning the possibility of a match between evidence and test samples.

This paper presents results that provide quantitative validation of the need for reproducing as closely as possible the angles that existed when an initial mark was made by a suspect tool. These results have been obtained from a test set of screwdriver tips that have been characterized quantitatively using surface measurement techniques, and the results analyzed using comparisons based on a standard statistical measure of similarity, implemented in the program mentioned above.

Experimental

The test set for this study involves 44 screwdriver tips obtained from Omega Company and manufactured sequentially so to be as nearly identical as possible. Test marks for study were produced in lead by dragging the tip across a small sample plate at fixed angles of 30°, 60°, and 85°. A small jig was manufactured to maintain the angles as nearly as possible to these values (Figure 1). All test marks were produced by the same experienced forensic toolmark examiner.

Once produced the surface roughness of the samples was characterized using a surface profilometer, a device that can provide height measurements as a function of location on the two-dimensional sample surface, thus producing a measured 3-dimensional surface profile if desired. The profilometer

Date Received: December 1, 2006

Peer Review Completed: September 19, 2007

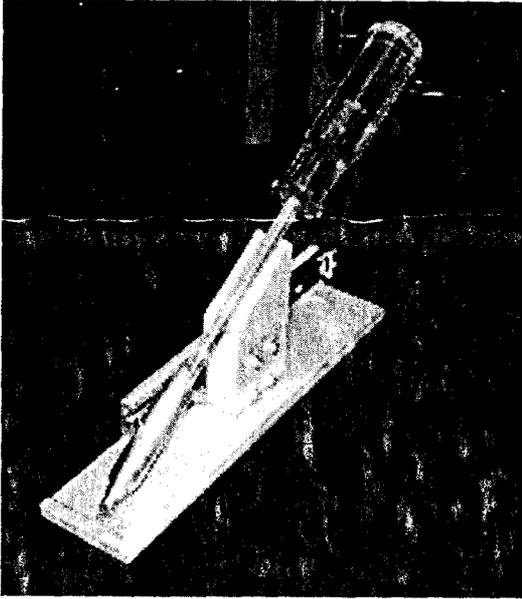
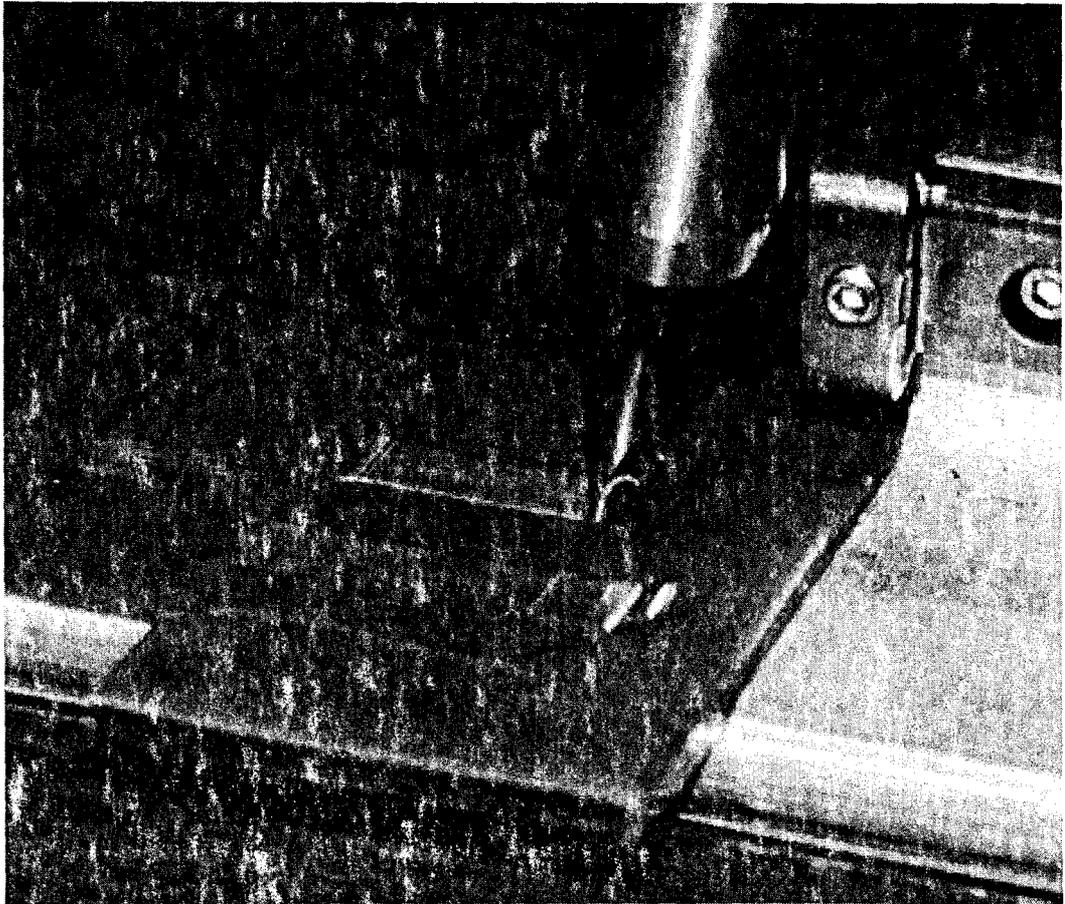


