Document Title:	Improve the NIBIN System by Providing Examiners a Capability to Match Infrared Images of Firing Pin Impressions and Deformed Bullets as well as Accurate Large Database Searches
Author:	Francine Prokoski, Ph.D.
Document No.:	237967
Date Received:	March 2012
Award Number:	2009-DN-BX-K262

Improve the NIBIN System by Providing Examiners a Capability to Match Infrared Images of Firing Pin Impressions and Deformed Bullets as Well as Accurate Large Database Searches

DOJ Grant # 2009-DN-BX-K262

Author:
Dr. Francine Prokoski
SED Technology LLC

Abstract

This grant effort was focused on improving the National Integrated Ballistic Information Network (NIBIN) System by providing examiners with advanced capabilities to:

- identify cartridge casings by using infrared (IR) images of firing pin impressions (FPI),
- identify bullets by using IR images of land impressions
- perform accurate high-speed search of a large database to identify fired cartridge casings.

This report summarizes successful accomplishment of those goals. Reviewers of the draft version of this report found the research effort provided a compelling proof of concept demonstration of the effectiveness of IR imaging for toolmark identification, and made a very significant contribution by showing an underlying scientific basis that can be quantified and statistically measured. The use of IR imaging offers reliable and significant improvement in case linkage supporting forensic firearms examinations.

The project was divided into three separate but related efforts:

- Determine the persistency of firing pin impressions (FPI) by collecting 1,000 cartridge cases fired in each of 8 different firearms
- Conduct a Proof of Principal using IR to accurately match fired bullets bearing minimal damage collected from sample firearms
- Build a large database of cartridge cases and be able to locate siblings. This database contains multiple fired cartridge cases from an unknown number of firearms

Historically, examiners have attempted to link shooting incidents using evidence from a current shooting to that from a past shooting. The volume of evidence forced the introduction of technology. By the late 1980s computer-based case linkage systems became a reality. These advances improved evidentiary analysis and fostered the production of first generation ballistics-type workstations for case linkage and the networking of visible light systems serving the majority of forensic laboratories. The predominant system today is the Integrated Ballistics Identification System (IBIS), networked under the umbrella of the National Integrated Ballistic Information Network (NIBIN).

Under this grant, two-dimensional infrared (2D/IR) and three-dimensional infrared (3D/IR) analysis of firing pin impressions (FPI) was performed. The persistence of FPIs, imaged with IR cameras, demonstrated IR imaging provides details of deep firing pin impressions that cannot be seen using visible light. Use of IR imaging demonstrated firing pin impressions remain relatively constant for a selection of firearms and ammunition during the sequential firing of 1000 rounds of ammunition. 3D/IR analysis of cartridge cases from certain firearms may also provide a reliable determination of firing order among sibling casings collected from a shooting scene. That could be a significant aid to crime reconstruction.

IR imaging demonstrated a high degree of effectiveness for the comparison of fired bullets (pristine and minimally damaged) to identify siblings fired from the same firearm. Additional comparison techniques included three-dimensional pattern recognition and feature metrics. Techniques for comparing bullets proved similar, producing a high degree of confidence and

significantly reducing the number of false positives compared to reported performance of current NIBIN workstations.

Improved performance was achieved by the use of IR imaging and an advanced pattern matching engine called FlashCorrelation®. This approach to pattern matching results in rapid locally-controlled image correlations based on the superior detail in IR images and the application of FlashCorrelation®. The ability to image, store, and accurately identify sibling cartridge cases from a large database of infrared images was clearly demonstrated. When a sibling was present in the database of 2000 Glock 9mm casings, its Match Value was ranked #1 for 99.5% of tests and was in the top three positions 100% of the time.

A database of fired cartridge cases was expanded to determine the scalability of the infrared image comparison techniques. A statistically significant sampling of infrared images was utilized for comparison purposes. Fired cartridge cases were primarily collected from common handguns: Glock and Hi-Point. The Hi-Point company provided fired cartridge cases and bullets from more than 100 handguns. Personal contacts allowed access to several hundred Glock cartridge cases that were entered into an established database involving many Glock firearms with known linkage to other fired cartridge cases. AR-15 and SKS rifles were also selected for examination and analysis.

Table of Contents

1. Executive Summary:

1.1. Background: The National Academy of Sciences (NAS) 2009 report, "Strengthening Forensic Science in the United States" outlines the need to improve the scientific foundations of the forensic disciplines dependent on qualitative analyses and interpretation of observed patterns. The current SED project seeks to improve measurement validity, accuracy, and reliability of firearms/toolmark identification through the scientifically rigorous application of thermal infrared imaging.

1.2. Nature of Spectral Emissivity Relative to Toolmarks: Every object with a temperature above absolute zero (-273^{o}C) continuously and spontaneously radiates thermal energy. The amount radiating from each particular point on a surface is determined by material composition, temperature, and other properties. Those factors combine to determine the local emissivity, a measure of how efficiently the surface radiates heat. Infrared cameras that are sensitive at thermal wavelengths produce images that contain evidence of all emissivity variations on surfaces within their field of view. In most applications, emissivity variations must be calibrated out of the sensor data. In other applications, mapping the emissivity variations is the objective, as it is for firearms/toolmarks identification.

Action of a tool against a substrate, leaving a toolmark pattern, creates variations in the surface texture which distort surface heat emissions. Local surface emissivity is a measure of the efficiency with which a portion of the surface radiates heat. Variations in emissivity can occur at each pixel in an infrared image, providing a potentially very high resolution toolmark replicator. Given the same tool acting in an identical manner against the same or equivalent substrate, resulting marks would be expected to appear more similar when imaged with an infrared thermal (IR) imager than with a visible light (VL) camera; IR imagers are essentially insensitive to lighting variations that can obscure significant details in VL images. The result is higher correlation between siblings IR images.

At the microscopic level, reproducibility of firearms-induced toolmarks on ammunition components may be limited by variations in the molecular structure of the material in each individual firearm and ammunition component, temperature changes induced in the ammunition by the action of the firearm, changes in vibration of the firearm during the firing of successive cartridges, manufacturing variations in ammunition shape and material composition and other physical inconsistencies due to inexact manufacturing procedures combined with irreproducible scenarios. Continuous changes occur each time a weapon is fired due to movement of internal components, accumulated wear, buildup of residue, or effects of cleaning. These minute changes to the firearm with each use, coupled with inconsistencies in ammunition, target conditions, and human performance, can cause broad differences in appearance of fired cartridge cases and bullets.

Infrared sensors have advantages as toolmark imagers; they are insensitive to visible light radiation and do not produce illumination-induced artifacts such as glare and shadow

effects. This also removes image variations due to lighting adjustments by human technicians, which was found to occur in 38% of NIBIN imagings and is considered a significant source of variations that produce match errors.

The science and engineering behind infrared imaging is well-established, with textbooks and peer-reviewed publications providing a scientific basis for its application to toolmark identification. The science of infrared metrology provides a scientific foundation for toolmark analysis and identification. Toolmark elements can be related to specific characteristics of ammunition and firearm components and their movements. Toolmark details can be predicted to the extent specific information is known about each ammunition and firearm component [composition, surface finish, shape and size], each firearm action [force, angle, movement, constraints, and timing], and its firing history. Toolmark changes as a result of firearm or ammunition component changes or firing can likewise be predicted. In reality, complete and perfect information on firearm and ammunition is not known; modeling generation of a specific toolmark must provide for variations in component parameters. Large scale investigations facilitated by automated emissivity map collection provide opportunities for statistical studies on individual, sub-class, and class characteristics. Use of Receiver Operating Characteristics (ROC) analysis of computerized toolmark matching quantifies its contribution to examiner decisions.

1.3. Analytical Proofs of Concept: Project investigations included performing reliability analysis metrology using 2D/IR imaging of ammunition components and 3D/IR models; detection and extraction of toolmark features from each; calculation of quantitative feature characteristics; comparison of the feature sets from two components; and matching against large databases of feature sets. Three areas of demonstration and evaluation were in the current SED effort which followed an earlier project directed toward showing the conceptual feasibility of using three-dimensional infrared imaging:

- Investigate the level of detail and persistence of firing pin impression features and characteristics produced by new firearms through 1000 consecutive firings with a single type ammunition. Eight firearms were investigated, including various makes, models, actions, and rifling types. General conclusions were that:
  - FPI details are more persistent and reliable than breechface marks when imaged by infrared cameras – provide a better foundation for identification of fired cartridge cases
  - FPI details as imaged by infrared cameras provide sharply defined indication of characteristics – firing pin rotation and shape that can aid in determining possible make and model of firearms used at a crime scene
  - The reliability and persistence of FPI and breechface toolmarks in infrared images introduces new class and individual characteristics for comparison

- Demonstrate methods to use infrared imaging for capturing and comparing toolmarks from pristine and minimally damaged bullets. Techniques demonstrated could be applied to each of the six handguns considered; the two rifles require the design of custom mechanisms to hold the fired bullets. General conclusions were:

  - FlashCorrelation® processing of infrared image frame sequences captured as a bullet was rotated before the infrared camera in every case automatically detected each land, computed the angle of twist, and produced a 3D surface model that was correctly aligned and matched to other bullets fired from the same weapon
  - Fired bullets with undeformed bases that include a concavity can be securely held in a proper position throughout their image capture by use of minimal vacuum
  - Extended focus bullet images were produced from multiple IR imaging rotations performed at minutely different focus distances.

- Compile a large database of infrared emissivity mappings from fired cartridge cases with known origins. Demonstrate methods for assessing the reliability of the mapping process, predicting accuracy of identifications, and estimating the probability a given level of similarity between two cartridge cases fired in different firearms could occur by chance. Rates: True Match, True Non-Match, False Match, False Non-Match (ROC) analysis was applied to a toolmark identification decision system that made quantitative pair-wise comparisons of infrared emissions maps from more than 2000 cartridge cases fired in Glock pistols.

ROC analysis provides a systematic method for quantitatively evaluating the performance of a decision-making system without knowing the decision algorithms involved. It provides methods for segmenting a decision process into two or more components, and separately analyzing each. For our application, we separate toolmark identification into computerized and manual processes, where NIBIN is the computerized screening tool and the Examiner performs the final manual decision process. ROC analysis is being applied to only the computerized process at this time.

Under the current project SED presented a systematic approach to establishing quantitative probabilities of identification that support Association of Firearm and Tool Mark Examiners (AFTE) criteria and utilize the knowledge accumulated by experienced examiners.

Criteria for fired cartridge case and bullet identification (AFTE Criteria for Identification) are based on manual detection of similar patterns in two items viewed under a comparison microscope aided by manual manipulation of incident lighting. Modern manufacturing techniques commonly eliminate hand finishing steps that produced individual characteristics previously used for visual identification and automated identification based on visible light imaging (e.g. fired Glock bullets). That suggests consideration of new toolmark comparison screening methods (such as IR imaging) that do not replicate human viewing of a magnified visible light image.

1.4. Quantified Performance Measures

Infrared imaging provides quantitative Match Value correlation values from comparisons of casings and bullets. ROC analysis generates an optimal threshold value for the binary decision as to whether a pair of images constitutes a match. However, to compare match performances of MTW and NIBIN requires use of rank ordering.

Performance Measures used included both approaches:

For Rank Ordering:

- **P[Sibling=#1]**: Percentage of matches in which a Sibling Image is ranked #1; normally applied to full comparison of every image in one database against every image in another.
- **Excess Search Count:** Number of non-Sibling images with higher Match Value than a Sibling Image
- **Average Search Count:** Average number of images reviewed by Rank order, to find sibling

For Threshold Comparison:

- **Threshold:** Selected Match Value from 0 to 1 used to designate Candidate Siblings
- **Accuracy:** For a given Threshold, percent of database images correctly labeled Sibling or Non-Sibling when above or below Threshold
- **ROC Curve:** Plot showing dependence between two of the four Match Outcome Rates (True Positive, True Negative, False Positive, False Negative)

Toolmark comparisons involve individual curvilinear features that can be measured and areas of impression whose features are indistinct but for which some measurements can be made; maximum depth, swells where primer material has been displaced, dragging of firing pin and other effects. Depth profiles across the primer of a cartridge case base can extract quantitative toolmark features. Comparison of the resulting waveforms can utilize standard signal processing software, but its use must be tailored to the particulars of the firearm and ammunition used.

Recent articles recommend toolmark identification become like DNA analysis; be based on established science to compare evidence and use statistical analyses from large populations to calculate a quantitative match probability. There are key differences between the disciplines: The makeup of DNA is not changing but new makes and models of firearms, and new manufacturing techniques must be incorporated into the choice of identifying features and the statistical distributions of their characteristics, toolmark examiners testify about their manipulation and viewing of the evidence items under a microscope; DNA experts don't visually compare DNA samples - they trust accepted analytical instruments; peer reviewed publications, competition by companies performing DNA analysis, and funding by government agencies fueled advancement and acceptance of DNA analysis. It is the gold standard of identification methods. This contrasts with the sole ballistic toolmark comparison system technology promoted by the Department of Justice (DOJ) for the NIBIN system the past 12 years.

The Mikos Forensic Toolmark Workstation (MTW) used in the current project performs automated imaging and comparison of 3D toolmarked items including cartridge cases and pristine test fired and damaged bullets. Innovations incorporated into the MTW:

- Use of thermal imaging to eliminate lighting-induced artifacts and eliminate the need for subjective adjustments to lighting position and intensity.
- Use of CNC (computerized numerical control) positioning of the item to be imaged relative to the imaging sensor(s).
- Method for producing a 3D surface model from range gated sequences of 2D image frames from a fixed focus camera.
- Detection of toolmarks as abrupt changes in spectral emissivity which appear in the thermal image as abrupt changes in apparent temperature.
- Exploitation of the very shallow depth of focus of IR imagers with microscope lenses to create precise inverted 3D models such as of firing pin impressions.
- Method for creating extended focus 2D images containing continuous toolmarks of varying depth.
- Various methods for comparing resulting 2D images and 3D models to find high-probability matches.
- Method for cross-spectral matching of infrared-derived emissivity maps against legacy visible light images in the NIBIN databases.

Performance testing of the MTW using ROC analysis produced identification accuracies consistently above 0.995 and cumulative match characteristics placing true siblings in #1 rank position 99.8% of the time in multiple tests involving 800 cartridge case images from more than 350 Glock 9mm firearms. Results attest to the efficacy of the particular image processing algorithms used with the IR images collected. They do not necessarily predict the reliability of IR imaging of bullets, or of cartridge cases from other firearm and ammunition types.

MTW incorporates computerized numerical control (CNC) for positioning the item under test and the IR imager, allowing a precise sequence of 2D/IR slices to be collected, generating a high resolution 3D/IR surface model in less than 30 seconds. Although IR imager and CNC desktop controller technology is rapidly advancing, the reliability of 3D/IR models produced using the current commercially available off-the-shelf (COTS) subsystems is an important benchmark to be determined.

Infrared-based identification demonstrated advantages over other imaging and matching technologies. Fully automated scans of cartridge cases require less than 10 seconds. Bullet scans require 30 seconds or less depending on size and degree of damage. The only human function is inserting and removing items to be scanned. Evaluation testing with 600 cartridge cases fired in Glock pistols showed excellent performance for imaging and matching capability. ROC analysis produced accuracy and reliability measurements, and calculated the number of rank ordered cartridge cases an examiner would need to consider before finding a match, under various policies, and with 0 to 8 siblings in the database. Corresponding measurements were produced for systems that used visible light imaging, and for those that compared infrared images of a target toolmark against legacy visible light image databases. In all tests, IR-IR matching was most accurate, with IR-VL matching being more accurate than VL-VL.

1.5. Implications for US Criminal Justice Policy and Practice: The well-established science of infrared signature analysis potentially offers the scientific foundation sought for toolmark analysis and identification. If the hypotheses of this project are proven, each element of a toolmark can be related to specific characteristics of ammunition and firearm components. Toolmark details can be predicted to the extent specific information is known about the ammunition and firearm. Toolmark changes can likewise be predicted. Since complete and perfect information on firearm and ammunition is not known, modeling must provide for variations in component parameters. Large scale investigations facilitated by automated emissivity map collection provide opportunities for statistical studies on individual, sub-class, and class characteristics. Use of ROC analysis together with the Match Value, produces a probability that two components have a common origin. That would move toolmark identification in the direction of DNA analysis.

Published studies have documented significant false negative errors generated by NIBIN [third party testing reported in Ballistic Imaging] and high percentage of collected evidence that is not analyzed [reported by recent surveys] equates to an expensive cache of potentially crime-solving evidence not exploited. The reliability of IR toolmark images produces consistent quantitative correlations between fired components regardless of the size of a database. Searching for components with common origin does not require comparisons and rank ordering of the entire database. Identification assessment based on the simplicity and robustness of quantitative pair-wise correlations provides substantial improvement to network performance under conditions of: multiple simultaneous users inputting data and requesting searches, multiple component databases with overlapping contents and different update schedules, and unknown numbers of siblings present. Reduction in false positive candidates over the number currently selected by NIBIN can reduce the labor burden on examiners, improve timeliness of response to NIBIN queries, reduce operating costs, and expand case management capacity without needing to increase labor resources and costs.

Although it has been demonstrated that entering more evidence into NIBIN results in finding more hits, some users have reported that the time currently required for additional NIBIN entries is not justified by the expected increased return. The high throughput, minimal labor requirement, and resulting reduced cost of IR image collection and matching is expected to improve the false negative error rate of NIBIN. ROC analysis can be applied to selecting policies regarding NIBIN use, including analyzing the optimal strategy for reviewing rank ordered candidate matches.

Toolmark image collection and identification is a forensic process performed by military police, intelligence agencies, and local and federal law enforcement. The same techniques used to identify cartridge cases from one shooting to another can be used in analysis of bomb components from IEDs and other explosive devices. The ability to match toolmarks on portions of explosive devices found at different target locations provides a linkage. The ability to identify the toolmarks to tools found provides evidence linking occupants to events. Mark-to-mark and mark-to-tool identification provides important assistance to investigating authorities.

2. Main Body

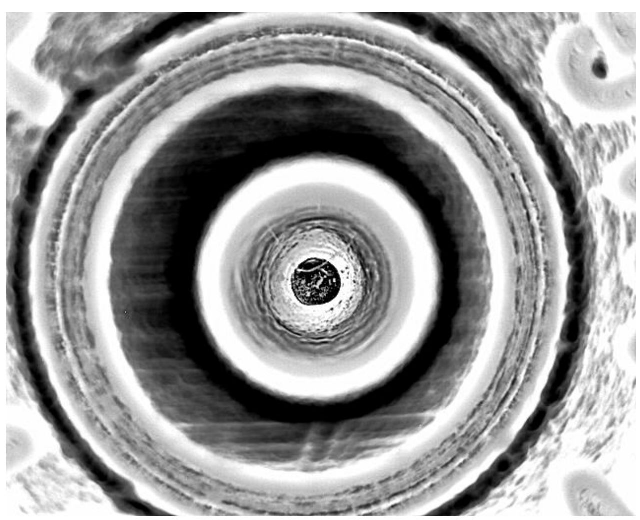
2.1. Determining the persistency of firing pin impressions after multiple firing

2.1.1. **Introduction:** The February 2009 National Academy of Sciences (NAS) report, ***Strengthening Forensic Science in the United States: A Path Forward*** outlined the need to improve the scientific foundations of the forensic disciplines, particularly those dependent on qualitative analyses and expert interpretation of observed patterns. It is an assessment in response to a request to address certain issues related to computerized imaging ballistics technology

2.1.1.1. **Statement of the problem:** Among its recommendations in that report, NAS outlined the need to improve the scientific foundations of the forensic disciplines, particularly those dependent on qualitative analyses and expert interpretation of observed patterns. The 386 page report details a five year assessment by the National Research Council (NRC) in response to a request from the National Institute of Justice (NIJ) of the U.S. Department of Justice to address certain issues related to computerized imaging ballistics technology. This NAS Committee was tasked to assess the feasibility, accuracy, and technical capability of a National Ballistics Database and, essentially, "assess the feasibility, accuracy and reliability, and technical capability of developing and using a national ballistics database as an aid to criminal investigations."

During a previous DOJ grant titled; "The Use of Infrared Imaging, a Robust Matching Engine, and Associated Algorithms to Enhance Identification of Both 2D and 3D Impressions", SED found that firing pin impressions (FPI) contain details that cannot be reliably reproduced by visible light imaging due to glare and shadows created by the illumination required to produce visible light images (see Figure 2.2.1).

Figure 2.2.1: IR Image of Firing Pin



Similarly, FPI details seen by an examiner using a visible light microscope have variations induced by the illumination required. Such variations reduce the reliability of automated toolmark comparisons and increase the time required by examiners to perform visual comparisons. What was not known was the persistency of the firing pin impressions considering multiple firings

2.1.1.2. **Literature citations and review:** See Section 3 for consolidated reference list.

2.1.1.3. **Statement of hypothesis or rationale for the research:** Our expectation is that IR images of firing pin impressions will not vary significantly over successive firings. The use of thermal infrared imagers and microscopes eliminates the need for illumination and thereby reduces apparent variations in captured and viewed FPI details. The reliability such imaging affords to automated comparison of cartridge cases is expected to improve the accuracy of fired cartridge case identification and reduce examiner workload. Prior research comparing visible light images of successive cartridge casings fired in new firearms showed greater variation among the initial 30 or 100 images compared to later firings. A similar effect is expected to occur in IR images; although IR images of multiple casings fired from the same firearm will have less variability than the corresponding VL images, initial groups of casings fired will display more variability than casings with a firing order above 100.

2.1.2. **Methods**

2.1.2.1. IR images are replicas of the spectral emissivity, or "texture map" of the casing surface, are not generated or influenced by visible light, and involve no user operations that could introduce variables into collected images. Sibling cartridge cases therefore produce IR images with substantially greater similarity than their visible light images. Given a collection of cartridge cases fired from known firearms, statistical analyses of the shapes, sizes, and positions of FPI features will be used to assess the benefit of using FPI features from IR images for eight firearms selected to represent ones commonly associated with crime investigations..

2.1.2.2. To validate the hypothesis, SED acquired eight (8) firearms and fired 1,000 cartridges in each firearm. Every 100th case was imaged in IR and compared to determine if the impression had changed and if the underlying algorithm can still be matched to other images of cartridge cases fired from the same firearm.

Firearms used are listed below with the brand of ammunition used for each::

- Beretta 92FS M9A1 9mm Luger (Federal)
- Raven Arms MP25 .25 Auto (Remington)
- Hi-Point CF380 .380 Auto (Winchester)
- H&K P2000 SK Subcompact .40 S&W with polygonal rifling (Federal)
- Glock 17 9mm Luger with polygonal rifling (PMC)
- Sig-Sauer P-226 .40 S&W (Speer Lawman)
- AR-15 (5.56) (PMC)
- SKS Eastern bloc 7.62x39mm (Wolf)

2.1.2.3. **Characteristics of firing pin impressions:** Several factors may be used to characterize and compare casings from their 2D and 3D image models, from IR cameras. Some of the same factors can be used with models from VL cameras.

