# Optimization of a Statistical Algorithm for Objective Comparison of Toolmarks

Ryan Spotts
*Ames Laboratory*

L. Scott Chumbley
*Iowa State University and Ames Laboratory*, chumbley@iastate.edu

Laura Ekstrand
*Ames Laboratory*

Song Zhang
*Ames Laboratory*

James Kreiser
*Illinois State Police*

# Optimization of a Statistical Algorithm for Objective Comparison of Toolmarks


## Abstract

Due to historical legal challenges, there is a driving force for the development of objective methods of forensic toolmark identification. This study utilizes an algorithm to separate matching and nonmatching shear cut toolmarks created using fifty sequentially manufactured pliers. Unlike previously analyzed striated screwdriver marks, shear cut marks contain discontinuous groups of striations, posing a more difficult test of algorithm applicability. The algorithm compares correlation between optical 3D toolmark topography data, producing a Wilcoxon rank sum test statistic. Relative magnitude of this metric separates the matching and nonmatching toolmarks. Results show a high degree of statistical separation between matching and nonmatching distributions. Further separation is achieved with optimized input parameters and implementation of a “leash” preventing a previous source of outliers—however complete statistical separation was not achieved. This paper represents further development of objective methods of toolmark identification and further validation of the assumption that toolmarks are identifiably unique.



## Keywords

forensic science, toolmark, algorithm, statistical comparison, pliers, striae, quasi-striated, shear cutter

## Disciplines

Forensic Science and Technology | Materials Science and Engineering


## Comments


This is the peer-reviewed version of the following article: Spotts, Ryan, L. Scott Chumbley, Laura Ekstrand, Song Zhang, and James Kreiser. "Optimization of a statistical algorithm for objective comparison of toolmarks." *Journal of Forensic Sciences* 60, no. 2 (2015): 303-314, which has been published in final form at DOI: 10.1111/1556-4029.12642 This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving. Posted with permission.

Ryan Spotts[1], L.Scott Chumbley[1], Jim Kreiser[2], Max Morris[1], Laura Ekstrand[1], Song Zhang[1]

Optimization of a Statistical Algorithm for Objective Comparison of Toolmarks

Iowa State University, Ames Laboratory

Ames, Iowa

[1]Ames Laboratory, Iowa State University, 2220 Hoover, Ames IA 50011

[2]Illinois State Police, Retired, 3112 Sequoia Dr., Springfield, IL 62712

**ABSTRACT:** A statistical algorithm used in comparing 3D optical data from screwdriver marks has been previously demonstrated to correctly identify known matches and non-matches to a reasonable degree. This study builds upon work comparing shear cut toolmarks using 50 sequentially manufactured slip-joint pliers. Unlike screwdriver marks, the pattern produced in this instance changes as one moves from the region of initial shear cut to final separation of the wire. Thus, the studied toolmarks are described as quasi-striated, and constitute a more difficult test of the applicability of the statistical algorithm. The algorithm's performance was comparable to results seen for completely striated marks and found to be a function of two user-defined parameters. An effect related to incorrect matching of opposite ends of two sets of data previously identified as a source of outlier data points was solved. Results from this analysis and solution to the "Opposite End" problem are presented.

**KEY WORDS:** forensic science, statistical comparison, algorithm, toolmark, pliers, striae, shear cutter

In recent history, the legitimacy of scientific testimony has been questioned in several court cases – specifically *Daubert v. Merrell Dow Pharmaeuticals, Inc.* This challenge has had profound implications in the field of firearm and toolmark examination and resulted in many studies conducted to validate the practice of comparative forensic examination. The primary validation needed is the basic assumption of toolmark examination – every tool has its own unique surface that will leave a unique mark. Recent tools examined by researchers include various types of pliers (4-5), screwdrivers (1-4), and chisels (3).

