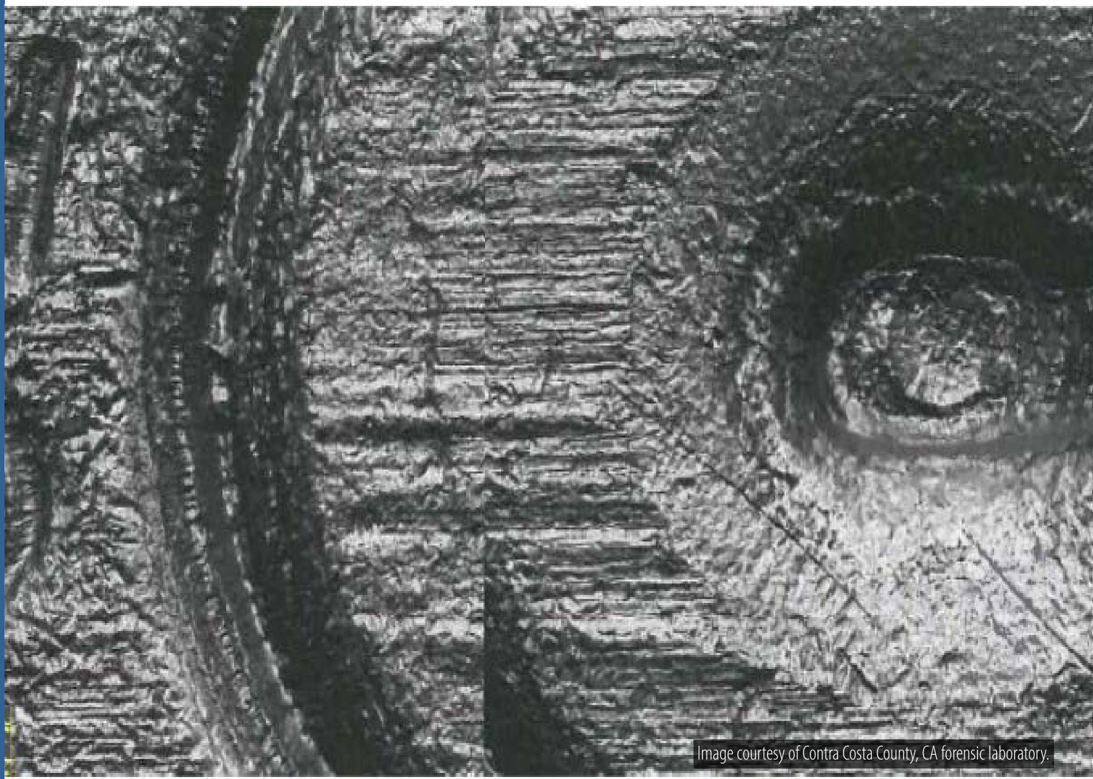
National Institute of Justice

Forensic Optical Topography Working Group Meeting

FINAL REPORT



Meeting Date: March 17–18, 2015

Report Date: April 17, 2015

National Institute of Justice
Office of Investigative and Forensic Sciences
810 Seventh Street, N.W.
Washington, D.C. 20531
FY2011 Award #2011-DN-BX-K564

Forensic Technology Center of Excellence

The Forensic Technology Center of Excellence (FTCoE) is a collaborative partnership of RTI International and its FEPAC [Forensic Science Education Programs Accreditation Commission]–accredited academic partners: Duquesne University, Virginia Commonwealth University, and the University of North Texas Health Science Center. In addition to supporting the National Institute of Justice's (NIJ's) research and development (R&D) programs, the FTCoE provides testing, evaluation, technology transition assistance, and other services for use by crime laboratories, forensic service providers, law enforcement, and other criminal justice agencies whose mission is to combat crime. NIJ funds the FTCoE to transition forensic science and technology to practice (Award Number 2011- DN-BX-K564).

The FTCoE is led by RTI, a global research institute dedicated to improving the human condition by turning knowledge into practice. With a staff of more than 3,700 providing research and technical services to governments and businesses in more than 75 countries, RTI brings a global perspective. The FTCoE builds on RTI's expertise in forensic science, innovation, technology application, economics, data analytics, statistics, program evaluation, public health, and information science.

Contents

Tables

Figures

1. OVERVIEW

The Forensic Technology Center of Excellence (FTCoE) at RTI International, in partnership with the National Institute of Justice (NIJ) and the National Institute of Standards and Technology (NIST), convened the Forensic Optical Topography Working Group Meeting on March 17 and 18, 2015 at the NIST campus in Gaithersburg, MD. The meeting included researchers and practitioners with a wide range of experience and knowledge regarding forensic applications of microscopy.

This working group seeks to establish the applicability and validity of optical topography to forensic investigations and to produce publications or training materials that can be accessed by the entire forensic community and that will provide guidance to practitioners on applications and recommendations for further research, development, and capacity assistance. Primarily, the working group will examine optical topography instruments, methods, data systems, and analysis from a practical perspective for ballistic and tool mark identification.

The meeting participants considered current technologies for optical topography, including the requirements for systems that may be deployed in crime laboratories. The extension of current ballistic identification methods to topographic methods was also examined. Participants noted the value of the comparison microscope to identification decisions in current practice and determined that it was unlikely that optical topography would supplant the comparison microscope as the primary tool for the forensic examiner in the near term. Instead, optical topography is likely to be a confirmatory tool or a method to examine very difficult comparison cases.

Participants reviewed current and past efforts to implement optical topography in the crime laboratory, including the application of confocal microscopy. These systems demonstrated that optical topography may permit the examiner to distinguish among consecutively manufactured firearms, although a great deal of work remains to establish accepted examination protocols. For example, data filters and matching algorithms will play a key role in the application of optical topography, but the validation of these analytical tools remains an issue.

NIST and the Federal Bureau of Investigations (FBI) have begun to collect reference data from test fires relevant to a wide range of firearms, including various combinations of research data, crime laboratory data, and instrument types. They have developed an XML data standard under the OpenFMC framework that could be used for interoperable sharing of ballistic identification data on a national basis. Many issues remain, including commercial acceptance, practitioner acceptance, the ability of laboratories to handle the large data files and computational load associated with topographic systems, and the collection of sufficient data to inform the development of analytical models.

The working group agreed that it would be advisable to hold a practical review of examination methods at the FBI Laboratory, which houses several optical topography instruments. The review would seek to establish current consensus concerning the application of optical topography to ballistic identification with respect to examiner practices, instrument requirements, training, and analysis. A small subgroup will meet at the FBI Laboratory at a date to be determined to complete this task.

2. GOALS AND OBJECTIVES

This working group will review the various technologies associated with the collection of optical topographic data, i.e., three dimensional (3D) data from surfaces using optical means, including confocal microscopy, interferometry, and focus variation. The working group confines its work to firearm identification and tool mark comparison, although it will briefly review other forensic applications.

The goals of the working group are as follows:

- Determine how to improve practitioner access to optical topography instrumentation and methods.
- Determine how optical topography may improve the ability to individualize firearms and tool marks and provide a more objective and reliable basis for forensic comparisons. These comparisons are expert-based. Although we do not yet have a way to develop a probabilistic framework at this time, we can try to determine how these methods can be used to improve the ability of the individual examiner to quantitate their findings.
- Stimulate future research and development in optical topography, and identify gaps in the research portfolio.

The National Institute of Justice (NIJ) has been funding research in optical topography for over 6 years. Nonetheless, few state and local laboratories have implemented the technology. The working group will consider specific issues relevant to the gap between research and practice, including

- Is this an area that NIJ should continue to fund? Research has shown that this technology can be important, but is there a practical application?
- What are the barriers that prevent transition of optical topography from research to forensic practice?
- What information and resources are needed by labs to implement these systems?
- What investment is needed at the laboratory to make the transition happen?
- What are the specifications of current systems? Is the technology mature enough to be used in practical applications?
- Do we have enough knowledge of the accuracy of the techniques?
- Is the technology ready for casework?
- Will the instrumentation cost decrease as the technology becomes more mature?

What is the current state of research in the area of optical topography? Where will the technology be in 3 to 4 years? The group will also examine the extension of current procedures for comparisons based on two-dimensional (2D) image data to 3D topographic images. As part of this latter task, the group will develop process maps that capture current and proposed comparison methods, including aspects related to data interpretation, such as baseline correction. In addition, the group will examine ways to improve the ability of NIST and affiliated organizations to collect reference data that may be used to validate mathematical approaches, although mathematical and statistical analysis will not be emphasized in the current phase of the working group.

3. TOPICS DISCUSSED

Topics discussed by the Forensic Optical Topography Working Group are presented in the following subsections. Presentations for some of these topics can be found at the following links:

- Elements of Firearm and Tool Mark Identification and Relation to 3D Imaging
- Considerations for Optical Topography in Forensic Science
- Firearm and Tool Mark Identification Reference Data and Collection

3.1 Elements of Firearm and Tool Mark Identification and Relation to 3D Imaging

Firearms identification is the forensic science discipline that identifies bullets, cartridge cases, or other ammunition components as having been fired from, or worked through the action of, a particular firearm to the practical exclusion of other firearms. A gun is a type of tool that produces marks on bullets and cartridge cases because the metals in bullets and cartridge cases are necessarily softer than those in the firearm. Firearms identification is based on the principle that most of the tool marks that firearms produce on bullets and cases are characteristic of the individual firearm. For example, tool marks are left on the bullet from the barrel rifling, resulting in striations in the land and groove impressions.

