2014

Objective forensic analysis of striated, quasi-striated and impressed toolmarks

Ryan Edward Spotts
Iowa State University

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

Part of the Materials Science and Engineering Commons, and the Mechanics of Materials Commons

Recommended Citation
Spotts, Ryan Edward, "Objective forensic analysis of striated, quasi-striated and impressed toolmarks" (2014). Graduate Theses and Dissertations. 14299.
http://lib.dr.iastate.edu/etd/14299
Objective forensic analysis of striated, quasi-striated and impressed toolmarks

by

Ryan E. Spotts

A thesis submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Materials Science and Engineering

Program of Study Committee:
Scott Chumbley, Major Professor
Song Zhang
Larry Genalo

Iowa State University
Ames, Iowa
2014

Copyright © Ryan E. Spotts, 2014. All rights reserved.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS ......................................................................................... iv

ABSTRACT .............................................................................................................. v

CHAPTER 1  THE SHIFT TO OBJECTIVITY ....................................................... 1
   Toolmark Identification: A Brief History .......................................................... 1
   Identifying a Match: Current Methodologies ................................................. 5
   Toolmark Identification: The Need for Objectivity ....................................... 8
   Progression to Objective Analysis of Toolmarks ........................................... 9
   Foundation of Current Research ................................................................... 12
   References ........................................................................................................ 16

CHAPTER 2  OPTIMIZATION OF A STATISTICAL ALGORITHM FOR OBJECTIVE COMPARISON OF TOOLMARKS ................................................................. 19
   Introduction ..................................................................................................... 19
   Experimental Methodology ........................................................................... 22
   Results ............................................................................................................. 33
   Discussion ....................................................................................................... 42
   Summary and Conclusions .......................................................................... 46
   Acknowledgments ......................................................................................... 46
   References .................................................................................................... 47

CHAPTER 3  ANGULAR DETERMINATION OF TOOLMARKS USING A COMPUTER GENERATED VIRTUAL TOOL ................................................................. 49
   Introduction .................................................................................................... 49
   Experimental Methodology ........................................................................... 52
      Sample generation ...................................................................................... 52
      Surface characterization ........................................................................... 53
      Noise reduction ........................................................................................ 54
      Data comparison ....................................................................................... 55
   Results ............................................................................................................ 60
   Discussion ....................................................................................................... 62
   Summary and Conclusions .......................................................................... 66
   Acknowledgments ......................................................................................... 67
   References .................................................................................................... 67
<table>
<thead>
<tr>
<th>CHAPTER 4</th>
<th>OBJECTIVE ANALYSIS OF IMPRESSED CHISEL TOOLMARKS</th>
<th>70</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Introduction</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>Experimental Methodology</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>Results</td>
<td>77</td>
</tr>
<tr>
<td></td>
<td>Discussion</td>
<td>79</td>
</tr>
<tr>
<td></td>
<td>Summary and Conclusions</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>Acknowledgments</td>
<td>81</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>82</td>
</tr>
</tbody>
</table>

| CHAPTER 5 | SUMMARY AND CONCLUSIONS                          | 84 |
ACKNOWLEDGEMENTS

I would like to thank Dr. Scott Chumbley, Dr. Song Zhang, Dr. Max Morris, Dr. Larry Genalo, Jim Kreiser and Laura Ekstrand for helping me towards the completion of my thesis. All of you played an integral part in this research either through technical guidance or the building of the tools necessary to perform the analyses. Special thanks are also needed for Adam Froeschl of the Wilde Tool Company Inc. for aiding the procurement of sequentially manufactured tools paramount for the conducted research.

I especially thank my parents (Steve and Kristie) and my grandparents (Ardelle and Roger Spotts and Joan and Richard Thayer) for all of their support throughout my undergraduate and graduate career.
ABSTRACT

Following the 1993 Daubert v. Merrell Dow Pharmaceuticals, Inc. court case and continuing to the 2010 National Academy of Sciences report, comparative forensic toolmark examination has received many challenges to its admissibility in court cases and its scientific foundations. Many of these challenges deal with the subjective nature in determining whether toolmarks are identifiable. This questioning of current identification methods has created a demand for objective methods of identification – “objective” implying known error rates and statistically reliability. The demand for objective methods has resulted in research that created a statistical algorithm capable of comparing toolmarks to determine their statistical similarity, and thus the ability to separate matching and nonmatching toolmarks. This was expanded to the creation of virtual toolmarking (characterization of a tool to predict the toolmark it will create).

The statistical algorithm, originally designed for two-dimensional striated toolmarks, had been successfully applied to striated screwdriver and quasi-striated plier toolmarks. Following this success, a blind study was conducted to validate the virtual toolmarking capability using striated screwdriver marks created at various angles of incidence. Work was also performed to optimize the statistical algorithm by implementing means to ensure the algorithm operations were constrained to logical comparison regions (e.g. the opposite ends of two toolmarks do not need to be compared because they do not coincide with each other). This work was performed on quasi-striated shear cut marks made with pliers – a previously tested, more difficult application of the statistical algorithm that could demonstrate the difference in results due to optimization. The final research conducted was performed with
pseudostriated impression toolmarks made with chisels. Impression marks, which are more complex than striated marks, were analyzed using the algorithm to separate matching and nonmatching toolmarks. Results of the conducted research are presented as well as evidence of the primary assumption of forensic toolmark examination; all tools can create identifiably unique toolmarks.
CHAPTER 1. TOOLMARK EXAMINATION: THE SHIFT TO OBJECTIVITY

Toolmark Identification: A Brief History

What is a toolmark? A toolmark can be loosely defined as the resulting deformation remaining after two objects have made physical contact. Toolmarks can occur with any objects including interfacing components within firearms and with objects traditionally thought of as tools such as pliers, hammers and screwdrivers. The resulting deformation, primarily localized to the softer object, can occur in two forms: striated and impressed. Impressed toolmarks occur when two objects are perpendicularly forced together with no lateral movement, while striated toolmarks are more dynamic in nature and require parallel movement (with respect to the substrate) during mark creation (1). The underlying assumption of forensic toolmark examination is that each tool contains a unique surface topography such that its resulting toolmarks are identifiably unique. The reasoning for this assumption is the many cumulative factors that contribute to the variation of tools. Some of these factors include: metallurgy (grain distribution, relative hardness etc.), manufacturing instrument precision, wear of tooling, random particulate size distribution in abrasive operations, functional use of the tool and more (2). The combination of these factors results in an extremely low probability that two tools could create the exact same toolmark. However, while the probability of two tools truly creating the same toolmark are extremely low, the discerning power of the human eye or apparatus built by humans are the limiting factors of toolmark identification.

The first noted case where toolmark analysis linked a person to a crime occurred in London in 1835 (3). An observant policeman working on a homicide case (servant suspected
of murdering homeowner) observed a mold mark on the lead projectile used to murder the victim. This mark was unique to the mold used to create the projectile. Further, the paper wadding used during the shooting was found to be similar to newspaper in the servant’s room. The observations made by the policeman were crucial in determining the guilty party.

It must be noted that the policeman in the first case identified what would eventually be defined as subclass characteristics. These subclass features are defined as features that are created during the manufacture of a tool and unique to the process (4). The lead projectile mold had an irregularity such that it imparted this irregularity to all of the projectiles created by it – thus allowing the determination of where the projectile was manufactured. Asides from subclass characteristics, there are also class and individual characteristics. Class features are features common to all tools of a particular group. For example, hammer face impressions from hammer faces with diameters of three centimeters will always leave impressions that are approximately three centimeters in diameter. Individual characteristics are features imparted to an individual tool through functional use or during manufacture that are unique to the tool; one example would be the individual striations caused by an abrasive sharpening process (4). The individual characteristics, the morphological patterns of the minutiae of surface peaks and valleys, are what examiners use to make most identifications.

Early forensic toolmark investigations were mostly limited to class characteristics. One example of this was with the famous death of Thomas “Stonewall” Jackson who died as a result of a gunshot wound in 1863 during the American Civil War. The investigation determined that Jackson was killed by one of his own troops because the caliber and shape of the projectile matched Confederate forces and was not consistent with the caliber used by Union forces (3).
Around the year 1900 in Germany the first noted occasion of striation matching occurred when Professor Krockel published two papers on his method of toolmark identification (5). The matching of individual features such as striations is critical to being able to determine correct identifications.

The next U.S. advancement in forensic toolmark examination came in 1900 and 1902 when the first documented occasions of what would eventually come to be known as individual characteristics were mentioned. An article written by Dr. Albert Hall discussed measurement of marks on bullets caused by rifling. Although his work was not continued, the understanding of the importance of the fine scratches that occurred on bullets was significant. Two years later a US court allowed expert testimony containing photographs of both evidence and test fired bullets (3).

In 1912, Professor Balthazard in France created a procedure in which he carefully took photographs of both evidence bullets and bullets fired from a suspect firearm. Markings (individual characteristics) between the bullets were compared for identification. The same photographic procedure was later applied to cartridge casings: firing pin, breech face, ejector and extractor marks (3). This work represented a significant leap forward for comparative forensic science due to the methodic comparison of individual characteristics.

Comparative forensic examination received its next big jump forward in 1918 when the Arizona State Supreme Court (Paul v. Hadley) determined evidence from ballistic testing was admissible and legitimate. An attorney named A. J. Eddy had conducted experiments firing projectiles through both a suspected murder weapon and several same-caliber firearms – all evidence was photographically documented including the fatal projectiles. Eddy’s
testimony of his testing convinced the jury that each firearm left discernable features on fired projectiles and that the fatal projectiles could be linked to the murder weapon (3).

The comparison microscope, a primary tool still used by examiners today, was adopted for use in firearm identification by Philip Gravelle of the newly established U.S. Bureau of Forensic Ballistics in 1925 (3). The comparison microscope significantly aided forensic examinations by allowing higher discernibility between striations.

The year 1930 saw the publication of an article by Luke May titled “The Identification of Knives, Tools and Instruments” (5). Mr. May used controlled lighting and comparisons of photomicrographs of individual features to declare positive toolmark identifications.

In 1942, Burd and Kirk published an important paper where various toolmark creation parameters were investigated (6). One important finding in their analysis is that toolmarks created by the same tool, at different angles, may appear to be nonmatching toolmarks. This finding demonstrates the significance of making accurate exemplar toolmarks for comparison to crime scene evidence.

The next couple decades saw the establishment of many crime labs, standardized forensic investigation methodologies and the initial meetings of what would become the Association of Firearm and Toolmark Examiners (AFTE) as the field of firearm and toolmark analysis began to flourish. Forensic examiners became crucial in many high level investigations such as the John F. Kennedy, and Martin Luther King Jr. assassinations (3).