Screwdriver marks are among the most studied due to their uniform and continuous striae. They have been previously characterized using stylus profilometry, confocal microscopy, and optical profilometry in various attempts designed for the identification of matching and/or nonmatching toolmarks (1-4). The results from these types of studies typically show that striae may be successfully objectively compared using a computer algorithm with relatively high accuracy. For example, in a previous study by the authors using a statistical algorithm, marks from fifty sequentially manufactured screwdriver tips were successfully separated as to differentiate between matching and nonmatching pairs to a reasonable degree of accuracy (1).

Studies of more difficult marks, which are not regularly striated, have also been conducted. Pliers are one type of tool known to create quasi-striated marks, where the toolmark changes along the length of the mark. Cassidy, one of the first to study sequentially manufactured pliers (5), found the toolmarks produced to be unique because the broaching process used to manufacture the plier teeth was performed in a direction perpendicular to the striae it would create. His study established the uniqueness of marks created by pliers, however, the analysis was not based on use of a computer algorithm. More recently Bachrach et al. studied tongue-and-groove pliers using a statistical algorithm for the objective comparison of marks created in brass, galvanized steel, and lead (4). This study found that marks could be compared when made on different media but with less accuracy than marks made on the same media.

Petraco et al. has studied impression marks, i.e. marks made through indentation (3). Results from the comparisons of chisel marks used to make impression marks proved inconclusive. Matching the discontinuous groups of striations was beyond the capabilities of the analysis software at that time.

The results discussed in this paper investigate the applicability of the algorithm employed in (1) to quasi-striated marks created by slip-joint pliers. Slip-joint pliers were chosen because this type of plier had not been studied previously. Initial results from the first investigation were submitted to the *Association of Firearm and Toolmark Examiners Journal* by Taylor Grieve, Dr. Scott Chumbley, Jim Kreiser, Dr. Max Morris, Laura Ekstrand, and Dr. Song Zhang titled “Objective Comparison of Marks from Slip-Joint Pliers” – henceforth referred to as (Grieve). Initial results having shown promise, this paper presents results on marks made by 50 sequentially manufactured slip-joint pliers. While initial results revealed the algorithm could correctly identify a large majority of matching and nonmatching pairs, some algorithm parameter values and options that work well for clearly striated marks are not optimal in the present setting. Two distinct deficiencies hindering algorithm operation were noted. The goal of this study was to investigate optimization of the parameters best suited for analysis of quasi-striated marks.

The first deficiency addressed involved parameters that affect the degree of statistical separation in the results. While separation was seen using the parameters employed for fully striated markings, better results could be obtained by changing the operational parameters of the algorithm. The second deficiency noted was concerning what the authors have termed the “Opposite End Problem”. This problem manifests itself when, in a small number of cases, the algorithm declares a “match” from two data sets which are known to be non-matching. Observation of the raw data files shows that the opposite ends of the two sets of toolmarks being compared are identified as the matching region. Such a match is physically impossible and results due to the inability of the algorithm to successfully complete the validation procedure when confronted with similar topography at opposite ends of the data sets. This possibility was first noted during research on regularly striated screwdriver toolmarks.

The current study involves complete analysis of fifty pliers, using new parameter values and an option that accounts for the "Opposite End" problem. The results of the study, including a description of the statistical algorithm, are discussed below.

**Experimental Methodology**

Fifty sequentially manufactured slip-joint pliers were obtained from Wilde Tool Co., Inc. The tools were manufactured sequentially using the same dies for the entire batch so the pliers would be as identical as possible. It is common knowledge within the field of study that the manufacturing process significantly affects the toolmarks that are created (6, 7). Thus it is relevant that the process is described.

