The author(s) shown below used Federal funding provided by the U.S. Department of Justice to prepare the following resource:

Document Title:	**A Metrology Foundation for 3D Ballistic Imaging**
Author(s):	**Michael T. Stocker**
Document Number:	**303986**
Date Received:	**December 2021**
Award Number:	**2016-DNR-6257-2**

Final Technical Report – Cover Page

Federal Agency and Organization Element: Department of Justice, Office of Justice Programs

Federal Award Number: Interagency Agreement #1608-683-68

Project Title: A Metrology Foundation for 3D Ballistic Imaging

PD/PI Name, Title, Contact Info: Michael T. Stocker, Physical Scientist, National Institute of Standards and Technology, 100 Bureau Drive Mailstop 8212, Gaithersburg, MD 20899

Submission Date: September 15, 2021

DUNS: EIN:

Recipient Organization: National Institute of Standards and Technology, 100 Bureau Drive Mailstop 8212, Gaithersburg, MD 20899

Recipient Identifying Number: N/A

Project Grant Period: Start: 01/01/2017 End: 06/30/19

Grant Award Amount: $534,500.00

Signature of Submitting Official: Michael T. Stocker

Table of Contents

Abbreviations

AAFS – American Academy of Forensic Science
AFTE – Association of Firearm and Toolmark Examiners
ANSI – American National Standards Institute
ANAB – ANSI National Accreditation Board
ASB – AAFS Academy Standards Board (ASB)
A2LA – American Association for Laboratory Accreditation
FATM – Firearm and Toolmark
FBI – Federal Bureau of Investigation
FTU – Firearm and Toolmark Unit
ISO – International Organization for Standardization
LCM – Light Comparison Microscope
NA – Numerical Aperture
NIST – National Institute of Standards and Technology
NPL – National Physical Laboratory
OSAC – Organization of Scientific Area Committees
OTF – Optical Transfer Function
SPC – Statistical Process Control
TWG3D2T – Technical Working Group for 3D Toolmark Technologies
VCM – Virtual Comparison Microscopy

Keywords: 3D ballistic imaging , firearm and toolmark analysis, metrological characteristics, quality control, reference materials, topography, virtual comparison microscopy

1.0 Statement of Problem

This report describes the development of methods and reference artifacts to facilitate the introduction and quality assurance of three-dimensional (3D) toolmark topography measurement in forensic labs. The forensic discipline of firearm and toolmark examination has been around and largely unchanged for the last 100 years. The Association of Firearms and Tool Mark Examiners (AFTE) defines a tool as "An object used to gain mechanical advantage. Also thought of as the harder of two objects which when brought into contact with each other, results in the softer one being marked" [1]. Firearms are considered tools in that their machined surfaces impart toolmarks to the cartridge case and bullet during firing. The current toolmark examination process involves a trained firearm examiner comparing the microscopic toolmarks on two samples using a specialized light comparison microscope (LCM) that provides a simultaneous view of the two samples. The goal of this evaluation is to determine whether toolmarks on the compared surfaces were made by the same firearm or tool. This evaluation is currently subjective. In 2009, the National Academies issued a report [2] expressing concern on the lack of objectivity in various forensic pattern evidence disciplines, including visual toolmark comparison.

The report accelerated research into methods for objective toolmark comparisons and quantitative measures for the weight of the evidence. One promising development identified in the 2009 report, as well as the 2008 National Academies report on ballistic imaging, is 3D ballistic imaging. 3D ballistic imaging yields a representation of the 3D surface topography of the toolmarks instead of a 2D reflectance microscopy image (from LCM) that is affected by lighting and variations in color and reflectivity. 3D ballistic imaging promises a higher reproducibility of the actual toolmark topography data of interest which may benefit both visual and objective numerical comparisons. Additional 3D surface topography related research has included ballistics reference databases [3], identification algorithms [4-12], and characterization of the weight of evidence [13-17], among many others.

The field of surface texture analysis has seen the introduction of a variety of three-dimensional (3D) optical measurement methods and instruments possibly suitable for firearm and toolmark (FATM) applications. Currently, 3D ballistic imaging is used in ballistic search [18, 19] during the investigations.. Here 2D and 3D toolmark images are submitted to a database, after which the system returns a ranked list of possible matches that are then visually compared with the submitted images to generate possible investigative leads. Currently, potential matches require confirmation by an examiner using traditional LCM of the actual samples. As research progresses, the goal is for 3D imaging and data analysis to mature to where it could eventually supplement an examiner's conclusion with an objective measure of toolmark similarity and weight of evidence, such as an applicable error rate or likelihood ratio. In the interim, 3D surface topography measurement has enabled another useful tool for firearm examiners, Virtual Comparison Microscopy (VCM) [20]. VCM is a new approach for conducting visual examinations of ballistic evidence. It relies on measuring the actual 3D surface topographies of the samples being evaluated and on visually comparing these 3D sample topographies with a special viewing and comparison software package, enabling visual comparisons in a manner similar to LCM.

Successful implementation of 3D surface topography in FATM analysis requires a sound metrological foundation to assure validity of the acquired data. Many of the respective requirements are currently not in place. In this context, it is important to note that the characterization of optical 3D topography measurement methods for general purpose surface texture analysis is a rapidly evolving field with unresolved challenges.

2.0 Rationale for the Research

Forensic laboratories considering optical 3D surface topography measurement for objective comparison and VCM face challenges with respect to cost, training, and quality assurance. The primary aim of this research was to address quality assurance issues that currently represent barriers to entry for these methodologies. 3D optical surface topography microscopes are complicated instruments and evaluating and validating their performance is not trivial. Understanding microscope performance is important in the selection, validation, and monitoring of these instruments. The two key research components of this project are: 1) the development of prototype physical reference artifacts for key metrological characteristics of 3D surface topography microscopes, and 2) establishing guidance on how to implement a quality assurance system to monitor and ensure the quality of topography data acquired for FATM analysis. The elements of the research presented below work towards establishing a necessary metrology foundation for implementing 3D methodologies in forensic case work.

3.0 Project Design

The two-year project began with a landscape survey to assess the state of the art of 3D surface topography metrology and its application to FATM. There were four components to this survey. First, a survey was conducted of the different optical measurement methods incorporated in 3D surface topography instruments. Next, major commercial vendors were identified that offered instruments capable of being utilized in FATM analysis. This effort covered both general-purpose optical surface topography microscopes and systems designed for FATM examination. Next, various documentary standards efforts were reviewed that offer guidance to practitioners on implementing these types of technology. These standards efforts covered both general-purpose applications and FATM examination. Finally, a survey was conducted of commercial physical reference artifacts available for performance evaluation, calibration, and quality control of 3D surface topography measurement instruments and methods.

With the survey completed, work commenced on three tasks. The first task was to identify key metrology characteristics to help guide practitioners regarding what characteristics of 3D surface topography microscopes should be considered when purchasing and maintaining (quality control) these instruments for FATM applications. A metrology characteristic is defined in ISO 25178-600 as a characteristic of measuring equipment which may influence the results of measurement. Next, work began on the core component of the research, to explore solutions for physical reference artifacts that can be used in a forensic laboratory to characterize and check the key metrology characteristics. Here our focus was on prototypes that can be measured by any instrument considered a candidate for FATM applications. Last, quality assurance protocols were developed for setting up a quality assurance system in a forensic laboratory to ensure that surface topography measurements remain in statistical process control (SPC). These quality assurance protocols were developed in collaboration with the Federal Bureau of Investigation (FBI) Firearm and Toolmark Unit (FTU), which has been actively researching the utilization of 3D optical surface topography measurement methods for FATM analysis.

4.0 Landscape Survey of 3D Surface Topography Measurement Methods for FATM Applications

4.1 Summary of Optical Measurement Methods

A survey of 3D optical surface topography instruments currently marketed for FATM analysis was conducted to understand the similarities and differences in capabilities and performance. Three categories of methods, described in ISO 25178 Part 6 [21], are capable of measuring surface topography characteristics: line profiling, areal topography and area-integrating methods. Line profiling provides a

2D cross section profile of the surface topography. In practice, its application is limited to the analysis of striated toolmarks such as those found on the land engraved areas of bullets. Area integrating methods provide some areal summary characteristic of the surface topography, such as average roughness, and are not suitable to provide 3D image data of surface topography. Of the methods, instruments falling into the areal topography family are finding use in the application of FATM analysis. These instruments are predominantly based on one of four optical methods for generating areal surface topography data: confocal microscopy, interferometry-based microscopy, focus-variation microscopy, and photometric stereo. Leach provides a comprehensive review of these techniques [22]. In the context of FATM analysis, Vorburger et al. [10] provide a relevant summary. We provide a brief description of each method below and in section 5.0, summarize some of the limitations of the methods that served as guidance in the design of the prototype reference artifacts.

All four methods utilize an imaging detector to capture the scene created by the imaging optics of the instrument. The first three methods (focus-variation microscopy, confocal microscopy, and interferometry-based microscopy) all involve moving the sample through focus and identify pixels with a specific surface height relative to the objective. However, each is inherently different in design and manner in which the through-focus data is analyzed to create the surface topography. Confocal microscopy relies on the use of physical pinholes within the microscope that mask light reflected from a surface area that is not in focus. Thus light reflected by the surface will only generate a strong intensity signal on a pixel of the imaging detector when the surface area corresponding to that pixel is in focus. Interferometry-based instruments work on the principle of the optical interference of light, in which the light within the microscope is split into a measurement beam (reflected off the surface to be measured) and a reference beam (reflected off a reference surface in the microscope) and are recombined to form a pattern of optical fringes. The through-focus behavior of these fringe patterns (e.g., fringe position and contrast) can be analyzed to generate surface topography. The focus-variation method is a simpler optical approach, as it is based on a conventional bright field microscope design. This method measures topography by identifying at each height of the sample the pixels that are in focus, based on evaluating the change in the sharpness (contrast) of a set of near-neighbor pixels in the image stack. The fourth approach, photometric stereo, is inherently different than the first three. It does not involve acquiring through-focus data. Instead, photometric stereo relies on acquiring images of the sample under different known illumination conditions, typically generated by activating different point lights. The differences in the images are then analyzed to estimate the surface topography.

4.2 Instrument Vendor Summary

3D optical surface topography instruments marketed for use in FATM analysis can be further divided into two categories. First, there are vendors who offer general purpose 3D microscopes that are used in a variety of industries (e.g., semiconductor, automotive, aerospace, energy, etc.). These microscopes generally have the advantage of a selection of magnifications and large measurement volumes. Accordingly, these instruments can accommodate a variety of sample holders, including the ability to integrate a pallet holder for the automatic sequential measurement of multiple samples (e.g., multiple cartridge cases). Some general-purpose microscope manufacturers have begun creating software geared towards FATM analysis applications. Second, some instrument vendors have designed microscopes specifically tailored for FATM analysis applications. Typically, these microscopes are focused on ballistic samples (spent cartridge cases and bullets). One advantage of these microscopes over most of the general-purpose microscopes is that they come packaged with FATM-relevant software for carrying out tasks such as VCM, objective numerical comparison, and, in some cases, weight of evidence estimations. Another advantage is that they typically have fixtures to facilitate measurement of ballistic samples. Some instruments also have a rotary axis to facilitate the measurement of bullets. Regardless of category, some microscopes are single-mode (e.g., a microscope that only uses focus variation), some allow for the selection of a measurement mode, and some use two methods simultaneously to generate the topography. Table 1 below contains examples of vendors[1] who offer optical 3D surface topography microscopes considered candidates for FATM analysis applications.

General Purpose Microscope Vendors	Topography Method
Alicona (Bruker)	Focus Variation, Interferometry
Leica Microsystems	Confocal
Nanofocus	Confocal
Sensofar	Confocal, Interferometry, Focus Variation
Zeiss	Confocal
Zygo	Interferometry

a)

Application-specific Microscope Vendors	Topography Method
Balscan	Focus Variation, Photometric Stereo
Cadre Forensics	Photometric Stereo
Evofinder	Focus Variation, Photometric Stereo
Pyramidal Technologies	Interferometry
Ultra Electronics Forensics Technology	Photometric Stereo

b)

Table 1 – Partial list of vendors who offer optical 3D surface topography microscopes that may be suitable for FATM analysis. a) list of general-purpose microscope vendors, b) list of FATM application-specific microscope vendors.

4.3 Standardization Summary

[1] Table 1 represents an incomplete list of vendors that the authors were aware of at the time of drafting this report. The manufacturers listed in Table 1 may offer microscopes based on other topography methods than indicated in the "Topography Method" column. Certain commercial equipment, instruments, or materials (or suppliers, or software, ...) are identified in this paper to foster understanding. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose.

A survey of documentary standards and best practice guides relevant to the optical measurement and use of 3D surface topography for FATM analysis was conducted to understand the current state of guidance and standardization. Two general categories for documentary standards were identified and are described below: 1) documentary standards for optical 3D surface topography measurement and instrumentation and 2) documentary standards for establishing a laboratory's qualifications and competence as a forensic science service provider.

4.3.1 Description and Use of 3D Surface Topography Methods

Documentary standards efforts regarding the characterization and use of 3D surface topography instruments can be broken into two sub-categories. The first consists of efforts to standardize the specification, measurement, and analysis of areal surface texture, being carried out within the International Organization for Standardization (ISO) Technical Committee 213 and national standards organizations, such as the ASME B46 committee. Here areal surface texture is a term used that is to a large extent equivalent to 3D surface topography. The second are efforts to standardize and offer guidance regarding the acquisition and application of 3D surface topography in the forensic discipline of FATM analysis. This is being carried out within two complementary bodies: 1) the Organization of Scientific Area Committees (OSAC) [23] and 2) the Technical Working Group for 3D Toolmark Technologies (TWG3D2T) [24].

ISO is working on a series of areal surface topography standards, ISO 25178. These standards address items such as terms/definitions, surface texture parameters, classification of methods for measuring surface texture, material measures (physical reference standards) for calibrating instruments, and metrological characteristics of areal surface topography instruments. This ISO series offers standardization for the general field of areal surface topography, but not specific language and guidance tailored to FATM analysis. These standards are written for experts in (surface) metrology, making interpretation difficult for FATM practitioners. These are barriers for successful implementation of these standards in a forensic laboratory setting.

Additional documentary standard efforts (OSAC and TWG3D2T) are underway aimed at providing guidance that forensic practitioners need to successfully implement 3D surface topography acquisition and analysis. The OSAC is an initiative to strengthen forensic science sponsored by the National Institute of Standards and Technology (NIST) and the Department of Justice (DOJ). It is a collaborative effort of over 500 forensic science practitioners and researchers with the goal to develop consensus-based forensic science specific documentary standards. Its scope covers all major forensic science disciplines, including FATM analysis. The FATM subcommittee divides their efforts into six task groups: 1) Firearms Examination, 2) Novel Technology, 3) Uncertainty of Measurement, 4) Training, 5) Terminology, and 6) Criteria for Identification. The Novel Technology task group is addressing standards for emerging technology (e.g., 3D imaging and comparison algorithms). This task group has worked on four standards; 1) Standard for Topography Comparison Software for Firearm and Toolmark Analysis, 2) Standard for Implementation of 3D Technologies in Forensic Laboratories for Firearm and Toolmark Analysis, 3) Standard for 3D Measurement Systems and Measurement Quality Control for Firearm and Toolmark Analysis, and 4) Virtual Microscopy. The first three have been submitted to the American Academy of Forensic Science (AAFS) Standards Board (ASB) for approval, and the last one is in development. The three drafts sent to ASB can be viewed on OSAC's Firearm and Toolmark Subcommittee website [25}.

