NISTIR 7266

Proceedings of the 2nd NIST LADAR Performance Evaluation Workshop – March 15 - 16, 2005

National Institute of Standards and Technology
Technology Administration, U. S. Department of Commerce

NISTIR 7266

Proceedings of the 2nd NIST LADAR Performance Evaluation Workshop: March 15 – 16, 2005

Geraldine S. Cheok, Editor
Building and Fire Research Laboratory

October 2005

U. S. DEPARTMENT OF COMMERCE
Carlos M. Gutierrez, Secretary
TECHNOLOGY ADMINISTRATION
Michelle O'Neill, Acting Under Secretary of Commerce for Technology
NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY
William A. Jeffrey, Director

ABSTRACT

In recent years, there has been a rapid growth of LADAR (laser detection and ranging) applications and a growing number of LADARs that are commercially available. As a result, there is a corresponding need to evaluate these instruments. This need provided the stimulus for the workshops conducted at the National Institute of Standards and Technology (NIST). The first was held on June 12-13, 2003 and the second workshop was held on March 15-16, 2005. Based on the first workshop, the second workshop focused on three topics: terminology, instrument characteristics that were of interest to the LADAR community, and test protocols. This report presents the proceedings from the second workshop.

Keywords: Laser scanning, LADAR, performance evaluation, standardization, terminology, test protocols, workshop.

ACKNOWLEDGEMENTS

The author would like to thank all the participants for their time and for sharing their thoughts at workshop.

Special thanks are extended to all speakers for their excellent presentations: Mr. John Palmateer of Boeing, Dr. Nell Sedransk formerly with NIST, Mr. Tom Greaves of SparPoint, and Dr. Wolfgang Böhler of i3mainz.

Also, very grateful thanks to Dr. David Gilsinn, Mr. Alan Lytle, Dr. Kamel Saidi, Dr. Nell Sedransk, and Dr. Christoph Witzgall for their invaluable help in setting up the workshop and for their insights on the workshop break-out issues. The author would also like to thank Mrs. Michele Abadia-Dalmau for her help in organizing the workshop.

DISCLAIMER

Certain trade names and company products are mentioned in the text or identified in an illustration in order to adequately specify the experimental procedure and equipment used. In no case does such an identification imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the products are necessarily the best available for the purpose.

The opinions expressed in the presentations by non-NIST authors are those of the non-NIST authors and are not necessarily the opinions of NIST.

POLICY

It is NIST's policy to use the International System of Units (SI). However, some of the units used in the workshop presentations made by the invited speakers are in U.S. customary units because of the intended audience.

CONTENTS

1.0 INTRODUCTION

A LADAR Calibration Workshop was held at the National Institute of Standards and Technology (NIST) on June 12-13, 2003. Prior to 2003, NIST was using LADARs (laser detection and ranging) in several research projects, and during the conduct of this work, the need to evaluate and determine the uncertainty of the measurements arose. This need was paralleled in the laser scanning industry which experienced both a rapid growth of LADAR applications and a growing number of commercially-available LADARs. Current LADAR applications include quality control, urban planning, mapping, surveying, autonomous vehicle navigation, global climatological monitoring, bathymetry (measurement of water depth), and homeland security. The objectives of the 2003 workshop were:

- to provide a forum for sharing and discussing current efforts in LADAR calibration
- to determine the type of performance evaluations and test protocols required
- to explore potential plans for the establishment /operation /location of a LADAR test facility
- to identify the physical requirements of such a facility.

The 2003 workshop participants expressed the need for standardization of testing and reporting and a neutral evaluation facility. Three common themes which emerged during the workshop revolved around the need for:

- definitions for commonly used terminology
- standard targets/artifacts with known reflectivity
- standard protocols for performance assessment/evaluation

A set of definitions for commonly used terms would *i*) help clarify manufacturers' specifications to enable meaningful comparisons between various commercially available instruments and *ii*) encourage uniform guidelines for manufacturers' specifications, testing, and reporting.

Based on the findings of the 2003 LADAR workshop [1], NIST has initiated a small, indoor, artifact-based facility for evaluating LADARs. Efforts included procuring a high resolution scanner, developing potential test artifacts, and acquiring appropriate laboratory space[1], and determining necessary modifications to the space. In determining the types of artifacts, input from LADAR users and manufacturers was essential since the type of artifacts are dependent upon the application. For example, if a geometric model is sought, accurate physical dimensions are important, or if the ability to discern a small object or feature is needed, knowledge of instrument resolution is important. As determined in the 2003 workshop, the latter example points out the need for definitions for commonly used terminology – in this case, what is meant by "resolution"?

---

[1] The initial space allocated for the indoor performance facility is adequate though not ideal in terms of size or environmental control. Continued efforts to improve the facility will be pursued.

A second workshop was held at NIST on March 15-16, 2005 to solicit input from LADAR manufacturers, end-users, and researchers on commonly used terminology and their definitions, measurements of interest, and types of artifacts for use in a performance evaluation facility. The objectives of the workshop were to:

- review and modify preliminary draft definitions of commonly used terms
- determine the types of measurements and levels of accuracy needed
- determine what artifacts are needed to evaluate the measurement

The proceedings from this second workshop are presented in this report.

2.0 WORKSHOP AGENDA

March 15, 2005

7:45 – 8:30	Registration
8:30 – 8:40	Welcome
8:40 – 8:55	Introduction: U.S. Measurement System and NIST – Jim St. Pierre, Acting Division Chief, Materials and Construction Research Division
8:55 – 9:25	“LADAR Angular Resolution” – John Palmateer, Technical Fellow, Boeing
9:25 – 9:45	“Terminology” – Nell Sedransk[2], Division Chief, Statistical Engineering Division
9:45 – 10:15	Break-out groups: Topic 1: Terminology
10:15 – 10:30	Break
10:30 – 12:30	Break-out groups
12:30-1:30	Lunch
1:30 – 2:30	Reconvene - Discussion of Topic 1
2:30 – 3:00	“A Survey of Industry Expectations” – Tom Greaves, Senior Analyst, Spar Point Research
3:00 – 3:15	Break
3:15 – 4:45	Break-out groups: Topic 2: Types of Measurements required
4:45 – 5:30	Reconvene - Discussion of Topic 2
5:30	Adjourn for Day 1

---

[2] Dr. Sedransk has since left NIST.

March 16, 2005

8:00 – 8:30	“Testing Laser Scanners - Experiences at i3mainz” – Wolfgang Böhler, Professor, i3mainz, Institute for Spatial Information and Surveying Technology
8:30 – 10:00	Break-out groups: Topic 3: Artifacts required to evaluate a particular type of measurement.
10:00 – 10:15	Break
10:15 – 11:00	Continue break-out session
11:00 – 12:30	Reconvene - Discussion of Topic 3
12:30 – 1:00	Wrap-up – Next Steps
1:00	Adjourn

3.0 SUMMARY OF BREAK-OUT DISCUSSIONS

The break-out sessions were recorded (audio only) and the material in the following sections is based on transcriptions of these tapes and on notes taken by various individuals in each group.

3.1 MARCH 15, 2005 – BREAK-OUT SESSION 1

The main topic of discussion for the first break-out session was the draft terminology (provided in Appendix B. Included in Appendix B are some suggested terms received via email from participants who could not attend the workshop). In particular, the participants were asked to review the definitions for accuracy, uncertainty, and resolution. Other terms that were to be discussed were the acronyms LADAR and LIDAR (light detection and ranging), and the participants were also asked to suggest terms that should be added to the list.

It was suggested that the term range "accuracy", as used in most manufacturers' specifications, is not appropriate and the term "uncertainty" should be used instead. This suggestion was received with mixed feelings because:

- The marketplace understands the term accuracy and it is commonly used.
- The term uncertainty implies "doubt" and therefore conveys a negative connotation. Situations where positive connotations are important include a court of law where the implication of doubt is seized upon. Similarly, it is very difficult to explain to a jury what is meant by uncertainty whereas jurists would understand the term "accuracy".

3.1.1 Group 1

Group participants:

John Palmateer (Chair)
Jack Cothren
Eric Hoffman
Steve Phillips
Wolfgang Böhler
Jim Filliben
Geoff Jacobs
Scott Ackerson
Marc Cheves
Brent Gelhar
Alan Lytle

It was suggested that terminology has already been established by the surveying community and manufacturers of survey equipment and that this could be used as a starting point. Another source of definitions for terminology is the engineering statistics handbook which may be found at: http://www.itl.nist.gov/div898/handbook/glossary.htm.

The suggested definitions, as taken from the VIM (International Vocabulary of Basic and General Terms in Metrology, ISO [International Organization for Standardization] 1993, Ref. 2), were not familiar to some in the group who were more familiar with definitions used in the surveying community. The VIM defines general terminology such as uncertainty and accuracy. It does not define application/technology terms such as LIDAR and volumetric accuracy nor does it define test protocols. In Europe, ISO standards are often adopted as the national standard.

The point was made that if the objective of the workshop efforts was to standardize test protocols, terminology that is globally accepted (i.e., adopted by ISO and ANSI) is necessary, and the use of such terminology from the start would ease the transition from converting "best practices/guidelines" to standards. Also, a single set of terminologies allows for consistent interpretation of test results whether from an accredited lab, a national lab, a manufacturer, or an end user. The general feeling within the group was that it didn't really matter which definitions were adopted – just so long as there is a single set that is adopted and accepted by everyone.

The issue of accuracy generated a lot of discussion. The development of test protocols to determine the uncertainty of the range measurements was acceptable to the group. A comment was made that "accuracy" would ultimately be defined in terms of the tests defined in a standard test procedure. As standard procedures are fairly generic, the reported accuracy may not be applicable to objects/targets that are different from that specified in the standard test. Basically, this was the crux of the problem: the accuracy as stated in the manufacturer's specifications is often used as the accuracy for <u>all</u> situations (applications). At issue, also, is the fact that the field conditions are different from those in the lab under which the tests were performed, and therefore, the uncertainties are unknown. Thus, it was felt a collection of terms is not sufficient to define the accuracy of a system. A test protocol which encompasses a range of parameters such as object/target material, curvature, and environmental conditions would provide some metric of the performance of the system. It was suggested that in view of the numerous factors affecting the range uncertainty, an experiment design (orthogonal, fractional, factorial designs) should be considered, and the test protocols should have a good statistical basis.

The issue of the large number of variables came up several times in the discussions, and the consensus was that it would be impossible to develop test protocols to account for all possible variables (environmental conditions; object/target type, reflectivity, material and texture; instrument set-up; point density; etc.) and combinations of these variables. A suggestion was made that a computer model could be developed to allow a user to enter relevant information so that the uncertainties could be predicted for whatever application/conditions arose. The basis of such a model would be the results obtained from a characterization of the instrument.

A point was made that there should be a differentiation between the uncertainty of a point measurement and a target uncertainty. The need for this distinction is due to the fact that an average of multiple measurements is used to determine the range to a target whereas only one measurement is used to determine the range to a point.

In addition to the uncertainty of the range measurement ("instrument uncertainty"), the determination of the "task specific uncertainty" is required – "how accurately can I measure the distance from point A to point B in the 3D model?" This accuracy, "project accuracy" as suggested by someone in the group, is a combination of the errors from image registration, calibration, and other procedures used to create the end product.

A comment was made that the issues, questions raised, and discussions in this workshop were similar to those raised by people using laser trackers several years ago. For laser trackers, a group was formed consisting of vendors/manufacturers, end-users, and researchers to draft U. S. national standards for laser trackers. The terminology is made up of 1) a descriptive terminology

– a description of what is meant by certain things and 2) testing protocols – how do you determine accuracy. A question was raised regarding the impetus behind the laser tracker effort – was it driven by vendors, was a consortium formed? There are only a few organizations within the United States that are authorized to create national standards. The process requires a balance of interest between vendors/manufacturers, users, and academia/researchers and if there is sufficient interest, a work statement is developed and a standards working group is formed under the auspices of an authorized standards organization.

The discussion then centered on what was the best path forward? Should NIST do it or should it be someone else? Many universities and firms have developed their own test facilities and test protocols and have conducted their own tests. However, objectivity of some of the findings and validity of the test procedures were flagged as important considerations. Access to facilities was, likewise, a point of concern.

The required level of accuracy of the reference instrument was also discussed. Currently, standard survey instruments such as total stations are being used to verify LADAR measurements. The accuracies of some LADARs are on the same order of magnitude as total stations. The existence of a facility where reference measurements of an order of magnitude higher than the LADAR measurements would be beneficial. However, the facility would have to be "football field" size. A comment was made that there were facilities for verifying GPS (global positioning system) receivers and totals stations that may also be used for evaluating LADARs. The question was then asked – "Can a facility the size of this room (L x W x H: 12.5 m x 9.4 m x 3.6 m) give you the information you need?" The answer was "some." This led to the inquiry of whether the results could be scaled?

There was also discussion on the need for some sort of factory-floor/field validation process to ensure that the instrument is functioning properly. Similarly, some companies have developed some QA (quality assurance) protocols to ensure valid measurements – e.g., specifying the required number of data points or setting a minimum number of points, specifying the number of control points.

When asked how the current instruments were being evaluated, the answer was surveyors were used to check the measurements at the time of purchase and for each project. However, the quality of the survey/surveyors varies. A comment was made that land surveyors were licensed whereas industrial surveyors were not. Another comment was made about the need to have a method to enable fair comparisons between survey measurements and LADAR measurements.

The following objections were made against the use of the acronym LADAR:

- The term is not used in the marketplace (nor in Germany) except in the military
- LADAR is similar to the term LIDAR which suggests an instrument just for ranging; whereas LADAR provides 3 measurements – range, and 2 angles (intensity is commonly reported also).

The term laser scanner was suggested but this may exclude triangulation-based devices and instruments that use structured light. The term 3D scanner was suggested as the output of these devices is 3D coordinates. Another term that was suggested was 3D imaging because one can recognize an object by examining a point cloud as you would in a photograph.

Another suggestion was to include some terminology regarding point density as it affects the uncertainty of the model. Other participants felt that this should be accounted for in the initial project planning – determine the number of scans and point density to achieve a desired level of accuracy. The group also felt that there was a lot of confusion with regards to the term “resolution.” A descriptor such as “range” or “object” was suggested for resolution, i.e., range resolution or object resolution. Other terms suggested for addition to the terminology list were beam divergence, beam spot size, rated conditions, limiting conditions, and registration (of multiple data sets).

