Objective Comparison of Toolmarks from the Cutting Surfaces of Slip-Joint Pliers

By: Taylor Grieve[1], L. Scott Chumbley[1], Jim Kreiser[2], Max Morris[1], Laura Ekstrand[1], Song Zhang[1]
[1]Ames Laboratory, Iowa State University, 2220 Hoover, Ames IA 50011
[2]Illinois State Police, Retired, 3112 Sequoia Dr., Springfield, IL 62712

Keywords: algorithm, known match, known non-match, profilometry, statistical comparison, slip-joint pliers, toolmark comparison, striae

ABSTRACT

Experimental results from a statistical analysis algorithm for objectively comparing toolmarks via data files obtained using optical profilometry data are described. The algorithm employed has successfully been used to compare striated marks produced by screwdrivers. In this study, quasi-striated marks produced by the cutting surfaces of slip-joint pliers were examined. Marks were made by cutting both copper and lead wire. Data files were obtained using an optical profilometer that uses focus variation to determine surface roughness. Early efforts using the comparative algorithm yielded inconclusive results when the comparison parameters used were the same as those employed successfully for screw-driver marks. Further experiments showed that the algorithm could successfully be used to separate known matches from non-matches by changing the comparison parameters. Results are presented from the analysis of the copper wires.

Introduction

In the last twenty years, several different court cases, have called into question the validity of scientific testimony, including those relating to firearm and toolmark examinations. As a result, recent research has sought to justify a basic assumption made by forensic examiners: each tool makes its own unique mark. Many different tools and their marks have been examined in the research setting including screwdrivers [1-4], tongue and groove pliers [4, 5], and chisels [3].

Striated screwdriver marks have been well studied and characterized by stylus profilometry and confocal microscopy. These characterizations have been used to analyze potential matches and non-matches via statistical validation in several different studies [1-4]. In general, the results have shown that striated marks can be compared objectively using computer algorithms with a fairly high success rate. Studies of somewhat irregular marks also exist, although to a lesser extent. Cassidy first published a study on the examination of toolmarks from sequentially manufactured tongue and groove pliers, as they are frequently used to twist off doorknobs to break into buildings [5]. This study, while not based in statistical validation, did establish that the tongue and groove pliers only produce individual characteristics due to the teeth being broached perpendicular to the direction of the striated mark. Bachrach et al. more recently examined the marks produced by the application of tongue and groove pliers to different materials (lead, brass and galvanized steel) and used statistical comparisons to objectively compare the marks [4]. Bachrach et al. found the marks from tongue and groove pliers could readily be compared when made on the same media. However, the empirical error rate increased when comparing marks made on different media. Chisel marks have been evaluated by Petraco et al. [3], but the patchy striated chisel marks used in this research proved too difficult for the developed suite of software currently in use to provide useful information during comparison. Thus, while a small body of work exists on less than perfectly striated marks, the results are somewhat disappointing at this time.

In a previous study [2], fifty sequentially manufactured screwdriver tips and their corresponding marks made at different angles were examined and compared using a statistical algorithm to determine the strength of evidence of a positive match between a mark and the tool that made it. This algorithm has been used extensively to evaluate the evenly striated marks of screwdrivers; however, it has not yet been used to evaluate less striated marks or impression marks. As a first step toward investigating the applicability of the current algorithm, quasi-striated marks such as those made by slip-joint pliers when cutting wire were examined. Slip-joint pliers were chosen since no studies currently exist on this subject to the authors' knowledge. Additionally, they were expected to produce a more difficult mark for analysis, due to the manner in which cutting occurs. When cutting a wire with slip-joint pliers, the mark produced reflects both striations from the actual cutting, and smearing due to shearing of the material during the process. This results in a mark that is not continuous from the beginning of the cut to the end. Thus, the surface topography that exists at the initial cut edge of the mark could vary substantially from what is seen at the final cut edge.

Date Received: March 29, 2013
Peer Review Completed: December 27, 2013

Experimental

For this experiment, 50 pairs of sequentially manufactured slip-joint pliers were purchased from Wilde Tool Co., Inc. so as to be as nearly identical as possible. It is well known that the manufacturing process greatly affects the resulting toolmarks a tool makes due to the surface features imparted on the tool during manufacturing [6, 7]. For this reason, a detailed description of the way the pliers used in this study were manufactured is in order.

All of the plier-half blanks examined in this study were hot forged from the same die, followed by cold forging from the same forging die. Following forging, holes were punched to seat the fastener, i.e. the bolt that will hold the two halves of the pliers together. At this point, a difference is introduced in the blanks. On slip-joint pliers, one half of the pliers has a small hole, while the other half has a larger, double hole (the "slip-joint") allowing the user to gain a better grip when using the pliers on a larger-diameter object (see **Figure 1**). Once the plier holes were punched, the teeth and shear cutting surfaces were created using a broaching process. It is this machining method that creates the scratch minutiae on the surface of the plier halves responsible for producing the characteristic toolmark that is of interest in forensic examinations.

The plier halves for this study were cut on two separate broaching machines; halves with the smaller hole were all broached on one machine, while the halves with the double hole were broached on a second. At this point in the process the manufacturer stamped numbers 1-50 on each plier half as they were finished being broached. Thus, the 50 pairs could be assembled with confidence that they were actually made sequentially. After broaching, both halves were given the same heat treatment and shot peened to surface harden the metal. The long, flat surface was then polished and the pliers were assembled and gripped. As a final step, the company branded the double-hole side of each pair of pliers. For the purposes of this study, each half of the pliers was designated as either Side A or B, with Side B being the branded half of the pliers (see **Figure 1**).