The establishment of AFTE in 1969 and its annual training seminars provided a great deal of standardization for the development of firearm and toolmark examiners. This culminated with the publication of the first AFTE Glossary, a book containing standardized
nomenclature, pertinent formulas for mathematical calculations (bullet trajectory, energy, spin etc.) and relevant chemical formulas (3). This organization of people and materials ensured further consistency between investigations and has increased the uniformity in how evidence is analyzed (and what constitutes positive and negative identifications).

This brief introduction to the history of firearm and toolmark examination explains many of the stepping stones in the early history of toolmark examination (which dealt mostly with firearms). However, the methodology of how toolmark identifications must now be described.

Identifying a Match: Current Methodologies

It must be stated that forensic examiners are trained professionals. They must pass training, examinations and obtain certifications before they are allowed to perform real work. Institutions such as the Association of Firearms and Tool Mark Examiners (AFTE) ensure consistent casework and proficiency by all personnel by standardizing certain aspects of training.

All forensic investigations begin once a crime has taken place. Before evidence is submitted to a laboratory, a crime scene investigator must obtain all relevant information. For example, background information like a toolmark’s location on a window sill could indicate a lever-action was applied to a tool – useful information when exemplar toolmarks are made.

Suspect tools and toolmarks from the crime scene are then submitted to an examiner after all of the crime’s details have been collected (the detail collection can also be an ongoing process). The examiner then begins a process of interpretation and evaluation (4). Often times the evidence is examined using a comparison microscope, equipment capable of
allowing an observer to view two objects independently side-by-side under magnification. Often times a suspect tool will first be compared to a toolmark. Once those observations are recorded, the suspect tool may be used to create exemplar toolmarks for comparison to the evidence toolmark. The examiner takes the utmost precautions to preserve the evidence during this process – often times lead is the substrate used for exemplar marks due to its low hardness and high ductility. An examiner must consider a variety of factors during their portion of an investigation. These factors include: quality of available toolmark, amount of detail present (are enough individual characteristics present?), do the observed features in the toolmark correspond to features on the tool, were the individual features on the tool produced in a manner that allows uniqueness (grinding operation) or a manner where individual features could be difficult to discern (turning). Other important details are the substrate containing the relevant deformation (a soft substrate may have preserved more of the acting tool’s detail than a hard substrate which may have caused some tool deformation) and how the toolmark was created – e.g. at what angle/s and under what forces did the tool cause the deformation.

For an examiner to determine a tool matches a toolmark, several things must be in agreement. The toolmark should exhibit class characteristics consistent with suspect tool. Under microscopic examination, striae or impression evidence should form a consistent pattern match – meaning striae/impressions of approximately the same height/depth and width should occur at approximately the same location relative to other corresponding features. An example of this is shown in Figure 1. AFTE has formed a (subjective) theory of identification that is used as a rule by most examiners. For toolmarks to be matching they should show sufficient agreement; sufficient agreement exists when the level of agreement
between two toolmarks is greater than the best known agreement between toolmarks made by similar but nonmatching tools, and consistent with the level of agreement observed between toolmarks made by the same tool (7).

![Figure 1: Matching striae example](image)

**Figure 1: Matching striae example (8).**

When an examiner finishes their evaluation they render an opinion with documentation of their evidence. Their conclusion is usually limited to four options: identification (if sufficient corresponding detail is present), inconclusive (not enough evidence to make a determination), elimination (enough detail is present to determine evidence does not agree) and unsuitable (evidence is not sufficient to make any determination) (4). Typically when identification is determined (with confirmation from secondary examiners), since it cannot be truly known whether another tool created the evidence, it is said that the probability of another tool having created the evidence is so exceedingly low that it should be a negligible possibility. While brief, this explanation of the current identification methodologies shows the subjective but reasonable nature of toolmark analysis.
Toolmark Examination: The Need for Objectivity

The method of toolmark identification by finding reasonable agreement between evidence characteristics was accepted until a 1993 landmark Supreme Court case of Daubert v. Merrell Dow Pharmaceuticals. This court case, while not directly dealing with toolmarks, set the legal standard for scientific testimony admissibility. For scientific testimony to be admitted and considered valid four criteria must be met: 1) the science must be testable 2) the error rates of the process must be known or estimable 3) the science has received peer review and publication and 4) has general acceptance in the relevant scientific community (3). While this ruling was reasonable, objective and pragmatic for keeping “junk” science out of the courtroom, it profoundly affected the field of forensic toolmark examination. It is difficult to know or estimate the error rate and quantify how identification was made with the subjective nature of pattern-matching. What level of pattern agreement constitutes a positive identification? It is a difficult conclusion to express with words or numbers. It is also entirely reasonable for an unsavvy person on trial to dislike a process wherein their conviction may rely on another human essentially stating “I believe the evidence matches the suspect’s tool” without a statistical argument.

The response to the Daubert ruling resulted in an abundance of research providing evidence of underlying toolmark analysis assumptions (tool/toolmark uniqueness), more objective forms of toolmark analysis and attempts to estimate examiner error rates. For example, a summary paper of various systematic studies has shown an estimated trained examiner error rate of ~1% (9). While the examiner proficiency studies provide confidence in the current methodologies, these systematic studies cannot realistically encompass the large variety of evidence and tools an examiner may be expected to analyze and do not replace the
need for objectivity. Ultimately the Daubert standard has set the appropriate expectation of statistical accuracy to coincide with forensic evidence.

**Progression to Objective Analysis of Toolmarks**

With the Daubert ruling, extensive research has been performed to objectively analyze toolmarks. For this research to be conducted, an automated means of individual characteristic analysis needed to be created. While the Daubert ruling occurred in 1993, the first instrumental form of forensic surface measurement was created in 1958. The director of the Oakland Police Department Crime Lab, John Davis, proposed the use of a Striagraph (essentially a stylus profilometer) for forensic surface texture analysis. The Striagraph never went past the research phase but it was certainly the beginning of instrumental analysis of toolmarks.

A stylus profilometer was used in 2007 by Faden et al. to examine toolmarks created by sequentially manufactured screwdriver tips at a variety of different angles (10). A number of observations were made: almost all comparisons contained relatively high correlation values, matching toolmark comparisons contained higher levels of correlation than nonmatching comparisons and matching comparisons performed between toolmarks created at different angles revealed that the toolmarks are dissimilar enough to be consistent with nonmatching toolmarks. Faden’s objective analysis indicates a correlation statistic could be used to determine identifications but current data would indicate not to the same degree of accuracy as subjective methods. Small scratch marks visible on samples from this work also demonstrate why noncontact methods of measurements are preferred in forensics.
Another technological progression that occurred before the Daubert ruling was the creation of DRUGFIRE in 1989. This system digitized markings on bullets and cartridge casings for manual comparison within a shared database between crime labs. The stored images saved time by allowing rapid digital comparison for sorting potential matches that could later be confirmed using a comparison microscope. This system was eventually made obsolete by the Integrated Ballistics Identification System (IBIS) at the beginning of the 21st century. Similar to DRUGFIRE, IBIS captured 2D images of bullet and cartridge casing markings to include in a shared database (NIBIN). IBIS represented another significant advancement because it used a mathematical metric to compare entered data and determine potential matches between cases, allowing even more rapid connection of multiple crimes (3). Two issues with IBIS are i) that the mathematical similarity metric is unknown because the equipment and comparison routine used are made by a private manufacturer not subject to peer review and ii) the analysis is still ultimately subjective. While the system is useful for pointing examiners towards potential matches, an examiner’s subjective assessment is still required to confirm identification. 2005 and 2006 saw the introduction of the BulletTRAX-3D and BrassTRAX-3D systems. These systems are similar in functionality to IBIS, but utilize a confocal microscope for 3D digitization of markings on bullets and cartridge cases (3). Once again the system suffers from the same disadvantages as IBIS (all three made by Forensic Technology Inc.) in that it uses an unknown similarity metric and ultimately relies on examiner assessment for identification – although all three mentioned systems have been significantly useful tools while research is underway on many forms of objective assessment. It should be noted that the first automated systems dealt primarily with ballistic toolmarks because firearms are used in the majority of crime.
Bachrach et al. used a confocal microscope and a correlation similarity measure for automated and objective analysis in 2010 (11). Results from Bachrach’s work indicate that while his correlation methodology is capable of determining matching and nonmatching toolmarks, the toolmark substrate can affect the similarity between matching toolmarks and the angle of toolmark creation can affect the degree of similarity. The results reported a minimum identification error rate of 0.11% when screwdriver marks were made at the same angle and the largest error rate of 49.50% when toolmarks’ creation angle differed by 30 degrees – the highest angular difference tested (11).

Wei Chu (12, 13) et al. has investigated both congruent matching cells and consecutive matching striae as objective criteria for identification. Consecutive matching striae are based on empirical observations made by Biasotti during many comparisons between bullets (14). In 1997, Biasotti published recommended objective criteria based on these observations that nonmatching toolmarks will not contain six or more consecutive matching striae or two groups of three or more consecutive matching striae (15). Wei used this proposed conservative criteria, a cross correlation function, and empirically determined critical values to determine the amount of consecutive matching striae on bullet land marks (produced by sequentially manufactured parts). The results of his work revealed the probability of no land marks matching between two matching bullets (out of 6 land marks on each bullet) was around 1.98% (12). Wei also performed work with congruent matching cells on cartridge cases. The idea involves digitally dividing a sample into many small areas. A cross correlation function is used to find a set number of cells that match – match denoting a certain value of correlation between cells that spatially correspond. Once again empirical parameters were used, and no false positive identifications occurred (13). A confocal
microscope was also used for data collection in both investigations. Wei’s objective methods of identification have shown good results and in one instance an estimated error rate approaching that of current subjective methodologies.

Petraco et al. has applied machine learning to toolmarks on cartridge cases and screwdriver marks (16). A confocal microscope was used to capture topographical information from the toolmarks. A cross correlation function was used to align samples during comparison to the region of best agreement while principal component analysis, canonical variate analysis and support vector machine were used to analyze the profiles in various combinations. Final - unaugmented - results show the lowest estimated error rate of 1.1% for cartridge cases and 0.6% for screwdriver marks (16).

The named examples are a summary of research being performed at the current time. It is clear that significant progress is being made towards a variety of truly objective methods of toolmark identification, but it is also clear there is more research to be done.