Depth Profile of the FPI along a vector:
- Shortest chord between the firing pin deepest penetration and edge of the primer
- Chord designated by specific toolmark features for certain firearm-ammunition
- Maximum depth varies by primer hardness

Pattern Formed by the Firing Pin Aperture (primer flow back)
- Size (Beretta 92 is large)
- Shape (Glock is rectangular)
- Location relative to FP deepest penetration

Size and Shape of FP Circumference
- Hemispherical, oval, rectangular, circular, etc.
- Length of axes

Pattern Formed by FP Tip
- Polished, smooth, bumpy
- Type of patterns (dimples, non-parallel lines, rough, etc.)

2.1.2.4. **Image Processing Approaches:** When comparing images of different cartridge cases fired from the same firearm, or from different firearms of the same make and model, three image processing approaches are used:

2.1.2.5. **Firing Pin Rotation:** For firearms whose firing pins rotate with respect to the breechface, possible geometric variations in the position and orientation of firing pin impressions relative to those characteristics of breechface, shearing, extractor, ejector, chamber, and other toolmarks must be considered. Frames are selected from the imaging sequence which provides best focus of specific FPI features considered useful in establishing individual characteristics for the firearm and ammunition used. Each selected frame is then compared against the corresponding frame from another casing in accord with rules for assigning an overall correlation value for the match of FPIs. Alternately, a single extended focus frame can be derived from the selected frames for each casing's FPI. Rotation is performed to best align FPI features with a reference standard orientation. Standardized extended focus frames from different cartridge cases are compared to produce an overall correlation value for matching FPIs.

FPI comparison may be sufficient for matching two cartridge cases, or comparison of additional features may be needed. The same process is used of selecting best-focused frames for specific characteristics, and producing an overall correlation value for matching each feature. An additional rule is employed to yield an overall match value between two cartridge cases which incorporates all the specific feature comparisons.

In general, areas that are not in sharp focus do not significantly affect correlation between frames. Cutting those areas can provide wider separation distances between sibling and non-sibling cartridge cases, but substantially reduces the residual number of pixels on which the overall match value is computed. Rather than eliminating areas not in sharp focus, they can be changed to random values, or set to a constant value.

2.1.2.6. **Firing Pin Deflection:** For firearms whose firing pins do not rotate but vary in the location of their FPI relative to other toolmarks, a similar approach is used for processing and comparing image sequences from two cartridge cases: FPI features are separately matched and may be followed by matching of other features. Extended focus frames may be computed, and areas not in sharp focus may be cut, changed to random values, or set to a constant value.

2.1.2.7. **Constrained Firing Pin:** For firearms that maintain constant relative orientation and location of FPI, breechface, shearing, and other identifying toolmarks, producing and comparing one Extended Focus Image of each casing may be adequate for accurate identification. If manufacturing or other incidental marks appear prominent in selected frames, they should be filtered out prior to forming the Extended Focus Image.

2.1.2.8. **Consistency of Infrared Toolmarks:** In this task SED investigated the use of infrared imaging to determine the consistency of toolmarks on cartridge cases through firing of 1000 rounds. Eight firearms were considered. Each was to use a single make and model ammunition. However, two firearms purchased included test rounds from the manufacturers using different ammunition than we had selected. Also, .380 Auto caliber ammunition was difficult to obtain, resulting in our use of two types. It was therefore possible to provide examples of the use of infrared imaging with different types of ammunition.

IR images are the result of both emissivity and thermal variations across the imaged surface. Disturbances to an item's surface texture by striated mark or impression creates emissivity differences that appear in the infrared image as abrupt jumps in temperature even in the absence of any thermal change. High sensitivity of current commercial IR cameras provides ability to image very fine toolmarks on metal and other surfaces. Because infrared imaging requires no illumination, it eliminates shadows, reflections and other lighting-induced variations and artifacts associated with visible light imaging. No ring light or oblique lighting is required for imaging in this portion of the electromagnetic spectrum.

As a result, striated marks and impressions yield highly detailed images based on emissivity variations alone. Surface geometry relative to the infrared camera axis is an important emissivity factor. It is addressed by provisions for avoiding cold finger reflections when aggregating images of cartridge case cylinders and bullets. Detailed IR images can be rapidly captured and compared by fully automated systems because no subjective assessment of lighting or position is required. The use of fixed focus optics having shallow depth of focus, driven by very precise CNC (computer numerically controlled) positioning, produces sequences of image frames from which precise 3D surface models are generated automatically in near real time.

2.1.2.9. **Use of IR image sequences:** CNC control of focus distance, synchronized with frame capture of IR image sequences, captures precisely spaced 2D image slices used to automatically construct 3D surface models of FPI. IR image sequences provide a competing technique for 3D surface modeling with significant advantages in cost and throughput. The ability to automate toolmark imaging and metrology provides a level of consistency that cannot be matched by imaging that requires subjective human intervention. A version of Mikos Forensic Toolmark Workstation (MTW) was used to automate imaging and comparison of fired cartridge cases. It incorporates provisions that insert a unique bar coded identifier into each cartridge case to eliminate manual errors in marking, reading, and transcribing identifiers. The insert does not damage the cartridge case and can be removed. However, while installed it provides a keyed positional reference to the receiver that holds and rotates cartridge case during image collection. This provision establishes three-dimensional axes and a reference origin for the cartridge case. Independent re-imaging of each cartridge case produce quantitative measures of the reliability of MTW toolmark capture and eliminate random positional variations to expedite sensitivity analyses of different processing algorithms. MTW design was directed toward providing a quantitative method for toolmark comparison that would withstand rigorous scientific scrutiny and allow upgrades to its algorithms for feature replication and correlation.

2.1.2.10. **Match Value Algorithm**: A Match Value algorithm was employed which compared templates of breechface and primer shearing marks extracted from replicated toolmarks of pairs of cartridge cases. It correctly assigned all sibling pairs significantly higher Match Values than any nonsibling pair when used with a database of 262 images reflecting 140 firearms. Under the current NIJ effort, the Match Value calculation also includes comparison of firing pin impression details. A significantly larger database of Glock cartridge cases was used, and it included control parameters of particular interest based on studies conducted by other organizations. These include: total ammunition fired through each firearm, type of ammunition used, and number of intervening firings between two cartridge cases being compared. Statistical distributions of Match Values related to the controlled parameters provide measures of their influence on the similarity of toolmarks produced from the same firearm and from different firearms.

2.1.2.11. **MTW use in Evaluating Infrared Toolmark Identification.** By eliminating the need for subjective manual adjustment of lighting and focus, fully automated image collection of a frame sequence suitable for 3D surface modeling was performed in ten seconds. While imaging speed may not be a concern in normal NIBIN system mode, it is a factor when using an MTW to populate large databases for statistical analyses within limited time and cost resources. The same MTW can be used to capture toolmarks from both fired bullets and cartridge cases. For cartridge cases, it automatically rotates and translates the specimen into standard position, then captures a sequence of topographic slices that comprise a 3D surface model. Matching is automatically performed against a designated database using a combination of algorithms that consider breech face, shearing, and firing pin marks. For bullets, overlapping image frames are collected along the bullet's long axis around a full revolution at a sequence of focus steps. The precision of the CNC controller and sharpness of the toolmark emissivity features facilitate automated integration of the frames into a continuous 3D surface model. For efficiency, land impression areas are automatically determined and a set of extended focus images are produced; one for each land area for each bullet. This processing is performed for each bullet in the database at the time the bullet is imaged. Two bullets are compared by correlating their land impressions using FlashCorrelation®. Relative shifts equal to the number of lands produce the maximum correlation when corresponding lands of sibling bullets are aligned. Intermediate shifts produce lower correlation values, resulting in a sinusoidal Match Value signature as a function of relative rotational orientation.

2.1.3. **Results**

2.1.3.1. **Persistence of FPI details through 1000 firings was determined.** In addition, specific features and characteristics peculiar to each type of firearm were noted. Among the metrics used to determine persistence of features in three-dimensional IR images of FPI through 1000 specimens were the relative depths at which specific features appeared in focus. Depth is established from the frame number in the imaging sequence which is correlated to the focus distance between IR camera and cartridge case. Both two and three-dimensional toolmark features were produced and analyzed by the Forensic Toolmark Workstation.

The ability of infrared imaging to document key features of firing pin impressions known to be associated with the firearms used is a minimum requirement for demonstrating its utility. Addressed were the level of detail and consistency with which those key features appear in the IR images. Other features were designated that are consistently produced by firing pin actions as represented in the IR images; this included depth profiles obtained from IR image sequences between appearances of specific features in specific frames of the sequence.

MTW Use. The MTW was used to replicate toolmarks from the cartridge cases. It automatically rotated and translated the specimen into standard position, then captured a sequence of topographic slices that comprise a 3D surface model. Matching was automatically performed against a designated database using a combination of algorithms that consider breech face, shearing, and firing pin marks. The precision of the CNC controller and sharpness of the toolmark emissivity features facilitate automated integration of the frames into continuous 3D model.

Match Value Formulation. In earlier studies Glock-fired 9mm Luger caliber cartridge cases, the strongest discriminator between sibling and non-sibling pairs was detailed 3D renditions of the firing pin impression. However, in the current project involving only a single Glock firearm, greater variation from successive firings was seen in FPI features than in breech face and primer shearing marks.

Prior to initiating those massive imaging efforts, selection of imaging system parameters must be concluded particularly what optics should be used to extract the highest total forensic value from the imagery obtained. High enough magnification to obtain details of the firing pin impression may be warranted, even if that requires imaging each casing twice in order to also obtain breechface, shearing, firing pin aperture, and other marks. That determination will be based on the results included in this report. Other firearms used in this study showed influences of firing order on FPI image details similar to those shown in the following pages.

The results of the effects of firing order on details of firing pin impressions (FPI) were extracted from 3D/IR imaging of cartridge cases from three of eight firearms used in the project: Glock 9mm, Hi-Point 380, and AR-15 rifle. Most of our effort has been directed towards these three firearms for various reasons. Through personal contacts we have had access to several hundred Glock cartridge cases over the past two years, and we have established a database involving many Glock firearms with known linkage to cartridge cases. That complements the data collected in this project, which includes many cartridge cases from a single firearm. Also through personal contacts, we have received helpful support from the Hi-Point company to establish a large database of cartridge cases from more than 100 of their guns. Thousands of additional cartridge cases collected from Glock and Hi-Point firearms are available for future imaging and analyses to investigate the impact of variations in firearm, type of ammunition, and total firings. Sufficient data can be extracted to generate statistical distributions of quantitative values for characteristic Glock and Hi-Point features. Comparison of two fired casings will now involve:

- calculating the correlation Match Value between extended focus images of the two casings

- determining the accuracy of a match decision using that Match Value as threshold, based on numerous comparisons of each of the two casings with siblings and nonsiblings.

- determining the probability that the Match Value is a chance occurrence, based upon cross-correlation values from a much larger database of extended focus images of fired casings from the same make and model of firearm and ammunition.

2.1.3.2. **Use of FPI Imagery to Identify Cartridge Cases** fired from a new 9mm Glock Pistol. Cartridge cases fired in a Glock pistol usually are readily distinguished by a rounded-rectangular impression from the firing pin aperture (FPA), plus primer shearing marks along the short sides of the FPA rectangle. A firing pin impression (FPI) will be located mostly within the area bounded by the FPA. Locations of both the FPA and FPI can move within the primer area for different firings of the same Glock, and portions of the FPA can extend beyond the primer. Each casing is physically transformed into standard 3D orientation prior to being imaged in order to facilitate comparisons between imagery of an unknown casing and images of other cartridge cases previously collected. The standard convention produces imagery having the long sides of the FPA rectangle horizontal and the drag mark pointing to 3:00. Using a higher degree of magnification, greater detail can be seen of the deepest portion of the FPI. However, resulting restriction to the field of view would eliminate the references used to standardize the orientation of the cartridge case. Therefore, a level of magnification was selected so each image would contain a feature with known dimensions (diameter of the primer), a reference for assuring the base of the cartridge case is parallel to the camera lens (roundness of the primer), FPA, and available breechface impression marks.

In addition, partial inclusion of the headstamp in every image provides a means to confirm the separate identities of cartridge cases being compared. Cartridge cases from the Glock and other firearms used in this study produced highly consistent features throughout the 1000 cartridges fired by each firearm. Variation in the rotational position of the headstamp serves as an indicator that duplicate images have not mistakenly been included in place of comparing different cartridge cases. Also, the headstamp provides ready indication of any change in ammunition type, which occurred a few times in this study due to problems with supply availability and when factory-fired cartridge cases were provided which were different from the type selected for the project. Imaged headstamp features provide an embedded indicator, within the IR imagery, of ammunition change.

Depth Profiles from Glocks: An Extended Focus IR image of each cartridge case is produced by combining best focused segments from various frames in its image sequence. In particular, this will include frames in which FPI details, primer shearing marks, and firing pin aperture impression are in sharp focus. Producing extended Focus IR images permit 2D storage and processing for database comparisons. Resulting images are illustrated in figures 2.1.1, 2.1.2 and 2.1.3.

Figure 2.1.1

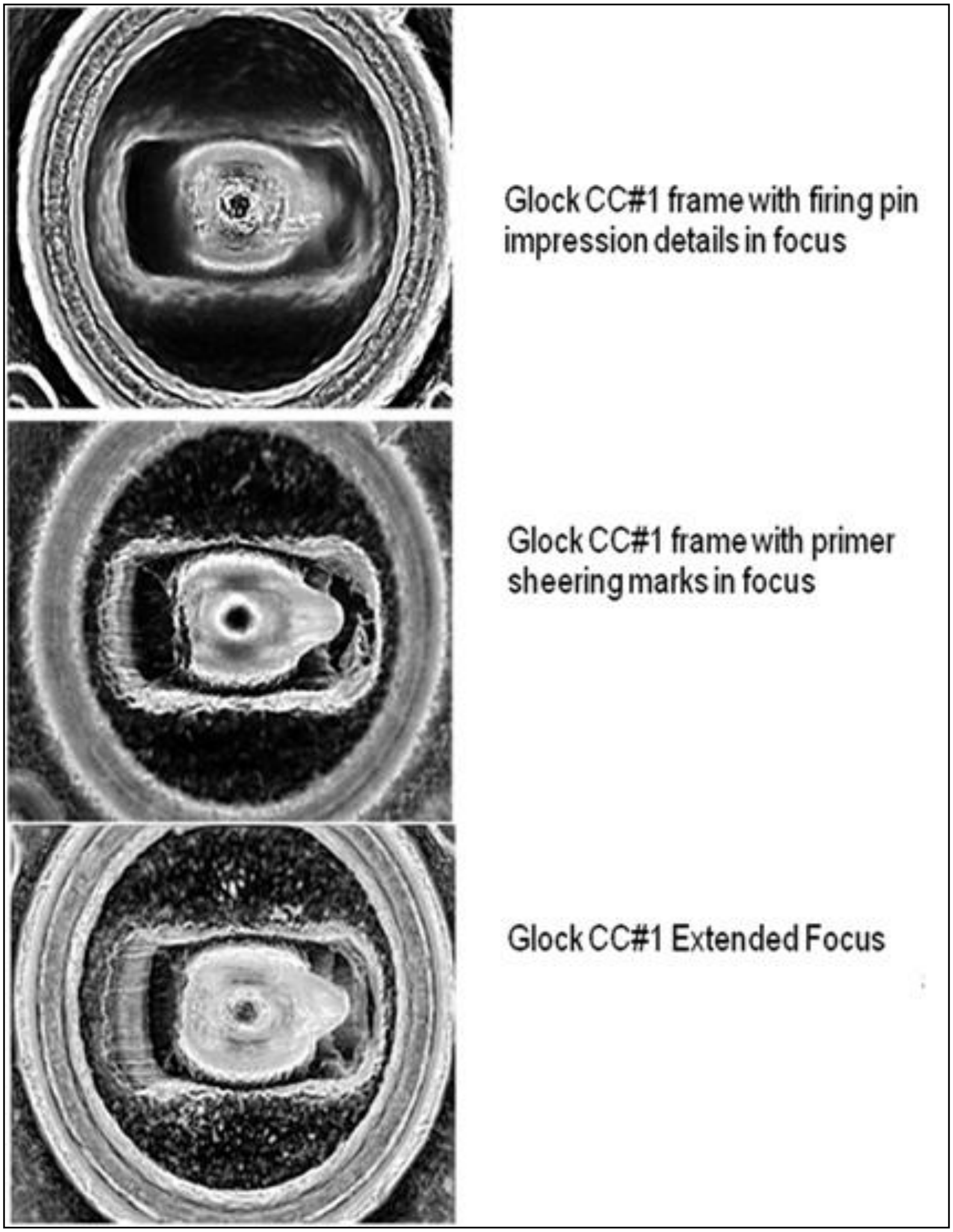

Glock CC#1 frame with firing pin impression details in focus
Glock CC#1 frame with primer sheering marks in focus
Glock CC#1 Extended Focus

Figure 2.1.2 Firing Pin Impressions from Glock 9mm

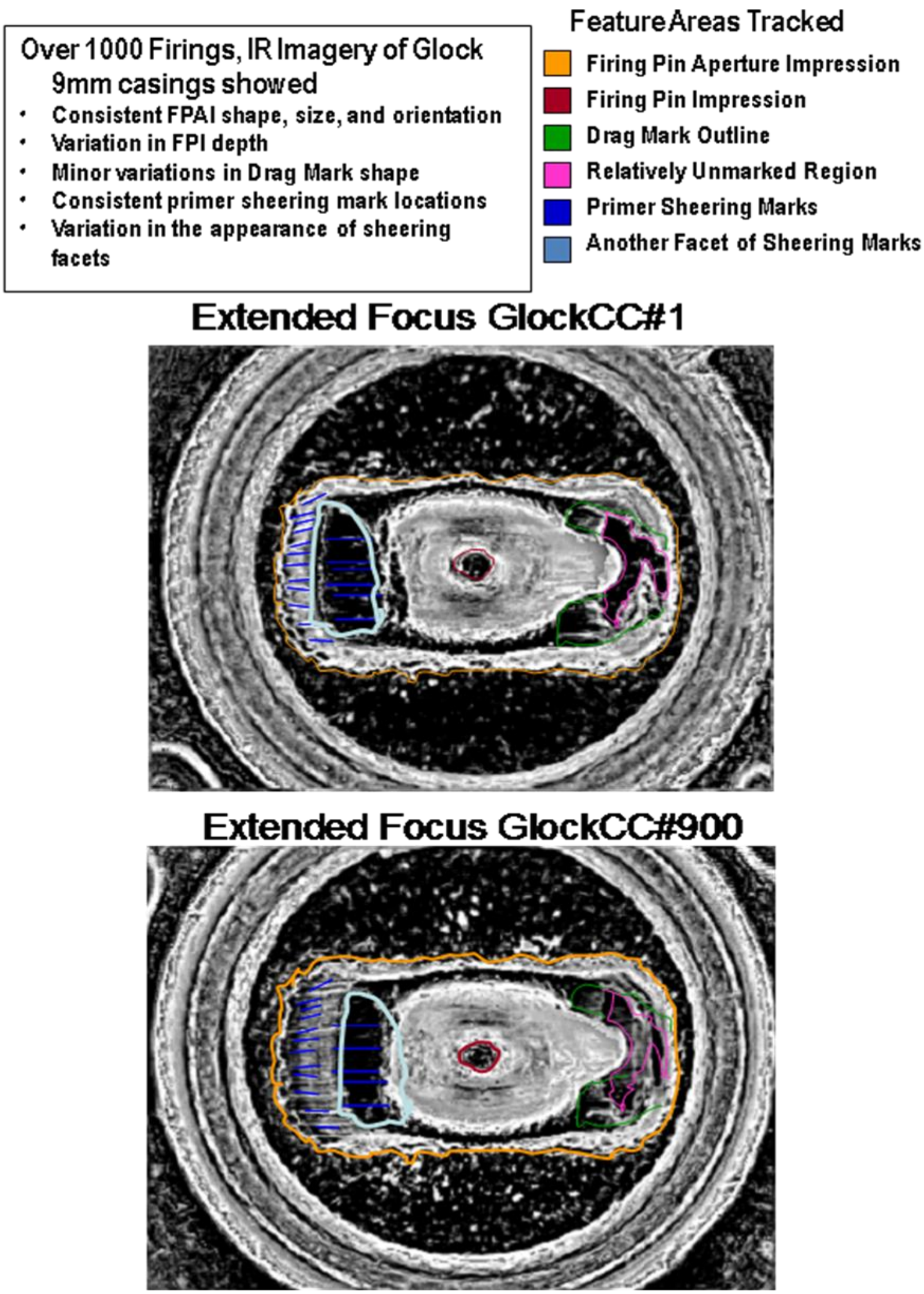
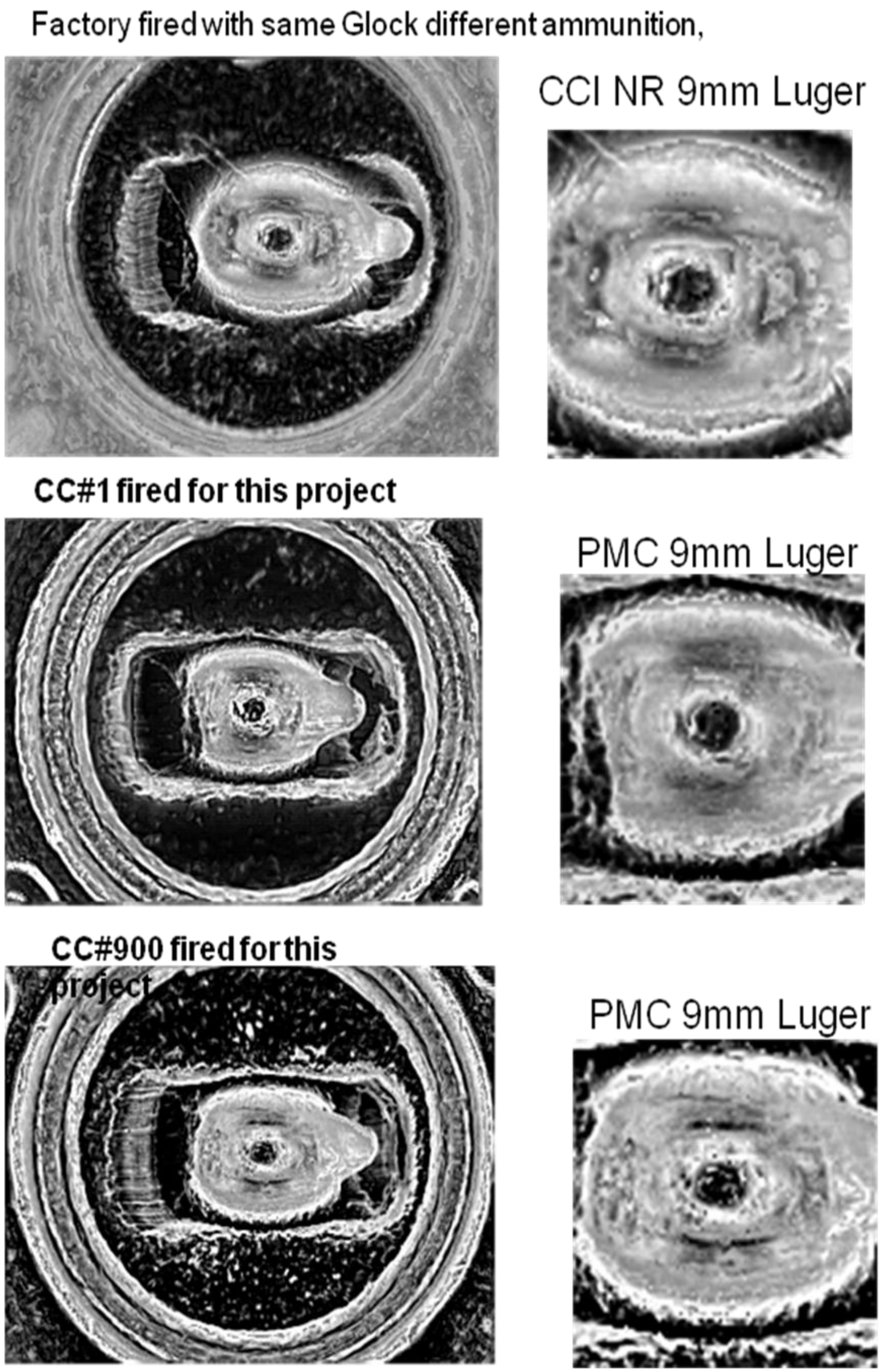


Figure 2.1.3 Effect of Different Ammunition on Glock FPI



Reliability of FPI Depth Profiles: In a previous study, the use of cutlines was evaluated for toolmark identification with IR cameras. In particular, cutlines perpendicular to the direction of travel were drawn through breechface and primer shearing marks. Fourier Transform Analysis was used to compare the sample waveforms from an unknown toolmark to a waveform database. The approach was found to be fast and effective, and was recommended as a screening technique against large databases. In this study, cutlines through in-focus FPI images were investigated as to their reliability, and effectiveness in cartridge case comparisons that use IR images of FPI. Waveforms extracted from manually selected frames without any control of ambient conditions were found sufficiently reliable to be the comparison method in identification studies with small databases. See figures 2.1.4 and 2.1.5.