To make the samples, copper wire of 0.1620" diameter and lead wire of 0.1875" diameter were obtained and cut into two-inch lengths with bolt cutters to distinguish the ends from the cuts made by the pliers. Next, the cut lengths of wire were placed centered in the plier jaws on the cutting surface with pliers side B facing down. Alternating shear cuts of lead and copper were made with each pair of pliers for a total of 21 cuts. All odd-numbered cuts were lead samples; all even-numbered cuts were copper. The total number of copper samples thus obtained was 1000, with 500 cuts in contact with Side A, and 500 cuts with side B.

Figure 1: Slip-joint pliers in their unfinished and finished states. From left to right: plier halves (single and double hole) before broaching; an example flat side of pliers that will be polished; finished and labeled pliers (sides A and B).

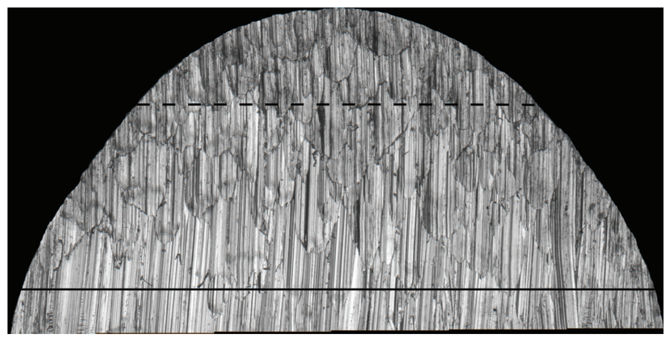

Figure 2: Areas of cut copper wire examined during comparisons. Dashed line is referred to as the "short edge," the solid line is referred to as the "long edge"

For the purpose of this study, only the copper samples were evaluated. Each cut mark surface was scanned optically with an Alicona Infinite Focus G3 profilometer at 10x magnification to acquire the surface geometry of the mark. An example of a typical scan is shown in **Figure 2**. The toolmark is seen to be quasi-striated, i.e. parallel linear striae do exist but the consistency of these striae clearly varies across the surface of the cut mark.

When the data are acquired, noise spikes occur around the edges of the mark where the cut surface drops off because there is no surface here for the profilometer to scan. This noise is non-informative for the matching process, and is not desirable in the data file. Therefore, the raw data are processed using a computer routine to remove the extraneous noise spikes. This process is referred to as a cleaning routine and does not affect the data that characterizes the cut surface. An example of a raw versus a cleaned data file can be seen in **Figures 3 and 4**.

All raw data files contained trended data. Simply put, due to the manner in which the data were collected, the line profile of a mark data file had an increasing linear trend in the z direction moving from one side of the mark to the other. Such a trend is common when using profilometers since the surface analyzed is rarely exactly parallel with the direction of scanning. Because the files were a rectangular collection of 3D data (shown in the raw data of **Figure 3**), trending was corrected by subtracting a plane matching that of the trended data from the file. To accomplish this, the detrending routine selects left and right diagonal points from the data (approximately 40 on each side, 80 in total) and uses a linear least squares method to fit the appropriate plane for the data. It then subtracts the fitted plane from the data to achieve an appropiately leveled data file for comparison.

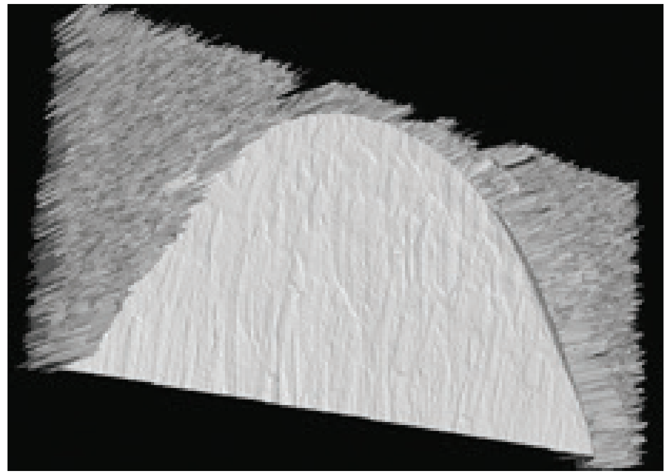

Figure 3: Raw data

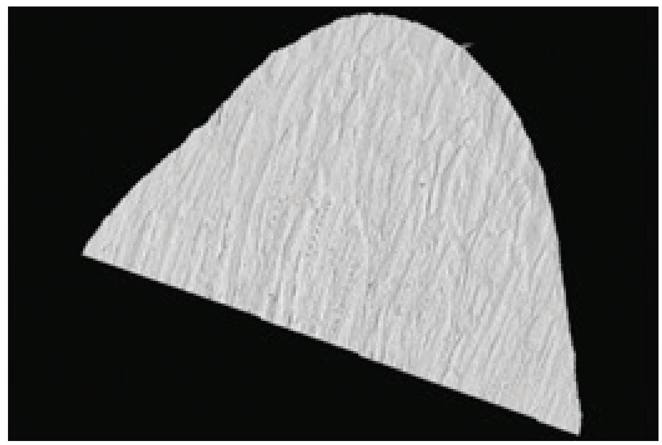

Figure 4: Cleaned data with noise spikes removed

Comparisons between the marks were made using the previously described algorithm [2]. The comparisons were divided into two different groups, those made close to the end of the mark, as designated by the solid line in **Figure 2**, and those made close to the start of the mark, shown by the dashed line in **Figure 2**. From this point on, the dashed line data will be referred to as the short edge and the solid line data as the long edge. These mark locations were chosen to examine differences between the beginning of the cut, where the mark has short and variable length striae, and the end of the mark, where the striae are longer and appear to be more regular.

Each side of the pliers was considered to be a separate data set, the assumption being, as confirmed by forensic examiners, each side acts as a different cutting surface. Given there are 50 pairs of pliers, with two sides for each pair of pliers and ten replicate cuts for each side of each pair of pliers, the total number of samples possible for examination came to 1,000 discrete data sets.