A final question raised by the group was what will be done with the glossary/where will it reside. The glossary, in part, could potentially end-up as part of a standard for LADARs. Clarification was then sought about NIST’s role as it was stated earlier in the workshop that NIST does not create standards. NIST’s role is to facilitate the establishment of standards. The output/results of the efforts by NIST and by ISPRS (International Society of Photogrammetery and Remote Sensing) can be used to develop standard test protocols through the auspices of ANSI (American National Standards Institute) or ASME (American Society of Mechanical Engineers) or ISO (International Organization for Standardization). A comment was made that even if the test protocols did not follow the formal standardization route, specification of terminology through operational tests is needed. For example, range uncertainty is defined as “**…**” and it is measured *this* way under *these* conditions. If these test protocols are well specified, they can be used by anyone anywhere, and it will not have to be conducted at any particular facility.

3.1.2 Group 2

Group participants:

Jim Flint (Chair)
Carl Adrian
John Battaglia
Bruce Borchardt
Tom Greaves
Ted Knaak
Stefan Leigh
Gregory Lepere
Joe Liadsky
Kamel Saidi
Dan Sawyer
Nicholas Vandapel
Chris Witzgall

The discussion started with the meaning of the term “accuracy” – how is it defined? A definition offered by one of the group members was that accuracy was defined as the difference between the measurement and the “truth”.

EDITOR’S NOTE*: The accuracy of a measurement is defined by the VIM (International Vocabulary of Basic and General Terms in Metrology, [2]) as the closeness of the agreement between the result of a measurement and a true value. It is also noted in the VIM definition that “accuracy” is a qualitative concept and the term “precision” should not be used for “accuracy.”*

The difference between the measurement and the "truth" is measurement error. Measurement error is a measure of accuracy, e.g., the smaller the measurement error, the more accurate the instrument.

The term precision was defined as the tightness or looseness of a group of measurements around what is being measured (measurand).

EDITOR'S NOTE*: Precision is defined by ASTM E456-02 as the closeness of agreement between independent test results obtained under stipulated conditions. It is also noted that "precision" does not relate to the true value and the measure of precision is usually expressed in terms of imprecision and computed as a standard deviation of the test results. Less precision is reflected by a larger standard deviation.*

Figure 3.1 illustrates the difference between "accuracy" and "precision."

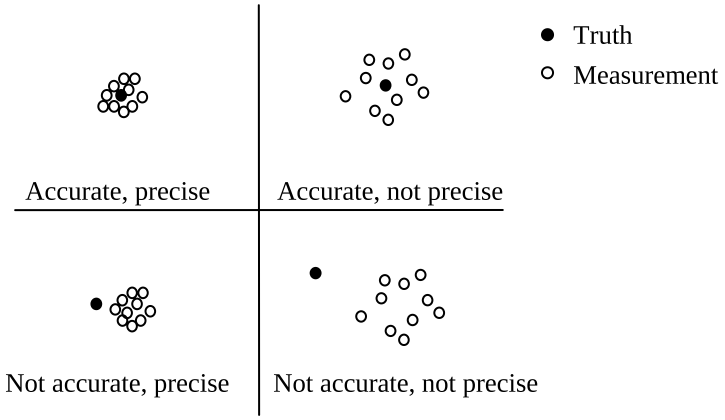


Figure 3.1. Accuracy vs. Precision.

Another issue that arose was the required accuracy of the reference measurement. It was mentioned that the accuracy of the reference measurements currently used in field work is on the same order as what is being evaluated. A point was made that the level of "accuracy" and "precision" required was dictated by industry and customer demands.

A point was made that definitions by themselves are insufficient to give a user the required information about an instrument. For example, an issue important to end users is - what is meant when a manufacturer specifies that the instrument accuracy is ± 2 cm? Does it mean that every measurement is within ± 2 cm of the truth? Does this hold for the entire range – from the minimum range to the maximum range? The general consensus was that the latter case was not possible and that the published specifications are based on results obtained for optimal conditions. The group felt that "validity conditions" or rating conditions had to be specified under which the specification was valid; two important conditions being range and temperature. Additionally, in determining the uncertainty of the range, it is important to be able to determine the uncertainty as a function of the angle of incidence of the laser with the target. It was suggested that multiple shapes should be used in the evaluation process and that a consensus on the geometry of the shapes and targets were necessary. Basically, in addition to definitions, test protocols are necessary to “define” how the uncertainty is determined or measured.

Related to the validity conditions, the concept of maximum permissible error (MPE) was discussed. A manufacturer would provide these values based on a wide range of tests that encompass a subset of the measurements that can be obtained with the instrument. Using standard test procedures, users can verify their instruments by comparing their results with the MPE.

A question arose about systems which included other means (such as on-board sensors) to augment the accuracy of the LADAR data. Should the specification be for the whole system or just for the LADAR? For example, if the LADAR data were combined with photogrammetry, the lateral resolution of the system may be improved because of the better edge detection capability from photogrammetry than if the lateral resolution were determined based solely on LADAR data. It was suggested that providing the specifications for both the LADAR alone and for the system would be appropriate as the instrument may be purchased without the additional on-board sensors.

There was also discussion about the need to develop protocols for evaluating the 3D models generated using LADAR data. It was argued that the advantage of LADARs was the ability to capture huge amounts of data which readily lends itself to the generation of models. It was also argued that the generation of models involves more than one point and therefore, the accuracy of the model is improved by the averaging of many points. There was also discussion about the need to evaluate the whole instrument or the system – hardware and software – as software could be used to improve the accuracy of the results. This led to the suggestion that the software used to obtain the specified accuracy should be identified in the specifications.

A suggestion was made to group the instruments: 1) Object scanner – small, medium, large object or 2) distance scanner – short, medium, long range. These groupings would aid a user in the selection of an instrument in that instruments of similar capabilities need be evaluated.

With regards to the acronym LADAR, this acronym is not commonly used in the laser scanning industry. However, it was suggested that LADAR be defined as terrestrial LIDAR where LIDAR commonly refers to aerial scanning. Another distinction between the acronyms LADAR and LIDAR was pointed out: in additional to aerial scanning, LIDAR is also used for remote sensing of gases/aerosols in the atmosphere.

In terms of useful information to include in the specifications, the validity conditions under which the specifications hold is important information. Results obtained under the optimal conditions and the extreme conditions would also be beneficial and would allow a user to extrapolate for conditions in between.

Considerable discussions concerned the fact that some instruments automatically take several measurements but report only the mean. Was it fair to compare the precision and/or timing of such instruments to those of instruments reporting single measurements? It was expressed that protocols for determining range accuracy and the time to acquire a scan were needed. It was suggested that knowing the beam characteristics would be useful. Another characteristic was the ability of an instrument to resolve two targets. For geometric objects, in addition to shape, the object material is also important. Currently, measurements from a total station are used as a basis for comparing LADAR measurements as the total station measurements have already been admitted in Federal and State courts. Additionally, the use of a total station is familiar and accepted by the marketplace.

With regard to the term “resolution”, a suggestion was made that resolution was point spacing at a given distance and that the term resolution should not be used as it could be confused with how it is used in photography which is 2D whereas LADAR is 3D. There appeared to be some confusion as to what was meant by resolution and some thought that accuracy and precision were synonymous with resolution.

In the discussion of angular resolution, it was recognized that angular resolution is not just the encoder resolution but also include some factors such as scan speed, laser pulse rate, and laser spot size. Some felt that angular resolution and spatial resolution essentially yielded the same information as the number of points at a given distance depends on the angle step size specified. Information on the minimum angle step size, i.e., “how many points can be obtained over a given angle?”, was thought to be useful. From an end users perspective, some examples of resolutions include: fracture/crack detection – what size crack and at what distance, size and shape of objects, and terrain features such as a crater. A question was asked about similar tests or artifacts for laser trackers. Point-to-point links were used with the manufacturer providing the equation which specifies the uncertainty of any point-to-point link. It was suggested that protocols could be developed for different applications, e.g., crack detection, but the manufacturers and users have to decide what subset of tests are representative of the needs of all users.

A source of error that needed to be accounted for is the uncertainty in locating targets since targets (e.g., spheres) are commonly used to register different scans.

An issue that was brought up but not discussed was eye-safety.

***EDITOR’S NOTE**: In the group discussion, a point was made about the need to compare a manufacturer’s report to a NIST report. This point was quickly corrected by another participant – NIST does not and will not publish any reports comparing instruments. NIST will help develop standard protocols that will be submitted to a standards organization for adoption: the types of protocols will be decided by the requirements of the laser scanning industry. A manufacturer may then report results as conforming to this XYZ standard, and a user may verify these results by conducting the same tests per the XYZ standard in an in-house lab or in an accredited facility.*

3.1.3 Group 3

Group participants:

Steve Hand (Chair)	*Gerry Cheok*	*Allen Cheves*
Les Elkins	*Tyler Estler*	*Dave Gilsinn*
Juergen Gittinger	*Mark Klusza*	*Dirk Langer*
Dave Ober	*Fred Persi*	*Omar-Pierre Soubra*
Bill Stone		

The session started off with a discussion of the two documents referenced in the draft terminology: the VIM and the GUM (Guide to the Expression of Uncertainty in Measurement, [3]). It was pointed out that NIST has a version of the GUM which may be downloaded from the following website: http://nvl.nist.gov/nvl2.cfm?doc_id=86 (Technical Note 1297 or TN 1297). A simple example of range uncertainty was discussed – measure the range to a distant target. The range is equal to the best estimate of the range corrected for all recognizable systematic errors (e.g., refractive index). However, there are uncertainties associated in the determination of the systematic errors and guidance on how to calculate/combine these uncertainties is given is the GUM. In the example, the best that can be said about the range measurement is that the range lies in the interval: best estimate of the range ± 2U where U is the uncertainty. Associated with this statement is a level of confidence in the measurement – the “2” in the previous statement is a coverage factor, *k*, and defines an interval having a level of confidence of 95 %.

Measurement error was also discussed in that error (measurement value – true value) is unknown as the true value is unknown. However, measurements from an instrument that has lower uncertainty (more accurate) than another (e.g., interferometer vs. LADAR) can for all practical purposes be used as the true value. In regards to the simple example which was a 1D measurement, a question was raised regarding the fact that LADAR is inherently 3D and can the uncertainty be propagated to 3D. Using volume as an example of a 3D measurement, it was noted that calculating the volume of a precision sphere was quite different than calculating the volume of a pile of dirt because what is meant by the volume of dirt has to be clearly defined – does it include voids and air pockets – and the propagation of the uncertainty would likely be more complicated.

A method for determining the uncertainty of a spherical measurement – angle, angle, range – for a particular instrument was briefly described. The method involved the use of tooling balls placed throughout the measurement volume of the instrument. Measurements to each of the tooling balls are made from the same location but with the scanner rotated through various angles and from various other locations, and the inter-ball distances from all locations are calculated. The mean and residuals from the mean of the inter-ball distances can then be determined and from these, a conservative estimate of the uncertainty of the system can be determined. A method to verify the angle uncertainties was through the use of a commercially available software package. The range uncertainty can be verified with an interferometer. The use of tooling balls for long range instruments would not be possible. The size of objects, in an object field, would have to be such that there will be a sufficient number of data points on the object to derive the center of the object.

Another method of determining range uncertainty was using a planar target in a calibration facility where the "true" range can be obtained and as a result, the systematic error can be determined over the entire range of the instrument. The noise would be the distribution of data points about the mean. The determination of the angular uncertainties would be more complicated as it involves both the range uncertainty and the beam spot size which is dependent on distance. It was pointed out that the material of the target affects the uncertainty of the measurements.

The uncertainty of a single point vs. the uncertainty of the model was discussed. The latter uncertainty is more of a concern for a user than a manufacturer.

With regards to the VIM definition of accuracy, a suggestion was made to replace "true value" with "reference value" in the definition. Caution was advocated in attempting to re-write the definitions of universally adopted terms. A suggestion was made to drop the use of the term accuracy, but it was felt that the term is too well entrenched in the marketplace and it would be very difficult to change it. A point was made that there would be a transition period from the use of the term "accuracy" to "uncertainty" as occurred in the CMM (coordinate measuring machine) community. This change was brought about by the development of standards for performance evaluation and was driven by manufacturers and users.

In terms of the reference measurements against which the LADAR measurements are to be compared, the question of whether the reference instrument should be 1, 4, or 10 orders of magnitude better than the LADAR instrument was asked.

The terms "repeatability" and "reproducibility" were also discussed. Reproducibility was felt to be an important characteristic due to the varied environmental conditions encountered in the field and the skill level of the operators. Knowledge of the repeatability or noise is necessary; however, knowledge of both systematic and random components of uncertainty is needed. The uncertainty from random components or noise can be reduced by averaging and systematic errors will have to be determined in a calibration facility involving the use of length standards.

The topic of color was also discussed as more systems have a means of capturing the color data in a scene. This being the case, a suggestion was made that there should be a means of testing the color and verifying the registration of the color information to individual points in the point cloud.

Resolution was briefly discussed. A differentiation between range and lateral resolution had to be made and a definition of lateral resolution was needed. Lateral resolution would have to be quantified as a function of range.

A suggestion was made to add the term "compensation" to the list of terminology. There were opposing opinions as to whether compensation was equivalent to calibration. One view was that compensation was a software correction to the instrument to adjust to the working environment whereas a calibration was physical adjustments to the instruments performed by the manufacturer.

In summary, the general feeling was that the word “accuracy” will not disappear from the vocabulary and will still be commonly used. It was felt that what was more important than having standard definitions of accuracy or uncertainty was having standard tests to “define” the accuracy of an instrument. Without these standard tests, manufacturers will continue to report values based on their own test protocols. This leads to the potential for a customer to be mistaken in thinking one instrument is more accurate than another when that may not be the case as a fair comparison is not possible without standard tests (e.g., the uncertainty of an instrument may be 2 mm using one protocol and 10 mm using another protocol). Additionally, guidelines should be given when reporting the uncertainty, e.g., 1U or 2U, so that there would be no confusion in interpreting the specification.