- For Glock 9mm using PMC Luger Ammunition, Order of Firing has a measurable effect on the depth profile of the fired primer. Furthermore, changes in topography of the fired cartridge case due to firing order can be greater than changes due to different ammunition as was the case in the current project.

- Accuracy of FPI comparisons based on 2D/IR extended-focus representation showed no degradation due to order of firing in this sample investigation. However other comparison algorithms, including those based on visible light imaging, should consider the potential influence of firing order on quantitative measures used to select candidate matching cartridge cases.

- The degree of change in depth profile as a function of firing order decreased for the final 100 firings compared to the initial 100 firings. The influence of firing order is expected to decrease after some number of initial firings for each firearm type. For certain firearm-ammunition combinations, 3D/IR analysis may provide reliable determination of firing order among sibling cartridge cases collected from one or more locations.

Figure 2.1.4 Consistency of Depth Profile in Repeated Imaging

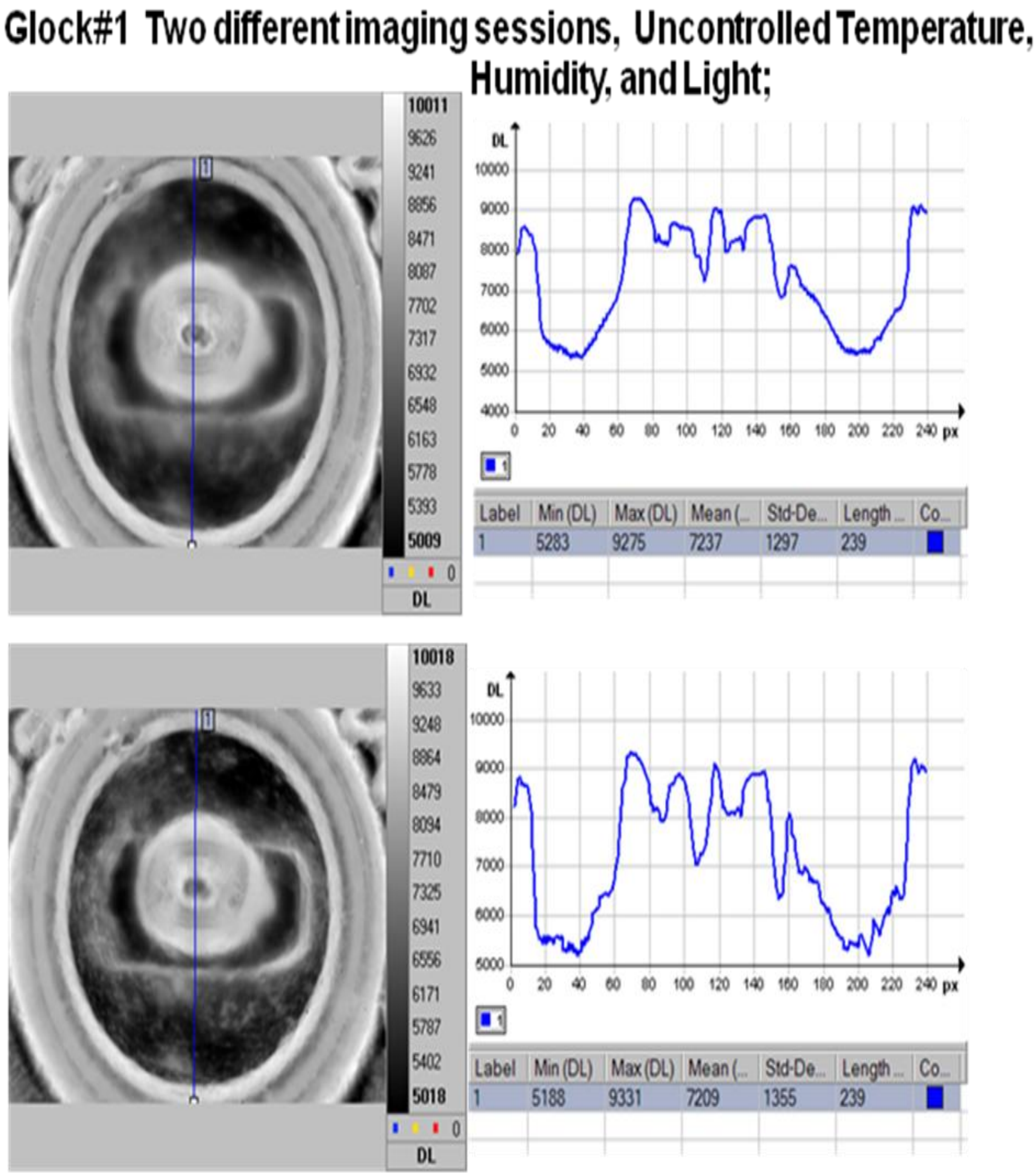
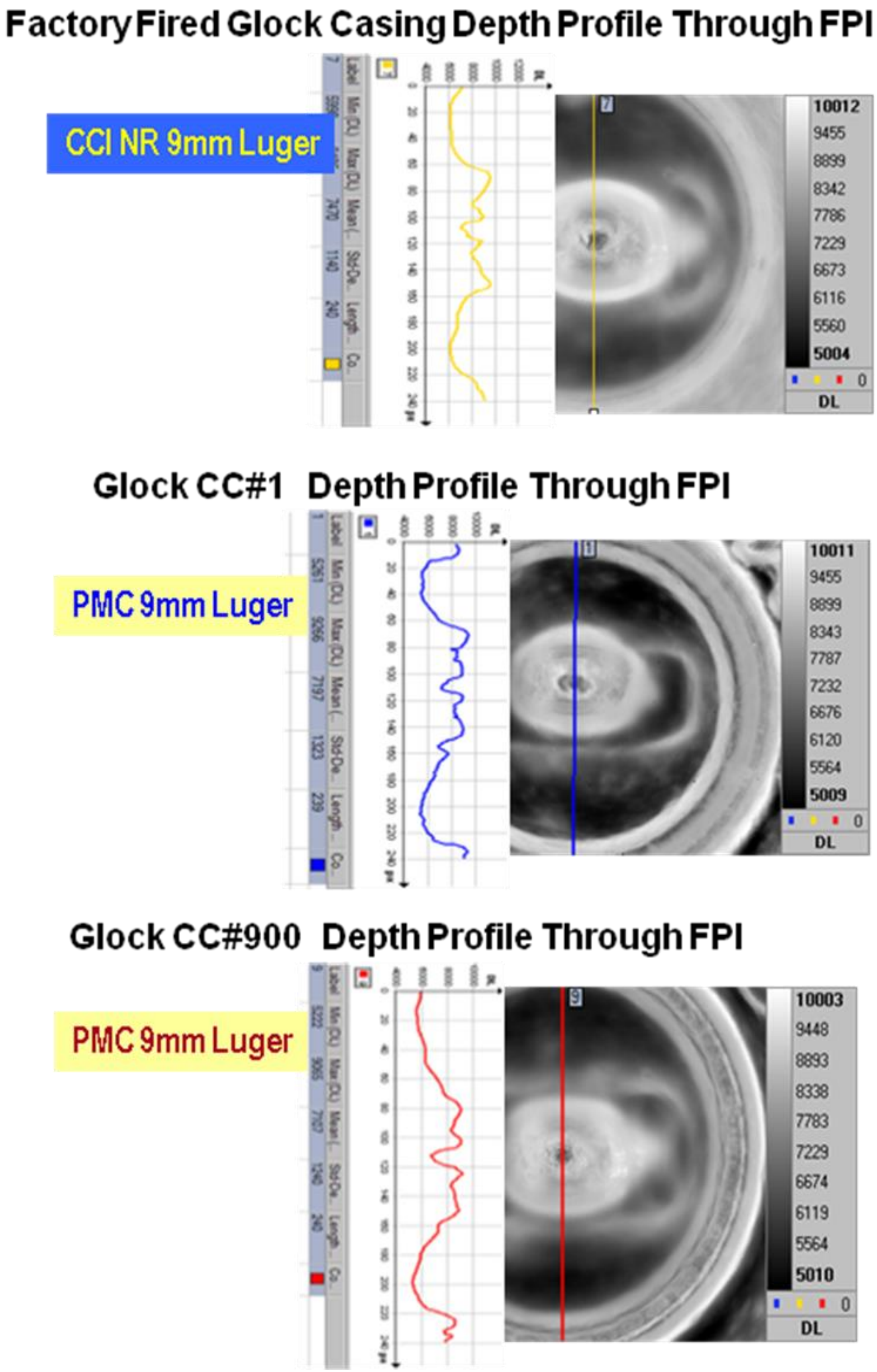


Figure 2.1.5 Influence of Ammunition and Firing Order on FPI Depth Profiles



2.1.3.3 Depth Measurements of Firing Pin Impressions in AR-15 223 Rifle

Fired PMC casings were collected from a new AR-15 223(5.56) rifle. Samples from the first 600 rounds were imaged and compared using the Mikos forensic toolmark workstation. The particular objective of this study was to determine whether fining order affected the firing pin impression as captured by 3D/IR imaging.

In order to establish a reference basis for depth analysis from 3D/IR imaging, the firing pin was removed from the rifle and measured, and FPI depth of a fired casing was also measured using traditional methods.

AR-15 Firing Pin

Firing Pin Tip Diameter

The tip of the firing pin from the project AR-15 rifle was measured using two different techniques. The first was the use of a mechanical caliper with a dial indicator from Brownells, and the second was employing the "air gap" method using a Leica UFM – IV Comparison Microscope along with a Mitutoyo digital caliper. Both readings were .0590 inches (1.499 mm). Both readings were taken a point beyond the ogive of the firing pin tip in an area at which a representative maximum diameter of the actual tip could be found.

Firing Pin Impression Depths and Apparent Diameters

Five fired cartridge cases having firing order 2-6 were used to measure a representative range of firing pin impression depths. The cartridge cases were arbitrarily designated and marked "A" through "E". Firing pin impression (FPI) depth measurements were made using a Federal Miracle Movement depth gauge. The apparent FPI diameters were taken using a Leica UFM – IV Comparison Microscope along with a Mitutoyo digital caliper. The results follow:

Fired Cartridge Case	FPI Depth (in/mm)	Apparent FPI Diameter (in/mm)
A	.0195/.496	.0533/1.355
B	.0250/.635	.0574/1.458
C	.0220/.559	.0565/1.435
D	.0255/.648	.0564/1.432
E	.0245/.622	.0583/1.482

Variations in measured depth and diameter were 16.3% and 5.4% respectively. FPI "apparent diameter" is defined for this purpose as where (in visible light) the FPI appears to present a maximum diameter exclusive of the circular downwardly curved surface adjacent to and surrounding the impression. Minimum resolvable difference of the IR imager used is 0.025mm. Reliability of measurements from 3D/IR analysis should be compared in future expansion of the current project.

Metrology of Fired Cartridge Cases in IR Imagery

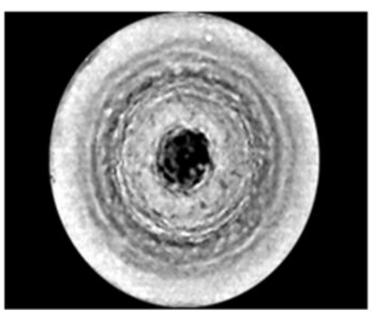
Optics Magnification Factor

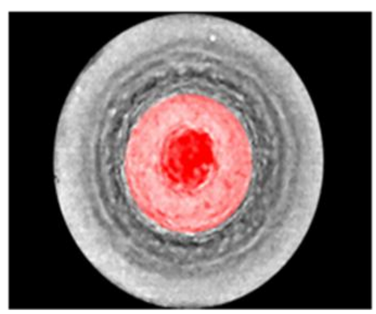
IR optics selected for viewing FPI details have very narrow field of view, as do magnifier lenses in general. IR optics present the additional challenge of severely limited depth of focus. In the Mikos forensic toolmark workstation, these two factors underlay the range gating method for producing 3D models from rapidly-stepped motion of a casing or bullet by a miniature CNC controller. This provides rapid 3D imaging of the total surface while also maximizing details extracted from the FPI.

Deconstructing Toolmarks from IR Imagery

When features within the primer area are in focus, the remainder of the casing base is generally out of focus and is cut from images so that measurement errors are not introduced by focus blur.

The left image below shows the primer area of an AR-15 spent casing from an IR imager having a 320x240 detector array and 4.3x3.2 mm Field of View. Measurement of the firing pin diameter made directly on the firing pin – just shy of 1.5mm – correspond to the diameter of the red area in the right image below, which coincides with a circular feature in the FPI impression area. Measuring the diameter of similar features in images of the AR-15 casings from the SED rifle would provide a comparison of the accuracy of this particular measurement from IR digital imagery vs. use of a comparison microscope and vs digital calipers.



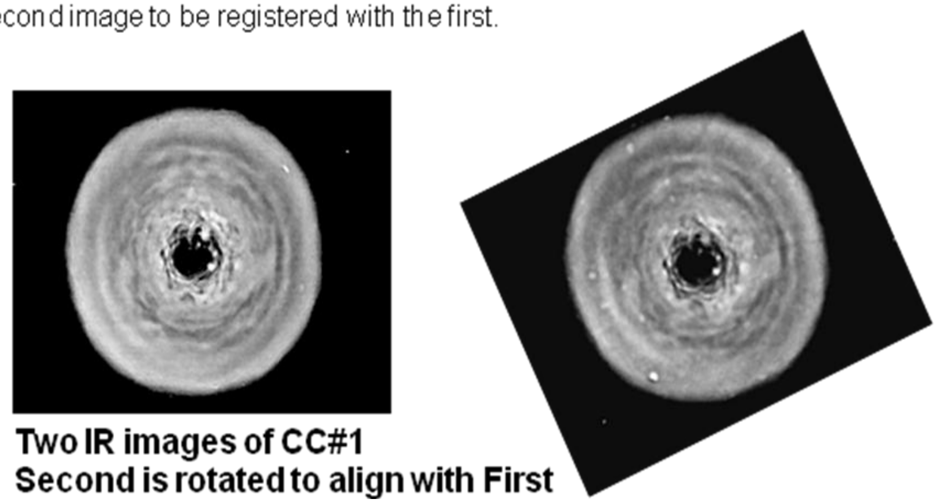


Rifle casings were not transformed into standard 3D orientation prior to initial imaging. Magnification was selected to include the complete primer diameter within the image boundary so that each image would contain: a feature with known dimension (diameter of the primer), reference for assuring the base of the casing is parallel to the camera lens (roundness of the primer), and the complete FPI.

Reliability of FPI Imaging

3D/IR analysis to compare FPIs uses structural and topographic features derived from thermal data collected by IR cameras. No controlled environment is required; ambient temperature and humidity, as well as lighting variations, do not measurably affect the accuracy of quantitative correlation between two casings' toolmark characteristics.

3D/IR imaging is performed automatically by the Mikos forensic toolmark workstation without manual adjustments required. Repeatability of produced imagery is illustrated below. Two independent scans of the same casing produced these images. FlashCorrelation® postprocessing rotated the second image to be registered with the first.

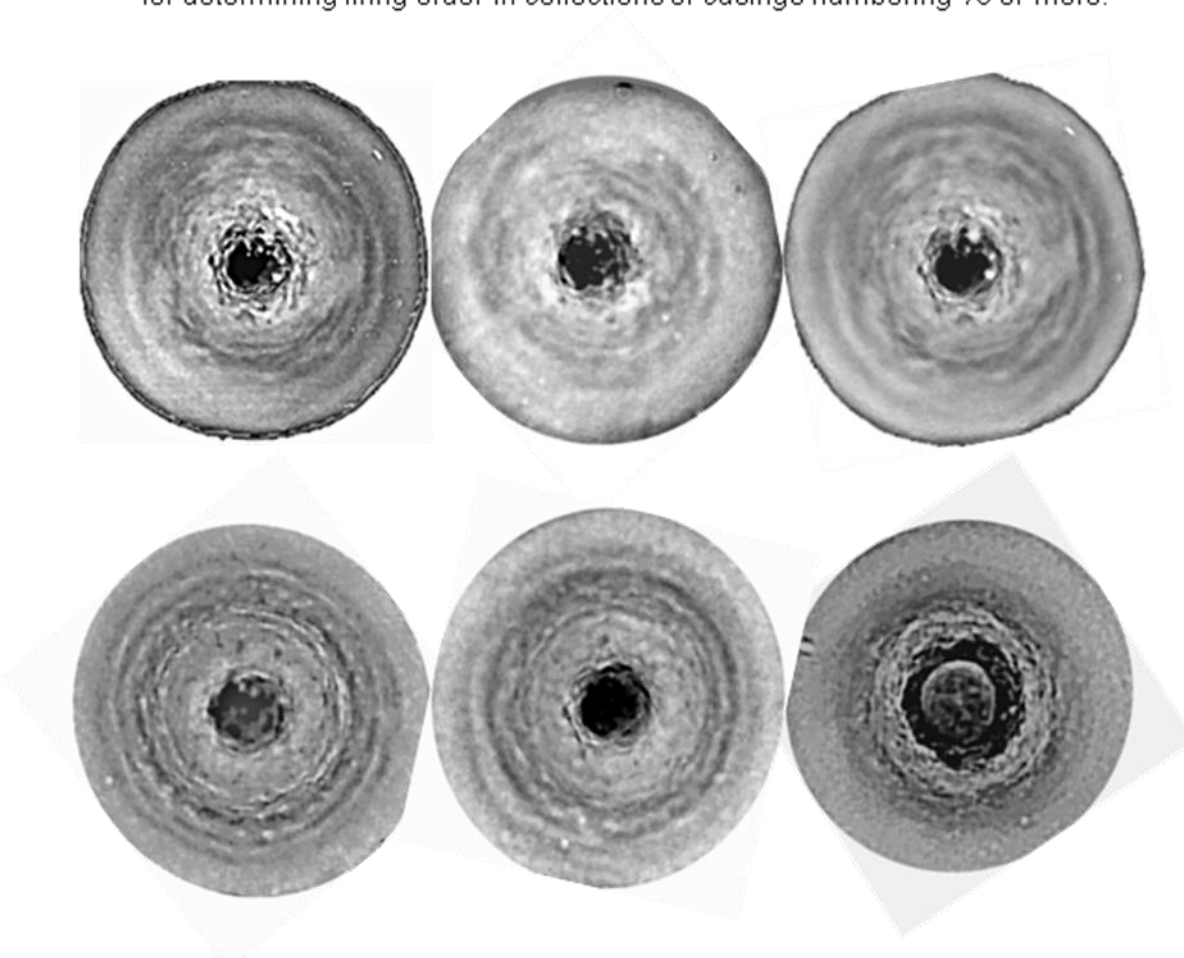


Two IR images of CC#1
Second is rotated to align with First

AR-15 Results

(1) *The first 100 casings fired through the new AR-15 rifle had highly similar individual details. Casings with later firing orders had fewer common details with the initial casings, but generally correlated well with casings having firing orders within +/- 50. FPI images below show the primer area of casings with firing orders from 1 to 900. While some features can be seen in common for several firing numbers, higher magnification or an image processing methodology would be needed to extract a set of features that would persist for more than 100 firings. Further investigation is needed on FPI variations from the AR-15 rifle.*

(2) *An apparent trend toward fewer fine details in later firings may be useful for estimating the total number of rounds fired from this rifle type. More detailed analyses, including higher magnification imaging, may provide the capability for determining firing order in collections of casings numbering 10 or more.*

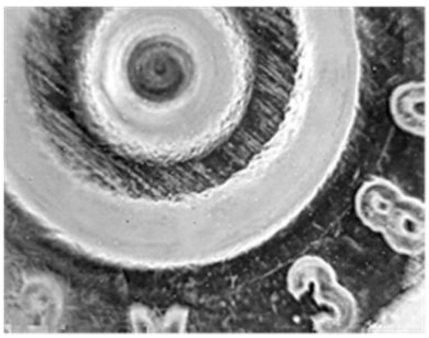


2.1.3.4 Variations in FPI with Firing Order

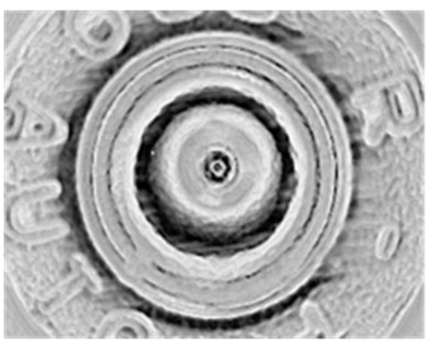
Variations in 3D/IR Firing Pin Impression Imagery with Firing Order from HighPoint CF380

HighPoint 380 Imaging

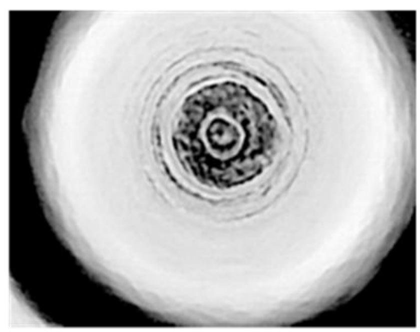
Before collecting the IR image sequence of the primer, as the casing is moved from the loading position of the Mikos forensic toolmark workstation, IR frames are captured in which portions of the headstamp and breechface marks are in focus. Breechface impressions are used by the workstation to standardize the rotational orientation of the casing. Different optics were considered for this project.