Results

A sampling format was set up to compare three different groups of data: known matches, known non-matches from the same pair of pliers (i.e. different sides), and known non-matches from different pairs of pliers. The comparison setups are as follows:

Set 1: Compare known matches. These should be marks from the same side of pliers. Comparisons were made between marks 2 and 4 and between marks 6 and 8 for each side of the pliers, side A and side B.

Set 2: Compare known non-matches from the same pair of pliers. Comparisons were made between side A and side B for marks 10, 12, and 14.

Set 3: Compare known non-matches from different pairs of pliers. The samples were divided into 12 groups of four, each numbered consecutively, e.g. tools 1-4, 5-8, etc. Comparisons were made for both side A and side B. **Table 1** shows an example comparison setup for the first group of pliers.

The same algorithm used in an earlier work for striated marks [2] was applied in this study to examine the quasi-striated marks made by the slip-joint pliers. The algorithm has two primary steps: Optimization and Validation. During the Optimization step, the regions of best agreement between the two marks are determined by the maximum correlation statistic, or "R-value." The size of the region is assigned by the user and is hereafter referred to as the "Search Window." The second step of the algorithm, Validation, uses both rigid and random window shifts to verify the regions chosen in the Optimization step indeed correspond to a true match. These

Comparison	Plier number	Side	Mark number	Plier number	Side	Mark number
A	1	A	16	2	A	16
B	3	A	16	4	A	16
C	1	A	18	4	A	18
D	2	A	18	3	A	18
E	1	A	20	3	A	20
F	2	A	20	4	A	20

Table 1: Comparisons for Set 3, Group 1

windows are hereafter referred to as the "Valid Windows" and their width is also user determined. The R-values in this step must clearly be lower than the R-value in the Optimization step, as the highest R-value has already been calculated. However, in the instance where a true match exists, the R-values associated with the rigid shift valid windows should be larger than those associated with the random shift valid windows; the assumption being, if an excellent match exists at one location then very good matches should exist at any number of corresponding locations. If true, this is indicative a true match does exist. Conversely, rigid window shifts do not produce systematically larger R-values than random shifts in the case of a true non-match, since the high values found during the Optimization step exist due to random chance rather than any physical relationship between the items being compared. Further discussion of this algorithm can be found in the literature [2].

Originally, the size of the search and valid windows were set at the comparison software's default 200 and 100 pixels, respectively, and the comparisons were conducted with samples from the first 20 pairs of pliers. This setup produced 400 different comparisons for the long and short edge comparisons. When a comparison is made, indication of a true match is found when the T1 value of the statistic returned is relatively high. Little or no relationship between the marks results in T1 values centered near 0.

Results of these early comparisons can be found in **Figures 5 and 6**. In these box plots, the bold line in the middle of the box represents the median, the lower quartile by the bottom line of the box, and the upper quartile by the top line of the box. The whiskers are one and a half times the difference between the upper and lower quartiles. Any outliers outside the whiskers are denoted by dots. In these plots, known matches are in the comparisons designated Set 1, while Sets 2 and 3 show comparisons between known non-matches from different sides of a pair of pliers and non-matches between different pairs of pliers, respectively. It is evident that with these window sizes, the success of identifying known matches

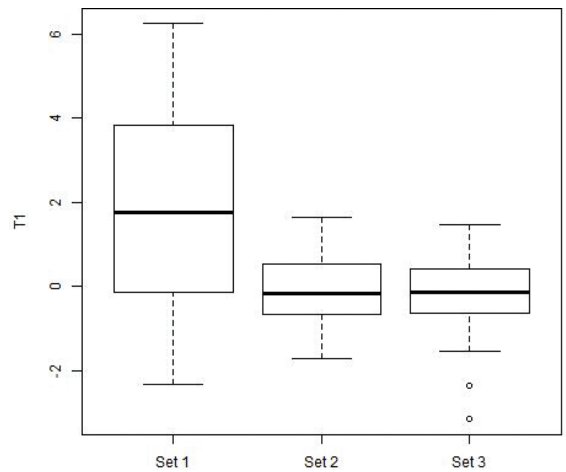

Figure 5: Original data comparisons for short edge

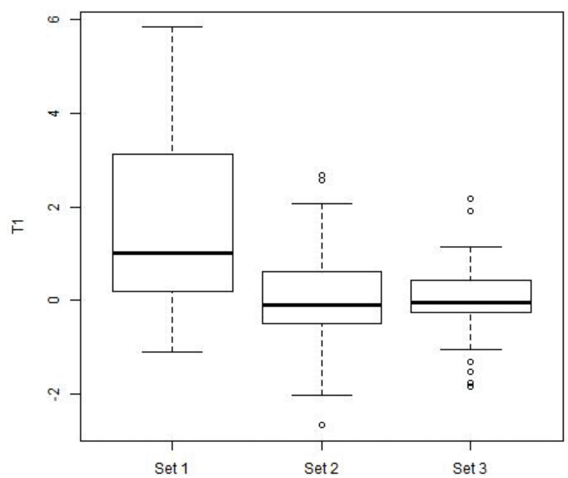

Figure 6: Original data comparisons for long edge

was relatively low, there being little separation between the returned T1 values of known matches and non-matches.

From the minimal success of the first attempt at matching the plier marks, several changes were decided upon for further comparisons. First, the data shown in **Figures 5 and 6** compared trended data. This was corrected in subsequent comparisons. Second, it was decided to vary the window size for all plier mark samples. The initial values used were chosen simply because they had proven effective for comparison of fully striated marks. A series of experiments was conducted within each plier comparison set where the window sizes were varied to evaluate the effect window size has on the resulting T1 value. In other words, the question asked was: Does the size of the window play a large role in the discrimination between known matches or known non-matches? In this series of experiments Search and Valid windows were assigned four different values. The Valid window was always half the size of the Search window. Search windows were set at values

100, 200, 500, and 1000 pixels, respectively, to examine the effects of one smaller Search window and two larger Search windows. These new settings were extended to all 50 pairs of pliers and their corresponding toolmarks in the copper wire, bringing the total number of comparisons to 3,952.