**Foundation of Current Research**

Machine learning, cross correlation functions, and other forms of similarity metric have been used to create objective methods of toolmark identification. However current investigations by this research group have not yet been mentioned. After the analysis of striated screwdriver marks by Faden et al. it was clear that a new form of statistical measure would be required to better separate matching and nonmatching toolmarks. In 2010, a new algorithm created by Dr. Max Morris was published in (17) under Chumbley et al. This new algorithm utilized a Mann-Whitney U-statistic (henceforth referred to as T1) as a nonparametric measure of similarity. The algorithm in many aspects seeks to mimic
examiner behavior during an examination. A region of best agreement is first found between two, two-dimensional topographical profiles. After this step is completed correlations are computed in both equidistant and randomized distances from the region of best agreement. The underlying hypothesis is that truly matching marks will still correlate at equidistances and not correlate during randomized distance computations, while nonmatching marks will have similar correlation during both equidistant and randomized computations because their correlation is a random event. Results from this study validate the effectiveness of the algorithm through comparison of screwdriver marks made by 50 sequentially manufactured screwdriver tips. Toolmarks were created at 30, 60 and 85 degree angles. Different combinations of toolmarks were compared. Matching toolmarks created at the same angle were computed to have the highest T1 values (consistent with matching toolmarks and relatively high correlation) while matching toolmarks created at largely different angles, nonmatching marks from the opposite edge of the same tool tip, and completely nonmatching toolmarks tended towards negative or zero T1 values (consistent with nonmatching toolmarks and relatively low correlation). One part of the study also enlisted trained examiners to evaluate a select subset of toolmarks, and their results were compared to those obtained by the algorithm. Results from this second part of the study found that trained examiners still outperformed this form of objective assessment (17).

After the success of comparing striated toolmarks, the intended capability of the designed algorithm, the next investigation compared quasi-striated shear cut marks made by pliers. Quasi-striated marks were composed of groups of regular striae, without regularity through the entire mark. It was expected to be a more difficult test of the algorithm. This investigation also observed the effect of varying different combinations of algorithm
parameter values and ratios. Results from this study revealed that the algorithm could separate matching and nonmatching shear cut plier marks, and also found that the degree of statistical separation could be increased by using increased algorithm parameter values (18).

The final area of former research being built upon involves the creation of a virtual toolmarking capability (19). This capability was developed to objectively and directly compare tool topography to the toolmark it could create. Six same-brand screwdriver tips were used to create 34 corresponding toolmarks using both edges on each tool tip. Toolmarks were created at angles of 45, 60 and 85 degree angles relative to the substrate. The topography of each screwdriver tip was digitized using a focus variation optical profilometer. An algorithmic procedure was used to rotate the digitized tool edge topography to a relevant angle of interest and then construct a virtual toolmark – a digital approximation of the toolmark the tool topography would be expected to generate at the given angle. Using the algorithm developed in (17), toolmarks were compared to virtual toolmarks at a range of different angles. The algorithm was able to correctly identify all toolmarks except in the case of two false negative matches (no false positives were recorded). Similar to the prior work in (17), T1 values were highest (indicating relatively high correlation) when the tool and toolmark creation angle were matching (19). This significant development in the capability to predict the toolmarks created by tools opened a wide range of possibilities. Virtual toolmarking could aid examiners by allowing estimation of toolmark parameters, speeding analysis by generating exact exemplar toolmarks, serving as an aid for examiner training and more.

The research performed in the partial fulfillment of this thesis builds upon the mentioned body of work. Chapter 2 continues the work performed in (18) by optimizing the
algorithm used in that study. This optimization primarily includes the addition of a leash to ensure the algorithm can only find regions of best agreement in portions of compared toolmarks that can physically correspond. Inclusion of the leash prevents an issue named the “Opposite End Problem” where the algorithm would identify toolmarks as matching even though the region of best agreement between marks could not physically correspond (e.g. the opposite ends to two toolmarks do not match). Full details of algorithm operation are given in Chapter 2 as well as the results of the optimization as the algorithm is reapplied to quasi-striated shear cut marks made by pliers.

Chapter 3 continues the effort put forth in (19) to validate virtual toolmarking via a blind study. Screwdriver marks were created at a variety of different angles by former forensic examiner Jim Kreiser while the remaining researchers were kept blind to the correct tool-toolmark combinations. Using only the comparative algorithm and virtual toolmarking, correct identifications were made. This study aided the validation of virtual toolmarking (and the adequacy of the algorithm) to correctly predict the striated toolmark a tool would create at a variety of different angles.

With the success in separating quasi-striated marks and identification of striated toolmarks, the algorithm was next applied to impression toolmarks. Chapter 4 details the investigation of impression toolmarks made using chisels. Although the statistical algorithm was designed for striated toolmarks, it was thought to be applicable in this scenario because the chisels were sharpened using an abrasive wheel, thus they impart impression marks with a striated appearance. Results from this investigation into pseudostriated impressions are presented.
References


7. http://projects.nfstc.org/firearms/module09/fir_m09_t05_05.htm


CHAPTER 2. OPTIMIZATION OF A STATISTICAL ALGORITHM FOR OBJECTIVE COMPARISON OF TOOLMARKS

A paper accepted by the Journal of Forensic Sciences

Ryan Spotts¹, B.S.; L. Scott Chumbley¹, Ph.D.; Laura Ekstrand¹, M.S.; Song Zhang¹, Ph.D.; James Kreiser², B.S.

¹Ames Laboratory, Iowa State University, 2220 Hoover, Ames IA 50011
²Illinois State Police, Retired, 3112 Sequoia Dr., Springfield, IL 62712

Introduction

In recent history, the legitimacy of scientific testimony has been questioned in several court cases – specifically Daubert v. Merrell Dow Pharmaceuticals, 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 assumption of toolmark examination – every tool has its own unique surface that will leave a unique mark.

Screwdriver marks are among the most studied due to their uniform and continuous striae. They have been previously characterized using stylus profilometry and confocal microscopy 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 between matching and nonmatching pairs to a reasonable degree of accuracy (2).
Studies of other tools that produce striations have also been conducted. Pliers are another type of tool that can create a variety of marks. Cassidy, one of the first to study sequentially manufactured pliers (5), found toolmarks produced by plier teeth – such as when a burglar would twist off a door knob to enter a building – 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. This study established the uniqueness of marks created by plier teeth; however the analysis was based on logical reasoning and not backed by mathematical analysis. More recently Bachrach et al. studied tongue and groove pliers marks created on brass pipe, galvanized steel pipe, and lead rope (4). Test marks were made using a singular tooth from the pliers to create a striated mark. This study found that marks could be compared when made on different material but with less accuracy than marks made on the same material.

Petraco et al. has studied striated chisel marks (3). The test marks created by the chisels were striated but discontinuous, resulting in patches of striations. Unfortunately the nature of the created marks was too difficult for the employed software to analyze.

While regularly striated marks have received the majority of research attention, the extension of mathematically based studies to other forms of toolmarks is also highly desirable. The results discussed in this paper investigate the applicability of the algorithm employed in (2) to quasi-striated marks created by slip-joint pliers. This type of plier mark was chosen for two reasons. Firstly, the type of mark produced, termed a shear cut, presents a more difficult pattern for identification than a fully striated mark. Secondly, pliers such as these and other tools that produce shear cut marks are routinely used by criminals to steal
copper from construction sites. A July 30, 2013 report on CNBC stated that copper theft in the US has become a 1 billion dollar industry (6). Thus, objective examination and identification of shear cut toolmarks is becoming increasingly important in law enforcement.

An initial investigation on slip-joint pliers was conducted by Grieve (7). These results having shown promise, this paper presents results on shear cut marks made by 50 sequentially manufactured slip-joint pliers. While the initial results revealed the algorithm could correctly separate a large majority of matching and nonmatching pairs, some algorithm parameter values and options that work well for regularly 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 marks described as quasi-striated.

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 nonmatching. 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, which is integral to the operation of the algorithm, when confronted with similar topography at opposite ends of the data sets. This possibility was first noted during research on regularly striated screwdriver toolmarks (1, 2).
The current study involves complete analysis of fifty pliers using various parameter values and an option that accounts for the “Opposite End” problem. The results of the study, including a brief description of the statistical algorithm, are discussed below.

**Experimental Methodology**

Fifty sequentially manufactured slip-joint pliers were obtained from Wilde Tool Co., Inc. It is common knowledge within the field of study that the manufacturing process significantly affects the toolmarks that are created (8, 9). Thus, although the manufacturing process and test sample creation was previously described (7) it deserves restatement.

The pliers start as pieces of steel that are hot forged into 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 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.

Figure 1: Slip-joint pliers in finished and unfinished states. The ‘A’ and ‘B’ sides were labeled as shown for every plier, with ‘B’ appearing on the branded side.

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 the 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 the Association of Firearm and Toolmark Examiners as “opposed jawed cutters whose cutting blades are offset to pass by each other in the cutting process” (10). 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 the author who is a retired forensic examiner.

Figure 2: Shows the exact location on the pliers used to shear cut wire samples.

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 surfaces on the ‘A’ and ‘B’ sides of the sample were scanned and analyzed. A schematic showing the process is shown in Figure 3.
Figure 3: The left photo shows an example wire sample mid-shear cut, revealing how the toolmark gains its angle. The right photo shows the ‘B’ side sample, also revealing why the created mark is not completely circular. A similar ‘A’ side sample also exists but is not shown in the right photo.

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 Infinite Focus Microscope G3 (Alicona). Scans were completed at 10x magnification with a two micron vertical resolution. An example image obtained using the IFM is shown in Figure 4. Similar to the initial study (7), 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 of parallel striae that are not continuous along the length of the mark.
Figure 4: Example scan of a sheared copper wire. The dashed line is the short edge while the solid line is referred to as the long edge.

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 background generated when making an IFM scan. 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 (11), but a brief description of these follows. Firstly, the 2D image texture and quality map from the IFM operating software were used to remove those points that were too dark or had a poor quality value. These points cause spikes in the visual 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 data points to exclude from the analysis.

![Uncleaned data vs. detrended and cleaned data](image)

**Figure 5:** Uncleaned data is on the left while the detrended and cleaned data is shown on the right.

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 (7), the comparative algorithm used was discovered to have the same limitation that prevented it from operating effectively in certain instances in (2). 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. Figure 6 schematically shows this process.

Figure 6: Generalized example of the iterative optimization step.

Once the region of highest correlation is found during the optimization step, two shifts are applied and compared – random shifts and rigid shifts. This is the validation step. The size of the “validation” windows that are shifted are user-determined. During the rigid shift step a user-defined window is moved a set distance from the best correlation window on each trace and the correlation at that point is calculated. For the random shift step the same size window is moved randomly calculated distances from the best correlation window for the two comparison scans and the correlation is again calculated. An example of rigid and random shifts 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 correlation values to the random shift correlation values by the algorithm produces the statistical values to be mentioned.
In the initial investigation (7), 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.