Rifling impressions are an indication of the source of the barrel. In part, the type of barrel may be identified because of variations among manufacturers with respect to the design of firearms.

Also, each barrel may be different from others of the same type and, therefore, may be individualized based on imperfections left in the barrel during the manufacturing process and visible at the microscopic scale. The firearm examiner may use these tool marks to identify class characteristics—for example, the type of firearm that fired a bullet. The examiner may also identify marks that permit the individualization to a particular firearm. Tool marks are generally present as either striated or impressed tool marks.

Striated tool marks are scratches or scrape marks that appear as parallel lines, or striae, along the bullet, cartridge, or cartridge case. Impressed tool marks are formed when the cartridge or bullet are forcefully pressed against another surface and give the appearance of being stamped into the metal. Striae arise from rifling marks on bullets, while impressions are left by the impact of the breech face and firing pin on a cartridge case. Forensic examiners use comparison microscopes to identify impressed and striated tool marks.

The comparison microscope was a major development in forensic firearm identification and consists of two microscopes joined together by an optical bridge. The comparison microscope is a minimum requirement to perform firearm and tool mark identification. The decision concerning whether a particular impression or striae is a match to another impression or striae is a subjective judgment based on the human examiner's pattern-matching ability and experience. This subjective identification can lead to questions, such as the following:

- How much similarity is enough for identification?
- How small a shard can you get to match with practical certainty?
- How does human cognition factor in?
- How many points of comparisons are required to confirm an identification?
- Should examiners report an identification based on a sufficient number of points of comparisons?
- What is the minimum criteria required to confirm identification, etc.?

In 1992, the Association of Firearm and Tool Mark Examiners (AFTE) adopted the "Theory of Identification," which was later updated in 2011 and reads[1]

1. The theory of identification as it pertains to the comparison of tool marks enables opinions of common origin to be made when the unique surface contours of two tool marks are in "sufficient agreement."
2. This "sufficient agreement" is related to the significant duplication of random tool marks as evidenced by a pattern or combination of patterns of surface contours. Significance is determined by the comparative examination of two or more sets of surface contour patterns comprised of individual peaks, ridges, and furrows. Specifically, the relative height or depth, width, curvature, and spatial relationship of the individual peaks, ridges, and furrows within one set of surface contours are defined and compared to the corresponding features in the second set of surface contours.
3. Agreement is significant when the agreement in individual characteristics exceeds the best agreement demonstrated between tool marks known to have been produced by different tools and is consistent with agreement demonstrated by tool marks known to have been produced by the same tool. The statement that "sufficient agreement" exists between two tool marks means that the agreement of individual characteristics is of a quantity and quality that the likelihood of another tool making the mark is so remote as to be considered a practical impossibility.

Currently, the interpretation of individualization/identification is subjective in nature, founded on scientific principles and based on the examiner's training and experience. Traditionally, firearms examiners have used their inherent cognitive ability to recognize and compare patterns of consecutive matching striae (CMS) between known and unknown tool mark specimens and determine if sufficient agreement for identification exists based on their recollection of the best known non-matching agreement they ever observed (from their training or other experience). Using this "pattern matching" approach, if the striated agreement exceeds this non-quantified threshold in the mind's eye of the examiner, then, assuming the absence of subclass influences, the examiner may conclude there is sufficient agreement for identification of the two marks to a single source. Although this method is subjective, the identification criteria used in pattern matching can be very accurate, although error rates are not well understood. In 1997, conservative quantitative consecutive matching striae (QCMS) criteria for the identification of striated tool marks was proposed by Biasotti and Murdock. Using QCMS, sufficient agreement is defined as two (2) runs of three (3x) CMS or one (1) run of six (6x) CMS for 3D marks and two (2) runs of five (5x) or one (1) run of eight (8x) CMS for 2D marks, assuming the absence of subclass influences (Modern Scientific Evidence: The Law and Science of Expert Testimony [West Group, 1997 & Supps., 1999-2001]). In actual empirical research, in 3D tool marks, one (1) group of four (4x) CMS is the highest CMS run that has been observed. In practice, examiners have observed that these thresholds exceed the best agreement observed in known non-matching tool marks. Thus, some examiners consider the QCMS thresholds to provide objective criteria for their discipline. The QCMS thresholds were designed to be conservative based on empirical research, which is why the criteria is higher for shallow (2D) or other difficult marks. Various statistical scientists are examining the firearms identification process. This work may benefit from an examination of QCMS criteria. For example, whether striae are considered a match by a human examiner or an algorithm derived from optical topographic data, QCMS may be used to derive a statistical representation of a match. Also, optical

[1] Association of Firearm and Tool Mark Examiners, " Theory of Identification as it Relates to Tool Marks: Revised", *AFTE Journal*, Vol. 43, No. 4 (Fall 2011), p. 287.

topography may be used to improve the ability of examiners to match striae, thus improving their ability to make identifications in difficult cases. As adopted in 1992, the range of conclusions was preceded by: "The examiner is encouraged to report the objective observations that support the findings of tool mark examinations. The examiner should be conservative when reporting the significance of these observations."[2] The conclusions of identity are based on the comparison of individual characteristics made after eliminating the possibility of subclass influence. In other words, the presence of some CMS observed between two firearm-produced tool marks may be due to the type of firearm and not indicative of the individuality of one firearm versus another of the same type. The ranges of conclusions are identification, exclusion, inconclusive, and no value, as described below:

- Identification: Agreement of a combination of individual characteristics and all discernible class characteristics, where the extent of agreement exceeds that which can occur in the comparison of tool marks made by different tools and is consistent with the agreement demonstrated by tool marks known to have been produced by the same tool.
- Inconclusive:
  - Some agreement of individual characteristics and all discernible class characteristics, but insufficient for an identification.
  - Agreement of all discernible class characteristics without agreement or disagreement of individual characteristics due to an absence, insufficiency, or lack of reproducibility.
  - Agreement of all discernable class characteristics and disagreement of individual characteristics, but insufficient for an elimination.
- Elimination: Significant disagreement of discernible class characteristics and/or individual characteristics.
- Unsuitable: Unsuitable for examination.

It is important to note that the word "inconclusive" was never intended to be used alone without an explanation in a report of examination, as outlined in a, b, or c, above.

3.2 Considerations for Optical Topography in Forensic Science

3.2.1 General Observations in Conventional Optical Microscopy vs Topographical Microscopy

The implementation of optical topography may present difficulties to crime laboratories. In one case of a laboratory that wanted to purchase a confocal microscope, the facility was located on the 4th floor, where there were issues with vibration from a nearby highway that would have made the system unusable. The crime lab could not obtain access to the ground floor, so they developed a plan to prepare a room to be vibration-free, but the cost of implementing the plan would have been $100,000. Thus, the full cost of deploying a confocal microscope included the system itself, training, and substantial facility costs. These ancillary costs are a significant barrier to the broad adoption of optical topography in forensic science.

[2] AFTE Committee for the Advancement of the Science of Firearm and Tool mark Identification. (2011, June 14). *AFTE Response to the 25 Questions related to firearms and tool mark examinations promulgated by the RDT&E IWG.* Retrieved March 27, 2015, from www.AFTE.org: http://afte.org/downloads/RDT&E%20IWG%2025%20Questions%206.14.11%20-%20AFTE%20Response%20w%20cov%20let.pdf

In conventional optical microscopy, the optical image contrast is measured two dimensionally as I (x,y) and is predominantly a function of slope, shadowing, multiple reflections, optical properties, directional illumination, and, indirectly, the local height variations. In topographical microscopy methods such as interferometric and confocal microscopy, the instrument can measure variations in height, i.e., Z (x,y), directly and independent from illumination and shadowing effects. However the data can be subject to signal-to-noise issues, distortions, and data dropouts.

3.2.2 Classification of Surface Topography (3D) Measurement Methods[3,4]

General considerations may be found for the selection of surface topographic methods in the literature, including ISO Standard 25178-6.[3,4] Surface topography measurement methods include Line Profiling, Areal Topography, and Area Integration. Line profiling methods include contact stylus scanning, phase shifting interferometry, circular interferometric profiling, and optical differential profiling. Areal topography includes contact stylus scanning, phase shifting interferometry, coherence scanning interferometry, confocal microscopy, confocal chromatic microscopy, structured light projection, focus variation microscopy, digital holography microscopy, angle resolved scanning electron microscopy (SEM), SEM stereoscopy, scanning tunneling microscopy, atomic force microscopy, optical differential profiling, point autofocus profiling, and photometric stereo. Area integrating methods of surface texture measurement include total integrated scatter, angle resolved scatter, parallel plate capacitance, and pneumatic area integration.