Figure 1. a) Macroscopic view of the jig used to produce samples. b) Close-up of toolmark being produced.



might be considered a descendent of the Striagraph, first developed and used by Davis in 1958 [2]. For the purpose of this study, data were collected along a linear trace as nearly perpendicular to the direction of striae as possible, essentially producing a measure of the height, depth, and width of the striae produced. A series of ten parallel traces were run for each mark produced. The length of each trace spanned the entire width of the mark and consisted of 9600 data points (i.e., surface heights) recorded at uniformly spaced points along the trace. For these scans the linear step along each trace was 0.277 microns. The measured height at each point is accurate to within ± 0.005 microns. A typical optical comparison is illustrated in Figure 2a; data obtained from the left-hand sample and generated using the profilometer is displayed in Figure 2b. For each tool, four separate toolmarks were produced using each side of the tool, at each of the three angles examined, with 10 separate profilometer traces recorded from each mark. Thus, for each side of the screwdriver tip, a total of 120 profilometer traces are available, encompassing the three angles used. However, in order to simplify the analysis, only two of the four samples produced with each tool were used, and only a single profilometer trace from the 10 made of each sample was included in the comparisons.

The resultant data were analyzed using a computer program written to compare and match such data sets in a rapid and objective manner. This program implements an algorithm that sequentially compares relatively short segments from two different traces, essentially mimicking the actions of a forensic examiner. The routine execution is such that each segment in one trace is compared to all such segments of the same length (i.e. number of pixels) in the other trace.

The results presented below are based on sides A and B of the 44 tips in the collection. All six angular comparisons were made, namely, 30-30, 30-60, 30-85, 60-60, 60-85, and 85-85 degree comparisons. For each angular comparison, 88 mark comparisons, i.e., pairs of marks made by the same side of the same tip, were examined.

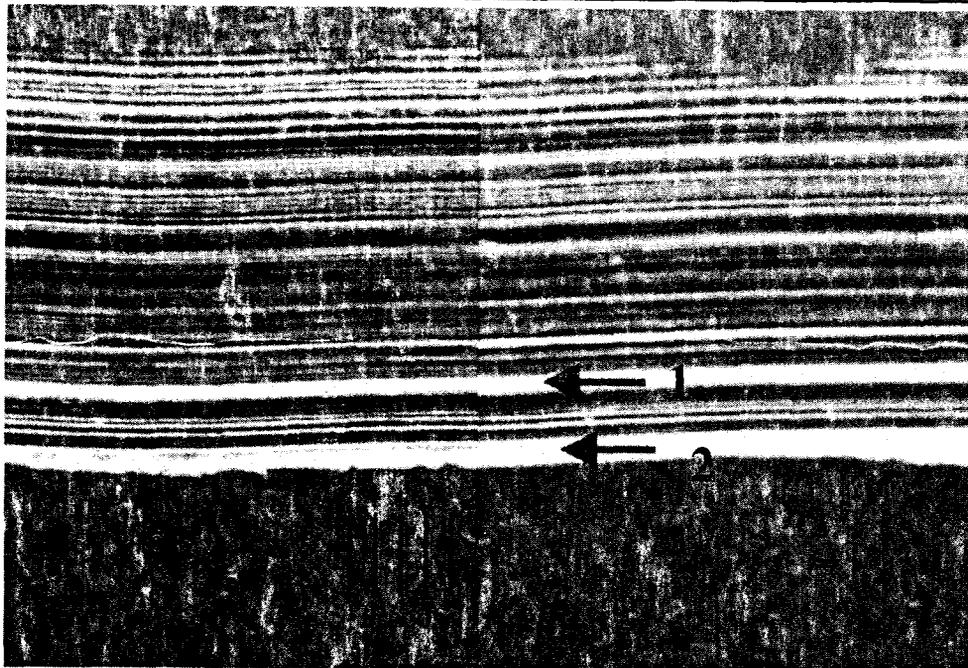
Comparisons were carried out by comparing all consecutive data windows of length 273 pixels from the third profilometer trace of one sample, to all consecutive data windows of the same length from the third trace of the other sample. Earlier experiments [3] had suggested this to be a reasonable value for window size, being roughly 0.2 mm out of the total trace length of 7 mm. The Pearson correlation coefficient [4] was computed from the two data series associated with every pair of windows compared; the two windows yielding the largest correlation were identified as the "best match" for the two samples. The maximum correlation found in each comparison is the "index of similarity" analyzed in this paper; these values are presented graphically in boxplot format below.