3.2 MARCH 15, 2005 – BREAK-OUT SESSION 2

The topics of discussion for this break-out session were basic instrument characteristics that were of interest and LADAR applications.

Some suggested instrument characteristics were:
- Range uncertainty as a function of target reflectivity, range, and angle of incidence
- Pointing/angular uncertainty
- Repeatability
- Reproducibility
- Correlation
- Instrument resolution : Range (depth), Horizontal, Vertical

Some suggested applications were:

- Quality control - extracting geometric dimensions
- Volume determination
- Object identification
- Pose determination
- Ground truth determination
- Vegetation assessment
- Mobility – autonomous navigation, crash avoidance
- Interference checking

Among the varying applications, the participants were asked to determine generic tasks common to these applications, i.e., what measurement needs to be evaluated.

The participants were also asked to rank the characteristics and to suggest test protocols for the top one or two characteristics.

3.2.1 Group 1

Clarification of the difference between “Repeatability” and “Reproducibility” was sought. The group felt that both these characteristics were important. Reproducibility was important because of variation in the environmental conditions and operators (training and skill level). Repeatability was also important to help determine systematic errors.

The object/target characteristic was discussed. The material of the object was thought to be important because of the absorption/penetration of the laser into the object. A suggestion was made that due to the numerous materials encountered in the field, it was impractical to have a facility provide these materials; rather, it should fall onto the users to conduct their own tests. The object color was also another important characteristic – standard photographic paper charts were suggested. The specular characteristic of an object affects the measurements – highly specular objects (mirrors) made poor targets. A comment was made that to ensure the quality of measurements in a project, decisions were made prior to scanning as to the point density and the maximum usable range (points greater than a certain range would result in a point spacing beyond the required project resolution and would be deleted or ignored).

Pointing/angular uncertainty – there was some discussion on this topic but no consensus was reached.

A question was asked about temperature effects and if they were compensated within the instrument – how it is done is often proprietary information. Other considerations were temperature of the object [from field experience, because of the size of the objects scanned, the difference between scanning an object when it is hot and when it is cold could be as much as 152 mm (6 in)] and the temperature gradients surrounding these objects.

One participant stated that the level of accuracy that is commonly required is about ± 6 mm (1/4 in) which is similar to that required for conventional surveys. The requirement would be for the entire area, typically about 91 m x 152 m (300 ft by 500 ft), and would include alignment (registration) errors. A point was made that for a 100 m piece of aluminum at 30 ºC (a change of temperature of 10 ºC), would yield a change in length of about 22 mm (0.87 in) which is much greater than the required ± 6 mm (1/4 in). However, because the tolerances in the fit-up and other processes are large enough, the level of accuracy of ± 6 mm (1/4 in) seems to be “working” and seems to be accepted. Given this “equality” of both techniques, traditional survey vs. laser scanning, it was suggested that traditional survey seems to result in about a 10 % rate of rework while laser scanning gives a more complete picture of the project due to the high density of points gathered. With regards to quantifying the project accuracy, a point was made that the accuracy can only be assessed at target locations because of the “truth” is unknown elsewhere.

The issue of registration error was discussed. It was agreed that this was a software issue and there were mixed feelings as to whether test protocols, in terms of standard data sets, should be developed to determine registration error. A separation of the hardware and software performance was desired by some. It was argued that vendors would benefit from having the level of accuracy of the work process, hardware and software, known. This would also allow a vendor to recommend software, in which he had confidence, to complement the hardware and if a problem arises, the problem could then likely be attributed to the hardware (if the software was already validated). The group agreed that “registration accreditation” was important because scan registration is necessary for most projects. The issue of other software testing was suggested but no consensus was reached.

The laser spot size was another characteristic that was felt to be important.

3.2.2 Group 2

The type of measurements required can be classified by the type of scanner: long range (> 100 m), medium range (6 m to 100 m), and short range (0 m to 6 m). This separation into three categories was thought to be necessary because the intended end-products from these classes of scanners are vastly different. If the intent is to scan a small object to obtain micron level accuracy, a long range scanner would not be used and would likely not perform will under those conditions.

The use of standardized targets (with known reflectivity ranging from 10 % to 90 %) and the geometric shapes that would encompass most of the objects encountered in the field were discussed. Cylinders of various diameters could be used to represent pipes which are often found in the field. Other geometric shapes include spheres, cubes, and pyramids and 2D targets with cut-out holes. Surface roughness could also be a factor – coarse, medium, and fine. Color and reflectivity of the object surface was also considered an important factor. The objects would be scanned “straight on” or at an angle. Some of these objects could potentially be heavy and/or large and therefore be difficult to move around. A suggestion was made to have these objects set-up and to scan the objects from pre-defined locations by moving the scanner to these locations.

There was also a brief discussion on the need for in-field calibration or a quick check that the instrument is within tolerance.

With regards to the effects of environmental conditions (temperature, humidity) on measurements – it was felt that it would be easier to begin with optimal conditions and then vary the conditions. Some felt that environmental conditions did not affect the measurements significantly for ranges of less than 50 m and some felt that environmental conditions were a big factor for ranges up to 25 m in a power plant. The possible effects of scanning hot surfaces such as desert sand, heated pipes, and road surfaces were also discussed.

The group felt that knowledge of the following basic characteristics were important: range uncertainty (ranked highest) – systematic and random errors, repeatability, reproducibility, point uncertainty, and instrument resolution as a function of distance. There was some discussion as to what was meant by repeatability and reproducibility.

EDITOR’S NOTE*: A clarification of the terms “repeatability” and “reproducibility” is offered.*

Repeatability *is defined by the VIM as the "closeness of the agreement between the results of successive measurements of the same measurand carried out under the same conditions of measurement. Repeatability conditions include:*

- *the same measurement procedure*
- *the same observer*
- *the same measuring instrument, used under the same conditions*
- *the same location*
- *repetition over a short period of time"*

Reproducibility *is defined by the VIM as the "closeness of the agreement between the results of measurements of the same measurand carried out under changed conditions of measurement. The changed conditions may include:*

- *principle of measurement*
- *method of measurement*
- *observer*
- *measuring instrument*
- *reference standard*
- *location*
- *conditions of use*
- *time."*

If the same measurement procedure and the same instrument model were used, the changed conditions that could potentially have a significant impact on LADAR measurements are observer/operator and conditions of use. It is also expected that the standard deviation of the measurements under repeatability conditions would be greater than the standard deviation of the measurements under reproducibility conditions.

Some instruments allow for multiple-shots or the averaging of several measurements. Therefore, in the determination of the range uncertainty, the uncertainty of a single shot should be determined in addition to the uncertainty for multiple-shots and the time for the measurement should also be reported. The group also discussed the ranges at which the uncertainties should be reported – optimum (dynamic range) or maximum.

Other types of measurements that were of interest to the group were geometric dimensions, object identification, object shape, object orientation, and relative distances – point-to-point measurements.

3.2.3 Group 3

The discussion began with the use of scaled artifacts – small artifacts that could be used at shorter ranges scaled up so that they could be “seen” by the longer range instruments. The artifacts (basic geometric shapes such as spheres or tetrahedrons) would be used to see if one could determine the volume, surface area, or extract edges and perimeters. There could then be standard artifacts to which secondary artifacts would be traceable.

The issue of performance metrics for the instrument vs. the system which includes the software for post-processing arose. It was felt that the software and the instrument could not really be de-coupled. Some instruments have post-processing software included with the instrument and there are other stand-alone commercial software packages. It was suggested that the results from the LADAR data using the different software packages could be given for purposes of comparison. This then led to the issue of file format for which there is currently no standard format. The generic output such as XYZ and intensity was insufficient because there may be other data dealing with the quality of the data that are available but reported differently amongst the manufacturers. It was felt that the issue of standard file format was important and should be brought up with the rest of the workshop participants.

As an initial step in evaluating scanners, a simple test suggested was to just scan a flat piece of steel or have a “standard” wall. The flatness of the wall and the deviations from that wall could be determined and used for comparison purposes. Consensus will have to be reached on wall texture, wall color, angle of incidence, distance of scanner to wall (several would be necessary), and height of scanner relative to the centroid of the wall. A three sided wall with different colors on each side mounted on a turntable or a single wall with different elements integrated in it were a couple of potential artifacts suggested.

On the topic of pointing/angular uncertainty, a definition was sought as to what it meant. It was felt that it was basically encoder accuracy. If that were the case, it was suggested that it is a calibration issue that was the responsibility of the manufacturer and his responsibility to sell a calibrated system. It was pointed out that any system would drift over time and had to be re-calibrated. If angular uncertainty meant the ability to resolve angles with the system, then artifacts could be constructed to test this aspect. A question posed was the ability to separate range uncertainty from angular uncertainty and resolution. A bar with tooling balls could be used – measure it head on and it would be purely range uncertainty and if you rotate the bar 90º, it would mostly be an angular motion.

When the group was asked to rank the list of characteristics – one vote was for range and angular uncertainty as number 1 and 2, respectively. Another vote was neither as a customer is mainly interested in the 3D uncertainty. It was argued that trackers and LADARs were spherical measurement devices. Thus, reporting range uncertainty and angular uncertainty in the specifications is all that would be necessary. Without these uncertainties, a table would have to be given to specify the uncertainties of the Cartesian coordinates as a function of range. For another participant, the angular/pointing uncertainty was more important than the range uncertainty.

There was some discussion about the coupling of range and angular uncertainties and beam spot size. The range uncertainty would be larger at edges due to the mixed pixel effect[3]. The issue again arose that if these characteristics are interdependent, is there a need to characterize them individually or is it only necessary to quantify the ability of the instrument to, for example, detect edges, to create a surface or to derive a volume. It was pointed out that feature resolution would allow for the evaluation of the combination of these three characteristics and it would be an important characteristic to quantify. There are many aspects of feature resolution, but as a starting point, edge detection would be useful.

The clarification of the term correlation was sought. Correlation, or more specifically, auto-correlation determines if measurements close in time are related.

EDITOR'S NOTE: *The following description of autocorrelation is offered and is taken from the NIST Statistic Handbook (http://www.itl.nist.gov/div898/handbook/eda/section3/eda35c.htm). "The autocorrelation function is defined as:*

$$r_k = \frac{\sum_{i=1}^{N-k} \left(Y_i - Y\right)\left(Y_{i+k} - \bar{Y}\right)}{\sum_{i=1}^{N} \left(Y_i - \bar{Y}\right)^2}$$

The autocorrelation function can be used to answer the following questions

1. *Was this sample data set generated from a random process?*
2. *Would a non-linear or time series model be a more appropriate model for these data than a simple constant plus error model?*

Randomness is one of the key assumptions in determining if a univariate statistical process is in control. If the assumptions of constant location and scale, randomness, and fixed distribution are reasonable, then the univariate process can be modeled as:

$$Y_i = A_0 + E_i$$

where E_i *is an error term.*

If the randomness assumption is not valid, then a different model needs to be used. This will typically be either a time series model or a non-linear model (with time as the independent variable)."

An example of how correlation could occur was given. Some scanners re-focus the laser beam depending on the range to obtain a better energy return. As the scanner beam crosses an edge,

---

[3] A mixed pixel or phantom point occurs when the laser beam is split at an edge of an object, i.e., part of the beam hits the object with the remaining portion hitting another object behind the first object. The reported range for that point is an interpolated value that lies between the two objects.

the scanner has to re-focus the beam. The refocusing is done mechanically, especially in older systems, and it is not an instantaneous action and there is lag time. Generally, scan speed is often set as high as possible, and there may be insufficient time between measurements for the instrument to re-set itself if the lag time is large enough. A similar phenomenon may occur in the electronics when scanning across a surface where there is a bright area next to a dark area or vice versa.

When measuring the scanner speed, there are many issues to consider – does the instrument only measure on the up and down swing or up swing only, and what are the start and start times. A system may look very good on paper for a single measurement but may be not so for a continuous scan.

Based on the discussion, correlation was added to the list of characteristics of interest. However, the terms "effective data rate" or "dynamic response time" were suggested in place of correlation. An opinion was voiced that "correlation", "effective data rate", or "dynamic response time" would not convey the same physical insight that range or angular uncertainty would. There were mixed feelings that "effective data rate" and correlation were similar in definition. It was generally accepted that the faster a scan, the higher the uncertainties of the measurements. However, if autocorrelation exists, autocorrelation tends to reduce the noise.

In summary, it was decided that range and angular uncertainty tied at rank 1 and feature resolution would rank as number 3. Correlation and effective data rate should also be characterized.

3.3 MARCH 16, 2005 – BREAK-OUT SESSION 3

For the third break-out session, the participants were asked to suggest test protocols for determining:

- Range uncertainty:
- Angular uncertainty
- Feature Resolution
- Performance evaluation – point to point accuracy

Although it was recognized that the instruments generally fell into three categories: short range (2 m to 25 m), medium range (25 m to 100 m), and long range (> 100 m) and the details of the protocols would be different, it was decided by the participants that at this stage general protocols were sufficient.

3.3.1 Group 1

The technique used by surveyors to validate their instruments was suggested as a possible technique. This technique involved benchmarks set-up throughout a site where the distances and errors between the benchmarks are known. The technique, however, allows a user to verify if an instrument is functioning properly, and it is not a calibration of the instrument. The question was then posed as to whether one would want a calibration facility or a verification facility – facility to test how close an instrument is to a manufacturer's specifications. The group felt that a verification facility was the desired facility; it was felt that instrument calibration would be instrument specific and would therefore be the manufacturer's responsibility as long as there was a way to verify the results.

The types of targets used in the verification test were discussed. Would the target be supplied by the manufacturer or would some generic target be used? Depending on the laser used, some targets are better than others for that instrument. Having the manufacturer supply the target and having generic (real world) targets was felt to be an ideal situation.

The question then was what was considered a target? Suggestions included spheres, flat plane, and well ground steel plate (or some material more commonly found in the field) with a matrix of coarsely spaced holes drilled into it. There were mixed feelings about the use of spheres – the advantage of spheres was that regardless of the location of the scanner to the sphere, there should always be near-normal angles of incidence on the spheres. A comment was also made that there were few applications for locating machined holes. For want of a standard target, regulation soccer balls, inflated to regulation pressure, mounted on a tripod have been used to measure distances from 500 m to 800 m. It was suggested that the dimensions of the targets would have to be specified – something on the order of a square foot.