As Figure 8 illustrates, through random chance opposite ends of a mark are occasionally selected as having the regions of highest correlation between marks for the selected window size. 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 pair of windows to correspond. 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.
Figure 8: Example of computer algorithm false positive (T1 = 8.09). Opposite ends of the toolmark were misidentified as a positive match – indicated by the “high” value for the T1 statistic (See text).

To address this problem a “leash” was applied to the search window of the original algorithm (2) during the optimization step, the purpose being to limit the comparison distance between profiles. In this case 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. Where contextual information exists concerning the shape of a mark (such as
exists for a distinctive mark like those used) this in no way affects the objective performance of the algorithm.

Figure 9: Generalized example showing possible ranges for the iterative search window.

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 clearly finds a reasonable location for best correlation but now computes a low value for the Wilcoxon Rank Sum test statistic (T1), indicating a nonmatching pair.

A Wilcoxon Rank Sum test statistic (centered and scaled to have a nominal standard deviation 1) 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.
Figure 10: Plier comparison from Figure 8 after implementation of the leash.

Comparison is now consistent with the expected result for nonmatching toolmarks (T1 = 0.63).

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 probability of observing a small T1 statistic will decrease. 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.
Currently there is not a definitive T1 value that perfectly defines when a match or non-match pair has been confirmed. This is due to the nature of the data – the comparisons are not independent events. Even if a definitive value were created for this study, it would likely not be applicable to other toolmark comparisons due to inherent differences in toolmark variability between different tools. Statements of correct or incorrect “identification” in this paper are qualitative – when the matching pairs consistently have significantly larger T1 values than nonmatching pairs it is fair to state that the algorithm is correctly separating (“identifying”) the majority of the pairs. A more advanced statistical argument is necessary to truly state whether an individual comparison was correctly identified.

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 comparison types were performed using data from both the long and short edges as defined in Figure 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.
Table 1: Example methodology for Set 1.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Plier Number</th>
<th>Side</th>
<th>Mark Number</th>
<th>Plier Number</th>
<th>Side</th>
<th>Mark Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>A</td>
<td>2</td>
<td>1</td>
<td>A</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>A</td>
<td>6</td>
<td>1</td>
<td>A</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>B</td>
<td>2</td>
<td>1</td>
<td>B</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>B</td>
<td>6</td>
<td>1</td>
<td>B</td>
<td>8</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>A</td>
<td>2</td>
<td>2</td>
<td>A</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>A</td>
<td>6</td>
<td>2</td>
<td>A</td>
<td>8</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>B</td>
<td>2</td>
<td>2</td>
<td>B</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>B</td>
<td>6</td>
<td>2</td>
<td>B</td>
<td>8</td>
</tr>
</tbody>
</table>

Set 2: Comparing known nonmatching pairs. Data for Set 2 were created by comparing marks made by different sides of the same pair of pliers. 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.

Table 2: Example methodology for Set 2.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Plier Number</th>
<th>Side</th>
<th>Mark Number</th>
<th>Plier Number</th>
<th>Side</th>
<th>Mark Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>A</td>
<td>10</td>
<td>1</td>
<td>B</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>A</td>
<td>12</td>
<td>1</td>
<td>B</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>A</td>
<td>14</td>
<td>1</td>
<td>B</td>
<td>14</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>A</td>
<td>10</td>
<td>2</td>
<td>B</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>A</td>
<td>12</td>
<td>2</td>
<td>B</td>
<td>12</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>A</td>
<td>14</td>
<td>2</td>
<td>B</td>
<td>14</td>
</tr>
</tbody>
</table>

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.
Table 3: Example methodology for Set 3.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Plier Number</th>
<th>Side</th>
<th>Mark Number</th>
<th>Plier Number</th>
<th>Side</th>
<th>Mark Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>A</td>
<td>16</td>
<td>2</td>
<td>A</td>
<td>16</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>A</td>
<td>16</td>
<td>4</td>
<td>A</td>
<td>16</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>A</td>
<td>18</td>
<td>4</td>
<td>A</td>
<td>18</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>A</td>
<td>18</td>
<td>3</td>
<td>A</td>
<td>18</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>A</td>
<td>20</td>
<td>3</td>
<td>A</td>
<td>20</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>A</td>
<td>20</td>
<td>4</td>
<td>A</td>
<td>20</td>
</tr>
</tbody>
</table>

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.

Observation of Figure 11 shows that the 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.
Figure 11: Results for search and validation window sizes of 200 and 100 pixels. A) Current long edge results B) current short edge results C) prior long edge results and D) prior short edge results (7). Note the difference in scale between the current data and early published results.

Results from the initial investigation that used the original algorithm are shown in Figure 11 for comparison. The same source data were used in both experiments. Note that the leash included as a fix to the Opposite End problem has resulted in a substantial improvement of algorithm performance 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 statistical separation of the quartiles for the long edge data.
Data Sets 2 and 3 are still centered approximately on zero, and there 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. Results from both long and short edges were similar to each other.

![Figure 12: Data Set 1 results for a 2:1 window ratio. A) Long edge. B) Short edge.](image-url)
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 did return a result for the 2:1 ratio analysis. The algorithm returned a result more than 98% of the time.

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.

Figure 15: Data Set 1 results for a 4:1 window ratio. A) Long edge. B) Short edge.
Figure 16: Data Set 1 results a 6:1 window ratio. A) Long edge. B) Short edge.

Figure 17: Data Set 2 results for a 4:1 window ratio. A) Long edge. B) Short edge.

Figure 18: Data Set 2 results a 6:1 window ratio. A) Long edge. B) Short edge.
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. The median values for the Set 1 data ratios were still qualitatively 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

The results presented add further credence to the basic assumption involved in 
toolmark identification, namely, that all manufactured tools are unique due to the machining 
processes used in their manufacture. This uniqueness is transferred to toolmarks as the tool 
is employed. Use of advanced characterization methods and computer algorithms can, to a 
large degree, allow objective comparison and identification of a series of toolmarks.

When the research transitioned from regularly striated to quasi-striated marks, it 
became apparent that a parameter optimization of the algorithm employed is necessary for 
different tools. This optimization led to improved results. The algorithm used 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 were obtained correctly.
In the most ideal scenario there would be 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 window sizes employed the algorithm typically returns a “correct” answer for most of the window combinations. Note also that the majority of these outlier points stem from larger window sizes.

Table 4: Comparison of returned T1 values for selected outlier data points.

<table>
<thead>
<tr>
<th>Plier Comparison</th>
<th>Search and Validation window size (pixels)</th>
<th>Data Variation for Specific Comparisons</th>
<th>Data Set</th>
<th>Match</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100-50</td>
<td>200-100</td>
<td>500-250</td>
<td>1000-500</td>
</tr>
<tr>
<td>P3A16 - P32A16</td>
<td>0.60</td>
<td>0.75</td>
<td>-2.72</td>
<td>5.91</td>
</tr>
<tr>
<td>P27B16 - P28B16</td>
<td>0.18</td>
<td>-0.09</td>
<td>5.45</td>
<td>0.33</td>
</tr>
<tr>
<td>P34A18 - P35A18</td>
<td>0.03</td>
<td>-0.64</td>
<td>0.51</td>
<td>5.45</td>
</tr>
<tr>
<td>P2B18 - P2B18</td>
<td>0.59</td>
<td>2.08</td>
<td>5.96</td>
<td>0.03</td>
</tr>
<tr>
<td>P20A14 - P20B14</td>
<td>-0.53</td>
<td>-0.16</td>
<td>3.99</td>
<td>0.32</td>
</tr>
<tr>
<td>P4A12 - P4B12</td>
<td>-1.16</td>
<td>-0.02</td>
<td>7.36</td>
<td>-4.22</td>
</tr>
<tr>
<td>P4A12 - P4B12</td>
<td>-0.66</td>
<td>-0.54</td>
<td>0.40</td>
<td>5.66</td>
</tr>
<tr>
<td>P18A12 - P18B12</td>
<td>1.02</td>
<td>-0.02</td>
<td>1.53</td>
<td>5.53</td>
</tr>
<tr>
<td>P30B2 - P30B4</td>
<td>0.14</td>
<td>1.20</td>
<td>3.62</td>
<td>6.11</td>
</tr>
<tr>
<td>P37B2 - P37B4</td>
<td>-0.51</td>
<td>0.37</td>
<td>3.33</td>
<td>-4.31</td>
</tr>
<tr>
<td>P23B6 - P23B8</td>
<td>1.74</td>
<td>-1.15</td>
<td>-0.36</td>
<td>6.22</td>
</tr>
</tbody>
</table>
If the underlying hypothesis behind the application of the T1 statistic is that matching pairs will have more correlation than nonmatching pairs, one might assume that this also holds true if one uses more search and validation window combinations. A simple experiment was done to see the effect of 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 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 considerably above a zero T1 value for matching pairs and remains near zero for nonmatching pairs.

![Figure 21: Average T1 values obtained when combining multiple window sizes for analysis.](image-url)
While the above observation is interesting it is recognized that providing any statistical merit to such an analysis would require substantial further development of the arguments on which the algorithm is based. This is necessary since correlations based on different algorithm parameter values will have different distributions, both in matching and nonmatching cases. Such a task is not anticipated at this time for this research group; however it is suggested as a possible field of endeavor for the reader.

Observation of Table 4 suggests that each unique type of toolmark will most likely require a study to determine what the best set of operating parameters is for that particular mark, and that any attempt to operate the code using a “one size fits all” mentality is insufficient for this algorithm. Future work will investigate whether an automated method to quickly evaluate multiple search and validation window combinations is possible. Such an enhancement would greatly speed analysis as new and more complicated toolmarks are examined.

Finally, while the quasi-striated marks examined in this study involve an added complexity when compared to the regularly striated marks previously examined, they are still less complex than, for example, impression marks. As the toolmark to be analyzed becomes more and more complex it is becoming increasingly likely that development of a truly robust objective algorithm will involve moving from a linear pixel-to-pixel comparison of the data to one that involves an area comparison. While this would represent a major shift as it relates to the operation of the current algorithm, exploratory efforts using this line of approach are already underway by other research groups (12).
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 in comparing regularly striated marks. The algorithm was optimized and applied using a leash option to the search and validation windows 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 non-match). A high degree of separation in the data was observed although sufficient statistical separation was not achieved. While the results have improved, more work is needed to increase the robustness of the identification process. Future improvements to the analysis method may involve automated means to examine combinations of search and validation windows quickly or, more radically, changing the process to compare areas rather than single linear files.