Three distinct comparative data sets were generated. To test true matches, separate marks generated from the same side of each tip were compared as described above. To test true nonmatches, marks generated from different tips selected at random were compared. The third data set was examined as an informal test of the hypothesis that different sides of a screwdriver tip produce fundamentally different marks, as maintained by forensic examiners. In this comparison the profilometer data obtained from side A of a tip was compared to data from side B of the same tip. The B data was reversed before comparison, the assumption being that if sides A and B were mirror images of each other, reversing the data should result in excellent matching. If each side truly is independent, the data should approximate that obtained for the true nonmatches. If a relationship does exist between sides A and B of a tip, the data might be expected to approximate that obtained for the true matches.

Results and Discussion

Figure 3 summarizes the results of the six angular comparisons made in this study for the true match data set. The 88 maximum correlations collected for each angle pair are displayed in the column for that comparison; a box encompassing the 25th to 75th percentiles is displayed to give a pictorial representation of the spread in the data, with the 50th percentile marked as a dark line. The "whiskers" extending on dashed lines beyond the edge of the box cover all points within an interval of length 1.5 times the difference between the 75th and 25th percentiles, while the dots represent individual data values that lie even further from the central half of the distribution [3].

Relative to the values generally observed in correlations calculated from data, all calculated Pearson correlations, whether from samples produced by the same tool or different tools, were quite large (i.e., very close to 1). This is to be expected since each correlation is the largest, from among many thousands of correlation coefficients calculated, in the comparison of a single pair of samples. While it might be expected that the correlations between matching samples should be larger than those between nonmatching samples, it must be born in mind that these indices will all be very close to one, and should not be interpreted relative to the critical values usually used in declaring correlations to be statistically significant. In fact, since many features of the comparison can potentially influence the value of the maximum correlation found, this index should not be interpreted as a direct indicator of the likelihood that two samples were made using the same tool. The authors have developed a more intricate statistical analysis that corrects this flaw via a "validation step," and this will describe that fully in a future publication. For the present study, the maximum correlation coefficient was only used as a comparative measure under tightly controlled experimental