The results of these comparisons can be found in **Figures 7 through 12**. Observation shows that the T1 value increases dramatically with increasing window size. While known non-matches return values centered around zero regardless of window size, the T1 value for known matches increases from just slightly over zero to an average of 6.36 and 6.09 for the largest window size for the long and short comparisons, respectively. However, the data range increases as well. At the larger window sizes, numerous outliers exist and failure of the algorithm occurs in some cases, especially for the short edge comparisons.

The large number of observed failures directly results from the constraints placed on the way the Search and Valid windows are chosen and compared. One of the standard conditions under which the algorithm operates is the Search and Valid windows are never allowed to overlap. In some cases, especially with the short edge comparisons, the shorter length of line from which data can be selected and compared results in far fewer data points for comparison. This problem is exacerbated as the window sizes increases. For larger sizes, there simply is not enough data available to meet these conditions in all instances. Thus, this stipulation can cause the algorithm to return no T value.

Table 2 summarizes the instances in which the algorithm failed to return values. It can be clearly seen that the return rate decreases with the shorter line profiles as the window size increases. As a reference, Set 1 has a total of 200 comparisons, Set 2 has 150 comparisons and Set 3 has 144 comparisons.

Long edge comparisons				
Set	100-50	200-100	500-250	1000-500
1	0	1	1	1
2	0	1	3	3
3	0	2	3	5
Short edge comparisons				
Set	100-50	200-100	500-250	1000-500
1	0	0	1	9
2	1	0	3	19
3	1	0	3	24

Table 2: Cases in which the algorithm returned no T values for each window size

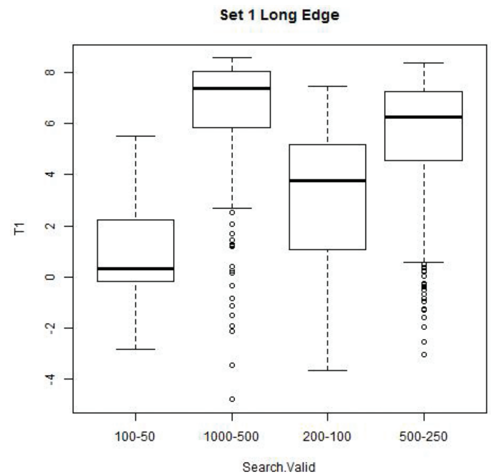

Figure 7: Long edge comparisons: Known matches from the same set of pliers

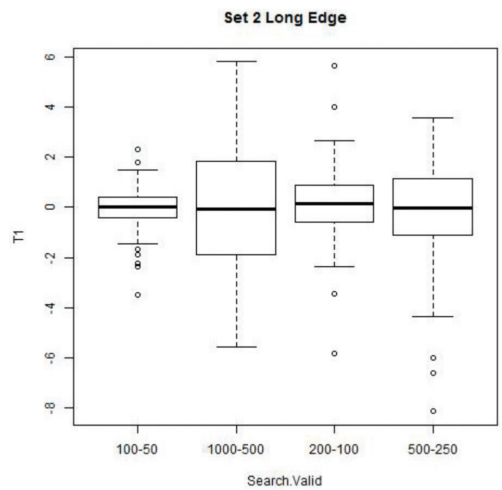

Figure 8: Long edge comparisons: Known non-matches from the same set of pliers

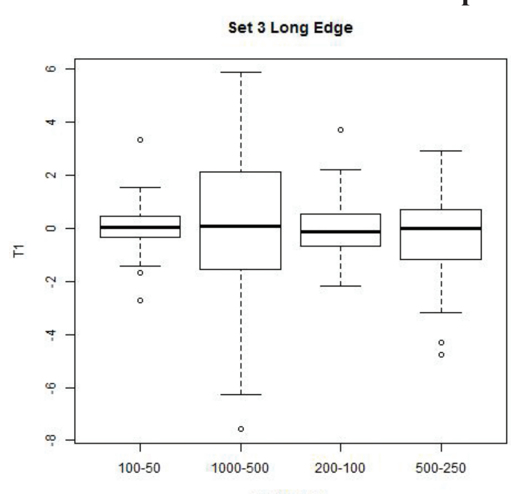

Figure 9: Long edge comparisons: Known non-matches from different sets of pliers

As a first attempt at a solution, two additional window ratios were examined: 4 to 1 and 6 to 1. It was hoped that by limiting the size of the Valid windows, less spread in the data would be seen. For each new ratio, four different window sizes were chosen and the algorithm was run again following sets 1, 2, and 3 at both the long and short edge locations on the mark. For these exploratory tests, the data were limited to pliers 1-25; the assumption being the abbreviated data set would be representative of the full 1-50 pliers data. Results of this examination can be found in **Figures 13 through 24**. This set of parameters does indeed appear to have a significant effect in reducing the number of outliers and spread of the known matches (i.e. Set 1) as compared to the 2:1 ratio data. A slight degradation in the maximum values obtained was seen for the known matches. Less change is seen in the results for the known non-matches (Sets 2, 3). Average values still were centered around zero and spread seemed to increase somewhat in some cases for the known non-matches.

Discussion

When using the developed algorithm, ideally the data should show a clear separation between T1 values for known matches as opposed to known non-matches, with no overlap occurring, even when considering outliers. While elimination of overlap in the outliers has not been achieved with the current algorithm, it is clear that a high degree of separation is seen in the majority of cases when the search parameters are adjusted from the defaults used for the (comparatively) well-striated screwdriver marks. This suggests that the current algorithm is more robust than it initially appeared, and could be suitable for discrimination if performance can be enhanced and the spread in the data can be decreased to produce complete separation between known matches and non-matches. These tests also indicate the size of the Search and Validation windows can have a critical role in determining when a match can be discriminated from a non-match. Since the size and number of Valid windows is user defined, future work must involve a series of experiments to determine what operation parameters are best suited for each individual class of marks. For example, the relatively small Search and Valid window sizes that worked well for screwdriver marks were inadequate for the plier marks. However, increasing the Search and Valid window size proved effective in producing a clear separation between known matches and non-matches for slip-joint pliers and changing the size ratio has an effect on the spread of the data.