IR image frame of HP380 casing breechface and headstamp areas

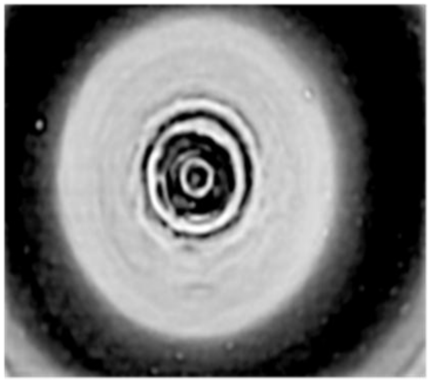


HP380 CC#1 frame with FPI in focus. Larger field of view covers almost the total casing base.

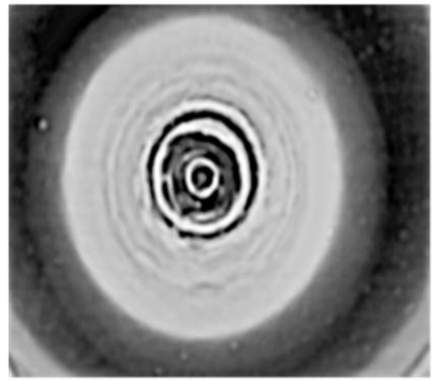


HP380 CC#44 frame with FPI in focus. Higher magnification optics do not resolve breechface toolmarks to provide standard rotation reference.

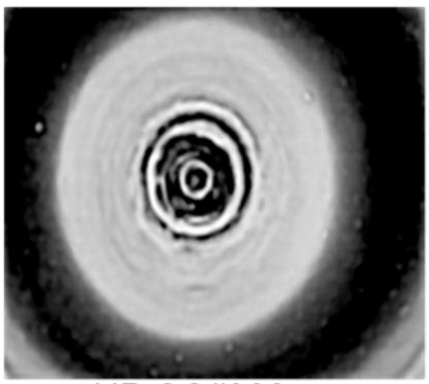
Winchester ammunition was primarily used in the Highpoint 380. Details of the Firing Pin Impression were consistent through at least the initial 300 rounds. Position of the FPI relative to distance to the primer edge was consistent through all rounds. RP ammunition was used for some initial rounds. FPI details were of different size and shape..



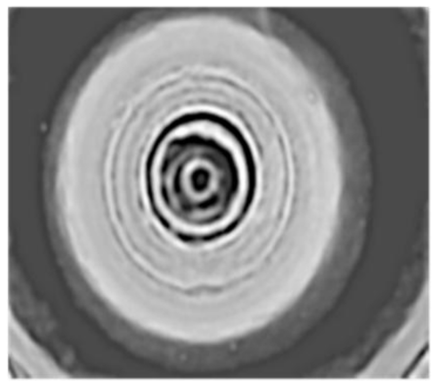
HP CC#1



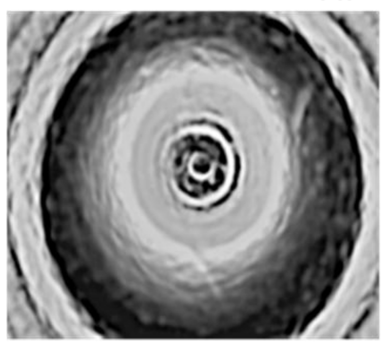
HP CC#200



HP CC#300



HP CC#500



CC#100 RP

For the Hi-Point .380 ACP caliber pistol using Winchester Ammunition, order of firing has a minimal effect on the firing pin impression for the initial 300 or more firings in a new firearm.

Change to RP ammunition produced significant changes in firing pin impression details. No attempt was made under this project to develop methods for comparing firing pin impressions from different ammunition. Any method that compared IR images from different ammunition types would need to consider the material composition of each type ammunition since that affects the spectral emissivity and would therefore induce changes to the infrared images which might require adjustment of the matching algorithm.

2.1.4. Conclusions

2.1.4.1. **Infrared images produce consistent depth profiles of FPI.** The MTW performance evaluation applied the following systematic approach to tests involving a large number IR and VL images of Glock 9mm cartridge cases. This generic approach could apply to other sample populations with other characteristics. The similarity measure was calculated for all pair-wise combinations of sample image templates. If the sample images are truly representative of the database, consider the feature extraction algorithm and the distribution of the number of features per template. Consider the match value distributions: compare AC, SC, and CC variances and the overlap between distributions. Adjust the match value algorithm.

2.1.4.2. **Implications for policy and practice.** A number of studies have found that forensic science experts are vulnerable to cognitive and contextual bias that leads them to make erroneous identifications. Recognizing those tendencies, courts might be expected to require that forensic evidence admitted in criminal trials be (1), based on a reliable scientific methodology that derives accurate findings from the evidence and (2), does not depend on human interpretation that might be flawed by bias, or by the lack of robust standard procedures. Historically, courts have not frequently challenged toolmark evidence, but more challenges are occurring. While court testimony is the purview of certified examiners, preliminary screening of database images to find potential matches has been developed as an automated computer system function over the past fifteen years. IBIS, the current automated approach used in the NIBIN system, uses visible light imaging which is known to create possible illumination-induced artifacts.

In addition, subjective adjustments to incident lighting, which is necessary to illuminate image details, negatively impacts the reliability of images used by that system to select candidate siblings. As a result of lighting variations, more candidates are selected than might otherwise be necessary. The resulting increased caseload might further lead to erroneous identifications by the examiner. The intent of automated forensic matching systems is to find the best matches between a database of images and a current item of evidence, and display the selected images to an examiner who makes the final determination as to whether a true match exists.

The goal is to minimize the workload of the examiner without sacrificing accuracy. Currently, the only fully automated forensic matching systems are for cartridge cases. The IBIS system has approximately one million records of cartridge cases although the full database is not generally searched. Third party testing of IBIS with a variety of firearms and ammunition has found negative error rates on the order of 35% and positive error rates on the order of 10%. Although emphasis has publically been put on the need to reduce the workload on examiners by reducing the number of false positives candidate matches presented, the cost impact on the criminal justice system from the high rate of false negatives is arguably greater. False negatives represent the failure of the system to recognize links among evidence it already has processed. Furthermore, decreasing the number of hits realized by the system works to decrease the perceived payoff to local law enforcement on their cost to house the system. Applying better search algorithms to the current NIBIN database might find new matches which could be important clues to solving more cases faster. Prior testing in 1998 of FlashCorrelation® (FC) vs. Drugfire, FC proved to have a much higher sibling match then Drugfire. In fact FC had an 80% higher probability of matches in the top 2% of rank ordering.

> Drugfire was the first computer based imaging system used in support of firearms examiners. It was developed by the FBI laboratory in 1989 and consolidated into the NIBIN system at a later date.

2.1.4.3. **Implications for further research.** The success to date using IR images is encouraging. MTW incorporates computerized numerical control (CNC) for precisely positioning the item under test and the IR imager, allowing a precise sequence of 2D/IR slices to be collected and used to generate a high resolution 3D/IR surface model in less than 30 seconds. Although both IR imager and CNC desktop controller technology is rapidly advancing, the reliability of 3D/IR models produced using current COTS subsystems is an important benchmark to be determined under the proposed effort. Also to be determined are the reliability and precision of striated, shearing, and impression toolmarks extracted from the 3D/IR models.

2.2. Conduct a Proof of Principal using IR to accurately match fired bullets bearing minimal damage

2.2.1. **Introduction:** The ability to determine sibling from a database of fired bullets has been less then effective and has proven to be a very difficult problem to solve. The current NIBIN system has very few bullets in file and seldom registers a bullet match. Providing a capability to more accurately match bullet siblings would significantly enhance law enforcements ability to solve cases.

2.2.1.1. **Statement of the problem:** Classical methods of bullet imaging and identification have involved a rather low power comparison microscope and an experienced human examiner looking at two bullets side by side. Appearance changes due to lighting strength and angle are complications for each of the imaging methods traditionally used. Approaches to 3D imaging have been aimed at eliminating dependence on visible light. 3D imaging technologies considered in recent studies by NIST and others have included:

- Stylus profilometer
- 3D virtual comparison microscope
- atomic force microscope
- confocal microscopy
- photometric stereo
- laser profilometry
- laser triangulation
- white light interferometry

In each case, a topographic profile was extracted from a single line across the surface of all lands. Bullets were compared through signal waveform analysis of their profile waveforms. All these methods have limitations and are lighting dependent. A better solution needs to be found and IR offers strong potential.

A recent NIST study involved four non-contact, non-destructive methods for firearms identification:

- vertical scanning interferometry
- point laser profilometry
- confocal microscopy
- focus-variation microscope

Summary critiques were:

- **Vertical scan interferometer:** is suited to rough surfaces
    - Note: identifications are usually made on examination of the striations in the LEAs since the engraved marks are more consistent and reproducible and therefore less variable from shot to shot
    - GEAs offer information but are not generally critical
    - Accuracy and precision of this method are excellent
    - But standard 50X lens and working distances of less than 1mm make imaging difficult
    - Also, technique cannot obtain data from an incline greater than 70 degrees

- **Point laser profilometer:** provides excellent depth resolution to 10nm; lateral resolution is limited by laser spot size
    - 1.7um typical smallest spot size which NIST contends is not sufficient lateral resolution for ballistic samples
    - Also, technique is very slow and specular artifacts found at the LEA/GEA transitions on actual bullets are not repeatable
    - Noise appears as large spikes;
    - This technique was deemed by NIST to be not appropriate

- **Confocal microscope**: can have high speed with excellent resolution vertical and lateral.
  - One type is the Nipkow disk used inside the FTI system.
  - However, the inability of the sensor head to image steep slopes raises chance of missing critical areas such as the sides of firing pin impressions, transitions from LEA to GEA, and sides of large striations.
  - Data collected across the NIST standard bullet LEA5 had large gaps and noise; generally this is considered due to the angle of reflection of the laser beam and numerical aperture of the lens

- **Focus-variation microscope:** Highly capable of imaging the transitional slopes between LEA and GEA without artifacts
  - Uses 10X lens
  - System can image up to 89 degree slopes and so is good for damaged specimens
  - Had very good lateral resolution even at low magnification; and vertical resolution that easily meet proposed criteria
  - This technology is slower than confocal system; but now has greater working distances which helps analysis of deep FPI
  - Also, cost is much less than confocal systems evaluated

NIST concluded that the confocal and focus-variation principles are most appropriate for continuing development. Confocal have an order of magnitude greater vertical resolution but are unable to image steep slopes Therefore NIST concludes that the greatest potential in firearms ID is offered by the focus-variation microscope. Its major advantages include:

- large working distance
- imaging up to 89 degree slopes
- could possibly reduce need to manipulate samples and could reduce sample prep time

Specifications that NIST deemed important to exploitation involved working distance of the lens relative to the sample surface, vertical and lateral resolution, and maximum angle of the surface relative to the sensor head that can be imaged. For a successful ballistics imaging tool the summary further specified:

- Acceptable vertical and lateral resolution such as 0.1um vertical and 1um lateral
- Acceptable working distance to enable measurement of deep impressions and badly deformed items
- Reasonable speed of data acquisition
- Ability to image steep transitions such as from land to groove areas and sides of firing pin impressions
- Rotary option to image cylindrical samples

- Utility for other purposes besides firearm ID

NIST developed its Standard Reference Materials SRM2460 standard bullet and is developing the standard casing for use in maintaining calibration of those 3D equipments. The imaging resolution of the technologies considered by NIST is on the order of 1 micron vs. the 3 to 4 micron resolution of the current MTW system. An investigation of sensor resolution in determining bullet identification sensitivity and specificity for each imaging technique, including spectral emissivity mapping, should be performed. Comparison of the benefits of spectral emissivity mapping relative to its traditional approaches to 3D imaging should be considered by NIST. See table 2.2.1

Table 2.2.1

· Specification details for each of the 3D principle types evaluated

	Vertical Scanning Interferometer	Point Laser Profilometers	Confocal Microscopes	Focus-Variation Microscope
Number of Systems Evaluated	1	4	2	1
Light Source	White light	Laser	Laser or White light	White light
Objective Lens Magnification	1x to 50x	Typically N/A (TTL 0.5x to 2.0x)	10x to 100x	10x to 100 x
Working Distance (mm)	7.4 (at 10x)	4 to 38	10.1 to 0.3	23.5 to 3.5
Resolution (µm)				
Vertical	0.03	0.01 to 0.05	0.01 to 0.001	0.1 to 0.01
Lateral	Not stated	1 to 30	3.1 to 0.12	1.1 to 0.4
Max. Surface Angle (°)	13.1 (higher for non-specular surfaces)	70	70 (85 in future)	90

2.2.1.2. **Literature citations and review:** See Section 3.0 for a consolidated reference index.

2.2.1.3. **Statement of hypothesis or rationale for the research:** SED's hypothesis is that using a quality infrared camera and appropriate microscopic lens, control hardware and underling algorithms can provide a much better method of matching sibling bullets from a database.

2.2.2. **Methods and Results:**

2.2.2.1. **Collection:** The methods employed for collection of samples began with the collection of pristine undamaged bullets test fired into a Kevlar collector (Bullet Catcher) and bullets bearing minimum damage test fired into solid wood material to produce slightly damaged bullets. SED collected bullets using the bullet catcher (see Figure 2.2.1 below) and fabricated wooden targets.

Figure 2.2.1 – Bullet Catcher

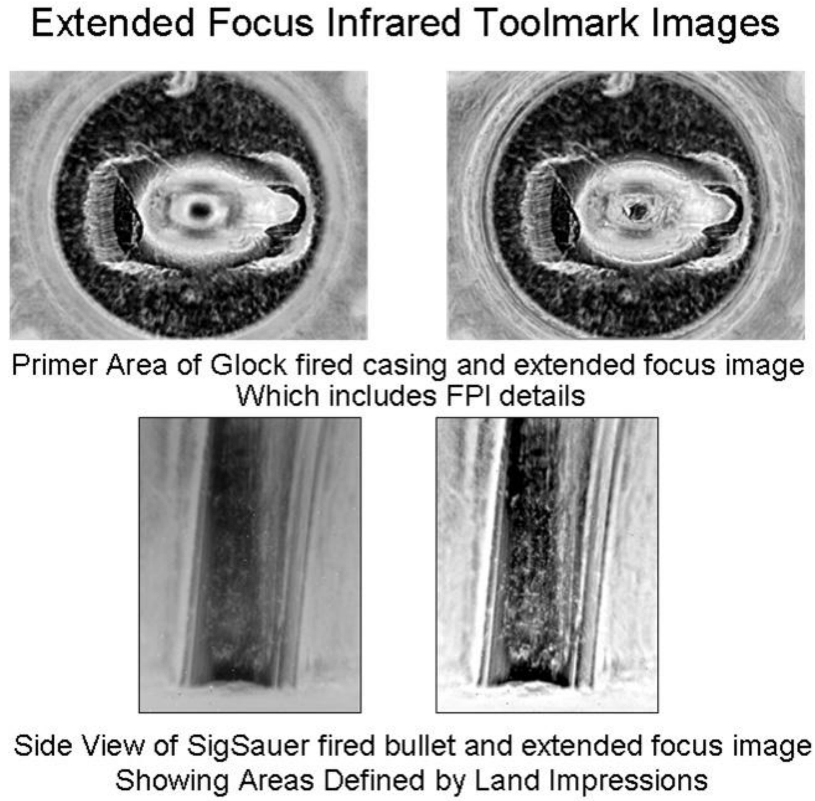
Each of the eight firearms was fired five times into the bullet catcher. The first two bullets and related cartridge cases were bagged separately as reference sets. The next three were bagged together. Each firearm was also fired into wooden targets consisting of pine boards from 5” to 10” in total thickness. At least two bullets were extracted for each firearm. Additional calibers were fired into the bullet catcher including: five .45 Auto caliber and two 9mm Luger caliber. Each was again fired five times, with the identifier of the bullets being the firearm serial number.

2.2.2.2. **Underlying Technology: Extended Depth of Focus IR Images:** The basic concept of extended focus images is to construct a single image in which all portions are in focus, starting with a stack of images that each have certain areas in focus. The resultant image represents what would be obtained from a camera that had an extended depth of focus capability, using images taken by a camera with a shallow depth of focus at different focal lengths. The processing is especially helpful with microscope imagery where minute change in focus can produce substantial changes to image features. Benefits from using extended focus images include:

- more compact file storage for large image databases
- faster searches for high correlation comparisons
- better visualization of image comparisons
- systematic method for combining macroscopic and microscopic details for comparisons
- ability to use lower cost optics and cameras
- more accurate comparisons between images taken with different camera systems

Infrared imagers and optics inherently have much shallower depth of focus than visible light systems. By automating precise stepper motor control of focus distance while collecting a sequence of infrared image frames, the MTW obtains depth-encoded 2D image slices representing a 3D toolmark model, and also obtains the component images needed to construct extended focus 2D composite images such as those shown in Figure 2.2.2.

Figure 2.2.2 Extended Focus Casing and Bullet Images



MTW processing is tuned to optimize use of IR imagery which is inherently greyscale. Various open source methods for producing extended focus images provide good results for toolmark imagery in general. They commonly use wavelet transforms to select focused areas in each component image and perform post-processing to maintain topology of the composite image while avoiding noise accumulation and image saturation. The topographic information is assumed to be contained in greyscale renditions of general imagery, and various methods have been developed to convert RGB to greyscale while preserving maximum contrast and intensity variations to highlight significant features in the greyscale presentation. Color information can then be reapplied to the composite extended focus image. Following a similar approach, the MTW combines infrared and visible light extended focus images to produce a two-color image that can be compared against both legacy NIBIN databases and newer collections.

NOTE: ImageJ plugin: Extended_Depth_Focus.jar Version 03.03.2004
References: Brigitte Forster, Dimitri Van De Ville, Jesse Berent, Daniel Sage, Michael Unser, "Extended Depth-of-Focus for Multi-Channel Microscopy Images: A Complex Wavelet Approach" Proceedings of the Second 2004 IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI'04), Arlington VA, USA, April 15-18, 2004

Measurement of surface temperature using infrared imaging is a well-established discipline of surface metrology, with its basis in the sciences of: thermodynamics, metallurgy, optics, infrared detector design, electromagnetic and spectral analysis. Common infrared cameras use a focal plane array of detectors to produce a corresponding array of grey scale values in which each pixel represents an “apparent temperature” measurement of the corresponding region on the surface. Calibration of the camera and surface is essential to determine the actual temperature from the apparent temperature. The key calibration factor is the spectral emissivity of the surface, which is a measure of how efficiently the surface emits heat. Its primary components are the texture of the surface, its material composition, and the angular relationship between the surface and the camera axis.

The huge impact of surface texture on apparent surface temperature makes infrared imagers excellent toolmark detectors regardless of actual surface temperature. In operational use, identification of firearm-induced toolmarks from infrared images does not require any temperature measurements. We cite the historical use of spectral emissivity for precise temperature measurement in order to introduce the underlying science, generally accepted procedures, and vast body of published peer-reviewed textbooks and other publications which can be used to support the scientific basis for this new application. The novel use of spectral emissivity for toolmark identification brings scientific rigor to areas of toolmark detection, characterization, comparison and identification. Infrared images are sometimes referred to as “spectral emissivity maps” or simply “emissivity maps” in order to emphasize that forensic toolmark applications of infrared imaging are directed toward topographic analysis and do not require precise determination of true temperatures.

Each of the components of spectral emissivity provides a benefit for toolmark identification. Both impressed and striated toolmarks create textural disturbances seen as apparent temperature variations in infrared images. The material composition of a bullet may be determined, or reduced to a small selection, by comparing its apparent temperature to its contact surface temperature to obtain its emissivity. Comparisons of toolmarks on fired bullets can utilize the variation in angular emissions to locate land areas of revolving bullets by their high contrast frames of imagery.

Infrared cameras that have large arrays of very sensitive detectors, plus microscope lenses, produce images having highly detailed representations of any changes in surface material or texture. When viewing a toolmarked surface, such as a cartridge case or bullet, the infrared image clearly shows striations and impressions as well as more subtle disturbances in surface finishes due to fingerprints, primer sealants, gunpowder and other residue.

Toolmarks made by a firearm on an ammunition component produce textural disturbances of the component’s surface that appear to be hotter than surrounding surface area when viewed by an infrared camera. Apparent temperature changes in the IR image are due to changes in spectral emissivity of the disturbed areas.

SED refers to IR images of toolmarks as "spectral emissivity maps" or simply "emissivity maps" in order to emphasize they are not calibrated images of toolmark temperatures, and in order to highlight important distinctions from visible light images of toolmarks. In particular:

- No illumination is required. Unlike visible light images which record reflections of light off the toolmarked item, emissivity maps are insensitive to visible light.

- Spectral emissivity maps are recordings of heat emissions continuously and spontaneously emitted by any object whose temperature is above absolute zero.

- Lighting-induced problems in visible light imaging do not occur in emissivity mapping. These include: speckle, glare, shadow, reversal of lands and grooves, loss of detail due to saturation or inadequate light.

- Toolmark features in emissivity maps are sharply defined and their dimensions relate directly to the actual feature dimensions measured on the toolmarked item; toolmark features in visible light images have less-defined feature edges whose apparent dimensions are influenced by lighting strength and angle.

Spectral emissivity maps of striated marks are generally characterized by abrupt emissivity changes at their edges, whereas impression marks have more gradual edge changes. Further analysis of emissivity maps is required to determine the utility of measurements from them that are not available from visible light images. For example, the slope of emissivity change along and across toolmark edges may be important characteristics of class, subclass, or individual features relative to determining make and model of the source firearm, or establishing linkage to a particular firearm or sibling casing or bullet. Transition slopes within certain features, such as breechface impressions or primer shearing, may vary with firing number for certain types of firearm.

Aspects of infrared imaging that are frequently considered serious limitations in some applications, in fact lead to major benefits when the application is high precision toolmark identification:

- **Huge Impact of Spectral Emissivity transitions on infrared camera images** provides an extremely sensitive, reliable, fast, high fidelity method for recording detailed toolmark images
- **Very Shallow Depth of Focus** facilitates range gating to create high precision extended-focus 2D images and 3D surface models for minimal additional cost and complexity.
- **Longer IR Detector Integration Time** combined with high magnification optics increases IR image sensitivity to vibration and other sources of motion blur. The need to eliminate vibration and motion induced by human operators, coupled with the desire to minimize size and weight of the toolmark workstation, lead to

development of the fully automated compact workstation that performs rapid image collection and database comparison.