Some of the key limitations for profiling instruments are the spatial resolution, lateral range bandwidth limits as well as the range of vertical and horizontal resolution. Some instruments are limited by their ability to discern very steep slopes, which may show up as artifacts in a 3D image. Therefore, the maximum measureable slope is a critical requirement for surface topographical systems. Similarly, systems must minimize dropouts and outliers and permit the quantitation of these artifacts.

There are standards that should be used by the examiner to ensure the calibration and traceability of their measurements, including NIST sinusoidal reference standards 2073a, 2074, and 2075.

For firearms identification, the relevant field of view ranges from a few microns to a few millimeters. In general, meeting participants believe that the relevant horizontal length scale for optical topography in firearms identification is one micron, although much reference data are based on a horizontal resolution of 3.25 microns, including NIST's own breech face data. This resolution is based on optical limitations.

The choice of a system must include consideration of cost and speed. In most cases, surface topography systems require some type of scanning in the z-direction (depth), which slows data collection considerably in comparison to traditional, 2D microscopy.

3.2.3 Methods Relevant to Firearm and Tool mark Identification

Excellent and detailed discussion of optical topography methods are available, including (insert Vorburger presentation from the web).

[3] ISO Standard 25178-6 (2010), *Classification of Methods for Measuring Surface Texture*.
[4] T.V. Vorburger et al., Int. J. Adv. Manuf. Technol. **33**, 110 (2007).

Coherence Scanning Interferometry (CSI),[5] also called vertical scanning interferometry or scanning white light interferometry, measures changes in interference signal strength across a surface. In CSI, the object of interest has height features (h) that vary according to the surface of the object. The object is then scanned mechanically, which provides a continuous, smooth scan of the interference objective along the optical axis in the z direction. While the object is scanned, the intensity data, I, is recorded for each image point. As an object is scanned vertically along the optical axis, the interference varies and surface heights are inferred by observing where the interference effect is the strongest. A strength of CSI is vertical resolution of approximately 3 nm with lateral resolution that is comparable to confocal microscopy at approximately 1µm; however, CSI is limited in its ability to optically resolve steep, sloped surfaces. One application of CSI in firearms research is ALIAS, from Pyramidal Technology in cooperation with Heliotis (Switzerland), where CSI was coupled with a high-speed camera to capture topographical images.

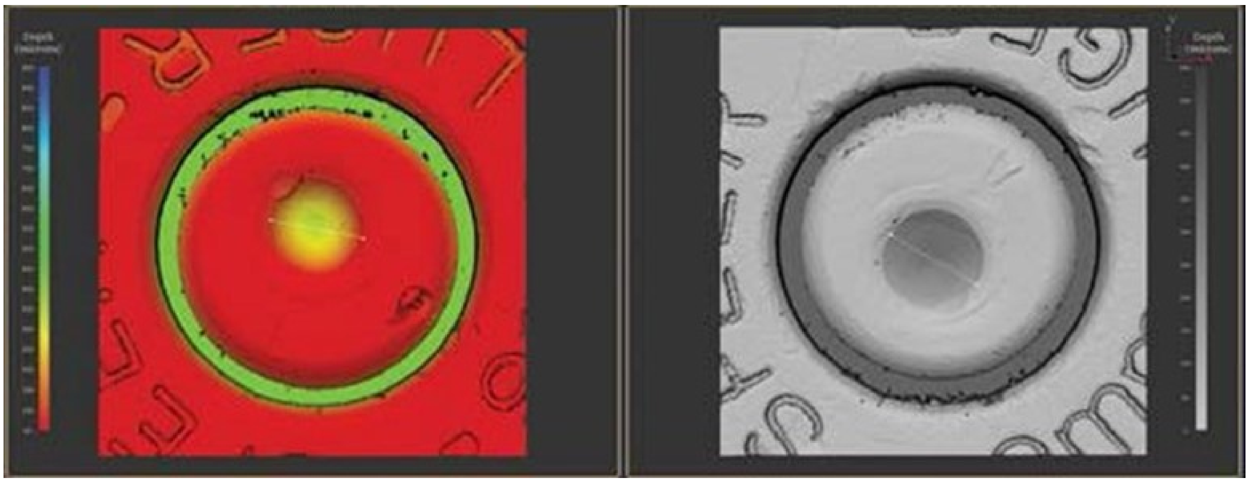


Figure 1. Topographical image of breech face and firing pin impressions[6]

Confocal microscopy[7] includes disc scanning confocal microscopy, laser scanning confocal microscopy, and programmable array confocal microscopy. Confocal microscopy works by assembling a series of thin slices of the cross section of the object of interest, taken along the vertical axis. Once assembled, these thin sections can build a very detailed 3D image of the object of interest. The strength of confocal microscopy lies in the vertical resolution of approximately 3 nm and lateral resolutions of approximately 1µm, both similar to CSI. The limitation of this method is that the signal decreases and can become unreliable for surfaces with steep slopes, with dropouts and outliers present at 15º.

[5] ISO Draft International Standard (DIS) 25178-604, *Geometrical product specification (GPS) – Surface texture: Areal – Part 604: Nominal characteristics of non-contact*.

[6] www.pyramidaltechnologies.com

[7] ISO New Work Item 25178-607, ISO/Technical Committee 213/ Working Group 16 , *Geometrical product specification (GPS) – Surface texture: Areal – Part 607: Nominal characteristics of non-contact (confocal microscopy) instruments.*

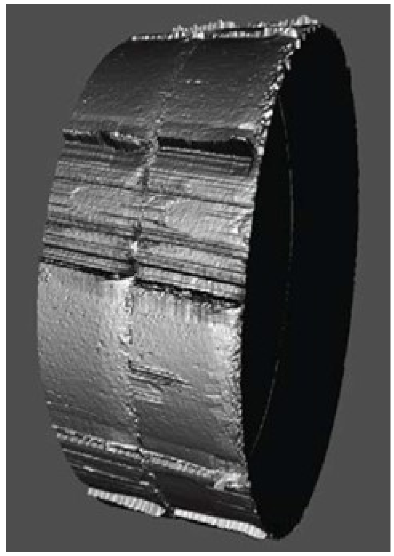


Figure 2. Topographical image of a pair of fired bullets[8]

Focus Variation Microscopy[9] is specifically designed for surface metrology and can be used to characterize surface texture. Focus variation works by moving the shallow depth of focus of the optical system over the object of interest while continuously scanning the surface to produce a 3D model. As the optical lens changes in distance from the object, the variation in sharpness or focus is measured and used to determine depth. The advantages of Focus Variation are that the image is produced in true color and that steep sloped surfaces can be measured. The limitations of focus variation are the vertical resolution of approximately 100 nm and the lateral resolution of several pixels.

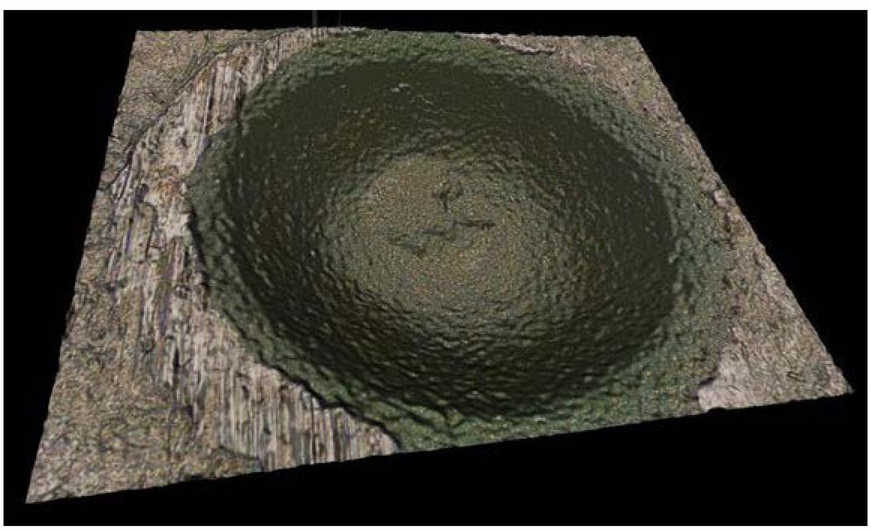


Figure 3. Firing pin impression on cartridge case measured with focus variation. Overlay of reflectance and topography images.

[8] From P. Murphy et al., Three-Dimensional Virtual Comparison Microscope for Bullets, http://www.forensictechnology.com/Portals/71705/docs/technote_3dvcmbullets_20100429.pdf(May, 2010).