Data sampling was then discussed. For a planar target, the two options were to scan the target and using standard algorithms, extract the target center or to take a single shot aimed at the center of the target. The latter option posed some obstacles as some instruments are not set-up to take single shots, i.e., have the ability to point the instrument and take single measurements of that point without moving off that point, and finding the center of the target. It was suggested that the use of spheres [e.g., 102 mm (4 in) diameter, matte surface] would eliminate the need to physically locate the center of the target. It was pointed out by one of the participants that for determining range error and range uncertainty, flat devices were always used. Again, the availability of both types of targets, spheres and flat plates, would be ideal. The ability to evaluate an instrument using targets of various shapes and materials in a 50 m, temperature controlled facility is highly desirable.

There was also discussion on the need to test various angles of incidence – have a planar target on a rotation stage, aim the scanner to the center of the target, and rotate the target through a series of known angles. The targets would generally be mounted on a tripod. This, in turn, led to the discussion of how does one find instrument center or zero. There was some discussion on how this could be accomplished but no resolution was reached. However, if relative distance was desired, finding instrument zero would not be an issue.

The distance intervals to set the targets were discussed – were 10 points sufficient/insufficient over 50 m? Should the distance intervals be linearly spaced? The possibility exists that even if many intervals were specified, problem spots may still be missed. It was felt that the spacing depended on the physics of the device – time-of-flight or phase-based (mode hopping). An understanding of the device and potential failure modes would be necessary before specifying the target spacing. This, however, could lead to non-standard requirements and open the way for claims of unfairness. A suggestion was made to let the manufacturer set the spacing. An opinion was voiced that a manufacturer would likely choose intervals to emphasize the instrument's good characteristics.

Angular uncertainty was the next topic of discussion. There were two definitions of angular uncertainty. First, there is the angular uncertainty of two sets of points – find the centroids of the point sets, get the distances, and calculate the angle between the two centroids. Second, there is the angular uncertainty of an individual measurement as there are situations where single point measurements are used rather than a collection of points as is the case when creating a model.

Protocols to determine the first definition of angular uncertainty are fairly straightforward. The protocols would involve setting up targets throughout the measurement volume and measuring them from different positions and backing out the range, azimuth, and elevation. Protocols to determine the second of angular uncertainty was not quite as clear and no resolution was reached. A point was made that there may be cross-coupling between the range uncertainty and the angular uncertainty.

Feature resolution was the third topic of discussion. The target (Figure 1) used by Wolfgang Böhler in his experiments and shown in his presentation generated interest in the group. The artifact provided a rapid means of qualitatively comparing instruments by visual inspection of the point cloud. The experiments were conducted at ranges of 6 m and 22 m – experiments at longer and different ranges were proposed.

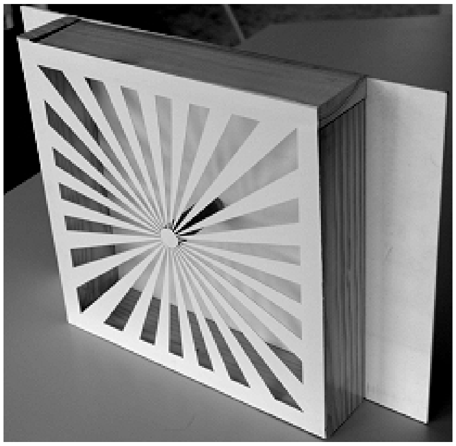


Figure 1. Wolfgang Böhler's target used for resolution experiments.
(Courtesy of W. Böhler)

Because of the large number of factors involved, good experimental design was deemed to be an important consideration. First, a well defined goal has to be set – what is the question to be answered? The second task is to determine the number of factors, K, of interest. The third task is to decide the number of runs, N, that is practical in terms of time and cost. The factors included environment (temperature, humidity, and pressure), spot size, angular increment/step size, angle of incidence, target distance, target material, target surface, target size, target geometry, and software processing. With regards to range, possible goals were:

1) comparison of instruments based on range uncertainty
2) determination of factors affecting range uncertainty
3) optimization of factors for range uncertainty
4) determination of range uncertainty of LADARs – results in a "single" value that states the uncertainty of the instrument under given conditions
5) modeling of range uncertainty of LADARs – develop mathematical function relating the factors to range uncertainty
6) determination of robustness of range measurements

A comment was made that an optimization experiment may be useful. A rebuttal to this was that vendors already optimize their own systems.

EDITOR'S NOTE: *The intent of the protocols is closest to goal No. 4 with the ultimate goal being goal No. 5. Other instrument characteristics will require similar protocol development.*

The funding for these efforts and the determination of what was practical and achievable were felt to be very important issues. The Construction Metrology and Automation Group of NIST has a lab (L x W x H: 13.3 m x 4.9 m x 4 m) which will be used to develop and refine artifact-based protocols. Some initial funding was obtained after the first workshop to start the facility – it is not the size of the "Houston Astrodome" but it is a start. Funding was also made available to obtain a LADAR with a range uncertainty of 100 µm. The Large-Scale Coordinate Metrology Group of NIST has a tape tunnel where 1D experiments up to 50 m are possible. Reference measurements could be made using an interferometer or a laser tracker. A comment was made that unlike total stations (1D devices), LADARs were scanners meant to cover large areas.

For object resolution, what is being done by some is to accumulate data for an object (e.g., pipes) by conducting thousands of scans of the object where the parameters include size (diameter), color, and material. Numerical analysis, based on these data, can then be used to predict the uncertainty of a measurement. Spot size and point density are also seen as factors affecting object resolution. Again, validity conditions have to be specified as long range scanners are not meant to function as short to medium range scanners, e.g., a long range scanner would not be used to measure pipes at 500 m. It was important to have protocols for all ranges of scanners – short, medium, and long.

3.3.2 Group 2

The group felt that the protocols would fall into two categories – test protocols to verify the manufacturer's specifications and protocols for applications – object identification, edges.

The determination of the range uncertainty would involve the use of planar targets with varying reflectivities and varying angles of incidence. The target had to be large enough [1 m x 1m or (3 ft by 3 ft)] to be able to be scanned at 200 m to 300 m.

The issue of multiple measurements was raised and the need to report the amount of time needed to obtain the measurement – as it was believed that increased accuracy is achieved at the expense of time (speed of measurement). There were basically two issues involving the measurement rate. The first being instruments which permit dwelling on/staring at a point to improve the accuracy. The second issue involved varying measurement/sampling rate of the different instruments: Is the measurement rate specified or is it left to the manufacturer to decide? Integral to the second issue was the point density which is directly proportional to the user specified angular increment/step size.

With regards to the first issue, a suggestion was made to define the sampling strategy – scanning the target and fitting a plane through the points or getting single point measurements. Fitting a plane through a set of points and averaging of several measurements would reduce the noise; however, these practices may not be representative of how the instrument is commonly used in the field. Nonetheless, knowledge of both uncertainties, single point and multi-shot, was felt to be important.

With regards to the second issue, uncertainties associated with fastest measurement rate and the slowest measurement rate would be reported along with the scan times. This will allow an end user to extrapolate the uncertainties and scan times for the particular application.

Several scan intervals were discussed:

1. 0 m to 50 m, 10 m intervals; 50 m to 100 m, 25 m intervals; > 100 m, 50 m or 100 m intervals.
2. 0 m to 50 m, 0.5 m intervals; 50 m to 100 m, 1 m intervals
3. 0 m to 20 m, 1 m intervals; 20 m to 50 m, 5 m intervals; 50 m to 100 m, 10 m intervals

It was felt that some information may be missed or lost by using the larger range intervals and hence the tight range intervals as suggested in the second option. The third option was suggested as many things “happen between 1 m and 20 m.”

Environmental conditions – room lighting, temperature – should also be specified. For evaluations outdoors, the environmental conditions will have to be recorded.

The resolution of the readout/display and reported measurement was brought up as a potential issue. The internal word length is commonly not reported along with the other specifications. The question was how many digits were significant.

Another issue that the group raised was the set-up of the instrument. How do you know where the center of the instrument is so that you can place it precisely over a known benchmark? It was felt that the manufacturer would have to specify this, i.e., instrument set-up

Results that the group felt should be reported were: truth (“true reference value”), the average measured range, the systematic error (average measured range – truth), and the noise (standard deviation of the measured range), and the measurement/scan time.

The protocol for angular uncertainty was the second topic of discussion. The group members were unsure of what was meant by angular uncertainty. Was it the uncertainty associated with calculating the angle between two points/objects or the uncertainty associated with pointing the instrument, i.e., if instructed to point at 90° azimuth and 0° elevation, is the instrument pointing there or at 89.5° azimuth and 0.1° elevation? (*EDITOR'S NOTE: The former uncertainty will be referred to as the angular uncertainty and the latter as pointing uncertainty in this section*).

Several protocols were suggested to evaluate the angular uncertainty. One involved the use of two targets separated by a known distance. The angle between the two targets can be calculated and this value is compared to the angle reported by the instrument. A second method suggested was the use of a link/length artifact of known length. A third was the placement of targets at known locations through the measurement volume. The use of a rail system to move the targets in 3 orthogonal directions – vertically (up/down), horizontally (towards and away from the scanner), horizontally (sideways) – was also suggested.

The group felt that the entire horizontal field-of-view (FOV) should be evaluated. This could be done by scanning the targets in one position, and then repeating the measurements by rotating the scanner head through several specified values to encompass the full horizontal FOV. This would identify potential encoder problems in any particular region.

Suggestions for planar targets included a black and white checkerboard, black targets, white targets, silver, and other colored targets. There was also discussion of the types of dual-purpose targets, i.e., extraction of dimensional data and extraction of object shape. The slotted artifact (Figure 1) that was presented by Wolfgang Böhler was a potential artifact. Target texture was also important – textures normally found in the real world such as concrete cinder block or brick.

A suggestion was made to also use targets that were expected to "cause problems" with the measurements. The group decided that this parameter should be evaluated in a separate protocol.

A summary of the group's discussion follows:

- Range uncertainty (2 m to 1 km)
  - What is reported
    - Ground truth range
    - Measured range
    - Systematic error (Measured range – truth)
    - Noise (Standard deviation around the average)
    - Scan time
    - Sample size, *n*
  - Scan planar surface of known reflectivity and of sufficient size to contain the beam spot, and under known lighting conditions
  - Measurement mode
    - Scan the target – no staring
    - Single point
  - Scan rates

        - Laser at maximum pulse rate
        - Scan at manufacturer recommended rate
    - Parameters
        - Angle of incidence (45º, normal, -45º)
        - Different target reflectivities
    - Range intervals
        - 1 m for ranges between minimum range and 20 m
        - 5 m for ranges between 20 m and 50 m
        - 10 m for ranges between 50 m and 150 m
- Angular uncertainty
    - Absolute and relative angular uncertainty
    - Rail system with multiple pairs of targets at fixed horizontal distances and spaced at known intervals vertically (10 m high) – V-shaped.

- Feature Resolution
    - Target
        - Reflectivity
        - Color
        - Texture
    - Lighting conditions
    - Problematic targets, e.g., water – does one get any return from water
    - Set of primitive artifacts
    - Set of complex artifacts

3.3.3 Group 3

A manufacturer in the group offered to share the general procedure that was used to determine the accuracy of that company's instrument. The values are provided in tabular format for different ranges. The targets used are Kodak grade tape with different reflectivities. The planar target is measured with a single shot, four shots, nine shots, 16 shots and maybe 25 shots. The time for measurements obviously increases as the number of shots increases. The target is placed on a motorized bench that extends 50 m and the LADAR measurements are compared to the values obtained from the motorized bench. A total station is used to provide the reference measurements for ranges greater than 50 m. The instrument is set-up over a point and centered over that point in a manner similar to that for a total station. It is assumed that the center of the instrument would then reside over this point. The procedure to find the zero location of the scanner is proprietary information (as with most manufacturers). Upon purchase of an instrument, a user is provided with a procedure to verify that the scanner is functioning properly.

Following this description, the discussion touched on several topics, among them being the issue of the different software used to process the data. It was felt that the resulting measurement from the hardware which has been processed by software, whether it is software that came along with the hardware or some other commercial package, should be what is reported.

A method to determine the angular uncertainty was to set up targets around the scanner. Scan the targets, rotate the scanner and re-scan the targets, and the procedure is repeated. A comparison of how the targets align for the different rotated positions will give an indication of the angular uncertainty.

The group agreed that for determining range uncertainty up to 100 m, a rail system would be appropriate with an interferometer to provide the reference measurements. A suggestion was made to have the manufacturer supply the targets as they would know what worked best for their systems. This was, however, felt to counter the goal of standard test protocols. Possible targets were flat plates and spheres where the target sizes will have to be specified in the protocols. Other targets such as corner cubes and tooling balls could be sub-parts of the protocols. It was suggested that developing protocols for the case of a controlled environment would be the simplest case and should be the starting point.

The manner in which the measurements are obtained is also important – single point or scanning where several points are averaged. Some instruments have the capability of pointing the scanner at a specified location and acquiring a range to that point while others do not have that capability. Therefore, separate protocols for single point, stare mode measurement and scan surfaces measurements were necessary.

One concept put forward for discussion for a scan measurement protocol was to position a planar target at different ranges along a rail. Scan the targets at the different ranges and fit planes through each set of points. Select a point on the last plane, and fit a line through that point and the scanner origin. Distances from that line to the intersections of the various planes will then be used as the "measured" distances and the angles between the line and the planes can be compared to the incidence angles.

For the scanning mode evaluation, a planar target with a rear-mounted corner cube would be used. The reference measurement would be made with an interferometer located at the opposite end of the rail to the scanner.

With regards to the interferometer, a couple of points were made. One point was that for distances greater than 50 m to 60 m, the interferometer beam starts to grow in diameter. In real world conditions, the beam is "going to walk around" and there will be a lot of noise. The Australians have an outdoor facility where a big corner cube was used that worked over 700 m – a suggestion was made to find out how they did that. The other point was that for a 100 m facility, very good information about the temperature and pressure is required throughout the length of the facility.