The OSAC documents provide a high-level summary of requirements but lack details necessary for their actual implementation. For example, an OSAC document may state that a forensic laboratory using an optical 3D surface topography microscope must have it calibrated and demonstrate measurement traceability for the 3 axes of the instrument, but will stop short of providing enough detail enabling a practitioner to fully and successfully carry out the measurements and analyses required. This is addressed by the complementary efforts of the TWG3D2T. The purpose of the TWG3D2T is to provide guidance and recommendations (e.g., in the form of best practice guides) to the FATM community on instrument performance assessment, virtual comparison microscopy, measuring practices, SOPs, and quality assurance at a level that will enable practitioners to successfully implement 3D technology. For example, in addition to requiring that the 3 axes of the instrument be calibrated and the laboratory demonstrate measurement traceability, additional detail would be given suggesting what type of material measures (physical reference artifacts) should be used and how to perform and analyze the measurements. This standards group consists of the following seven subcommittees: 1) Membership, 2) Webmaster, 3) OpenFMC/Interoperability, 4) Quality Assurance, 5) Statistics, 6) By-laws, and 7) Deployment Assessment/Validation. The efforts of the TWG3D2T are in the beginning stages, formalizing its structure and mission in the fall of 2020.

4.3.2 Description of Documentary Standards for Forensic Laboratory Accreditation

Forensic laboratories offer a variety of services spanning the different forensic disciplines. A laboratory may choose to pursue accreditation in order to demonstrate their ability to carry out specific types of testing, measurement, and calibration. Most forensic disciplines have come under some level of scrutiny since the 2009 NAS report [2], prompting many laboratories to seek formal accreditation. Forensic laboratories can seek accreditation to two different standards: 1) ISO 17020 [26]and 2) ISO 17025 [27]. A laboratory may seek accreditation to one or both of these standards, depending on the services they provide. A laboratory that performs inspections or comparisons between gathered evidence, such as comparing fingerprints from a subject with those found at a crime scene, may seek accreditation to ISO 17020. In this case, a forensic scientist would utilize professional judgement to make a determination regarding whether or not the samples are from the same source. ISO 17025 is the relevant standard when forensic laboratories are using measurement instruments to quantify materials and samples (e.g., measurements are performed). In this case, measurement uncertainty would need to be determined and measurement traceability demonstrated for measurements from these instruments.

4.4 Commercially Available Material Measures

A survey of physical artifacts for the calibration and validation of 3D surface topography instruments was conducted. In ISO 25178 Part 70 [28], physical reference artifacts that are used to characterize, calibrate, and verify metrology characteristics of optical 3D topography microscopes are referred to as “material measures”. There are numerous commercially available material measures that could be used for calibration, validation, and quality control. These commercially available material measures are available for characterizing the various metrological characteristics defined in ISO 25178 Part 600 or for checking the associated performance parameters [29], including but not limited to amplification coefficient, linearity deviation, measurement noise, topographic spatial resolution, and maximum measurable local slope. Our review of commercially available material measures did not identify artifacts that could be applied to all candidate microscopes marketed for FATM analysis.

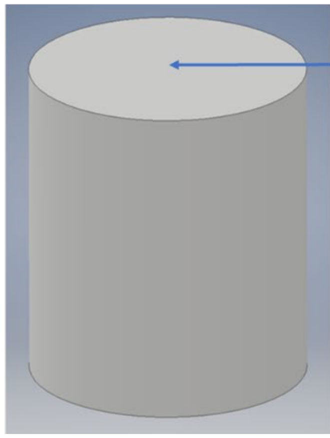
For example, various vendors offer chrome-on-glass type artifacts that contain binary features useful for evaluating different metrology characteristics (e.g., lateral scale, linearity, and resolution). The artifact surfaces have a mirror-like finish and the feature steps are discrete and small. Chrome layers on these types of targets are very thin, often less than 100 nm. Several instruments marketed for FATM measurements, e.g., those based on focus variation and photometric stereo, cannot measure these smooth binary features. Some artifacts are made via standard silicon processes, but this also creates surfaces that are extremely smooth and not measurable on all instruments.

5.0 Artifact Design for Measurability

Above, optical methods, commercial microscopes, and commercially available material measures were briefly summarized. Considering these three together, with respect to measurability, we identified key design factors of reference artifacts that would be considered in exploring designs of material measures for use on surface topography microscopes used in FATM analysis. Achieving measurability across the spectrum of surface topography microscopes available for FATM analysis applications was the most challenging aspect of designing and fabricating artifact prototypes. The strengths and weaknesses of the optical methods discussed in 4.1 had to be factored into the designs. Five key factors were considered in our artifact designs; 1) form factor and size of the artifact so that it can be measured on FATM application-specific microscopes, 2) presence of dimensional scales relevant to FATM samples, 3) sufficient nominal surface roughness, 4) surface normals within the numerical aperture (NA) of the optical system, and 5) artifact features that do not contain vertical transitions. We expand on each of these design considerations below. Two additional issues contributed to the challenging nature of this project. First, the determination of instrument requirements for forensic applications is an active topic of research, both for virtual comparison microscopy (VCM) and automated comparisons. Second, the definition of areal instrument performance parameters, validation and calibration procedures, and artifacts is still being standardized within ISO.

The first design consideration was simply ensuring that the material measures developed are able to be mounted on or inside the various candidate optical surface topography microscopes for FATM analysis. General purpose optical surface topography microscopes are able to accept a wide variety of samples because they typically have an open table to mount samples on. As described above, FATM application-specific optical surface topography microscopes are available for FATM analysis. These microscopes often have mounts designed specifically for holding ballistic samples, and sometimes they may only accept ballistic samples. To address this constraint, the metrology targets must be fabricated and presented in a format that mimics a ballistic sample. Conceptually it would look like the sample shown in Figure 1a; a cartridge case with integrated reference step heights to evaluate the accuracy of measured surface heights. Practically, our choice was to have the metrology targets reside on a cylindrical blank that has the same dimensions as a cartridge case. The blank form factor (Figure 1b) chosen is equal to the diameter of a 12-gauge shotgun shell. We recognized that there may be a few commercially available instruments whose fixtures can only accept bullets and cannot accept the larger diameter 12-gauge shells, but did not address that issue in this research.

a)



b)

Figure 1 – a) Concept for a step height standard on the headstamp area of a cartridge case, b) sample blank cylinder having the same diameter as a 12 gauge shotgun shell that will incorporate metrology targets. The outside diameter of this blank is 21.5 mm.

The second design consideration was to design features with dimensions relevant to the types of features analyzed in FATM applications. Our designs were primarily guided by the size of the regions on ballistic samples that contain toolmark features whose relative location is individualizing, as well as the size and height of individualizing toolmark features themselves. Individualizing toolmarks are found on a variety of fired cartridge components, such as the land engraved areas on bullets, and the breech face and firing pin impression marks on cartridge cases. The respective regions of interest and features have dimensions ranging from micrometers to millimeters.

Another limitation of certain optical topography measurement methods is the requirement that the measured surface contains sufficient surface roughness or contrast in order to be measured [30]. The third design consideration was to produce prototype artifacts with some nominal amount of surface roughness (or another appropriate contrast mechanism) to enable measurability across all of the candidate microscopes. Confocal and interferometric surface topography microscopes are usually better suited to measure surfaces that are extremely smooth and highly reflective. Focus variation microscopes and photometric stereo microscopes that do not use a gel, usually require some minimal amount of roughness in order to measure the surface effectively.

The fourth design consideration was to have features with surface normals that enable reflected light to be collected by the system. Interferometric, confocal, and focus-variation microscopes typically distribute and collect a cone of light rays whose half angle is determined by the Numerical Aperture (NA) of the objective. Many firearm toolmark topography measurements are made at low magnification. For example, the measurement of breech face marks on the primer have shown to produce sufficient detail for identification at 10X magnification [13, 31]. 10X magnification objectives typically have lower NA values, which significantly limits the allowed local slope of a measurable surface area. A typical quality 10X microscope objective has an NA of around 0.3, which translates to a collection half angle of approximately 17.5 degrees. A specular surface feature with a local slope larger than this angle will not reflect light back into the microscope and will therefore not be imaged. Surfaces oriented beyond the objective's NA can sometimes be measured if some degree of surface roughness is present, causing a diffuse reflection. [32] Some focus variation microscope vendors allow for a light source to be placed outside of the objective, typically resulting in the capability of measuring larger slopes for surfaces with some minimum roughness.

The fifth and final design consideration was to avoid surface steps with near vertical slopes, such as those found on artifacts manufactured using semiconductor lithography methods. Of the optical surface topography methods, photometric stereo relies on estimating the local surface normal vectors. The respective slopes are then integrated to obtain an estimate of the surface topography. For a vertical step, the measurement system is typically unable to obtain the sidewall information required to estimate the height difference between the surface steps. Depending on the application of advanced techniques, a surface topography microscope based on photometric stereo may have difficulties measuring and reconstructing a surface with vertical steps [33, 34].

6.0 FATM Relevant Metrological Characteristics

Stated above, a metrological characteristic is defined as a characteristic of measuring equipment which can influence the results of measurement. ISO 25178 Part 600 provides a comprehensive list of metrological characteristics for common optical methods to measure areal surface topography, including the four described above. The draft standard ISO 25178 Part 700 provides procedures to measure the metrological characteristics. Manufacturers often publish specifications of their instruments. These specifications include system specifications and performance specifications. System specifications are relatively straightforward uncontroversial facts about a microscope system. These include such things as stage travel, stage speed, overall dimensions, and the objectives (magnifications) available. Performance specifications are more complicated and are based on metrological characteristics. More specifically, a performance specification is a limit value associated with placed on a given metrological characteristic. For example, measurement noise is a metrological characteristic defined in ISO 25178-600. It is defined as the "noise added to the output signal occurring during the normal use of the instrument". A microscope manufacturer may conduct tests to evaluate measurement noise and publish a limiting value as part of a set of instrument specifications that they make available to customers. Under normal operation, a customer should expect the measurement noise to always be less than or equal to the published measurement noise value from the manufacturer.

A complete performance specification should include a description of the measurand (exactly what characteristic or performance parameter is being measured), acceptable environmental conditions, measurement procedure and artifacts (if any), data processing (including averaging and filtering), and reporting. Most of these can be specified by a simple reference to an applicable standard, such as ISO 25178 Part 700. When the measurement conditions are not explicitly stated, it is difficult to equitably compare performance specifications from microscope to microscope. For noise, the instrument performance parameter typically depends on environmental conditions and the characteristics of the measured surface. Therefore, the standard also defines instrument noise; internal noise added to the output signal caused by the instrument if ideally placed in a noise-free environment.

For this project, we list the metrological characteristics (shown in Table 2) defined in ISO 25178-600 that have bearing on the accuracy of FATM topography measurements. Each metrology characteristic in Table 2 is briefly described below. In the right column of Table 2, a red "X" indicates whether a material measure exists that can be used to evaluate the metrological characteristics identified in the left column on all of the various surface topography microscopes being marketed for FATM applications. Except for the last two (measurement stitching and measurement noise), there is a need for new physical artifacts to evaluate the metrological characteristics. That was a core part of the project. A goal of future work is to provide guidance on minimum instrument performance specifications for FATM applications, typically as maximum permissible errors. The artifact prototypes explored in this project enable the instrument user to evaluate conformance to the specifications.

Metrological Characteristics	Physical Artifact
1. Amplification Coefficient (in X and Y axes)	X
2. Amplification Coefficient (in Z axis)	X
3. Linearity Deviation (in X and Y axes)	X
4. Linearity Deviation (in Z axis)	X
5. Topographic Spatial Resolution	X
6. Maximum Measurable Local Slope (convex)	X
7. Maximum Measurable Local Slope (concave)	X
8. Maximum Measurable Local Slope (planar surface)	X
9. Flatness Deviation	X
10. Measurement Noise	SRM 2461
11. Measurement Stitching	SRM 2461

Table 2 – List of metrology characteristics to address for FATM analysis

Line items 1 and 2 in Table 2 are the amplification coefficients for the 3 axes of a 3D instrument. Line items 3 and 4 are the respective linearity deviations. All these characteristics affect the ability to accurately measure the height of features and their relative lateral position. To understand the characteristics, the instrument response function must be defined. ISO 25178-600 defines the response function as *the relation between the actual quantity (instrument input) and the measured quantity (instrument output)*. Figure 2 illustrates an instrument response function (curved dotted black line). ISO 25178 defines the amplification coefficient as *the slope of the linear regression line obtained from the instrument response function*. The linear fit to the instrument response is shown as the dotted blue line in Figure 2. It is the slope of this linear fit that is the amplification coefficient. Ideally (but never achieved), a measurement instrument would have a one-to-one correspondence of instrument input to instrument output. This perfect linear response is illustrated as the solid blue line in Figure 2. For a perfect instrument, the amplification coefficient equals 1. In practice, the actual amplification coefficient of an instrument may have a scale error, which describes the linear approximation of the errors in the instrument response function, i.e., the error that is proportional to the measured value. Sometimes the amplification coefficient is referred to as the scale factor. Residual errors in the response function are described by the linearity deviation. ISO 25178 defines linearity deviation as *the maximum local difference between the line from which the amplification is derived and the response function*.

The amplification factor and linearity deviation relate to the parts of the microscope (scales) that provide the X-, Y-, and Z-coordinates of measured points in a given 3D topography data set. For the X- and Y-axes, the scale factor is typically determined by the mean spacing of the pixels in the camera system and the effective magnification of the microscope. The respective linearity errors are often determined by aberrations in the imaging optics that distort the image laterally. For the Z axis of a confocal, interferometric, and focus variation microscope, both the scale factor and linearity deviation are typically determined by the Z-axis displacement measurement system embedded in the instrument, such as an encoder or capacitance gauge, with additional contributions by, for example, hysteresis and angular errors in stage motion.

Estimation of the amplification coefficient of an axis is closely related to estimating the non-linearity error of that axis. A common approach for the z-axis is to first measure an artifact with a certified step height to obtain an estimate of the amplification coefficient. This is typically a larger step, as the respective measurement error is often dominated by the error in the amplification coefficient. The error in the amplification coefficient is then estimated as the difference between the measured value and certified value of the step divided by the certified value. Next, one or more heights are measured at different locations in the axis measurement range of interest. The linearity error is then estimated as the maximum (absolute) error in the measured heights after adjustment for the error in the amplification coefficient.

The amplification coefficient and linearity deviation of the X- and Y-axes are metrological characteristics that affect the lateral errors in the position of measured features. These metrology characteristics can be evaluated by measuring a grid of features with certified values for their relative position. After measurement, the measured grid is translated, rotated, scaled in the X-direction, and scaled in the Y-direction to minimize the distance between the measured and nominal features in a least squares sense. The two estimated scale factors represent the inverse of the X and Y amplification coefficients. The residual errors in the X- and Y-direction are the mapping errors of the instrument. They describe the lateral distortion of an image after correction for the difference in the X- and Y-axis amplification coefficients. The maximum (absolute) mapping errors in the X- and Y- position of the measured features are the linearity deviations in X and Y.