[9] ISO Final Draft Int. Std. (FDIS) 25178-606 *Geometrical product specification (GPS) — Surface texture: Areal —Part 606: Nominal characteristics of non-contact (focus variation).*

Some other novel 3D imaging techniques with forensic application include photometric stereo[10] and chromatic confocal microscopy.[11]

3.2.4 Calibration Issues and Standards

Table 1. Types and names of measurement standards

Type	Name
A	Depth measure standard
B	Tip condition measurement standard
C	Spacing measurement standard
D	Roughness measurement standard
E	Profile coordinate measure standard

At what length level at x,y,z is it relevant to calibrate these instruments? At what level would they have to be the same to say they truly match? Calibration should correspond to the length scale of the features that one wishes to image. As outlined in Table 1, Current Standard Reference Materials (SRM) were developed for the calibration of the Integrated Ballistic Identification System (IBIS), which is used for database searches of unknowns. These SRMs are for two automated systems, not comparative systems such as the comparison microscope or optical topography. They are relevant to checks on methods, not instruments. Of course, the level of calibration is also limited by the type of optical microscope that is used.

Daubert challenges may be aimed toward questioning the subjective aspects of the comparisons and conclusions rendered by the firearm examiner; however, such subjective assessments may be sufficient and even superior to the quantitation that could be provided by optical topography. In any case, calibration and instrumentation issues remain open questions in this regard.

Overall, the important properties of topographical microscopes for firearm and tool mark analysis are vertical and lateral resolution; maximum measurable slope (important for measuring firing pin impressions); minimization and quantification of dropouts and outliers; cost; and speed.

3.3 Ballistic and Tool Mark Identification Reference Data and Collection

Instruments for confocal microscopy, focus variation microscopy, photometric stereo microscopy, coherence scanning interferometry, and stylus profilometer all give x,y,z data (µm). However, all instrument manufacturers save the data in their own proprietary format, which prevents interoperability, interlaboratory comparison, and certain types of analysis. To improve interoperability, NIST has created an Open Forensic Measurement Consortium (OpenFMC) with the goal of establishing a standard file format for the exchange of 3D forensic topography measurements. The primary point of contact for the OpenFMC is Alan Zheng of NIST. OpenFMC has standardized the use of the XML 3D Surface Profile (X3P) format defined in the ISO 25178-72 standard (OpenGPS). Currently, Sensofar and Cadre Research are the only companies participating in the OpenFMC.

[10] M.K. Johnson et al., "Microgeometry Capture using an Elastomeric Sensor", ACM Trans. Graph. **30**, 4, Article 46 (July 2011), DOI = 10.1145/1964921.1964941.
[11] ISO 25178-602: 2010 *Geometrical product specifications (GPS) — Surface texture: Areal — Part 602: Nominal characteristics of non-contact (confocal chromatic probe) instruments.*

OpenFMC a binary file format that contains four records:

1. Header, data types, and axes definition,
2. Metadata regarding the instrument and user,
3. Profile data (x,y,z), and
4. Checksum of the xml-document.

Metadata can include any type of information, including user-defined fields that would enable the use of X3P in crime laboratory information systems. The consortium has developed open-source read/write function converters to enable the adoption of the standard by commercial vendors. Currently, file metadata include information on the firearm, ammunition, measurement conditions, and other elements useful for the development of a reference collection relevant to scientific and statistical study. The XP3 format contains undefined fields that allow for any additional information to be collected in a fifth record. The XP3 format exists on most instruments, so many laboratories can utilize it already, though adoption is not universal.

X3P is compressed data to minimize the storage size. For example, an ASCII file would be about 50 MB, whereas X3P files are 25 MB. With X3P, cross-modality matching the same test fire measured on two different systems can be compared. To test the interoperability of this file system, Cadre research successfully imported NIST Nanofocus confocal data into their TopMatch matching correlation software and were able to correctly identify all test fires using their TopMatch GelSight system. Thus, interoperability between collection methods and laboratories is possible, even when the optical topography instruments are very different.

3.4 NIST Ballistics Tool Mark Database

NIST has established a reference collection of optical data for firearms at www.nist.gov/forensics/ballisticsDB. This database is an open–access, research database containing reflectance microscopy images as well as 3D surface topography tool mark data. This database was created to improve the transition of 3D surface topography from research to application and to improve method development and validation, as well as to allow for development of uncertainty estimates for objective ballistic identification. This database is also being used to validate new algorithms. The horizontal resolution in breech face impression images is 3.25 microns in the database, which is limited by the optical methods used.

The database contains a vast array of test fires collected from consecutively manufactured slides, consecutively manufactured barrels, and persistence/decay studies, as well as collected test fires from different firearms and ammunition. The FBI actively collaborates with NIST in the development of the reference collection. The FBI is conducting an extensive collection of data to add to the database, including test fires from each of the firearms in its collection. This includes 1,038 different models. The FBI has a broad set of optical collection systems (Focus Variation, Confocal, Interferometry [PSI, VSI] Holography, Fringe Projections), all of which will be used in its collection activities. NIST would like to add more crime laboratory sets and firearm types to its collection. NIST also intends to create a limited access database that can be used to validate algorithms for interpretation of firearms identification, such as data search algorithms.

The primary point of contact for the NIST database is Dr. Alan Zheng, who also provides leadership for the OpenFMC.

3.5 Applications of Optical Topography

3D imaging technologies have been utilized in many industries, including semiconductor, materials, paper, energy, optics, forensics, and other fields that are concerned with surface texture analysis. Surface texture consists of micro-roughness, roughness, waviness, and form:

- **Micro-roughness** is the finest component of the surface texture and is defined as the set of high frequencies or smallest wavelengths (2.5µm, -8 µm) resulting from either sampling noise or the microscopic relief of the structure of the material.
- **Roughness** is defined as wavelengths ranging from 20 µm to 500 µm and can vary rapidly, depending on the horizontal position. The roughness is decisive for the texture and gives an indication of the nature of the materials, production process, and the machining methods used. However, roughness is not unique; it's the extraction of the roughness and the characterization of those features that gives the object uniqueness.
- **Waviness** is defined as wavelengths ranging from 0.5 to 2.5 mm and varies slowly depending on the horizontal position. Waviness generally results from vibrations between the work piece and the machining tool. The roughness is superimposed on the waviness.
- **Form** has the longest wavelength, similar to the wavelength of the object, so it must be removed to analyze the surface texture of the object of interest.

In order to examine the area of interest or the part of the surface to be visualized, it becomes important to filter out the finer irregularities (roughness) and noise (micro-roughness). The proper use of filtering permits the exploitation of areas of interest. There are two types of filtering: low-pass (waviness) filtering and high-pass (roughness) filtering. The most common form of filtering consists of separating data frequencies (or wavelengths) into two parts, the first one encompassing the long wavelengths or low frequencies (waviness), the other one encompassing the short wavelengths or high frequencies (roughness). The waviness and roughness phenomena are separated mathematically.

Algorithms include thresholds ("cut-off") on each end of the scale of feature wavelength to determine the amount to which roughness or waviness is removed from a set of data. The quality of the separation depends both on the type of filter and the cut-off value.

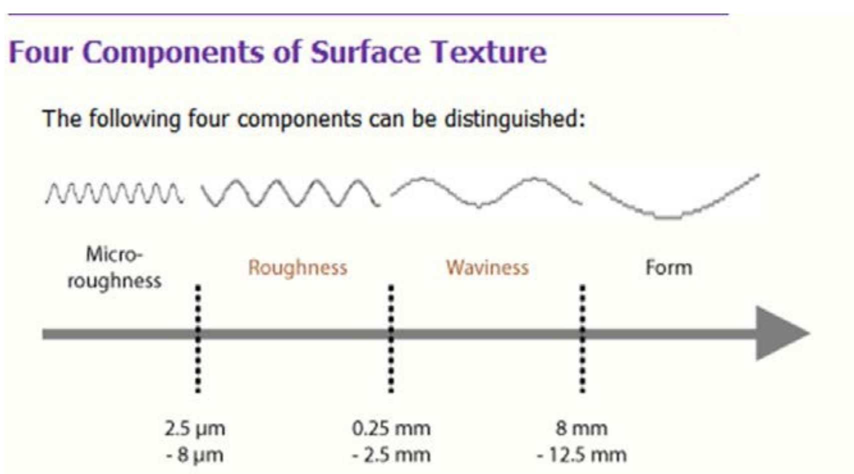


Figure 4. Four components of surface texture[12]

[12] Image from Mountains Map Help Topics.

The FBI utilizes a filter template that does all the calculations automatically after the examiner selects an area of interest. Once the examiner selects an area of interest, they then pick an area between the shoulders of a striae or impression to remove calculation artifacts that would arise from including the shoulder of the feature under examination. The goal is that the area should include 18 peaks and valleys in the calculation. After the noise (micro-roughness) is removed, then the form is removed through the use of a quadratic polynomial filter. In these images, one will observe a background "arc" that can be removed through this quadratic polynomial filter. Higher-order polynomials may be needed if the background arc is more complex. For example, a cubic spline may fit well to the background and permit improved background subtraction. After background correction, roughness is filtered. The examiner then applies the same filters to both bullets under comparison, the software picks the best fit for the two final surfaces, and the profile is extracted.