To accommodate the longer range scanners, the group discussed the possibility of establishing an indoor or outdoor kilometer range facility. An indoor facility will be very expensive and an outdoor facility would require setting up weather stations. A question was posed on the use of compact range with mirrors. It was felt that this may work for coherent laser radars but may not work for instruments using other technology. Existing potential sites (e.g., David Taylor model basin, the abandoned Super Collider site in Texas) were discussed and dropped from consideration for various reasons. It was concluded that an indoor, kilometer long facility was too cost prohibitive to construct.

The outdoor kilometer facility was then discussed as being more feasible. Permanent monuments would be set up throughout the facility and there would be a “built-in” ability/method (e.g., differential GPS) to perform a calibration of the monument locations. The monuments would have kinematic mounts to hold the targets in position. Agreement on the size, color, and shape of the target will have to be reached – possible targets include planar targets and very large spherical mounted retroreflectors (SMRs). To correct for the environmental conditions, one method would be to set-up weather stations to monitor the temperature, humidity, and pressure to determine the index of refraction. Another method to determine the index of refraction is to constantly monitor the range measurements using an ADM (absolute distance meter) between monuments. The refractive index can then be backed out since the changes in the ADM measurements are known and the “true” values are known. The refractive index obtained in this manner is integrated over the range of interest. It was suggested that since none of the commercially available instruments has this capability, the measurements acquired by the instruments being evaluated be corrected by the refractive index obtained using the ADM method. This would be a way of normalizing the measurements in terms of the environment.

A note was made that the determination of the range uncertainty, as discussed in the previous paragraphs, would be of secondary interest to some users. The primary interest of these users would be the accuracy of the resulting 3D models.

The first requirement discussed was the size of the facility required to determine the angular uncertainty – is a facility 100 m x 100 m facility required? Will a smaller facility suffice if the scanner was rotated? The response to the latter question was that there were disadvantages to this procedure. The definition of angular uncertainty was sought – is it angular pointing accuracy or is it the ability to resolve two targets? The latter definition was felt to be what was sought. The idea was that the same indoor facility for evaluating the range uncertainty would be used – 100 m facility with a rail system. A target similar to what Wolfgang Böhler used in his experiments (Figure 1) could be placed on the rail and scanned at various ranges. It was felt that although the target yielded interesting qualitative results, a target that would yield a quantitative result would be more useful.

To obtain a quantitative value, a simple concept suggested was to scan two plates, placed side by side, and the output from the test would be 1) is there one plate or two plates, 2) what is the gap between the plates, and the 3) distance from the scanner to the plates. The test set-up would consist of two plates on a translation stage to allow lateral separation of the plates and to allow for measurement of the gap between the plates. The combined assembly would be placed on a rail to allow movement towards and away from the scanner. A third plate would be located behind the two plates to act as a back plane. This test set-up could be scaled up for use in an outdoor facility and could also be used for range resolution with minor modifications. Modifications would have to be made so that the separation of the plates is parallel to the axis of the rail, i.e., rotate the translation stage 90°.

Some issues discussed in relation to this test include a possible bias since it is known *a priori* that two plates would be moving apart. A second issue was the use of software to aid in the gap detection. An example of this is the creation of a polygon model from the data points. By setting the triangle length equal to the grid size (point spacing), a gap would be detected when the model broke into two pieces. Another example is the use of edge detection algorithms, either internal to the scanner or external to the scanner, to detect the edges of the plates. A third issue dealt with data cleaning. How do you eliminate outliers – is the manufacturer allowed to clean the data? The group felt that whatever software is supplied by the manufacturer should be used – edge detection algorithms or data processing – as it was the system that was being evaluated. A fourth issue was the specification of the step size or the point spacing.

A suggestion was also made to use a pie shaped artifact to narrow down the range at which there may be a breakdown in the angular resolution. The advantage of this is that it would be a simple test and only one scan need be taken as opposed to many scans. It would also be useful to know how well one can define angular resolution using this test and if that would be sufficient.

In the course of the discussions, a NIST report on Performance Analysis of Next-Generation LADAR for Manufacturing, Construction, and Mobility [4] was mentioned. Interest in obtaining this report was expressed – the report may be downloaded from the following website: http://www.bfrl.nist.gov/861/CMAG/publications/index.htm.

4.0 SUMMARY

This chapter summarizes the various ideas/concepts that were discussed in the three break-out sessions at the workshop.

4.1 WORKSHOP SUMMARY

The second NIST workshop on LADAR performance evaluation centered on the following objectives:

1. Discussion and modification (addition or deletion) of the terms to the draft terminology.
2. Determination of instrument characteristics of interest and levels of accuracy required
3. Development of test protocols and artifacts required to evaluate instrument characteristics that are of most interest to the participants.

A statement was made in the workshop that NIST does not set standards. Standards are set by organizations that are authorized to do so and are consensus based – organizations such as ANSI, ASTM, ASME, etc. The objective of NIST's effort is to facilitate the development of standards for LADAR performance evaluation. As alluded to in Chapter 1, the use of and applications for LADARs are growing rapidly. The availability of performance standards for LADARs is expected to not only benefit the construction industry, where they are primarily being used, but other industries as well, e.g., manufacturing, defense, law enforcement.

With regards to terminology, the general feeling among the participants was that it did not matter which definitions (e.g., VIM, GUM, ASTM, survey definitions, etc.) were adopted so long as a set of definitions was adopted and accepted by the laser scanning community. There was some unfamiliarity with the VIM and GUM among the participants. A point was made that if standardization was the desired outcome of the current efforts then definitions that are adopted by organizations such as ISO and ANSI should be used. The participants felt that the replacement of the term "accuracy" with "uncertainty" was not likely to occur as the term accuracy was too entrenched in the marketplace and was better understood by the users, sales force, and managers. On a similar note, in the CMM community, such a changeover did occur but over a period of time.

It was also felt that in addition to formal definitions of terms, operational tests are required to specify how one determines "measurement uncertainty" or "resolution". The terms accuracy, precision, and resolution created some confusion as they have often been used interchangeably. The general consensus was that resolution was not encoder resolution but was instrument resolution in terms of depth (range) resolution, horizontal resolution, and vertical resolution. For example, for horizontal resolution – how far apart do two objects have to be laterally separated before one can tell that there are indeed two objects and not one. Some additional terms suggested for inclusion to the list were beam divergence, beam spot size, rated conditions, limiting conditions, registration, and compensation.

The general feeling among the participants was that the acronym LADAR was unfamiliar to most in the construction/manufacturing communities and was not commonly used outside of the military community. Other terms were suggested were laser scanning (may preclude triangulation-based or structured light devices) and 3D imaging. Until a term is agreed upon, NIST will continue to use the acronym LADAR to mean ground-based laser radars/laser scanners.

There were mixed feelings about the need to characterize an instrument. There were some who felt that it was necessary to determine the instrument characteristics such as range uncertainty while model or project accuracy was more important to others. It was not an "either/or" situation, and most felt that knowledge of both instrument characteristics and project accuracy was necessary.

Knowledge of instrument characteristics was felt to be important because when purchasing an instrument, users want a way to fairly compare instruments. Answers to basic questions with regards to manufacturer specifications such as:

- Does the uncertainty[4] hold for the entire range of the instrument?
- Under what conditions does it hold – only at 20 ºC?
- For what objects – smooth white wall, brick wall, highly specular object?
- Was this accuracy achieved at the highest scan speed or at the lowest scan speed?

are required to enable fair comparison between instruments and an educated decision as to the best instrument for a particular application. Additionally, the test protocols used by the different manufacturers to determine the uncertainties differ, and the resulting uncertainties would be different for different test protocols; thus, one-to-one comparisons of stated specifications for different instruments is not possible.

Knowledge of the project accuracy was important to users/service providers and facility owners. The basic questions to be answered are:

- How accurate is the 3D model?
- How do you verify the accuracy (in terms of the uncertainty of geometric dimensions, point-to-point distance, etc.) of the model?
- How do the 3D models derived from LADAR data compare to those obtained using other methods?

With regards to instrument characteristics, range uncertainty was ranked the highest. Other characteristics that were also felt to be important were reproducibility, repeatability, angular uncertainty, and instrument resolution. Other characteristics discussed were data rate and correlation.

---

[4] Most manufacturers currently use the term "accuracy" when the more correct term for this specification is "uncertainty". It should also be noted that the measurement conditions are often not provided.

In terms of developing test protocols for range uncertainty, the parameters considered to have the most affect on range uncertainty include environmental conditions (temperature and humidity), target reflectivity, target material, angle of incidence, and distance to target. It was generally agreed that a planar target with standard (known) reflectivity would be used. The size of the target will have to be agreed upon. Range uncertainties for ranges of 50 m to 60 m could be determined in an indoor facility; range uncertainties for ranges greater than this would likely have to be determined in an outdoor facility as an indoor facility would not be economically feasible. Issues that have to be considered when developing the protocols include:

- The inability of some scanners to acquire a single measurement, i.e., they can only operate in scanning mode. Protocols would then have to be developed for single shot uncertainty and for scanning mode where a planar target would be scanned and a plane fitted through the points. In scanning mode, the point density would have to be specified.
- The ability of some scanners to average measurements. This can be done by scanning a scene multiple times and averaging the values or by dwelling on a point. Averaging would increase the time to acquire a scan and decrease the noise. Therefore, it was suggested that the scan or measurement time needs to be reported.
- The determination of the distance intervals for placing the targets. Would the intervals be "constant" or would a more random spacing be required? With regards to the latter, it was felt that knowledge of the physics of the instrument was needed to ascertain potential fail points, e.g., occurrence of mode hopping. The danger of this would be the development of non-standard or instrument-specific protocols. It was generally felt that for ranges between 1 m and 20 m, the spacing of the targets need to be smaller.
- The inability to de-couple hardware and software. It was felt that the measurement as processed by the software, either provided by the manufacturer or other commercially available packages, should be used and the software that was used should be reported.
- The ability to center an instrument over a benchmark. It was decided that the manufacturer would have to provide instrument set-up procedures.
- How the reference measurements are obtained. Currently, LADAR measurements are compared to measurements from traditional survey instruments such as total stations. It is anticipated that the accuracy of LADAR devices will likely equal the accuracy of traditional survey instruments in the near future.

The definition of angular uncertainty generated some discussion. Was angular uncertainty "how accurately is the angle between two points determined?" or was it "how accurately can the instrument be pointed, i.e., if the angle between measurements was specified to be 0.05°, did the instrument go to 0.05° or did it go to 0.04°?" Several protocols were suggested for the former definition. However, it was felt that developing protocols for the latter definition would be much more difficult. Angular uncertainty, as defined by the latter definition, provides information that would allow for a direct means of quantifying the uncertainty of a point in Cartesian coordinates as these values are directly derived from the azimuth angle, elevation angle, and range – quantities that are measured by the instrument.

With regards to model or project accuracy, the use of scaled artifacts to allow for indoor and outdoor evaluations was suggested. These artifacts could include spheres, tetrahedrons, and cylinders. The material and shape of these artifacts will have to be decided and agreed upon. There was only limited discussion as to what constituted project accuracy. For example, was it point-to-point accuracy, was it the ability to identify an object or features on an object, or was it the ability to detect misaligned components. Currently, a control network of benchmarks is used to infer or to extrapolate project accuracy. The adequacy and/or improvement of this approach needed further investigation. The specification of the required point density and the number of scans are two other means that are used to help ensure that the required project accuracy is achievable.

Finally, a method was needed to allow a user to perform field verification of an instrument to assure that it is functioning properly. This was an issue that some felt fell in the prevue of the manufacturer while others felt that a simple artifact (e.g., ball bar, SMRs) could be used or a verification protocol could be specified.

In the general workshop discussions, the topic of standard data format and metadata was discussed. Most of the participants felt that a standard data format would allow for interoperability of the hardware and software.

At the conclusion of the workshop, the following next steps were agreed upon:

- Terminology
    - Glossary of terms – edit and send out to participants
    - Send a link to TN 1297 to participants (see Section 3.1.3)
    - Look at terms used for similar instruments – total stations, laser trackers, GPS – and those used in the ASPRS (American Society of Photogrammetry and Remote Sensing) effort
- Look at the process to develop GPS and CMM standards: what is the process for formal standardization of test protocols?
- Develop test protocols for range uncertainty, angular uncertainty and send out for comments
- Contact and request information from manufacturers on how they assess their range uncertainties or how they assess the performance of their system for reliability
- Investigate potential artifacts
    - Wolfgang Böhler's artifact with modifications
    - Pipes, spheres, etc.
- Publish test results and disseminate among workshop participants
- Organize volunteers with different instruments for evaluating
- Conduct another workshop in early 2006

4.2 RELATED EFFORT

Following the March 15-16, 2005 workshop, NIST was asked to participate in an effort to assess the accuracy of 3D models generated based on laser scanning of some GSA-owned buildings. A workshop sponsored by NIST and FIATECH (www.fiatech.org) was held at NIST on July 13, 2005 to determine industry interest and input. The overarching goals of NIST/FIATECH workshop and of NIST's efforts are the development of consensus based procedures for assessing the:

- accuracy of field measurement technologies and equipment,
- quality and completeness of the data and meta data,
- quality of work processes and procedures, and
- accuracy of the modeling software

The NIST/FIATECH workshop objectives were to determine:

- Current practices in quantifying resulting model accuracy
- Scope of field demos
- Project plan

The group developed an outline of a document that listed the tasks/steps necessary to enable the assessment of the accuracy of 3D models. The outline of the document is presented below:

- Guidelines for Readers
    - Manufacturers
    - Users
    - Owner/Operators
    - Procurement Contractors
- Field Measurement Technology
    - What do customers need?
    - Levels of accuracy required for different types of applications
- LADAR Technology Introduction
    - What is the current state of technology?
    - Why are we looking at it?
    - What is the potential cost savings based upon data quality, time, safety, etc.
- LADAR Quality Assurance
    - What are the sources of error during the data capture and analysis process? (Calibration, field survey control, scanning, registration, modeling)
    - Discussion of Error Budgeting
    - Discussion of Error Sources
        - Calibration
        - Field Surveying
        - Scanning
        - Registration
        - Modeling
    - Methods of Controlling Error (Optimum Work Processes)
        - Calibration
        - Field Surveying
        - Scanning
        - Registration
        - Modeling
- Verification

    - Calibration
    - Field Surveying
    - Scanning
    - Registration
    - Modeling
    - As-built
- Standards and Specifications
    - Concepts for End-Product Performance Specifications
    - LADAR data and meta-data standards
    - ISO 9000 Certification
    - Accreditation / Licensing
- Glossary

REFERENCES

1. **Cheok, G. S., Ed. [2003]**, "Proceedings of the LADAR Calibration Facility Workshop, June 12-13, 2003," *NISTIR 7054* (CD-ROM), National Institute of Standards and Technology, Gaithersburg, MD, October.