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. Also thank you to Dr. Max Morris for the critical review of this study and technical guidance.
This work was supported by the National Institute of Justice and the U.S. Department of Energy (DOE), Office of Science, Basic Energy Sciences, Materials Science and Engineering Division. The research was performed at the Ames Laboratory, which is operated for the U.S. DOE by Iowa State University under contract number DE-AC02-07CH11358. Funding was provided by award number 2009-DNR-119 from the National Institute of Justice.

References


CHAPTER 3. ANGULAR DETERMINATION OF TOOLMARKS USING A COMPUTER GENERATED VIRTUAL TOOL

A paper accepted by the Journal of Forensic Sciences

Ryan Spotts\textsuperscript{1}, B.S.; L. Scott Chumbley\textsuperscript{1}, Ph.D.; Laura Ekstrand\textsuperscript{1}, M.S.; Song Zhang\textsuperscript{1}, Ph.D.; James Kreiser\textsuperscript{2}, B.S.

\textsuperscript{1}Ames Laboratory, Iowa State University, 2220 Hoover, Ames IA 50011
\textsuperscript{2}Illinois State Police, Retired, 3112 Sequoia Dr., Springfield, IL 62712

Introduction

Recent history has seen scientific testimony challenged in numerous court cases since the \textit{Daubert v. Merrell Dow Pharmaceuticals, Inc.} ruling. In particular, comparative forensic examinations such as firearms and toolmark examination have been increasingly challenged by attorneys due to the perceived subjective nature of the examination and the mistaken impression that there is a lack of scientific studies aimed at addressing its primary assumption, namely, that all tools produce a unique toolmark. Scores of studies have been conducted and published in the Association of Firearm and Tool Mark Examiners (AFTE) Journal to validate comparative forensic examination (1, 2), and give credence to the assumption that every tool contains a unique surface topography capable of creating a unique mark if used against another surface. A number of more recent studies using objective computer-based algorithms have shown toolmarks created by the same tool will statistically be more similar than toolmarks created by similar tools, even when the tools are manufactured sequentially (3, 4, and 5). Nichols has written a thorough literature review that responds to many criticisms of forensic examination (1).
The large body of work that currently exists has not stopped judicial challenges, however, since in many instances published studies still rely on what is considered a subjective assessment. Even though current methods have been shown to produce false identification rates of approximately 1% (1), ideally what is desired is an entirely objective analysis that can provide known error rates that are similarly low.

Current research has focused on the use of computers and algorithms to increase the robustness of tool-toolmark identifications. Several different approaches have been made. Chumbley has used a nonparametric Mann-Whitney U statistic (referred to as T1) to compare three-dimensional topography data obtained using an optical profilometer (4). Toolmarks made using screwdriver tips at various angles were compared. Results from this study were in agreement with experiential evidence from forensic examiners. T1 values were largest (e.g. increased probability of a matching pair) when toolmarks created by the same screwdriver tip edge and angle were compared. An angular dependence for screwdriver toolmarks was also found by Bachrach et al. (6). It was concluded from this study that screwdriver toolmarks created at different angles can appear to be completely different toolmarks. Specifically, results from this study showed comparisons of screwdriver marks at 15 and 45 degree angles made on the same medium had an error rate of approximately 50% (6).

Wei, C. et al. has proposed the use of correlation cells for rapid ballistic identifications (7). Correlation cells, select regions on a surface, can incorporate three-dimensional topography data for quick matching between marks. An initial test utilized cartridge cases fired from ten sequentially manufactured pistol slides. Three-dimensional topography data was obtained using a confocal microscope and analyzed under the
constraints that three out of three correlation cells must show positive or negative correlation for known and unknown matches respectively to be declared. Initials results of this methodology yielded no false positive or negative identifications (7).

Petraco has applied machine learning to analyze striation patterns on cartridge cases fired from 9mm Glock pistols and screwdriver marks made in lead (5). A confocal microscope was used to measure three-dimensional topography data from each toolmark. There were 162 profiles measured from 24 Glocks and 290 profiles measured from 58 screwdriver edges (29 screwdrivers). Simulated mean profiles based on the real data resulted in 720 total profiles from the Glocks and 1740 total profiles from the screwdrivers available for analysis. Initial algorithmic identifications on subsets of the total collected data found an error rate of 2.5% for the Glock toolmarks with a 95% confidence interval of 1.3 to 3.2% and an error rate of 6.5% for the screwdriver marks with a 95% confidence interval of 3.5 to 10%. Further refinement of the pattern recognition process through increasing the analyzed data sets reduced error rates to 0.03% and 0.01% respectively (5). The results were good and demonstrated the capability of pattern recognition in toolmark analysis.

The aforementioned studies, and the study to be presented, show that computer based systems can be valuable aids in forensic examinations. In this study the ability to specifically characterize a tool surface and relate that surface to the mark that tool could be expected to generate under any given set of conditions is demonstrated. The analysis involves a completely objective assessment of the tool surface which is then used to create a computer generated “virtual tool” that can be manipulated at will by the forensic examiner to create any number of “virtual toolmarks”. When combined with a statistical algorithm for making comparisons, it then becomes possible not only to directly relate a tool to a toolmark but also
to predict with a high degree of accuracy the conditions that existed (specifically, angle of attack) when the toolmark was made.

**Experimental Methodology**

*Sample generation:* A blind study was devised where six sequentially manufactured screwdriver tips and one randomly selected screwdriver from another manufacturer (treated as an unknown) were used by a forensic examiner to create a series of 20 toolmarks. Researchers were kept ignorant of the correct tool-toolmark combinations throughout the experiment. Toolmarks were created on approximately 1.5 by 1.5 inch lead plates using the jig shown in Figure 1. Different fixtures allow toolmarks to be made at angles ranging from 30 - 85° in 5° increments. For this study the examiner was instructed to use any angle they wished as well as either side of the screwdriver. This allowed 168 possible combinations of tool and angle used to create marks when both sides of the screwdriver are considered.

![Figure 1: Photograph of jig used to create toolmarks. The 85° angle holder is in use and the direction of the manually applied force is indicated.](image)
Ultimately twenty toolmarks were made. Toolmarks were labelled with randomly generated three digit ID tags and sent to the research group. An answer key containing the correct combinations of tool-toolmarks was kept in a sealed envelope until the research group presented their identifications.

*Surface characterization:* After receiving the tools and toolmarks, the surface topography was obtained from the six screwdriver tips and the created toolmarks using an optical profilometer (Alicona InfiniteFocusSL). This equipment employs focus variation to scan and obtain accurate three-dimensional topography data. Focus variation works through a precisely controlled z-axis that is able to bring varying portions of the surface into focus. When an object is in focus, the object sharpness (function of light returned to sensor) is at a maximum. The sharpest data for each pixel is then used to construct the three-dimensional topography (8). An approximately 2x7 mm² area of topography data were obtained from each toolmark. The most completely toolmarked areas of the substrate were selected for analysis. An example of one lead plate with a labeled toolmark showing the area scanned is shown in Figure 2.

![Figure 2: Example toolmark used in the study.](image)
Each side of the screwdriver tips was scanned at a 45° angle relative to the vertical axis of the infinite focus microscope objective. An apparatus was used to hold the screwdriver tips at precisely the same angle for each scan. Figure 3 shows an example of the portable scanning equipment and the apparatus holding a screwdriver tip.

![Figure 3: Scanning equipment and apparatus for holding screwdriver tip (45° angle defined).](image)

Each data set (from both toolmark and screwdriver tip) had a horizontal pixel resolution of 3.914 μm and vertical resolution of 1.007 μm. Scans were completed at a 10x magnification. Approximately 5 minutes were needed to obtain data from each sample.

**Noise reduction:** Any method of automated data collection will contain noise due to random variables. In using an optical system noise (e.g. artifacts such as small spikes or holes) can be generated by imperfections that greatly alter the normal scattering of light collected from the surface that is used for generating an image. To fix this issue and to eliminate unneeded data a cleaning procedure is required.

Before an automated statistical algorithm can be used the operator must ensure that the data to be compared only contains information relevant to the comparison. For example, Figure 4 shows the complete scan obtained from a toolmarked surface. Since the goal is to
compare the toolmark, not the unmarked surface of the lead plate on which the mark is made, extraneous data at the edges of the scan must be removed. Software named Mask Editor was developed to allow manual cleaning of each data set. A “painting” tool allows the user to manually paint over (mask) areas of unneeded data without altering the data itself; painted regions are simply ignored in further analysis. The software is programmed to find the largest contiguous unmasked region, so the masking process is not tedious.

![Figure 4: Toolmark 556 (dimensions approximately 7x2mm) during masking process.](image)

After masking, spike artifacts on the screwdriver tips are dealt with through the use of a seventh-order polynomial that was applied to each row and column within a data set. Any data points with a depth 100 μm greater or less than the predicted value were removed. This threshold was determined through experimentation. Any small holes (20 pixels or less in diameter) in the data were then filled using linear interpolation. Toolmarks were detrended to remove small angular differences (relative to the z-axis of the optical profilometer) that occur when scanning multiple samples (9).

Data comparison: Once cleaned, the data sets were suitable for comparisons using a developed software suite and the previously developed algorithm (4). Termed the Mark ANd Tool Inspection Suite (MANTIS) this software was developed to allow comparisons of toolmarks to actual marks or virtual marks generated from three-dimensional tool
topography. A full overview of the virtual mark generation procedure is given in (10), and the reader is directed there for complete technical details. Briefly, the software uses the tool data set that was scanned and cleaned at a known angle to generate a virtual surface that can be rotated to any angle one wishes to investigate. The edge of the screwdriver may be thought of as being analogous to a mountain ridge. If one were to superimpose a coordinate system, X would be the direction perpendicular to the ridge, Y the direction along the ridge, and Z the height above sea level. As one hikes along the ridge you are always at the highest point on the ridgeline (in an XZ plane), even though you encounter small peaks and valleys on your walk. So it is with the toolmark. The left side of Figure 5 shows a reconstructed virtual image of the screwdriver edge and it is clear that as you move along the “ridge” in the Y direction, 2D cross-sections (XZ planes in the shown coordinate axis) taken at any point will produce a profile similar to that shown below the virtual image. By finding the highest topographical point in all possible 2D cross-sections moving along the Y direction, a virtual “effective” topography profile (YZ data) is constructed of the ridgeline, that captures all the peaks and valleys encountered as you proceed along the crest.