The state crime laboratory in Alabama conducted a series of studies of confocal microscopy using the Sensofar system with the intent to apply confocal microscopy in case work. They did not apply the technique in actual cases, but they did complete extensive studies demonstrating the effectiveness of confocal microscopy to firearm identification, including the following:

- 10 Ruger Barrel Study – 10 consecutively manufactured barrels - a little overlap because this study did not focus on the best area, and it was representative of only one land striae.
- Adjudicated casework – this was a much larger population with a slight overlap again indicative of only one land.
- Speed comparison (SRM 2460) – evaluated the difference between 1x scanning speed and 4x speed using a 20X objective; this results in the use of 4 micron slices and resulted in a good correlation factor (did 10 scans); up to 12x were attempted but higher speeds were not as effective; a 4x speed corresponds to a 2-minute scan.
- Angle difference: striae left – evaluation of how can positioning of bullet under the microscope affect the result. Even shifted, the results are still good. Software gives a good visual so you can also see that the shift has occurred, and they are not aligned. This is different than rotation, but even with rotations the results are good and the visual can show you that there is a shift, so it's readily visible that this is a bad scan.
- Also evaluated Nose Up and Nose Down shifts in angle rotation. Optical topography systems and evaluation algorithms are highly sensitive to sample tilt.
- Razor blade study with University of Central Oklahoma.

Additional studies include evaluating subclass carryover and the use of waviness as an exclusion characteristic to distinguish barrels. The data are preliminary right now, but show potential as exclusion characteristic instead of a match criteria.

3.6 Tool Mark Examination

The Ames Laboratory/Iowa State University (AL/ISU) research team currently utilizes an optical system manufactured by Alicona, which was selected over other instrumentation (e.g., stylus profilometer, laser confocal) because it allows a 90-degree scanning angle. The system is very portable and offers 4 µm resolution in x and y, and 1 µm resolution in the z direction. AL/ISU has conducted several studies, including comparison of striated tool marks created by screwdrivers and comparison of quasi-striated tool marks with pliers. The research team is also using computer simulation methods, including portable prototype development and virtual tool mark creation, to provide examiners with new capabilities. The portable prototype is based on the Alicona optical profilometer and will allow everything from simple visual comparisons to comparisons that include statistical information. Everything is written in open- source software and designed to be interoperable with researchers utilizing different analytical algorithms. The system looks like a simple plate-to-plate visual comparison window, but each side can be manipulated independently and different files can be loaded into each window. The topography is graphed, and correlation can be calculated between the two windows. The system also allows the user to load screwdriver tip files to generate a virtual mark that can be compared to the physical mark seen in the window. Then, the program can calculate a correlation value to statistically verify the match. These marks can be manipulated to change the angle so that they can be calculated over any angular range and statistically compared to the real mark.

In a blind angular prediction study involving 20 tool marks, all tips correctly related to their marks, with 14 out of 20 within 5 degrees, and the remainder within 10 degrees.

Future research at AL/ISU will focus on the ability to generate a better virtual mark, developing advanced algorithms to address ever more complex tool marks, and continuing to develop and enhance the prototype by adding additional algorithms and capabilities.

3.7 Effect of Instrumental Variability on Analysis Algorithms

Currently, ballistic imaging techniques are based on comparative methods. Moving forward, the goal will be quantification of the probative value of firearm and tool mark evidence with the ability to apply weight to the evidence, such as likelihood ratios (LRs) or other techniques such as reporting the p value.

Probative value depends on the different sources of variability, the set of circumstances considered, statistical method, and assumptions. The sources of variability include natural variability from the tool, (type brand make, manufacturer, manufacturing process, raw material), evolution of the tool over time, the use of the tool and the variability within the tool mark, as well as the analytical variability from the make and model of the analytical instrument, instrument settings, calibration, and the operator. Sources of variability can also include differences between statistical models (philosophy, design, and assumptions), hypotheses tested by the model, the sample size, and the computational methods used to estimate the parameters of the model. For example, is the statistical model a test against a family of firearms?; just Beretta firearms?; or a specific type or model of Beretta?, etc. These are just some of the many sources of variability that affect the weight of the evidence.

When making a characterization of tool mark features, different methods of acquisition may measure different features, including the width and angle of the land and grooves. This directly impacts the ability of the examiner to match striae or impressions and to make comparison decisions. In other words, as the examiner moves from the comparison microscope to data-based representations of optical topography, the types of distortion will also change. The comparison microscope introduces instrumental variation and subjective bias from the human examiner. The optical topographic microscope introduces distortions associated with the acquisition methods and post-processing. During acquisition post-processing, the aim is to measure raw data with the highest resolution possible, least instrumental interpretation, and the least distortion.

Hypotheses can be common source or specific source. In common source, multiple objects are tested to determine if the objects originate from a common source. The source may be unknown or unavailable for resampling. A typical example is a determination of whether or not two bullets found at one or more crime scenes were fired by the same firearm when you have two traces but no donor or source of the tool marks. This tests the level of similarity, but not the specificity of the features. Specific source analysis is a test of whether or not multiple objects originate from a single specified source. This would occur when the source is available for resampling or can be entirely characterized based on the available samples. A typical example of specific source is a test to determine whether or not a bullet originates from a seized firearm. This tests both the level of similarity and the specificity of those features. When a sample from the putative source entirely characterizes that source, both the common source and the specific source approaches result in the same probative value. An example of this would be the standard DNA profile; however, this is not true for most evidence types even when variability of the control samples is relatively reduced (e.g., fingerprints, firearms)

The aim of integrating statistics to quantify probative value is to quantify the “true probative value” of a trace. Assuming that we can completely define and characterize the different sources of variability of the evidence, there exists a likelihood ratio that can help one select between two alternative hypotheses. The question becomes: is there a metric for achieving the likelihood ratio? One can’t quantify the real likelihood ratio because of the computational assumptions that are necessary to simplify the likelihood ratio calculation, therefore the divergence is calculated to account for all the sources of variability. The likelihood ratio can be calculated based on the ratio of two likelihoods or by a plug-in estimation. The ratio of two likelihoods is based on the unrealistic implication that the likelihood structures are known under both alternatives. Plug-in estimates are the most common approach. The plug-in estimate approach makes assumptions on the stochastic process that gave rise to the objects and uses plug-in estimates for the parameters.

Another statistical approach to quantify the probative value of the evidence is the Bayes Factor. The Bayes Factor is a statistically rigorous summary of the value of the evidence, and is the gold standard if calculated properly. The convergence properties of typical computational methods used to calculate marginal likelihoods for the Bayes Factor are unstable unless there is a very large sample size. Plug-in likelihood ratios and Bayes Factors are very difficult and often impossible to calculate for complex evidence forms such as pattern evidence. The high dimensionality of the random vectors, heterogeneity of the random vectors, and unknown likelihood structures make it more difficult to calculate. These difficulties have led several researchers to attempt to simplify the problem by taking advantage of biometric technology to use scores to estimate the likelihood ratio. Different score-based likelihood ratios have been proposed in the literature, including the non-anchored approach, trace-anchored approach, print-anchored approach, and the asymmetrical approach. However, it is important to understand that score-based methods quantify the probative value of the score, not the probative value of the trace. The area of interest is the probative value of the trace and how the probative value of the score and the probative value of the trace relate to each other. The mathematical calculation is not everything; how that information is reported, documented, and verbally testified to in court is a major issue as well. Is it possible to test the convergence of different types of score-based LRs if there is trace and control material that are normally distributed with known distributions.

In ***Figures 5 and 6***, for various examples of both common and specific source LR analysis, the x-axis value is the real likelihood ratio and the y-axis is the score-based likelihood ratio. The line indicates the path the points should follow if the two likelihood ratios were the same. In some cases, these likelihood ratios will roughly correspond. In every case, there is random variation between the two calculations because information is lost when the original data sets are reduced to raw score numbers. The figures depict situations that vary between specific and common sources and difficulty of examination/comparison.

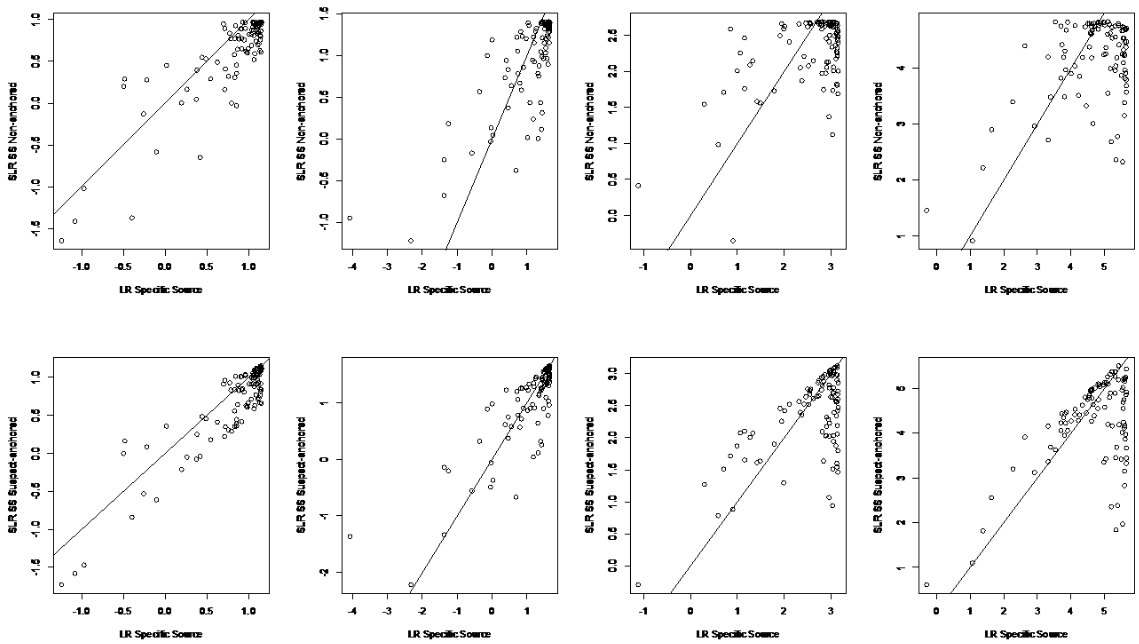


Figure 5. Specific source distribution

In short, it is important to measure and analyze raw data as much as possible to reduce the possibility of deviation. Staying in the feature space maintains the probative value of the evidence. Current NIJ-funded research is examining the effect of various parameters on the likelihood ratio calculation, including the effect of the physical system, instrumental/analytical systems, and statistical computation approach. Each type of variable affects the "ideal" likelihood ratio.