9A3022

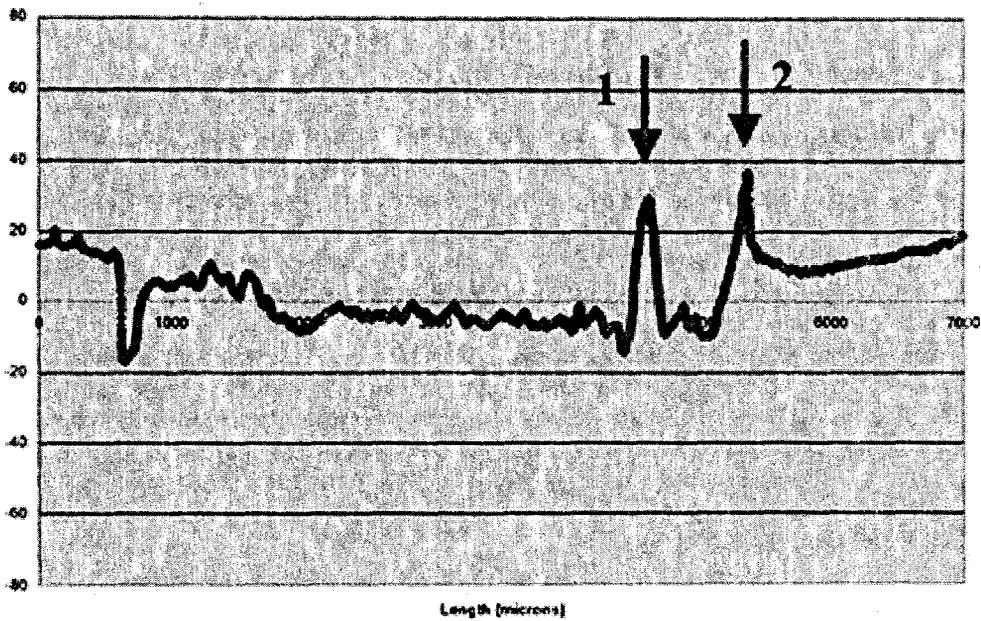


Figure 2: a) Photomicrograph from a comparison microscope showing match region with profilometer traces just visible as thin black vertical lines. (The positions of a few of the more easily seen marks are indicated.) b) profilometer scan from one of the traces. Arrows indicate distinctive features.

conditions.

Figure 4 shows these correlation values are, in fact, large even for samples that do not match. However, close examination of Figures 3 and 4 reveal a number of interesting features. Figure 3 clearly shows a separation exists in the true match data when comparing pairs of marks produced at equal angles to comparisons made when the angles are not equal. Comparisons made between known matches where the angles vary by 25-30 degrees (i.e. the 30-60 and 60-85 data comparisons) yield maximum correlation values that are very similar to those found in comparing true nonmatch sample pairs. This provides independent, statistical data in support of assertions long maintained by forensic examiners that marks must be made at nearly the same angle to produce an identifiable matching pattern in the striae.

Figure 5 shows the results for comparisons made between sides A and B of the same tip. No separation of data is seen between equi-angular comparisons for this data, even though the compared marks were produced by the same screwdriver tip. It is interesting that the opposite sides of the tip produce roughly the same results as those from the true nonmatch set, in other words, the two sides do indeed appear to be distinctly different, supporting the test hypothesis of side individuality. This assertion is more clearly illustrated in Figure 6, where the means of the angular comparisons for the three data sets are plotted.

There are additional observations that can be made from these plots. The correlations for 85-85 degree comparisons tend to be greater than those for 60-60, which, in turn, tend to be greater than those for 30-30 for all comparisons made in this study. For example, the 50th percentile median line increases as a function of angle from 0.9927 for the 30-30 comparisons to 0.9953 for the 85-85 comparisons in Figure 3. A similar trend is seen in Figure 4 and 5 for the true nonmatch and different side data. This observation is believed to be related to the quality of the mark that can be produced by employing the jig used in this study. During production of the marks it was noticed that it was much easier to produce a mark of the entire width of the screwdriver tip at the higher angles than at the lower angles. This is only logical since, for a given applied force, the resolved force component perpendicular to the plane of the lead test medium increases as the sine of the angle of attack. This results in more complete marks being available for determining a match. As more information is available, the quality of the match might be expected to increase. However, there is also the possibility that as the number of striae increases, the chance of finding a random segment that matches well from two different screwdriver tips also increases. Thus, in both cases, the average correlation value increases. This observation reinforces the assertion that

the correlation values computed as described here cannot be interpreted as in the usual case of a single pair of matched random samples, or as direct evidence of the "quality" of match, but rather as a general index of pattern similarity that is influenced both by the tool used in making each mark and by other factors (such as the tool angle) as well.

Comparisons that produced the most extreme data values (those plotted as individual points beyond the whiskers in the box plots) were reviewed by the author who is an experienced examiner [5]. Using a comparison microscope, the corresponding pairs of profilometer traces were examined for abnormalities. In the case of the most extreme outliers it was found that due to the poor quality of the markings, no match was possible. For points nearest to, yet still outside, the box plots, observations fell into two categories. In some instances regions within the sample where a match could be made were seen, yet often this region was not within the area sampled by the profilometer scan. In other cases, border-line matches were visible, which might prompt an experienced examiner to make additional markings to improve the quality of the mark before declaring a positive match. The conclusion drawn from these observations is that the algorithm does a very good job of comparing scans for the regions examined, with the obvious shortcoming being that it does not have the flexibility available to examiners to make more marks or examine other areas.