2. *International Vocabulary of Basic and General Terms in Metrology*, International Organization for Standardization, Genève, Switzerland, 1993.

3. **American National Standards Institute [1997]**, *U. S. Guide to the Expression of Uncertainty in Measurement*, ANSI/NCSL Z540-2-1997.

4. **Stone, W.C., Juberts, M., Dagalakis, N., Stone, J., Gorman, J. [2004]**, "Performance Analysis of Next-Generation LADAR for Manufacturing, Construction, and Mobility," *NISTIR 7117*, National Institute of Standards and Technology, Gaithersburg, MD, May 2004.

5. **W. Boehler, M. B. Vicent, A. Marbs [2003]**, " Investigating Laser Scanner Accuracy," presented at the XIXth CIPA Symposium at Antalya, Turkey, 30 Sept. – 4 Oct., 2003. http://www.i3mainz.fh-mainz.de/publicat/cipa2003/laserscanner_accuracy.pdf

APPENDIX A: LIST OF WORKSHOP PARTICIPANTS

Last Name	First Name	Company	Address	Phone Number	E-mail Address
Ackerson	Scott	Dimensional Photonics	103 Terrace Hall Avenue Burlington , MA 01803	(781) 272-4545 x225	scott.ackerson@dphotonics.com
Adrian	Carl	FBI	Room 1170, 2501 Investigation Parkway Quantico, VA	(703) 632-8206	carl.adrian@ic.fbi.gov
Battaglia	John	Bentley - CloudWorks	20201 Century Blvd., Suite 420 Germantown, MD 20874	(240) 780-4452	John.Battaglia@bentley.com
Böhler	Wolfgang	i3mainz i3mainz at the specialist area 1 geo computer science and measurement	Holzstr. 36 55116 Mainz, Germany	+49 - 6131 - 2859-671	boehler@geoinform.fh-mainz.de
Borchardt	Bruce	NIST - MEL/PED	100 Bureau Drive, Stop 8211 Gaithersburg, MD 20899-8211	(301) 975-3469	bruce.borchardt@nist.gov
Cheves	Allen	Cheves Media LLC	905 W 7th St #331 Frederick MD 21701	(301) 620-0784	allen.cheves@chevesmedia.com
Cheves	Marc	Cheves Media LLC	905 W 7th St #331 Frederick MD 21701	(301) 620-0784	marc.cheves@chevesmedia.com
Cheok	Gerry	NIST – BFRL / CMAG	100 Bureau Drive, Stop 8611 Gaithersburg, MD 20899-8611	(301) 975-6074	cheok@nist.gov
Cothren	Jack	Univ. of Arkansas	Department of Geosciences Center for Advanced Spatial Technology 12 Ozark Hall University of Arkansas Fayetteville, AR 72701	(479) 575-6790	jcothren@cast.uark.edu
Elkins	Les	Spatial Integrated Systems	7564 Standish Place, Suite 100 Rockville, MD	(301) 610-7965 x107	lelkins@sisinc.org
Estler	Tyler	NIST - MEL / PED	100 Bureau Drive, Stop 8211 Gaithersburg, MD 20899 - 8211	(301) 975-3483	tyler.estler@nist.gov

Last Name	First Name	Company	Address	Phone Number	E-mail Address
Filliben	Jim	NIST – ITL/ SED	100 Bureau Drive, Stop 8980 Gaithersburg, MD 20899-8980	(301) 975-2855	James.filliben@nist.gov
Flint	Jim	Laser Geomatics Bohanan Huston	7500 Jefferson NE Albuquerque, NM 87109	(505) 798-7864	jflint@bhinc.com
Gelhar	Brent	Optech	100 Wildcat Road Toronto, Ontario Canada M3J 2Z9	(416) 661-5904	brentg@optech.ca
Gilsinn	Dave	NIST - ITL/MCSD	100 Bureau Drive, Stop 8910 Gaithersburg, MD 20899 - 8910	(301) 975-3491	david.gilsinn@nist.gov
Gittinger	Juergen	iqvolution	Konigsallee 35 71638 Ludwigsburg, Germany	+49-7141-6889-581	Juergen.Gittinger@iqvolution.com
Greaves	Tom	Spar Point Research	85 Constitution Lane, Suite 2E Danvers, MA 01923-3630	(978) 774-1102	tom.greaves@sparllc.com
Hand	Steve	MagLev, Inc	1905 Technology Center 1100 Industry Road, Box 11 McKeesport, PA 15132	(412) 948-1905 x 6097	shand@maglevinc.com
Hoffman	Eric	Quantapoint	Arbor Professional Centre 275 Curry Hollow Road, M100 Pittsburgh, PA 15236	(412) 653-0100 x203	ehoffman@quantapoint.com
Hong	Tsai	NIST - ITL / ISD	100 Bureau Drive, Stop 8230 Gaithersburg, MD 20899-8230	(301) 975-3444	hongt@nist.gov
Jacobs	Geoff	Leica Geosystems HDS	4550 Norris Canyon Road San Ramon, CA 94583	(925) 790-2317	Geoff.Jacobs@hds.leica-geosystems.com
Klusza	Mark	Bitwyse	27 Congress Street, Suite 406 Salem, MA 09170	(978) 745-2300 x 24	mark.klusza@bitwyse.com
Knaak	Ted	Riegl USA	7035 Grand National Drive, Suite 100 Orlando, FL 32819	(407) 248-9927	tknaak@rieglusa.com
Langer	Dirk	Z & F	1 Library Place, #203 Pittsburgh, PA 15110	(412) 469-9210	langer@zf-usa.com

Last Name	First Name	Company	Address	Phone Number	E-mail Address
Leigh	Stefan	NIST – ITL / SED	100 Bureau Drive, Stop 8980 Gaithersburg, MD 20899-8980	(301) 975-2856	stefan.leigh@nist.gov
Lepere	Gregory	Trimble	7401 Church Ranch Boulevard Westminster, CO, 80021	(720) 887-4317	Gregory_Lepere@Trimble.com
Liadsky	Joe	Optech	100 Wildcat Road Toronto, Ontario Canada M3J 2Z9	(416) 661-5904	joe@optech.on.ca
Lytle	Alan	NIST - BFRL / CMAG	100 Bureau Drive, Stop 8611 Gaithersburg, MD 20899 - 8611	(301) 975-6048	alan.lytle@nist.gov
Ober	Dave	MetricVision	8500 Cinder Bed Rd. Suite 150 Newington, VA 22122	(703) 541-2642	david_ober@metricvision.com
Palmateer	John	Boeing	P.O Box 3707, Mail Stop 4E-71 Seattle, WA 98124-2207	(253) 931-4636	john.w.palmateer@boeing.com
Persi	Fred	Quantapoint	Arbor Professional Centre 275 Curry Hollow Road, M100 Pittsburgh, PA 15236	(412) 653-0100	fpersi@quantapoint.com
Phillips	Steven	NIST - MEL/PED	100 Bureau Drive, Stop 8211 Gaithersburg, MD 20899 - 8211	(301) 975-3565	steven.phillips@nist.gov
Saidi	Kamel	NIST – BFRL / CMAG	100 Bureau Drive, Stop 8611 Gaithersburg, MD 20899 - 8611	(301) 975-6069	kamel.saidi@nist.gov
Sawyer	Dan	NIST - MEL / PED	100 Bureau Drive, Stop 8211 Gaithersburg, MD 20899-8211	(301) 975-5863	daniel.sawyer@nist.gov
Shackleford	William	NIST – MEL / ISD	100 Bureau Drive, Stop 8230 Gaithersburg, MD 20899-8230	(301) 975-4286	william.shackleford@nist.gov
Sedransk	Nell	Formerly NIST – ITL / SED			
Soubra	Omar-Pierre	Trimble / MENSI	7401 Church Ranch Boulevard Westminster, CO, 80021	(720) 887-4317	Omar_Soubra@Trimble.com
Stone	Bill	NIST - BFRL/CMAG	100 Bureau Drive, Stop 8611 Gaithersburg, MD 20899 - 8611	(301) 975-6075	william.stone@nist.gov

Last Name	First Name	Company	Address	Phone Number	E-mail Address
Vandapel	Nicolas	Carnegie Mellon University Robotics Institute	NSH 5000 Forbes Avenue Pittsburgh, PA 15213	(412) 268-2991	vandapel@muck.ius.cs.cmu.edu
Witzgall	Christoph	Christoph Witzgall	20 Walker Ave Gaithersburg, MD	(301) 330-0129	christoph.witzgall@nist.gov

APPENDIX B: DRAFT TERMINOLOGY

B.1 DRAFT TERMINOLOGY PRESENTED AT THE WORKSHOP

An initial collection of draft terminology was distributed to the participants in the first NIST LADAR workshop in the summer of 2004. Comments received are incorporated into the revised draft terminology given below.

Accuracy of measurement [VIM 3.5]:

Closeness of the agreement between the result of a measurement and a true value of the measurand

NOTES:

1. “Accuracy” is a qualitative concept.
2. The term “precision” should not be used for “accuracy”.

Calibration [VIM 6.11]:

a set of operations that establish, under specified conditions, the relationship between values of quantities indicated by a measuring instrument or measuring system, or values represented by a material measure or a reference material, and the corresponding values realized by standards.

Notes:

1. The result of a calibration permits either the assignment of values of measurands to the indications or the determination of corrections with respect to indications.
2. A calibration may also determine other metrological properties such as the effect of influence quantities
3. The result of a calibration may be recorded in a document, sometimes called a calibration certificate or a calibration report.

Conventional true value (of a quantity) [VIM 1.20]:

Value attributed to a particular quantity and accepted, sometimes by convention, as having an uncertainty appropriate for a given purpose.

EXAMPLES

a) At a given location, the value assigned to the quantity realized by a reference standard may be taken as a conventional true value;
b) The CODATA (1986) recommended value for the Avogadro constant,
N_A: $6.022\,136\,7 \times 10^{23}$ mol^{-1}

NOTES:

1. “Conventional true value” is sometimes called assigned value, best estimate of the value, conventional value or reference value. “Reference value”, in this sense, should not be confused with “reference value”.
2. Frequently, a number of results of measurements of a quantity are used to establish a conventional true value.

Deviation [VIM 3.11]:

Value minus its reference value

Error (of measurement) [VIM 3.10]:

Result of a measurement minus a true value of the measurand.

NOTES:

1. Since a true value cannot be determined, in practice a conventional true value is used.
2. When it is necessary to distinguish “error” from “relative error”, the former is sometimes called “absolute error of measurement”. This should not be confused with the “absolute value of error”, which is the modulus of error.

Measurand [VIM 2.6]:

Particular quantity subject to measurement

EXAMPLE	vapor pressure of a given sample of water at 20° C.
NOTES	The specification of a measurand may require statements about quantities such as time, temperature and pressure.

Repeatability (of results of measurements) [VIM 3.6]:

Closeness of the agreement between the results of successive measurements of the same measurand carried out under the same conditions of measurement.

NOTES:

1. These conditions are called repeatability conditions.
2. Repeatability conditions include:
    - the same measurement procedure
    - the same observer
    - the same measuring instrument, used under the same conditions
    - the same location
    - repetition over a short period of time

3. Repeatability may be expressed quantitatively in terms of the dispersion characteristics of the results.

Reproducibility (of results of measurements) [VIM 3.7]:

Closeness of the agreement between the results of measurements of the same measurand carried out under changed conditions of measurements.

NOTES:

1. a value statement of reproducibility requires specification of the conditions changed.
2. The changed conditions may include:
    - principle of measurement
    - method of measurement
    - observer
    - measuring instrument
    - reference standard
    - location
    - conditions of use
    - time.
3. Reproducibility may be expressed quantitatively in terms of the dispersion characteristics of the results.
4. Results are here usually understood to be corrected results.

Standard uncertainty (GUM 2.3.1):

Uncertainty of the result of a measurement expressed as a standard deviation.

Uncertainty of measurement [VIM 3.9]:

Parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand.

NOTES:

1. The parameter may be, for example, a standard deviation (or a given multiple of it), or the half-width of an interval having a stated level of confidence.
2. Uncertainty of measurement comprises, in general, many components. Some of these components may be evaluated from the statistical distribution of the results of series of measurements and can be characterized by experimental standard deviations. The other components, which can also be characterized by standard deviations, are evaluated from assumed probability distributions based on experience or other information.
3. It is understood that the result of the measurement is the best estimate of the value of the measurand, and that all components of uncertainty, including those arising from systematic effects, such as components associated with corrections and reference standards, contribute to the dispersion.

Additional concepts [GUM 2.2.4]:

- A measure of the possible error in the estimated value of the measurand as provided by the result of a measurement;
- An estimate characterizing the range of values within which the true value of a measurand lies.

True value (of a quantity) [VIM 1.19]:

Value consistent with the definition of a given particular quantity

NOTES:

1. This is a value that would be obtained by a perfect measurement.
2. True values are by nature indeterminate.
3. The indefinite article “a”, rather than the definite article “the”, is used in conjunction with “true value” because there may be many values consistent with the definition of a given particular quantity.