The pliers start as pieces of steel that are hot forged into "identical" half blanks. Each half blank was then cold forged once again using the same die for every piece. After forging, the first difference between half-pairs was introduced. Fifty halves received were punched to create a small hole, while the other fifty halves received a double-hole punch – allowing future users to better hold a wider dimension range of objects. The gripping teeth and shear surface were next created with a broaching process. Two broaching machines were used in the production of the pliers. Plier halves with the double-hole punch went to one machine, while pliers with a single hole punch went to the other. During this separation time the manufacturer stamped the numbers 1-50 on the plier halves so the correct sequence could be ensured. The broaching process on the shear surface created the characteristic nature that is of interest for this study.

After broaching, both halves of each plier were given the same heat treatment and shot peened to strengthen the material and increase the surface hardness. The flat side regions were next polished and the double-hole punch half was branded with the company logo. The plier half with the double-hole punch and company logo was labeled as the 'B' side for every plier pair, and the other side 'A'. An overview of pliers from unfinished to finished states is shown in Figure 1.

Wire test samples were created by using bolt cutters to cut 2” samples from wire spools. The bolt cut ends were marked using a permanent marker so they could not be confused with plier shear cut surfaces. Diameters of the wire used were 0.1620” for the copper and 0.1875” for the lead. Test marks were made by shear cutting the copper and lead wire. Shear cutters are defined by AFTE as “opposed jawed cutters whose cutting blades are offset to pass by each other in the cutting process” (8). Since the shear face was used on the pliers to make the samples, by definition the created marks are shear cutting marks. Figure 2 pictorially shows the exact location used on the pliers to create the toolmarks.  Each shear cut was made by placing the sample between the shear surfaces – with the ‘B’ surface always facing downward – marking the sides ‘A’ and ‘B’ corresponding to which plier shear surface would be acting on that section of the wire. Thus, two samples, one ‘A’ and one ‘B’, were created with each shear cut. The samples were shear cut alternating between copper and lead until ten of each sample type were created. This resulted in 2,000 total samples: 1,000 samples for both copper and lead with half of each coming from each side of the pliers. For consistency, every sample was made by retired forensic examiner Jim Kreiser.

When the wire is mechanically separated, the two surfaces of the shear edges move past each other. The resultant action is therefore a combination of both cutting the surfaces and a shearing action of the edges as they move through the material. The result is two surfaces being created on each half of the separated wire sample, comprising both shear cut and impression markings, roughly at 90˚ to each other with both being ≈ 45˚ to the long axis of the wire. Only the shear cut surface on the ‘A’ and ‘B’ sides of the sample were scanned and analyzed. A schematic showing the process is shown in Figure 3.

The scope of the current study included only the copper samples, leading to a total sample size of 1,000. To obtain the surface data from the samples, each piece was scanned using an Alicona G3 Infinite Focus Microsope. Scans were completed at 10x magnification with a two micron vertical resolution. An example shear cut surface scan after noise reduction is shown in Figure 4. Similar to the initial study by

Grieve, the data was taken from two locations. The long edge (solid line in Fig. 4) is near where the shear cut began and the short edge (dashed line) is nearer where the shear cut ended. Striae near the beginning of the shear cut are longer and more regular than striae near the end of the shear cut, so it is important to observe the results at both locations. It is clear from viewing the figure that the pliers created a quasi-striated surface – a surface consisting of groups or parallel striae that are not continuous along the length of the mark.

An example of the scanned data from the infinite focus microscope prior to and post noise reduction is shown in Figure 5. The cut surface is embedded in irregular spiky noise which arises from the sample's edge and the scanner background. This spiky noise must be removed so that it does not interfere with the statistical analysis. To remove this noise, the authors used a combination of automated cleaning algorithms and manual cleaning. The automated cleaning algorithms are described in more detail in (9), but a brief description of these follows. First, the 2D image texture and quality map from the scanner were used to remove those points that were too dark or had a poor quality value. These points cause spikes in the scanner output. Next, a seventh-order polynomial fitting was generated for each row of the data. For each point, the discrepancy between the measured and predicted depth was computed, and points with a discrepancy of 100 microns or greater were discarded. This process was repeated for the columns. Finally, small holes (less than 20 pixels in diameter) were filled through linear interpolation. To remove any remaining spikes and the sides of the cut wire, the authors used a visual painting program to paint over noisy regions. The computer algorithm then interpreted the painted areas as points to exclude from the analysis.