Summary and Conclusions

This study has examined a series of toolmarks produced in soft lead plates by screwdriver tips that have been dragged across the surface at various angles. The surface roughness of the resultant marks was measured and the height, depth, and width of the resultant striae were analyzed using a computer search / match algorithm. From the results of this study the following conclusions can be drawn:

1. Comparison of toolmarks involving true matches, true nonmatches, and marks made from different sides of the same tool, evaluated using the maximum Pearson correlation computed from data segments, all produce high correlation values. This suggests that correlation value alone is a poor means of determining when an actual match exists.
2. A significant separation in correlation values is observed between true-match and true-nonmatch pairings when toolmarks produced at the same angle are compared. This observation suggests that it should be possible to identify true matches by means of a computer algorithm, if maximum correlation values can be effectively calibrated.
3. Matching of toolmarks made from different sides of the

same tip produce no separation in data and are similar to data from true nonmatches. This supports the hypothesis that different sides of a screwdriver act as different tools when producing toolmarks, and is also in agreement with observations of forensic examiners.

4. Maximum correlations computed from matching samples, but prepared at different angles, were similar to maximum correlations computed from nonmatching samples at the same different angles. That is, there is effective separation of true match and true nonmatch comparisons ONLY when the tool angles are equal. This supports the accepted practice of examiners where efforts are made to duplicate the angle of attack used when an evidentiary toolmark was made.

Acknowledgments

The authors would like to thank Stan Bajic and David Baldwin of the Midwest Forensic Resource Center for use of their comparison microscope, and Matt Besser for training and assistance with the stylus profilometer. This work was supported by the National Institute of Justice under contract 2004-IJ-R-088, John Hihn contract manager, and was performed in part at Ames Laboratory, which is operated under contract No. W-7405-Eng-82 by Iowa State University with the US Department of Energy.

References

1. Grzybowski, R.A. and Murdock, J.E., "Firearm and Toolmark Identification-Meeting the Daubert Challenge," *Journal of the Association of Firearms and Toolmark Examiners*, Vol. 30, No. 1, Winter 1998.
2. Davis, J.E., "An Introduction to Toolmarks, Firearms and the Striagraph," Charles C. Thomas—Publisher, Springfield, Illinois, 1958.
3. Faden, D, unpublished research.
4. Vardeman, S.B. and Jobe, J.M., *Statistical Quality Assurance for Engineers*. John Wiley and Sons, New York, 1999.
5. Kreiser, J., Toolmark Examiner for the State of Illinois, retired.