- **Relative Cost of Infrared Toolmark Workstations** include the cost of infrared cameras that are typically 5-10 times the cost of standard 2D visible light cameras. However, that cost differential is more than offset by the level of automation provided by the IR workstation, which requires no trained operator. Furthermore, upgraded toolmark workstations performing 3D modeling and comparison capability using IR cameras have lower equipment cost than traditional 3D visible light technologies such as SEM. For each toolmark imaged, the IR toolmark workstation stores both a 2D extended focus IR image and a sequence of image slices from which the system can compose, manipulate and display a 3D view of each database item in response to an examiner operating virtual comparison microscope controls for simultaneously viewing a current physical item and a 3D digital model of a database item that may have common origin. This ability to integrate live and animated 3D presentations depends upon IR images being consistent and independent of visible light variations.
- **Relative Cost/Effectiveness of IR Toolmark Identification** based on performance tests to date is expected to significantly improve the efficiency and effectiveness of current systems; virtually eliminating false negative searches while dramatically reducing false positive matches provided to examiners, providing faster turnaround to investigators, reducing the cost of sole source maintenance contracts, while providing continued use of legacy databases of visible light images.

Infrared images of fired bullets provide visualization of textural disturbances to the bullet surface caused by contact with the firearm and materials it encounters. Detectors that are sensitive in the infrared electromagnetic spectrum are affected not only by temperature differences, but also by changes in texture, material composition, surface finish, humidity, and angle of sensor relative to the surface. Spectral emissivity, which refers to the efficiency with which a surface emits its heat, is a direct measure of influences that disturb the surface of a bullet.

Although normally considered a sensor for measuring temperature, infrared cameras actually detect surface emissivity measurements and convert them into temperature measurements only when other information is known about the surface material and condition. For forensic bullet identification, temperature is assumed to be constant across the bullet surface, and variations in spectral emissivity are therefore ascribed to either changes in material composition, surface texture, or trace materials. Composition changes would include a change in hardness due to variations in gilding metal. Trace elements that could be detected on the bullet surface by an IR camera include blood and human tissue, particles of construction materials, gun oil, and probably any trace materials found by other techniques and large enough to be resolved by the IR optics. Emissivity variations caused by trace evidence or by change in material composition can in general be distinguished from firearms-induced toolmarks by their edge characteristics. Striated toolmarks create abrupt crosswise changes to emissivity values associated with changes in depth of the mark.

2.2.2.3. **Bullet Imaging and Matching:** An MTW workstation was used to generate a series of overlapping emissivity maps that over-specified a 3D surface model for each bullet. The bullet stood on its base centered on a rotating platform controlled by the CNC subsystem of the MTW. An infrared camera with microscope optic was positioned with a horizontal principal axis that was perpendicular to the bullet principal axis. The field of view of the camera encompassed all of the undamaged extent of the bullet plus the portion of the rotating platform in contact with the bullet base. Various intervening mounts were evaluated that could be assigned to each bullet for purposes of applying a bar coded identifier. Knurled brass elements with a protrusion that fit the hollow under the base of the bullet performed best. They lifted the bullets off the rotating platform to provide imaging access to the full edge of the bullet base, provided reference markings that could be used to establish three-dimensional axes and definitive origin for each bullet, and helped maintain the centered and upright position of the bullet as it was rotated.

The frame sequence for each of two bullets being compared are subjected to processing that produces a closed 360^o 3D surface model replicating toolmarks and other topographic variations that can be replicated from the bullet. Called “sleeves”, the digital 3D models are automatically rotated into the best relative orientation and position defined by the total number of toolmark feature pixels that are overlaid minus the number of toolmark feature pixels that do not overlay. Assigning increased weight to pixels from longer features generally improves and speeds selection of the best relative positioning, with weighting used for both overlay and non-overlay conditions. Differently weighting each pixel by its location within the length and width of a toolmark feature tends to reduce false positive matches but unduly complicates the specification of feature edges whose continuity vary in different bullets fired from the same firearm.

The task of bullet identification is primarily determining whether two bullets have been fired through the same barrel. If so, then their respective class characteristics will be the same: caliber, the direction of twist of the rifling, the number of land and groove impressions, and the widths of the land and groove impressions. So the automatic detection, counting, and characterization of class characteristics is first performed on each bullet analyzed in order to quickly eliminate impossible matches.

Once the digital model replicas of toolmarks from a pair of bullets are aligned and registered, each is cut, flattened, and processed to produce 2D extended focus emissivity mappings which are then compared to each other and/or to a reference database using one or more of the methods developed for general 2D toolmarks such as screwdriver scrapes.

The primary tool used to do the matching is FlashCorrelation® which is a high speed optical correlater using frame to frame pattern recognition. Other methods were explored and can offer some additional capabilities. These include:

- 2D BarCode Generation and Comparison
- Pattern Recognition
- 1D BarCode Comparison from CutLines
- Signal Waveform Analysis

2.2.2.4. **Application of the Technology**: This section of the report summarizes applications for infrared imagery in the identification of fired bullets. Bullets were fired from six handguns and two rifles into soft material to produce mild deformations. Image sequences were collected using an automated CNC (computer numerical control) that guided the focus distance and frame intervals of the imager. Frame rates from 25 to 100 frames per second were collected. The camera field of view encompassed the full width and partial height of each bullet. For all bullets except those fired from the AR-15 rifle, each bullet was mounted on a riser that permitted unobscured view of the edge of the bullet base. Sequential frames were collected as the bullet on its riser rotated while centered on a platform. One or more rotations were imaged for each bullet. Most of the risers utilized a knurled or faceted design that provided geographical reference landmarks for stitching together individual frames or partial frames into a continuous 3D surface model of the bullet.

Uses for infrared imaging of bullets have been developed to improve measurement validity, accuracy, and reliability of ballistic toolmark identification while supporting our original hypotheses: First, that surface texture variations associated with firearm-induced toolmarks on cartridge cases and bullets, when imaged using infrared cameras, provide a scientific basis for toolmark identification. Second, that the reliability of emissivity mapping provides for meaningful quantitative correlations between surfaces being compared, with analyses of statistical distributions supporting likelihood estimates that surface toolmarks have a common origin. Third, that emissivity maps, being true representations of surface topography, can be compared against other representations such as three-dimensional visible light images. The infrared image recordings used to construct three-dimensional emissivity maps can also be rendered as Extended -focus two-dimensional images and compared against legacy NIBIN images. Figure 2.2.3 illustrates the repeatability of IR images, showing independent images of a 9mm Luger caliber bullet fired from a Glock pistol.

Independent IR Images of a 9mm Luger caliber Bullet Fired from a Glock Pistol Illustrate Reliability of IR Images

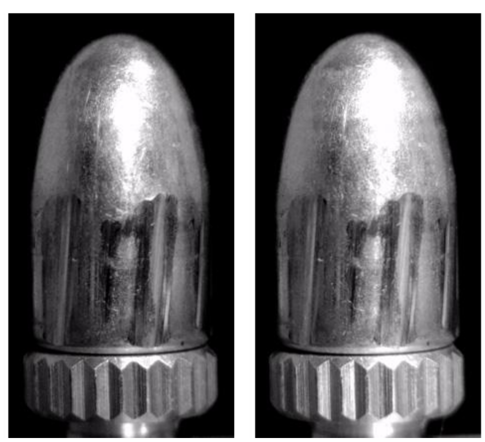
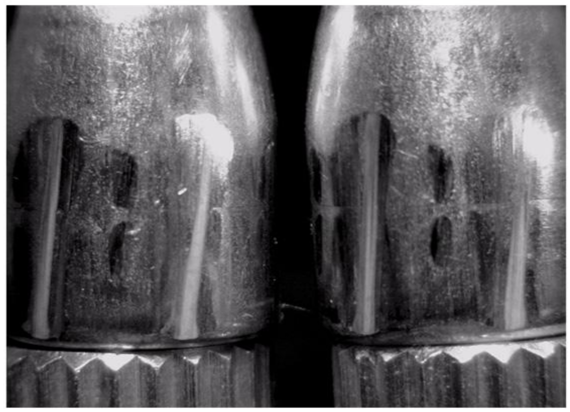


Figure 2.2.3

Figure 2.2.4 shows the similarity of IR bullet images in spite of polygonal rifling in this simultaneous imaging of two corresponding areas of two bullets fired from the same Glock pistol.

Figure 2.2.4

IR Image of Two 9mm Luger caliber Bullets Fired from the Same Glock Pistol Showing Corresponding Land Areas



In Figure 2.2.5, both bullets have been rotated the same amounts to display other similarly marked areas. In turn, common rotations are applied and successively exposed areas are compared between two bullets to compute an overall correlation.

Figure 2.2.5

IR Image of Two 9mm Luger caliber Bullets Fired from the Same Glock Pistol Showing Corresponding Land Areas

Figure 2.2.6 shows three rotational positions of a single Glock bullet to center three different feature areas.

Figure 2.2.6

Bullets Are Compared by Correlating Frames in which Their Land Impression Areas are Centered

At varying magnification, Figure 2.2.7 presents raw frames collected from two bullets fired from the same Beretta 92FS pistol.

Figure 2.2.7

Two 9mm Luger caliber bullets from the same Beretta 92FS pistol imaged in IR

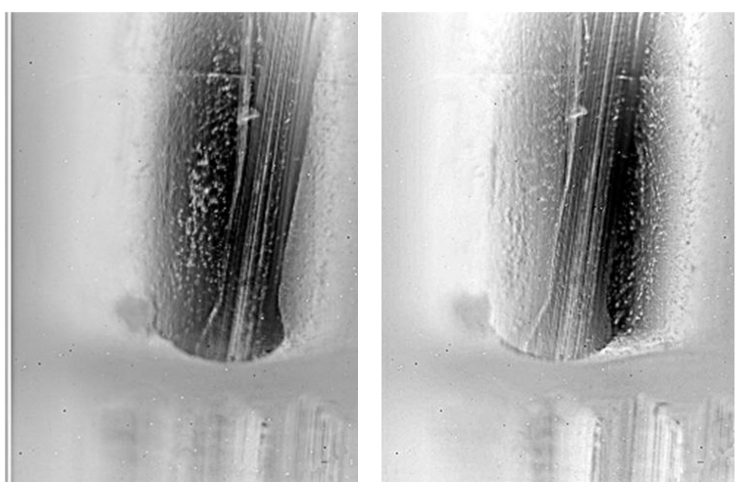


Figure 2.2.8 presents land areas of a Remington bullet fired from the Glock

Figure 2.2.8

Rotational IR Image Frames of a 9mm Luger caliber Remington Bullet Fired from a Glock Pistol

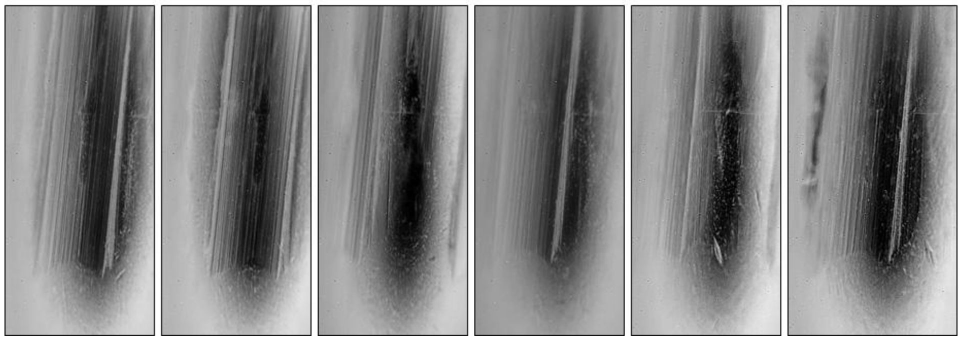
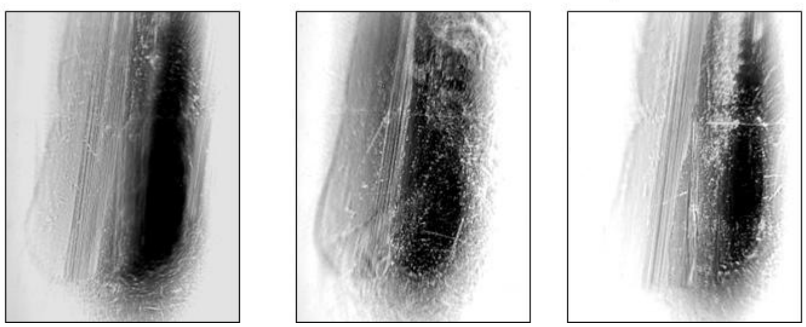


Figure 2.2.9 presents areas providing the highest correlations between bullets fired from the same Glock. The three areas shown are believed to represent three of the six land impressed areas on the bullet.

Figure 2.2.9

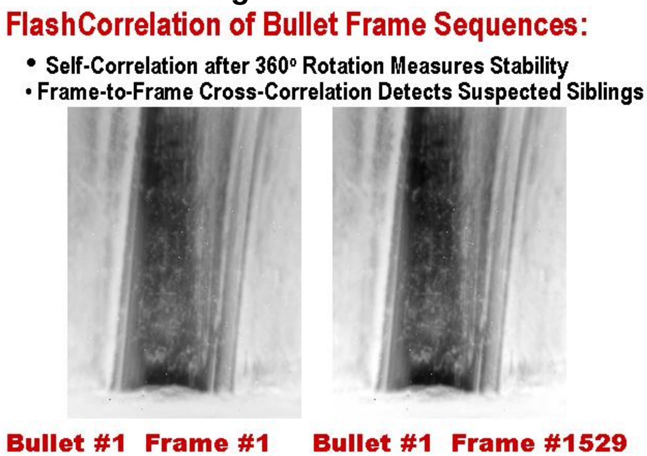
IR Images of Rotating Bullet Bearing Marks of Value for Correlation Purposes



2.2.2.5. **FlashCorrelation® used to match bullet images:** Striations can be automatically detected by the MTW from emissivity mappings. The complete treelike structures or skeletonized versions can be compared. Following the steps of CMS analysis used by experienced examiners, software algorithms can establish regions within the mappings, count the number of corresponding striae within each region for a given pair of bullets, and thereby perform large scale statistical tests on the reliability of CMS-guided decisions on common origin.

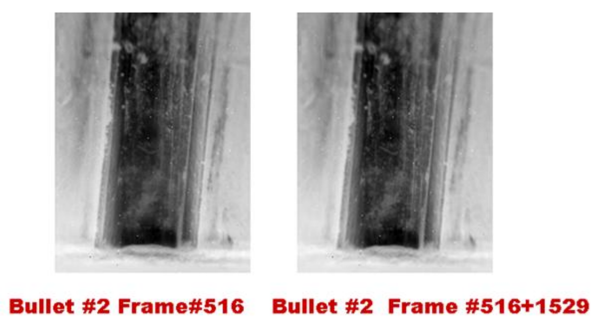
FlashCorrelation® is a generalized matching engine especially designed for optical processing. In Figure 2.2.10 an initial frame from the image sequence of Bullet #1 is processed against each of the 500 frames of the sequence from Bullet #2. The frame with the highest correlation becomes the initial frame of the Bullet #2 sequence. Corresponding frames between the two sequences can now be compared. Extended Focus frames generated from the sequences can be compared.

Figure 2.2.10



FlashCorrelation® is performed between the initial frame from the image sequence of Bullet #1 and the other frames of the same sequence for the purpose of finding which frame completes the first 360° rotation and subsequent rotations, and whether there is a shift due to mechanical slip in the rotating mechanism or other cause that affects reliability of the image capture. For the indicated sequence, a first rotation required 1529 frames. This type of self-calibration provision helps to maintain the precision and reliability of the CNC subsystem. Figure 2.2.11 identifies corresponding images of bullet #2.

Figure 2.2.11

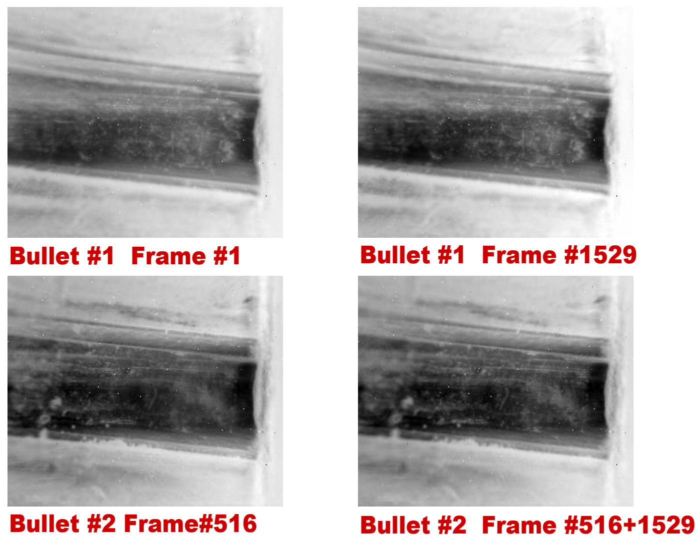


The above calibration determined that a full bullet rotation of bullet #1 under the parameters set required 1529 frames. Bullet #2 is now placed on the rotating platform and FlashCorrelation® is used to determine the best matching of its sequence of frames. It is also 1529 frames from the initial frame of Bullet #2. That is an indication that there is similarity between the sizes of Bullets #1 and #2. A frame-by-frame correlation value is computed which indicates the quantitative measure of toolmark correspondence across the entire surface of the bullet. See Figure 2.2.12.

Figure 2.2.12

FlashCorrelation of IR Bullet Image Sequences:

- **Self-Correlation after 360° Rotation Measures Stability**
- **Frame-to-Frame Cross-Correlation Detects Suspected Siblings**



2.2.3. **Conclusions:** Extended focus infrared images of land impression areas on fired bullets are expected to provide an accurate, reliable, and cost-effective basis for establishing common origin of bullets. This approach appears to be applicable even to firearms that employ polygonal rifling, which cannot be reliably identified by other bullet analysis techniques. The extended focus process currently used by the MTW has been shown capable of detecting and comparing some but not all of the land impressed areas on a few bullets fired from the same Glock pistol. Further study is warranted on the ability of this method to uniquely link fired bullets to different firearms, which was not a consideration in this study.

2.2.3.1. **Discussion of findings.**

- Methods employed in exploitation of emissivity mappings for bullet identification proved to be simple in concept and execution.
- Correct decisions of common origin or different origin were made in every case considered, which numbered less than 100.
- Additional testing with the parameters listed above and others will be performed during the remainder of this funded effort and will be disseminated on the project website.

2.2.3.2. **Implications for policy and practice.** The results have shown that if an Infrared based system is eventually made available to examiners and labs, the ability to match sibling bullets will become a reality and a totally enhanced capability provided to law enforcement in case management and solving crimes of violence.

2.2.3.3. **Implications for further research.** Some of the areas ripe for future research include:

- Comparisons across different types of ammunition
- Comparisons of bullets fired with 100 to 500 intervening firings
- Comparisons of bullet pairs with a specific number of intervening firings, but differing as to gun maintenance and type of ammunition,

2.3. Build a large database of cartridge cases and locate siblings

2.3.1. **Introduction:** The National Academy of Sciences (NAS) released its NIJ-funded report, ***Strengthening Forensic Science in the United States: A Path Forward*** in February of 2009. Among its recommendations, NAS outlined the need to improve the scientific foundations of the forensic disciplines, particularly those dependent on qualitative analyses and expert interpretation of observed patterns. The ability to image, store, and accurately identify fired sibling cartridge cases from a large database of infrared images was one of the goals of this grant. Earlier work demonstrated a procedure for using infrared imaging for firearms/toolmark comparisons that virtually eliminated false positive matches and had an accuracy of 99% in tests of cartridge cases fired in Glock pistols. An objective of the current grant is to expand the size of the database to determine the scalability of the infrared image comparison techniques used. The target is several thousand images to produce a statistically significant result. The investigation involved techniques for standardizing the imagery [as to scale, orientation, histogram, feature enhancement], determining which features had information content and which could be ignored, and selecting techniques to speed matching against large databases.

2.3.1.1. **Statement of the problem:** The NAS Report included the statements that: "Underlying the specific tasks with which the committee was charged is the question of whether firearms-related toolmarks are unique: that is, whether a particular set of toolmarks can be shown to come from one weapon to the exclusion of all others." One of the conclusions was - "***Finding: The validity of the fundamental assumptions of uniqueness and reproducibility of firearms-related toolmarks has not yet been fully demonstrated.***"

2.3.1.2. **Literature citations and review:** See Section 3. For a consolidated list of references.

2.3.1.3. **Statement of hypothesis or rationale for the research:** A database of fired cartridge cases was expanded to determine the scalability of the infrared image comparison techniques. A statistically significant sampling of infrared images was utilized for comparison purposes. Fired cartridge cases were primarily collected from common handguns: Glock and Hi-Point. The Hi-Point company provided fired cartridge cases and bullets from more than 100 handguns. Personal contacts allowed access to several hundred Glock cartridge cases that were entered into an established database involving many Glock firearms with known linkage to other fired cartridge cases. AR-15 and SKS rifles were also selected for examination and analysis.

2.3.2. **Methods:** A large database of infrared emissivity mappings was compiled from fired cartridge cases with known origins. The methods for assessing the reliability of the mapping process, predicting accuracy of identifications, and estimating the probability a given level of similarity between two cartridge cases fired in different firearms could occur by chance were evaluated.

Receiver Operating Characteristic (ROC) analysis was applied to a toolmark identification decision system that made quantitative pair-wise comparisons of infrared emissions maps from more than 2000 fired cartridge cases fired in Glock pistols. Different systems in similar applications can be compared on the basis of these distributions. The evaluation was used to determine the accuracies of identifications from comparing IR mappings; Infrared to Infrared (IR2IR), comparing visible light images to visible light images (VL2VL), and comparing IR images to visible light images (IR2VL). In particular, an IR2VL decision system can be compared to a VL2VL system.

2.3.3. Results

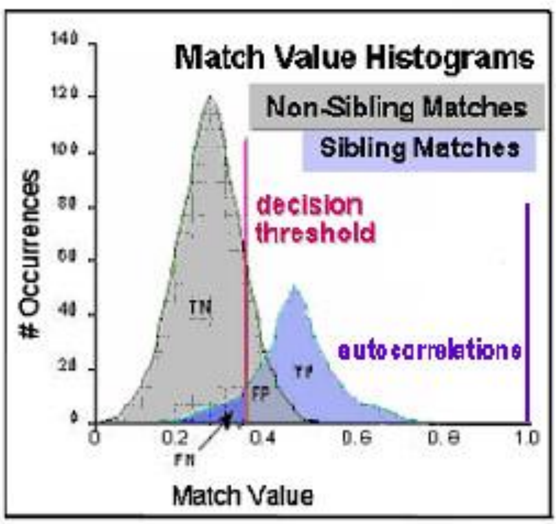
2.3.3.1. **MTW Performance Comparisons.** ROC analysis (noted above) was previously used in small scale tests to judge overall quality of MTW toolmark matching based on infrared emissivity mappings, visible light images, and a mixture of the two types. All three studies utilized the same methods to produce feature templates and the same FlashCorrelation® matching engine. Only breechface marks were compared. While IR2IR matching based on firing pin impression details has shown superior accuracy, the effect of multiple firings between comparison images had not been tested prior to the current project. Comparing firing pin impressions of fired sibling cartridge cases indicates much greater similarity in IR images than in VL images. This suggests that incorporating firing pin details into the match value algorithm would have improved IR2IR performance but not VL2VL or IR2VL.