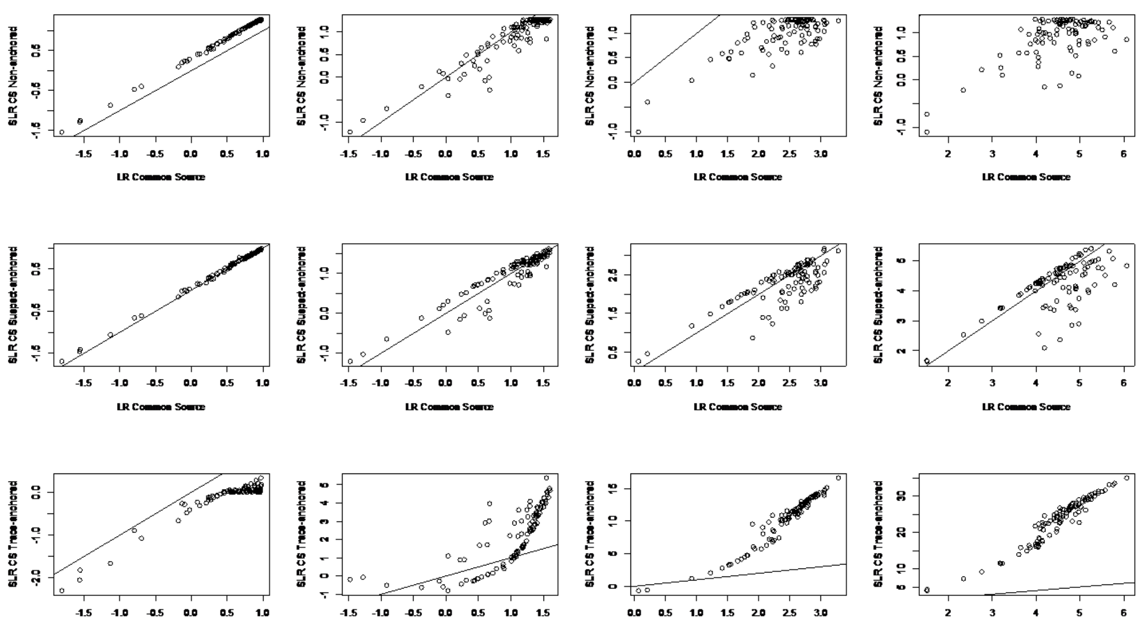


Figure 6. Common source distribution

The common source distribution in Figure 6 shows the difference between the suspect anchored to the trace anchored at the bottom where the scatter increases. The rarer the source becomes as it moves to the right, the more scatter that is present. If the source is rare the result can be an underestimation of the value of the evidence, and if it is common the value of the evidence can be overestimated.

Unfortunately, it is difficult to know which direction it is moving. Research has tested this empirically with handwriting and found that the probative value of the score is not the probative value of the trace. This can be conservative in some circumstances, but not in all, so research is needed to evaluate when this results in a conservative circumstance versus when the power of the evidence is overestimated. This is also complicated because the less sample that is present and available for testing, the more this relationship degrades.

While there may be a true likelihood ratio for any piece of evidence, that likelihood ratio is only an estimation, and different approximations of the likelihood ratio will provide different probative values. Current NIJ research projects address the convergence. Research has shown that score-based methods do not converge and cannot be used to quantify probative value. There are fields (e.g. AFIS scores) that are currently using score-based methods, although these methods should be limited in their application. However, these could be used in a binary decision engine or database search, as long as it is understood that the score does not imply the probative value of the source.

Other demonstrated methods exist that use scores to approximate likelihood ratios:

- Neumann et al., 2012, Quantifying the weight of fingerprint evidence, RSS Series A
- Guharay et al. 2012, Algorithm for spectroscopic data analysis and outlier detection, DTRA/NSF/NGA Algorithm Workshop, San Diego CA.
- Chumbley et al. 2013, Final Report for NIJ Award 2009-DN-R-119
- Saunders et al., 2015, Final report for NIJ Award 2009-DN-BX-K234.

Quantifying the weight of the evidence can only be done in the feature space of the evidence and needs to rely on untransformed raw data.

3.8 Considerations for the Field, Instrumentation: Types and Costs, Statistical Methods, Training, Integration with Crime Laboratory Operations

3.8.1 Panel Discussions

If a score-based method causes the distortions in the results shown, it is possible that any applied filter could distort the data, and would then effect the probative value of the evidence. A score-based method will distort the evidence and undermine the probative value. The score is any quantity that is calculated between two objects of interest using an algorithm that calculates relative value between the objects. Scores have been used to search and retrieve information from a database, but they cannot be used for probative value in court. Every filtering algorithm will introduce distortions, so it is important to understand the effect of the filtering on the power of any post-filter comparison. A filter may be acceptable if it has been demonstrated that it does not affect the comparison, particularly if it can be reduced to a likelihood ratio analysis.

Courtrooms and judges ask for and want statistics; numbers that do not currently exist. The Judge is interested in probative value of the trace and that value is not currently calculated; therefore, the way examiners testify has to be very specific. The way the field currently works versus how we want the field to work is very different. The answer may be to quantify the measurement process and guide the examination process but abandon the match score. An exception may be cold case investigation, which might need a score approach. In difficult cases, there may be more differences than there are similarities, and a match score would be unreflective of the value of the evidence. In other words, the examiner attempts to find consecutive matching striae, but it is expected that there is variation from one tool mark to the next, even with identical tool and evidence pairs. Therefore, there will be differences between matching evidence. A comparison is based on examination of positively-matching CMS or impressions without regard for non-matching marks. The field is based on the assumption that non-matching features are unimportant, though this assumption has not been fully tested through research. It is possible that the NIST reference collection could serve as a basis for this type of testing, as well as an examination of the statistical power of QCMS criteria for various types of situations (e.g., firearm manufacturing method, type of round, etc.).

During the discussion, two distinct views were expressed over how optical topography should be used. There are several ways to apply optical topography in current practice:

1. Replace current practice completely with optical topography, with comparison done in the computer. This option is not possible at this time, given the current state of technology, analysis, and understanding of methods. Practitioners rely on the comparison microscope, which has been proven an effective tool for making comparison judgments.

2. Optical topography may supplement current practice by giving the examiner a method to look more closely at striae and impressions that may be difficult to compare under the comparison microscope. This would allow for the examination of difficult comparisons, particularly in cases in which emerging manufacturing methods are making it more difficult to distinguish consecutively manufactured barrels. It would be critical to establish methods to ensure that practice conforms to standards concerning the use and incorporation of optical topographic examination in particular cases, especially because of the extended time it requires to apply the method currently.

3. Optical topography could be used as a confirmatory tool in which the examiner conducts a traditional analysis, finds a match or non-match, and then uses optical topography as a check on the result. The system could be used to produce a match score or likelihood ratio, and the examiner's comparison microscope-based decision would have no effect on the score. This approach would require validation and approval by ASCLD/LAB, at a minimum, though it is anticipated that the results would be used as a basis of court testimony. Again, the laboratory would need to run all data or a predefined subset in the optical topography system to provide an objective analysis of the examiner's decision. The examiner would testify that, based on his or her knowledge, training, skill, and ability, he or she believes that this gun fired this bullet, and would present pictures to prove that and indicate that an independent instrument confirmed his or her conclusion.

Optical topography could be applied on an experimental basis only so that methods and experience could be developed to inform one or more of the above scenarios.

Some members of the working group expressed concern over any process of examining a specific source and then using a common source score to make a decision, and then using a statistic to confirm. This process might boil down to having the examiner make a decision and then the information would be put it in the machine to get a number to support decisions that were considered favorable. They felt that the examiner must rely on the tool to make a decision instead of making a decision then using the tool. Theoretically, this concern could be alleviated if there were specific conditions under which optical topography and statistical matching would be applied.