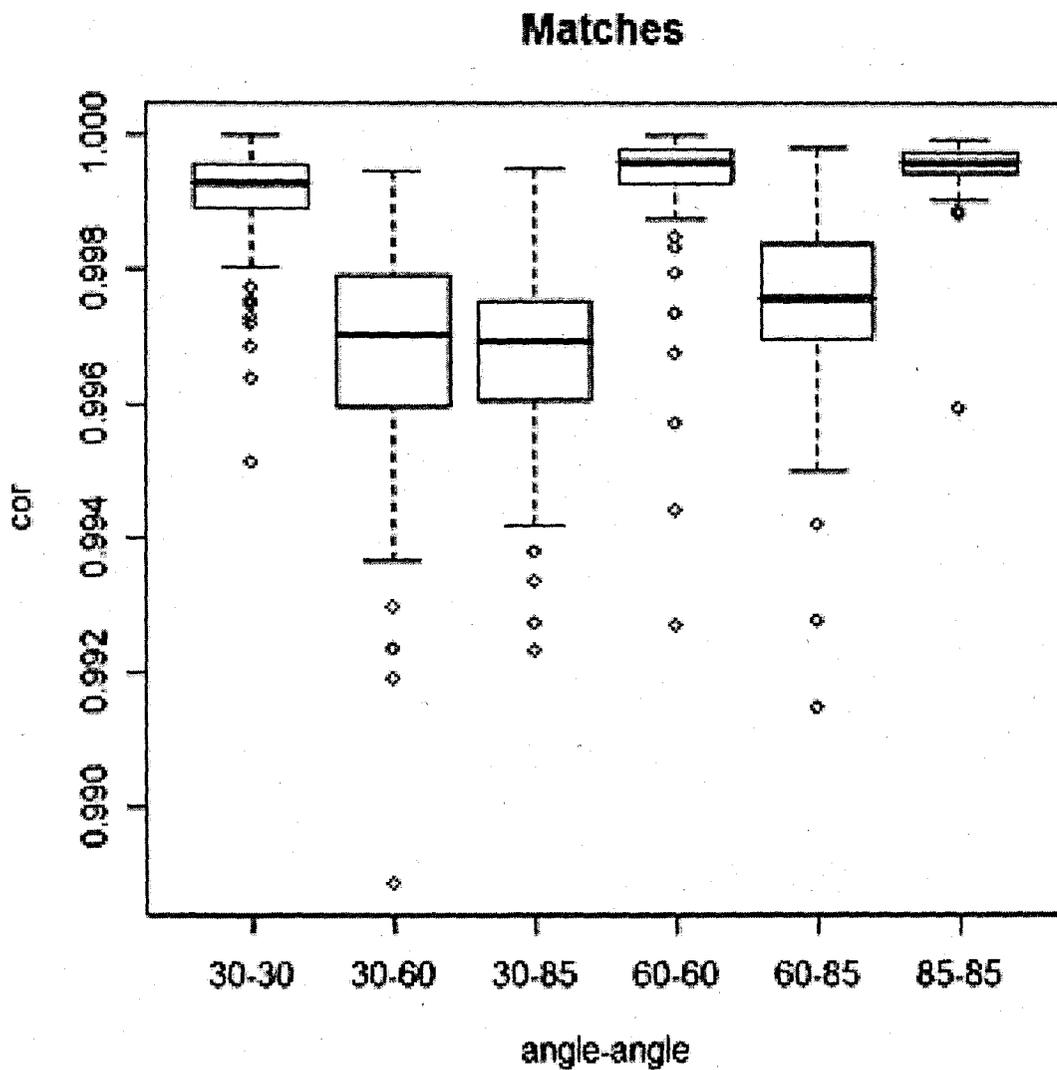


Figure 3: Summary of all results for the six angular comparisons for the true match data set.