Some systems have strong quality-control filters that will not allow poor images to be accepted. Eliminating poor images can decrease both false match and false non-match rates. Two identical systems can produce different ROC curves based on the strictness of the quality-control filters. With the exception of arbitrary quality control policies, ROC curves do not depend upon decision system policies, but only upon the basic distinctiveness and repeatability of the match values resulting from the feature extraction and comparison methods of that decision.

2.3.3.2. **Match Value Distributions**. Identification system performance is determined by the distinctiveness and repeatability of feature templates generated by the system. Regardless of what comparison method is used, identification can be modeled as a binary match/no-match [or sibling/nonsibling] decision based on thresholding a scalar match value. Accuracy of the system can be predicted by analyzing the distributions of match values produced from pair-wise comparisons.

Through controlled testing, we can establish three distributions of match values: 1) Match values from comparing images of siblings (ammunition components known to have been fired from the same firearm), which is called the "self-correlation or SC" distribution. 2) Match values from comparing images of nonsibling ammunition components known to have been fired by different firearms, which is called the "cross-correlation or CC" distribution. 3) Match values from comparing different independent image captures of the same ammunition component, which is called the "auto-correlation or AC" distribution and shows the variability in match values due to variations in system behavior when the toolmarked item is unchanged. We did not collect multiple independent target images. Therefore AC distribution in IR2IR and VL2VL identification is a line at 1.00 that represents the perfect correlation of each target image with itself. The effect of autocorrelations on system performance measures, and treatment of cross-spectral autocorrelations are discussed later in this report. For the current discussion, the AC distribution can be considered included in the SC distribution. If we establish a policy of using a "decision threshold" Match Value to distinguish nonsibling vs. sibling pairs, errors would inevitably occur to the extent there is overlap between the SC and CC distributions. No threshold could cleanly separate the two distributions. See Figure 2.3.1

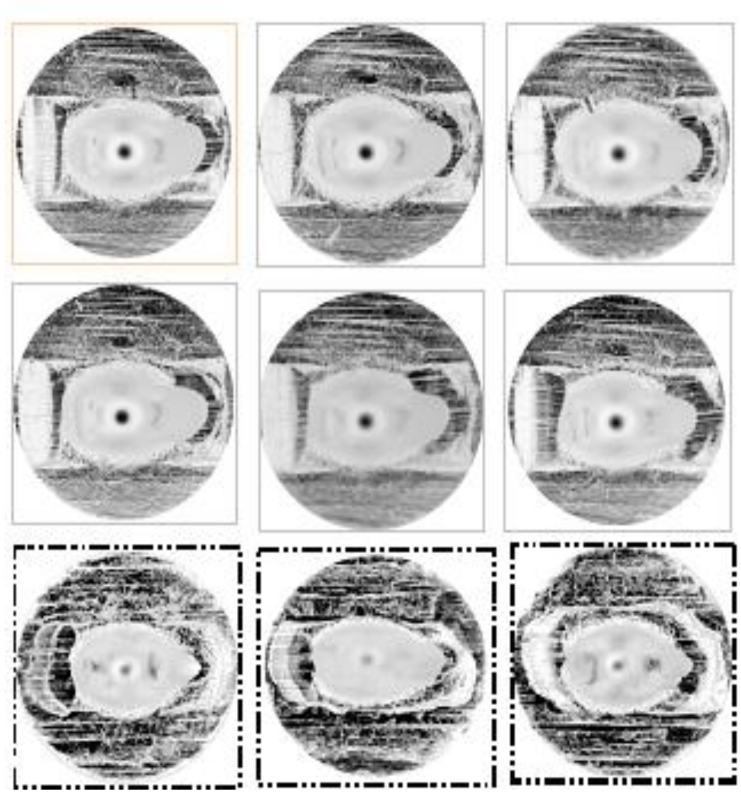
Figure 2.3.1 Match Value Histograms



2.3.3.3. **Match Value Characteristics**. A performance test was conducted using emissivity mappings from 262 fired 9mm Luger caliber cartridge cases fired in Glock pistols. The number of fired siblings each mapping had in the database was unknown at the start of the test, but was later determined to range from 0 to 5. Each of the IR mappings was used as a target image and matched against the full 262 mapping database.

Figure 2.3.2: Match Value Characteristics

		tan is Target			silver is Sibling			
	TOP EIGHT MATCHES FOR TARGET							
	WITH POSITION OF SIBLINGS FROM OBSERVATION							
Rank		1	2	3	4	5	6	7	8
Template	1072	1159	1164	1092	1160	1079	617	629	611
FeatureCount	10518	10549	10194	10436	9928	10815	10507	10651	10293
MatchCount	10518	3838	3695	3740	3345	3052	2791	2819	2664
MatchValue	1.000	0.357	0.356	0.352	0.330	0.276	0.260	0.259	0.253



The eight highest match values for cartridge case image IR1072 as shown in Figure 2.3.2, include its autocorrelation (which should always equal 1.000 and is computed as a check), followed by five siblings, followed by nonsiblings. Results imply the optimal Match Value threshold for sibling/ nonsiblings decision is between 0.260 and 0.276 for this Match Value algorithm, firearm and type of ammunition. Initial match values for Glock sibling pairs in all tests ranged from 0.271 to 0.524 for siblings and from 0.150 to 0.332 for nonsiblings using the current algorithm. Match Value distributions for sibling and nonsiblings IR pairs were both normal but the nonsibling pair variance was twice that of sibling pairs. However, the sibling pair population was only 0.5% the nonsibling pair population. At this time we assume no significant difference in variances of the two distributions so that we can use the simple ROC analysis equations given below. Studies using l visible light images have found matches of 25-28% between toolmark striae made with different screwdrivers and bolt cutters of the same brand and model. Biasotti's 1955 study found matches in 15 to 20% of the striae on bullets fired from different .38 Special Smith & Wesson revolvers.

Due to the known potential for finding high percentages of matching lines in known nonsibling toolmarks, a simple percentage of matching lines test is not sufficient to establish a common origin. Rather, the Match Value must represent a measure of similarity between corresponding toolmark locations. We note that the current MTW match value algorithm is not equivalent to a simple matching lines percentage, and that it has aspects in common with the use of consecutive matching striae (CMS) guidelines.

Studies using traditional visible light images have found matches of 25-28% between toolmark striae made with different screwdrivers and bolt cutters of the same brand and model. Biasotti's 1955 study found matches in 15 to 20% of the striae on bullets fired from different .38 Special Smith & Wesson revolvers. Due to the known potential for finding high percentages of matching lines in known nonsibling toolmarks, a simple percentage of matching lines test is not sufficient to establish a common origin. Rather, the Match Value must represent a measure of similarity between corresponding toolmark locations. We note that the current MTW match value algorithm is not equivalent to a simple matching lines percentage, and that it has aspects in common with the use of consecutive matching striae (CMS) guidelines.

- **Autocorrelation Match Values.** Sensitivity analyses can determine the impact of repeated firings, different ammunition, or different firearm on match values and the threshold value that best divides siblings from nonsiblings to achieve desired system performance. Two or more independent mappings of each casing added to the database should routinely be performed to properly assess the reliability of the system process and provide a value for autocorrelations in the ROC analysis. The inclusion of autocorrelations noticeably impacts ROC performance curves on small-scale tests. Since some of the sensitivity analyses we will be performing involve small numbers of samples, such as cartridge cases from different types of ammunition, we adopted a standard policy to perform independent re-imaging.

- **Accuracy as a function of Match Values.** The following charts show the calculations of identification accuracy as a function of match value, with and without inclusion of autocorrelations. The difference caused by treatment of autocorrelations is apparent at match value extremes where there are few data points. When autocorrelations are included, the total number of comparisons is 1600, of which 40 are autocorrelations with match value of 1.00. In the lower charts, autocorrelations are deleted and the total number of comparisons is 1560 of which 40 are self correlations with match values ranging from 0.310 to 0.500. Accuracy vs. Match Value graphs are useful in optimizing a decision strategy only if the chart is also considered. For example, maximum system accuracy (highlighted) occurs when the match value threshold is 0.330 to 0.320. That generates less than ½% False Positives but produces a True Positive Rate of only 65%, meaning 35% of true hits would be missed. A better strategy might be a match value threshold of 0.300 which produces 10.86% False Positives but identifies 82.5% of true hits. The FPR and TPR columns in the accuracy vs. match value charts indicate the tradeoff between those two values. As illustrated by the histograms of CC and SC distributions, adjusting the threshold line improves one of the values at the expense of the other. These two columns are used to produce the ROC curves discussed below in Table 1, IR2IR with and without AC Match Values Calculation of Accuracy vs. Match Value.
- **Decision Strategy in Toolmark Identification.** NIBIN/IBIS can be considered a component of a decision system for toolmark identification. Its input is ammunition components and metadata entered by human operators and its output is a rank-ordered list of database images based on a proprietary algorithm. A firearms/toolmark examiner is needed to make the actual decision. The examiner

uses the NIBIN rank ordering as a guide as he examines images and corresponding physical specimens to find any true match to a particular target component. See Figure 2.3.3 The NIBIN + Examiner decision system makes a decision for each database item, for each target. The overall toolmark identification decision system also includes policies directed by management and strategies implemented by examiners. For example, a policy can state that the top 10% of matches will be examined in all cases, or that all images producing a match value above 0.25 will be examined regardless of rank order. Other policies could be that images with match values below 0.10 are never examined, or that target images producing no match values above 0.15 can be assumed to have no sibling images in the database. Policies could direct treatment of damaged specimens and poor quality images, and specify the number of hours of training examiners must receive. Policies can include different procedures depending on cost/benefit analysis of expected results from each. Several policies are generally in effect at the same time, and they impact the performance of the overall decision system. An examiner's strategy or approach to assessing a candidate match may differ from that of another examiner even though they reach the same conclusion and adhere to the same policies. Of course a broader definition of the toolmark identification decision system would include training and experience of the examiners, proficiency testing of technicians, maintenance of the NIBIN system, and consideration of individual factors such as visual acuity, levels of concentration and fatigue. See Figure 2.3.3 and Table 2.3.1

Figure 2.3.3

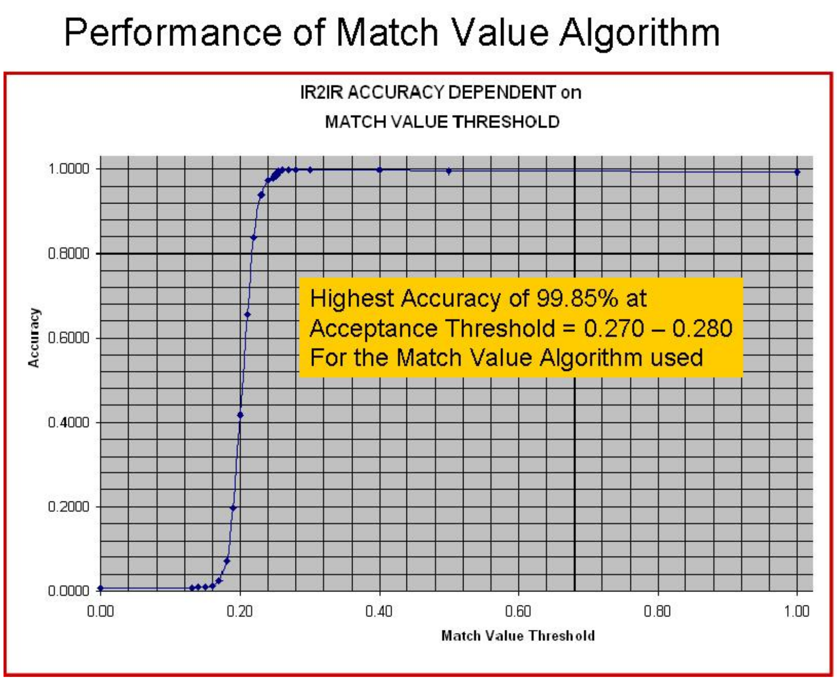


Table 2.3.1

IR2IR with and without AC Match Values
Calculation of Accuracy vs Match Value

IR2IR 40 9mm Casings from 20 Glocks 2 Firings Each Includes Autocorrelations=1.00

Match Value Threshold	Auto-Correlation	Self Correlation	Cross-Correlation	TOTAL< Threshold	TOTAL> Threshold	TN= CC<	FP= CC>	TP= ACL> +SC>	FN = AC< +SC<	FPR= FP/(FP+TN)	TPR= TP/(TP+FN)	ACC= (TP+TN)/TOTAL
0.100	0	0	0	0	1600	0	1520	80	0	1.0000	1.0000	0.0500
0.110	0	0	0	0	1600	0	1520	80	0	1.0000	1.0000	0.0500
0.150	0	0	1	1	1599	1	1520	80	0	0.9993	1.0000	0.0506
0.170	0	0	1	2	1598	2	1519	80	0	0.9987	1.0000	0.0512
0.180	0	0	2	4	1596	4	1518	80	0	0.9974	1.0000	0.0524
0.200	0	0	22	40	1560	40	1502	80	0	0.9741	1.0000	0.0140
0.220	0	0	114	223	1377	223	1411	80	0	0.8635	1.0000	0.1168
0.240	0	0	252	643	957	643	1129	80	0	0.6371	1.0000	0.3904
0.250	0	0	256	899	701	899	877	80	0	0.1938	1.0000	0.5275
0.260	0	0	282	1181	419	1181	621	80	0	0.3446	1.0000	0.6100
0.270	0	0	188	1369	231	1369	339	80	0	0.1985	1.0000	0.8104
0.280	0	1	96	1466	134	1465	151	79	1	0.0934	0.9815	0.9104
0.290	0	3	33	1502	98	1498	55	76	4	0.0354	0.9500	0.9639
0.300	0	0	15	1517	83	1513	22	76	4	0.0143	0.9500	0.9639
0.320	0	3	2	1528	72	1520	2	72	8	0.0013	0.9000	0.9938
0.330	0	1	0	1529	71	1520	0	71	9	0.0000	0.8815	0.9944
0.340	0	0	0	1529	71	1520	0	71	9	0.0000	0.8815	0.9944
0.350	0	2	0	1531	69	1520	0	69	11	0.0000	0.8625	0.9931
0.360	0	1	0	1532	68	1520	0	68	12	0.0000	0.8500	0.9925
0.370	0	6	0	1538	62	1520	0	62	18	0.0000	0.7750	0.9888
0.400	0	1	0	1544	56	1520	0	56	24	0.0000	0.7000	0.9850
0.420	0	3	0	1549	51	1520	0	51	29	0.0000	0.6315	0.9819
0.450	0	0	0	1552	48	1520	0	48	32	0.0000	0.6000	0.9800
0.480	0	3	0	1557	43	1520	0	43	37	0.0000	0.5315	0.9169
0.500	0	2	0	1560	40	1520	0	40	40	0.0000	0.5000	0.9150
1.000	40	0	0	1600	0	1520	0	0	80	0.0000	0.0000	0.9500

IR2IR 40 9mm Casings from 20 Glocks 2 Firings Each Omit Autocorrelations=1.00

Match Value Threshold	Auto-Correlation	Self Correlation	Cross-Correlation	TOTAL< Threshold	TOTAL> Threshold	TN= CC<	FP= CC>	TP= SC>	FN =SC<	FPR= FP/(FP+TN)	TPR= TP/(TP+FN)	ACC= (TP+TN)/TOTAL
0.100	0	0	0	0	1560	0	1520	40	0	1.0000	1.0000	0.0256
0.180	0	0	2	4	1556	4	1518	40	0	0.9974	1.0000	0.0282
0.190	0	0	14	18	1542	18	1516	40	0	0.9883	1.0000	0.0368
0.200	0	0	22	40	1520	40	1502	40	0	0.9741	1.0000	0.0506
0.210	0	0	69	109	1451	109	1480	40	0	0.9314	1.0000	0.0915
0.220	0	0	114	223	1337	223	1411	40	0	0.8635	1.0000	0.1571
0.230	0	0	168	391	1169	391	1297	40	0	0.1684	1.0000	0.2194
0.240	0	0	252	643	917	643	1129	40	0	0.6371	1.0000	0.3169
0.250	0	0	256	899	661	899	877	40	0	0.1938	1.0000	0.5171
0.260	0	0	282	1181	379	1181	621	40	0	0.3446	1.0000	0.6629
0.270	0	0	188	1369	191	1369	339	40	0	0.1985	1.0000	0.8061
0.300	0	0	15	1517	43	1513	22	36	4	0.0143	0.9500	0.9835
0.310	0	1	5	1523	37	1518	7	35	5	0.0046	0.9315	0.9923
0.330	0	1	0	1529	31	1520	0	31	9	0.0000	0.8815	0.9942
0.340	0	0	0	1529	31	1520	0	31	9	0.0000	0.8815	0.9942
0.370	0	6	0	1538	22	1520	0	22	18	0.0000	0.7750	0.9885
0.390	0	2	0	1543	17	1520	0	11	23	0.0000	0.7125	0.9853
0.400	0	1	0	1544	16	1520	0	16	24	0.0000	0.7000	0.9846
0.420	0	3	0	1549	11	1520	0	11	29	0.0000	0.6315	0.9814
0.430	0	1	0	1550	10	1520	0	10	30	0.0000	0.6250	0.9808
0.450	0	0	0	1552	8	1520	0	8	32	0.0000	0.6000	0.9195
0.460	0	2	0	1554	6	1520	0	6	34	0.0000	0.5750	0.9182
0.480	0	3	0	1557	3	1520	0	3	37	0.0000	0.5315	0.9163
0.490	0	1	0	1558	2	1520	0	2	38	0.0000	0.5250	0.9156
0.500	0	2	0	1560	0	1520	0	0	40	0.0000	0.5000	0.9144
1.000	0	0	0	1560	0	1520	0	0	40	0.0000	0.0000	0.9144

Receiver Operating Characteristic (ROC). ROC analysis was originally developed sixty years ago to model human radar operators who needed to make decisions as to whether each blip on a radar screen was "target signal", "non-target noise", or "unclear - insufficient evidence" [analogous to the choices made by toolmark examiners.] In the past ten years, ROC analysis has been successfully applied to more complex decision-making human tasks such as playing video games and making accurate medical diagnoses based on different assortments of medical imaging and blood test results. ROC analysis has also successfully been used to analyze the performance of systems apart from that of their human operators. That is the application for which we used ROC analysis in the current evaluation of MTW and for which we are advocating its future use.

Our primary objective is to evaluate the absolute performance of a NIBIN-function system that uses infrared emissivity mappings. A secondary objective is to evaluate its relative performance compared to systems using visible light images. Both the IBIS and MTW systems produce a rank ordered list of database images as a guide to the firearms/toolmark examiner. Both systems provide a capability for the examiner to review the database images in rank order. MTW offers emissivity maps layered on visible light images obtained from LED illuminating sensors within the IR camera. In addition to virtual display of a split screen as it would appear in a comparison microscope, the MTW provides a vertical slide that adjusts the image displayed from 100% emissivity map to 100% visible light image, or an intermediate mix. The blended mix of overlaid IR and VL layers applies to all open images, providing additional comparison content to the examiner. This feature can make moving between the MTW display and the comparison microscope an easier transition for the examiner. Some examiners may eventually prefer viewing the 100% IR images; however we expect the primary mode of choice will be variable blending of the IR and VL images.

As the examiner changes the displayed mix from 100% IR to 50-50% to 100% VL, s/he should develop increasing confidence in the MTW matching capability. That confidence building is essential to reducing the number of candidate matches an examiner reviews. Increasing use of the variable blending mode will indicate an increasing level of user acceptance of the MTW. Some decision systems have quality-control modules that will not allow poor quality images to be accepted as target or database entries. Eliminating poor images can decrease both false positive and false negative rates. Two identical devices can show different performance levels based on the strictness of the quality-control module. Analogously, the same system will have different performance depending on the quality of the imagery. A key aspect of ROC analysis is that, with the exception of quality control provisions of a toolmark comparison system, ROC curves do not depend upon system decision policy, but upon the basic distinctiveness and repeatability of the feature templates produced and match values calculated. Therefore this analysis method can be used to compare the effectiveness of different systems' feature extraction and template matching. A non-dimensional number is needed to compare two unrelated systems using a common and basic technical performance measure.

Basic ROC Analysis - Receiver Operating Characteristic (ROC) Curve. Binary Decision Strategy produces either True or False assignment of a database image to one of two categories. In toolmark identification we define them as Sibling or Nonsibling. TPR is the rate of correctly assigning images of sibling cartridges to the Sibling category. TNR is the rate of correctly assigning images of nonsiblings to the Nonsibling category. FPR is the rate of falsely assigning nonsiblings to the Sibling category, and FNR is the rate of falsely assigning images of siblings to the Nonsibling category.

- **Sensitivity** is the proportion of sibling images assigned to the Sibling category. It is the True Positive Rate. The sensitivity is how good the test is at picking out siblings. Sensitivity gives us the proportion of siblings found by the decision strategy, relative to the number of siblings in the database. TPR = TP / (TP + FN).