How does DNA differ? DNA relies on human beings to do the mixture interpretation and then the numbers go into a machine but the quantitative part is much simpler because a 5 allele is a 5 allele. DNA can stay in the space of individual characteristics without worry about class characteristics in the same way.

Mathematical algorithms would give the examiner numbers to show the uncertainty of their testimony. That's good from a research view but not practical in a crime laboratory setting. The biggest issue may be when an examiner comes to an inconclusive decision, but the machine finds a match. Perhaps the solution is to use error rates as opposed to a likelihood ratio. This would be changing the score to be a binary score between x and y so it's above a threshold or not. Once you use the score to make a decision, one would not report the score but only the error rate of the score.

From the perspective of the practitioner what's in it for the examiner?

1. Definitive answers in inconclusive comparisons
2. Get ID comparison results in a more objective way – we work with objective information but the interpretation is a subjective process based on objective criteria. It is the visual comparative examination of topographical features of two different tool marks. The use of CMS is an important element. Tool mark examination is a skill—it depends on the examiner's cognitive ability and training to build awareness of uniqueness—and science—the validated premise that tool marks can have unique, reproducible striae patterns that can, in most instances, be identified to the tool that created them. This is where we are now, but all examiners would recommend a way to make more sound identifications.
3. Help increase the percentage of correct identifications. There are false positive for striated and impressed tool marks, with a variable error rate based on examiner proficiency. Usually, errors arise from poor assumptions about known, non-matching commonalities. Thus, optical topography may provide an objective basis to improve the rate of correct identifications. Further, optical topography is a powerful research tool to improve the field's understanding of class characteristics and commonalities that are not related to firearm-to-firearm variability.

The field needs to improve proficiency testing. Retention of information is an issue. At the end of training, an examiner may be proficient but 3 months later would they start making mistakes again. AFTE standards have only recently been improved to ensure that those who fail proficiency testing are removed from the list of certified examiners.

Improved proficiency testing may be based on extracting from examiners what makes them say that there is a match. We could then use that to determine a metric to apply on a comparison to determine the difficulty of that comparison. Where is the line where an examiner should always be able to get the right answer? The solution might be to use replicas like they do in Europe. They represent the full range from easy to very difficult. However, CTS error rate would climb and AFTE doesn't want that. Europe has higher error rates. Europe stumbled when subclasses where introduced. They stumbled but they learned. CTS provides a proficiency test with some blind aspects but they are not truly blind tests. Anyone can take a CTS exam, regardless of training or experience. Certification of examiners is better now because you have to notify AFTE if you fail a proficiency test and certification can be suspended immediately. Ideally, proficiency tests would reflect validation testing that meets ASCLD-LAB and ISO requirements.

What gives good firearm tool marks? The best tool marks from the range of guns in the middle range of quality. High quality guns are difficult to distinguish because of subclass characteristics and poorly made guns are less likely to leave reproducible marks. It is unknown how optical topography may address these gaps. The FBI is collecting test fires from all the firearms that they have. In addition, NIST is also adding low quality test shots, firing shots with no maintenance (dirty guns) and then cleaning them and test firing. All of this information will be represented in the NIST reference collection.

3.9 State of the Industry: Overview of Available Instrumentation

Presentations were given at the meeting on the available instrumentation by for firearm and tool mark identification. The following bullets provide links to those presentations for review:

- Gelsight TopMatch–GS 3D
- Alicona InfiniteFocusG5
- Leica (Multiple Models)
- IBISTrax HD3D.

3.10 Challenge of Subclass Characteristics

During the working group meeting, John Murdoch presented an example of the extreme difficulties that can arise due to subclass characteristics based on an examination of consecutively manufactured Ruger P95 and LC9 barrels from the Ruger plant in Prescott, AZ. Initial, blind examination of test fires from these barrels demonstrated that traditional comparison-microscope-based methods will not reliably individualize in certain cases, as there were no discernable differences among three of the barrels.

Subclass characteristics are manufactured tool marks that repeat virtually unchanged on a series of consecutively produced items that have been made by the same tool. The definition of subclass characteristics, as outlines in the AFTE glossary, is as follows: “Discernible surface features of a manufactured item that are more restrictive than class characteristics in that they:

1. Are produced incidental to manufacturer
2. Relate to a smaller group source (a subset of a class to which they belong) and
3. Can arise from a source that changes over time.”

When these tool marks are present on or near the working surfaces of tools, the tool marks they produce can be mistaken for individual working surface features. Therefore, subclass influences must be recognized to ensure that the tool marks they produce will not be used for identification purposes. It is important to note that although subclass tool marks maybe present near the working surface of the tool, they may either have no influence on the individuality of tool marks made by this working surface or edge because of their position or the manner in which the tool is used. Subclass characteristics are not always present on manufactured tool working surfaces.

Subclass features are those that carry on through the entire length of the rifled bore and may be characteristic of other members of the subclass, such as consecutively manufactured barrels. In order to evaluate subclass, the examiner should make a cast of the barrel bore and compare breech end to the muzzle end to find the tool marks that carry all the way through

We do not know how many sequentially manufactured guns can carry the same subclass. Sometimes, tool marks are present at the crown that can override the subclass characteristics. The best marks for identification are often on the heel of copper-jacketed bullets, in the land impressions. NIST does have some bullets with extensive subclass carryover and will begin working on them and evaluating if filters could be used to eliminate the subclass features. If a gun is fired with a lot of lead bullets, the rifled bore can lead up. Typically when a laboratory receives a gun, they fire it as they receive it. Then, the laboratory cleans the gun and fire progressive test shots after cleaning to see if there is any change in the quality or quantity of the firearm-produced individual tool marks.

In any case, comparison microscopes are very fast tools for sifting through data. Optical topography will not fill that need soon. The FBI feels that confocal microscopy could fill a need as a blind verification tool. Ultimately, optical topography could be a comparison tool when combined with an algorithmic approach.

Other issues include: interoperability, calibration, and interlaboratory comparison. Significant work is required to establish the “math” of optical topography, including error rates, filtering, and other calculations. There is insufficient information that has been established to know how to make an instrument fail and how would know when it fails. In other words, the examiner needs an objective approach to maintenance and calibration issues. For casings, correlation algorithms are still under development. Finally, work flow must be established.

3.11 Research Needs in the Application of Confocal Microscopy to Ballistic Imaging

3.11.1 Research to Date

Some instances of early firearm and tool mark research is shown in Table 2. Sources of more recent research[13] and a description of what they reviewed is provided in Table 3.

Table 2. Early Firearm and Tool Mark Research

Type	Event
Earliest Topographical Analysis and Comparison of Bullet Stria	**1958** – John E. Davis Oakland Police Department Crime Laboratory, An Introduction to Tool Marks, Firearms and the Striagraph.
Early Computer Based Bullet Comparisons:	**1978** – Geoffrey Garner, Scanning Electron Microscope used to make surface profile measurements on 13 bullets fired from the three 0.38 special caliber revolvers and “signatures” were mathematically compared. **1988 to 1993** – Tsuneo Uchiyama, Bar code like lines were calculated from the image and can be counted by the computer **1998** – De Kinder, et al. Non-optical method that measured the surface topography of the striae in land impressions. The instrument used was an infrared laser surface topography scanner. The data was captured on a digital sensor. This non-optical approach is an early effort to eliminate problems of a surface reflection, surface curvature and illumination differences.

(continued)

[13] Additional reference to research related to forensic firearms and tool mark identification can be found in AFTE Committee for the Advancement of the Science of Firearm and Tool mark Identification. (2011, June 14). *AFTE Response to the 25 Questions related to firearms and tool mark examinations promulgated by the RDT&E IWG.* Retrieved March 27, 2015, from www.AFTE.org: http://afte.org/downloads/RDT&E%20IWG%2025%20Questions%206.14.11%20-%20AFTE%20Response%20w%20cov%20let.pdf

Table 2. Early Firearm and Tool Mark Research (continued)