Since it is impossible to scan every sample at precisely the same angle relative to the equipment, it is necessary to correct for this sample angle using a process called detrending. To detrend the data, linear least squares was used to fit a plane to the data. To make this process faster and less sensitive to noise, only 80 points were used in the plane fitting. These points were selected in an "X" pattern that evenly covered the majority of the sample surface. Once the plane fitting was obtained, the plane was subtracted from the surface data to remove the global surface angle.

When employed in the initial study by Grieve, the comparative algorithm used in (2) was discovered to have a flaw that prevented it from operating effectively.  For a more complete discussion of the algorithm the reader is referred to (2).   Briefly, the algorithm works in two major steps, the optimization and validation steps. An iterative "search" window of user-determined size (in pixels) is held stationary on Trace 1 while the correlation to a same size window is calculated over the entirety of Trace 2. The window is then shifted one pixel over on Trace 1 and the process is repeated. This is done until the two regions of best correlation are found. This process is shown schematically in Figure 6.

Once the region of highest correlation is found during the optimization step, two shifts are applied and compared – random shift and rigid shift. This is the validation step. The size of the "validation" windows that are shifted are user-determined. The rigid shift moves from the best correlation window on each trace a set distance and calculates the correlation. The algorithm also moves from the best correlation window a random distance and calculates the correlation. An example rigid and random shift are shown in Figure 7. The number of rigid and random shifts employed is also user defined; for the purposes of this study the number was set at 50.  Comparison of the rigid shift values to the random shift values allows the likelihood of a true match being found to be assessed.

In the initial investigation by Grieve, outlier data points were observed to stem from the algorithm misidentifying the opposite ends of marks as a positive match. One example of this is shown in Figure 8. The solid line orthogonal to the shear cutting direction is the location the profiles were compared along and the dashed lines represent coordinate axes.

Through random chance opposite ends of a mark occasionally have the regions of highest correlation between marks. Clearly, given that the shape of the shear cut wire specifies a definitive left and right side of the shear cut, it is physically impossible for this to occur. In investigating the cause for the false match it was discovered that in cases where the region of highest correlation between two marks occurs at the end of the scan profile, the validation routine used by the algorithm to ascertain the quality of the comparison cannot function properly. When a “match” is found near the end of a scan profile the space needed to successfully accomplish the rigid shifts and complete the validation step does not exist. This results in an incorrect validation, and a “match” being declared when in fact a non-match may exist.

To address this problem, the original algorithm (2) was modified so that during the optimization step a “leash” is applied to the search window. The leash limits the comparison distance between profiles; the comparative correlation is no longer calculated over the entirety of Trace 2 for each iteration of the search window but only to a certain percentage of the entire distance. Figure 9 shows schematically an example of how the leash limits the search range for the region of highest correlation. Leashing the search window makes it impossible for the algorithm to identify regions far from each other on the real surface as matching. The current version of the leash is set as a percentage of the total length of the trace. The leash was set at 80% for this analysis. Figure 10 shows the same plier comparison as Figure 8 but after the leash was implemented. The algorithm now calculates a low T1 value for this particular comparison, consistent with the expected result for nonmatching test marks.

A Wilcoxon Rank Sum test statistic (centered and scaled to have a nominal standard deviation 1, and labeled T1 here) is calculated during the validation step and is what is returned by the algorithm. The T1 statistic is determined by comparing the results of rigid and random shifts. Matching marks should have relatively high correlation after a rigid shift if they are truly similar, and lower correlations during random shifts. The magnitude of the T1 statistic is affected by how much the rigid and random shifts differ. High rigid shift correlation and low random shift correlation would result in a high T1 value – indicating a matching pair – while the opposite scenario would result in a low or negative T1 value indicating a nonmatching pair. The reason many shifts are applied is because random chance may allow a few random shift windows to have a high correlation. As more shifts are applied to a matching pair, the expected trend would be an average rigid shift correlation that would increase and an average random shift correlation that approaches zero. As more shifts are applied and those correlations separate, the T1 statistic will increase. As more shifts are applied to a nonmatching pair, the expected trend would be average rigid and random shift correlations that become closer in value – resulting in  a T1 statistic value near zero.