- **Specificity** is the proportion of nonsibling images assigned to the Nonsibling category. This is synonymous with the True Negative Rate: TNR = TN / (TN + FP). See Figure 2.3.4

Figure 2.3.4:

	Match Pair are Siblings	Match Pair not Siblings	
Match Value $\geq \Phi$	TP	FP	Total $\geq \Phi$
Match Value $< \Phi$	FN	TN	Total $< \Phi$
	Sum is Total #Siblings	Sum is Total - #Siblings	Total Comparisons

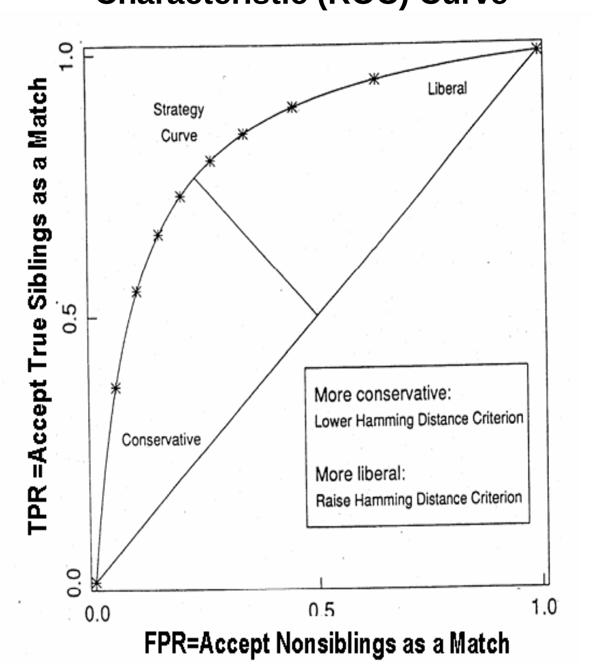
The sensitivity and specificity of a decision system depend on the decision threshold. The position of the threshold will determine the number of true positive, true negatives, false positives and false negatives. In signal detection theory, a receiver operating characteristic (ROC) curve, is a graphical plot of the fraction of true positive decisions (TPR = true positive rate) vs. the fraction of false positive decisions (FPR = false positive rate) as the criterion that defines positive and negative decisions changes. ROC analysis provides tools to select possibly optimal decision strategies and to discard suboptimal ones independently from (and prior to specifying) costs and values associated with incorrect and correct decisions. Most decision systems can be represented as a closeness function by which two candidate measures can be compared. Although we don't know the feature extraction and similarity scoring algorithms used within IBIS workstations, the fact that IBIS produces an ordered list of all database images means that we can represent its decisions using ROC analysis. We consider NIBIN workstations, both IBIS and MTW, to be binary decision systems. For each database image, a decision is made as to whether it is sufficiently similar to a target image that it should be consider a possible match (P or positive class), or it is too dissimilar to be considered a possible match (N or negative class).

There are four possible outcomes from a binary classifier. If the decision is P and the image is truly from a sibling component to the target component, the decision is called a ***true positive*** (TP); however if the image is actually from a nonsibling component, the decision is a ***false positive*** (FP). Conversely, a ***true negative*** decision occurs when the image of a nonsibling component to the target is rated a non-match, and a ***false negative*** decision occurs when the image from a true sibling component to the target is rated a non-match. The similarity measure between each pairing is called its Match Value. Each particular threshold value of the similarity function generates a set of outcomes dependent upon the match value between the target image and each image in the database.

This analysis is performed for every target image and the results plotted as an ROC curve such as show below. The best possible decision strategy would yield a single point in the upper left corner or coordinate (0,1) of the ROC curve below, representing 100% ***sensitivity*** (all true positives are found) and 100% ***specificity*** (no false positives are found). The (0,1) point is also called a ***perfect classification***. Points on the diagonal line connecting left bottom and top right corners represent random guessing on a binary decision, which is likely to be correct 50% of the time.

Receiver Operating Characteristic (ROC) Curve. The diagonal line in the Receiver Operating Characteristic (ROC) Curve, divides ROC space into areas of good or bad decision systems. Points above the diagonal line indicate good decision results, while points below the line indicate wrong decision results (but inverting the decisions can be a good decision strategy). Summary statistics used to interpret ROC curves include: 1) Intercept of ROC curve with the line at 90 degrees to the random guess line. 2) Area between the ROC curve and the random guess line. 3) Area under the ROC curve, or "AUC". The AUC is equal to the probability that the decision system will rank a randomly chosen sibling pair match higher than a randomly chosen nonsibling pair match. 4) d' (pronounced "d-prime") is calculated from the means of the distribution of sibling match values and distribution of nonsibling match values, and their standard deviations, under the assumption that both distributions are normal with the same variance. Under these assumptions, it can be proved that the shape of the ROC depends only on d'. See Figure 2.3.5

Figure 2.3.5: Receiver Operating Characteristic (ROC) Curve



The d' measure is a function of match value distributions that is commonly used to roughly characterize a decision system. As given in the equation below, d' is the ratio of the distance between the means of the sibling and nonsibling distributions divided by the conjoint measure of their standard deviations.

$$d' = \frac{\left\| M_{Siblings} - M_{Nonsiblings} \right\|}{\sqrt{(SD^2_{Siblings} + SD^2_{Nonsiblings})/2}}$$

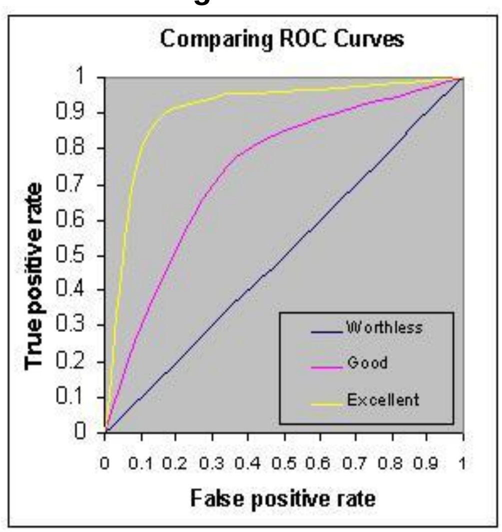
A ROC curve summarizes several aspects of a decision system:

- It shows the tradeoff between sensitivity and specificity (any increase in sensitivity will be accompanied by a decrease in specificity).
- Curves that are more accurate follow left-hand and top borders of the ROC space more closely
- Curves that are less accurate follow the 45-degree diagonal of the ROC space more closely
- The slope of the tangent line at a cut point gives the likelihood ratio (LR) for that value of the test.
- The area under the curve is a measure of test accuracy.

Figure 2.3.6, Comparing ROC Curves, shows three ROC curves that represent excellent, good, and worthless tests plotted on the same graph. The accuracy of the test depends on how well the test separates siblings from nonsiblings for the Target images. Accuracy is measured by the area under the ROC curve. An area of 1 represents a perfect test; an area of 0.5 represents a worthless test. A rough guide for classifying the accuracy of a diagnostic test is the traditional academic point system:

- 0.90 – 1.00 = excellent
- 0.80 - 0.90 = good
- 0.70 - 0.80 = fair
- 0.60 - 0.70 = poor
- 0.50 - 0.60 = fail

Figure 2.3.6:



The area measures **discrimination**, that is, the ability of the test to correctly classify database images. AUC is the percentage of randomly drawn pairs for which the calculated match value correctly determines whether or not they are sibling images.

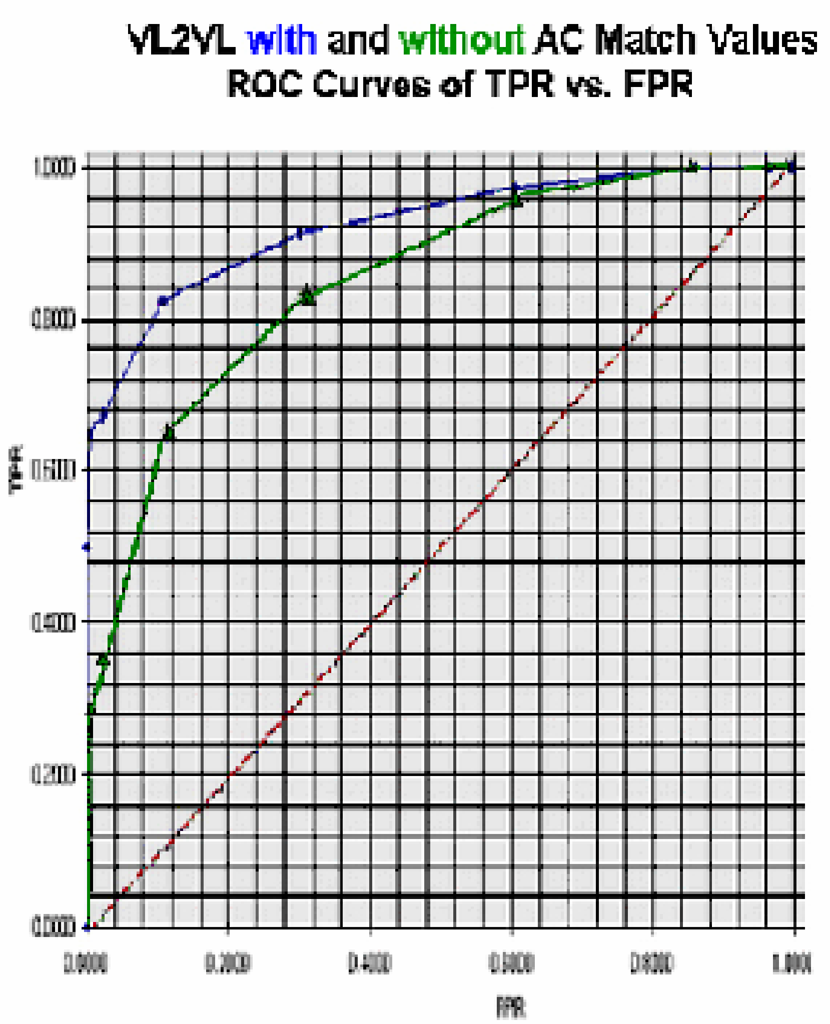
Although the ROC curve gives a good summary of the performance of a one-class classifier, it is difficult to compare systems from their ROC curves; especially if the curves cross. A general comparison can be made from derived measurements such as the Area under the ROC curve (AUC). Larger values indicate a better separation between sibling and nonsibling pairs. In actual decision system operation, a specific threshold must be chosen. That means that only one point of the ROC-curve is used. It can therefore happen that for a specific threshold, a system with lower AUC could produce higher performance than a system with higher AUC whose ROC curve crosses.

Application of ROC Analysis to Toolmark Identification. Match value distributions from ROC analysis are related to the "sufficient agreement" requirement in the AFTE Theory of Identification. According to AFTE: "*Agreement is significant when it exceeds the best agreement demonstrated between tool marks known to have been produced by different tools (*CrossCorrelations in the ROC analysis*) and is consistent with the agreement demonstrated by tool marks known to have been produced by the same tool (*SelfCorrelations in the ROC analysis*). The statement that "sufficient agreement" exists between two tool marks means that the agreement is of a quantity and quality that the likelihood another tool could have made 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."* In addition to being a tool for quantitative evaluation of MTW performance, ROC analysis provides a framework for comparing performance of the current NIBIN system against proposed upgrades. It also lays some ground for developing infrared spectral emissivity mapping as a scientific basis for quantitative toolmark identification in accord with traditional AFTE criteria for identification.

MTW Performance Comparisons. Receiver Operating Characteristic (ROC) analysis was previously used in small scale tests to judge overall quality of MTW toolmark matching based on infrared emissivity mappings, visible light images, and a mixture of the two types. All three studies utilized the same methods to produce feature templates and the same FlashCorrelation® matching engine. The goal of the evaluation was to determine the accuracies of identifications from comparing IR mappings (IR2IR), comparing visible light images (VL2VL), and comparing IR mappings to visible light images (IR2VL). Only breechface marks were compared. While IR2IR matching based on firing pin impression details has shown superior accuracy, the effect of multiple firings between comparison images had not been tested prior to the current project. Also, SED has not yet developed satisfactory feature extraction techniques for visible light firing pin details. Comparing firing pin impressions of sibling cartridge cases finds much greater similarity in IR images than in VL images. This suggests that incorporating firing pin details into the match value algorithm would have improved IR2IR performance but not VL2VL or IR2VL. Some systems have strong quality-control filters that will not allow poor images to be accepted. Eliminating poor images can decrease both false match and false non-match rates. Two identical systems can produce different ROC curves based on the strictness of the quality-control filters.

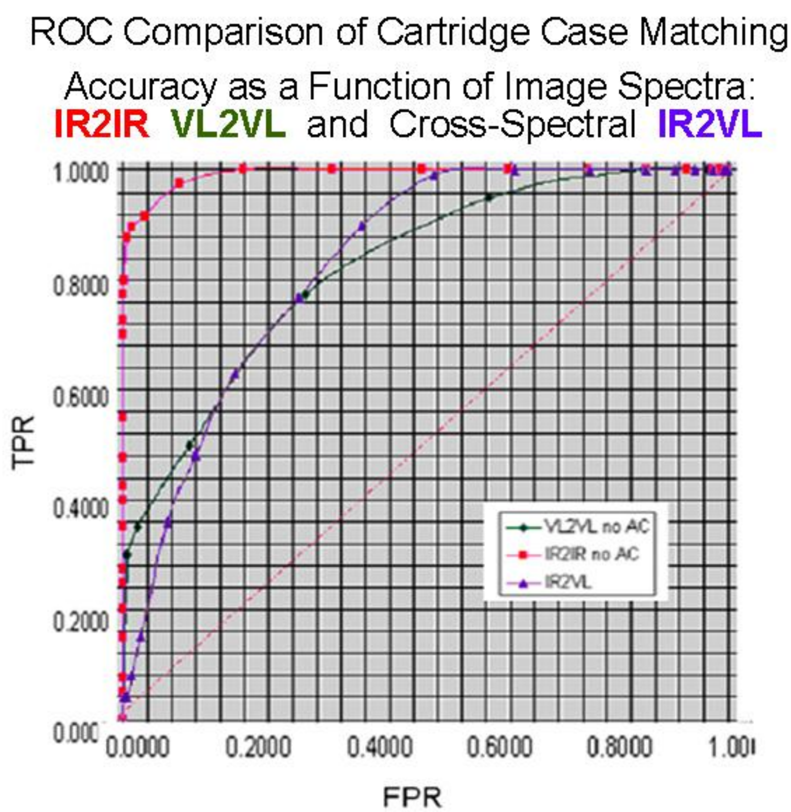
With the exception of arbitrary quality control policies, ROC curves do not depend upon decision system policies, but only upon the basic distinctiveness and repeatability of the match values resulting from the feature extraction and comparison methods of that decision system. Different systems in similar applications can be compared on the basis of these distributions. In particular, an IR2VL decision system can be compared to a VL2VL system (See figure 2.3.7).

Figure 2.3.7



Comparison of ROC Curves – AUC. AUC for IR2IR is above 99%, indicating excellence performance. Without autocorrelations, AUC is lowered by about 0.1%. AUC for VL2VL is 82% with autocorrelations included and 74.6% which is generally considered good to very good performance. At a given match value, inclusion of autocorrelations increases identification rate an average of 2%. ROC analysis assumes independence of matches. In the databases, twenty Glocks were each fired twice. After removing the autocorrelation, our assumption is that each target is related to a single feature template in the database for IR2IR and VL2VL matches. With autocorrelations included, the average match value increases and the number of related comparisons doubles. See Figure 2.3.7, VL2VL with and without AC Match Values ROC Curves of TPR vs. FPR. Accuracy vs. Match Value for IR2VL has a greater dependence on the treatment of sibling IR targets and sibling VL database entries. At a given match value, ID accuracy varies 4% depending on treatment. This is more clearly seen in the expanded portion of the IR Accuracy vs. Match Value graphs for IR2VL. Effect on error rates is shown in the ROC curves where the highest performance is when SC match values are omitted and a slightly lower performance results from omitting AC match values. Including AC and SC match values decreases performance under all decision thresholds; the opposite effect from IR2IR and VL2VL matching. Each IR target is related to its sibling IR target, its corresponding VL image and the VL's sibling image. The increased number of dependent match values reduces the performance assessment under this simple application of ROC analysis. Testing with large populations will reduce the influence of how AC and SC matches are treated in cross-spectral matching. See 2.3.8

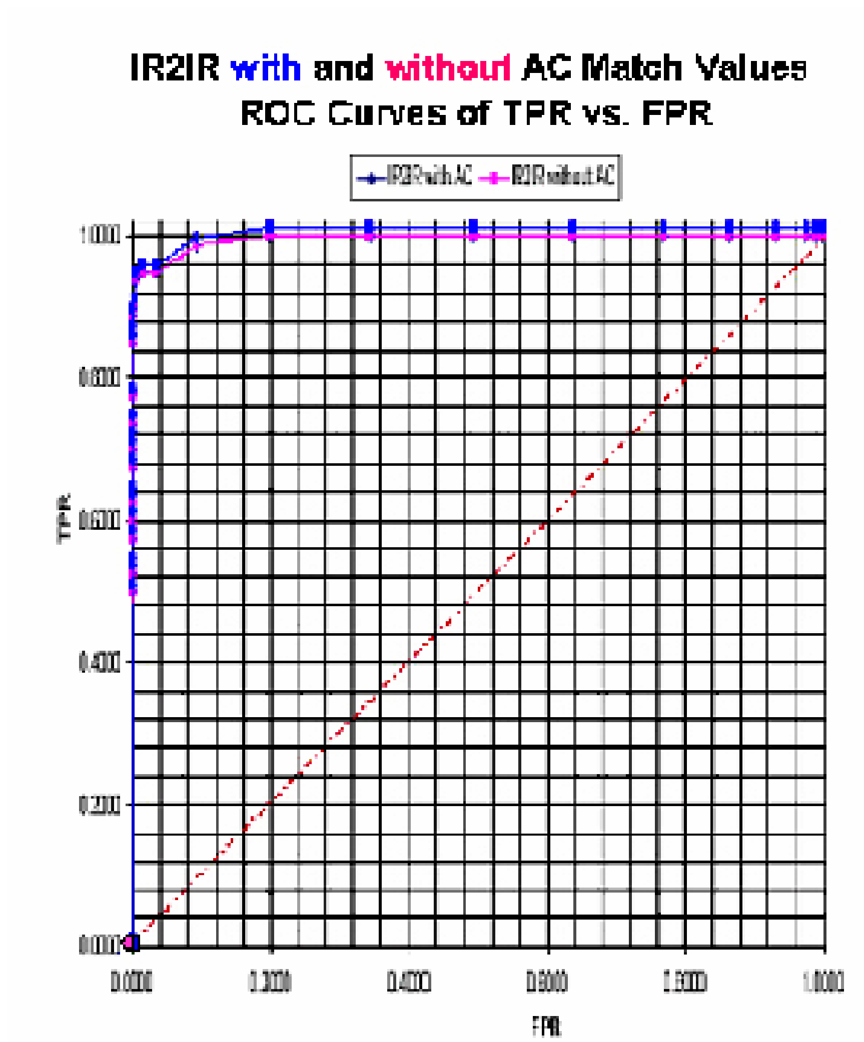
Figure 2.3.8



AUC for IR2VL appears higher than for VL2VL, but the curves cross when plotted together, as shown, and the underlying statistical differences interfere with a clear comparison based on ROC curves. See Figure 2.3.9, IR2VL with and without inclusion of Corresponding VL Match values ROC Curves of TPR vs. FPR. Regardless of the decision threshold, cartridge case identification based on spectral emissivity mapping was significantly more accurate than identification based on visible light imagery. When the rate of false positives was 0%, true positive rates were 50% for VL imagery and 95% for IR. 100% true positive rate was attained at 85% false positive rate for VL and 10% false positives for IR. ROC curve peak, intersection of ROC curve and 90°

line to the random guess diagonal, occurred at 96% true positive and 4% false positive rates for IR; 85% true positive and 16% false positive for VL. Cross-spectral matching of IR target images against a database of VL images was less accurate than IR2IR matching. However, at identification rates of 95% and above, IR2VL was more accurate than VL2VL matching.

Figure 2.3.9



Comparison of ROC Curves – d'. Calculation of d' provides a single quality value for each of the three tested modes of MTW operation. Using the 40 visible light images to produce 1560 match values (after the 40 AC 1.00 values are omitted), of which 40 were self-correlations, d' for VL2VL matches was 1.142. Corresponding matching of infrared images produced IR2IR d'=4.742. That indicates significantly stronger performance for IR2IR compared to VL2VL identification, as also indicated by AUC and Accuracy vs. Match Value considerations. As the d' calculations on the following page indicate, greater separation between sibling and nonsibling match value distributions for IR2IR identification is anticipated from the three time greater difference in mean of the CC and SC distributions. See Figure 2.3.9, Cross Spectral Match Performance Compared to Same Spectrum Matching.

d' for IR2VL was computed by removing the corresponding VL image to each IR target. This is therefore a measure of how well the MTW identified a match between an IR target image and the VL image taken of the IR target's sibling. Omitting VL sibling images instead would simply measure how well the MTW identifies IR and VL images taken of the same cartridge case. That is an important study in itself and, as with studying autocorrelations from multiple independent images of each target and database image, should be performed on large enough databases to support assumptions inherent in the analysis approach used. See Table 2.3.2, Calculation of d’ for IR2IR VL2VL IR2VL, below.