Outliers are seen in all the data sets, both known match and known non-match. Examination of these data files points to a consistent problem with the current state of the algorithm, which the authors refer to as the "opposite end" match problem.

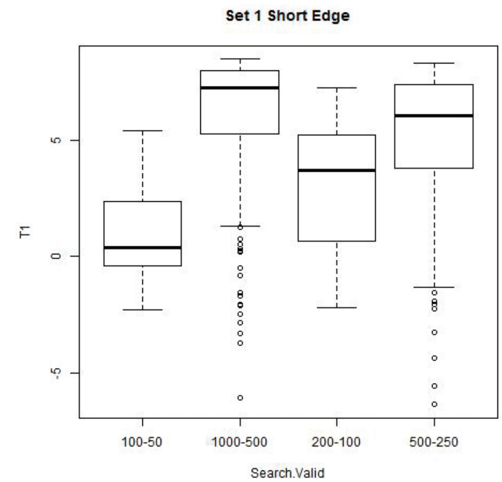

Figure 10: Short edge comparisons of Set 1: Known matches from the same set of pliers

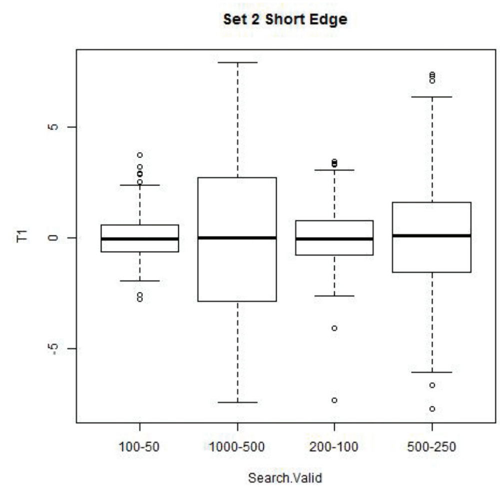

Figure 11: Short edge comparisons of Set 2: Known non-matches from the same pair of pliers

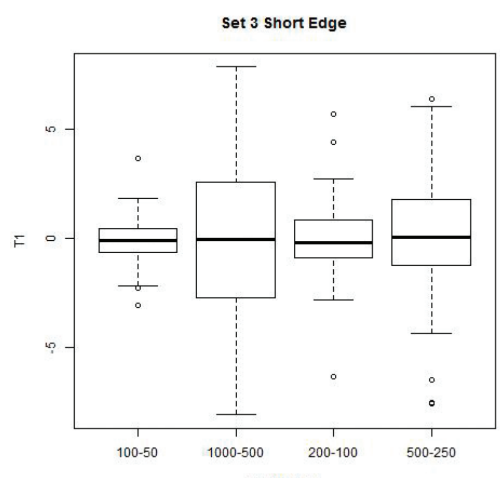

Figure 12: Short edge comparisons of Set 3: Known non-matches from different pairs of pliers