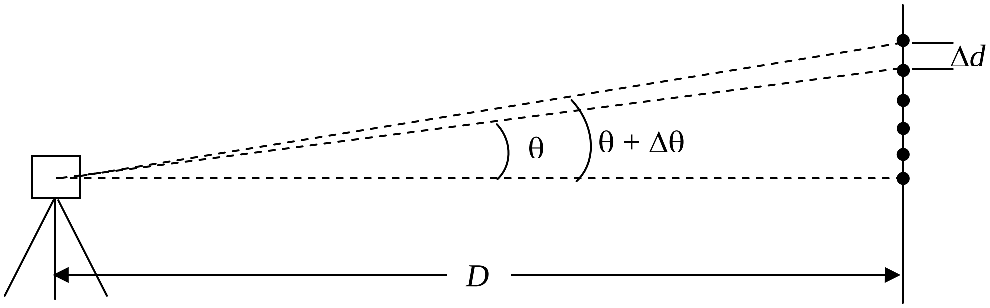
LADAR Terminology and Acronyms

APD

An acronym for "Avalanche Photo Diode".

Angular Increment or Step Size

The change in an angular measurement, $\Delta\theta$, where $\Delta\theta = \theta_i - \theta_{i-1}$. This value, specified by the laser scanner manufacturer, is typically a minimum value and gives an indication of the achievable point density. A smaller angular increment results in a denser point cloud. The angular increment can be used to determine the distance between pixels/points, e.g.,



Distance between points = $\Delta d \cong D\,(1 + \tan^2 \theta)\,\Delta\theta$

CCD

An acronym for "Charge Coupled Device". An imaging sensor where individual pixels in an array are allowed to transport, store, and accumulate optically-generated charge carriers to defined sites within the device. CCD principles, combined with on-chip timers for each pixel can be used to create a time-of-flight focal plane array.

Flash LADAR

A LADAR system comprised of a broad field illumination source (commonly a laser, but for close proximity it can be a bank of LEDs) and an FPA detector, such that the range image is acquired simultaneously in one burst. This allows for the achievement of high frame rates which is critical for real-time applications.

FOV

An acronym for "Field of View". The angular coverage of a scene and the units normally associated with FOV are degrees, e.g., The LADAR has a FOV of 300° (horizontal) x 80° (vertical).

FPA

An acronym for "Focal Plane Array". A 2D "chip" in which individually addressable photo sensitive "pixels" can be accessed.

Frame

A frame is equivalent to a region of interest where data is to be acquired. The size of the frame is generally user specified with the maximum size of a frame equivalent to the FOV of the instrument. In scanning instruments, the point density in a frame is determined by the angular increment (generally user specified) in both the horizontal and vertical directions.

Frame Rate

The number of frames that can be acquired per second. For example, if 10 frames could be acquired in one second, the frame rate would be 10 Hz. This is generally a metric that is mainly applicable to real-time systems such as flash LADARs, since most commercial scanning LADARs have update rates on the order of minutes and are dependent on the laser pulse repetition rate, selected FOV, and selected angular increment.

In the case of scanning systems, an appropriate description of the frame rate should include the time, FOV, and the angular increment. For example, the time to acquire a frame for an FOV of 360° x 90°, and an angular increment of 0.3° (horizontal and vertical) is 160 s.

Instrument Resolution

Horizontal Resolution: The smallest lateral distance separation between two objects at a given distance that is discernible.

Range/ Depth Resolution: The smallest distance separation (along the scan axis) between two distinct objects that can be detected in a single return.

Vertical Resolution: The smallest vertical distance separation between two objects at a given distance that is discernible.

LADAR

An acronym for laser detection and ranging. A LADAR is a system that is used to obtain multiple distance measurements of a scene. These measurements, several thousand to several million, are commonly referred to as a “point cloud”. The distances are measured by measuring the time-of-flight of a laser pulse, the phase difference of a laser pulse, or by triangulation. A LADAR is also referred to as a laser radar or a laser scanner and is also commonly associated with ground-based laser radars/laser scanners.

LIDAR

An acronym for light detection and ranging. A LIDAR is a laser system that is typically associated with long-range, remote sensing applications. Example applications include atmospheric monitoring and airborne scanning.

Mixed Pixels/Phantom Points

Mixed pixels or phantom points result when a laser hits the edge of an object. When this occurs, the laser beam is split and part of it is reflected off the object and the other part is reflected off another object beyond the first object. Typically, the reported range measurements in such cases fall between the ranges to these two objects; hence, recording points that do not exist and are referred to as mixed pixels or phantom points.

B.2 ADDITONAL TERMS

Included in this section are some terms that were suggested by participants in the effort who could not attend the workshop. Some of the terms were suggested without definitions.

Angular resolution:

> The ability to resolve two objects on adjacent sightlines, and is a function of spatial sampling interval, the laser beam width, angular quantization and possibly other factors.

Effective instantaneous field of view (EIFOV)

> An angular resolution measure that accounts for all factors influencing resolution (see Angular Resolution).

Error: (of registration, that is, of fit to discrete target features and or cloud to cloud fit algorithms)

Spatial performance:

> Error defined as the square root of the average of the set of squared differences between dataset coordinate values and coordinate values from an independent field survey dataset of higher accuracy for identical well-defined control points

Optical center:

Spot Size:

Spatial resolution:

Control:

Fiducial targets vs ***discrete*** targets:

Well defined control points:

> Well-defined points are those that are a visible or recoverable reference common to both the independent source of higher accuracy and on product itself (point-cloud).

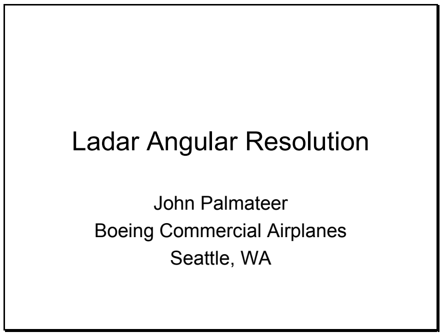
Uniformly distributed well-defined control points:

> Minimum of 20 uniformly distributed "well-defined" points or control points at intervals of at least 10 % of the diagonal distance across the data set and at least 20 % of the control points in each quadrant of the dataset allowing one point in 20 to fail the threshold given in product specifications.

APPENDIX C: WORKSHOP PRESENTATIONS

C.1 LADAR ANGULAR RESOLUTION BY JOHN PALMATEER

Slide 1



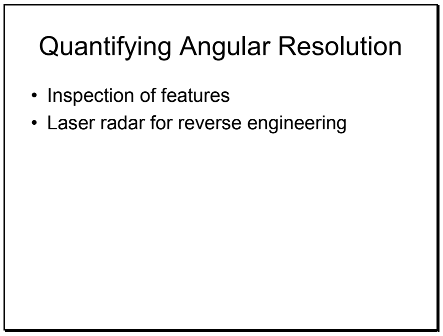
Ladar Angular Resolution

John Palmateer
Boeing Commercial Airplanes
Seattle, WA

John Palmateer is a Boeing Technical Fellow specializing in metrology. John has been involved with the development and implementation of 3D metrology starting with photogrammetry, theodolites, tracking interferometers, structured light measurement equipment and laser projectors.

Current interests include development of numerical processes for making 3D measurements more robust. This includes development of better programming tools in the Metrology GUIs that are currently on the market, as well as processes for fitting measurements to warped parts.

Slide 2



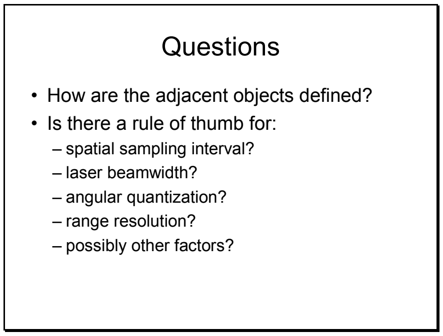
Quantifying Angular Resolution

- Inspection of features
- Laser radar for reverse engineering

My company has been looking at laser radar since the late 1980s for use in Tooling. The difficulty with this, and other 3D metrology system, was their difficulty to use. Recent developments in Metrology GUI has dramatically changed this.

Use in inspection often requires measurements that are capable of identifying objects. Likewise, reverse engineering requires measurement fidelity that allows items to be defined.
This talk is hopefully designed to stimulate conversation on this topic and perhaps lead to more research and answers.

Slide 3

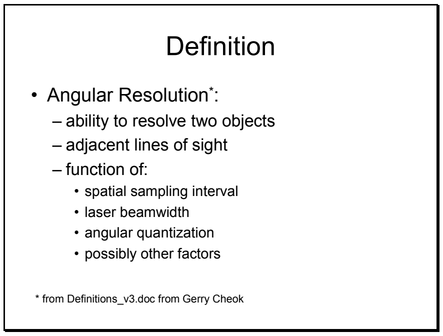


Questions

- How are the adjacent objects defined?
- Is there a rule of thumb for:
  - spatial sampling interval?
  - laser beamwidth?
  - angular quantization?
  - range resolution?
  - possibly other factors?

So the question stemming from the resolution issue is how are objects recognized and identified? Are there simple rules that can be applied in the shop to aid personnel in making sure they are getting the needed data? Some of the parameters are listed. The question is how do they impact the ability to identify an object? What are the interactions between these factors? And of course, there will be equipment specific characteristics that impact these (e.g., MetricVision averaging measurements while moving).

Slide 4



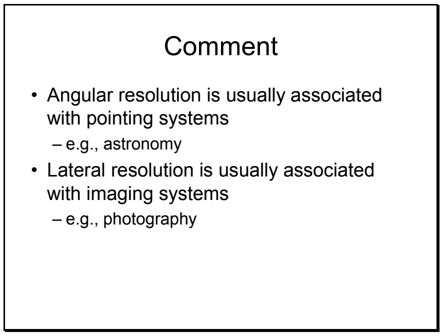
Definition

- Angular Resolution*:
  - ability to resolve two objects
  - adjacent lines of sight
  - function of:
    - spatial sampling interval
    - laser beamwidth
    - angular quantization
    - possibly other factors

* from Definitions_v3.doc from Gerry Cheok

This is the definition provided by Gerry in the material for the Workshop. It is a good starting point for discussion.

Slide 5



Comment

- Angular resolution is usually associated with pointing systems
  - e.g., astronomy
- Lateral resolution is usually associated with imaging systems
  - e.g., photography

This comment may be out of place, but from experience in the field, resolution is in the eye of the beholder.

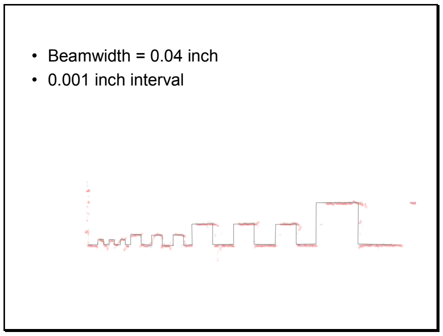
Slide 6

Possible Starting Point

- 3D grating
  - akin to USAF resolution chart
    - groups of three objects
    - doubling in size from group to group

Work on Ophthalmic holography centered on the use of a USAF resolution chart. From discussion at the previous Workshop, this grating was fabricated. The small tooth size was limited by the size of the mill bit. The smallest size tooth is 0.020 inch. It has been suggested that stacked sheet metal may be a method of getting smaller teeth sizes. A problem with remedial math caused the block to be too short.

Slide 7



- Beamwidth = 0.04 inch
- 0.001 inch interval

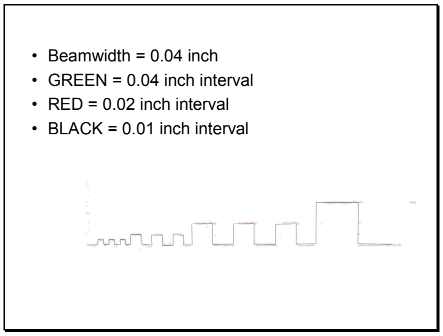
It should be noted that I had no way to confirm the beam diameter. This value was taken from the manufactures equations for beam diameter versus range.

The teeth shown are at the small end of the resolution block. The smallest tooth shown is 0.020 in. The next are (0.040, 0.080, and 0.160) in.

The points shown are dense measurements, 0.001 inch spacing. The points on the smallest teeth show some of the variation that occurs as a result of multiple reflections in corners, and, I think, coherent interference at the surface.

Also quite evident are many measurements that are an average across two surfaces. These are points that could be easily taken out, but not reconstructed, using a filter (one was developed for structured light measurement systems but not implemented here).

Slide 8



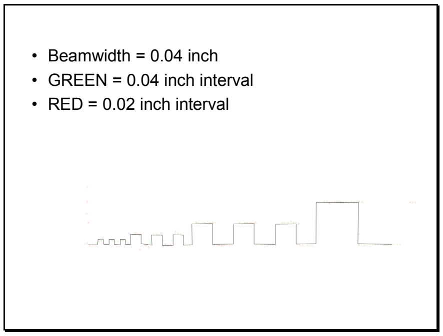
- Beamwidth = 0.04 inch
- GREEN = 0.04 inch interval
- RED = 0.02 inch interval
- BLACK = 0.01 inch interval

Again the small tooth end of the resolution block. Several scans are superimposed. There appears to be a calibration problem with the interval size. It appears that the true interval size is half the entered value.

Nonetheless, the smallest teeth are recognizable, although a bit rounded. Note that the laser radar see to the bottom of the teeth, even though the spot size is nominally twice the size of the teeth.

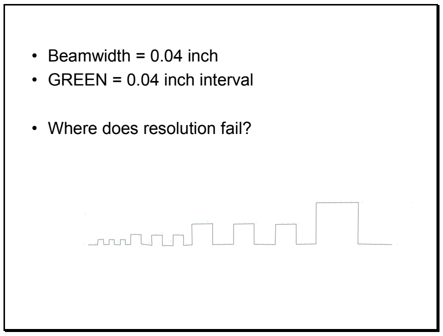
It is also clear from this collection of points that the surface interaction between the aluminum and laser is significant. As will be seen a mere convolution between the 3D grading and the round Gaussian beam should not see the bottom of the smallest teeth – but it does. Additionally, one can see other interactions, such as on the 4th tooth from the right – something that might plot out as a vertical asymptote.

Slide 9



- Beamwidth = 0.04 inch
- GREEN = 0.04 inch interval
- RED = 0.02 inch interval

Slide 10



- Beamwidth = 0.04 inch
- GREEN = 0.04 inch interval

- Where does resolution fail?

One can see the smallest teeth, or at least the depth change that marks these teeth, but cannot discern the teeth.