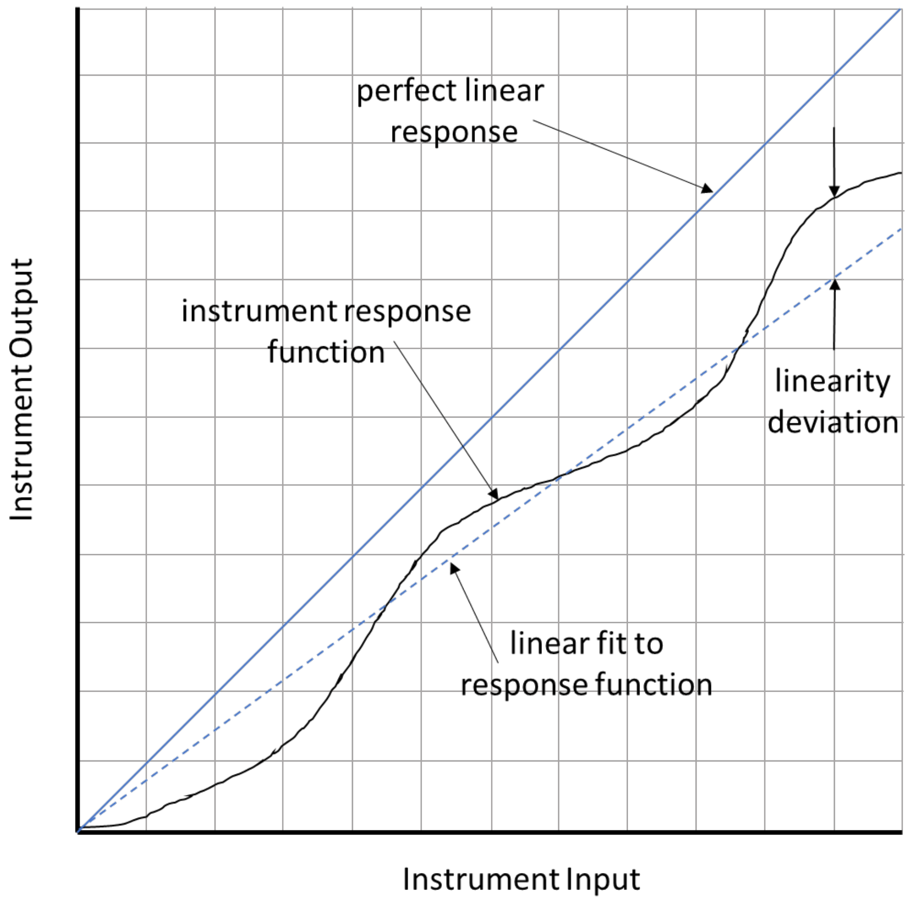


Figure 2 – Generic illustration of an instrument response function. The perfect linear response of an instrument is indicated with the solid blue line. The actual instrument response function is shown with the solid curved black line. The slope of the linear fit (dotted blue line) to the instrument response function is defined as the amplification coefficient. The linearity deviation (indicated in upper right corner) is the maximum difference between the response function and the linear fit to it.

Line item 5 is the topographical spatial resolution. ISO 25178-600 defines this metrology characteristic as the ability of a surface topography measuring instrument to distinguish closely spaced surface features. The standard defines several parameters and functions that may be used to quantify spatial resolution. In optical instruments, there exist different theoretical criteria [35] for defining optical resolution, e.g., Sparrow criterion and Rayleigh criterion. These criteria relate to the optical system's ability to distinguish (resolve) two closely spaced point sources. They are proportional to an optical system's illumination wavelength and inversely proportional to the system's numerical aperture. The lateral resolving capability of an imaging system can also be described by the optical transfer function (OTF) [36]. The OTF expresses numerically how an optical system handles different spatial frequencies that are present in the scene being imaged. This is often evaluated by the ability to detect contrast on samples with alternating light and dark areas. As these areas get closer together, contrast decreases. In surface topography, lateral resolution needs to be determined differently, instead considering how accurately height information is measured as surface features get closer together. The Instrument Transfer Function (ITF) describes the instrument's height response as a function of the spatial frequency of the surface topography. Ideally, it describes the measured amplitude of a sinusoidal profile (grating) as a function of its frequency. However, as noted by ISO 25178-600, the ITF for several optical topography instruments may be a nonlinear function of surface height. The ISO standard lists several parameters that can be used to summarize key aspects of the ITF. One example is the lateral period limit, the wavelength of a sinusoidal topography profile at which the height response of the ITF falls to 50%. Another approach to evaluate topographical spatial resolution is to measure an artifact that has an array of features of consistent height, but spaced at various pitches.

Line items 6 through 8 relate to the maximum measurable local slope. ISO 25178 defines this metrological characteristic as *the greatest local slope of a surface that can be assessed by the probing system*. It addresses the ability of an optical surface topography microscope to measure surface areas that are not normal to the optical axis of the microscope. When local surface angles exceed a surface topography microscope's maximum measurable slope, the microscope can no longer acquire data. This is often observed as dropouts in the measured data. The maximum measurable slope is affected by surface roughness. Some of the light scattered by rough surfaces may enter the microscope and provide sufficient signal for measurement. We make a distinction between evaluating the microscope's slope measuring capability on convex spherical surfaces, concave spherical surfaces, and flat surfaces.

Line item 9 is flatness deviation. ISO 25178 defines this metrology characteristic as *the deviation of the measured topography from an ideal plane*. In other words, it considers how well an optical surface topography microscope is capable of measuring an ideally flat surface. For some instruments, this deviation is related to the quality of the internal areal flatness reference of the instrument. In other cases, it can be caused by uncorrected or residual optical aberrations in the instrument optical setup. Although flatness deviations are affected by setup and instrument drift, they are to a large extent reproduced in every measurement. In general, this flatness deviation can be evaluated by measuring an optical flat. Some microscopes may require some nominal roughness or contrast on the flat surface.

Line item 10 is measurement noise. In ISO 25178, two different types of noise are defined: instrument noise and measurement noise. Instrument noise is defined as *internal noise added to the output signal caused by the instrument if ideally placed in a noise-free environment*. Measurement noise is defined as *noise added to the output signal occurring during the normal use of the instrument*. Measurement noise is the preferred metrological characteristic to check, as it represents noise encountered in the user's environment during actual operation. Measurement noise is assessed by measuring the same surface at least two times and evaluating the repeatability of the reported surface heights. Measurement noise is affected by the type of surface being measured. Measuring toolmarks on a cartridge case or bullet, such as the NIST standard cartridge case or bullet, yields noise values for surfaces relevant to FATM analysis.

Line item 11 is image stitching. While not addressed directly in ISO 25178, it is an important aspect of surface topography microscopes to evaluate. Any time the area to be measured is larger than the microscope's field of view (FOV) at a given magnification, image stitching is required. In this case, the microscope is programmed to measure smaller overlapping regions of the area. Stitching software then merges the overlapping topography images into a single topography image. For FATM applications, image stitching is utilized on many instruments for measuring features containing individual toolmarks. For example, breech face impressions on the primer measured at 10X magnification will often require a 3x3 stitch to capture the entire breech face impression. Image stitching can introduce errors in the measured surface topography. A key possible error source are errors in the relative lateral position of the images to be stitched. Errors in this relative position cause errors in the relative lateral position of features on the stitched image. They may also result in the blurring of features in the overlapping areas. One possible test to detect such errors is to measure the same surface twice, with, for example, the sample shifted by half a FOV during the second measurement. The two stitched images are then registered and subtracted. For surfaces with fine features, the difference image may show areas with relatively large differences between the two stitched images due to differences in the position of the same features on both images.

7.0 Physical Reference Artifact Prototype Design and Fabrication

Evaluating the metrology characteristics described in the previous section requires suitable artifacts. A major effort was undertaken to fabricate these artifacts utilizing several different silicon processing techniques. Eventually this effort was halted due to several technical barriers. This effort is briefly described in section 7.1 Section 7.2 describes our final approach towards design and fabrication of physical reference artifacts that address the metrology characteristics described above.

7.1 Exploring Solutions with Silicon Processing

A focused effort was undertaken to explore physical reference artifact prototypes that could be fabricated with various silicon fabrication techniques, motivated by ease, cost and accuracy of these fabrication techniques. We designed (Figure 3) and produced a photomask with metrology targets that address most metrology characteristics listed above. The photomask was used to create a silicon wafer with these metrology targets. The features that were produced included: Section A) individual line/space pairs ranging from 1 µm to 8 µm line width (at 1:1 and 1:2 ratios of peak width and valley width), Section B) chirped line/space array with line widths ranging from 0.1 µm to 20 µm, Section C) large and small checkerboard patterns consisting of 25 µm and 50 µm squares (at 1:1 and 1:2 peak width and valley width ratios), Section D) circular, radially chirped, wavelength standard, Section E) Siemens star, and Section F) cross hatch patterns. These features were intended to address the following metrological characteristics and associated specifications with one artifact; amplification factor (scale) of the X and Y axes, mapping errors and linearity deviation of the X and Y axes, and lateral resolution of the X and Y axes.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

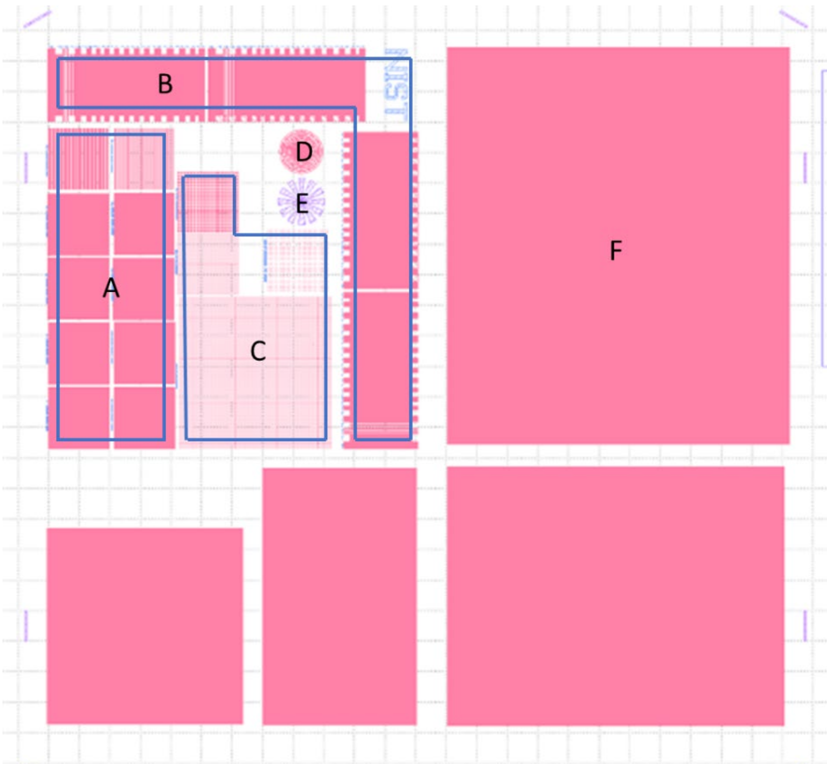


Figure 3 – Photomask design incorporating various metrology targets for evaluating lateral scale (amplification factor), mapping errors and lateral linearity deviation, and lateral resolution.

Some instruments cannot measure the ultra-smooth low roughness silicon surfaces. To address this, we developed a novel approach to introduce local contrast (differences in reflectivity) through rapid thermal annealing (RTA) of thin gold coatings. This effort was conducted in collaboration with researchers from the Center for Nanoscale Science and Technology (CNST) at NIST. Some success in this approach was achieved, improving the ability of a focus variation microscope to measure the smooth silicon surfaces. Figure 4 is a scanning electron microscope (SEM) image of a gold coated silicon wafer after RTA treatment. The image shows the small micrometer-level gold islands produced by the RTA process. These islands improve a focus variation microscope's ability to measure the surface by adding local changes in reflectivity that the focus algorithm can detect. The height of the gold islands is on the order of a few nanometers and does not significantly affect the estimation of instrument performance parameters for our application.

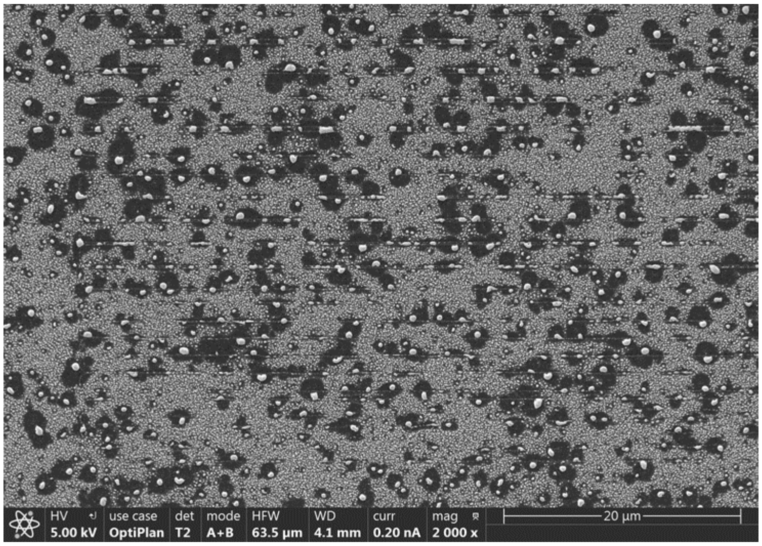
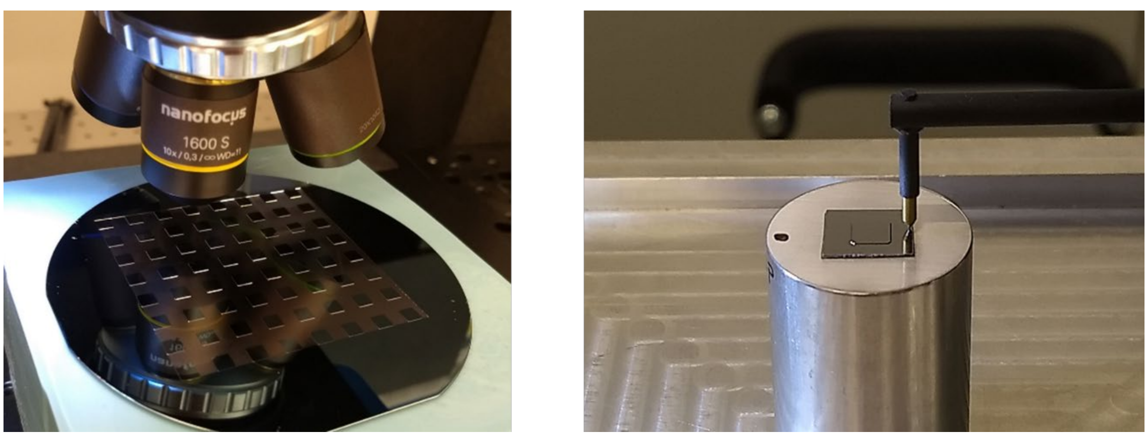


Figure 4 – SEM image of RTA prepared silicon surface with gold island formation

In addition, we explored bulk silicon micromachining methods to fabricate artifacts for evaluating the scale and linearity of the Z-axis and the maximum measurable slope. Bulk silicon micromachining [37] is based on anisotropic etching of crystalline silicon. Etching is influenced by various factors including the choice of etching solvent, crystal orientation, and doping concentration of the substrate. This technique enables fabrication of deep etches of the silicon substrate (tens to hundreds of micrometers) as well as inclined surfaces. A nominal 100 µm step height was fabricated using this approach (Figure 5). We never proceeded to making a maximum measurable slope artifact with this technique before a decision to abandon development efforts in silicon was reached.



a) b)

Figure 5 – Step height standard prototype fabricated using bulk silicon micro machining techniques, a) complete wafer with an 8 x 8 array of nominally 100 µm step heights, b) one 100 µm step height diced and assembled to a cylindrical blank

Ultimately, the effort to fabricate artifacts in silicon was halted. Several technical barriers warranted a change in approach. First, the heights of the silicon features using standard silicon processing are only on the order of a few micrometers. Second, most standard silicon processing methods create features with vertical sidewalls. Both of these issues present a measurement challenge for some instruments and could not be overcome. Last, we were unable to apply the gold RTA approach to surfaces produced by the bulk silicon micromachining technique, leaving these types of sample unmeasurable by some instruments. For these reasons, we moved onto other designs and fabrication methods for artifact prototypes to address the metrology characteristics discussed above. However, the metrology targets developed here may still have utility in evaluating performance of some microscopes. They can also be used to evaluate the 2D imaging capabilities of the 3D surface topography microscopes. However, metrology characteristics determined using the 2D imaging capability of a 3D microscope do not necessarily translate into the metrology characteristics for 3D imaging.