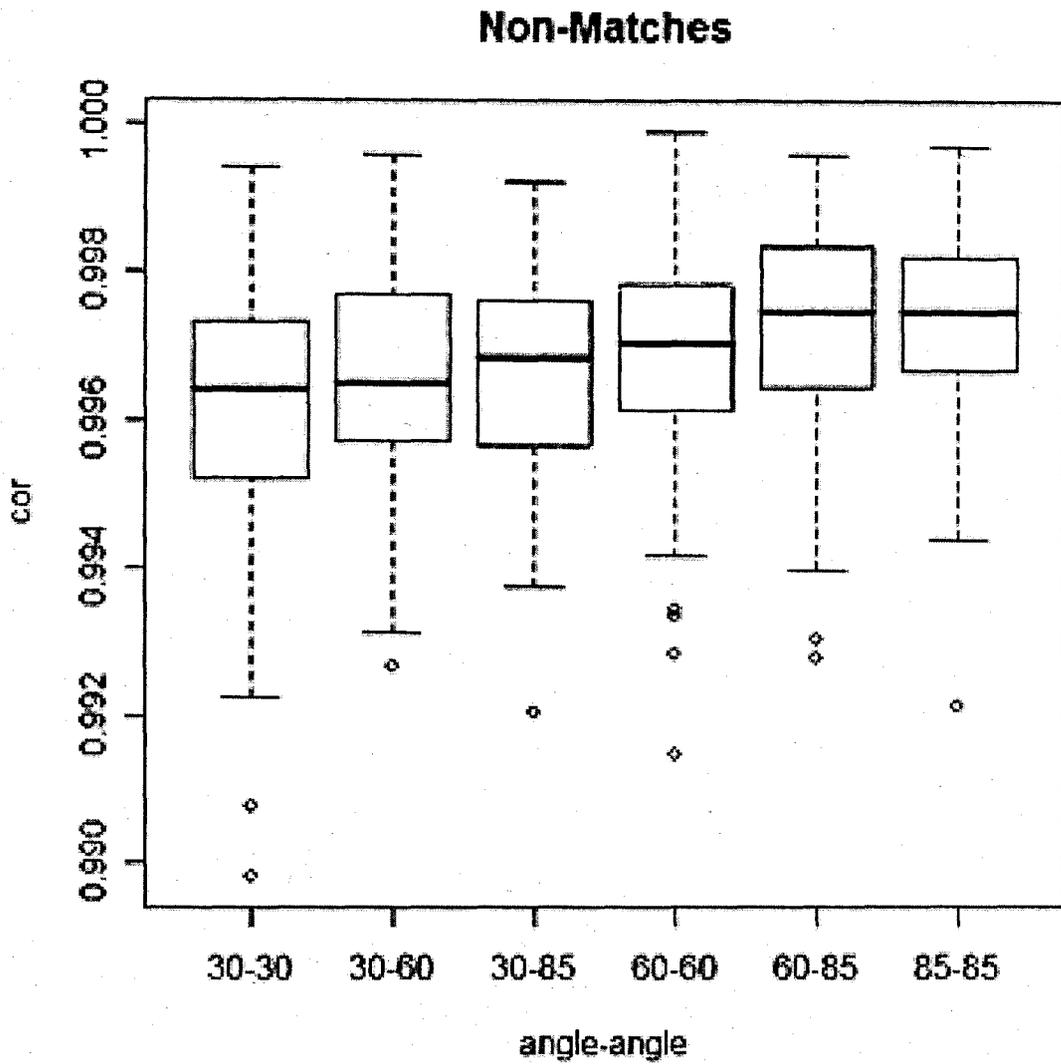


Figure 4. Comparison of true non-matches.