**Results**

The data from the fifty sequentially manufactured pliers were compared using three different types of comparisons resulting in three sets of data. All three comparisons types were performed using data from both the long and short edges as defined in Fig. 3.

Set 1: Comparing known matching pairs. Data for Set 1 were created by comparing marks made by the same side of the same pliers. Comparisons were made between marks 2 and 4, as well as marks 6 and 8 for both sides of each plier. An example of the methodology for comparisons in Set 1 are best described in a tabular format; an example of the comparison order through two pliers is shown in Table 1.

Set 2: Comparing known nonmatching pairs. Data for Set 2 were created by comparing marks made by different sides of the same plier. This set could confirm that both sides leave a unique mark. Comparisons were made between sides 'A' and 'B' for marks 10, 12 and 14. An example of the methodology for comparisons in Set 2 is shown in Table 2.

Set 3: Comparing known nonmatching pairs. Data for Set 3 were created by comparing marks from the same side of different pliers. Marks 16, 18 and 20 were compared between different pliers for both sides. An example of the methodology for comparisons in Set 3 is shown in Table 3.

Search and validation window sizes of 200 and 100 pixels, respectively, were used as part of the initial analysis. These window sizes had been previously used for successful matching of screwdriver toolmarks (2). The results for all three data sets are shown in Figure 11 presented as box and whisker plots. The solid black line represents the median value of the comparisons. The upper and lower bounds of the box represent quartiles, and the whiskers are within one and a half times the difference of the quartiles. Outliers are denoted by circular dots. A T1 statistic close to zero indicates little or no correlation between the data sets (i.e., a nonmatching pair) while a larger positive value would indicate a correlation exists between the two data sets being compared. Set 1 involves comparisons between known matching pairs while Set 2 and 3 were known nonmatching comparisons.

Observation of Figure 11 shows that the modified algorithm performs reasonably well for the quasi-striated plier marks. Both the long and short edge comparisons show significantly higher T1 values for the known matches of Set 1 than for the known non-matches of Sets 2 and 3.

Data from Grieve's initial investigation that used the original algorithm are shown in Figure 11 for comparison. Note that the leash included as a fix to the Opposite End problem has resulted in a substantial improvement of the modified algorithm over the original version. The median value for Set 1 has been increased from a T1 value of approximately 1.5 – 2 to a T1 value of 4, and even more importantly there is now complete statistical separation of the quartiles for the long edge data. Data Sets 2 and 3 are still centered approximately on zero, and there also is a net reduction in the total number of outliers.

Since the quasi-striated marks produced by the plier shear cuts are far less regular than the previous screwdriver marks studied, experiments were conducted to determine the effect window size (i.e. search and validation) may have on the results. The 2:1 window size ratio was maintained for this second round of analysis with the search windows set to 1000, 500, 200 and 100 pixels with corresponding validation window sizes of 500, 250, 100 and 50 pixels.

The results of these experiments for the short and long edges are shown in Figures 12, 13 and 14. Figure 12 shows that except for the smallest window size (100-50), Set 1 always has a median value well above zero, with the median increasing as window size increases from approximately 4.0 (200-100) to 7.5 (1000-500).  In contrast, the median values for Sets 2 and 3 always hover near zero as expected for nonmatching pairs, regardless of the window size. An apparent increase in data spread and number of outliers is also observed with increasing window sizes as do the number of outliers. Results from both long and short edges were similar to each other; the short edge had fewer outliers but an increased amount of data spread.