7.2 Artifact prototype designs and fabrication methods

The core portion of the research involved exploring designs and fabrication methods for physical reference artifacts measurable by the various 3D instruments being marketed for FATM analysis. Prototype designs and fabrication methods for artifacts that address the metrology characteristics discussed above are presented in this section. More standard metal-working approaches (including micro-machining), in addition to a novel micro-indentation technique, were pursued in place of the silicon fabrication processes discussed previously.

7.2.1 Prototype artifact for X and Y axis Amplification Coefficient, Mapping Deviations, and Linearity Deviation

The overall design consists of an array of features that fills the FOV at low magnification (e.g., 10X or 20X). The application of this artifact follows the procedure outlined in ISO 25178-700:

1) Measure the surface topography

2) Estimate the X- and Y-position of each feature in the measured 3D surface image

3) Translate and rotate the set of measured features, and adjust the scale in X- and Y-direction, to minimize the distance between the actual and nominal feature positions in a least squares sense. The thus estimated scales are the inverse of the amplification coefficients in X and Y. The residual errors in the position of the measured features are the mapping deviations of the instrument

4) Obtain the linearity deviation of the X- and Y-axis as the maximum value of the respective mapping errors.

Several types of features were considered, including an array of conical features and an array of hemi-spherical features. For conical features, laser ablation [38] was pursued to machine micrometer size conical features. Eventually, a decision was made (based on ease of analysis) to go with hemispherical features. For hemispherical features, two different approaches were considered: endmill micromachining and a novel micro-indentation approach. Our final prototype fabrication method utilized the micro-indentation method due to the better surface quality.

The overall artifact prototype design consists of an 8x8 array of concave hemispherical features. The features were designed to have a diameter of 100 µm at the intersection with the artifact plane. The features are spaced at 150 µm nominal pitch. In addition, the maximum sidewall angle of the hemispherical features is intended to stay below 22 degrees to stay within the numerical aperture of low magnification optical systems. Fiducial lines are incorporated in the artifact to guide the user to the array located at the center of the artifact. The features would be generated on a cylindrical blank that has the same diameter as a 12-gauge shotgun shell. A sample of the engineering drawings are shown in Figure 6.

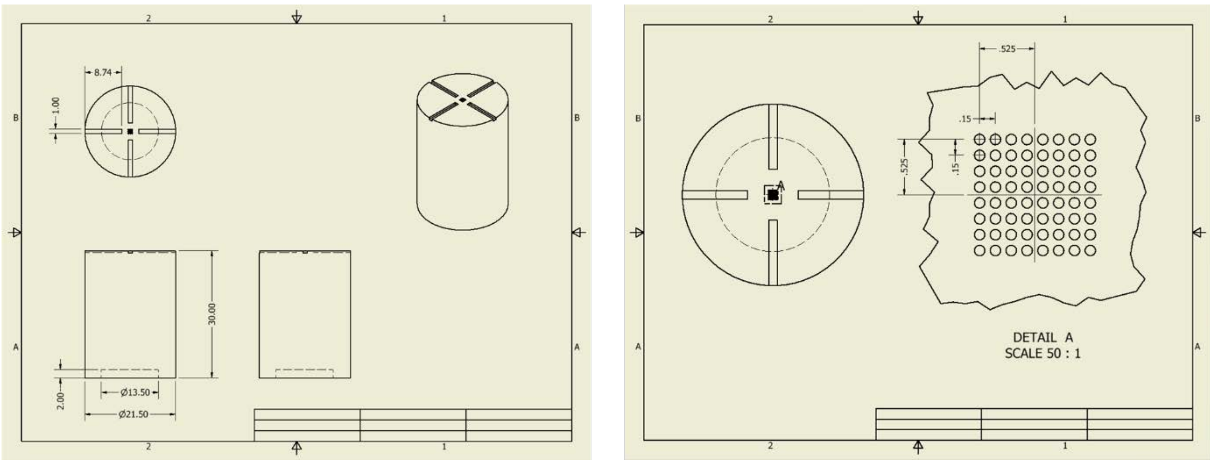


Figure 6 – Engineering drawings for an 8x8 array of hemispherical features to estimate lateral scale, mapping deviations, and linearity deviations

An initial prototype of the 8x8 array was fabricated using specialized hardness indenter instrumentation at NIST. This effort was in collaboration with researchers from the Mechanical Performance Group within the Material Measurement Laboratory at NIST. A custom 160 µm indenter was designed and purchased from a commercial vendor, and utilized in the fabrication. Figure 7 shows an initial prototype made in aluminum and a sample topography measurement. The spherical geometry produced from the micro-indentation process is quite suitable for distance metrology. Preliminary analysis on a subset of four micro-indentations (Figure 8) reveals good sphericity. The S90 number reported in the figure represents the standard deviation of the signed distance of the measured surface points to the best fit sphere. The surface points considered are those within 90% of the radius of the intersection circle of the sphere with the artifact plane (indicated by the red circles). It follows that the form error of the four indentations, expressed as a range corresponding to four standard deviations, is on the order of 200 nm.

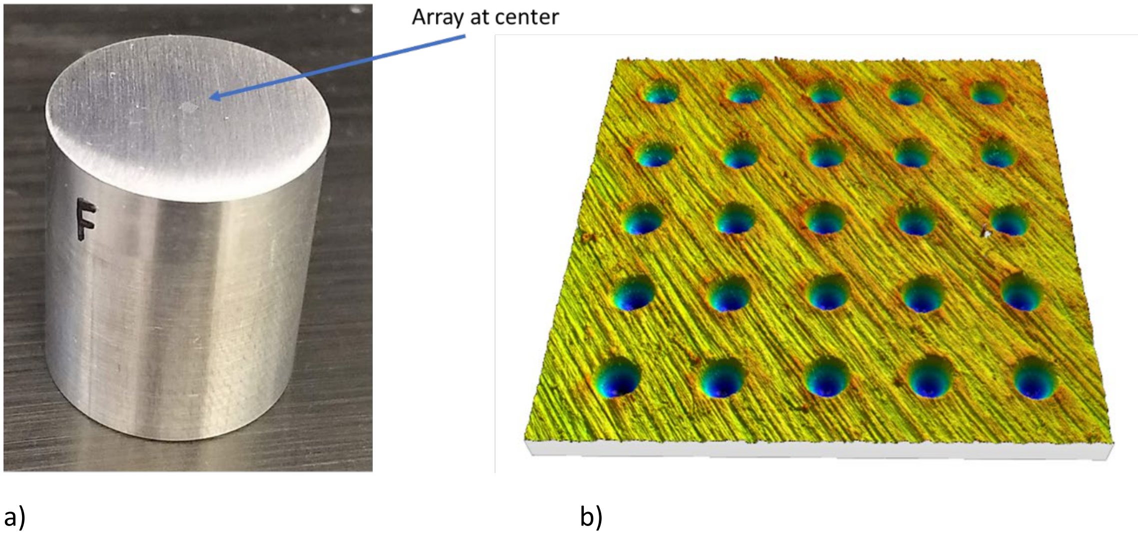


Figure 7 – a) XY lateral 8x8 array prototype with array located at center, b) sample optical topography measurement of the array

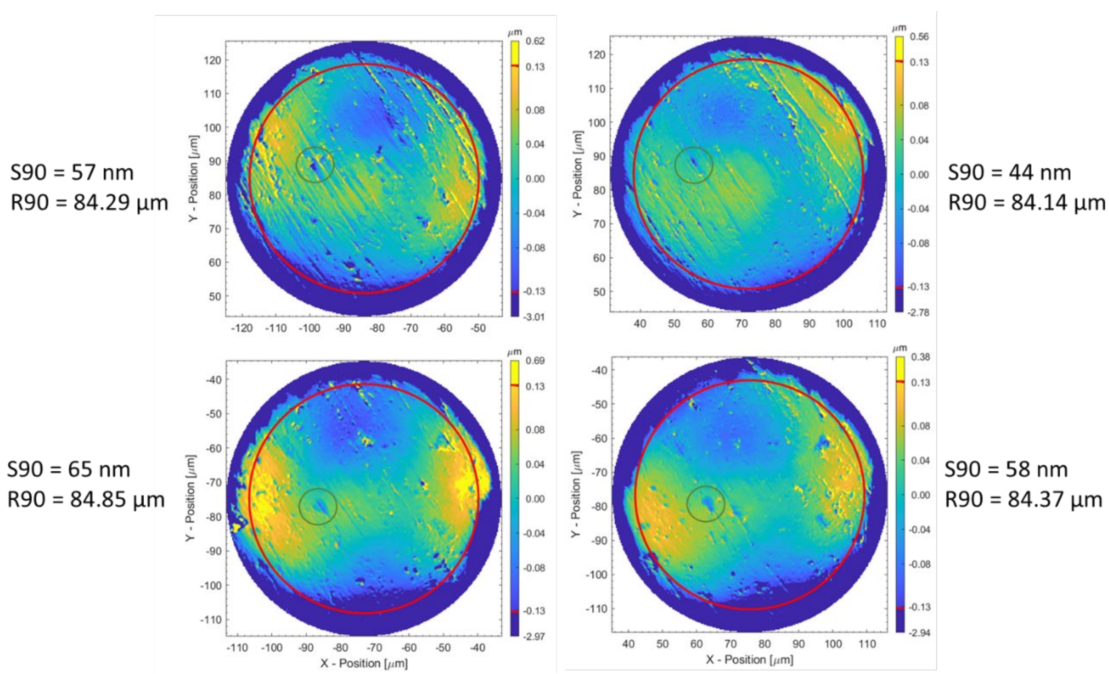
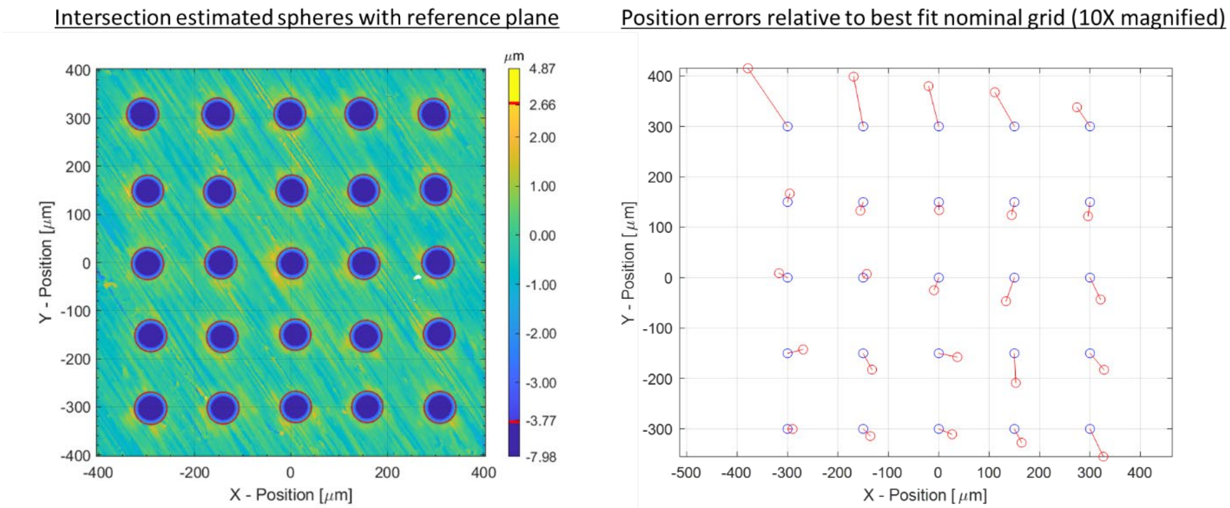


Figure 8 – Measured topography of 2x2 subset of 8x8 array of spherical micro-indentations. The S90 number is the standard deviation of the signed distance of the surface points relative to the best fit sphere. The points considered are those within the large red circle, representing 90% of the radius of the intersection circle of the sphere with the artifact plane. The R90 number is the 90% radius.

We developed analysis software to estimate the positions of the spherical indentations in surface measurement data of the artifact. The software then calculates the instrument X- and Y-axis amplification coefficients, mapping errors, and linearity deviations. Data shown in Figure 9 was acquired at 20X magnification using a confocal surface topography microscope. This type of data will be obtained by practitioners using this artifact in a laboratory setting to calibrate, verify, and monitor the X- and Y-

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

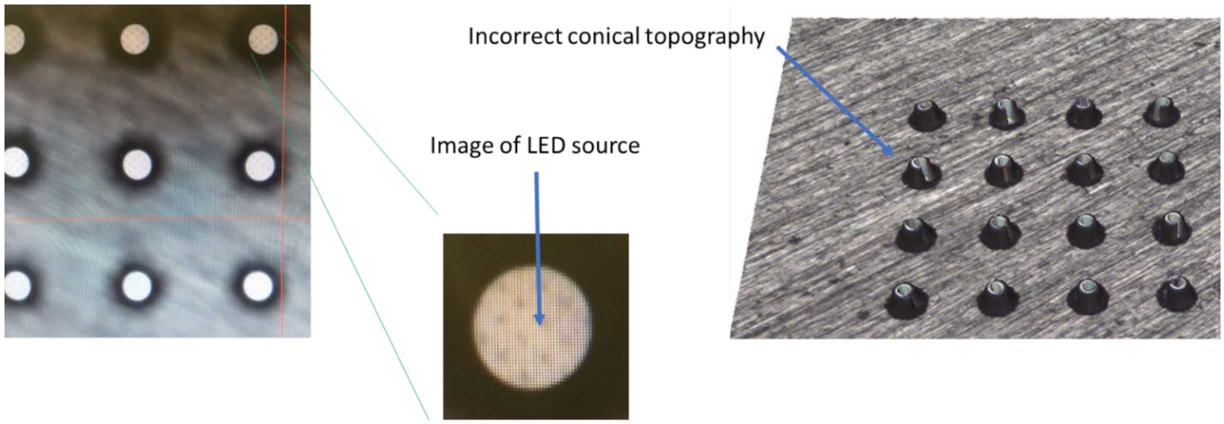
axis scales, mapping deviations, and linearity deviations of their microscopes. For demonstration purposes, we show a plot of the positional analysis in Figure 9b comparing the measured position of the hemispherical features to the nominal positions. The maximum error in this positional analysis was 14 µm, and mainly represents errors in the prototype fabrication process of the uncalibrated artifact.



a) b)

Figure 9 – Positional analysis on a 5x5 subset of the 8x8 array, a) plot showing the measured surface topography with the 5x5 grid of spherical indentations, b) positional analysis comparing the measured position of the indentations to their nominal positions. The deviations (shown with red lines and circles) are magnified by a factor of 10.