Suppose one now tilts the screwdriver. While peaks and valleys still exist in the Y direction, the surface topography that results due to the tilt means that the location of the crest may change. It is possible that for any specific XZ plane what was once a low point below the crest now becomes the highest point of the ridge, as shown in Figure 5 (right side). In other words, the YZ data generated at the first angle is altered when tilting to the second angle.
Figure 5: Generic screwdriver tip with highest topographical point found for one cross-section at two different angles.

In creating a virtual mark the highest points at any angle (i.e. the crest of the ridge containing both peaks and valleys) are assumed to be the first points to contact another surface and, therefore, are the primary cause of the striae observed in a toolmark. By taking the effective virtual tool topography at the crest of the ridge at any angle and virtually “dragging” it against a virtual substrate (i.e., expanding the YZ data in the ‘X’ direction) a virtual toolmark is generated and it is this topography that is compared to the physical toolmark. The process virtually mimics the actions performed by a forensics examiner – using a test mark created by marking a substrate with a suspect tool to compare to an evidence mark. This first level approximation does not account for material properties, applied force or the other two rotation axes; it was assumed the effective topography was fully transferred and the virtual mark resulting from this process is “basically” the inverted
effective tool topography. The complexity of the virtual marking process will allow for higher level approximations in the future where the sliding action of the virtual tool against the virtual substrate has a large impact on the results.

Figure 6 shows an example of a virtual toolmark created by characterizing the tooltip as compared to data obtained from the corresponding physical toolmark. The solid line crossing the physical toolmark represents the path trace that produced the profile data shown.

Figure 6: a) Virtual mark generated at 75°. b) Corresponding physical toolmark created at 75°.

The MANTIS software was used to generate virtual toolmarks at angles from 30 to 85° at 5° increments for each screwdriver tip and perform comparisons to physical toolmarks using the statistical algorithm discussed in (4). For this study three combinations of user-determined parameters were used. Multiple parameter sets were used to see if the results
varied due to different user-input. Parameters that were varied include the pixel widths of search and validation windows as shown in Table 1.

**Table 1: Algorithm input parameters**

<table>
<thead>
<tr>
<th>Parameter Set</th>
<th>Search Window</th>
<th>Validation Window</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>500</td>
<td>200</td>
</tr>
<tr>
<td>2</td>
<td>500</td>
<td>500</td>
</tr>
<tr>
<td>3</td>
<td>800</td>
<td>800</td>
</tr>
</tbody>
</table>

The algorithm outputs a single T1 statistic for each comparison. Three parameter combinations, twelve angles, twelve effective screwdriver tip edges (N.B. the unknown screwdriver was not examined) and twenty toolmarks led to 8,640 T1 values. Based on observations in (10, 3 and 4), a maximum T1 value is expected to occur when there is a statistical likelihood of a match. This will only occur if the correct combination of virtual mark, angle and toolmark are compared. The T1 value is also expected to decrease as comparisons are made of similar marks made at angles varying by greater than 10° (4, 11).

A heuristic critical value was used to determine whether the T1 output corresponded to a matching tool-toolmark combination. It is known that for matching pairs, the statistical distribution of T1 depends on many factors, and cannot be derived analytically based on what is currently known. For nonmatching pairs, observations show that the distribution of T1 is closer to that suggested by simple theory (approximately normal with zero mean and unit variance), but cases have occurred where this does not agree with experimental data. Hence for this study, a value was heuristically chosen. A T1 value greater than 6 was heuristically treated as the critical identification criteria – If the standard asymptotic distribution theory held, this would correspond to incorrectly identifying a nonmatching pair as matching with
an extremely low probability. The toolmark and creation angle were identified as matching by the maximum T1 value above the critical value.

Results

Figure 7 shows the identification process graphically using both matching and nonmatching tool-toolmark combinations from the results of this study. For reference, screwdriver tips are named using a simplified scheme (e.g. T20A where the number can range from 20 to 25 for the 6 different tips and either an A or B is present to differentiate between the two tool edges) and toolmarks were named using randomized 3 digit ID codes. T1 values from this example show the nonmatching pair fluctuated between approximately -4 and 3, but never above the critical value of 6 – consistent with a nonmatching comparison. The shown matching pair T1 values were above the critical value from 30 to 45° with the maximum value occurring at 35°. These results would indicate that the algorithm had identified toolmark 408 as being created by screwdriver tip T20A at 35°. The algorithm was able to identify the correct tool-toolmark combination over an angular range of -5 to +10° (30 to 45°). This methodology was repeated for each parameter set over all possible combinations to determine the matching pairs.

![Figure 7: Statistical output using Parameter Set 2 as a function of angle.](image-url)
Results from the identifications are shown in Table 2. Utilizing only the algorithm and virtual marking, every toolmark (20/20) was correctly paired with the screwdriver tip edge that created it, with toolmarks from the additional unknown screwdriver being identified through exclusion. No false positives occurred during identifications. On average the maximum T1 value occurred at an angle 6.12 degrees less than the answer key creation angle. The average range in Table 2 refers to the average angular range that the algorithm was able to identify matching tool-toolmark combinations – for parameter set 3 the algorithm identified matching tool-toolmark combinations on average from 11.7° less than the maximum T1 value to 13.6° greater than the maximum T1 value. This means the algorithm calculated a T1 value greater than the critical value over a total range of 25.3 degrees.

Table 2: Tabulated results from the study

<table>
<thead>
<tr>
<th>Toolmark</th>
<th>Answer Key</th>
<th>Parameter 1</th>
<th>Parameter 2</th>
<th>Parameter 3</th>
<th>Avg. Angle Mismatch (°)</th>
</tr>
</thead>
<tbody>
<tr>
<td>420</td>
<td>T23B at 75°</td>
<td>T23B at 70°</td>
<td>T23B at 70°</td>
<td>T23B at 70°</td>
<td>-5.0</td>
</tr>
<tr>
<td>408</td>
<td>T20A at 40°</td>
<td>T20A at 35°</td>
<td>T20A at 35°</td>
<td>T20A at 35°</td>
<td>-5.0</td>
</tr>
<tr>
<td>787</td>
<td>T24A at 55°</td>
<td>T24A at 45°</td>
<td>T24A at 45°</td>
<td>T24A at 45°</td>
<td>-10.0</td>
</tr>
<tr>
<td>556</td>
<td>T22B at 65°</td>
<td>T22B at 55°</td>
<td>T22B at 60°</td>
<td>T22B at 60°</td>
<td>-6.7</td>
</tr>
<tr>
<td>821</td>
<td>T21A at 45°</td>
<td>T21A at 35°</td>
<td>T21A at 35°</td>
<td>T21A at 35°</td>
<td>-10.0</td>
</tr>
<tr>
<td>983</td>
<td>T25A at 45°</td>
<td>T25A at 35°</td>
<td>T25A at 35°</td>
<td>T25A at 35°</td>
<td>-10.0</td>
</tr>
<tr>
<td>872</td>
<td>T20B at 60°</td>
<td>T20B at 55°</td>
<td>T20B at 55°</td>
<td>T20B at 55°</td>
<td>-5.0</td>
</tr>
<tr>
<td>648</td>
<td>T25A at 60°</td>
<td>T25A at 55°</td>
<td>T25A at 50°</td>
<td>T25A at 55°</td>
<td>-6.7</td>
</tr>
<tr>
<td>552</td>
<td>Unknown at 30°</td>
<td>No match</td>
<td>No match</td>
<td>No match</td>
<td>-</td>
</tr>
<tr>
<td>514</td>
<td>T22A at 30°</td>
<td>No match</td>
<td>No match</td>
<td>T22A at 30°</td>
<td>0.0</td>
</tr>
<tr>
<td>416</td>
<td>T23A at 70°</td>
<td>T23A at 60°</td>
<td>T23A at 65°</td>
<td>T23A at 65°</td>
<td>-6.7</td>
</tr>
<tr>
<td>916</td>
<td>T21A at 70°</td>
<td>T21A at 60°</td>
<td>T21A at 60°</td>
<td>T21A at 60°</td>
<td>-10.0</td>
</tr>
<tr>
<td>409</td>
<td>Unknown at 70°</td>
<td>No match</td>
<td>No match</td>
<td>No match</td>
<td>-</td>
</tr>
<tr>
<td>423</td>
<td>T24B at 80°</td>
<td>T24B at 70°</td>
<td>T24B at 70°</td>
<td>T24B at 70°</td>
<td>-10.0</td>
</tr>
<tr>
<td>212</td>
<td>T20B at 80°</td>
<td>T20B at 75°</td>
<td>T20B at 75°</td>
<td>T20B at 75°</td>
<td>-5.0</td>
</tr>
<tr>
<td>394</td>
<td>T25B at 40°</td>
<td>T25B at 35°</td>
<td>T25B at 35°</td>
<td>T25B at 35°</td>
<td>-5.0</td>
</tr>
<tr>
<td>674</td>
<td>T21B at 40°</td>
<td>T21B at 35°</td>
<td>T21B at 30°</td>
<td>T21B at 30°</td>
<td>-8.3</td>
</tr>
<tr>
<td>448</td>
<td>T24B at 40°</td>
<td>T24B at 40°</td>
<td>T24B at 40°</td>
<td>T24B at 35°</td>
<td>-1.7</td>
</tr>
<tr>
<td>286</td>
<td>T23B at 40°</td>
<td>T23B at 35°</td>
<td>T23B at 40°</td>
<td>T23B at 35°</td>
<td>-3.3</td>
</tr>
<tr>
<td>616</td>
<td>T22A at 75°</td>
<td>T22A at 75°</td>
<td>T22A at 75°</td>
<td>T22A at 70°</td>
<td>-1.7</td>
</tr>
</tbody>
</table>

Avg. Range  -7.6°, +9.4°  -11.8°, +12.4°  -11.7°, +13.6°  Avg. Mismatch -6.12°
Discussion

Correctly identifying every toolmark provided validation of both the comparison algorithm used and the ability to create a virtual toolmark that accurately reflects what can be expected in real life. These results open up a number of possibilities for the future use of virtual markings, both in the area of basic science and applied research.

Although initial results are promising, deficiencies were observed in the results. The first deficiency is the clear bias of the maximum T1 value occurring at lower angles than the creation angle. Deflection of the jig was investigated as the potential cause since the direction of deflection would naturally lower the true creation angle. Figures 8 - 10 show photos captured from slow motion video recorded while creating additional toolmarks. The deflection was measured using photo imaging software on the photographs for toolmarks created at 30, 55 and 85°. Measurements revealed a deflection of approximately 2 to 4° can occur during toolmark creation caused by movement of the screwdriver tip in the holder and not deflection of the screwdriver tip holder. However, a 2 to 3° deflection during toolmark creation explains much of the apparent angular bias in the results. Since virtual marks were compared to toolmarks in 5° increments, and a 3° deflection occurring during toolmark creation was possible, it is not surprising that the virtual mark 5° less than the nominal creation angle would have higher correlation to the toolmark.