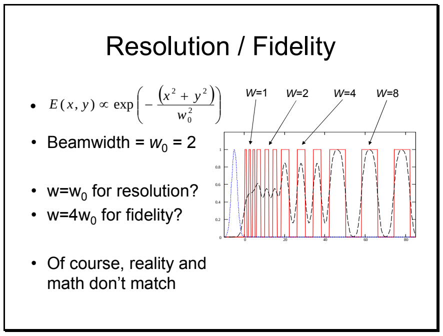
Slide 11

Traditional Resolution

- Resolved images are more than single lines of sight
- Resolution says little about identifying an object (fidelity?)

So just as with tradition definitions of resolution we can tell something is there, but we cannot tell what.

Slide 12



Resolution / Fidelity

- $E(x,y) \propto \exp\left(-\frac{(x^2+y^2)}{w_0^2}\right)$
- Beamwidth = w_0 = 2

- w=w_0 for resolution?
- w=4w_0 for fidelity?

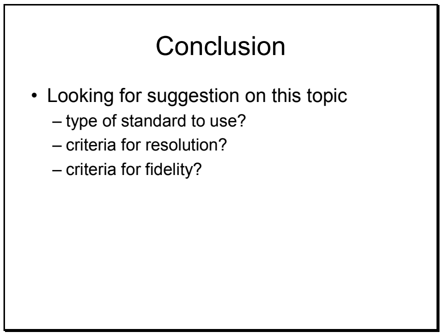
- Of course, reality and math don't match

A little math. A round gaussian beam is stepped over a 1D grating. Where the beam width is twice the smallest tooth size, there is no discernable signal. Once the beam with and tooth size are similar, there is sufficient change in return range to determine that something is there. Finally, when the tooth size is about 4X the beam width, the teeth are discernable.

Note the difference between this and reality. Even though the math did not see the bottom of the teeth when beamwidth is half the tooth size, the actual data sees the bottom and top of the teeth.

Caveat. since I could not measure the actual diameter of the laser, this conclusion could be in error.

Slide 13



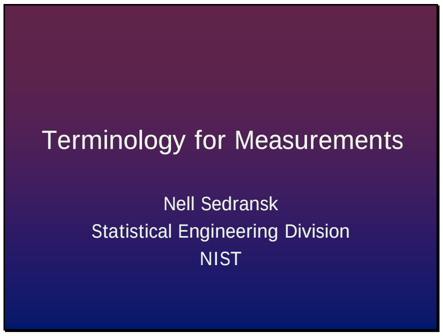
Conclusion

- Looking for suggestion on this topic
  - type of standard to use?
  - criteria for resolution?
  - criteria for fidelity?

Back to the question and generation of discussion: are there any better artifacts for lateral resolution and fidelity? What is the criteria for resolution? Signal 3dB down similar to photography? Rules of thumb for fidelity?

C.2 TERMINOLOGY FOR MEASUREMENTS BY NELL SEDRANSK

Slides 1 & 2



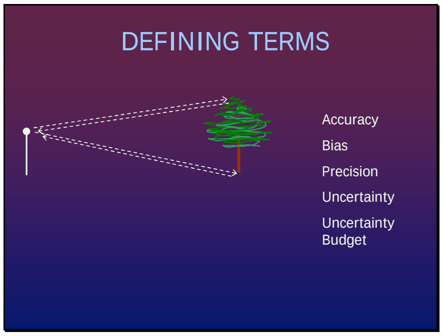
Terminology for Measurements

Nell Sedransk
Statistical Engineering Division
NIST

ESSENTIAL IDEAS

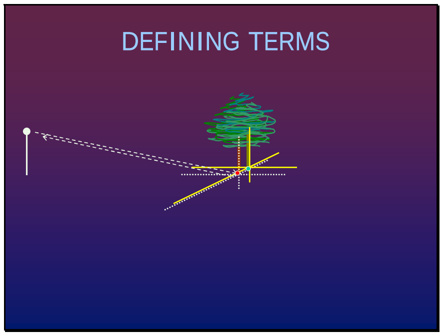
- Deciding What to Measure
- Understanding the Measurement
- Determining Properties of Measurements
  - Accuracy / Bias
  - Precision
  - Uncertainty

Slides 3 & 4



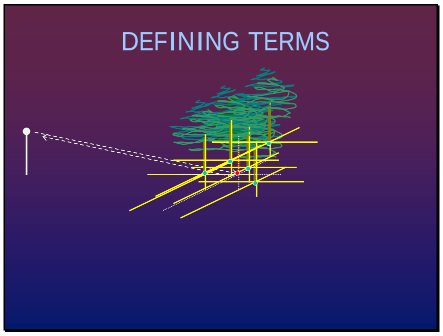
DEFINING TERMS

Accuracy
Bias
Precision
Uncertainty
Uncertainty Budget

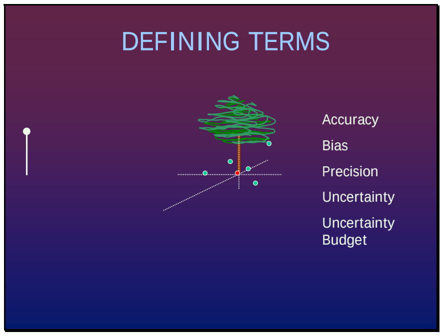


DEFINING TERMS

Slides 5 & 6



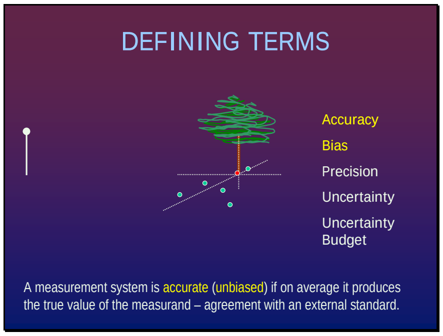
DEFINING TERMS



DEFINING TERMS

Accuracy
Bias
Precision
Uncertainty
Uncertainty Budget

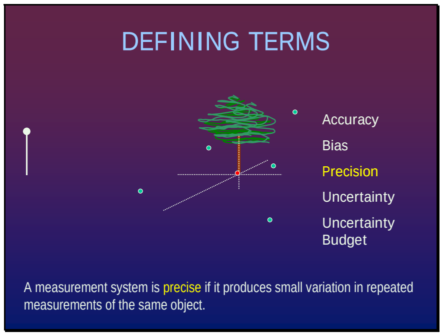
Slides 7 & 8



DEFINING TERMS

Accuracy
Bias
Precision
Uncertainty
Uncertainty Budget

A measurement system is accurate (unbiased) if on average it produces the true value of the measurand – agreement with an external standard.

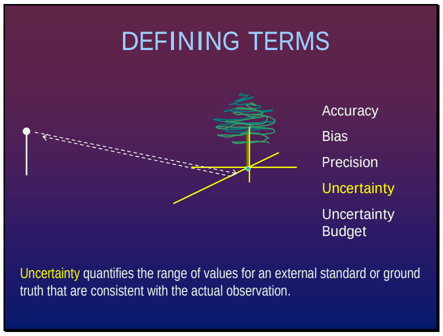


DEFINING TERMS

Accuracy
Bias
Precision
Uncertainty
Uncertainty Budget

A measurement system is precise if it produces small variation in repeated measurements of the same object.

Slides 9 & 10



DEFINING TERMS

Accuracy

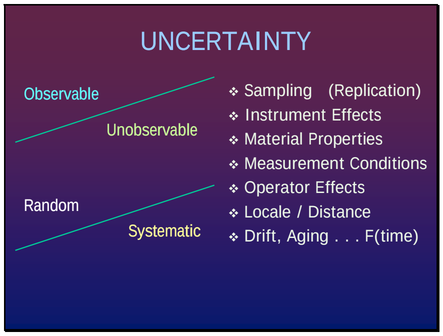
Bias

Precision

Uncertainty

Uncertainty Budget

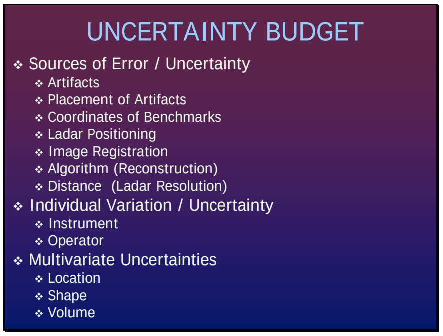
Uncertainty quantifies the range of values for an external standard or ground truth that are consistent with the actual observation.



UNCERTAINTY

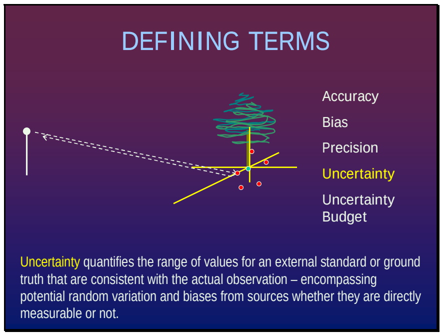
- Sampling (Replication)
- Instrument Effects
- Material Properties
- Measurement Conditions
- Operator Effects
- Locale / Distance
- Drift, Aging . . . F(time)

Slides 11 & 12



UNCERTAINTY BUDGET

- Sources of Error / Uncertainty
  - Artifacts
  - Placement of Artifacts
  - Coordinates of Benchmarks
  - Ladar Positioning
  - Image Registration
  - Algorithm (Reconstruction)
  - Distance (Ladar Resolution)
- Individual Variation / Uncertainty
  - Instrument
  - Operator
- Multivariate Uncertainties
  - Location
  - Shape
  - Volume



DEFINING TERMS

Accuracy

Bias

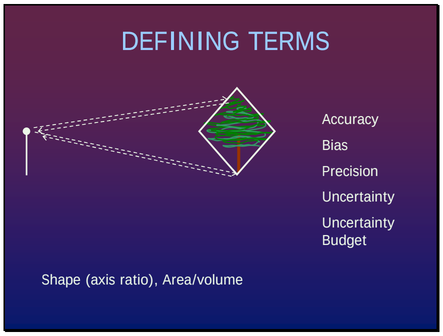
Precision

Uncertainty

Uncertainty Budget

Uncertainty quantifies the range of values for an external standard or ground truth that are consistent with the actual observation – encompassing potential random variation and biases from sources whether they are directly measurable or not.

Slides 13 & 14



DEFINING TERMS

Accuracy

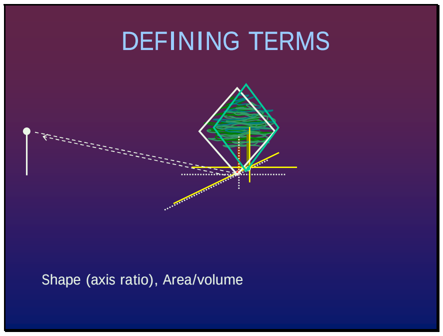
Bias

Precision

Uncertainty

Uncertainty Budget

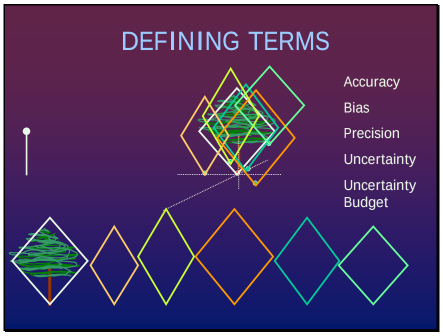
Shape (axis ratio), Area/volume



DEFINING TERMS

Shape (axis ratio), Area/volume

Slides 15 & 16



DEFINING TERMS

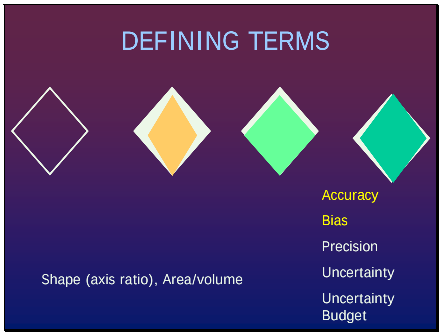
Accuracy

Bias

Precision

Uncertainty

Uncertainty Budget



DEFINING TERMS

Accuracy

Bias

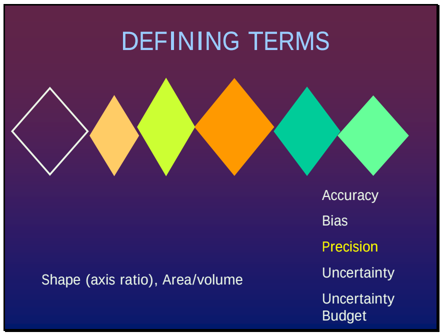
Precision

Uncertainty

Uncertainty Budget

Shape (axis ratio), Area/volume

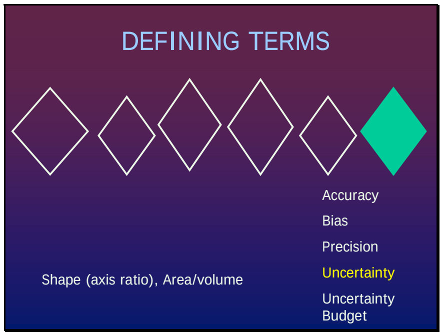
Slides 17 & 18



DEFINING TERMS

Accuracy
Bias
Precision
Uncertainty
Uncertainty Budget

Shape (axis ratio), Area/volume

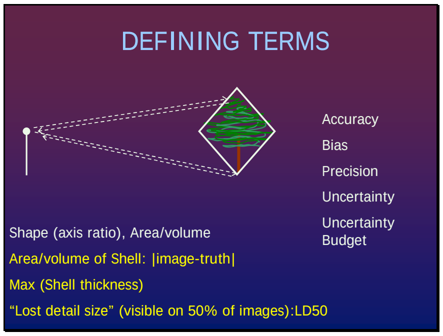


DEFINING TERMS

Accuracy
Bias
Precision
Uncertainty
Uncertainty Budget

Shape (axis ratio), Area/volume

Slide 19



DEFINING TERMS

Accuracy
Bias
Precision
Uncertainty
Uncertainty Budget

Shape (axis ratio), Area/volume
Area/volume of Shell: |image-truth|
Max (Shell thickness)
“Lost detail size” (visible on 50% of images):LD50

C.3 TESTING LASER SCANNERS: EXPERIENCES AT I3MAINZ BY WOLFGANG BÖHLER

The contents of Prof. Böhler’s presentation may be found in Ref. 5.