The prototype artifact shows promise to enable the quick evaluation of the X- and Y-axis amplification coefficients, mapping errors, and linearity deviations. However, the high-pressure indentation process generates a spherical geometry with very low roughness which, unfortunately, creates a measurement challenge for some microscopes that needs to be addressed. The near perfect hemispherical indentations are essentially concave spherical mirrors. On most microscopes (e.g., focus variation microscopes), the illumination optics are based on an illumination design known as Koehler illumination [39, 40], which produces a defocused image of the light source at the sample plane. However, for our artifact, the spherical indentations may result in a focused image of the illumination source above the sample surface, which can confuse the topography calculation. Spherical reflectors have a focal length of half their radius. In this case, with an indentation radius of 80 µm, the image would be formed 40 µm from the bottom of the micro-indentation. We observed this effect for artifact measurements on a focus variation microscope at 10X magnification. Figure 10a shows an example of the micro-indentation forming an image of the microscope's illumination source, in this case a LED. Figure 10b subsequently shows the incorrect positive protruding conical topography that the focus variation microscope's software calculates. They should be concave hemispheres. We evaluated this effect on three different microscopes: a confocal microscope, an interferometric microscope, and a focus variation microscope.



a) b)

Figure 10 – The optical power of the micro-indentations creates problems for some instruments to correctly measure and determine the topography, a) formation of an image of the microscope's illumination source, b) sample topography measurement incorrectly estimates positive protruding conical shapes.

Several methods were considered to mitigate this undesired effect. The first approach is to chemically etch the micro-indentation surfaces to reduce the optical power of the geometry. There will be a balance between the amount of etching performed and how much the spherical form and roughness of the micro-indentation is altered. Preliminary etching results using NaOH improved the ability of the low magnification 10X focus variation microscope to measure the sample. Additional chemical etching experiments will be explored. The second approach is to create micro-indentations with a larger diameter by using an indenter with a larger diameter. The larger diameter of the spherical indentation results in the undesired image of the light source to be created further away from the sample plane. This is expected to reduce or eliminate its effect on the images of the sample surface produced during the through-focus scanning motion. For a focus variation microscope, for example, the light source does not have to be coaxial (through the objective illumination). In that case, the formation of undesired images can be mitigated by choosing ring or other non-coaxial light sources.

7.2.2 Prototype artifact for Z Amplification Coefficient and Linearity Deviation

The overall design incorporates a row of four adjacent 2 mm by 2 mm planar pads, each offset vertically from one another, to create a total of three different step heights. The nominal step values are 100 µm, 50 µm, and 10 µm. Vertical surface discontinuities were avoided by including nominal 20° sloped transitions from one pad to the next. Nominal surface roughness was planned to be introduced through the native roughness from the manufacturing technique, chemical etching, or media blasting. Three manufacturing methods were explored to produce this prototype artifact: traditional end milling, diamond turning, and wire-electro-discharge machining (EDM). The end milling approach was abandoned because the surface roughness was too high, and it proved too difficult to accurately machine the sloped transitions between adjacent pads. Diamond turning is being considered as it enables the fabrication of highly accurate steps and surfaces. However, the fabricated surfaces are very specular, requiring some means to introduce artificial roughness such as relatively large feed rates during machining or etching. The initial method selected to fabricate the artifact prototype was wire EDM. Figure 11 shows engineering drawings for this sample prototype.

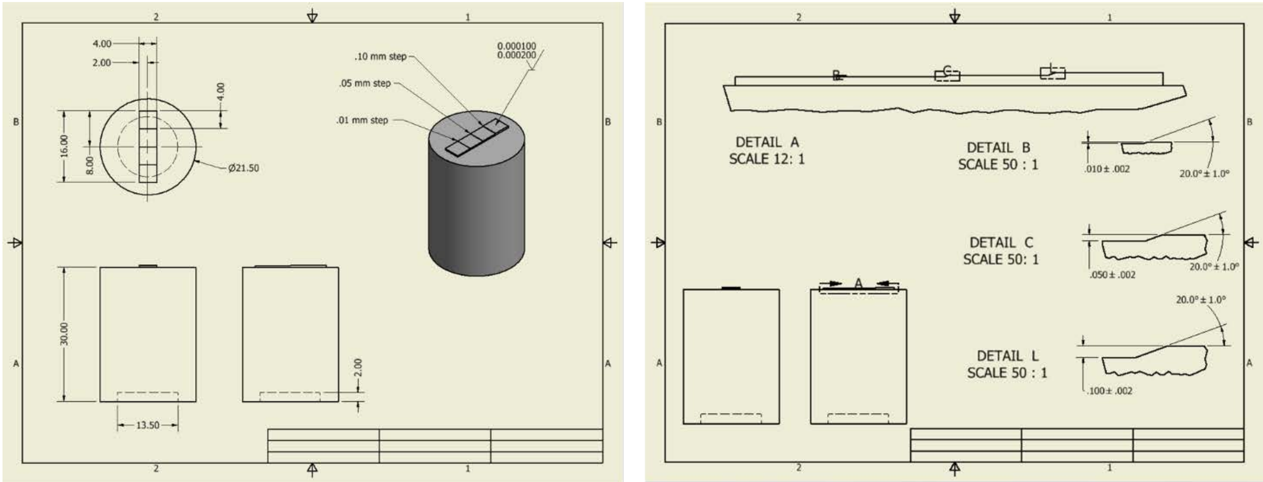
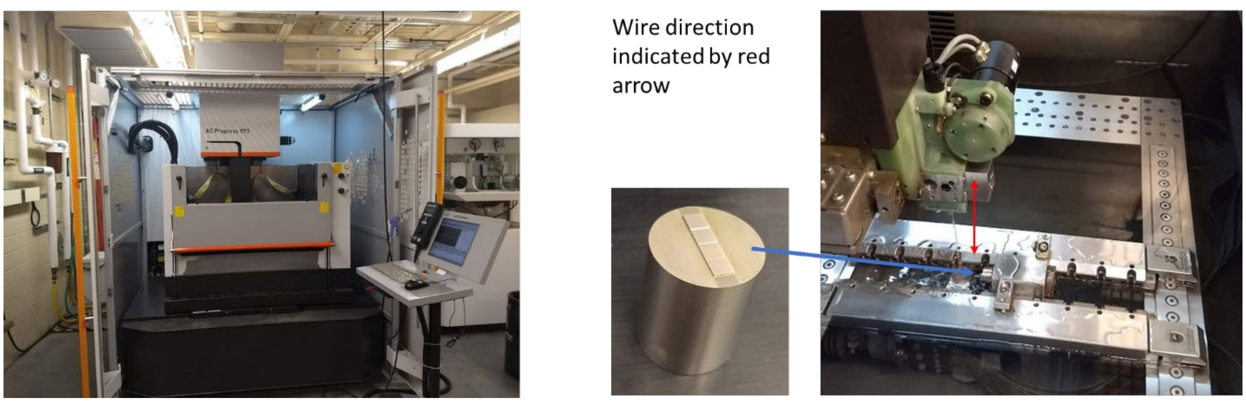


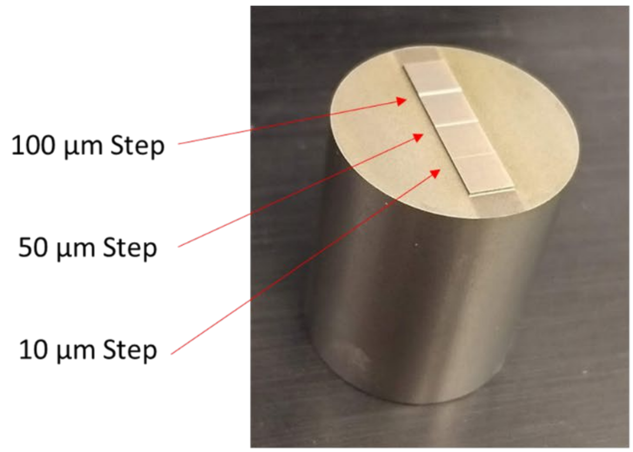
Figure 11 – Engineering drawings for the prototype design of the Z axis scale and linearity artifact

Wire EDM is a common metal machining method used in industry. The technique works by passing a high current through a small diameter wire. The wire is used to cut the desired shape through electrical discharges between the wire and the part that erode part material close to the wire (spark erosion). The sparks also erode wire material, which is therefore constantly being changed. A 0.1 mm diameter wire was used to cut the adjacent pads on the prototype. The NIST-owned wire EDM was made by Agie Charmilles. Figure 12 shows a picture of this wire EDM and the resultant step height prototype artifact. The target value for the surface roughness average, Ra, was between 100 nm and 200 nm Ra. Measured roughness values on the four pads produced between 100 nm and 150 nm Ra. Figure 13 shows the final prototype with the three steps and a sample topography measurement highlighting the slope transition between adjacent pads.

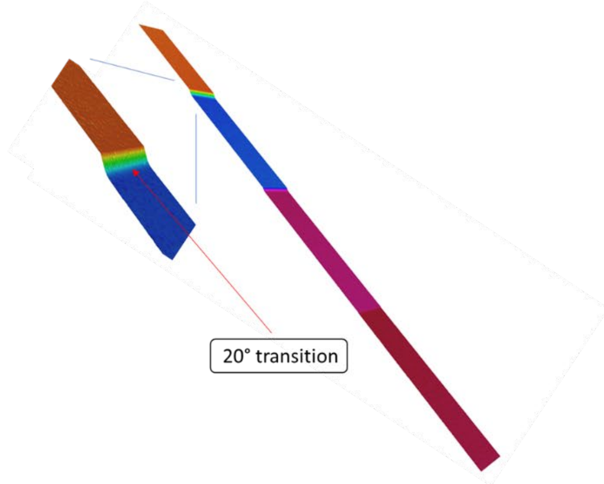


a) b)

Figure 12 – Wire EDM used to fabricate the step height prototype, a) picture of the wire EDM machine, b) view inside the wire EDM tank and the prototype after the cutting was finished. The red arrow indicates the direction of the wire.



a)



b)

Figure 13 – a) Prototype step height artifact for evaluating the Z-axis amplification coefficient and linearity deviation, b) focus variation-based sample topography measurement highlighting the 20° sloped transition between adjacent pads

7.2.3 Prototype artifacts for Maximum Measurable Local Slope

Three different prototype artifact designs were considered for evaluating maximum measurable local slope. The first design incorporates the use of small convex precision spheres. The second design is based on small concave hemispherical features. The last design involves machined flat planes at several different discrete nominal angles. The two different spherical geometries were considered because of the different measurement challenges posed by convex and concave surfaces. The planar discrete angle solution was considered to avoid possible challenges caused by the optical power of a spherical geometry, as demonstrated in the measurement of the spherical indentation array discussed above.

7.2.3.1 MMS artifact utilizing small convex sub-FOV spheres

The prototype design incorporates 3 precision 440 stainless steel spheres with different diameters. Three diameters are offered to provide one that is sub-FOV, depending on the magnification being used by the end user. The advantage of using spherical geometry to evaluate the MMS specification is that it provides a continuum of angles to evaluate from the pole of the sphere (0 degrees) to the equator (90 degrees). The three spheres have nominal diameters of 0.5 mm, 0.79375 mm, and 1.0 mm. The three spheres are fitted to a cylindrical blank having the same diameter as a 12-gauge shotgun shell. The cylindrical blank is machined from aluminum and has three hemispherical cups machined to mount the spheres. A fiducial mark is positioned in the center to assist the end user with locating the individual spheres during measurement. Figure 14 shows the engineering drawing and a sample of the prototype artifact.

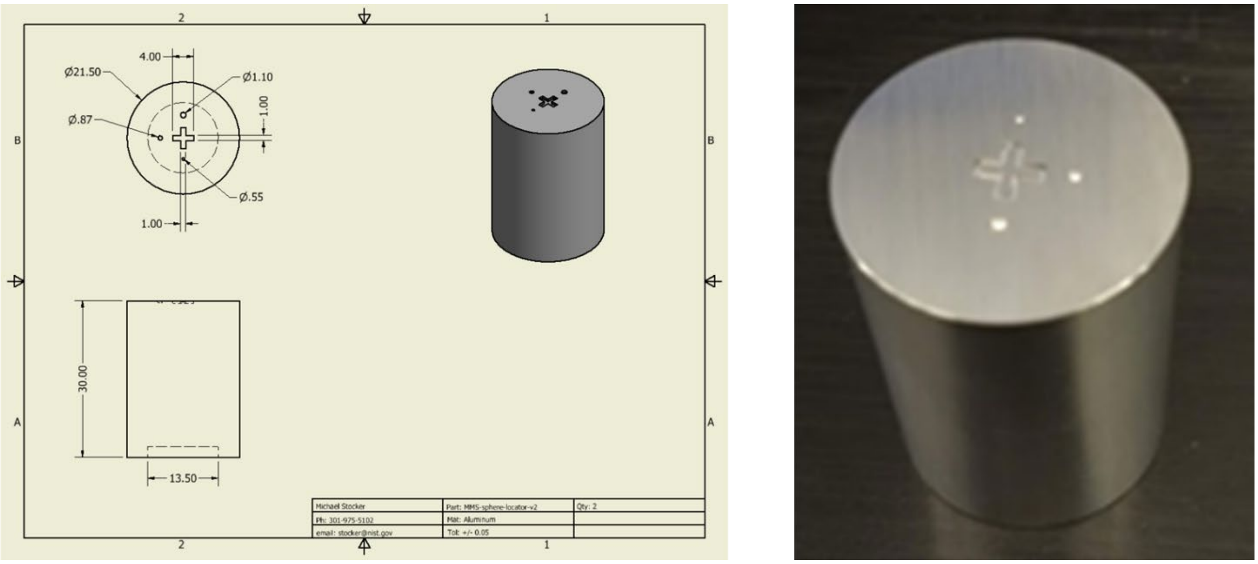
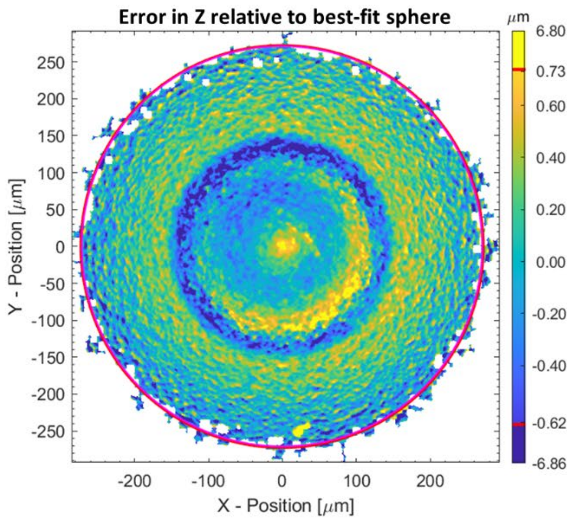
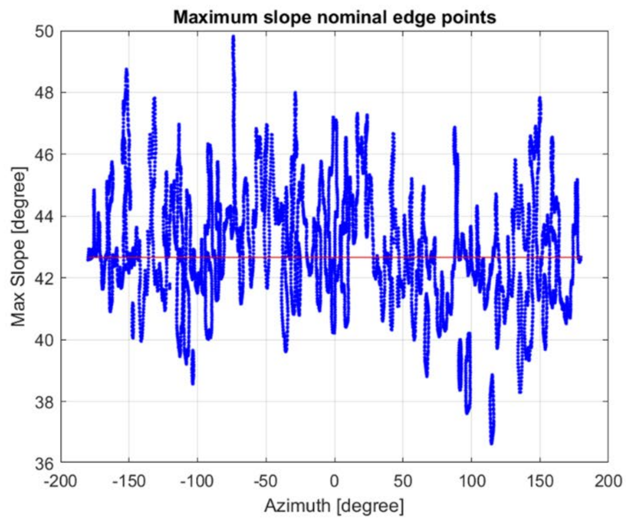


Figure 14 – MMS artifact prototype based on precision sub-FOV spheres, a) engineering drawing of the cylindrical blank, b) picture of the actual machined cylindrical blank.