In some cases during the analysis the algorithm would not return a result for every comparison. This is because the algorithm does not allow validation windows to overlap. Thus, as larger and larger window sizes are used it becomes more likely that the algorithm will run out of profile length, especially on short edges with large window sizes, and not return a T1 value. For a 2:1 ratio the algorithm did not return 6 values for the Set 1 short edge, 9 values for the Set 2 long edge, 13 values for the Set 2 short edge, and 19 values for the Set 3 short edge. These numbers should be compared to the total of 3,965 data comparisons that were performed for the 2:1 ratio analysis.

With a clear trend in the results due to window size, the effect of size ratio was next analyzed using both 4:1 and 6:1 search to validation window size ratios. Search windows were set to 800, 600, 400 and 200 pixels with corresponding validation window sizes of 200, 150, 100 and 50 pixels used for the 4:1 ratio experiment. Search windows were set to 750, 600, 450 and 300 pixels with corresponding validation window sizes of 125, 150, 75 and 50 pixels for the 6:1 ratio experiment. The results from these analyses are shown in Figures 15 - 20 for the short and long edges.

Observation of Figures 15 and 16 show that the clear trend of increasing T1 value with increasing window size observed for the 2:1 ratio holds true for both the 4:1 and 6:1 ratios for Set 1 comparisons. An increasing number of outliers were also observed with increasing window size. However, slight decreases in data spread, median value and number of outliers for Set 1 comparisons was also observed relative to the 2:1 ratio. The median values for the Set 1 data ratios were still significantly above a zero value, with medians approaching a T1 value of 6.

Set 2 and 3 comparisons, known non-matches, are shown in Figures 17-20. Observations for Set 2 and 3 comparisons showed a median value near zero regardless of window size, increasing data spread with increasing window sizes and no clear trend in the number of outliers. In general long edge comparisons had better results evidenced by the general decrease in data spread. The algorithm did not fail to return any results for the 4:1 and 6:1 ratios. This analysis contained 4,012 comparisons for each ratio.

**Discussion**

When the research transitioned from regularly striated to quasi-striated marks, it became apparent that a parameter optimization is necessary for different tools. This optimization lead to significantly improved results. The algorithm employed in this research was optimized to provide better results for the current set of toolmarks by experimentally changing window sizes and utilizing an option that limits errors due to the

Opposite End problem. While the leash restriction is effective, it should be realized that its effectiveness is only made possible by the introduction of contextual knowledge into the analysis. For the plier marks, the non-symmetric shape of the shear cut makes it easy to determine in which direction the scans should be analyzed. A more symmetric plier mark might be more difficult to orient properly to make use of the leash, involving a trained examiner to ensure the data was obtained correctly.

In the most ideal scenario there would complete data separation between known matching and nonmatching pairs, giving a clear indication of correlation, with no outliers in the data. Although ideal degree of separation has not been achieved, there is clearly a large majority of correctly identified toolmarks. Close examination of the outlying data points from both edges reveals that for these specific comparisons the algorithm produces a correct result for the vast majority of window combinations used. For example, consider Table 4 where several individual outlying points are shown. Incorrect matches, bolded and italicized, were found that were inconsistent with the algorithm results for other window sizes. Of the 12 different windows sizes employed the algorithm returns a “correct” answer in 10 or more instances. Note also that the majority of these outlier points stem from larger window sizes.

Observation of Table 4 suggests another improvement that might be made to the algorithm operation. In the current form the algorithm only uses one combination of search and validation windows at a time. This means that the current algorithm has a “one size fits all” mentality even though every mark is unique. It may be possible to further improve the algorithm performance by examining multiple search and validation window combinations simultaneously during comparison to find a better fit. If the underlying hypothesis behind this application of the T1 statistic is that matching pairs will have more correlation that nonmatching pairs, it should also hold true if one uses more search and validation window combinations. However, this will require substantial further development of the statistical arguments on which the algorithm is based, since correlations based on different algorithm parameter values will have different distributions, both in matching and non-matching cases.