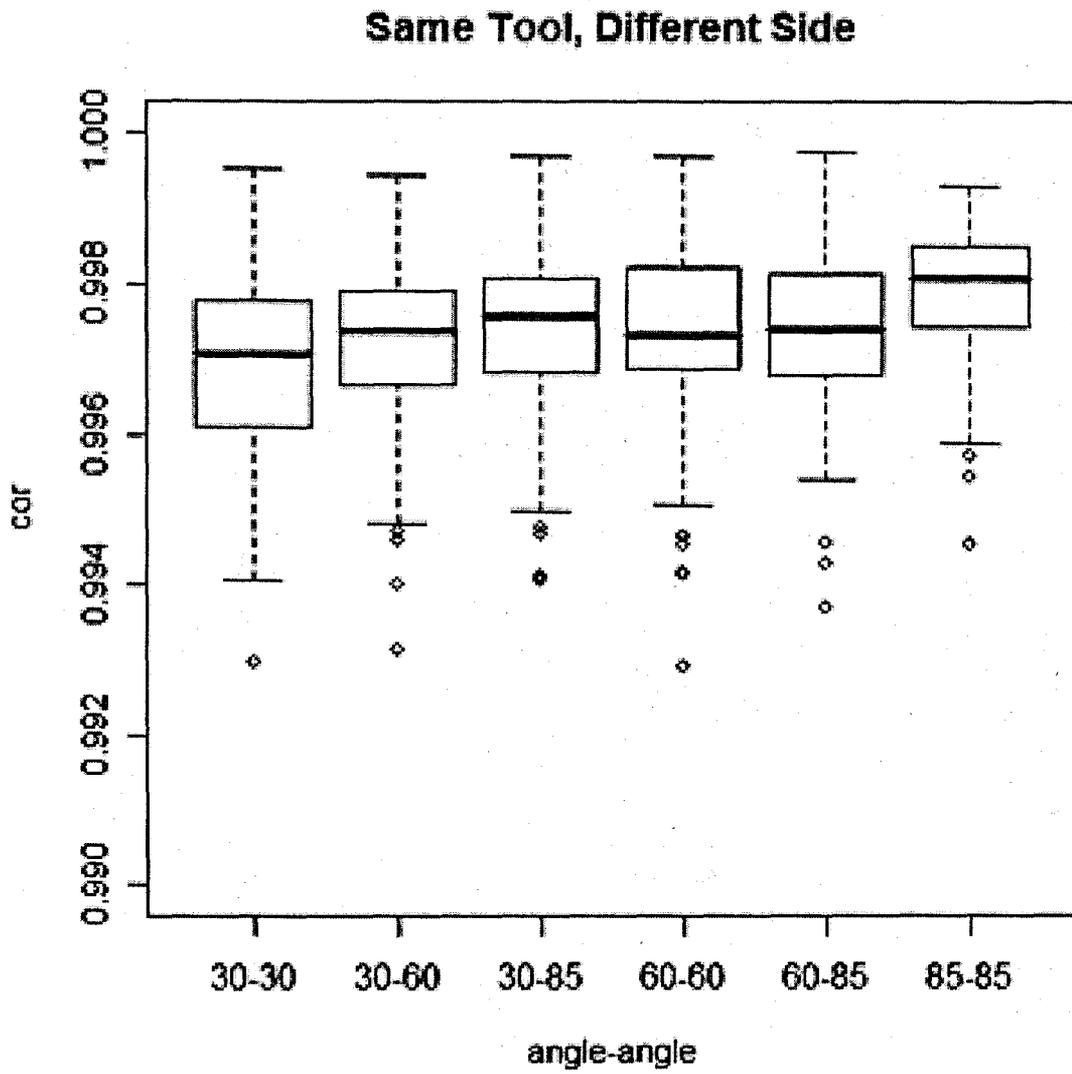


Figure 5: Comparisons of side A to side B for screwdriver tips 1-44.

Means of the Six Angular Comparisons

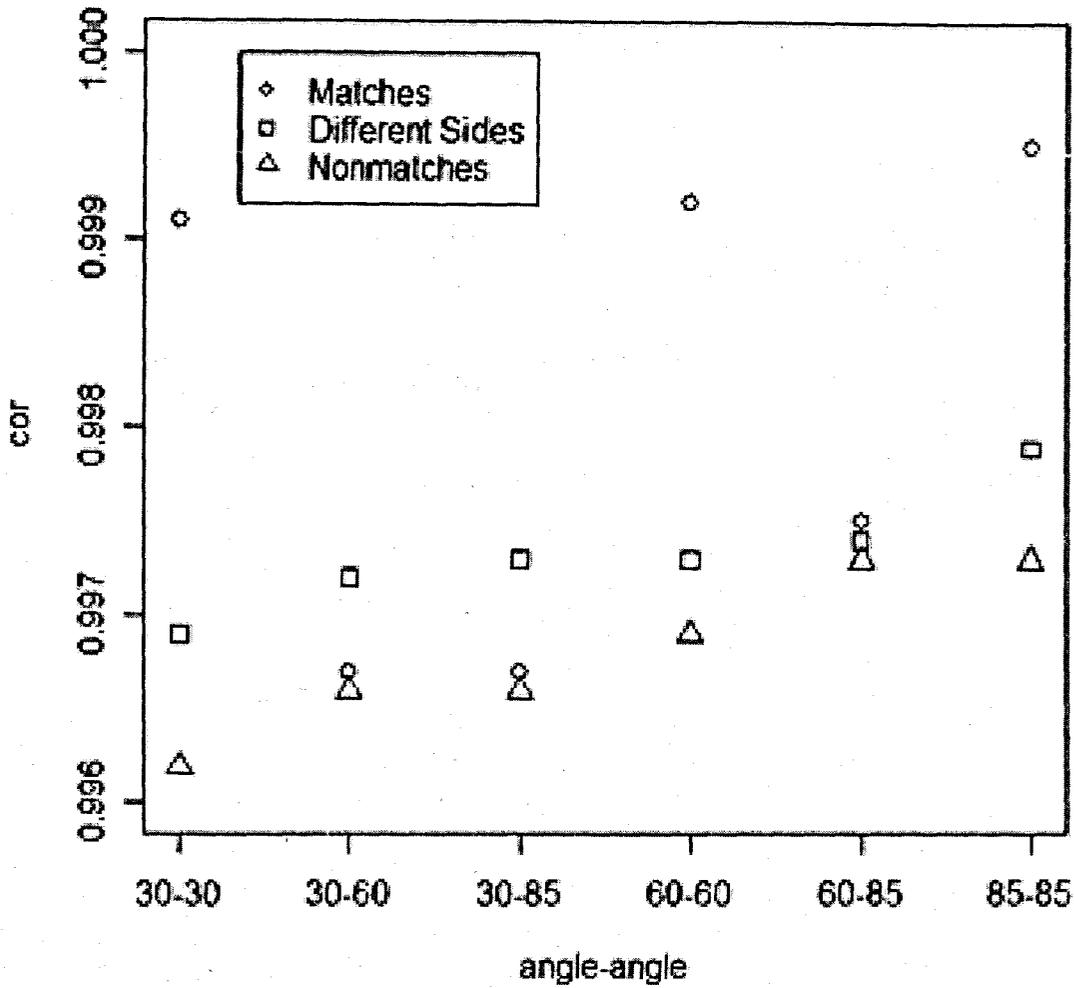


Figure 6: Plot showing the means of the three data sets (true matches, true non-matches, and Side A and B comparisons) examined in this study.