Preliminary measurements were made on a 0.8 mm sphere and an analysis routine was developed to evaluate the MMS specification. The sphere was an off-the-shelf precision grade 3 440 stainless steel sphere with nominal roughness of 25 nm Ra. No surface modifications were made to make the sphere measurable. The preliminary measurements were performed on a focus variation microscope with coaxial lighting at 20X magnification without using a polarizing filter to reduce reflections. Figure 15a shows the topography measurement and Figure 15b shows the estimated maximum measurable slope for different azimuth angles. In the preliminary analysis procedure, we first limit the evaluation to the largest contiguous area where the instrument reported measurement values. This implies that data islands at the edges of the measured area are not considered in the analysis. These data islands are often small and typically contain outliers as they are at the very edge of the instrument's capability to obtain data. Next, we estimate a best-fit least-squares sphere through the remaining data points. Here we limit the estimation to surface points whose lateral position is within 90% of the radius of the circle through the edge pixels. The respective circle was estimated using a robust least-squares estimation procedure that ignores data points with a large distance to the circle. Finally, we calculate for each edge pixel the respective slope of the sphere surface. These slopes are shown in Figure 15b. The surface deviations relative to the best fit sphere, shown in Figure 16a, show a dark blue circle that is likely an illumination artifact.



a)



b)

Figure 15 – Positional MMS analysis on a convex sphere with a diameter of 0.8 mm, a) plot showing the topography data with a best fit circle (in red) through the edge points of the data, b) plot showing the slope of the edge points with the slope at the edge circle shown in red.

To improve measurability, chemical etching methods were explored for the stainless-steel spheres. Experiments were conducted on 10 mm diameter, grade 10, 440 stainless steel spheres to make it easier to visualize when the etching solution was altering the surface roughness of the sphere. The etching solution used was Ferric Chloride ($FeCl_3$). A variety of $FeCl_3$ solution concentrations and etching times were tried. The etching results shown in Figure 16 were achieved with an approximate 2 molar $FeCl_3$ solution and an etch time on the order of 30 seconds. This etch time consists of a 10 second undisturbed submersion in the etch solution and an additional 20 to 25 seconds exposure during removal of the sphere from the etch solution before being rinsed with distilled water. The measured roughness value before and after etching were Sq = 135 nm and Sq = 599 nm respectively.

Figure 16 – Picture of 10 mm grade 10 stainless steel spheres before (right) and after (left) $FeCl_3$ chemical etch

7.2.3.2 MMS artifact utilizing small concave hemispherical features

Recognizing that optical surface topography microscopes may not measure surfaces on a concave spherical geometry the same way as surfaces on a convex geometry, we considered another prototype design for evaluating the MMS specification. Differences in measurement between the two geometries may occur due to unwanted imaging of the light source (see Section 7.2.1) as well as internal reflections. The design incorporates an array of hemispherical concave impressions. A preliminary prototype was fabricated using a liquid resin-based 3D printer. No preliminary measurements or analysis of this prototype were performed during the course of the research. Exploring the differences between the convex and concave surfaces with respect to evaluating the MMS specification will be pursued in future efforts. Figure 17 shows the mechanical drawing and the 3D printed prototype. This artifact contained 5 different diameter hemispheres: 0.5 mm, 1.0 mm, 1.5 mm, 2.0 mm, and 2.5 mm. Each diameter hemisphere was also located at different depths (e.g., equator at, above, and below the top surface). The prototype artifact is designed to simulate firing pin impression geometries, giving us the ability to eventually evaluate firing pin impression measurements as a function of diameter and depth of impression (aspect ratio).

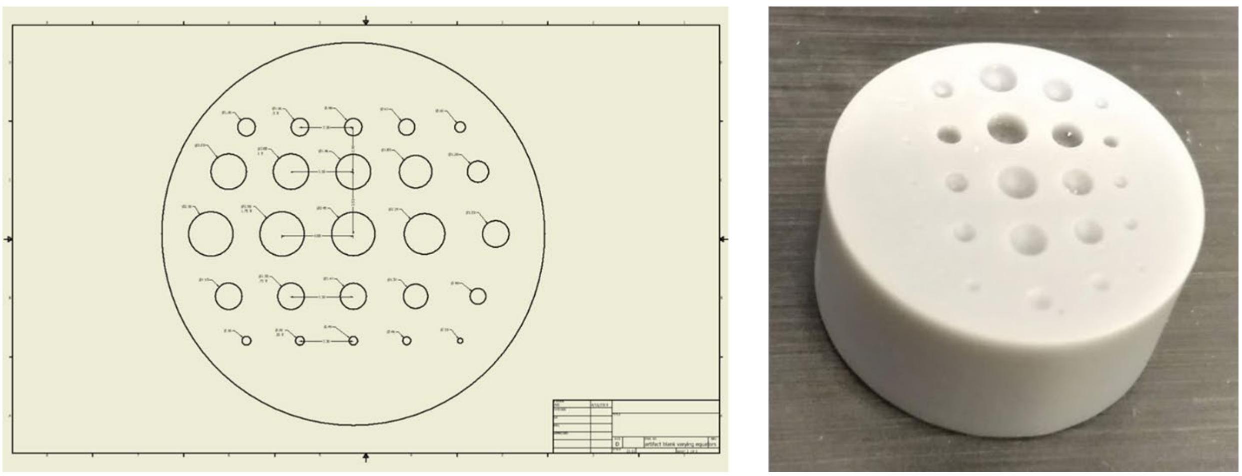


Figure 17 – Concave MMS prototype, a) – engineering drawing, b) 3D printed prototype

7.2.3.3 MMS prototype with flat discrete angles surfaces

In addition to the spherical geometry solutions explored for evaluating MMS, another prototype artifact was designed and fabricated based on flat planes at discrete angles. This design avoids any possible effects on the measurement from the focusing and image forming capability of the spherical geometries described above. The overall design includes 3 separate artifact parts, each with 2 planes at a discrete nominal angle, for a total of 6 discrete angles to evaluate MMS. The parts are machined out of stainless steel using wire EDM. The nominal discrete angles are 5 degrees, 10 degrees, 15 degrees, 30 degrees, 45 degrees, and 60 degrees. Figure 18a shows the three separate parts. Figure 18b shows a sample topography measurement of the six discrete angles obtained with a focus variation microscope using a 10X objective. The instrument was able to measure all planes.

Figure 18 – a) Three parts with two planes at different discrete angles, b) Sample topography measurement of the 6 angled planes and reference planes. In this measurement, all three parts are oriented side by side and imaged together.

7.2.4 Prototype artifact for Flatness specification evaluation

Depending on the optical method utilized in an optical surface topography microscope, different surfaces will be required to measure and evaluate flatness deviation. In general, some photometric stereo and confocal and interferometric based systems enable the use of a mirror surface. Other instruments will need a flat surface with some nominal roughness. The design for confocal, interferometric, and some photometric stereo microscopes incorporates an inexpensive commercial off-the-shelf high-quality protected aluminum coated mirror with a $\lambda/10$ flatness specification (at λ =633 nm wavelength). This mirror is attached directly to the cylindrical blank having the same size as a 12-gauge shotgun shell. Figure 19 shows a picture of the assembled prototype. No preliminary measurements or testing were performed on this prototype.

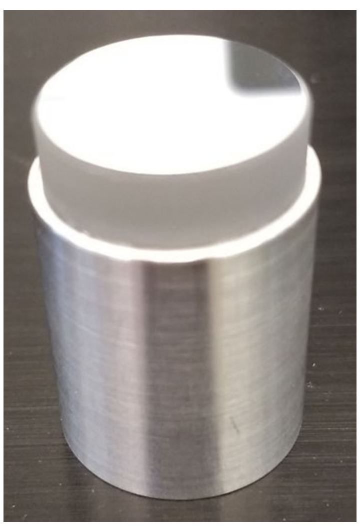


Figure 19 – Commercial off-the-shelf mirror attached to aluminum cylindrical blank for evaluating flatness deviation

7.2.5 Method for Evaluating Measurement Noise

As indicated in Table 2, a specific artifact was not needed to evaluate the measurement noise specification. This specification evaluation can be conducted on a variety of appropriate surfaces. The measurement noise estimate is affected by the characteristics of the surface measured. Flat surfaces typically yield the lowest values for measurement noise. However, the noise estimate is most useful when obtained for a surface that is representative of the surfaces that will be measured in practice. For focus variation and some photometric stereo based surface topography microscopes, the surface needs a minimal amount of roughness or contrast. This procedure can be performed on the NIST standard cartridge case (SRM 2461).

7.2.6 Method for Evaluating Measurement Stitching

As indicated in Table 2, a specific artifact was not needed to evaluate image stitching. In Figure 20, we show the results of applying the image stitching evaluation methodology described in section 6.0. In this case, the NIST SRM 2461 standard cartridge case was utilized to demonstrate the effect of image stitching error in the resultant stitched data set. The surface difference plot in in Figure 20a reveals the individual 9 FOVs that were acquired to create the stitched data set. If there were no image stitching related errors, the individual FOVs that make up the resultant stitched data set would not be easily observed in the difference plot. Figure 20b is shows the exact same data as Figure 20a, but with lines overlaid on the surface to highlight the location of the individual 9 FOVs.

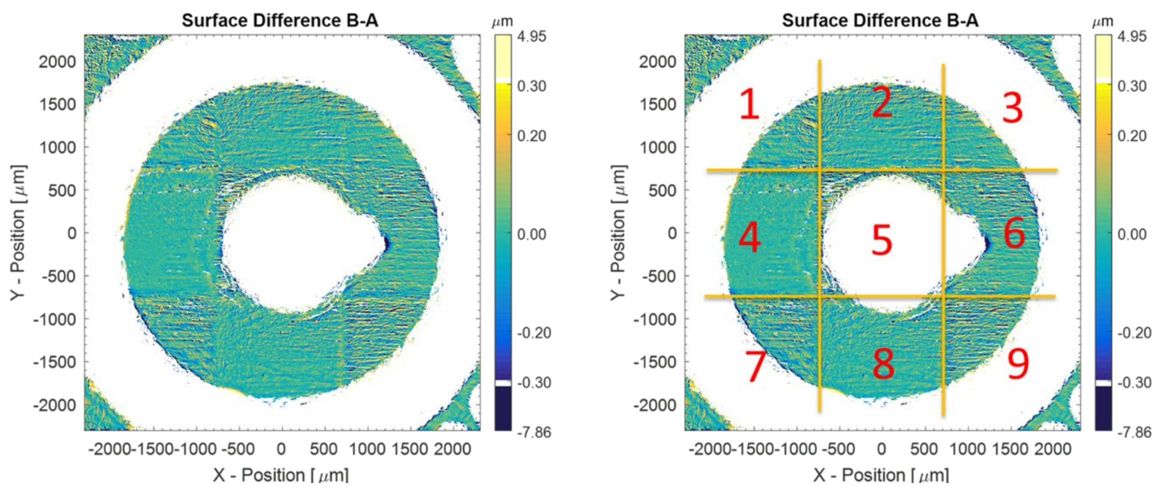


Figure 20 – a) Topography errors observed in 3 x 3 stitched data set of the NIST SRM 2461 standard cartridge case due to image stitching, b) yellow lines and red numbers highlight the 9 individual FOVs that were acquired to create the stitched data set

8.0 Quality Assurance Demonstration

The implementation of 3D optical surface topography measurement instruments in forensic laboratories for FATM analysis requires that appropriate quality assurance systems are in place to ensure that the measurement data obtained meets the requirements of the analysis. This may require calibration or performance evaluation procedures after installation and during operation. Guidance for implementing a quality assurance (QA) system for this purpose was not available to FATM practitioners at the beginning of this project (January 2017). Since then, OSAC has worked on drafting standards that begin to address quality assurance. The OSAC draft standard "Standard for 3D Measurement Systems and Measurement Quality Control" [25] provides basic QA guidance for implementing 3D technologies. The standard addresses issues such as 1) checking key instrument error sources (instrument scales) with a traceable reference standard, 2) ensuring that measurement values of a reference artifact and their uncertainty are consistent (overlap) with the certified values of the artifact and their uncertainty, 3) evaluating and documenting measurement uncertainties in accordance with the “Evaluation of measurement data – Guide to the expression of uncertainty in measurement” [41], and 4) performing periodic “check and recheck” measurements of reference standards to ensure continued data validity. The guidance provided in this OSAC document is relevant, and laboratories will need to follow the requirements set forth in it. Even with this document, however, a laboratory needs additional detailed information to fully implement a quality assurance system that will ensure that data obtained meets the analysis requirements.

Over the course of the research conducted for this project, we collaborated with the FBI Firearms and Toolmarks Unit (FTU) to develop practical guidelines for QA. The FBI has invested significant resources to research the application of 3D optical surface topography methodologies in FATM analysis. Our role in this collaboration is to provide guidance on metrology issues and assist in implementing quality assurance protocols for the various instruments they have in their laboratory. The prototype quality assurance protocols and their implementation data serve as a test case for the development of generic QA guidelines for 3D toolmark topography measurements in forensic labs. An overview of how the protocols are set up will be provided below. Eventually, the lessons learned will be applied to develop best practice guides within the TWG3D2T. The guides are expected to include application of reference artifacts developed by NIST as a result of this project.