**{The remainder of the discussion may be removed by the editor if it is deemed too speculative for the Journal. The purpose of the remaining discussion was meant to show that there may be more avenues available to further increase the robustness of the algorithm. As it stands now, if an examiner were to use this algorithm in an examination it would require a decent amount calibration for each toolmark examined to produce a significant result.}**

Exploratory analysis was done to see the effect using multiple search and validation window combinations simultaneously could have on separating known matches from known non-matches. The data from the approximately 6,000 possible long edge comparisons were used for this exploratory analysis. The average and maximum T1 values from all 12 search and validation window combinations were determined for each comparison. The results were organized by data set and are shown in Figure 21.

The average result returned by the algorithm when all window sizes are considered remains significantly above a zero T1 value for matching pairs and remains near zero for nonmatching. The maximum returned values have every T1 result above zero for matching pairs, with only a few outliers overlapping the median maximum T1 value of approximately 2 for nonmatching pairs. Clearly these results give strong evidence that the quasi-striated marks from the 50 sequentially manufactured pliers are unique and can be separated objectively.

The results from this exploratory analysis indicate that the algorithm may be improved if a statistically sound method of examining multiple search and validation window combinations simultaneously can be found. Another avenue worth consideration is simultaneously comparing more than one profile between marks at a time. As the algorithm currently works, a one pixel wide profile from each toolmark is compared at a time – one pixel is 0.8 μm in width. Having the algorithm compare within plus or minus 10 profile widths of the chosen profile may also end with improved results. It is likely that comparing more than 0.8 μm of width at a time would improve the results produced by the algorithm. Future work will examine the proposed methods to further improve the algorithm.

## Summary and Conclusions

This study was completed using 1,000 samples of copper wire shear cut into two pieces using 50 sequentially manufactured pliers. The resultant toolmark on the shear cut surfaces was quasi-striated in nature, consisting of groups of striations. Pairs of shear cut surfaces were objectively compared utilizing a statistical algorithm that had been previously successful comparing regularly striated marks. The algorithm was optimized and applied using a leash option to the search and validation windows in the algorithm to prevent incorrect identification related to matching at opposite ends of the comparison pairs from occurring. This resulted in a noticeable improvement in the analysis. Known matching pairs had large T1 values for the majority of comparisons (indicating a match) and known nonmatching pairs had near zero values for the majority of comparisons (indicating a none match). A high degree of separation in the data was observed although complete statistical separation was still not achieved. While the results have improved, more work is needed before the algorithm can be used as a tool by forensics examiners to increase the robustness of the identification process. Future improvements to the analysis method may involve utilizing combinations of search and validation windows simultaneously or changing the process to view more than one profile of data at a time.

## Acknowledgments

The research group thanks Adam Froeschl and Wilde Tool Co., Inc. for their aid in gathering sequentially manufactured slip-joint pliers. Their effort helped make this research possible. The research group also thanks the U.S. Department of Justice, National Institute of Justice, and the Midwest Forensics Research Center for their support. This support was provided through Interagency Agreement 2009-DNR-119 and through contract W-7405-Eng-82 for the Midwest Forensics Research Center and Ames Laboratory respectfully. Ames Laboratory is funded by Iowa State University under contract number W-7405-Eng-82 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 Toolmark 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 Toolmarks: 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.
8. AFTE. "Shear Cutters." *Glossary of the Association of Firearm and Toolmark Examiners*. 5th ed. Montreal: Forensic Technology, Incorporated, 2007. 182. CD.
9. Song Zhang, Laura Ekstrand, Taylor Grieve, et al., "3D data processing with advanced computer graphics tools", Proceedings of SPIE Vol. 8493, 849310 (2012) SPIE Digital Library