Type	Event
Faden, Chumbley et al., 2007	Striated tool marks produced by 44 sequentially produced screwdrivers to examine the profiles that were measured by a stylus profilometer. A mathematical algorithm was used in the comparisons between known match and known non-matched tool marks, including the variability of tip angles.
Bachrach, Koons, et al., 2010	Possibly the first study that used 3D surface profile measurement and comparison of striated tool marks produced by tools under different variable of angle, pressure, and materials. The acquisitions of the surface measurements was through the use of a "non- contact" method; confocal microscopy. The correlation to measure similarity of the profiles used open sourced statistical formulae.
Song, Vorburger, et al., 2012	Application of Cross Correlation Function (CCF) in NIST standard bullet comparison. Describes the production and validation of the Standard Reference Bullet (SRM) using surface "signature profiles" from fired bullets for the production of a virtual profile set. These profiles were used in the production of the SRM bullets. The method and formula used in the qualification of the bullets are the basis for a prototype objective mathematical comparison system for actual fired bullet tool marks.
Weller, Zheng, et al., 2012	Breech face marks on cartridge cases fired from 10 consecutively manufactured slides were measured by confocal microscopy. A total of 8010 comparisons using a 3D arial cross-correlation statistical algorithm which mathematically measures the similarity of know matching and known non-matching cartridge cases. There was no overlapping of scores from the matching and non-matching cases.
Petraco et al., 2012	The research report describes results of using confocal microscopy measurement and methods of mathematical based computer comparisons of striated and impressed tool mark surfaces. Statistical analyses of the methods were compared, and a web-based database was developed for interaction with other researchers.
Petraco et al., 2013	The authors report of the examination and comparison of 3D surface topography measurements from striated tool marks from screwdrivers and cartridge case firing pin aperture shear marks. The measurements were performed using confocal microscopy and the comparisons were performed by multivariate statistical methods; principle component analysis and support vector machine. Using these methods, an estimation of error was determined.
Chu, Thompson, et al., 2013	The concept of consecutive matching striae (CMS) numerical criteria was used as the basis of a matching model derived from tool mark striae from fired bullet land engravings. The bullets were fired from 10 consecutively rifled 9mm caliber barrels, and 15 unknown bullets were compared to the "knowns". The surface features were measured by confocal microscopy, uninformative features were automatically masked, and the remaining signature detail was compared between all the bullet land combinations totaling almost 13,000 comparisons. Mathematically, the formula is comparing surface profiles between two bullets and measuring the degree of similarity in the "match" position.

3.11.2 Research Needs

During its discussions, the Forensic Optical Topography Working Group identified a number of research needs in which optical topography may be developed or relevant to broader issues, including the following:

1. The field needs an improved understanding of the incidence of class and subclass characteristics in tool mark impressions. The relationship of manufacturing method, history of the tool, materials, type of striae or impression, and other variables can influence the extent to which the examiner can individualize a tool mark impression. Optical topography could be used to produce a wider understanding of the incidence of these characteristics. The NIST database will be a valuable resource in this regard.

2. Examiners rely on assumptions about the power of CMS that should be subject to review with regard to statistical power, relevance within classes and subclasses, and information available in optical topographic data. The examiner should have access to information regarding the statistical power of various numbers of CMS to inform their decision-making.
3. Statistical and data science experts need to provide a framework for the analysis of optical topographic data to provide a basis for its application in the crime laboratory. This should include an understanding of filtering and data analysis algorithms, statistical characterization, error analysis, instrumental analysis, and related issues.
4. Proficiency testing for tool mark examination should be improved to reflect various levels of expertise, including the ability to account for class and subclass characteristics and other confounding variables. Recruitment, training, and proficiency testing standards should be updated to reflect current research concerning tool mark examination and human factors.
5. The Forensic Optical Topography Working Group should develop guidance for the field with respect to the use of optical topography as a confirmatory tool in the forensic laboratory. This guidance should include standards concerning when optical topography might be warranted, the use of optical topography to provide a statistical basis for an examination; the use of optical topography when traditional comparison microscopy cannot provide a match; the standards under which optical topography data is acceptable; guidelines for the procurement and deployment of systems; training; and protocols for the examination process. The working group will establish baseline considerations in this regard during a hands-on workshop at the FBI Laboratory.

3.11.3 Other Discussions

Other topics discussed at the Forensic Optical Topography Working Group Meeting included the following:

- Future work with firearm and tool mark identification will have some challenges (e.g., factory-produced [pre-fired] tool marks on cartridges).
- Consecutively manufactured guns are typical in police officer shootings and military because these organizations purchase large batches of firearms at the same time.
- What needs to be done to implement new methods? Validation procedures and ISO requirements exist, so research needs to be tailored to answer those requirements.
- An interesting type of data would be to tweak IBIS to see how powerful or discriminatory subclass characteristics can be in matching, similar to what is done with DNA.
- Base rates per municipality based on how manufacturers are shipping guns will be needed.

4. SUMMARY

This report provides a summary of generalized comments and opinions from a diversified group of researchers, examiners, commercial providers, and technology experts concerning the current status of forensic optical topography. The Forensic Optical Topography Working Group covered a diverse set of considerations, including the status of tool mark examination, potential impact of optical topography on the field, research needs, considerations in analysis, development of standards and reference data collections, proficiency testing, options in optical topographic technology, data interoperability, and related items. The working group will continue to develop protocols for the application of optical topography as a confirmation tool to supplement current practice, with the understanding that much work needs to be done to establish optical topography as a primary instrument for tool mark analysis.

BIBLIOGRAPHY

AFTE Committee for the Advancement of the Science of Firearm and Toolmark Identification. (2011, June 14). *AFTE Response to the 25 Questions related to firearms and toolmark examinations promulgated by the RDT&E IWG.* Retrieved March 27, 2015, from www.AFTE.org: http://afte.org/downloads/RDT&E%20IWG%2025%20Questions%206.14.11%20-%20AFTE%20Response%20w%20cov%20let.pdf

Association of Firearm and Toolmark Examiners. (2011, Fall Volume 43 Number 4). Theory of Identification as it Relates to Toolmarks: Revised. *AFTE Journal*, p. 287.

Braga, A. (2011). Reconsidering the Ballistic Imaging of Crime Bullets in Gun Law Enforcement Operations. *Forensic Science Policy & Management: An International Journal*, 2:3, 105-117. http://www.tandfonline.com/doi/abs/10.1080/19409044.2011.613444#.VS-zFC6gs80

Braga, A., & Pierce, G. L. (2004, July 49:4). Linking Crime Guns: The Impact of Ballistics Imaging Technology on the Productivity of the Boston Police Department's Ballistics Unit. *Journal of Forensic Science*. http://www.researchgate.net/publication/8394737_Linking_crime_guns_the_impact_of_ballistics_imaging_technology_on_the_productivity_of_the_Boston_Police_Department%27s_Ballistics_Unit

Thompson, R. (2010). *Firearm Identification In The Forensic Science Laborataory.* Alexandria, VA: National District Attorneys Association.

X Zheng, J. S. (2014). Applications of surface metrology in firearm identification. *Surface Topography: Metrology and Properties*. http://iopscience.iop.org/2051-672X/2/1/014012/

Forensic Optical Topography Working Group Attendees

Gerald Laporte
National Institute of Justice
Gerald.Laporte@ojp.usdoj.gov

Danielle McLoud-Henning National Institute of Justice
Danielle.Mcleod-Henning@usdoj.gov

Nicole McCleary
RTI International
nmccleary@rti.org

Sue Ballou
National Institute of Standards and Technology
susan.ballou@nist.gov

Alan Zheng
National Institute of Standards and Technology
alan.zheng@nist.gov

Hans Soons
National Institute of Standards and Technology
soons@nist.gov

James Yen
National Institute of Standards and Technology
james.yen@nist.gov

John Lu
National Institute of Standards and Technology
John.lu@nist.gov

Derrick McClarin
Federal Bureau of Investigation
derrick.mcclarin@ic.fbi.gov

John Murdock
Contra Costa County (California) Office of the Sheriff, Forensic Services Division
jmurdock@so.cccounty.us

Dan Gunnell
Illinios State Police, Joliet Forensic Science Laboratory
gunneda@isp.state.il.us

Martin Ols
Bureau of Alcohol, Tobacco, Firearms and Explosives
martin.ols@atf.gov

Chris Saunders
South Dakota State University
christopher.saunders@sdstate.edu

Martin Baiker
Netherlands Forensic Institute
Baiker, Martin
m.baiker@nfi.minvenj.nl

John Brad Etter Alicona
brad.etter@alicona.com

Gretchen Falter Leica Microsystems
Gretchen.falter@leica-microsystems.com

Gregory Dutton
National Institute of Justice
Gregory.J.Dutton@usdoj.gov

John Morgan RTI International
jmorgan@rti.org

Jeri Ropero-Miller RTI International
jerimiller@rti.org

Robert Thompson
National Institute of Standards and Technology
robert.m.thompson@nist.gov

Brian Reneger
National Institute of Standards and Technology
thomas.renegar@nist.gov

Ted Vorburger
National Institute of Standards and Technology
theodore.vorburger@nist.gov

John Butler
National Institute of Standards and Technology
John.butler@nist.gov

John Song
National Institute of Standards and Technology
jun-feng.song@nist.gov

Eric Smith
Federal Bureau of Investigation
Erich.smith@ic.fbi.gov.

Eric Collins
Contra Costa County (California) Office of the Sheriff, Forensic Services Division
ecoll@so.cccounty.us

Ronald Nichols
Bureau of Alcohol, Tobacco, Firearms and Explosives
ronald.g.nichols@usdoj.gov

Chad Macziewski Iowa State
chadm@iastate.edu

Cedric Neuman
South Dakota State University
cedric.neumann@sdstate.edu

Jan de Kinder
Nationaal Instituut voor Criminalistiek en Criminologie, Belgium
Jan.DeKinder@just.fgov.be

Tom Calahan Zeiss
tom.calahan@zeiss.com

Serge Lévesque
FTI, Inc. (NIBIN/IBIS system)
Serge.Levesque@ultra-ft.com