In our collaboration, specific quality assurance protocols were provided for the different surface topography instruments owned by the FBI. To accomplish this, procedures were created for a quality system that includes conducting gauge repeatability and reproducibility (GR&R) [42]studies and setting up control charts for different metrological characteristics, depending on the instrument and the measurement application. At a minimum, the procedures ensured that all three scales of the instrument were checked at regular intervals. In addition, depending on the application, the procedures involved periodic measurements of either the NIST standard bullet (SRM 2460a) [43]or the NIST standard cartridge case (SRM 2461) [43]. These standard artifacts have a very specific and necessary role in a measurement assurance system. As the artifacts imitate real ballistic samples, they have toolmarks that exercise a surface topography instrument’s capability in the same way as real ballistic evidence samples. This provides an important overall check of the system. In contrast, the new artifacts developed in this study focus on the evaluation of individual metrological characteristics for calibration, uncertainty analysis, and diagnosis . The two SRMs are shown in Figure 21.

Figure 21 – a) NIST standard bullet (SRM 2460), b) NIST standard cartridge case (SRM 2461)

Here we provide an overview of a QA protocol developed for a confocal microscope within the FBI that is used to measure the land engraved areas of bullets at 20X magnification. The physical reference standards used in the FBI QA protocols were different than the physical reference standard prototypes described above, as the latter were not yet available. In setting up the QA system, three physical reference artifacts were used: 1) a custom step height standard (Figure 22), 2) a NIST SRM 2073a roughness standard (Figure 23), and 3) a NIST standard bullet (SRM 2460). The custom step height standard contained 3 different step heights (constructed from 4 adjacent gauge blocks wrung on an optical flat); 100 µm, 50 µm, and 10 µm. The 10 µm step was used as a scale check (amplification coefficient) for the z axis of the microscope. The NIST SRM 2073a is a sinusoidal profile roughness standard with certified values for roughness average (Ra) and lateral feature spacing (RSm), with corresponding nominal values of 3.0 µm Ra and 100 µm RSm. The NIST standard bullets (SRM 2460) were produced to have a certified high similarity of the land engraved striation profiles with a virtual reference profile. Specifically, NIST provides different virtual profiles for each of the six lands that are present on the bullet. The end user can compare their measurements of the land engraved profiles to these virtual profiles via the normalized cross-correlation function (Pearson correlation coefficient, CCF) or signature difference (SD) [43]. Only land #1 was tracked for QA purposes for the FBI QA protocol.

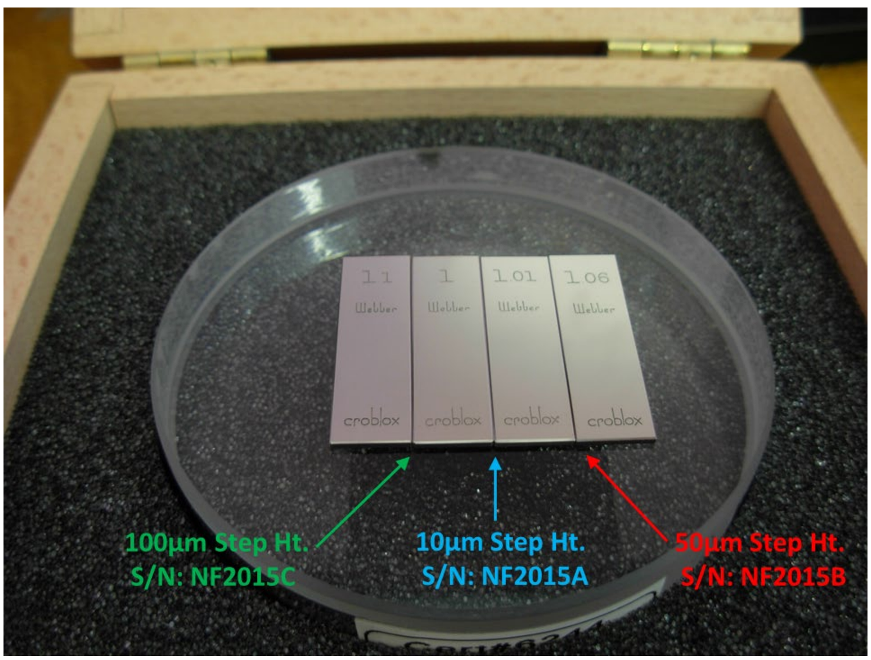


Figure 22 – NIST manufactured custom step height standard constructed from commercial chrome-carbide gauge blocks

Figure 23 – NIST SRM 2073a sinusoidal roughness specimen

The goal of a properly implemented QA system is to ensure that the measured topography data meets the requirements of the intended FATM analysis. As these requirements are not yet known for the various types of analysis, including comparisons performed by algorithms or VCM, we concentrate on ensuring that the data meets specific accuracy requirements and is in statistical control. In QA systems, a statistical tool known as a control chart [44] is used to monitor whether a given measurement process is in statistical control. Setting up a control chart involves first executing a gauge repeatability and reproducibility (GR&R) study to gather a sufficient set of baseline data. The assumption is that you are starting with a measurement instrument that is calibrated and stable. After the set of baseline data is taken, warning limits and control limits are established. At this point, the control chart can be used to monitor the measurement process. On an interval equal to when bullet measurements are taken for casework, the identified check standards should be measured and resultant values plotted on the control charts to determine if the instrument is in statistical control. The respective frequency is determined by the stability of the system, the cost of incorrect measurements, and the time available for any corrections. This QA system was designed to make check measurements before and after (recheck measurements) important data acquisitions (casework). If either the check or the recheck measurements indicate an out-of-control measurement process, then data acquired between those intervals should not be used. When this occurs, an investigation should be initiated to identify and correct the cause of the out-of-control condition.

There are many statistical rules regarding interpretation of control charts. This topic is beyond the scope of this report. At a basic level, measurements of the check standards are compared against the warning limits and the control limits. If a given check standard measurement exceeds a warning limit, it may require an evaluation of the measurement process to determine if there is a problem that needs to be addressed. If a given check standard measurement exceeds a control limit, it will require that the measurement process be halted and investigated for reasons for the out-of-control condition. Generally, “in-control” implies that all of the control measurements are within the control limits and are randomly distributed. Any type of obvious or systematic pattern of the data may indicate an issue that may merit further investigation.

The GR&R was designed to acquire 10 measurements back-to-back in 1 day to evaluate repeatability and 1 measurement per day for 10 days to evaluate reproducibility. In-between measurements the check standard was removed from the system to incorporate any setup effects in the analysis. The GR&R data provides a baseline of data, from which the mean line of in-control process data, warning limits, and control limits are calculated. The mean line of the in-control data is an average of the baseline reproducibility data taken once a day for at least 10 days. More reproducibility data can be taken to construct the control chart if desired. The warning limit is calculated by taking two times the standard deviation of this data and the control limits are calculated by taking three times the standard deviation. This is for a two-sided control chart, where you have a mean line and plus and minus warning and control limits. For similarity scores, such as the CCF for the bullet measurement, the control chart is one-sided. The highest CCF value that can be achieved is 100% and scores are always smaller than 100% (never greater than 100%). In this case, the mean line is still calculated the same way, but for warning limits equivalent to 95% probability, the multiplier is 1.645 and for the control limit equal to 99.7% probability, the multiplier is 2.83.

We show four sample control charts[2] that were generated according to the design described above; 1) standard bullet control chart (Figure 24), 2) z scale control chart (Figure 25), 3) x scale control chart (Figure 26), and 4) y scale control chart (Figure 27). The mean line, warning limits, and control limits for these sample control charts were calculated from a baseline set of data that included 30 measurements. These 30 data points are plotted on the control chart. Control data for monitoring the measurement process starts at point 31 for these sample control charts. The mean line for each control chart is indicated by the green line. The warning limits are indicated by the yellow lines. The control limits are indicated by the red lines. Measurement data taken for monitoring the measurement process are plotted in pink. For this test of the bullet measurement process on a confocal microscope, the control charts indicated that check measurements were in control for the standard bullet and for the x scale, but out-of-control data points were observed for the z scale and the y scale. This emphasizes the need to evaluate and check different aspects of a surface topography microscope to ensure it is generating quality topography data. Even though it was only a test of the measurement process control system, the out-of-control data points indicated a problem. Investigating the cause of the observed out-of-control data points in this particular case led to the replacement of the 20X microscope objective on the instrument.

[2] These are not official control charts utilized within the FBI FTU QA system. These were generated as a test of the QA system and are shown here for informational purposes.

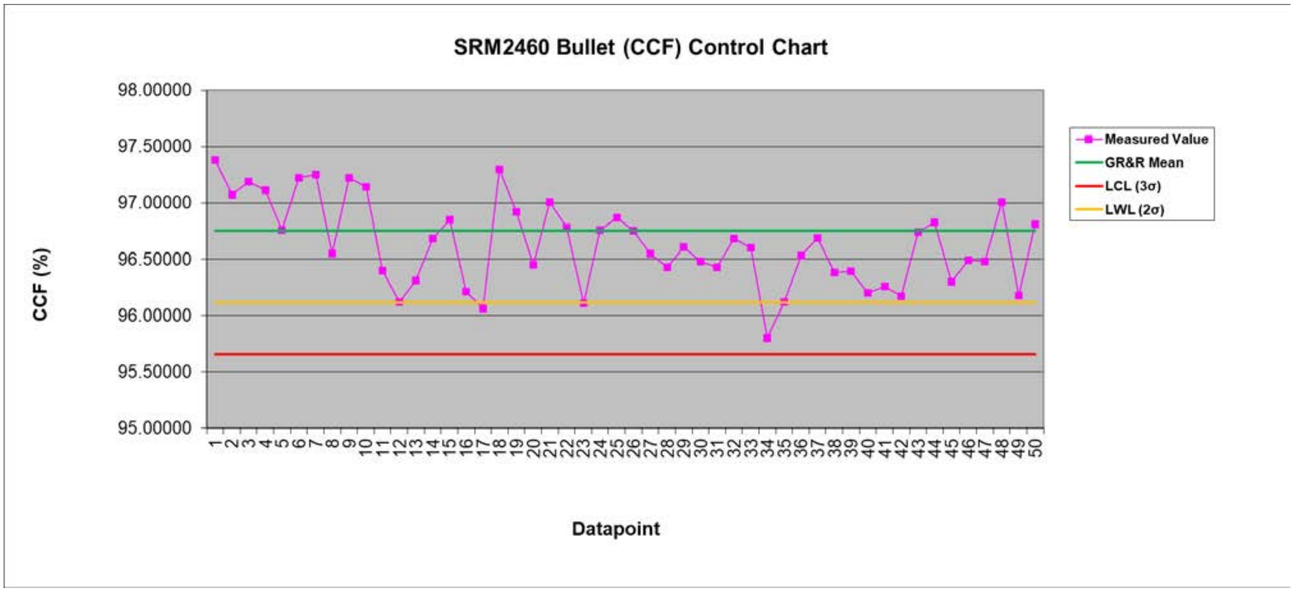


Figure 24 – Control chart for CCF using the NIST standard bullet

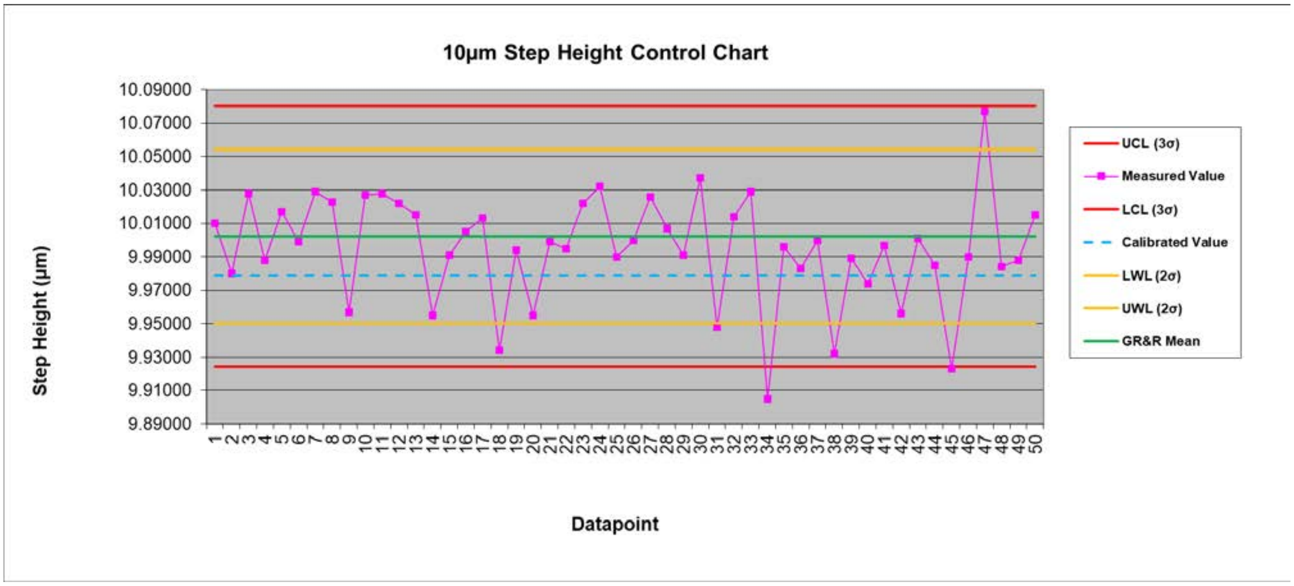


Figure 25 – Control chart for Z scale using a 10 µm step height standard

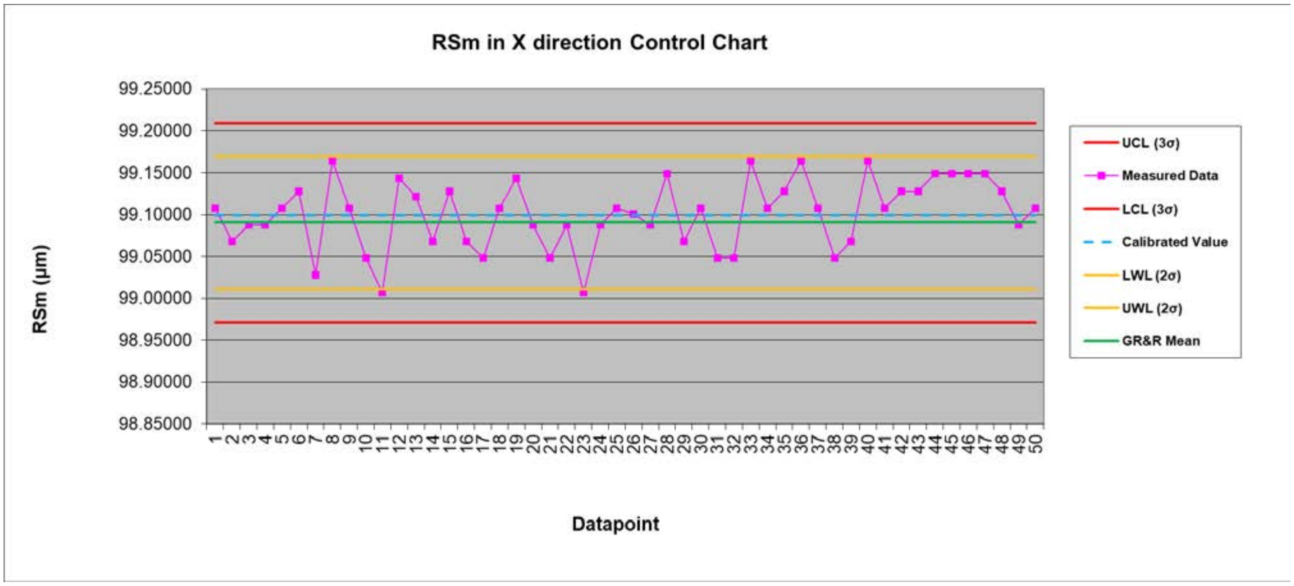


Figure 26 – Control chart for RSm in the x direction of the microscope using NIST SRM 2073a