Figure 8: Before and during toolmark creation, measured angles of 29.5° and 27.8°.
The second deficiency was observed with Toolmark 514. Two parameter settings failed to identify Toolmark 514 (two false negatives). Since this toolmark was nominally created at 30°, and compared to a 30° virtual mark, it was thought that deflection may have
played a role in this lack of identification. Toolmark 514 was compared to a 25° virtual mark; however T1 values were still well below the critical value for the two parameter settings. Inspection of the data revealed that for parameter settings 1 and 2, the regions of highest correlation were found between two prominent topography features that were not actually related. The larger sized parameter setting 3 was large enough to distinguish between the prominent topography features – resulting in the correct identification of the tool-toolmark combination. In other words, careful examination of the data is necessary when conflicting results are obtained for different parameter settings to ensure that the computer is truly conducting a valid comparison.

The final deficiency investigated were instances where the maximum T1 value occurred when the angle was 10° lower than the creation angle, as occurred for toolmarks 787, 621, 916, 983 and 423. Algorithm output variation was investigated as a possible cause of this deficiency in the results. The comparison algorithm utilizes random number generation during analysis to select toolmark profile regions for correlation computation (4). Due to the use of random numbers the same T1 value is not computed for repeated comparisons – there is a small amount of output variation.

To test this, Toolmark 787 and its corresponding screwdriver tip T24A were used for repeated comparisons. Toolmark 787 was created using a 55° holder. However the maximal T1 value was computed at 45°. To determine whether output variation was the potential cause of an additional 5° of error after error due to deflection, 50 comparisons were performed using Parameter Set 2 at 45, 50 and 55°. The results are presented in Figure 11 using box plots. The box plots are composed of a solid black line representing the median value, boxes representing the 1st and 3rd quartiles (bottom and top of the box respectively),
whiskers representing a maximum of 1.5 times the interquartile range and circles to represent outlier data points. The box plots show complete separation of each measurement, indicating output variation is not a likely cause of this deficiency.

![Box plots showing output variation of repeated comparisons.](image)

**Figure 11: Output variation of repeated comparisons.**

While every attempt is made to control the creation of the toolmarks, there are inherent fluctuations in the applied force during toolmark creation. Fluctuations in applied downward force are caused by the relative alignment of the jig to the lead plate. The screwdriver will not be aligned perfectly within the jig before each use and the lead plate thickness varies between samples (lead plates were purchased at the same nominal thickness). These factors result in fluctuations of the applied force and ultimately the topography of the created mark. This inherent variation may be the cause of some of the observed error.
Summary and Conclusions

Twenty out of twenty tool-toolmark combinations were correctly identified, and the marking angle reasonably estimated, in a blind study comparing virtual toolmarks created by analyzing a tooltip to actual toolmarks by means of an objective statistical algorithm. On average virtual marking estimated the angle of creation within approximately 5° of the true angle of creation. The results from this study indicate that toolmarks are most identifiable using the employed algorithm when made within approximately 10° of each other. Deficiencies in the results were addressed. The heuristically chosen critical T1 value, while useful, is not entirely defensible due to the dependence of the matching pair T1 statistical distribution on input parameters. It was also found that 2 to 3° of deflection occurs during toolmark creation. The deflection was the root cause of some of the estimated angular inaccuracy. If 3° of deflection occurred during toolmark creation, it is expected that the virtual mark 5° lower than the nominal creation angle would have higher correlation to the toolmark. In instances where the maximal T1 value occurred at an angle 10° off of the nominal creation angle it was found that inherent variation of the calculated T1 values was not a factor. It is likely that 5 of the 10° is due to deflection while the remaining 5° of angular mismatch was possibly due to inherent variation of the applied force during toolmark creation.

This first level study provided validation for the concept of creating “virtual toolmarks” as an aid in the identification process by directly relating a tool to a mark and in allowing determination of certain parameters related to the marking. Evidence that a fixed angular range exists over which toolmark identifications can occur was found - in agreement with prior studies that indicate an identification angular range exists (12). Virtual marking
could ultimately serve as a useful tool to aid forensic examiners in more accurately estimating toolmark angles, speeding analysis, as a training tool, and in obtaining basic information concerning perception of what does or does not constitute a match. Finally, this study provided further validation of the primary assumption of comparative forensic examination, namely, that even sequentially manufactured tools contain identifiably unique topographies.

Acknowledgments

The research group thanks the Omega Company for supplying sequentially manufactured screwdriver tips. Without help from companies such as Omega, this research would not be possible. The research group also thanks Dr. Max Morris for his manuscript review and technical guidance.

This work was supported by the National Institute of Justice and the U.S. Department of Energy (DOE), Office of Science, Basic Energy Sciences, Materials Science and Engineering Division. The research was performed at the Ames Laboratory, which is operated for the U.S. DOE by Iowa State University under contract number DE-AC02-07CH11358. Funding was provided by award number 2011-DNR-0230 from the National Institute of Justice.

References


CHAPTER 4. OBJECTIVE ANALYSIS OF IMPRESSED CHISEL TOOLMARKS

A paper accepted by the *Journal of Forensic Sciences*

Ryan Spotts, B.S.; L. Scott Chumbley, Ph.D.

Ames Laboratory, Iowa State University, 2220 Hoover, Ames IA 50011

**Introduction**

Challenges to comparative forensic science have created the need for an objective means of toolmark identification to aid forensic examinations. These challenges, such as the 1993 Daubert v. Merrell Dow Pharmaceuticals, Inc. court case over admissibility of expert testimony or the more recent 2010 National Academy of Sciences (NAS) report challenging a “lack” of scientific background, have had profound impacts on the field of comparative forensic examination and resulted in many validation studies. Many of the studies, including this one, provide evidence for the primary assumption of comparative forensic science – every tool has a unique surface topography such that it can leave an identifiably unique toolmark.

Prior research that has been conducted with chisels involved the creation of striated chisel marks. By AFTE (Association of Firearm and Tool Mark Examiners) definition, striated toolmarks are created by force and movement of a tool in a direction approximately parallel to the surface being acted upon (1). Petraco et al. used 5 sequentially manufactured chisels to make 50 striated marks at 30° in a lead medium. Topographical data were obtained using a confocal microscope. Toolmarks were created by dragging the chisels across the substrate while using a jig for angular control. However this process resulted in patches of continuous striations that were not suitable for analysis with the available methods (2).
Zheng et al. also created striated chisel marks. Ten sequentially manufactured chisels were used to create striated toolmarks by using a jig to drag the tools against a copper substrate with a controlled force at a 90° angle. Two known and unknown marks were created for each tool, resulting in 40 total toolmarks. A stylus profilometer was used to obtain the topographical information. Using the cross correlation function (CCF), the unknown toolmarks were correctly identified to the chisels that created them after establishing a critical CCF value. The critical value was obtained with the known marks by comparing known matching and nonmatching toolmarks (3).

While many toolmark studies, and the previous research on chisel toolmarks, have focused on striated marks, little has been done with impressed chisel toolmarks. By AFTE definition, impressed toolmarks are created by force and movement of a tool in a direction approximately perpendicular to the surface being acted upon (1). A criminal could leave an impressed chisel mark while breaking into a safe (4) or a locked desk drawer.

The current study involves the analysis of 50 sequentially manufactured cold chisels (chisels designed for metal working at room temperature) that were used to create 500 impression toolmarks. In this case, the impressions from the sharpened tool sides were analyzed and not the impression due to the working edge of the tool. Impressions made from a blunted edge are often what are observed by forensic examiners. Since the chisels in this study were acquired directly from the manufacturer, the chisels’ working edges had not yet been blunted by use. The toolmarks were compared using a statistical algorithm originally developed for the comparison of striated toolmarks. The algorithm, utilizing a Wilcoxon Rank Sum test statistic, has been used to statistically separate matching and nonmatching striated screwdriver marks (5) and quasi-striated plier marks (6) although complete
separation was not achieved. It was also observed that the degree of statistical separation achievable between matching and nonmatching toolmarks decreased with increasing toolmark complexity – from the striated screwdriver marks to the quasi-striated plier marks. Impression toolmarks are more three-dimensional in character than striated toolmarks and pose an even more difficult test for the algorithm. Results of this study will show the algorithm’s ability to statistically separate these matching and nonmatching impression marks.

**Experimental Methodology**

Fifty sequentially manufactured 1/4" cold chisels were obtained from Wilde Tool Company, Inc. To ensure sequentiality, the manufacturing process was observed by the researchers. It is well known that the manufacturing process influences the final tool topography (7), so a brief description of the relevant manufacturing processes is provided.

The material used to create the chisels is received as 5/16” cold rolled hexagonal bar. The bar is then sectioned into the necessary lengths to make the chisels. Each sectioned bar is hot forged on one end to create the functional end of the tool. After forging, the functional end (i.e. the eventual sharpened end) is trimmed to remove burrs and rough cut to the correct dimensions. The oxide scale is then abrasively blasted away. This is followed by an induction heat treatment and molten salt quench to obtain the final microstructure and desired mechanical properties. The chisels are then abrasively blasted a second time to remove residual salts or oxides. The functional end of the chisel is sharpened via an abrasive wheel. Gauges are set in the process line to ensure consistent sharpening angles between chisels. The final abrasive wheel sharpening imparts the functional tool topography that would be
expected to leave toolmarks at a crime scene. Since the chisels are held in place during
sharpening and the abrasive moves in one direction, the functional tool topography is linear
in nature.

Each chisel was placed into a numbered and enclosed plastic bag to ensure
sequentiality was maintained. Additionally, a unique identification number was punched into
the branded side of each chisel. Figure 1 shows an example of two chisels.

Figure 1: Example of two chisels used in the study. The impressed letters designate side
A, the unmarked side is side B.