Table 2.3.2:

Calculation of d’ for IR2IR VL2VL IR2VL

d' Calculation for IR2IR 40x39 one Sibling each Target

IR2IR	*SC only*	*CC only*		
			d' = (dif means)/SDfcn	
Mean	0.421853074	0.254961853		
Standard Error	0.007158222	0.000529938	dif means	0.166891
Median	0.417790357	0.257096295	sum of SD**2+SD**2	0.002476
Mode	0.380388703	0.233487606	sum/2	0.001238
Standard Deviation	0.045272574	0.020660803	SQRT	0.035189
Sample Variance	0.002049606	0.000426869	**d' for IR2IR**	**4.742764**
Kurtosis	-0.385658495	-0.067292968		
Skewness	0.749997571	-0.217829028		
Range	0.154361669	0.122463507		
Minimum	0.369745005	0.191368279		
Maximum	0.524106674	0.313831786		
Sum	16.87412296	387.5420161		
Count	40	1520	no autocorrelations included	
Confidence Level(95.0%)	0.014478871	0.001039488		

d' Calculation for VL2VL 40x39 one Sibling each Target

VL2VL	*SC only*	*CC only*		
			d' = (dif means)/SDfcn	
Mean	0.33302508	0.2840339		
Standard Error	0.009362707	0.00033598	dif means	0.048991
Median	0.307	0.283643423	sum of SD**2+SD**2	0.003678
Mode	0.308	#N/A	sum/2	0.001839
Standard Deviation	0.059214956	0.0130989	SQRT	0.042884
Sample Variance	0.003506411	0.000171581	**d' for VL2VL**	**1.142424**
Kurtosis	-0.384392591	0.182837583		
Skewness	1.146404796	0.062337628		
Range	0.173869555	0.089957799		
Minimum	0.278	0.241718976		
Maximum	0.451869555	0.331676775		
Sum	13.32100319	431.731528		
Count	40	1520	no autocorrelations included	
Confidence Level(95.0%)	0.018937862	0.000659033		

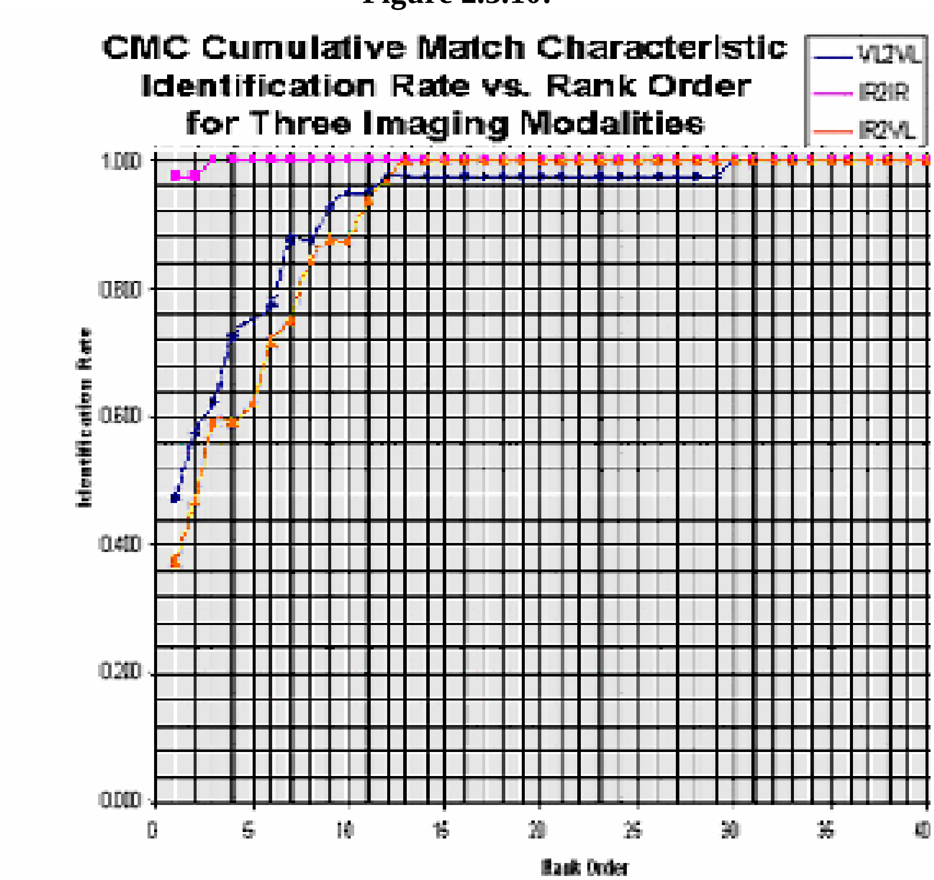
d' Calculation for IR2VL 40x39 Remove Sibling VL Template

IR2VL	*Sibling Only*	CC		
Remove Sibling Targets from Analysis			d' = (dif means)/SDfcn	
Mean	0.318800217	0.304831813		
Standard Error	0.001725372	0.00039411	dif means	0.013968
Median	0.317848551	0.304456475	sum of SD**2+SD**2	0.000355
Mode	0.3067938	0.258574915	sum/2	0.000178
Standard Deviation	0.010912213	0.015365237	SQRT	0.013326
Sample Variance	0.000119076	0.000236091	**d' for IR2VL**	**1.048203**
Kurtosis	0.478246484	-0.074141932		
Skewness	0.715556089	-0.02920661	Cross-Spectral matching has no	
Range	0.04960069	0.09681791	autocorrelations	
Minimum	0.300804939	0.256122643	Corresponding VL removed	
Maximum	0.350405629	0.352940553	Sibling VL matches considered	
Sum	12.75200868	463.3443564		
Count	40	1520		
Confidence Level(95.0%)	0.003489895	0.000773057		

A cost/benefit analysis would indicate the value of using local MTW workstations to sort all fired cartridge cases recovered at a crime scene to determine the number of firearms represented, link apparent sibling cartridge cases in the local database, and select the best quality images for NIBIN searches. Although such sorting was not a focus of our evaluation testing, it offers an additional application for MTW workstations. Accurate firearm count and batch sorting of cartridge cases by firearm occurred only with IR2IR matching.

Cumulative Match Characteristics (CMC) Curves. Another tool for analyzing the performance of a decision system, particularly meaningful for identification systems, is the cumulative match characteristic (CMC) curve. A CMC curve depicts the increase in the identification rate of the system with increase in the rank before which a correct match is found. CMC curves generally plot ID rate vs. the log of Rank order. Because our maximum rank was only 40, a linear scale was used. The smoothed and data point versions of the CMC curves from the three modes of 40x40 matches are shown below. Note that all three modes gave equal performance once Rank Order 30 was reached. The CMC presentation of identification results is equivalent to the standard method of presenting toolmark matching results. See Figure 2.3.10

Figure 2.3.10:



Comparison of CMC Curves. Vertical difference between IR2IR and VL2VL curves is a measure of performance Improvement from using infrared rather than visible light images as the basis for Rank Ordering candidate image matches as a screening filter for identification by a firearms and toolmark examiner. Reviewing the top three Ranked candidate matches yields 100% ID for IR2IR matching, 62% ID for VL2VL matching, and 58% ID for IR2VL matching.

Horizontal difference between curves is a measure of change in Examiner Time (# Rank Ordered Reviews) required to achieve a desired System ID Rate under different imaging modes. IR2IR matching identified 100% of hits at Rank 3; VL2VL matching required review to Rank 30 to find 100%; IR2VL matching required review to Rank 13.

IDR was 38% at Rank #3 and RO was 12 at IDR=95%. The last measure is most directly related to the goal of improved NIBIN performance: to identify 95% of the true matches required reviewing 1 IR image, or 12 VL images. See Figure 11, CMC Cumulative Match Characteristic Identification Rate vs. Rank Order for Three Imaging Modalities.

Developing a Toolmark Comparison Protocol. The MTW performance evaluation applied the following systematic approach to tests involving small databases of 40 IR and 40 corresponding VL images of 20 sibling pairs of 9mm Luger caliber cartridge cases fired in Glock pistols. This generic approach would apply to other sample populations with other characteristics.

1) Select a representative sample of toolmark images from the reference data base. Images could be VL or IR.
2) Number of samples needed to characterize a database depends on several factors including whether or not the database is expected to contain multiple sibling images, and whether or not it contains independent autocorrelation images for some or all of the database. The degree of independence of the sample images affects the predictive power of the analysis results.
3) Select and apply an algorithm that extracts toolmark features to produce a feature template for each sample image.
4) Select and apply a matching algorithm that compares two sample image feature templates and generates a similarity measure. Calculate the similarity measure for all pair-wise combinations of sample image templates.
5) Assign each generated Match Value to AC, SC, or CC distribution. Test that each distribution is normal and that the means of the AC, SC, and CC distributions decrease in that order.
6) Plot the ROC curve. Calculate AUC and d'. Decide whether the resulting decision system quality is acceptable. If not, consider the image quality of the sample images. If the sample images are truly representative of the database, consider the feature extraction algorithm and the distribution of

the number of features per template. Consider the match value distributions: compare AC, SC, and CC variances and the overlap between distributions. Adjust the match value algorithm. Once the overall decision system quality is acceptable, select the best match value threshold to meet policy objectives.

7) Plot the CMC curve. Assess whether the Identification Rate is satisfactory for the number of Rank Order reviews to be performed. Assess whether the Rank Order review policy provides an acceptable Identification Rate.
8) Perform a cost/benefit analysis of proposed upgrades, using ROC methods to predict system performance.

2.3.4. Conclusions

2.3.4.1. **Discussion of findings.** The method of toolmark identification described in this report was developed to evaluate whether and to what extent infrared imaging improved the performance of a toolmark identification system over the use of visible light imaging and if upgraded NIBIN workstations will be required to search legacy databases. A small demonstration test was performed that evaluated comparing IR emissivity mappings against a database of visible light images. The same corresponding IR and VL images were used as in the same-spectrum tests. For IR2VL d' =1.048 which is slightly below the value for VL2VL on overall performance. Below 95% identification rate, □RO averaged 1.4 between CMC curves for IR2VL and VL2VL with IR2VL having the lower performance. However, for identification rates of 95% and above, IR2VL produced the higher performance. To achieve 100% identification based on VL2VL matching required review of an additional 17 images more than IR2VL. Larger scale testing of IR2VL matching is required to confirm these results and address scaling to NIBIN-size databases.

2.3.4.2. **Implications for policy and practice.** A number of studies have found forensic science experts are vulnerable to cognitive and contextual bias that leads them to make erroneous identifications. A limiting factor is the access to searchable databases often produces false positive results which must be checked to ensure the accuracy of the methodology used to by the database to produce the subject sample image. Recognizing those tendencies, courts might be expected to require that forensic evidence admitted in criminal trials be the subject of a reliable scientific methodology that derives accurate findings from the evidence and does not depend on human interpretation that might be flawed by bias, or by the lack of robust standard procedures. Courts have not historically challenged toolmark evidence, but more challenges are occurring. While court testimony is the purview of certified examiners, preliminary screening of database images to find potential matches has been developed as an automated computer system function over the past fifteen years. IBIS, the current automated approach used in the NIBIN system, uses visible light imaging which is known to create possible illumination-induced artifacts. In addition, subjective adjustments to incident lighting, which is necessary to illuminate image details, negatively impacts the reliability of images used by that system to select candidate siblings. As a result of lighting variations, more candidates are selected than might otherwise be necessary. The resulting increased caseload might further lead to erroneous identifications by the examiner.

The goal of a reliable scientific methodology that derives accurate findings from the evidence is to minimize the workload of the examiner without sacrificing accuracy. Currently, the only fully automated forensic matching systems are for cartridge cases. The IBIS system has approximately one million records of cartridge cases although the full database is not generally searched. Third party testing of IBIS with a variety of firearms and ammunition has found negative error rates on the order of 35% and positive error rates on the order of 10%. Although emphasis has publically been put on the need to reduce the workload on examiners by reducing the number of false positives candidate matches presented, the cost impact on the criminal justice system from the high rate of false negatives is arguably greater. False negatives represent the failure of the system to recognize links among evidence it already has processed. Furthermore, decreasing the number of high probability associations produced by the system works to decrease the perceived benefit to local law enforcement on their cost to house the system. Applying better search algorithms to the current NIBIN database might find new matches which could be important clues to solving more cases faster. In prior testing dating back to 1998, FlashCorrelation® matching was demonstrated to identify 80% more hits at 2% than DrugFire – without using infrared images. Using IR mappings and current correlation methods could find cold hits for negligible cost.

2.3.4.3. **Implications for further research.** Recommendations for further research in this area should be described. MTW incorporates computerized numerical control (CNC) for precisely positioning the item under test and the IR imager, allowing a precise sequence of 2D/IR slices to be collected and used to generate a high resolution 3D/IR surface model in less than 30 seconds. Although both IR imager and CNC desktop controller technology is rapidly advancing, the reliability of 3D/IR models produced using current COTS subsystems is an important benchmark to be determined under the proposed effort. Also to be determined are the reliability and precision of striated, shearing, and impression toolmarks extracted from the 3D/IR models.

Additional Performance Evaluations. Alternative algorithms for toolmark feature replication and comparison should be considered. Distribution of a sample database of emissivity mappings could create interest from universities in developing advanced algorithms.

3. References

3.1. DOJ grant report

Final report for DOJ Grant # **2007-DN-BX-K169, titled;** "The Use of Infrared Imaging, a Robust Matching Engine, and Associated Algorithms to Enhance Identification of Both 2D and 3D Impressions"

3.2. Confusion of Subclass and Individual Characteristics

Churchman, J. The Reproduction of Characteristics in Signatures of Cooey Rifles. *R.C.M.P. Gazette*, 11(5), 133-140.

Skolrood, R. "Comparison of Bullets Fired From Consecutively Rifled Cooey .22 Calibre Barrels." *Canadian Society of Forensic Science Journal*, 8(2), 1975, 49-52.

Lomoro, V. "Class Characteristics of 32 SWL, F.I.E. Titanic Revolvers." *AFTE Journal*, 6(2), 18-21.

Lardizabal, P. "Cartridge Case Study of the Heckler and Koch USP." *AFTE Journal*, 27(1), 49-51; Thompson, E. "False Breech Face Id's." *AFTE Journal*, 28(2), 95-96;

Matty, W. "Lorcin L9MM and L380 Pistol Breechface Tool Mark Patterns." *AFTE Journal*, 31(2), 134-137; Lopez, L. and Grew, S. "Consecutively Machined Ruger Bolt Faces." *AFTE Journal*, 32(1), 19-24.

Coffman, B. "Computer Numerical Control (CNC) Production Tooling and Repeatable Characteristics on Ten Remington Model 870 Production Run Breech Bolts." *AFTE Journal*, 35(1), 49-54.

Nies, R. "Anvil Marks of the Ruger MKII Target Pistol – An Example of Subclass Characteristics." *AFTE Journal*, 35(1), 75-78.

3.3. Studies Concerning Tool Surface Changes with Firing Order:

Shem, R. and Striupaitis, P. "Comparison of 501 Consecutively Fired Bullets and Cartridge Cases From a .25 Caliber Raven Pistol." *AFTE Journal*, 15(3), 109-112.

Kirby, S. "Comparison of 900 Consecutively Fired Bullets and Cartridge Cases From a .455 Caliber S&W Revolver." *AFTE Journal*, 15(3), 113-126.

Ogihara, Y., *et al*. "Comparison of 5000 Consecutively Fired Bullets and Cartridge Cases From a .45 Caliber M1911A1 Pistol." *AFTE Journal*, 15(3), July 1983, 127-140.

Schecter, B., *et al*. "Extended Firing of a Galil Assault Rifle." *AFTE Journal*, 24(1), 37-45.
Dolleing, B. "Comparison of 4000 Consecutively Fired, Steel-Jacketed Bullets." Presented at the 53rd Annual AAFS Seminar, Seattle, WA, February 2001 .

Hall, J. "Consecutive Cuts by Bolt Cutters and Their Effect on Identification." *AFTE Journal*, 24(3), 260-272 .

3.4. Low Percentage of Matching Striations in Components having a Common Origin and High Percentage of Matching Striations when Origins are Different

Biasotti, Alfred A; Study of the Individual Characteristics of Fired Bullets, J. Forensic Sci 34, 44 (1959)

Butcher, S. and Pugh, D. "A Study of Marks Made by Bolt Cutters." *JFSS*, 15(2).

3.5. Influence of Total Rounds Fired on Changes to Toolmarks

Shem, R. and Striupaitis, P. "Comparison of 501 Consecutively Fired Bullets and Cartridge Cases From a .25 Caliber Raven Pistol." *AFTE Journal*, 15(3), 109-112.

Kirby, S. "Comparison of 900 Consecutively Fired Bullets and Cartridge Cases From a .455 Caliber S&W Revolver." *AFTE Journal*, 15(3), 113-126.

Ogihara, Y., *et al*. "Comparison of 5000 Consecutively Fired Bullets and Cartridge Cases From a .45 Caliber M1911A1 Pistol." *AFTE Journal*, 15(3), July 1983, 127-140.

Schecter, B., *et al*. "Extended Firing of a Galil Assault Rifle." *AFTE Journal*, 24(1), 37-45.

Dolleing, B. "Comparison of 4000 Consecutively Fired, Steel-Jacketed Bullets." Presented at the 53rd Annual AAFS Seminar, Seattle, WA, February 2001 .

3.6. Concepts of Uniqueness vs. Identity

Stoney, D. "What Made Us Think We Could Individualize Using Statistics?" *JFSS*. 31(2), April/June, 1991

3.7. The Role of Statistics in Toolmark Identification, and Utility of CMS Guidelines

Biasotti, A., "A Statistical Study of the Individual Characteristics of Fired Bullets." *JFS*, 4(1), Jan. 1959.

Brackett, J. "A Study of Idealized Striated Marks and Their Comparison Using Models." JFSS, 10(1), Jan. 1970, 27-56.

Blackwell, R. and Framan, E., "Automated Firearms Identification Systems AFIDS: Phase I." *AFTE Journal*, 12(4), Oct. 1980, 11-37.

Deinet, W. Studies of Models of Striated Marks Generated by Random Processes. *JFSS*, 26(1), Jan. 1981, 35-50.

Uchiyama, T., "A Criterion for Land Mark Identification." *AFTE Journal*, 20(3), July 1988, 236-251.

Miller, J. and Neel, M. "Criteria for Identification of Toolmarks Part III Supporting the Conclusion." *AFTE Journal*, 36(1), 7-38.

Stone, R. "How Unique are Impressed Toolmarks?" *AFTE Journal*, 35(4), 376-383.

Collins, E. "How 'Unique' are Impressed Toolmarks? An Empirical Study of 20 Worn Hammer Faces." Presented at the CAC Semi-Annual Seminar, May 2005, Oakland, CA.

4. **Dissemination of Research Findings:**

SED used a variety of methods to provide the firearms and toolmark community with current status and actions associated with this Grant. This included attendance at selected conferences, do briefings to communities of interest, providing status updates on the SED website, and producing a video showing the technology. These actions are shown below;

4.1.1 Attendance and booth displays at selected forensic Training Seminars. The technology at the booth contained information on the grant and selected demonstrations of the technology. The seminars attended included:

- Association of Firearm and Tool Mark Examiners (AFTE) 2010 Annual Training Seminar in Henderson, NV
- The 2010 International Association for Identification Annual Meeting in Spokane, WA
- American Academy of Forensic Sciences Annual Conference in February 2011 in Chicago, IL
- Association of Firearm and Tool Mark Examiners (AFTE) 2011 Annual Training Seminar in Chicago, IL

4.1.2 Presentations to selected forensic audiences

- Dr. Prokoski and Mr. Jack Dillon presented 30 minute presentations talking about the status and technology associated with this grant to AFTE 2010
- Stan Derr presented a poster at the DOJ Sponsored Impression Evidence Symposium in the Tampa Bay area
- Stan Derr and Jack Dillon presented comprehensive briefings to the University of Central Oklahoma's new Forensic Science Institute
- Stan Derr and Jack Dillon made presentations to the State of Oklahoma's Forensic Science Laboratory.
- Stan Derr presented a presentation to the Scientific Working Group for Firearms and Toolmarks (SWGGUN) in Columbus, OH

4.1.3 Provide on-going information via the SED website, www.sedllc.com. This includes periodic updates, announcement of planned conferences attendance, and preparation of a 10 minute video supporting the grant work. The video was updated several times and is now available for viewing on the website. Links to face book and twitter were added to increase information dissemination.

Appendix A List of Acronyms:

2D/IR	Two-dimensional infrared
3D/IR	Three-dimensional infrared
AC	Auto-correlation
ACP	Automatic Colt Pistol
AFTE	Association of Firearm and Tool Mark Examiners
AR-15	ArmaLite AR-15, a selective fire assault rifle adopted by United States armed forces as the M16 rifle. Because of financial problems, ArmaLite sold the AR-15/M16 design to Colt. AR-15 is a registered trade mark of Colt and refers only to the semi-automatic version of the rifle.
AUC	Area under the ROC curve
CC	Cross-correlation or Cartridge case; determine which from context
CMC	Cumulative match characteristics
CMS	Consecutive matching striae
CNC	Computerized numerical control
COTS	Commercially available off-the-shelf
DNA	Deoxyribonucleic acid. Deoxyribonucleic acid is a nucleic acid that contains the genetic instructions used in the development and functioning of all known living organisms (with the exception of RNA viruses).
DOJ	Department of Justice
FNR	False negative rate
FPA	Firing pin aperture
FPR	False positive rate
FPI	Firing pin impression
IBIS	Integrated Ballistics Identification System
IED	Improvised explosive device

IR	Infrared
IR2IR	Comparison of infrared image to infrared image
IR2VL	Comparison of infrared image to visible light image
LED	Light emitting diode
MTW	Mikos Forensic Toolmark Workstation
NAS	National Academy of Sciences (NAS)
NIBIN	National Integrated Ballistic Information Network
NIJ	National Institute of Justice
NIST	National Institute of Standards and Technology
NRC	National Research Council
ROC	Receiver operating characteristics
SC	Self-correlation
SED	SED Technology LLC
TNR	True negative rate
TPR	True positive rate
VL	Visible light
VL2VL	Comparison of visible light image to visible light image

Appendix B Glossary

Breechface	That part of the breechblock or breech bolt which is against the head of the cartridge or shotshell during firing
Bullet catcher	Steel tube approximately six feet long that is filled with Kevlar which slows then stops bullets so the bullets can be retrieved in near pristine condition.
Chamber	The rear part of the barrel bore that has been formed to accept a specific cartridge. Revolver cylinders are multi-chambered.
Comparison microscope	A device used to analyze side-by-side specimens. It consists of two microscopes connected by an optical bridge, which results in a split view enabling two separate objects to be viewed simultaneously; with proper and adjustable lighting, some are capable of rendering a 2D view of the 3D surfaces in a manner similar to that of the conventional comparison microscope.
Constrained Firing Pin	A firing pin that does not rotate.
CNC controller	Used with the Mikos Forensic Toolmark Workstation for positioning of samples for imaging
Drag mark	A microscopic mark having longitudinal striations produced by a firing pin moving laterally across a primer surface
Drugfire	First computer-based imaging system used in support of firearms examiners, developed by the FBI laboratory in 1989. Now superseded by the NIBIN system.
Emissivity	A measure of how efficiently a surface radiates heat.
Ejector	A portion of a firearm's mechanism which ejects or expels cartridges or cartridge cases from a firearm.
Extended focus images	Construct of a single image in which all portions are in focus, starting with a stack of images that each have certain areas in focus.
Extractor	A mechanism for withdrawing the cartridge or cartridge case from the chamber

Firing Pin Deflection	Occurs in firearms in which firing pin impressions vary in location due to the range of manufacturing design tolerances.
FlashCorrelation®	Advanced generalized matching engine especially designed for optical processing.
Impressions	Surface contour variations on an object caused by applying force without motion, or where the motion is approximately perpendicular to the plane being marked.
IR images	Images resulting from emissivity and thermal variations across an imaged surface.
Laser profilometry	Laser enabled metrological measurement of surface structure, roughness, etc.
Metrology	All theoretical and practical aspects of measurement.
Match Value	Quantitative result from correlation of two images.
NIBIN/IBIS	National Integrated Ballistic Information Network/Integrated Ballistics Identification System
Pixel	An acronym for picture element. The individual elements in a digitized image array.
Photometric stereo	**A** technique in computer vision for estimating the surface normals of objects by observing that object under different lighting conditions
Polygonal rifling	**A** type of gun barrel rifling where the traditional lands and grooves are replaced by "hills and valleys" in a rounded polygonal pattern, most often hexagon or octagon.
RGB	RGB refers to the three colors displayed on computer monitors, red, green, and blue. These three colors are combined to create the appearance of the rest of the spectrum.
ROC	Receiver Operating Characteristics (ROC) analysis provides a systematic method for quantitatively evaluating the performance of a decision-making system without knowing the decision algorithms involved. It provides

	methods for segmenting a decision process into two or more components, and separately analyzing each.
Shearing	Refers to the occurrence of a shear strain, which is a deformation of a material substance in which parallel internal surfaces slide past one another.
Sibling	A sample (bullet or cartridge case) created by the same parent (firearm) device; bullets or cartridge cases fired from the same firearm.
SKS Eastern bloc	SKS is a Soviet semi-automatic rifle chambered for the 7.62x39mm cartridge, designed in 1945 by Sergei Gavrilovich Simonov. SKS is an acronym for Samozaryadnyj Karabin Sistemy Simonova.
Sleeves	Digital 3D surface models
Spectral emissivity of a surface	A measure of how efficiently the surface emits heat measured at a particular wavelength.
Striations (striae)	Contour variations, generally microscopic, on the surface of an object caused by a combination of force and motion where the motion is approximately parallel to the plane being marked.
Stylus profilometer	Traditional tool to determine surface characterization however seldom used because it can be destructive to the surface.