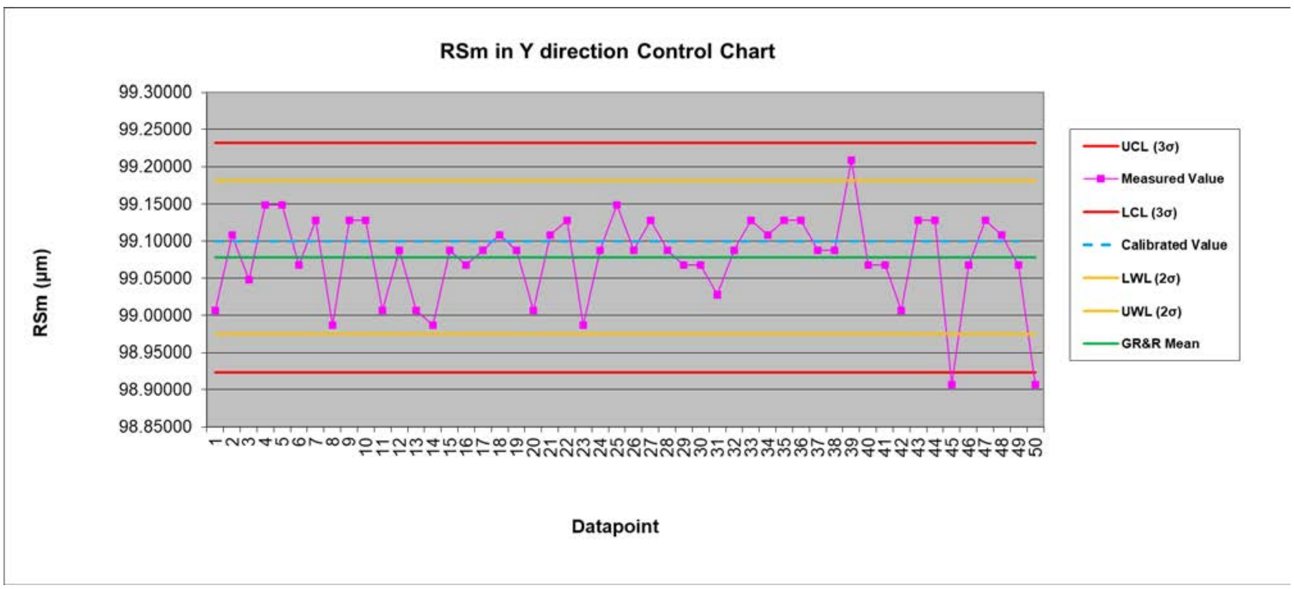


Figure 27 – Control chart for RSm in the y direction of the microscope using NIST SRM 2073a

9.0 Future Work and Implications of Research for the FATM Community

The research conducted during this project provides key building blocks to establish a solid metrology foundation for the successful implementation of 3D optical surface topography methodologies in the FATM forensic discipline. The prototype lateral array artifact (for X and Y scale, mapping deviations, and linearity deviations) and the prototype step height artifact (for Z scale and linearity) are currently in the NIST SRM Standards Development phase. In this phase we are refining the designs and are developing procedures for manufacturing and certification. Making NIST certified SRMs available to forensic laboratories to verify 3D microscope performance will save practitioners the task of researching what types of artifacts they will need to monitor and calibrate their 3D surface topography instruments. In addition, having artifacts that are measurable on the broad spectrum of surface topography instruments will facilitate uniformity in QA procedures and a more equitable comparison of microscopes available to the FATM community. The development of QA procedures and protocols specifically for FATM applications will also facilitate successful implementation of 3D methodologies in forensic laboratories. These QA protocols and procedures are currently being developed as best practice guides through the TWG3D2T.

One metrological characteristic listed in Table 2, but not addressed in this report with a proposed artifact design, is topographical spatial resolution. This metrological characteristic is important to evaluate and quantify for FATM analysis applications as it affects the level of detail in the measured surface topography data. Artifacts are readily available to characterize spatial resolution of 2D imaging systems. However, it is challenging to develop artifacts for characterizing the spatial resolution of various 3D topography measurement instruments, both for generic and FATM applications. Research in this area is ongoing [45].

10.0 Scholarly Products

Technical Presentations

Stocker, M., Zheng, A., Renegar, T., Soons, J., Thompson, R., "Performance Evaluation and Calibration Artifacts for 3D Ballistic Imaging", American Academy of Forensic Science 69th Annual Meeting, New Orleans, LA, February 2017

Journal Articles

Stocker, Michael T., et al. (2018) "Addressing Quality Assurance Issues in 3D Firearm and Toolmark Imaging| NIST." Journal of the Association of Firearms and Toolmarks Examiners, Vol. 50, pp. 104 – 111.

11.0 Participants and Collaborations

11.1 Participants

Four other co-principal investigators, all from the same recipient organization as the principal investigator (National Institute of Standards and Technology) contributed to this research project. First, a mechanical engineer, Dr. Johannes Soons, provided algorithm development and evaluation in additional to data analysis. Second, a physical science technician, Thomas Brian Renegar, provided measurements and mechanical design work for the project. Third, a mechanical engineer, Alan Zheng, provided consultation on overall project direction. Last, a forensic scientist, Robert Thompson, provided subject matter expertise on firearm and toolmark analysis to guide the research and development of the project.

Other NIST staff members who made significant contributions to the project include Dr. George Orji (photomask design), Dr. Ronald Dixson (atomic force microscopy), Rick Lake (machining), Casey Shatzley (machining), Eric Windsor (chemical etching), Dr. Lei Chen (lithography), Dr. Mark McLean (RTA of gold coatings), Dr. Kavuri Purushotham (scanning electron microscopy), and Dr. Pradeep Namboodiri (lithography).

11.2 Collaborations

During the course of the project, the FTU of the FBI provided measurement support from their diverse array of optical surface topography microscopes. In addition, they aided in development of QA protocols for the implementation of 3D instruments in forensic laboratories.

12.0 References

1. AFTE Glossary, 6th edition, Version6.110619
2. National Research Council. Strengthening forensic science in the United States: a path forward. National Academies Press, 2009.
3. Zheng, X., Soons, J., Thompson, R., Singh, S., Constantin, C., "NIST Ballistics Toolmark Research Database", J Res Natl Inst Stan, Vol. 125, No. 125004, Jan. 2020, https://doi.org/10.6028/jres.125.004
4. Chu, W., Thompson, R., Song, J., Vorburger, T., "Automatic Identification of Bullet Signatures Based on Consecutive Matching Striae (CMS) Criteria," Forensic Sci. Int., Vol. 231, No. 1, Sept. 2013, pp. 137-141.
5. Chumbley, L.S., Morris, M.D., Kreiser, M.J., Fisher, C., Craft, J., Genalo, L.J., Davis, S., Faden, D., and Kidd, J., "Validation of Tool Mark Comparisons Obtained Using a Quantitative, Comparative, Statistical Algorithm," Journal of Forensic Sciences, Vol. 55, No. 4, July 2010, pp. 953-961.
6. Ma, L., Song, J., Whitenton, E., Zheng, A., and Vorburger, T., "NIST Bullet Signature Measurement System for RM (Reference Material) 8240 Standard Bullets," Journal of Forensic Sciences, Vol. 49, No. 4, Aug. 2004, pp. 649-659.
7. Petraco, N. D., Chan, H., De Forest, P., Diaczuk, P., Gambino, C., Hamby, J., Kammerman, F.L., Kammrath, B.W., Kubic, T.A., Kuo, L., McLaughlin, P., Petillo, G., Petraco, N., Phelps, E.W., Pizzola, P.A., Purcell, D.K., Shenkin, P. , "Application of Machine Learning to Toolmarks: Statistically Based Methods for Impression Pattern Comparisons," NIJ Report 239048, National Institute of Justice, Washington DC, July 2012, <https://www.ncjrs.gov/pdffiles1/nij/ grants/239048.pdf>.
8. Song, J., "Proposed 'NIST Ballistics Identification System (NBIS)' Based on 3D Topography Measurements on Correlation Cells," AFTE Journal, Vol. 45, No. 2, Spring 2013, pp. 184-189.
9. Song, J., Whitenton, E., Kelley, D., Clary, R., Ma, L., and Ballou, S., "SRM 2460/2461 Standard Bullets and Cartridge Cases Project," Journal of Research of NIST, Vol. 109, No. 6, Nov.-Dec. 2014, pp. 533-542.
10. Vorburger, T.V., Song J., and Petraco, N., "Topography measurements and applications in ballistics and tool mark identifications," Surface Topography: Metrology and Properties, Vol. 4, No. 1, Mar. 2016, pp. 1-35.
11. Weller T., Brubaker M., Duez P., Lilien, R., "Introduction and Initial Evaluation of a Novel Three-Dimensional Imaging and Analysis System for Firearm Forensics," AFTE Journal, Vol. 47, No. 4, Fall 2015, pp. 198-208.
12. Lilien, R., "Applied Research and Development of a Three-Dimensional Topography System for Firearm Identification using GelSight," NIJ Report 248639, National Institute of Justice, Washington DC, February 2015, <https:// www.ncjrs.gov/pdffiles1/nij/grants/248639.pdf>.
13. Song, John, et al. "Estimating error rates for firearm evidence identifications in forensic science." Forensic science international 284 (2018): 15-32.

14. Song, John, et al. "Evaluating Likelihood Ratio (LR) for firearm evidence identifications in forensic science based on the Congruent Matching Cells (CMC) method." Forensic science international 317 (2020): 110502.
15. Petraco, N.D., Kuo, L., Chan, H., Phelps, E., Gambino, C., McLaughlin, P., Kammerman, F., Diaczuk, P., Shenkin, P., Petraco, N., and Hamby, J., "Estimates of Striation Pattern Identification Error Rates by Algorithmic Methods," AFTE Journal, Vol. 45, No. 3, Spring 2013, pp. 235-244.
16. Riva, F., and Champod, C., "Automatic Comparison and Evaluation of Impressions Left by a Firearm on Fired Cartridge Cases," Journal of Forensic Sciences, Vol. 59, No. 3, Feb. 2014, pp. 637-647.
17. Chu, W., Tong, M., & Song, J., "Validation Tests for the Congruent Matching Cells (CMC) Method Using Cartridge Cases Fired With Consecutively Manufactured Pistol Slides," AFTE Journal, Vol. 45, No. 4, Fall 2013, pp. 361-366.
18. Ultra Electronics Forensic Technology, Inc., https://www.ultra-forensictechnology.com/en/our-products/ballistic-identification/brasstrax
19. Bureau of Alcohol, Tobacco, Firearms, and Explosives, National Integrated Ballistics Information Network (NIBIN), https://www.atf.gov/firearms/national-integrated-ballistic-information-network-nibin
20. Duez, P., Weller, T., Brubaker, M., Hockensmith, R., Lilien, R., "Development and validation of a virtual examination tool for firearm forensics," Journal of forensic sciences, Vol. 63(4), 2018, pp. 1069-1084.
21. ISO 25178-6:2010(en) Geometrical product specifications (GPS) — Surface texture: Areal — Part 6: Classification of methods for measuring surface texture
22. Leach, Richard, ed. Optical measurement of surface topography. Vol. 8. Springer Berlin Heidelberg, 2011.
23. The Organization of Scientific Area Committees for Forensic Science, https://www.nist.gov/osac, Accessed February 22, 2021.
24. Technical Working Group for 3D Toolmark Technologies in Forensic Laboratories, https://twg3d2t.org/, Accessed on February 23, 2021
25. Firearms and Toolmarks Subcommittee, https://www.nist.gov/osac/firearms-toolmarks-subcommittee, Accessed February 23, 2021
26. ISO/IEC 17020 – Conformity Assessment – Requirements for the operation of various types of bodies performing inspection, 2012, www.iso.org
27. ISO/IEC 17025 – General Requirements for the competence of testing and calibration laboratories, 2005, www.iso.org
28. ISO 25178–70 2014 Geometrical product specification (GPS), —Surface Texture: Areal—Part 70: Material measures(ISO 25178:70:2014)
29. ISO 25178 part 600 2019 Geometrical Product Specifications, (GPS) – Surface Texture: Areal — Part 600: Metrological Characteristics for Areal-topography Measuring Methods
30. Giusca, Claudiu L., et al. "Practical estimation of measurement noise and flatness deviation on focus variation microscopes." CIRP Annals 63.1 (2014): 545-548.
31. Ott, Daniel, Robert Thompson, and Junfeng Song. "Applying 3D measurements and computer matching algorithms to two firearm examination proficiency tests." Forensic science international 271 (2017): 98-106.

32. Thomas, Matthew, et al. "Optical topography measurement of steeply-sloped surfaces beyond the specular numerical aperture limit." Optics and Photonics for Advanced Dimensional Metrology. Vol. 11352. International Society for Optics and Photonics, 2020.
33. Zhang, Qing, et al. "Edge-preserving photometric stereo via depth fusion." 2012 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2012.
34. Wang, Yinting, et al. "Detecting discontinuities for surface reconstruction." Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012). IEEE, 2012.
35. Leach, Richard, ed. *Optical measurement of surface topography*. Vol. 14. Berlin: Springer, 2011.
36. Foreman, Matthew R., et al. "Determination of the transfer function for optical surface topography measuring instruments—a review." Measurement Science and Technology 24.5 (2013): 052001.
37. Kovacs, Gregory TA, Nadim I. Maluf, and Kurt E. Petersen. "Bulk micromachining of silicon." Proceedings of the IEEE 86.8 (1998): 1536-1551.
38. Shirk, M. D., and P. A. Molian. "A review of ultrashort pulsed laser ablation of materials." Journal of Laser Applications 10.1 (1998): 18-28.
39. August Köhler, “A new system of illumination for photomicrographic purposes,” Mikroskopie, Vol. 10, pp. 433-440 (1893).
40. Sohn, Yeung Joon, et al. "Köhler illumination for high-resolution optical metrology." Metrology, Inspection, and Process Control for Microlithography XX. Vol. 6152. International Society for Optics and Photonics, 2006.
41. BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML, Evaluation of measurement data – Guide to the expression of uncertainty in measurement, Joint Committee for Guides in Metrology, JCGM 100, 2008, www.bipm.org
42. The NIST Engineering Statistics Handbook, “Gauge R & R Studies”, https://www.itl.nist.gov/div898/handbook/mpc/section4/mpc4.htm
43. Song, J., Whitenton, E., Kelley, D., Clary, R., Ma, L., and Ballou, S., “SRM 2460/2461 Standard Bullets and Cartridge Cases Project,” Journal of Research of NIST, Vol. 109, No. 6, Nov.-Dec. 2014, pp. 533-542.
44. The NIST Engineering Statistics Handbook, “What are Control Charts”, https://www.itl.nist.gov/div898/handbook/pmc/section3/pmc31.htm
45. Eifler, Matthias, Andrej Keksel, and Jörg Seewig. "Comparison of material measures for the determination of transfer characteristics of surface topography measuring instruments." *Surface Topography: Metrology and Properties* 7.1 (2019): 015024.