To create impression toolmarks, a jig was created to hold a chisel at a 90° angle
relative to a lead substrate. Two two-pound weights and a sliding plate guided by rails were
used to strike one end of the chisel and create an impression of the chisel’s functional end.
The weight was released onto the chisel from a height of 6” (the top of the jig’s rails were
designed to be 6” above the top of the chisel) resulting in an impact energy of 2.66 joules.
Ten impression marks were made with each chisel, resulting in 500 total impressions.
However, each chisel has two sides for a total of 1,000 toolmarks. Figure 2 shows an
example of the jig and impression marks in a lead substrate. The tool side facing the branded
surface of the tool was named the ‘A’ side of the tool while the opposite side was named ‘B’.
Toolmark topography information was obtained and digitized using an optical profilometer. Each impression was approximately a 1x10 mm area. A scanning magnification of ten times, vertical resolution of 0.997 microns and horizontal resolution of 3.91 microns was used during scanning. An example of a toolmark “as-scanned” is shown in Figure 3.
The algorithm used in this study, originally developed in (5), was designed to be used with approximately planar, striated toolmarks such as those created by screwdrivers. However, in Figure 3 it is clear that the toolmark has a ‘V’ cross section and there is extraneous data around the edges of the toolmark that should not be included in the analysis. To overcome these issues, internally developed software named MaskEditor was used to manually clean the data. MaskEditor contains a “painting” tool that allows a user to mask over unneeded data. The masking process, although manual, is not tedious because the program is designed to use the largest contiguous region of data. The unnecessary data is not deleted but simply ignored during analysis. MaskEditor was also useful for extracting the two planar regions of the toolmarks. The user could carefully mask over all but half of the toolmark, leaving just one planar section. The unmasked planar section is then leveled making it suitable for analysis. Using this procedure, each impression mark was cleaned twice to extract the ‘A’ and ‘B’ planar sides of the toolmark. The masking and extraction of one toolmark is shown in Figure 4 – the solid line along the planar section corresponds to the topographical profile shown adjacent to it.

With the data cleaning complete, the toolmarks were ready for analysis. Algorithmic comparisons were performed between two 2D profiles from approximately the center of each planar section. The computer algorithm statistically compares both profiles and outputs a single value - referred to as a T1 value. For full details on how the algorithm operates, the reader is referred to (5, 8). The T1 value, due to issues of the matching pair statistical distribution dependence on a variety of factors (9), cannot be used to declare a definitive match.
However, a large T1 value indicates a high degree of correlation (consistent with matching toolmarks) while low, zero and negative T1 values indicate little to no statistical correlation (consistent with nonmatching toolmarks). Three different algorithm parameter settings (500-500, 750-750 and 1000-1000 pixels) were used for comparisons. The parameter settings were the pixel-widths of the search and validation window sizes (5). The window sizes correspond to the length of the topographical profile that computations were performed on at a given point in time during algorithm operation. It is known that different parameter settings (and varying search and validation window size ratios) can affect the results, thus three different settings were used to ensure consistency (6).

Three types of comparisons were performed for this analysis. Set 1 comparisons were performed between all matching toolmarks. For example the first toolmark, Chisel 1 side ‘A’
replicate 1 (C1A1) was compared to its nine replicates. Next C1A2 (the second replicate) was compared to its nine replicates. This procedure was repeated until all ‘A’ and ‘B’ side matching replicates had been compared (4500 comparisons). Set 2 comparisons were performed between the opposite sides of the same chisel for all replicates – e.g. C1A1 compared to C1B1 and their replicates (5000 comparisons). These toolmarks are expected to be nonmatching and could demonstrate that each side of the chisel effectively appears to be a different tool. Finally, Set 3 comparisons were performed between all of the nonmatching toolmarks of sequential chisels – e.g. C1A1 was compared to C2A1 and their replicates followed by C3A1 being compared to C4A1 and their replicates and so forth. This procedure was repeated for Set 3 so ‘A’ to ‘A’, ‘A’ to ‘B’, and ‘B’ to ‘B’ side comparisons were completed (7500 comparisons). A total of 17,000 comparisons were performed for each parameter setting and a total of 51,000 statistical comparisons were performed in this analysis.

Results

The data from the comparisons were organized and Figures 5-7 show the box and whisker plots for comparison Sets 1-3 for each parameter setting. The solid black line represents the median value, while the lower and upper bounds of the box represent quartiles. The whiskers are within one and a half times the interquartile range and outliers are signified by unfilled circles.
Figure 5: Statistical results using the 500-500 pixel-width parameter setting.

Figure 6: Statistical results using the 750-750 pixel-width parameter setting.
Discussion

Data from Set 1 comparisons for all parameter settings were relatively high T1 values, consistent with a high degree of correlation and matching pairs of toolmarks. The high degree of correlation present for the majority of the Set 1 comparisons has caused the median, upper bound and upper whisker to visually blend together. Sets 2 and 3 exhibited low to negative T1 values, consistent with little to no statistical correlation and nonmatching toolmarks as expected. The results are similar for all parameter settings, but complete separation between matching and nonmatching toolmarks, excluding outliers, was only achieved with the 500-500 pixel-width parameter setting.

Many outliers were observed, with the majority of outliers occurring in Set 1 comparisons, and a systematic cause was not found. The root cause may be due to the fact that the applied algorithm was not designed for this type of toolmark. Two observations - listed below - were also made that could contribute to the number of outliers.
1. The jig used to hold the chisels cannot hold them perfectly orthogonal to the substrate for each replicate; there are a couple degrees of angular variation possible.

2. The sliding plate cannot strike the chisel perfectly orthogonally for each replicate. The sliding plate also has a couple degrees of angular variation.

The mentioned factors combine to cause variation in the toolmarks that are created. This variation results in variation of the correlation between toolmarks. It should also be noted that many of the Set 1 outlier comparisons had T1 values over 5, still indicating a high degree of correlation but not to the degree of the majority of matching pairs.

This study demonstrates that this statistical algorithm can be applied to “pseudostriated” impression marks. While the results were good, an algorithm designed specifically for impression marks can be expected to perform better. It is hypothesized that it is effective in this scenario because the abrasive wheel imparts a regularly striated mark onto the chisel; the chisel then impresses a copy of the topography into the lead substrate. While the chisel does not create a striated toolmark by AFTE definition in the lead, the original striations caused by the abrasive wheel are regular and adequately impressed into the substrate, allowing analysis by the algorithm developed for 2D striated toolmarks.

**Summary and Conclusions**

Fifty sequentially manufactured chisels were used to create 500 impression toolmarks. When both sides of the chisel are considered, 1000 unique toolmarks had been created. All impressions were made by releasing a weight onto the chisel and orthogonally impressing the functional end of the chisel into a lead substrate.
Three types of algorithmic comparisons were made between toolmarks: matching pairs, nonmatching pairs from opposite sides of the same chisel, and nonmatching pairs from varying sides of sequential chisels. Three different sets of algorithmic parameters were used to analyze the data (resulting in a total of 51000 statistical comparisons). All of the parameter sets performed similarly, achieving strong separation between matching and nonmatching toolmarks. The parameter set containing 500 pixel-wide search and validation windows did achieve complete separation (excluding outliers). Numerous outliers were present in the data, primarily concentrated in the Set 1 matching toolmark data. However many of the Set 1 outliers still indicated a high degree of statistical correlation but not to the same degree as the majority of matching toolmarks. A systematic cause of the outliers was not found, but multiple factors were observed that cause variation between replicate toolmarks that could affect the degree of correlation. Some outliers may also be due to the inexact use of this algorithm to analyze impressed toolmarks.

This study demonstrates that the algorithm can successfully separate pseudostriated impression toolmarks. However an algorithm specifically designed for impression marks would likely outperform the current methods. The ability of current methods to separate matching and nonmatching impression toolmarks provides evidence for, and further validates the primary assumption of comparative forensic science – even sequentially manufactured tools contain unique, identifiable topographies.

**Acknowledgments**

The research group thanks the Wilde Tool Company Inc., for supplying us with sequentially manufactured tools and allowing us to visit and observe the manufacturing process.
This work was supported by the National Institute of Justice and the U.S. Department of Energy (DOE), Office of Science, Basic Energy Sciences, Materials Science and Engineering Division. The research was performed at the Ames Laboratory, which is operated for the U.S. DOE by Iowa State University under contract number DE-AC02-07CH11358. Funding was provided by award number 2011-DNR-0230 from the National Institute of Justice.

References


CHAPTER 5. SUMMARY AND CONCLUSIONS

Chapter 1 explained how forensic toolmark examination not only arose as a field of science but how it has been a useful tool for determining guilty parties through the 20th century until today. Legal challenges have significantly impacted the field and created a driving force for the creation of objective methods of toolmark identification to aid the current subjective methods. Many academic attempts have been made to create these objective methods but research is ongoing.

Chapter 2 built upon prior research of a statistical algorithm applied to quasi-striated shear cut plier marks. These quasi-striated marks are not completely regular throughout the toolmark, and pose a more difficult test for the algorithm than the previously tested striated screwdriver marks. In this study, an optimized version of the algorithm was used to reanalyze the quasi-striated toolmarks. This optimization primarily included the implementation of a leash to solve the “Opposite End Problem”. This issue transpired when the statistical algorithm would find regions of best agreement between two toolmarks in locations that could not possibly physically correspond (e.g. the opposite ends of two toolmarks). The leash prevented this issue by only allowing the algorithm to search limited ranges between two toolmarks to ensure correspondence. Results of this implementation significantly improved the separation between matching and nonmatching toolmarks. The varying parameter settings and parameter ratios showed improvement but still not to the degree of regularly striated toolmarks.

Chapter 3 provided further validation for the developed virtual toolmarking capability. This capability aims to directly relate tools to their marks. Tool topography is digitized, rotated to an angle of interest, analyzed and then used to create a virtual toolmark.
The virtual toolmark can then be compared to a physical toolmark to determine if they match. A blind study was performed to validate this procedure. Twenty toolmarks were created by a research collaborator using seven screwdriver tips. All of the tools and toolmarks were digitized using an optical profilometer. Virtual toolmarks were then created by each tool tip at every possible angle (30-85°, 5° increments) and compared to all of the physical toolmarks. All twenty tool-toolmark combinations were correctly identified with an average angle prediction of 6.36° less than the nominal creation angle. This study validated the virtual toolmarking procedure and the algorithm’s ability to correctly identify matching striated toolmarks. Virtual marking may someday be able to aid examiner training and speed toolmark analysis by correctly estimating critical toolmark parameters such as the angle of creation.

Chapter 4 featured the analysis of pseudostriated impression marks created using chisels. Although the algorithm was designed for striated toolmarks, it was found to be applicable in this scenario because the chisel edge topography was striated from the grinding process used to sharpen it. Hence, although the marks were impressions, they were referred to as pseudostriated. Results of this analysis demonstrate that the algorithm was able to separate matching and nonmatching toolmarks to a high degree using three different parameter settings.

The work contained in this thesis is another step towards objective methods of forensic toolmark examination. All of the results also provide evidence for the primary assumption of toolmark analysis – all tools contain a topography that is capable of imparting an identifiably unique toolmark.