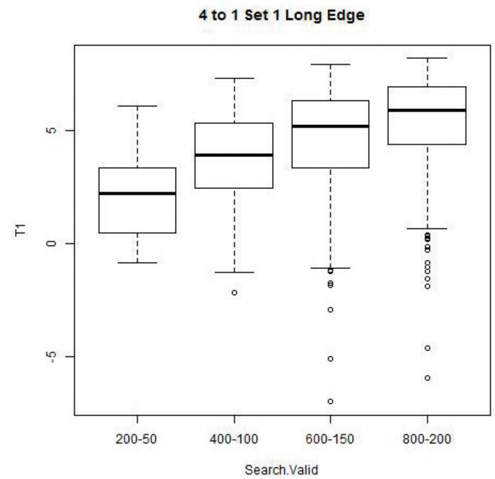

Figure 13: Results of varied ratio long edge comparisons

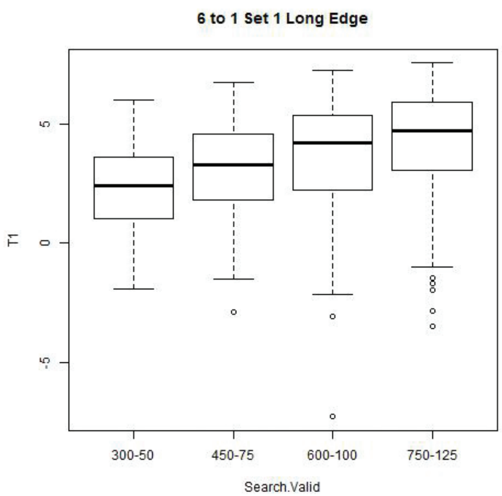

Figure 14: Results of varied ratio long edge comparisons

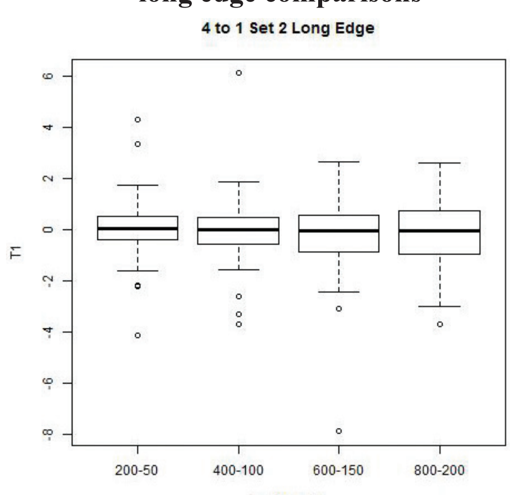

Figure 15: Results of varied ratio long edge comparisons

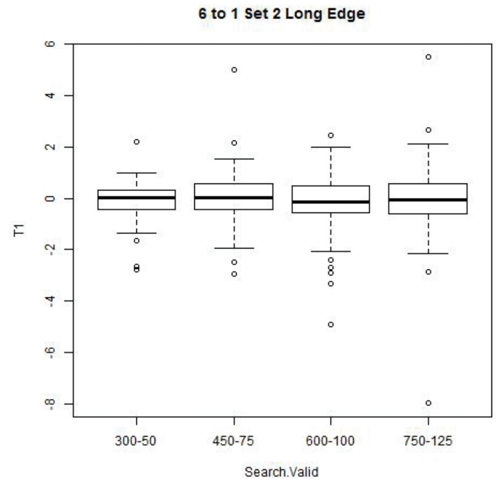

Figure 16: Results of varied ratio long edge comparisons

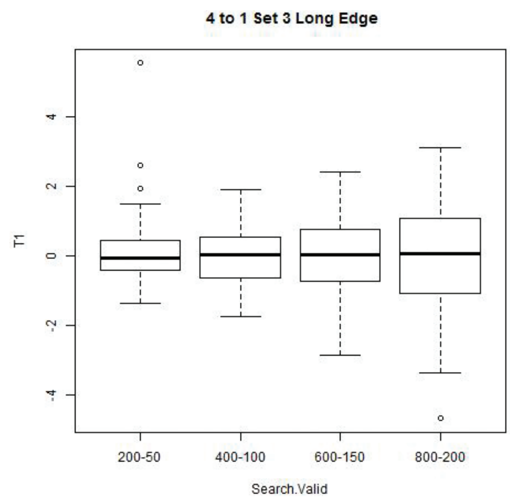

Figure 17: Results of varied ratio long edge comparisons

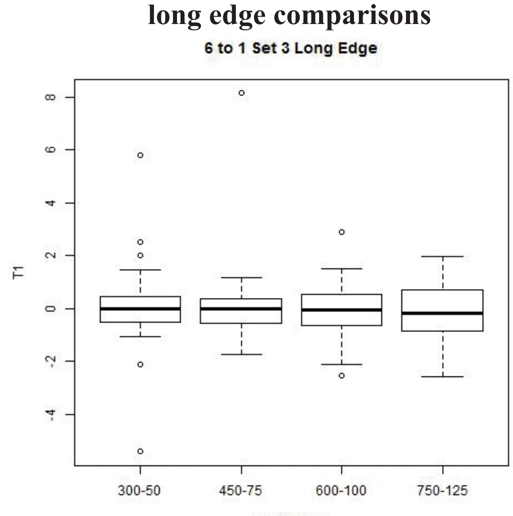

Figure 18: Results of varied ratio long edge comparisons

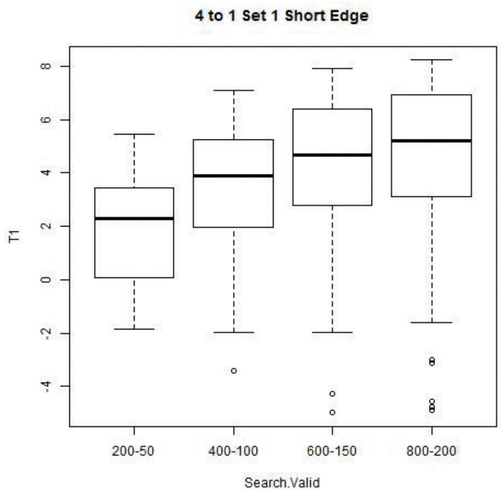

Figure 19: Results of varied ratio short edge comparisons

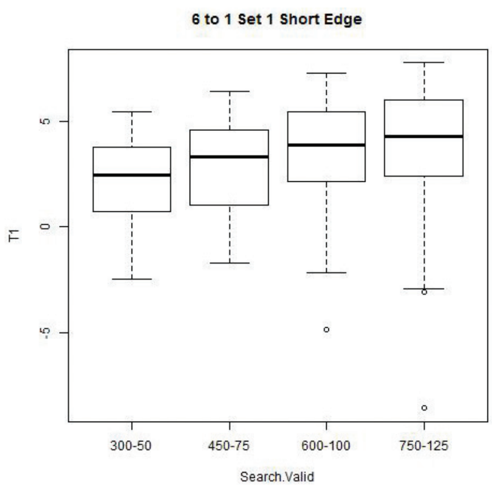

Figure 20: Results of varied ratio short edge comparisons

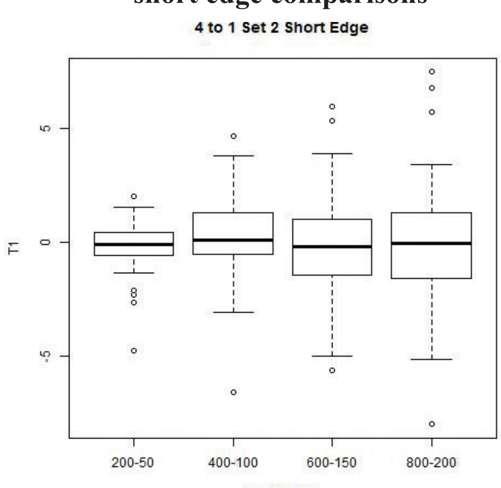

Figure 21: Results of varied ratio short edge comparisons

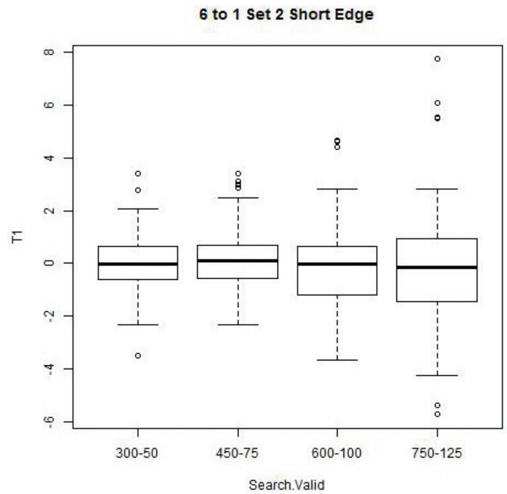

Figure 22: Results of varied ratio short edge comparisons

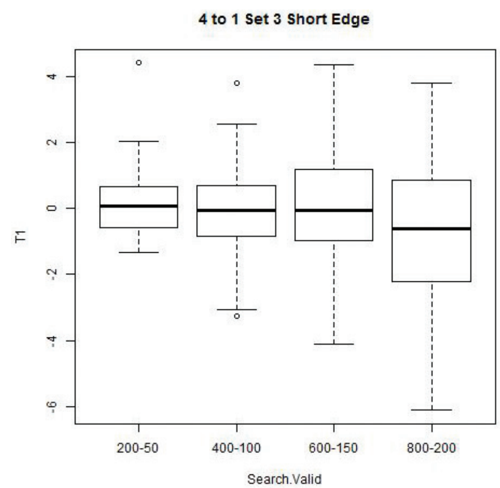

Figure 23: Results of varied ratio short edge comparisons

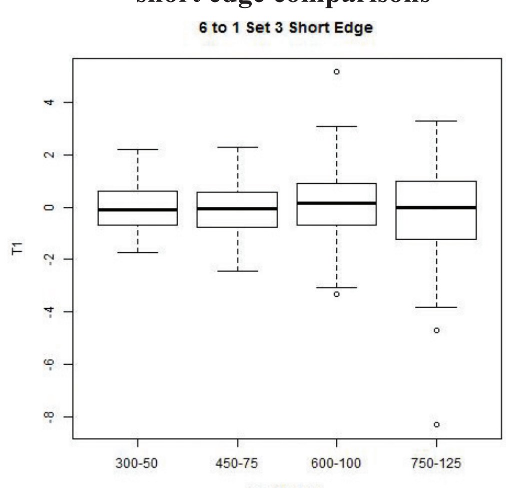

Figure 24: Results of varied ratio short edge comparisons

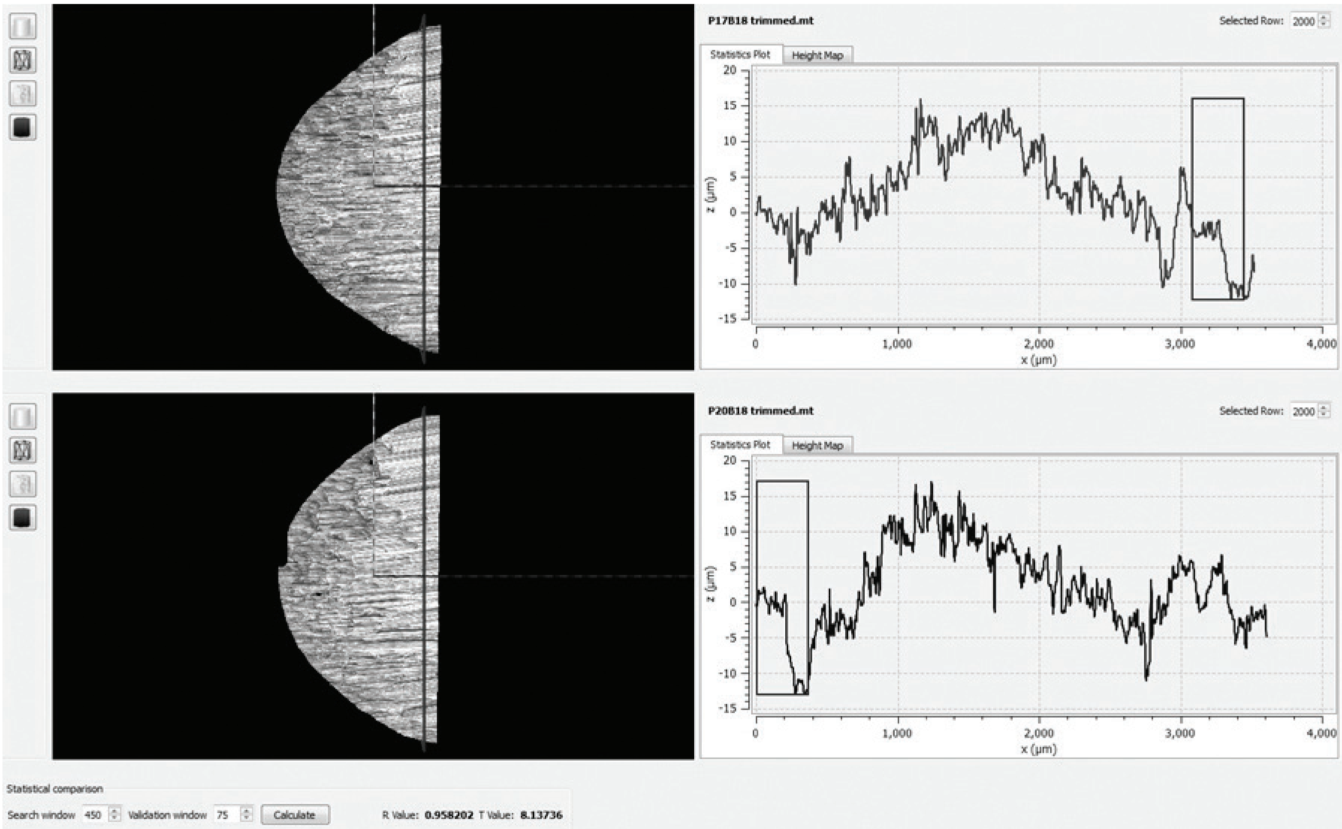

Figure 25: Incorrect opposite ends match for long edge comparison of known non-matches from different pairs of pliers. The search and valid windows were 450 and 75. T1 value is 8.137

This seems to be an area where further improvements can be made. In earlier work involving screwdriver comparisons [2], it was noted the algorithm often returned false match values, incorrectly identifying the match areas on opposite ends of the mark's cross-sectional profile. "Opposite end" matches appear to occur most often in known non-matches, however non-match values have been returned for known matches as well with similar opposite end match problems. In detrending the data, many of these problems have been eliminated; however, a few opposite end match problems still exist. One such example can be seen in **Figure 25** for a plier comparison datafile, which consists of detrended data. One data set is shown at the top while the second is shown at the bottom. Simple chance where the opposite ends of the mark have a very similar profile over the small area of the search window, as denoted by the box, has resulted in the computer declaring an excellent match. Obviously, such a match is physically impossible, no matter how good the numbers.

In its current form, the algorithm has maximum flexibility, allowing marks to be compared along a linear direction both forwards and backwards. Such a methodology requires no contextual information to be known about the mark. A fully striated mark may leave few clues as to what is the "left" side of the mark vs. the "right" side, as determined by how one holds the screwdriver, **Figure 26**. As shown by the bold arrows, pulling the screwdriver across the surface in opposite directions results in marks having the same general appearance, even though they are oriented 180 degrees to each other . While this situation is usually easily recognized by a trained examiner making a test mark, it is more of a problem for an automated system. To the machine, both situations

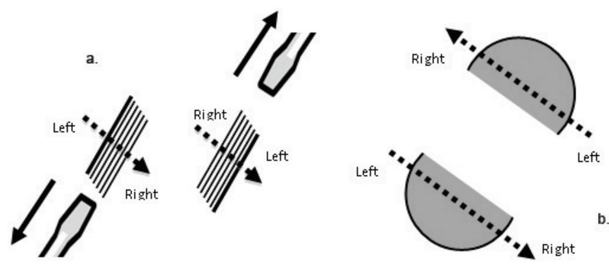

Figure 26: a) Fully striated marks hold few clues to "left" vs. "right for the automated scan as denoted by the dashed line. b) Cut wire sample scan directions are easily distinguishable by the macroscopic shape

result in a series of parallel lines. If the scan is constrained to run comparisons in only 1 direction (dotted line), this match may be missed since "left" could be viewed as "right" and vice versa. For this reason, the current algorithm is written to be as flexible as possible with comparisons run in both directions so it is not necessary to know which side of the mark was on the left and which was on the right as it was being made.

Determining the correct scanning direction is less of a problem for a cut wire, where contextual information such as "left" and "right" can be easily assigned due to the macroscopic shape of the object itself, as exemplified in **Figure 26(b)**. In this instance the situation is somewhat similar to distinguishing between class characteristics in a firearm examination.

Currently, each data file needs to be examined separately in order to determine whether an "opposite end" match has occurred. A screening option is being considered that will automatically determine whether an "opposite end" match has occurred and alert the user to this possibility. The user can then examine only those files so flagged and decide whether an incorrect match has occurred. Clearly, in this instance the examiner will have to use their contextual knowledge of the marks being compared to make this determination.

Summary and Conclusions

An objective analysis of 1000 cut copper wire samples produced using 50 sequentially manufactured pliers was carried out using an algorithm that had previously been used to successfully compare striated marks produced by screwdrivers. Early efforts using the algorithm produced inconclusive results when using the same parameters used successfully for the screwdriver marks. Further experiments showed changing the comparison parameters, specifically the sizes of the Search and Validation windows, could produce successful identification of known match/non-match comparisons. Future improvements to the algorithm are planned to screen the identified matched search windows in order to eliminate the possibility of clearly incorrect "opposite end" matches.

Acknowledgments

The authors are extremely grateful to Adam Froeschl of Wilde Tool Co., Inc. for making our unusual request for sequentially manufactured slip-joint pliers possible. This study was supported by the U.S. Department of Justice, National Institute of Justice, through the Midwest Forensics Research Center at Ames Laboratory, under Interagency Agreement number 2009-DNR-119. The Ames Laboratory is operated under contract No. W-7405-Eng-82 by Iowa State University with the U.S. Department of Energy.

References

[1] Faden, D., Kidd, J., Craft, J., Chumbley, L.S., Morris, M., Genalo, L., Kreiser, J., and Davis, S., "Statistical Confirmation of Empirical Observations Concerning Toolmark Striae," AFTE Journal, Vol. 39, No. 3, 2007, pp. 205-214.

[2] Chumbley, S., Morris, M., Kreiser, J., Fisher, C., Craft, J., Genalo, L., Davis, S., Faden, D., and Kidd, J., "Validation of Tool Mark Comparisons Obtained Using a Quantitative, Comparative, Statistical Algorithm," Journal of Forensic Sciences, Vol. 55, No. 4, 2010, pp. 953-961.

[3] Petraco, N. et al. "Application of Machine Learning to Toolmarks: Statistically Based Methods for Impression Pattern Comparisons," Document 239048, U.S. Department of Justice, July, 2012.

[4] Bachrach, B., Jain, A., Jung, S., and Koons, R., "A Statistical Validation of the Individuality and Repeatability of Striated Tool Marks: Screwdrivers and Tongue and Groove Pliers," Journal of Forensic Sciences, Vol. 55, No. 2, 2010, pp. 348-357.

[5] Cassidy, F. H., "Examination of Toolmarks from Sequentially Manufactured Tongue-and-Groove Pliers," Journal of Forensic Sciences, Vol. 25, No. 4, 1980, pp. 796-809.

[6] Miller, J., "An Introduction to the Forensic Examination of Toolmarks," AFTE Journal, Vol. 33, No. 3, 2001, pp. 233-248.

[7] Monturo, C., "The Effect of the Machining Process as it Relates to Toolmarks on Surfaces," AFTE Journal, Vol. 42, No. 3, 2010, pp. 264-266.